Confined random walks : explored territory and jump processes

Jérémie Klinger

SORBONNE UNIVERSITÉ

LPTMC

THÈSE DE DOCTORAT
de Sorbonne Université

Spécialité : Physique

École Doctorale n°564 : Physique en Île-de-France

réalisée

au Laboratoire de Physique Théorique de la Matière Condensée

sous la direction de
Olivier BENICHOU
et Raphaël VOITURIEZ

présentée par

Jérémie KLINGER

Sujet de la thèse :

Marches aléatoires en confinement : Territoire exploré et Processus de Saut

Confined Random Walks: Explored Territory and Jump Processes

Soutenue le 4 juillet 2023

Devant le jury composé de

M^{me}	Raffaella BURIONI	Rapporteuse
M.	Heiko RIEGER	Rapporteur
M.	Kirone MALLICK	Examinateur
M.	Julien RANDON-FURLING	Examinateur
M.	Grégory SCHEHR	Président du jury
M.	Raphaël VOITURIEZ	Co-Directeur de Thèse
M.	Olivier BENICHOU	Co-Directeur de Thèse, Invité

Remerciements

En premier lieu, j'aimerais remercier Heiko Rieger et Raffaella Burioni, qui ont accepté de rapporter mon manuscrit, ainsi que les membres de mon jury, qui chacun à leur façon, ont fortement influencé mon travail. Kirone Mallick - qui m'a en premier introduit aux concepts de la physique statistique; Julien Randon-Furling - dont la thèse a inspiré certaines des directions de recherche que j'ai prises, et enfin Grégory Schehr - qui, avec Claude Loverdo, m'a parrainé avec bienveillance et disponibilité tout au long de la thèse.

Je remercie chaleureusement mes directeurs de thèse, Raphaël Voituriez et Olivier Bénichou qui m'ont tour à tour guidé, encadré et encouragé, aussi bien dans les moments d'euphorie que dans les inévitables périodes creuses. J'ai énormément appris à leur côté, et leur enthousiasme sans faille et constante gentillesse m'ont permis de pleinement m'épanouir durant ces trois années. Si celles-ci se terminent aujourd'hui, il ne s'agit pour moi que d'un début.

Je remercie tous les membres du LPTMC, qui ont fortement contribué à rendre mon quotidien agréable et stimulant. Plus particulièrement, chez les *vieux* (une distinction somme toute très subjective), je remercie Aurélien Grabsch pour sa bonne humeur et sa capacité à me remonter le moral, Julien Vidal pour les belles journées d'escalade dont je ne suis parfois revenu indemne que grâce à lui, et Pascal Viot, qui m'a donné l'opportunité d'enseigner à ses côtés et dont la passion pour l'informatique s'est révélée contagieuse.

Chez les *jeunes* ensuite, je remercie Alex, Léo, Julien et Brieuc, avec qui j'ai eu la chance de partager un bureau - je n'aurais souhaité aucune autre compagnie. Je remercie également l'équipe grimpette Tim, Saverio et David, pour les nombreuses soirées passées à écumer les salles parisiennes, et les quelques allers-retours impulsifs à Fontainebleau. Enfin, Louise, Anna, Pierre, merci d'avoir été présents dans tous les autres moments.

Je remercie évidemment mes très chers amis, qui m'ont épaulé durant ces péripéties, parfois avec intérêt, souvent avec patience, mais toujours avec une oreille attentive. Merci aux plus et aux moins anciens, aux parisiens et aux londoniens, si j'en suis ici aujourd'hui c'est certainement grâce à toutes ces expériences partagées au fil des années. Vous m'apportez beaucoup, et je suis impatient de continuer à grandir avec vous.

Un grand merci à ma grande famille, qui m'a soutenu et périodiquement accueilli, aussi bien en Alsace qu'à Marseille, pour quelques bouffées d'air frais loin de Paris - une forme de *vacances* étonnement productives ! Une pensée particulière pour mes parents, Yann et Muriel qui, mieux que quiconque je crois, ont su me rassurer et m'apporter des réponses dans les moments de doute que j'ai pu traverser. Enfin, merci à mon petit frère Colin, dont la trajectoire, certes différente de la mienne, me rend aujourd'hui diablement heureux.

Mes derniers mots sont pour toi, Jeanne. Merci d'être là, et merci d'être toi. Merci pour tous ces projets que je ne saurais ni envisager seul, ni mener à bien; merci aussi pour notre petit havre routinier. Merci de m'avoir accompagné pendant cette thèse - tu m'as sauvé la mise plus d'une fois ! Enfin, merci d'avance pour la suite et toutes nos aventures à venir, j'ai hâte.

Contents

Chapter 0

Résumé des principaux résultats

Ce chapitre présente les principaux résultats obtenus au cours de cette thèse et peut être lu indépendamment.

0.1 Introduction

Omniprésents dans les systèmes physiques, les phénomènes de transport aléatoire s'observent à toutes les échelles temporelles et spatiales, et dans tous les domaines, de la biologie à la mécanique, en passant par la finance et la géologie. La diffusion de photons dans un milieu hétérogène [Rosenstock 1961, Savo *et al.* 2017], le mouvement d'un moteur moléculaire le long d'un brin d'ADN [Berg *et al.* 1981], le repliement d'une chaine polymérique [Lifshitz *et al.* 1978], l'évolution d'un actif financier [Black & Scholes 1973], la structure d'un vol d'étourneaux [Cavagna *et al.* 2010], les fluctuations d'un risque sismique [Matthews 2002]... sont autant de phénomènes qui peuvent être décrits par un mouvement aléatoire, et dont la compréhension est essentielle pour répondre à des questions très importantes en pratique. Par exemple, avec quelle probabilité le cours de la bourse atteint-il un certain seuil ? À quelle vitesse la transcription génétique, responsable de la synthèse de protéines nécessaires au bon fonctionnement de l'organisme, s'effectue-t-elle ? Ou encore, quelle est la distribution du temps d'attente entre deux séismes connaissant l'historique sismique d'une région donnée ?

Face à ces multiples interrogations, la communauté scientifique s'attelle depuis la fin du $19^{\text{ème}}$ siècle à catégoriser et modéliser ces phénomènes avec un objectif double : extraire des propriétés statistiques pertinentes pour interpréter les résultats empiriques, et tenter d'émettre des prédictions quantitatives afin de répondre aux questions pratiques soulevées. Deux siècles et demi plus tard, nous résumons, certes brutalement, les éléments essentiels à la modélisation et à l'interprétation d'un problème physique de transport aléatoire.

(i) En premier lieu, il convient de définir les règles du transport pour rendre compte de la dynamique aléatoire observée, et plus précisément d'identifier le processus stochastique sous-jacent qui décrit le mouvement de la particule considérée (nous utilisons ici le terme particule de manière générique, qu'il s'agisse d'un étourneau ou du prix d'une action).

Historiquement, c'est entre 1900 et 1905 que Bachelier [Bachelier 1900] et Einstein [Einstein 1905] introduisent pour la première fois le mouvement Brownien, afin de décrire le mouvement erratique de certaines particules expérimentales. Si leur modélisation repose sur le formalisme des équations aux dérivées partielles continues en temps et en espace, d'autres descriptions mathématiques parallèles émergent. Pearson propose un modèle de marche aléatoire en temps discret et espace continu [Pearson 1905] et, quelques années plus tard, Polya [Pólya 1921] se focalise sur les marches aléatoires sur réseaux hypercubiques, discrètes à la fois en temps et en espace. Notons que, bien que distincts, ces modèles partagent des similarités

fortes. Premièrement, tout trois sont *Markoviens*, c'est à dire que la dynamique stochastique ne dépend que de l'état de la particule à l'instant t donné, et non pas de l'historique de la trajectoire. Deuxièmement, il s'agit de processus *diffusifs*, pour lesquels le déplacement quadratique moyen $\langle x^2(t)\rangle$ de la particule (MSD pour *mean square displacement* en anglais) se comporte linéairement à grand temps : $\langle x^2(t)\rangle \sim t$.

Il est important de souligner que le MSD empirique d'un processus Markovien n'est pas toujours linéaire. Par exemple, certaines trajectoires d'animaux en quête de nourriture [Reynolds 2006] sont caractérisées par un MSD empirique sur-linéaire $\langle x^2(t)\rangle \propto t^\beta$ avec $\beta > 1$. L'introduction de nouveaux processus stochastiques pour correctement modéliser ces phénomènes de diffusion anormale [Bouchaud & Georges 1990] est alors nécessaire. Citons les vols de Levy [Levy 1937], un cas de marche aléatoire Markovienne dont les incréments sont distribués en loi de puissance $p(\ell) \propto \ell^{-(1+\mu)}$ avec $\mu \in]0, 2[$, conduisant à une dynamique *super-diffusive*, et un déplacement typique à grand temps $x_n \sim n^{1/\mu}$, par opposition au cas diffusif $x_n \sim \sqrt{n}$. À l'inverse, certains résultats expérimentaux, concernant la dynamique de matériaux vitreux par exemple, manifestent un comportement *sous-diffusif*, qui peut être correctement modélisé par des marches aléatoires en temps continu [Montroll & Weiss 1965, Monthus & Bouchaud 1996] (CTRW pour *Continuous Time Random Walk* en anglais), pour lesquelles la particule attend un temps aléatoire τ avant chaque déplacement, distribué selon une loi de puissance $p(\tau) \propto \tau^{-(1+\beta)}$ où $\beta \in]0, 1[$. Dans ce cas, le MSD est alors sous linéaire $\langle x^2(t)\rangle \propto t^\beta$.

Si les modèles Markoviens mènent souvent à une description simplifiée d'un phénomène physique, et permettent d'obtenir des résultats analytiques explicites, ceux-ci négligent complètement les effets de mémoire qui peuvent potentiellement influencer très fortement la dynamique de la particule observée. Parmi les modèles stochastiques non-Markoviens, citons l'exemple emblématique du mouvement Brownien Fractionnaire (FBM de l'anglais *Fractional Brownian Motion*) [Mandelbrot & Van Ness 1968], un processus gaussien non-Markovien caractérisé entièrement par sa fonction d'auto-corrélation $\langle x(t_1)x(t_2)\rangle \propto (|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H})$. En particulier, le FBM décrit correctement la dynamique sous-diffusive ($H < 1/2$) de polymères semi-flexibles. Enfin, dans un autre contexte, les marches à renforcement [Grassberger 2017] modélisent des phénomènes dont la dynamique à l'instant t dépend de l'historique complet de la trajectoire, et dont la distribution à n temps $(x(t_1), \ldots, x(t_n))$ n'est a priori pas gaussienne. Ce type de comportement a été observé dans certains systèmes naturels, chez les fourmis par exemple [Goss *et al.* 1989], dont le mouvement aléatoire est influencé par la géométrie du territoire visité précédemment.

(ii) Une fois qu'un processus stochastique pertinent pour décrire le mouvement aléatoire a été identifié, il est nécessaire d'introduire une *observable statistique*, c'est à dire une quantité probabiliste que l'on va chercher à évaluer afin d'en tirer des conclusions physiques. À différentes classes de questions physiques correspondent différentes observables, et nous distinguons deux grandes catégories.

Tout d'abord, les observables *dynamiques* permettent de décrire les propriétés statistiques de la trajectoire associée au processus aléatoire $x(t)$ à un instant déterministe t fixé. Si le MSD $\langle x^2(t)\rangle$, qui rend compte du caractère diffusif ou non du phénomène considéré, est l'observable dynamique la plus naturelle, d'autres observables permettent de quantifier l'efficacité du processus d'exploration. Par exemple, pour les marches aléatoires discrètes en temps et en espace, le nombre S_n de sites distincts visités après n pas [Dvoretzky & Erdös 1951, Wijland *et al.* 1997]

quantifie le taux de découverte de nouveaux sites. Notons que le calcul de la distribution de S_n est particulièrement pertinent pour déterminer la probabilité qu'une particule sur un réseau contenant des pièges distribués aléatoirement soit toujours en vie après n pas [Rosenstock 1961]. Pour les processus planaires en espace continu comme le mouvement Brownien 2D, l'aire de l'enveloppe convexe de la trajectoire [Majumdar *et al.* 2010a] constitue une observable dynamique alternative permettant de décrire l'étendue du territoire visité.

La seconde catégorie correspond aux observables de *premier passage*, qui caractérisent les propriétés statistiques de la trajectoire au bout d'un temps d'arrêt *aléatoire*, et sont associées à l'efficacité d'un processus de recherche de cible. En particulier, le temps de premier passage (FPT de l'anglais *first-passage time*) sur une cible fixée a reçu une attention particulière dès l'introduction du mouvement Brownien; en effet, de nombreux phénomènes physiques sont contrôlés par un événement de premier passage : cyclisation d'une chaine polymérique [Gooden *et al.* 1998], fixation d'une protéine [Berg *et al.* 1981] ou encore longueur d'une file d'attente [Asmussen 2003]. Par conséquent, la distribution du FPT a été beaucoup étudiée, aussi bien pour des marches discrètes sur réseaux et graphes [Montroll 1969, Haynes & Roberts 2008], que dans le cadre des processus confinés invariants d'échelle [Bénichou *et al.* 2010a] et, plus récemment, des processus confinés non-Markoviens [Levernier *et al.* 2018]. Notons enfin que les observables de premier passage ne sont pas uniquement associées au temps de premier passage, mais s'inscrivent également dans le cadre de la recherche préférentielle de cible. En particulier, la probabilité de splitting [Zoia *et al.* 2009] définie comme la probabilité d'atteindre une cible avant une autre, est une observable fondamentale, quantifiant par exemple la probabilité de fixation allélique [Wright 1931].

En résumé, la modélisation et l'interprétation d'un phénomène de transport aléatoire s'organise autour de deux axes : le choix d'un processus aléatoire précis - règles de saut et géométrie - et d'une observable stochastique associée, reflétant une propriété physique du système que l'on cherche à analyser. Dans ce manuscrit, nous nous appuyons sur ces deux aspects et proposons d'étudier deux problèmes distincts. Dans un premier temps, nous nous focalisons uniquement sur le second axe, et introduisons une nouvelle observable, le territoire visité avant atteinte d'une cible en confinement, que nous caractérisons pour la classe de processus stochastiques la plus vaste possible. Inversement, dans une seconde partie, nous nous concentrons sur un type de marche aléatoire particulière : les processus de saut unidimensionnels en confinement, pour lesquels, bien qu'introduits dès le début du 20$^{\text{ème}}$ siècle, peu d'observables ont été évaluées.

Partie I. Dans le cadre de marches aléatoires sur réseau, le nombre S_n de sites distincts visités après n pas, qui quantifie l'efficacité du processus d'exploration, a été longuement étudié; pour une marche normale sur un réseau 1D, la distribution complète de S_n est connue explicitement [Dvoretzky & Erdös 1951]. Dans le cas des réseaux hypercubiques en plus grande dimension, seul le comportement asymptotique à grand n des premiers moments de S_n a été déterminé [Jain & Pruitt 1971, Jain & Pruitt 1974]. Ces résultats ont été étendus au cas de marches de plus proches voisins sur graphes quelconques [Havlin & ben Avraham 1987], ainsi qu'aux marches de Riemann unidimensionnelles [Gillis & Weiss 1970, Mariz *et al.* 2001] qui effectuent des sauts distribués algébriquement [Hughes 1995].

Notons que ces résultats concernent essentiellement des marches aléatoires dans une géométrie

infinie; en effet, dans un espace confiné, le nombre de sites total est borné, et le marcheur finit invariablement par explorer l'intégralité du domaine confinant. Pourtant, la question du nombre de sites distincts visités reste légitime, pour peu que l'on définisse un temps d'arrêt physique auquel S_n est évalué. Par exemple, pour une particule confinée dans une boîte possédant une unique sortie, on pourra se demander quelle fraction du territoire est visitée avant que la particule ne s'échappe. En particulier, cette information est essentielle pour déterminer la probabilité que la particule réagisse avec une cible intérieure avant de trouver la sortie. Pour répondre à ce type de question, nous consacrons la première partie de cette thèse à la détermination du nombre $C(s_0)$ de sites distincts visités par un marcheur confiné issu de s_0, avant d'atteindre une cible fixée s_T. Nous insistons sur le fait que le temps d'arrêt correspondant à la découverte de la cible est intrinsèquement lié à la trajectoire d'exploration; $C(s_0)$ est donc une observable de *premier passage*.

Théoriquement, l'évaluation des propriétés statistiques du territoire visité $C(s_0)$ présente un certain nombre de difficultés techniques. En particulier, le statut (visité ou non) du site atteint au $n^{\text{ème}}$ pas, dépend de l'historique complet des sites visités, et la présence d'un site absorbant s_T modifie la statistique des trajectoires. Afin de contourner ces difficultés, nous avons montré que l'évaluation du nombre de sites distincts visités $C(s_0)$ se ramène à la détermination des probabilités de splitting associées au processus considéré.

Ce faisant, nous avons calculé la distribution exacte de $C(s_0)$ pour la classe des processus unidimensionnels avec span connecté, c'est à dire ne laissant pas de *trous*, parmi lesquels figurent la marche normale et la marche persistante par exemple. De manière plus générale, dans le cadre de processus invariants d'échelle et dans la limite de grand volume confinant, nous avons identifié différentes classes d'universalité, et montré que la distribution complète de $C(s_0)$ prend une forme de scaling explicite. Nos résultats couvrent en particulier les exemples classiques des marches de Riemann, ainsi que des marches de plus proches voisins sur réseaux fractals. En conclusion, nous soulignons que l'introduction et la détermination des propriétés statistiques du territoire $C(s_0)$ visité avant atteinte d'une cible ouvre la voie à une caractérisation plus fine des propriétés géométriques du processus d'exploration avant la sortie d'une enceinte confinante, et présente un intérêt expérimental fort, notamment dans le cadre de réactions chimiques limitées par diffusion.

Partie II. La seconde partie de cette thèse est consacrée non plus à une unique observable, mais à un unique type de processus stochastiques : les processus de saut unidimensionnels. Ces marches aléatoires sur $\mathbb{R}$ décrivent la position x_n d'une particule issue de x_0 après n pas, et dont les incréments aléatoires sont indépendants et identiquement distribués, de distribution commune $p(\ell)$ symétrique quelconque. En particulier, le modèle de la marche exponentielle [Van Kampen 1992] est un exemple typique de processus de saut, pour lequel $p(\ell) = \frac{1}{2}e^{-|\ell|}$. Ces processus sont particulièrement importants pour modéliser et interpréter les données expérimentales ou numériques, qui sont intrinsèquement discrètes, et ce de manière transverse, que ce soit dans le cadre de problèmes de transfert radiatif [Milne 1921], ou pour décrire le mouvement saccadé de bactéries [Koshland 1980].

Les propriétés statistiques de processus de saut en milieu infini, c'est à dire pouvant évoluer sur $\mathbb{R}$ tout entier, sont bien connues. Par exemple, la distribution de la position de la particule

au temps n s'écrit explicitement en fonction de $p(\ell)$. De la même manière les processus de saut en milieu semi-infini, c'est à dire stoppés dès le premier passage à travers 0 (*ie* dès que x_n devient strictement négatif), sont également bien caractérisés. La distribution de la position après n pas est explicite [Ivanov 1994], et la probabilité de survie, définie comme la probabilité de rester positif durant les n premiers pas, est connue exactement pour toutes valeurs de n et x_0 [Doney 2012, Majumdar *et al.* 2017]. Citons en particulier un résultat surprenant connu sous le nom de théorème de Sparre Andersen [Andersen 1954] : la probabilité de survie partant de $x_0 = 0$ est universelle et indépendante de $p(\ell)$.

En revanche, pour les processus de saut confinés dans un intervalle $[0, x]$, c'est à dire stoppés dès leur première sortie, il n'existe pas de résultats généraux concernant la distribution de la position de la particule après n pas, ni même quantifiant la probabilité de survie associée à l'intervalle. Nous soulignons que la difficulté technique réside principalement dans le fait que les équations intégrales définissant ces observables ne sont pas solubles pour des processus de saut arbitraires. Cependant, nous avons montré que dans la limite de grand intervalle $x \to \infty$, certaines observables *de premier passage* associées aux événements de sortie de $[0, x]$ peuvent être évaluées de manière générale, c'est à dire pour toute distribution $p(\ell)$.

Tout d'abord, en exhibant un lien entre les processus de saut confinés et semi-infinis, nous avons montré que la probabilité de splitting $\pi_{0,\underline{x}}(x_0)$, définie comme la probabilité que le marcheur sorte à travers x plutôt que 0, prend une forme asymptotique universelle, valable pour tout $p(\ell)$ et toute condition initiale x_0. Pour caractériser plus finement les événements de sortie, nous nous sommes ensuite intéressés à la distribution du temps de sortie de l'intervalle. Dans la limite asymptotique de grand temps et grand intervalle, nous avons montré que cette distribution prend également une forme universelle, indépendante des détails de $p(\ell)$. Enfin, nos résultats s'étendent naturellement à des processus de saut confinés en plus grande dimension, qui sont particulièrement pertinents pour décrire des situations physiques réalistes.

Les sous-sections suivantes sont dédiées à un résumé plus précis des résultats rassemblés dans ce manuscrit.

0.2 Territoire exploré avant atteinte d'une cible en confinement

Le nombre S_n de sites distincts visités par un marcheur aléatoire sur réseau durant ses n premiers pas est initialement introduit par [Dvoretzky & Erdös 1951] dans le cas de la marche normale unidimensionnelle, pour laquelle la distribution complète de S_n est calculable explicitement. Une dizaine d'années plus tard, Rosenstock propose son désormais classique modèle de pièges [Rosenstock 1961] : pour une particule évoluant sur un réseau quelconque où chaque site contient un piège avec une probabilité λ, quelle est la probabilité P_n pour que la particule ne rencontre aucun piège durant ses n premiers pas ? Bien que P_n s'exprime simplement en fonction de S_n

$$P_n = \langle (1-\lambda)^{S_n} \rangle, \tag{1}$$

son évaluation nécessite le calcul difficile de la distribution complète de S_n. Par conséquent, Rosenstock propose d'approximer P_n à l'aide du premier moment de S_n

$$P_n \simeq (1-\lambda)^{\langle S_n \rangle}, \tag{2}$$

et de nombreux travaux se focalisent alors sur le calcul des premiers moments de S_n dans diverses situations : réseaux hypercubiques [Jain & Pruitt 1971, Jain & Pruitt 1974], N marcheurs indépendants [Larralde *et al.* 1992], ou encore marches de Riemann sur réseau unidimensionnel [Gillis & Weiss 1970, Mariz *et al.* 2001].

Nous soulignons le fait que l'observable dynamique S_n, qui quantifie l'efficacité du processus d'exploration, a été majoritairement étudiée en géométrie infinie. En effet, dans une géométrie bornée, le domaine finit invariablement par être visité en entier. Cependant, dans le cas d'un domaine confinant possédant une sortie, il est légitime de se demander quel est le territoire exploré par le marcheur avant de s'échapper; rien ne garantit que le domaine sera visité en entier par exemple. Cette information est particulièrement pertinente dans le cadre d'applications chimiques, et permet par exemple de quantifier la probabilité de réaction d'un catalyseur avec un ligand avant que celui ci ne s'échappe de l'enceinte de réaction (voir figure 1).

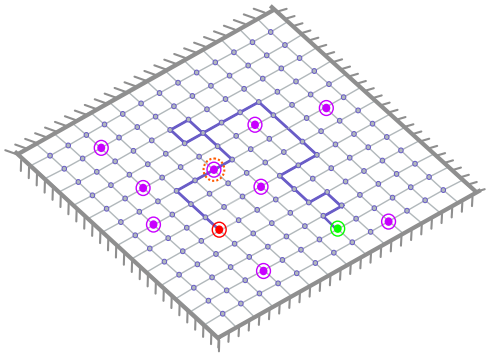


Figure 1: La particule issue du s_0 (point vert) évolue sur un réseau 2D borné, possédant une sortie en s_T (point rouge). Les sites réactifs (point violets) sont distribués aléatoirement sur le réseau. Dans cet exemple de trajectoire, l'un des sites réactifs est rencontré avant que la particule n'atteigne la sortie.

Afin de répondre à cette question, nous avons introduit une nouvelle observable *de premier passage* $C(s_0)$, définie comme le territoire - ou nombre de sites distincts - visité par un marcheur

aléatoire issu de s_0, avant d'atteindre une cible fixée s_T dans un espace confiné. Par conséquent, la probabilité de survie du marcheur dans un domaine contenant des pièges distribués avec probabilité λ sur chaque site, c'est à dire la probabilité que celui-ci s'échappe sans rencontrer de piège, est simplement donnée par:

$$P(\text{survie}|s_0) = \langle (1-\lambda)^{C(s_0)} \rangle. \tag{3}$$

À nouveau, le calcul de la distribution complète du territoire visité $C(s_0)$ est nécessaire pour déterminer $P(\text{survie}|s_0)$.

0.2.1 Marches unidimensionnelles connectées

En premier lieu, nous avons considéré le cas particulier des processus 1D sur réseau avec *span connecté*, définis comme suit : pour deux sites s_1 et s_2 visités, tout site intermédiaire s tel que $s_1 \leq s \leq s_2$ est nécessairement visité. Autrement dit, la découverte de nouveaux sites ne peut se faire que par plus proches voisins. Dans ce cadre, nous avons montré que le nombre de sites distincts visités avant d'atteindre la cible localisée en 0, est directement relié à la distribution $\mu(s_m|s_0)$ du site maximal s_m atteint avant le premier passage en 0, pour le même processus non-confiné. Sur un anneau contenant N sites (voir figure 2(a)), la distribution de $C(s_0)$ s'exprime alors simplement comme

$$P(C=n|s_0) = 1_{s_0+1\leq n}\ \mu(n-1|s_0) + 1_{n\geq N-s_0+1}\ \mu(n-1|N-s_0), \tag{4}$$

où les effets de périodicité sont pris en compte en considérant indépendamment les trajectoires de sortie horaires et anti-horaires. Pour faciliter l'évaluation de la distribution (4) de $C(s_0)$, nous avons exprimé la distribution du maximum $\mu(s|s_0)$ en fonction de la probabilité de splitting $\pi_{\underline{0},s}(s_0)$, définie comme la probabilité d'atteindre 0 avant s, partant de s_0:

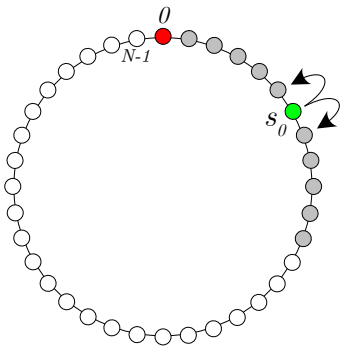
$$\mu(s|s_0) = \pi_{\underline{0},s+1}(s_0) - \pi_{\underline{0},s}(s_0). \tag{5}$$

Par conséquent, pour les marches aléatoires 1D confinées avec span connecté, le calcul de la probabilité de splitting $\pi_{\underline{0},s}(s_0)$ est suffisant pour obtenir la distribution complète du territoire visité avant d'atteindre 0. De plus, les équations (4) et (5) s'adaptent facilement au cas de processus unidimensionnels continus. En pratique, la probabilité de splitting est calculable pour un grand nombre de marches Markoviennes, et nous avons obtenu la distribution de $C(s_0)$ pour de nombreux processus discrets, comme la marche normale et la marche persistante [Weiss 1994], mais aussi continus, comme le Mouvement Brownien avec resetting [Evans & Majumdar 2011].

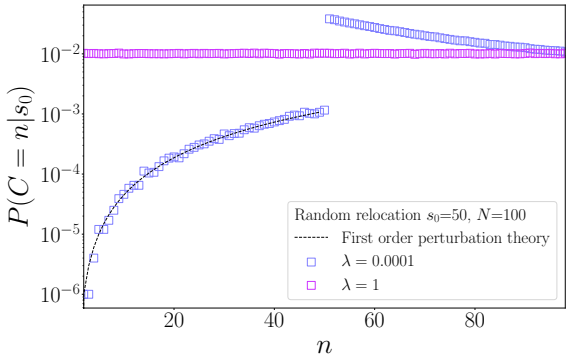
Il est important de souligner que la condition de span connecté est nécessaire pour pouvoir relier le territoire visité et le maximum. Pour aller plus loin, nous avons d'abord considéré le cas d'une marche normale symétrique se relocalisant uniformément sur l'anneau avec probabilité λ à chaque pas. Dans le cas d'une relocalisation systématique ($\lambda = 1$) la distribution de $C(s_0)$ est explicite :

$$P(C=n|s_0) = \frac{1}{N-1}. \tag{6}$$

Dans le cas $\lambda < 1$, le résultat (4) n'est plus valable, et nous avons montré que le calcul de la distribution de $C(s_0)$ peut se faire perturbativement en puissances de λ (voir figure 2(b)). Ensuite, pour s'affranchir des difficultés calculatoires liées aux détails microscopiques des règles de transport et dépasser le cadre des processus 1D avec span connecté, nous nous sommes intéressés à l'évaluation du territoire $C(s_0)$ dans la limite de grand volume confinant, en nous focalisant sur la classe générale des processus stochastiques invariants d'échelle.



(a) Trajectoire typique d'une marche normale contribuant à $P(C = 12|s_0 = 6)$ dans le sens anti-horaire. Partant de s_0, la particule visite tous les sites grisés avant d'atteindre la cible absorbante localisée en 0. En particulier, le lien entre $C(s_0)$ et le maximum atteint avant le FPT en 0 est manifeste.



(b) Distribution de $C(s_0)$ pour une marche symétrique se relocalisant aléatoirement avec probabilité λ. Dans le cas de la relocalisation systématique, la distribution de $C(s_0)$ est bien uniforme. Dans le cas $\lambda \ll 1$ nous obtenons une approximation perturbative de la distribution de $C(s_0)$.

Figure 2

0.2.2 Processus invariants d'échelle dans la limite de grand volume

Les processus invariants d'échelle [Havlin & ben Avraham 1987] forment une classe de processus stochastiques définis par la forme de scaling du propagateur en espace infini $P^{\infty}(\mathbf{x}, t, \mathbf{x_0})$, qui correspond à la probabilité pour que le processus soit à la position $\mathbf{x}$ à l'instant t partant de $\mathbf{x_0}$:

$$P^{\infty}(\mathbf{x}, t|\mathbf{x_0}) = \frac{1}{t^{\frac{d_f}{d_w}}} \Pi\left(\frac{|\mathbf{x} - \mathbf{x_0}|}{t^{\frac{1}{d_w}}}\right). \tag{7}$$

La dimension fractale d_f caractérise le domaine dans lequel le processus évolue, et plus précisément le nombre V de sites distincts accessibles à une distance donnée R, $V \propto R^{d_f}$. La dimension de marche d_w décrit quant à elle le déplacement typique du processus $x(t) \propto t^{1/d_w}$ aux temps longs, et Π est une fonction dépendant spécifiquement du processus considéré. Notons que de nombreuses marches aléatoires ont un comportement invariant d'échelle à grand temps, parmi lesquels la diffusion normale en d-dimensions, mais également la diffusion sur graphes fractals déterministes, comme le triangle de Sierpinski [Havlin & ben Avraham 1987] ou aléatoires, comme les clusters de percolation [Bouchaud & Georges 1990].

De manière générale, pour un marcheur aléatoire partant de s_0 sur un réseau Ξ, le territoire $C(s_0)$ visité avant atteinte d'une cible localisée en s_T se réécrit sous la forme exacte suivante

$$C(s_0) = \sum_{s\in\Xi} \mathbb{1}(s \text{ est visité avant } s_T|s_0). \tag{8}$$

Nous soulignons que la difficulté provient du fait que les variables indicatrices sont fortement corrélées, puisqu'elles sont évaluées le long d'une même trajectoire. Concentrons nous d'abord sur le premier moment $\langle C(s_0)\rangle$ du territoire visité, qui s'exprime uniquement en fonction des probabilités de splitting $\pi_{s_2,\underline{s_1}}(s_0)$, définies comme la probabilité que le marcheur issu de s_0 atteigne s_1 avant s_2:

$$\langle C(s_0)\rangle = \sum_{s\in\Xi} \pi_{s_T,\underline{s}}(s_0). \tag{9}$$

Dans la limite de grand volume, c'est à dire avec s_0 et s_T fixé, et le volume V du réseau tendant vers $+\infty$, certaines observables de premier passage associées aux processus invariants d'échelle comme le FPT [Condamin *et al.* 2005, Condamin *et al.* 2007b, Chevalier *et al.* 2011] adoptent un comportement universel, uniquement régit par le ratio d_w/d_f. C'est aussi le cas pour les probabilités de splitting, qui sont asymptotiquement données par [Bénichou & Voituriez 2014]

$$\pi_{s_2,\underline{s_1}}(s_0) \sim \begin{cases} \dfrac{A + B(r_{10}^{d_w-d_f} - r_{20}^{d_w-d_f} - r_{12}^{d_w-d_f})}{2(A - Br_{12}^{d_w-d_f})} & \text{si } d_w < d_f \\ \dfrac{A + B\log\left(\frac{r_{20}\ r_{12}}{r_{10}}\right)}{2\left[A + B\log(r_{12})\right]} & \text{si } d_w = d_f \\ \dfrac{r_{20}^{d_w-d_f} + r_{12}^{d_w-d_f} - r_{10}^{d_w-d_f}}{2r_{12}^{d_w-d_f}} & \text{si } d_w > d_f \end{cases} \tag{10}$$

où r_{ij} correspond à la distance entre s_i et s_j, R est l'échelle de longueur caractéristique du domaine confinant, telle que $R^{d_f} \propto V$, et A et B sont des constantes qui ne dépendent que de la fonction d'échelle Π définie dans (7) . En exploitant la décomposition (9), nous avons montré que le territoire moyen visité prend également une forme universelle dans la limite de grand volume, qui ne dépend que de d_w/d_f:

$$\frac{\langle C(r_s)\rangle}{V} \sim \begin{cases} \dfrac{1}{2} - \dfrac{B}{2Ar_s^{d_f-d_w}} & \text{si } d_w < d_f \\ \dfrac{1}{2} + \dfrac{B\log\left(\frac{r_s}{R}\right)}{2\left[A + B\log(R)\right]} & \text{si } d_w = d_f \\ \dfrac{d_f}{2(2d_f - d_w)}\left[\dfrac{r_s}{R}\right]^{d_w-d_f} & \text{si } d_w > d_f \end{cases} \tag{11}$$

En particulier, nous avons entièrement caractérisé la dépendance fonctionnelle du territoire moyen visité dans les paramètres géométriques du système, à savoir la distance source-cible r_s, et la

taille du système $V \propto R^{d_f}$. En outre, ce comportement universel n'est pas limité au premier moment $\langle C(s_0)\rangle$. En identifiant les contributions relatives des trajectoires s'approchant ou non du bord du système, nous avons identifié deux familles de classes d'universalité distinctes pour le comportement de la distribution complète de la fraction $\mathcal{C} = C/V$ du territoire visité

$$\begin{aligned} P(\mathcal{C} = x|r_s) &\underset{1\ll r_s\ll R}{\sim} 2\langle \mathcal{C}(r_s)\rangle + [1 - 2\langle \mathcal{C}(r_s)\rangle]\,\delta(x) \quad \text{si} \quad d_w \leq d_f \\ P(\mathcal{C} = x|r_s) &\underset{1\ll r_s\ll R}{\propto} \left[\frac{r_s}{R}\right]^{d_w - d_f} \left[\frac{1}{x^{\frac{d_w}{d_f}}} + A\mu_1(x)\right] \quad \text{si} \quad d_w > d_f, \end{aligned} \tag{12}$$

où $\mu_1(x)$ est une fonction dépendante du processus considéré, telle que $\mu_1(x) = o(x^{-d_w/d_f})$ dans la limite $x \to 0$.

Dans les cas dits *non-compact* et *marginal* $d_w \leq d_f$ (aussi appelés transients), la distribution de $\mathcal{C}$ est asymptotiquement quasi uniforme; dans la limite de grand volume confinant, tout les sites du domaine deviennent équivalents, et l'ordre de découverte est distribué uniformément. En revanche, le cas *compact* (ou récurrent) $d_w > d_f$ est plus difficile à caractériser. Cependant, nous décrivons précisément la dépendance géométrique de la distribution de $\mathcal{C}$ en r_s et R, et obtenons également la forme fonctionnelle précise de la distribution dans le régime de petite fraction explorée $x \ll 1$, pour lequel la décroissance de $P(\mathcal{C} = x|r_s)$ est algébrique, et ne dépend que de $d_w - d_f$. En particulier, nous insistons sur le fait que le résultat (12) donne accès à la dépendance en r_s et R de tout les moments de $\mathcal{C}$, et permet par exemple de quantifier l'importance de ses fluctuations. Ces résultats analytiques sont illustrés de manière extensive sur de nombreux processus invariants d'échelle compacts et non-compacts sur la figure 3.

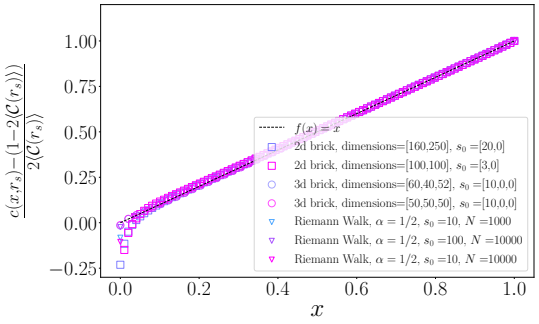
0.2.3 Statistiques jointes de temps et d'espace

Notons que la distribution du territoire visité avant l'atteinte de la cible s_T ne contient aucune information sur la statistique du FPT en s_T. Pourtant les deux variables sont intuitivement corrélées : une exploration plus longue mène probablement à un territoire exploré plus vaste. Afin de quantifier ce couplage, nous nous sommes intéressés à la distribution jointe $\sigma(s, n|s_0)$ du territoire visité s et du temps n de premier passage en 0. Nous soulignons que le calcul de cette distribution jointe donne accès à de nombreuses quantités importantes, comme les distributions marginales et conditionnelles du territoire visité et du FPT.

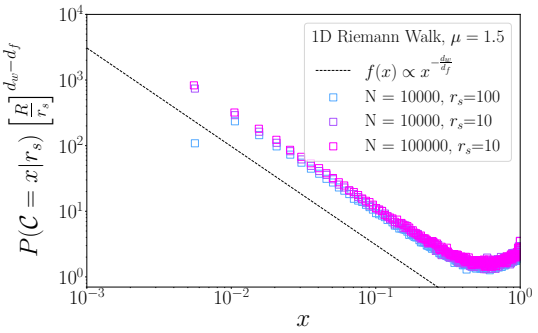
À l'instar de l'évaluation du territoire $C(s_0)$, nous nous somme d'abord focalisés sur l'élaboration d'une méthodologie pour évaluer systématiquement $\sigma(s, n|s_0)$ dans le cas exactement soluble des processus unidimensionnels avec span connecté. En introduisant la probabilité $F_{\underline{0},s}(n|s_0)$ de première sortie à gauche d'un intervalle (LETP pour *leftward exit-time probability* en anglais), définie comme la probabilité d'atteindre 0 pour la première fois au $n^{\text{ème}}$ pas sans avoir atteint s précédemment, la loi jointe $\sigma(s, n|s_0)$ se réécrit simplement en fonction de $F_{\underline{0},s}(n|s_0)$

$$\sigma(s, n|s_0) = F_{\underline{0},s+1}(n|s_0) - F_{\underline{0},s}(n|s_0). \tag{13}$$

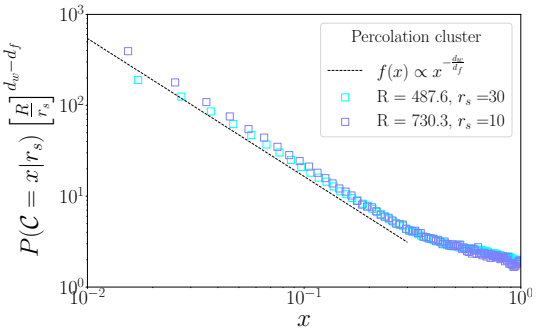
En d'autres termes, pour les processus 1D à span connecté, continus ou discrets, la détermination de la LETP est suffisante pour quantifier complètement le couplage entre temps de premier



(a) Distribution cumulative $c(x, r_s) = \int_0^x P(\mathcal{C} = u|r_s)du$ du territoire visité pour des processus marginaux (marche normale 2D) et non compacts (marche normale 3D, marche de Riemann 1D). Correctement rescalée selon l'équation (12) toutes les distributions se superposent. Le résultat dans la limite de grand volume est valable pour des domaines confinants de forme arbitraire.



(b) Distribution de $\mathcal{C}$ pour une marche de Riemann ($p(s_1 \to s_2) \propto |s_1 - s_2|^{-(1+\mu)}$) compacte avec $d_w = \mu = 3/2$ sur un anneau contenant N sites. Chaque courbe correspond à des valeurs de r_s et R différentes. Lorsque $x \to 0$, la distribution de $\mathcal{C}$ est algébrique et l'exposant est correctement prédit par l'équation (12).



(c) Distribution de $\mathcal{C}$ dans le cas de clusters de percolation 2D. Les distributions sont moyennées sur différentes paires source-cible et différent clusters générés à partir de grilles de taille 150^2 (bleue) et 200^2 (violet). L'échelle caractéristique R est également moyennée.

Figure 3

passage et territoire exploré.

En pratique, dans le cas de processus markoviens, la LETP s'obtient en partitionnant sur la position de la particule après le premier pas:

$$F_{\underline{0},s}(n|s_0) = \mathcal{L}_{s_0} F_{\underline{0},s}(n-1|s_0), \tag{14}$$

où l'opérateur $\mathcal{L}_{s_0}$ est défini par les détails microscopiques du processus considéré. Par exemple, dans le cas de la marche normale, $\mathcal{L}_{s_0} F_{\underline{0},s}(n-1|s_0) = \frac{1}{2}\left[F_{\underline{0},s}(n-1|s_0-1) + F_{\underline{0},s}(n-1|s_0+1)\right]$. En combinant (13) et (14) nous avons alors calculé la loi jointe du territoire visité et du temps de premier passage pour un certain nombre d'exemples représentatifs de processus markoviens 1D avec span connecté, comme la marche normale, la marche biaisée ou encore la marche normale avec resetting.

Afin de dépasser le cadre des processus Markoviens à span connecté, nous nous sommes ensuite intéressés à la loi jointe du territoire visité et du temps de premier passage pour des processus 1D invariants d'échelle, dans la limite de grand territoire et grand temps $n \to \infty$

et $s \to \infty$. Soulignons le fait que cette classe de processus englobe une multitude de modèles de transport unidimensionnels classiques, notamment des processus non-Markoviens comme le FBM [Mandelbrot & Van Ness 1968, Molchan 1999] évoqué en introduction, mais aussi le Random Acceleration Process [Bicout & Burkhardt 2000] ou encore des modèles de marches à *renforcement*, comme la Self Avoiding True Walk (SATW) [Sapozhnikov 1994].

Les propriétés de persistance de processus 1D invariants d'échelle, c'est à dire le comportement de la distribution $F_{\underline{0}}(n|s_0)$ du FPT en 0, ont fait l'objet de nombreux travaux [Bray *et al.* 2013]. En particulier, le comportement à grand n est caractérisé par une décroissance algébrique

$$F_{\underline{0}}(n|s_0) \underset{n\to\infty}{\sim} \frac{k(s_0)}{n^{1+\theta}}, \tag{15}$$

ou $k(s_0)$ est une fonction dépendant du processus considéré, et θ est appelé exposant de persistance. En introduisant la variable rescalée $\tau = n/s^{d_w}$, où d_w est la dimension de marche du processus, et en étendant une approche initialement développée dans [Levernier *et al.* 2018], nous avons montré que le loi jointe σ admet une forme de scaling universelle à grand n et grand s :

$$\sigma(s,n|s_0) \underset{\substack{s\to\infty \\ n\to\infty \\ \tau \text{ fixé}}}{\sim} \frac{h(s_0)}{s^{d_w\theta+1}} \frac{1}{s^{d_w}} f(\tau) \tag{16}$$

où h et f sont des fonctions qui dépendent spécifiquement du processus considéré. Nous avons alors illustré numériquement ce comportement asymptotique sur de nombreux processus invariants d'échelle Markoviens, comme les marches de Riemann, et non-Markoviens, notamment le FBM et le Random Acceleration Process. Nous insistons sur le fait que dans la limite n et s grand, le couplage entre temps d'exploration et espace exploré est entièrement contenu dans la variable d'échelle $\tau = n/s^{d_w}$. De plus, la fonction $f(\tau)$ correspond exactement à la distribution de τ conditionnée à la valeur du territoire exploré s. Par conséquent, la mesure de $f(\tau)$ pour une unique valeur de s est suffisante pour déterminer entièrement le comportement asymptotique de la loi jointe $\sigma(s,n|s_0)$.

Le résultat (16) présente un intérêt théorique en soi, mais permet également de finement caractériser les propriétés de persistance de certaines marches aléatoires complexes. En effet, dans certains cas, les fonctions $h(s_0)$ et $f(\tau)$ peuvent être déterminées explicitement, et la loi jointe donne alors accès aux distributions marginales asymptotiques du territoire exploré et du FPT en 0. À titre d'illustration, nous avons considéré le cas de la Self Avoiding True Walk (SATW), un modèle de marche à renforcement non-Markovien récemment utilisé pour décrire la dynamique de cellules [d'Alessandro *et al.* 2021]. Dans ce modèle, une particule évolue sur un réseau 1D, par sauts sur plus proches voisins. Si l'un des sites voisins n'a jamais été visité, il est choisi avec une probabilité β ($0 < \beta < 1$). Autrement, si les deux sites voisins ont déjà été visités, la particule effectue un saut symétrique. La dynamique stochastique dépend donc de la totalité du territoire visité précédemment, ce qui complique l'évaluation du temps de premier passage en zéro. En particulier, si l'exposant de persistance $\theta = (1-\beta)/(2\beta)$ a été récemment obtenu [Barbier-Chebbah *et al.* 2020], le préfacteur exact de la distribution asymptotique du FPT reste à déterminer. En calculant directement la loi jointe σ pour la SATW, nous avons obtenu le comportement asymptotique exact de $F_{\underline{0}}(n|s_0 = 1)$:

$$F_{\underline{0}}(n|s_0=1) \underset{n\to\infty}{\sim} \frac{\Gamma(\frac{2}{\beta}-1)}{\Gamma(\frac{1}{2\beta}-\frac{1}{2})} 2^{-\frac{1+\beta}{2\beta}} n^{-\frac{1-\beta}{2\beta}-1}, \tag{17}$$

valable pour toute valeur de β, ce qui illustre la pertinence de la loi jointe σ dans le cadre de problèmes de premier passage de processus à renforcement.

Naturellement, la distribution jointe des variables d'espace s et de temps n pour des processus invariants d'échelle en dimension d_f plus grande que 1 est une observable primordiale pour mieux comprendre le processus d'exploration avant atteinte d'une cible dans des situations physiques réelles. En combinant les méthodes utilisées pour déterminer le comportement asymptotique de $C(s_0)$ pour $d_f > 1$ et celui de $\sigma(s, n|s_0)$ dans le cas $1D$, nous nous attendons à identifier de nouveaux comportements universels, mais, à ce stade, il s'agit encore d'une question ouverte.

0.3 Processus de saut unidimensionnels en confinement

Dans cette seconde partie, nous ne nous focalisons plus sur une unique observable, mais sur un processus stochastique en particulier: les processus de saut unidimensionnels en confinement.

Un processus de saut unidimensionnel, ou *jump process* en anglais, est un processus stochastique discret en temps et continu en espace, décrivant l'évolution sur $\mathbb{R}$ de la position x_n d'une particule issue de x_0. La dynamique de la particule est Markovienne, et sa position à l'instant $n+1$ est donnée par $x_{n+1} = x_n + \xi_n$, où les variables ξ_i sont *i.i.d.* de distribution commune $p(\ell)$. Les processus de saut sont des modèles transverses, et apparaissent aussi bien en théorie des files d'attente [Asmussen 2003] et en modélisation financière qu'en biologie, pour décrire par exemple la dynamique saccadée des bactéries chemotactiques [Koshland 1980]. Nous soulignons que ceux-ci sont particulièrement pertinents pour modéliser des données réelles, expérimentales ou numériques, tandis qu'une description continue ne permet pas de capturer les effets propres à la nature discrète des jeux de données.

Dans cette thèse, nous nous focalisons principalement sur les processus de saut symétriques et continus. Plus précisément nous avons considéré l'ensemble des processus dont la transformée de Fourier [1] $p(k) = \int_{-\infty}^{\infty} e^{ik\ell} p(\ell) \mathrm{d}\ell$ se comporte à petit k comme

$$p(k) \underset{k\to 0}{=} 1 - (a_\mu |k|)^\mu + o(k^\mu), \tag{18}$$

et où le paramètre $\mu \in]0, 2]$, est appelé exposant de Levy, et caractérise les queues de la distribution $p(\ell)$. En particulier, dans le cas $\mu < 2$, $p(\ell)$ décroit algébriquement : $p(\ell) \propto \ell^{-(1+\mu)}$. Le paramètre a_μ correspond quant à lui à l'échelle de longueur caractéristique du processus.

Considérons tout d'abord le cas de processus de saut en milieu infini, qui peuvent évoluer sans contraintes sur $\mathbb{R}$ tout entier. La distribution de la position de la particule après n pas est donnée par le propagateur infini $G_\infty(x, n|x_0)$, qui s'écrit explicitement en transformée de Fourier et fonction génératrice [Hughes 1995]:

$$G_\infty(k, \xi) = \sum_{n=0}^{\infty} \xi^n \left[\int_{-\infty}^{\infty} e^{ik(x-x_0)} G_\infty(x, n|x_0) \mathrm{d}x \right] = \frac{1}{1 - \xi p(k)}. \tag{19}$$

[1] Pour alléger les notations, la Transformée de Fourier est uniquement indiquée par l'utilisation de la variable k.

Si l'absence de contraintes sur la dynamique de la particule permet d'exprimer facilement le propagateur infini en fonction de $p(k)$ (équation (19)), qu'en est-il des processus de saut semi-infinis, dont l'évolution s'arrête dès lors qu'ils traversent 0 ? En fait, dans ce cas, le propagateur semi-infini $G(x, n|x_0)$, correspondant à la probabilité que la particule soit en x après n pas sans jamais avoir traversé 0, a aussi été calculé exactement[Ivanov 1994] :

$$\sum_{n=0}^{\infty} \xi^n \left[\int_0^{\infty} \int_0^{\infty} e^{-sx-s_1 x_0} G(x, n|x_0) \mathrm{d}x \mathrm{d}x_0 \right] = \frac{G_0(s, \xi) G_0(s_1, \xi)}{s + s_1}, \tag{20}$$

où $G_0(s, \xi)$ est défini par

$$G_0(s, \xi) = \exp\left[-\frac{s}{2\pi} \int_{-\infty}^{\infty} \frac{\log\left[1 - \xi p(k)\right]}{s^2 + k^2} \mathrm{d}k \right]. \tag{21}$$

En outre, le résultat (20) permet d'obtenir la probabilité de survie en milieu semi-infini $q(x_0, n)$, définie comme la probabilité que la particule reste positive durant ses n premiers pas :

$$\sum_{n=0}^{\infty} \xi^n \left[\int_0^{\infty} e^{-sx_0} q(x_0, n) \mathrm{d}x_0 \right] = \frac{1}{s\sqrt{1-\xi}} \exp\left[-\frac{s}{2\pi} \int_{-\infty}^{\infty} \frac{\log\left[1 - \xi p(k)\right]}{s^2 + k^2} \mathrm{d}k \right]. \tag{22}$$

En particulier, la probabilité de survie $q(0, n)$ partant de $x_0 = 0$ est donnée par $q(0, n) = \binom{2n}{n} 2^{-2n}$, et est complètement indépendante de la distribution de saut $p(\ell)$. Ce résultat combinatoire important, connu sous le nom de théorème de Sparre Andersen [Andersen 1954] souligne précisément les effets discrets des processus de saut. En effet, pour tout processus continu en temps et non-smooth, la probabilité de survie $q^{(c)}(x_0, t)$ jusqu'au temps t est identiquement nulle pour $x_0 = 0$: $q^{(c)}(0, t) = 0$.

Intéressons nous maintenant aux processus de saut confinés dans un intervalle $[0, x]$, et tués (*ie* stoppés) dès la première sortie. Le propagateur confiné correspondant $G_{[0,x]}(u, n|x_0)$ satisfait une équation intégrale obtenue en partitionnant sur la position du dernier pas :

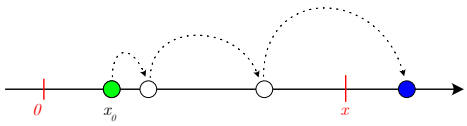
$$G_{[0,x]}(u, n|x_0) = \delta_{0,n}\delta(u - x_0) + [1 - \delta_{0,n}] \int_0^x p(u - y) G_{[0,x]}(y, n-1|x_0) \mathrm{d}y. \tag{23}$$

À l'inverse des cas infini et semi-infini, cette équation intégrale n'est pas soluble pour une distribution de saut arbitraire $p(\ell)$. En réalité, il s'avère qu'il existe *très peu* de résultats généraux pour des processus de saut confinés, principalement à cause de la structure des équations intégrales mises en jeu. En particulier, l'observable associée à la direction de sortie de l'intervalle n'est pas connue pour un processus de saut quelconque.

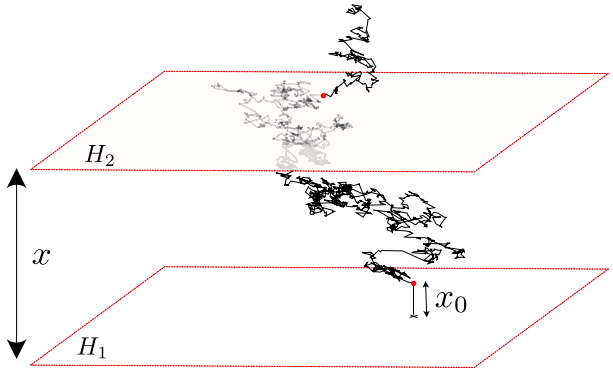
0.3.1 Sortie d'un intervalle

Nous nous sommes d'abord intéressés à la probabilité de splitting $\pi_{0,\underline{x}}(x_0)$ associée à un processus de saut arbitraire, et définie comme la probabilité de quitter l'intervalle par la droite, *ie* à travers x (voir figure 4(a)). Celle-ci satisfait une équation intégrale similaire à l'équation (23), obtenue en partitionnant sur la position du dernier pas :

$$\pi_{0,\underline{x}}(x_0) = \int_x^{\infty} p(y - x_0) \mathrm{d}y + \int_0^x \pi_{0,\underline{x}}(y) p(y - x_0) \mathrm{d}y. \tag{24}$$



(a) Trajectoire typique d'un processus de saut contribuant à $\pi_{0,\underline{x}}(x_0)$. Après 2 sauts à l'intérieur de l'intervalle, la particule issue de x_0 s'échappe de l'intervalle à travers x. Nous soulignons que ni le temps auquel la particule ne s'échappe ni la position d'arrivée ne sont contenus dans $\pi_{0,\underline{x}}(x_0)$.



(b) La diffusion de photons dans un milieu hétérogène est modélisée par un processus de saut isotrope en 3 dimensions. Dans la géométrie parallélépipédique, la probabilité de transmission, définie comme la probabilité que la particule travers H_2 avant H_1, correspond exactement à la probabilité de splitting du processus de saut projeté selon la direction x.

Figure 4

Sauf dans le cas de quelques distributions de saut spécifiques, par exemple $p(\ell) = \frac{\gamma}{2} e^{-\gamma|\ell|}$, la probabilité de splitting ne semble pas pouvoir être calculée à partir de l'équation (24). Afin d'obtenir des résultats universels, indépendants de la forme précise de $p(\ell)$, nous nous sommes intéressés à la limite de grand intervalle $x \to \infty$. Dans cette limite, nous avons montré que la probabilité de splitting $\pi_{0,\underline{x}}(x_0)$ est asymptotiquement équivalente à la probabilité de survie $q(x_0, n)$ évaluée à un instant $n{=}n_x$ correspondant au nombre typique de pas pour s'approcher de x, qui ne dépend que de x, μ et a_μ, et que nous avons explicitement déterminé. La probabilité de splitting prend alors la forme universelle suivante :

$$\begin{aligned} \lim_{x\to\infty} \left[\frac{\pi_{0,\underline{x}}(x_0)}{A_\mu(x)}\right] &= \frac{1}{\sqrt{\pi}} + V(x_0) \\ A_\mu(x) &= \left(\frac{a_\mu}{x}\right)^{\mu/2} 2^{\mu-1}\Gamma\left(\frac{1+\mu}{2}\right), \end{aligned} \tag{25}$$

où la fonction $V(x_0)$ est explicite, et s'exprime uniquement en fonction de $p(\ell)$. Nous soulignons le fait que le résultat asymptotique (25) est valable pour toutes les valeurs de x_0 telles que $0 \leq x_0 \ll x$, et capture précisément les effets discrets qui apparaissent dans le régime $0 \leq x_0 \lesssim a_\mu$, et qui ne peuvent être décrits par une approche continue. En particulier, nous soulignons un résultat important : la probabilité de splitting partant de $x_0 = 0$ est strictement non-nulle, et s'écrit simplement en fonction de a_μ, μ et x

$$\pi_{0,\underline{x}}(0) \underset{x\to\infty}{\sim} \frac{2^{\mu-1}}{\sqrt{\pi}}\Gamma\left(\frac{1+\mu}{2}\right)\left(\frac{a_\mu}{x}\right)^{\frac{\mu}{2}}. \tag{26}$$

En d'autres termes, $\pi_{0,\underline{x}}(0)$ ne dépend pas des détails de la distribution de saut $p(\ell)$ mais uniquement de son comportement à grand ℓ, et peut donc être facilement évaluée. Enfin, le calcul

de $\pi_{0,\underline{x}}(x_0)$ présente également un intérêt applicatif, notamment dans le cadre d'expériences de diffusion de photons en milieu hétérogène [Baudouin *et al.* 2014, Araújo *et al.* 2021], où la probabilité de splitting s'identifie naturellement à la probabilité de transmission des photons à travers un parallélépipède en 3 dimensions (voir figure 4(b)).

Pour caractériser plus finement les événements de sortie d'un intervalle, nous avons ensuite introduit la probabilité $F_{\underline{0},x}(n|x_0)$ de première sortie à gauche (LETP pour *leftward exit-time probability* en anglais), définie comme la probabilité que la particule traverse 0 pour la première fois à l'instant n, sans avoir quitté l'intervalle auparavant, et son homologue à droite, la RETP $F_{0,\underline{x}}(n|x_0)$. À nouveau, nous insistons sur le fait que l'équation intégrale qui définit la RETP

$$F_{0,\underline{x}}(n|x_0) = [1-\delta_{n,1}] \int_0^x p(y-x_y) F_{0,\underline{x}}(n-1|y)\mathrm{d}y + \delta_{n,1} \int_x^\infty p(y-x_0)\mathrm{d}y \tag{27}$$

n'est a priori pas soluble pour une distribution de saut arbitraire $p(\ell)$. En considérant la variable rescalée $\tau = \left[\frac{a_\mu}{x}\right]^\mu n$, et en exploitant les résultats nouveaux obtenus pour la probabilité de splitting, nous avons caractérisé pour la première fois le comportement à grand n et grand x de la RETP, qui prend une forme de scaling universelle

$$F_{0,\underline{x}}(n|x_0) \underset{\substack{x\to\infty \\ n\to\infty \\ \tau \text{ fixé}}}{\sim} \pi_{0,\underline{x}}(x_0) h_\mu(\tau) n^{-1}, \tag{28}$$

où $h_\mu(\tau)$ est une fonction d'échelle qui ne dépend que de μ. Nous insistons sur le fait que la dépendance en x_0 de la RETP dans le régime $0 \le x_0 \lesssim a_\mu$ est entièrement contenue dans la probabilité de splitting, et les effets discrets liés à la nature des processus de saut sont correctement décrits. En particulier, la RETP $F_{0,\underline{x}}(n|0)$ partant de $x_0 = 0$ est strictement positive pour tout n et x. De la même manière, nous avons montré que la LETP se réécrit en fonction de la distribution connue $F_{\underline{0}}(n|x_0)$ du temps de premier passage en 0 dans la géométrie semi-infinie

$$F_{\underline{0},x}(n|x_0) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} F_{\underline{0}}(n|x_0) g_\mu(\tau), \tag{29}$$

où g_μ est une fonction d'échelle ne dépendant que de μ, et la dépendance en x_0 de $F_{\underline{0},x}(n|x_0)$ est à nouveau correctement caractérisée dans tous les régimes $0 \le x_0 \ll x$. De plus, nous avons évalué h_μ et g_μ exactement pour $\mu = 2$, et asymptotiquement dans les régimes $\tau \ll 1$ et $\tau \gg 1$ pour $\mu < 2$. Nous résumons nos résultats dans la table 1 :

	$F_{0,\underline{x}}(n\|x_0) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixé}}}{\sim} \pi_{0,\underline{x}}(x_0) h_\mu(\tau) n^{-1}$	$F_{\underline{0},x}(n\|x_0) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixé}}}{\sim} F_{\underline{0}}(n\|x_0) g_\mu(\tau)$
$\mu = 2$	$h_2(\tau) = 2\tau\pi^2 \sum_{k=1}^\infty k^2(-1)^{k+1} e^{-k^2\pi^2\tau}$	$g_2(\tau) = 4\pi^{\frac{5}{2}}\tau^{\frac{3}{2}} \sum_{k=1}^\infty e^{-k^2\pi^2\tau} k^2$
$\mu < 2$	$h_\mu(\tau) \underset{\tau\ll 1}{\sim} \Gamma(\mu/2)\sin(\pi\mu/2)\pi^{-\frac{3}{2}}\sqrt{\tau}$	$g_\mu(\tau) \underset{\tau\ll 1}{\sim} 1$
	$h_\mu(\tau) \underset{\tau\gg 1}{\sim} \gamma_\mu \left[\lambda_1 2^\mu \tau\right] e^{-\lambda_1 2^\mu \tau}$	$g_\mu(\tau) \underset{\tau\gg 1}{\sim} \omega_\mu \left[\lambda_1 2^\mu \tau\right]^{3/2} e^{-\lambda_1 2^\mu \tau}$

Table 1

où les constantes γ_μ, ω_μ et λ_1 sont explicitées au chapitre 5. Nous insistons sur le fait que la table 1 fournit une description complète du comportement jusqu'alors inconnu de la RETP et LETP à grand n et grand x, et valable pour tout les processus de saut symétriques. En particulier, nos résultats caractérisent complètement le comportement spécifique des processus de saut à petit x_0, qui n'est pas décrit par une modélisation continue, et qui n'a pu être obtenu qu'à partir de la détermination de la probabilité de splitting $\pi_{0,\underline{x}}(x_0)$.

0.3.2 Statistiques d'extrêmes des processus de saut

Le calcul de la probabilité de splitting, de la RETP et de la LETP s'inscrit dans le cadre plus large des statistiques d'extrêmes (EVS de l'anglais *extreme value statistics*) des processus de saut, qui concernent les distributions jointes d'extremums et de temps d'atteinte d'extremums.

Les EVS de marches aléatoires ont été principalement étudiées dans le cas du mouvement Brownien. À titre d'exemple, les distributions du maximum $m(t)$ atteint par un Brownien libre de durée t, ainsi du temps d'atteinte t_m de ce maximum ont été initialement calculées par Paul Lévy [Levy 1937]; la distribution jointe de $m(t)$ et t_m a été étudiée dans [Majumdar *et al.* 2008, Mori *et al.* 2021] pour le Brownien libre et, entre autres, le pont Brownien. Citons également la distribution du temps d'atteinte du maximum avant le FPT en 0, qui a été calculée dans [Randon-Furling & Majumdar 2007].

Concernant les EVS de processus de saut généraux, seule les distributions $\mu(x|n,x_0)$ du maximum x atteint jusqu'au pas n, et $\rho(n_m|n,x_0)$ du temps d'atteinte n_m de ce maximum dans la géométrie infinie ont été obtenues [Majumdar 2010]. Dans ce contexte, nous avons calculé la distribution jointe $\rho_1(x,n_m|n,x_0=0)$ du maximum x et du temps d'atteinte n_m, qui s'écrit simplement comme $\rho_1(x,n_m|n,x_0=0) = G_0(x,n_m)q(0,n-n_m)$, où G_0 est le propagateur semi-infini partant de 0, défini précédemment (21). De plus, nous avons montré que cette loi jointe admet un comportement asymptotique universel qui ne dépend que de l'exposant de Levy μ. Soulignons en particulier le fait que la simplicité de ce résultat découle du caractère Markovien et discret en temps des processus de saut, qui permet un découpage trivial de la trajectoire autour du maximum x.

De manière générale, cette dernière remarque suggère que les statistiques d'extrêmes de processus de saut semi-infinis, pour lesquels il n'existe pas de résultats généraux, peuvent s'obtenir à partir d'un découpage similaire. En particulier, nous avons introduit une nouvelle brique élémentaire nécessaire à la description de trajectoires stoppées lors du premier passage en 0 : la *probabilité de strip* $\mu_{0,\underline{x}}(n)$, définie comme la probabilité que la particule partant de 0 atteigne son maximum x au $n^{\text{ième}}$ pas exactement, tout en restant positive durant ses n premiers pas (voir figure 5).

Bien que l'expression explicite de $\mu_{0,\underline{x}}(n)$ dépende du propagateur borné $G_{[0,x]}(u,n|x_0=0)$ et n'est donc pas calculable pour un processus de saut arbitraire, nous avons montré que la probabilité de strip prend une forme asymptotique universelle dans la limite n grand et x grand. Pour les processus de saut avec $\mu=2$, et en définissant $\tau = n/(x/a_2)^2$, la probabilité de strip s'écrit

$$\mu_{0,\underline{x}}(n) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixé}}}{\sim} 2\left[\frac{a_2}{x}\right]^2 \frac{1}{x}\pi^2 \sum_{k=1}^{\infty} k^2(-1)^{k+1} e^{-k^2\pi^2\tau}. \tag{30}$$

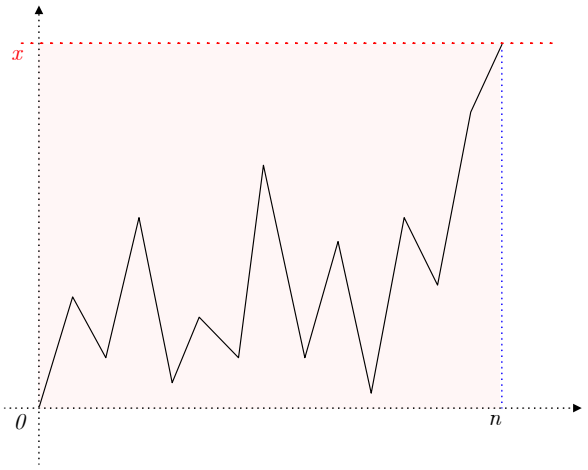


Figure 5: Trajectoire typique contribuant à la probabilité de strip $\mu_{0,\underline{x}}(n)$. En exactement n pas, la particule traverse la bande $[0, x]$, sans s'en échapper.

Pour les processus de saut avec $\mu < 2$, et en définissant $n_x = (x/a_\mu)^\mu$, $\mu_{0,\underline{x}}(n)$ prend la forme asymptotique suivante :

$$\mu_{0,\underline{x}}(n) \underset{1\ll n\ll n_x}{\sim} \frac{\mu}{\pi}\,\Gamma(\mu)\sin\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}. \tag{31}$$

En particulier, nous obtenons un résultat contre-intuitif : dans la limite $1 \ll n \ll n_x$, $\mu_{0,\underline{x}}(n)$ devient indépendante de n. Une multitude d'observables d'EVS pour les processus de saut peuvent alors être systématiquement évaluées en fonction des briques élémentaires que sont la probabilité de survie semi-infinie $q(x_0, n)$, la probabilité de splitting $\pi_{0,\underline{x}}(x_0)$, les LETP et RETP ainsi que la probabilité de strip $\mu_{0,\underline{x}}(n)$. La table 2 présente de manière synthétique certains des résultats obtenus dans le cas de processus généraux.

En conclusion, nous insistons sur le fait que les résultats présentés dans la table 2 ne sont que des exemples représentatifs d'observables d'extrêmes calculées au chapitre 6, et illustrent le rôle clé des nouvelles observables que nous avons introduites pour caractériser les statistiques d'extrêmes des processus de saut généraux.

0.3.3 Processus de saut isotropes dans la limite de grand volume

Les résultats asymptotiques obtenus pour les observables de premier passage (comme $\pi_{0,\underline{x}}(x_0)$ ou $F_{0,\underline{x}}(n|x_0)$) associées aux processus de saut confinés en une dimension peuvent en réalité s'étendre aux processus de saut isotropes en dimension plus grande que 1. Plus précisément, nous nous sommes intéressés au cas des processus de saut isotropes et continus, caractérisés par une distribution de saut $p(\mathbf{l})$ et dont le comportement de la transformée de Fourier à petit vecteur d'onde $\mathbf{k}$ s'écrit

$$p(\mathbf{k}) \underset{\mathbf{k}\to 0}{=} 1 - |a_\mu \mathbf{k}|^\mu + o(|\mathbf{k}|^\mu). \tag{32}$$

Comme dans le cas unidimensionnel, l'exposant de Levy μ caractérise les queues de la distribution, et a_μ est l'échelle de longueur caractéristique. Pour un processus de saut isotrope arbitraire en

observable	Expression exacte	$\mu < 2$ - Comportement asymptotique dans le régime $1 \ll n \ll n_x$	$\mu = 2$ - Comportement asymptotique dans le régime de scaling n/n_x fixé
Processus de saut en milieu infini			
$\mu(x\|n, x_0 = 0)$	$\frac{\mathrm{d}}{\mathrm{d}x} q(x, n)$	$\frac{n\mu}{\pi} \sin\left(\frac{\pi\mu}{2}\right) \Gamma(\mu) \left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}$	$\frac{1}{a_2\sqrt{\pi n}} e^{-\left[\frac{x}{a_2}\right]^2 \frac{1}{4n}}$ o
$\rho(n_m\|n, x_0 = 0)$	$q(0, n_m) q(0, n - n_m)$	$\frac{1}{\pi} \frac{1}{\sqrt{n_m(n-n_m)}}$ o	
$\rho_1(x, n_m\|n, x_0 = 0)$	$G_0(x, n_m) q(0, n - n_m)$	$\frac{1}{\pi} \sqrt{\frac{n_m}{n-n_m}} \frac{2\mu}{\pi} \sin\left(\frac{\pi\mu}{2}\right) \Gamma(\mu) \left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}$	$\frac{x}{a_2} \frac{1}{2a_2\pi n_m^{3/2}\sqrt{n-n_m}} e^{-\left[\frac{x}{a_2}\right]^2 \frac{1}{4n_m}}$
$\rho_2(x, n_m, x_f\|n, 0)$	$G_0(x, n_m) G_0(x - x_f, n - n_m)$	$\sqrt{n_m(n-n_m)} \frac{4\mu^2}{\pi^3} \sin^2\left(\frac{\pi\mu}{2}\right) \Gamma^2(\mu) \left[\frac{a_\mu^2}{x(x-x_f)}\right]^\mu \frac{1}{x(x-x_f)}$	$\frac{x}{a_2} \frac{x-x_f}{a_2} \frac{1}{4(a_2)^2\pi(n_m(n-n_m))^{3/2}} e^{-\left[\frac{x}{a_2}\right]^2 \frac{1}{4n_m} - \left[\frac{x-x_f}{a_2}\right]^2 \frac{1}{4(n-n_m)}}$
Processus de saut en milieu semi-infini			
$\mu(x\|x_0 = 0)$	$-\frac{\mathrm{d}}{\mathrm{d}x} \pi_{0,\underline{x}}(0)$	$\mu 2^{\mu-2} \Gamma\left(\frac{1+\mu}{2}\right) \frac{1}{\sqrt{\pi}} \left[\frac{a_\mu}{x}\right]^{\frac{\mu}{2}} \frac{1}{x}$	
$\rho_3(x, n_m\|x_0 = 0)$	$\mu_{0,\underline{x}}(n_m) \pi_{0,\underline{x}}(0)$	$\frac{2^{\mu-1}\ \mu\ \Gamma(\frac{1+\mu}{2})\ \Gamma(\mu)\ \sin(\frac{\pi\mu}{2})}{\pi^{\frac{3}{2}}} \left[\frac{a_\mu}{x}\right]^{\frac{3\mu}{2}} \frac{1}{x}$	$2\left[\frac{a_2}{x}\right]^3 \frac{1}{x}\pi^2 \sum_{k=1}^{\infty} k^2 (-1)^{k+1} e^{-k^2\pi^2 n_m \left[\frac{a_2}{x}\right]^2}$
$\rho_4(x, n_m, n_f\|x_0 = 0)$	$\mu_{0,\underline{x}}(n_m) F_{0,\underline{x}}(n_f - n_m\|0)$	$\frac{\mu}{\sqrt{\pi(n_f-n)}} \left[\frac{\Gamma(\mu)}{\pi} \sin\left(\frac{\pi\mu}{2}\right)\right]^2 \left[\frac{a_\mu}{x}\right]^{2\mu} \frac{1}{x}$	$4\left[\frac{a_2}{x}\right]^5 \frac{1}{x}\pi^4 \sum_{k,l=1}^{\infty} k^2 l^2 (-1)^{k+l} e^{-k^2\pi^2 n_m \left[\frac{a_2}{x}\right]^2 - l^2\pi^2 (n_f - n_m)\left[\frac{a_2}{x}\right]^2}$
$\sigma(x, n_f\|x_0 = 0)$	$-\frac{\mathrm{d}}{\mathrm{d}x} F_{\underline{0},x}(n_f\|0)$	$\frac{2\mu\sqrt{n_f}}{\pi^{\frac{3}{2}}} \Gamma^2(\mu) \sin^2\left(\frac{\pi\mu}{2}\right) \left[\frac{a_\mu}{x}\right]^{2\mu} \frac{1}{x}$	$2\pi^2 \left[\frac{a_2}{x}\right]^3 \frac{1}{x} \sum_{k=1}^{\infty} e^{-k^2\pi^2 n_f \left[\frac{a_2}{x}\right]^2} k^2 \left[2k^2\pi^2 n_f \left[\frac{a_2}{x}\right]^2 - 3\right]$

Table 2: Statistiques d'extrêmes pour des processus de saut en milieu infini d'une durée de n pas, et processus de saut semi-infinis stoppés dès le premier passage dans $\mathbb{R}^*_-$. La variable x correspond au maximum, n_m au temps d'atteinte du maximum et n_f au temps de premier passage à travers 0 (dans le cas de processus semi-infinis). Les expressions exactes sont valables pour tous les processus de saut symétriques et continus, et $n_x = (a/a_\mu)^\mu$. Les comportements marqués d'un o sont classiques, et connus dans la littérature.

confinement, nous avons considéré le cas général d'une observable $\mathcal{B}(\mathbf{x_0})$ *de premier passage*, définie comme une observable évaluée lors de la sortie du domaine confinant $\mathcal{D}$ de longueur caractéristique R, dans la limite de grand volume $R \gg a_\mu$ (voir figure 6). Dans le cas le plus simple, nous considérons que la particule sort du domaine dès le premier passage à travers la surface confinante Σ.

À titre d'exemple, $\mathcal{B}(\mathbf{x_0})$ peut aussi bien correspondre au temps de première sortie d'une enceinte quelconque, qu'à la position du processus de saut lors de sa sortie. Il est important de souligner que l'évaluation de telles observables en dimension plus grande que 1 présente un intérêt applicatif fort, que ce soit dans le contexte de la diffusion de photons avec $\mu < 2$ [Baudouin *et al.* 2014] ou $\mu = 2$ [Savo *et al.* 2017], ou encore pour mieux décrire certains mouvements bactériaux modélisés par des processus de type *Run and Tumble* [Van Kampen 1992, Patteson *et al.* 2015].

Dans la limite de grand volume confinant, l'approche classique pour évaluer $\mathcal{B}(\mathbf{x_0})$ consiste à identifier le processus limite continu en temps associé au processus de saut. Dans le cas $\mu = 2$, le mouvement de la particule à grand temps est asymptotiquement décrit par un mouvement Brownien [Redner 2001], et si $\mu < 2$, le processus limite est un processus α-stable isotrope de paramètre μ [Kyprianou & Pardo 2022]. Par conséquent, dans la limite de grand volume, et pour une position de départ $\mathbf{x_0}$ positionnée loin du bord absorbant,

$$P(\mathcal{B}(\mathbf{x_0}) = b) \underset{a_\mu \ll d_e \ll R}{\sim} P(\mathcal{B}^{(c)}(\mathbf{x_0}) = b), \tag{33}$$

où $\mathcal{B}^{(c)}(\mathbf{x_0})$ correspond à la même observable, évaluée pour le processus limite. En particulier, la distribution de $\mathcal{B}^{(c)}(\mathbf{x_0})$ est indépendante des détails du processus de saut, et, par conséquent, est souvent plus simple à calculer. Néanmoins, si la distance initiale d_e de la particule au bord absorbant Σ de $\mathcal{D}$ (voir figure 6) est petite par rapport à l'échelle de longueur a_μ du processus , la relation $\mathcal{B}(\mathbf{x_0}) \sim \mathcal{B}^{(c)}(\mathbf{x_0})$ n'est plus valable. En effet, de nombreuses trajectoires sont stoppées

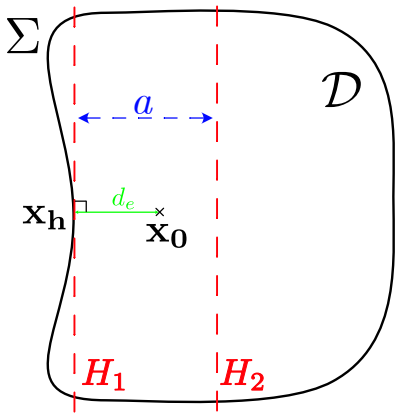


Figure 6: **Schéma de la méthode de *raccordement***. Le processus de saut isotrope part de x_0, situé à une distance d_e de l'enceinte Σ. Nous considérons ici une observable $\mathcal{B}(x_0)$ évalué lors de la première sortie de $\mathcal{D}$, c'est à dire dès que la particule traverse Σ. Les hyperplans H_1 et H_2 sont tangents à Σ en $\mathbf{x_h}$. Si la particule traverse H_2 avant H_1, l'observable $\mathcal{B}(x_0)$ est alors asymptotiquement équivalente à sa version continue $\mathcal{B}^{(c)}(x_0)$.

trop rapidement, avant que la limite continue ne soit atteinte, et les effets discrets du processus de saut ne sont plus négligeables. Notons que le cas $\mathcal{B}(x_0) = \pi_{0,\underline{x}}(x_0)$ a été traité en détail au chapitre 4; nos résultats montrent bien que l'approche continue n'est plus valable lorsque $x_0 \to 0$ puisque $\pi_{0,\underline{x}}(0)$ est strictement non nulle.

Afin de prendre en compte la présence d'une frontière absorbante proche du point de départ, nous avons développé une méthode de *raccordement* pour déterminer $\mathcal{B}(\mathbf{x_0})$ dans la limite de grand volume. En introduisant deux hyperplans H_1 et H_2 distants de a, avec $a \gg a_\mu$ au voisinage de la position initiale $\mathbf{x_0}$ (voir figure 6), nous avons montré que la distribution de l'observable $\mathcal{B}(\mathbf{x_0})$ se réécrit

$$P(\mathcal{B}(\mathbf{x_0}) = b) \underset{\substack{a_\mu \ll a \ll R \\ a_\mu \ll R_c}}{\sim} \pi^{\perp}_{0,\underline{a}}(d_e) \lim_{\mathbf{u} \to \mathbf{x_h}} \left[\frac{P(\mathcal{B}^{(c)}(\mathbf{u}) = b)}{\pi^{(c)}_{0,\underline{a}}(\mathbf{u})} \right], \tag{34}$$

où $\mathbf{x_h}$ est le projeté orthogonal de $\mathbf{x_0}$ sur Σ, R_c le rayon de courbure en $\mathbf{x_h}$ et $\pi^{\perp}_{0,\underline{a}}(d_e)$ est la probabilité de splitting du processus de saut associé au mouvement perpendiculaire à H_1 et H_2. Nous insistons sur le fait que le comportement discret de la particule est entièrement contenu dans la probabilité de splitting $\pi^{\perp}_{0,\underline{a}}(d_e)$, que nous avons complètement caractérisée dans la limite $a \gg a_\mu$ au chapitre 4. La méthode de raccordement est en particulier valable pour une particule partant strictement du bord de l'enceinte, puisque $\pi^{\perp}_{0,\underline{a}}(0)$ est strictement non nulle. Notons enfin que le résultat (34) est transparent dans le cas $\mathcal{B}(x_0) = F_{0,\underline{x}}(n|x_0)$: la dépendance en x_0 de $F_{0,\underline{x}}(n|x_0)$, donnée dans la table 1, est directement contenue dans la probabilité de splitting $\pi_{0,\underline{x}}(x_0)$.

Ensemble, les équations (33) et (34) permettent d'évaluer la distribution de $\mathcal{B}(\mathbf{x_0})$ dans la limite de grand volume $R \gg a_\mu$, et pour toutes conditions initiales $\mathbf{x_0}$, à condition que celle de l'observable $\mathcal{B}^{(c)}(\mathbf{u})$ associée au processus limite continu soit calculable. En pratique cette condition est satisfaite pour de nombreuses observables, aussi bien spatiales que temporelles. Pour des processus planaires avec $\mu = 2$ partant de la surface absorbante d'un disque, nous avons calculé la mesure harmonique, définie comme la distribution de l'angle de sortie du disque, ainsi

que la distribution exacte du temps de sortie du disque dans la limite de grand volume. Ces résultats, inaccessibles par une description purement brownienne, sont entièrement nouveaux et directement applicables aux modèles de particules de type Run and Tumble par exemple. Dans le cas $\mu < 2$, notre méthode de raccordement nous a permis de calculer entre autres (voir chapitre 7) le temps moyen de sortie d'un intervalle $[0, x]$ partant de $x_0 = 0$, et ainsi de compléter l'analyse asymptotique des temps de sorties $F_{0,\underline{x}}(n|x_0)$ et $F_{\underline{0},x}(n|x_0)$ effectuée au chapitre 5

En conclusion, nos résultats, valables dans la limite de grand volume confinant, permettent de déterminer de manière explicite un grand nombre d'observables associées à un événement de premier passage, pour des processus de saut quelconques en dimension supérieure ou égale à 1, à partir du calcul de la même observable pour un processus continu. Par conséquent, les nombreux résultats concernant les temps de premier passage et probabilités de splitting de mouvement browniens [Bénichou & Voituriez 2014] ou de processus α-stables [Getoor 1961, Kyprianou *et al.* 2014] s'adaptent naturellement aux processus de saut isotropes.

0.4 Conclusion

La description et l'interprétation d'un phénomène physique par un problème de transport aléatoire s'organise autour de deux axes : la détermination d'un processus stochastique rendant compte de la dynamique aléatoire empirique, et le choix d'une observable statistique répondant à une question physique précise. Ainsi, la propagation d'une épidémie peut être modélisée par une marche aléatoire sur un graphe (processus), et le temps de contamination complète de la population correspond alors au temps mis par le marcheur pour visiter l'intégralité des nœuds (observable). Considérons un autre exemple : la longueur d'une file d'attente peut être décrite par un processus de Poisson (processus); la probabilité que le caissier puisse un jour rentrer chez lui est alors donnée par la probabilité que ce processus atteigne 0 (observable).

S'appuyant sur ce découpage, ce manuscrit s'organise en deux parties, chacune reflétant l'un de ces deux axes. Dans une première partie, nous avons introduit une nouvelle observable, le territoire visité avant l'atteinte d'une cible par un marcheur en confinement, qui quantifie l'efficacité du processus d'exploration de l'espace, et caractérisé quantitativement sa statistique pour la classe de processus stochastiques la plus vaste possible. À l'inverse, dans une seconde partie, nous nous sommes focalisés sur un unique type de processus stochastiques, les processus de saut unidimensionnels en confinement. Bien qu'étudiés depuis longtemps, il existe peu de résultats analytiques les concernant, et nous avons proposé une méthode pour caractériser de nombreuses observables dans la limite de grand volume confinant.

Le nombre de sites distincts visités par un marcheur aléatoire au bout de n pas, qui quantifie la capacité du marcheur à explorer un territoire inconnu, a été majoritairement étudié dans une géométrie infinie. En effet, dans un domaine confiné, celui-ci finit toujours par être entièrement visité. Cependant, il est légitime de se demander quelle fraction d'un domaine fermé est visitée par un marcheur avant de s'échapper par une sortie fixée. Notons que cette question est particulièrement pertinente dans le cadre d'applications en chimie ou biologie par exemple. Afin d'y répondre, nous avons introduit l'observable $C(s_0)$, définie comme le territoire visité par un marcheur aléatoire confiné issu de s_0 avant d'atteindre une cible fixée s_T.

Nous avons tout d'abord considéré la classe des processus Markoviens, unidimensionnels et avec span connecté, c'est à dire ne laissant pas de trous dans leur trajectoire. Cette classe de processus englobe un nombre conséquent de marches aléatoires classiques sur réseau : la marche normale symétrique et biaisée, la marche persistante; mais aussi de processus continus comme le mouvement Brownien avec ou sans resetting. Dans ce cadre nous avons établi une relation exacte entre le territoire visité avant d'atteindre la cible située en 0, et les probabilité de splitting $\pi_{s_1,\underline{s_2}}(s_0)$, définie comme la probabilité que le marcheur atteigne s_2 avant s_1. Ce faisant, nous avons pu déterminer la distribution exacte du territoire visité avant d'atteindre 0 pour un grand nombre de processus.

Pour dépasser le cadre de la dimension 1 et du span connecté, nous avons ensuite considéré la classe plus générale des processus invariants d'échelle, dans la limite de grand volume confinant. En particulier, cette classe rassemble les diffusions sur réseaux fractals déterministes et aléatoires, mais également les processus avec saut à longue portée, comme les marches de Riemann. En exploitant le comportement asymptotique des probabilités de splitting pour les processus invariants d'échelle dans la limite de grand volume, nous avons identifié un comportement de scaling universel

du premier moment $\langle C(s_0)\rangle$ du territoire exploré, mais également de sa distribution complète. En particulier, nous avons entièrement caractérisé la dépendance du territoire visité dans les paramètres géométriques du système, à savoir la distance source-cible, ainsi que la taille du domaine confinant.

Enfin, il est manifeste que le territoire visité avant de trouver la cible s_T est intrinsèquement lié au temps mis pour l'atteindre. Pour décrire plus finement le processus d'exploration et quantifier les corrélations intuitives entre territoire et temps, nous avons considéré la loi jointe $\sigma(s, n|s_0)$ du territoire visité s et du FPT n en s_T, partant de s_0. Dans le cas de processus 1D Markoviens avec span connecté nous avons élaboré une méthodologie pour calculer explicitement σ, et l'avons appliquée aux différents processus évoqués ci-dessus. De plus, dans la limite s et n grand, nous avons montré que la loi jointe admet une forme de scaling universelle, valable pour tout les processus unidimensionnels, connectés ou non, Markoviens ou non. En particulier, nos résultats sont valables dans le cas de l'emblématique mouvement brownien fractionnaire.

Ces premiers résultats concernant le territoire visité avant l'atteinte d'une cible dans un domaine confiné ouvrent la voie à de nombreuses autres questions. On pourrait par exemple se demander si le comportement universel de la distribution jointe du territoire et du temps est toujours valable en dimension plus grande que 1. Si l'on s'attend à retrouver des classes d'universalité similaires à celles régissant le comportement de $C(s_0)$, cette question reste pour l'instant ouverte. De même, que se passe-t-il si la cible n'est pas fixée, mais est distribuée aléatoirement dans le système, ou si le domaine possède plusieurs sorties ? Ces extensions naturelles sont évidemment passionnantes d'un point de vue théorique, mais s'inscrivent également dans l'optique plus large de la modélisation de phénomènes et systèmes réalistes.

La seconde série de questions qui nous a intéressés concerne un processus stochastique en particulier, les processus de saut unidimensionnels (*jump process* en anglais). Ce modèle décrit l'évolution de la position x_n d'une particule issue de x_0, effectuant un saut aléatoire caractérisé par une distribution $p(\ell)$ à chaque pas de temps discret. Nous insistons sur le fait que ces modèles, introduits dès 1905 par Pearson, sont particulièrement bien adaptés pour décrire des données expérimentales ou numériques, qui sont intrinsèquement discrètes. A contrario, une description continue, comme le mouvement brownien, ne peut capturer les effets spécifiques liés à cette discrétisation. Notons d'abord que les processus de saut en milieu infini sont bien compris. Par exemple, la distribution de la position de la particule au $n^{\text{ème}}$ pas s'écrit explicitement en fonction de la distribution de saut $p(\ell)$. De la même manière, les processus de saut semi-infinis, c'est à dire stoppés dès le premier passage à travers 0, ont également été longuement étudiés; en particulier, la distribution de la position et la probabilité de survie au $n^{\text{ème}}$ pas s'expriment de manière générale pour tout $p(\ell)$ symétrique. En revanche, pour les processus de saut confinés, c'est à dire stoppés dès la première sortie d'un intervalle $[0, x]$, il n'existe pas de résultats généraux valable pour tout $p(\ell)$.

Afin de caractériser plus précisément ces processus confinés, nous nous sommes intéressés aux observables liées à la sortie de l'intervalle, dans la limite de grand intervalle $x \to \infty$. En premier lieu, nous avons considéré l'observable la plus naturelle, la probabilité de splitting $\pi_{0,\underline{x}}(x_0)$ définie comme la probabilité que le processus s'échappe en x plutôt qu'en 0. Dans la limite de grand intervalle, nous avons montré que la probabilité de splitting admet un comportement universel explicite, valable pour toute distribution de saut $p(\ell)$. Il est important de souligner que notre

résultat caractérise complètement la dépendance en x_0 de la probabilité de splitting, et capture précisément les effets spécifiques à la nature discrète du processus. En particulier, nous avons montré que la probabilité de splitting partant de 0 ne dépend que des queues de la distribution $p(\ell)$ et est strictement non-nulle, ce qui n'est pas possible pour un processus continu non-smooth.

Pour obtenir une information plus fine sur les événements de sortie de l'intervalle, nous nous sommes ensuite focalisés sur le temps de sortie à gauche et à droite, en introduisant la probabilité de sortie à gauche (LETP) $F_{\underline{0},x}(n|x_0)$, définie comme la probabilité que le processus sorte de l'intervalle au $n^{\text{ème}}$ pas et à travers 0, et son homologue à droite, la RETP $F_{0,\underline{x}}(n|x_0)$. Dans la limite de grand intervalle, nous avons montré que ces deux quantités admettent également un comportement asymptotique universel, qui ne dépend que des queues de la distribution. À nouveau, nos résultats quantifient précisément les effets discrets des processus de saut, et sont valables dans tout les régimes de position initiale x_0, en particulier $x_0 = 0$.

Enfin, la détermination de la probabilité de splitting ainsi que des LETP et RETP nous a permis d'étendre nos résultats pour les processus de saut dans deux directions. Tout d'abord, nous avons considéré les statistiques d'extrêmes, associées de manière générale à la distribution d'extremums et de temps d'atteinte de ces extremums. L'un de nos résultats principaux a été de montrer que les LETP, RETP et probabilités de splitting constituent des briques élémentaires essentielles à la détermination d'observables d'extrêmes. En particulier, nous avons calculé exactement et asymptotiquement un grand nombre de distributions jointes d'extrêmes et de temps d'atteinte, pour des processus de saut arbitraires. Ensuite, nous avons adapté nos résultats au cas de processus de saut isotropes arbitraires en dimension plus grande que 1. Pour ce faire, nous avons développé une méthodologie valable pour toute distribution de saut isotrope, permettant d'évaluer des observables liées à la sortie d'un domaine confinant, dans la limite de grand volume. À titre d'illustration, nous avons par exemple calculé la distribution du temps d'atteinte du bord d'un disque pour une particule de type Run and Tumble initialement située sur le bord du disque, qui ne peut être obtenue par une approximation brownienne.

En conclusion, nos résultats brossent un tableau complet du comportement asymptotique d'observables associées à la sortie d'un domaine confinant en dimension 1 et plus pour des processus de saut arbitraires. Plus précisément, nos formules se veulent exhaustives : pour un processus quelconque, la détermination du comportement des queues de $p(\ell)$ est suffisant pour évaluer le comportement de l'observable considérée. Pourtant, il reste encore beaucoup de questions sans réponses. Si la plupart de nos résultats sont valables uniquement dans la limite de grand volume confinant, l'évaluation de la première correction à ce comportement asymptotique, à partir du *big jump principle* [Vezzani *et al.* 2019] par exemple, permettrait de caractériser encore plus finement les processus de saut. Une autre direction qu'il nous tarde de creuser est celle de l'asymétrie du processus. Est-il possible d'identifier des comportements universels dans le cas où $p(\ell)$ n'est plus symétrique ? Nous savons que c'est le cas pour la probabilité de survie de processus semi-infinis [Majumdar *et al.* 2012], et espérons que la base méthodologique que nous avons proposée puisse s'étendre au cas asymétrique confiné.

Introduction

Ubiquitous in physical systems, random transport phenomena are observed at all temporal and spatial scales, and across numerous fields, from biology to mechanics, finance or geology. The scattering of photons in a heterogeneous medium [Rosenstock 1961, Savo *et al.* 2017], the motion of a molecular motor along a DNA strand [Berg *et al.* 1981], the folding of a polymeric chain [Lifshitz *et al.* 1978], the evolution of a financial asset [Black & Scholes 1973], the structure of a starling flight [Cavagna *et al.* 2010], the fluctuations of an earthquake hazard [Matthews 2002]... All these phenomena can be described by a random motion, and their understanding is essential to answer important questions in practice. For example, with what probability does a stock price reach a certain threshold? How fast does genetic transcription, responsible for the synthesis of proteins necessary for the proper functioning of the organism, take place? Or what is the distribution of the waiting time between two earthquakes knowing the seismic history of a given region?

Since the end of the 19^{th} century, the scientific community has been working to categorize and model these phenomena with a double objective: extracting relevant statistical properties to interpret empirical results, and making quantitative predictions to answer these practical questions. Two and a half centuries later, we summarize, admittedly brutally, the essential ingredients needed to model and interpret a physical problem of random transport.

(i) The first step is to define the rules of transport accounting for the observed random dynamics, and more precisely to identify the underlying stochastic process that describes the motion of the particle under consideration (where the term "particle" is used generically, whether it refers to a starling or a stock price).

Historically, Brownian Motion was first introduced in the early 1900 by Bachelier [Bachelier 1900] and Einstein [Einstein 1905], to explain the erratic motion of some experimental particles. While Brownian Motion is continuous in both space and time, other parallel mathematical descriptions have emerged. For instance, Pearson [Pearson 1905] developed a model of random walk in discrete time and continuous space, and several years later, Polya [Pólya 1921] investigated random walks on hypercubic networks that are discrete both in time and space. Importantly, although these models are distinct, they also share strong similarities. First, all three models are Markovian, meaning that the stochastic dynamics depend only on the state of the particle at a given time t, and not on its trajectory history. Furthermore, these processes exhibit diffusive behavior, with the particle's mean square displacement (MSD) denoted by $\langle x^2(t)\rangle$ behaving linearly at large times, such that $\langle x^2(t)\rangle \sim t$.

It should be noted that the mean squared displacement of Markovian processes does not always follow a linear behavior. For instance, some trajectories of foraging animals exhibit a superlinear MSD [Reynolds 2006] with $\langle x^2(t)\rangle \propto t^\beta$ and $\beta > 1$. To accurately model such anomalous diffusion phenomena [Bouchaud & Georges 1990], new stochastic processes need to be introduced. One such example is Levy flights [Levy 1937], which are Markovian random walks with power-law distributed increments $p(\ell) \propto \ell^{-(1+\mu)}$ and $\mu \in]0, 2[$, and result in super-diffusive dynamics with a typical large-time displacement $x_n \sim n^{1/\mu}$, in contrast to the diffusive case where $x_n \sim \sqrt{n}$. Conversely, subdiffusive behavior has been observed in experimental studies of glassy

dynamics, and can be properly modeled by considering continuous time random walks (CTRW) [Montroll & Weiss 1965, Monthus & Bouchaud 1996], for which particles wait for a random time τ before each move, following a power-law distribution $p(\tau) \propto \tau^{-(1+\beta)}$, with $\beta \in]0,1[$. In turn, CTRWs have a sublinear MSD $\langle x^2(t)\rangle \propto t^\beta$.

Markovian models often provide a simplified representation of physical phenomena, and allow for explicit analytical results. However, they completely disregard memory effects that could significantly impact the dynamics of the observed particle. In turn, non-Markovian stochastic models, such as the emblematic Fractional Brownian Motion (FBM) [Mandelbrot & Van Ness 1968], can account for such effects. More precisely, the FBM is a non-Markovian Gaussian process characterized entirely by its auto correlation function $\langle x(t_1)x(t_2)\rangle \propto (|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H})$, and is particularly suited to describe anomalous non-Markovian dynamics, such as those observed in semi-flexible polymers for instance ($H < 1/2$). As another example, *reinforcement walks* [Grassberger 2017] model phenomena where the dynamics at time t depend on the geometry of the complete trajectory, and whose n times distribution $(x(t_1), \ldots, x(t_n))$ is not necessarily Gaussian. Such behavior has been evidenced in certain natural systems such as ant motion for instance [Goss *et al.* 1989], where the ant movement is influenced by the geometry of the previously visited territory.

(ii) Once a stochastic process accounting for the observed random motion has been identified, it is crucial to define a statistical *observable*, which is a probabilistic quantity that can be evaluated to draw physical conclusions. Different classes of physical questions require different observables, which can be broadly classified into two main categories.

First, *dynamic* observables describe the statistical properties of the trajectory associated with the random process $x(t)$ after a fixed deterministic time t. Although the mean square displacement $\langle x^2(t)\rangle$ is the most natural dynamic observable, it only accounts for the diffusive or non-diffusive behavior of the particle under consideration, and other dynamic observables can be defined to quantify the efficiency of the exploration process. For instance, in the case of discrete random walks in time and space, the number S_n of distinct sites visited after n steps [Dvoretzky & Erdös 1951, Wijland *et al.* 1997] is a good metric to quantify the rate at which new sites are discovered. The computation of the distribution of S_n is especially relevant to determining the probability that a particle on a network containing randomly distributed traps remains alive after n steps [Rosenstock 1961]. For planar processes in continuous space such as 2D Brownian motion, the area of the convex envelope of the trajectory [Majumdar *et al.* 2010a] can serve as an alternative observable to characterize the extent of the visited territory.

The second category, which we refer to as *first-passage* observables, are related to the efficiency of a target search process, and quantify the statistical properties of the trajectory evaluated upon a random first-passage event. For instance, the first-passage time (FPT) on a fixed target has garnered considerable attention since the introduction of Brownian motion. Indeed, such FPTs are fundamental to many physical phenomena, such as cyclization of a polymeric chain [Gooden *et al.* 1998], protein binding [Berg *et al.* 1981] , or queue lengths [Asmussen 2003]. Consequently, the distribution of the FPT has been extensively studied, both for discrete walks on lattices and graphs [Montroll 1969, Haynes & Roberts 2008], as well as in the context of scale-invariant confined processes [Bénichou *et al.* 2010a] and more recently, non-Markovian confined processes [Levernier *et al.* 2018]. It is worth noting that first-passage observables are not limited to the FPT alone, but also encompass preferential target search observables. For instance, the

splitting probability, defined as the probability of reaching a given target before another, plays a crucial role in physics, such as in quantifying the probability of allele fixation [Wright 1931].

In brief, the study and interpretation of a stochastic transport phenomenon involve two essential aspects: the selection of a particular stochastic process, which comprises transition rules and geometry, and the identification of a relevant stochastic observable that reflects a physical attribute of the system of interest. In this work, we leverage this framework and tackle two separate problems. In the first part, we concentrate exclusively on the second aspect and introduce a novel observable - the visited territory before hitting a target in confinement - that we analyze for the most general class of stochastic processes possible. Conversely, in the second part, we narrow our focus to a specific type of stochastic process: one-dimensional jump processes in confinement. Despite being introduced in the early 20^{th} century, few observables for this system have been computed, and we aim to fill this gap.

Part I. In the context of lattice random walks, a fundamental measure of the exploration process is the number of distinct sites S_n visited after n steps. This quantity has been the subject of extensive research: the complete distribution of S_n is known for a normal walk on a 1D lattice [Dvoretzky & Erdös 1951], and the asymptotic large n behavior of the first moments of S_n have been determined [Jain & Pruitt 1971, Jain & Pruitt 1974] for higher-dimensional hypercubic networks. These findings have been further extended to nearest neighbor walks on arbitrary graphs [Havlin & ben Avraham 1987], and, more recently, to one-dimensional Riemann walks [Hughes 1995] that perform algebraically distributed jumps [Gillis & Weiss 1970, Mariz *et al.* 2001].

Note that the aforementioned results mostly pertain to random walks in an infinite geometry. Indeed, in confined spaces, the total number of sites is bounded, and the walker eventually explores the entire domain. Nevertheless, the question of the number of distinct sites visited is still relevant, provided that a physical stopping time is defined to evaluate S_n. For example, consider a particle confined in a box with a single exit. It is pertinent to investigate the extent of the territory visited by the particle before it reaches the exit, which is essential in determining the likelihood of the particle reacting with an interior target before escaping. Therefore, in the first part of this thesis, we determine the number of distinct sites $C(s_0)$ visited by a confined walker issued from s_0 before reaching a fixed target s_T. It is worth noting that the stopping time corresponding to the target discovery is intrinsically linked to the exploration trajectory. Hence, $C(s_0)$ is a *first-passage* observable.

From a theoretical perspective, determining the statistical properties of the visited territory $C(s_0)$ poses several technical challenges. In particular, the state (visited or not) of the site reached on the n^{th} step depends on the complete history of visited sites, and the presence of an absorbing site at s_T alters the trajectory statistics. To overcome these challenges, we establish a robust connection between evaluating the number of distinct sites visited $C(s_0)$ and determining the splitting probabilities associated with the given process.

More precisely, we devote chapter 1 to the evaluation of the explored territory $C(s_0)$ before reaching a target site for the specific class of Markovian 1D random walks with connected span, *ie* that leave no holes, for which we show that the determination of the splitting probability is necessary and sufficient to determine the complete distribution of $C(s_0)$.

We expand on this analysis in chapter 2 by investigating the behavior of $C(s_0)$ for the broader class of scale-invariant processes, which includes multi-dimensional normal walks, as well as diffusion on fractal graphs and walks with long-range jumps. In the large confining volume limit, we identify universality classes and study the asymptotic behavior of $C(s_0)$, thus extending our previous results beyond the 1D and connected span conditions, and providing a more general framework to analyze the explored territory before reaching a target site.

To further understand the exploration process before reaching the target s_T, we finally focus in chapter 3 on elucidating the interplay between the explored territory and the time taken to do so, which is not contained in the statistics of $C(s_0)$ alone. To that end, we introduce the joint distribution $\sigma(s, n|s_0)$ of the number s of visited sites and the FPT n, and uncover universal behaviors for 1D systems, that hold beyond the Markovian assumption.

Part II. In the second part of this thesis, we shift our focus to a specific type of stochastic process: one-dimensional jump processes. These random walks on the real line $\mathbb{R}$ describe the position x_n of a particle after n steps, starting from x_0, and whose random increments are independent and identically distributed, with common symmetric distribution $p(\ell)$. One typical example of a jump process is the exponential random walk [Van Kampen 1992], for which $p(\ell) = \frac{1}{2}e^{-|\ell|}$. Importantly, these processes are particularly relevant for modeling and interpreting experimental or numerical data, which are intrinsically discrete. They find applications in various transverse fields such as radiative transfer problems [Milne 1921] or for describing the saccadic motion of bacteria [Koshland 1980].

The statistical properties of jump processes in an infinite medium, which can evolve on the entire real line $\mathbb{R}$, are well known. For example, the distribution of the position of the particle at time n can be written explicitly as a function of $p(\ell)$. Similarly, jump processes on the semi-infinite line, *ie* those that are stopped as soon as they pass through 0, are also well characterized. The distribution of the position after n steps is explicit [Ivanov 1994], and the survival probability, defined as the probability of remaining positive during the first n steps, is known exactly for all values of n and x_0 [Doney 2012, Majumdar *et al.* 2017]. In particular, we mention a surprising result known as Sparre Andersen's theorem: the probability of survival starting from $x_0 = 0$ is universal and independent of $p(\ell)$.

However, for jump processes confined to a finite interval $[0, x]$, *ie* stopped at their first exit, there is a lack of general results regarding the distribution of the particle's position after n steps, and the survival probability associated with the interval. The main challenge lies in the fact that the integral equations defining these observables are not solvable for arbitrary jump processes. Nonetheless, we have demonstrated that for large interval size x, some observables linked to the exit events of $[0, x]$ can be assessed in a general way, regardless of the distribution $p(\ell)$.

As a starting point, we focus in chapter 4 on the splitting probability to exit the interval through x rather than through 0. By establishing a link between this bounded quantity and the semi-infinite survival probability, we show that the splitting probability can be asymptotically evaluated for all jump distributions $p(\ell)$ and initial positions x_0. In chapter 5, we aim to further refine the characterization of exit events by examining the rightward and leftward exit-time probabilities, and we provide a comprehensive analysis of their large time asymptotic behavior.

Building upon these results, we investigate in chapter 7 the extreme value statistics of

general jump processes. In particular, we show that the splitting probability and the exit-time probabilities appear as essential components to determine a variety of joint distributions of extremums and time at which they are reached, both for infinite and semi-infinite jump processes.

Furthermore, in chapter 7, we extend our findings to the more general class of confined isotropic jump processes in higher dimensions, and introduce a systematic methodology to evaluate first-passage observables related to exit event in the large volume limit. Our results are valid for any isotropic jump distribution $p(\boldsymbol{\ell})$ and starting position $\mathbf{x_0}$, and include quantities such as first exit-times of disks or overshoot distributions for heavy-tailed processes.

Part I

On the territory explored before reaching a target in confinement

Chapter 1

Confined 1D random walks

Contents

As mentioned in the introduction, the number of distinct sites visited by a random walker is a natural observable to quantify the *geometry* of the exploration and, as such, has received its fair share of attention over time.

The objective of this chapter will be twofold. First, we wish to give a comprehensive overview of known results regarding the territory explored by a random walker. We aim to highlight both important theoretical results as well as direct application examples, in particular in the realm of chemical sciences. Doing so, it will become apparent that the *explored territory for a confined walker before reaching a target* is missing from this picture, even in the simplest case of the symmetric random walk.

In the case of one-dimensional Markovian random walks with connected span, we develop a general framework to provide this missing piece of information, and illustrate the method on a variety of examples. In turn, the results obtained for this class of stochastic processes give precious hints to address more general processes, such as processes with non-connected span, or processes in higher dimensions, studied in chapter 2. Note that this first chapter essentially builds up on results from [Klinger *et al.* 2021].

1.1 Number of distinct sites visited - known results

The determination of the number S_n of distinct sites visited by a lattice random walker during its first n steps has been a long lasting and difficult question, initially introduced in [Dvoretzky & Erdös 1951] in the specific case of the one-dimensional nearest neighbor symmetric random walk. While the full distribution of S_n can be derived in that case, higher dimensional lattices are much harder to deal with, because of the inherent non-Markovian nature of S_n. Indeed, although the walk itself is Markovian, the precise set of visited sites up to time n is a hidden variable, and the status (visited or not) of the site reached on step $n+1$ is unknown.

Despite the mathematical difficulties arising in the investigation of S_n, even for simple hypercubic lattices, a variety of works have been produced in an effort to make progress, stirred by the strong relation between S_n and tangible physical phenomena. Of prime importance in this development is the celebrated Rosenstock trapping problem [Rosenstock 1961]. Consider a particle evolving on some lattice where each vertex contains a trap with probability $\lambda \in [0,1]$. Denote P_n the probability of surviving during the first n steps, *ie* of not encountering a single trap during the first n steps. In terms of S_n, P_n is simply given by:

$$P_n = \langle (1-\lambda)^{S_n} \rangle. \tag{1.1}$$

However, determining P_n is difficult, and requires the full distribution of S_n. As a first approximation, Rosenstock proposed to rewrite $P_n \simeq (1-\lambda)^{\langle S_n \rangle}$, launching a series of studies to determine the first moments of S_n across various geometries and walk rules [1]. Hypercubic lattices were first studied in [Jain & Pruitt 1971, Jain & Pruitt 1974] but we refer the reader to more recent texts for a global account of results on hypercubic lattice walks [Weiss 1994, Hughes 1995]. From there, two natural extensions emerge. The case of the territory explored by N independant walkers was first discussed in [Larralde *et al.* 1992, Weiss *et al.* 1992, Yuste & Acedo 1999] and modified environment or transport rules were investigated respectively in [Havlin & ben Avraham 1987] and [Gillis & Weiss 1970, Mariz *et al.* 2001].

Importantly, throughout these examples, the explored territory is evaluated at some arbitrary deterministic time n. While this is the right way to investigate the rate at which new sites are discovered, we stress that, in light of the initial trapping problem, it is natural to wonder about the territory explored *before some random stopping time*. In particular, for a confined random walker in a domain containing traps, one could wonder about the probability of escaping the domain, which occurs at a random first passage time, without encountering any traps. With the exception of [Yuste *et al.* 2013], in which the territory is evaluated at some external exponential clock, it seems that this problem has been largely overlooked in the community, and we here aim to fill that gap.

Focusing on bounded geometries, for which any target is found almost surely, we introduce the main observable of the three following chapters. For a given fixed target site s_T and a starting site s_0, we consider the territory $C(s_0)$ explored before reaching the target. We emphasize that the random stopping time at which the territory is evaluated is intrinsically linked to the *geometry* of the exploration, as the walker is killed upon reaching the target s_T. In turn, the determination of the full distribution of $C(s_0)$ yields the answer to a slightly modified Rosenstock trapping

[1] Typical fluctuations of S_n are usually studied alongside the first moment.

problem. Considering a particle evolving in a domain with uniform trap distribution and an exit located at s_T, the probability $P(\text{survival}|s_0)$ to exit the system without encountering any traps is simply given by the corresponding Rosenstock formula:

$$P(\text{survival}|s_0) = \langle (1-\lambda)^{C(s_0)} \rangle, \tag{1.2}$$

with immediate applications to determining the efficiency of chemical reactions in bounded living media for example.

1.2 Connected random walks - general methodology

As an introductory and exactly solvable case, we focus first on Markovian one-dimensional lattice random walks on a periodic ring of size N. Additionally, we require the walker to have a *connected span*, such that, for two visited sites s_1 and s_2, all intermediate sites $s \in [\![s_1, s_2]\!]$ are necessarily visited. In particular, walks performing *large jumps* to unvisited sites do not fall into that class. Unless otherwise specified, we consider that the random walker starts at position s_0 and that the absorbing target is located at position 0, as depicted in figure 1.1.

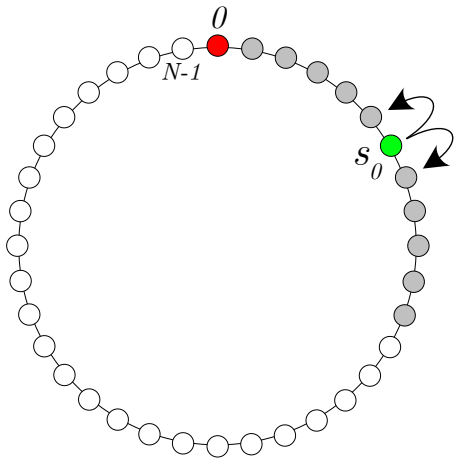


Figure 1.1: Symmetric nearest neighbor random walk in a bounded domain of N sites. Starting from s_0, the walker evolves until it reaches 0 for the first time. In this particular realization, all grayed-out sites have been visited, such that the explored territory is $C(s_0) = 12$ (by convention the target counts as visited).

1.2.1 Linking the maximum and the territory

For one-dimensional connected random walks (*ie* with connected span), we exhibit a direct link between the explored territory $C(s_0)$ and the *maximum* reached by the random walk before reaching zero.

Consider first a random walker starting from s_0 on the semi-infinite line, and stopped upon reaching 0 for the first time. Denoting $\mu(s|s_0)$ the probability that the farthest site reached is s (*ie* that the maximum reached is site s), the distribution of the number of distinct sites is simply given by:

$$P(C = n|s_0) = \mu(n-1|s_0), \tag{1.3}$$

where by convention we count both s_0 and 0 as visited sites. Consider now the same random walker on a periodic ring of size N. Partitioning over whether the target is reached in a clockwise or counter-clockwise manner, we obtain the complete distribution of C:

$$P(C = n|s_0) = 1_{s_0+1\leq n}\ \mu(n-1|s_0) + 1_{n\geq N-s_0+1}\ \mu^r(n-1|N-s_0) \tag{1.4}$$

where $\mu^r(.|s_0)$ is the distribution of the maximum for the same random walk with reversed rightward and leftward hopping probabilities. Note that for general symmetric random walks, $\mu^r(.|s_0) = \mu(.|s_0)$. Importantly, equation (1.4) indicates that the distribution of the number of distinct sites visited is completely determined by the distribution of the maximum in the semi-infinite geometry, provided it can be computed.

To that end, we introduce the splitting probability $\pi_{s_1,\underline{s_2}}(s_0)$ [Hughes 1995, Van Kampen 1992], defined as the probability for the walker starting from $s_0 \in [\![s_1, s_2]\!]$ to reach s_2 before s_1. Note that the splitting probability to reach s_1 before s_2 is given by $\pi_{\underline{s_1},s_2}(s_0) = 1-\pi_{s_1,\underline{s_2}}(s_0)$. In turn, for Markovian one-dimensional random walks, the distribution of the maximum is simply obtained as:

$$\mu(s|s_0) = \pi_{0,\underline{s}}(s_0)\pi_{\underline{0},s+1}(s). \tag{1.5}$$

Combining equations (1.4) and (1.5), we finally rewrite the distribution of C in terms of splitting probabilities only:

$$\begin{aligned} P(C = n|s_0) = {} & 1_{s_0+1\leq n}\pi_{0,\underline{n-1}}(s_0)\pi_{\underline{0},n}(n-1) \ + \\ & 1_{n\geq N-s_0+1}\pi_{\underline{N-n-1},N}(s_0)\pi_{N-n,\underline{N}}(N-n+1). \end{aligned} \tag{1.6}$$

Importantly, the determination of $P(C = n|s_0)$ has been reduced to the evaluation of a single process dependent quantity: the splitting probability $\pi_{s_1,\underline{s_2}}(s_0)$. We emphasize that equation (1.4) is valid for any lattice walk with connected span. In turn, it provides a straightforward methodology to determine the distribution of C, fully characterizing the geometry of exploration before reaching a target in confined spaces.

Beyond the Markovian hypothesis. Note that equation (1.5) only holds for Markovian Random walks. However, even for non-Markovian walks, the distribution of the maximum can be expressed as a function of $\pi_{\underline{0},s}(s_0)$. Indeed, partitioning $\pi_{\underline{0},s}(s_0)$ over the farthest site reached before hitting 0 yields:

$$\pi_{\underline{0},s}(s_0) = \sum_{s=0}^{\infty} \mu(s|s_0). \tag{1.7}$$

As a result, $\mu(s|s_0)$ is simply given by

$$\mu(s|s_0) = \pi_{\underline{0},s+1}(s_0) - \pi_{\underline{0},s}(s_0) \equiv D_s\pi_{\underline{0},s}(s_0), \tag{1.8}$$

and the derivation of $\pi_{\underline{0},s}(s_0)$ is still sufficient to compute $\mu(s|s_0)$ and obtain the complete distribution of $C(s_0)$.

1.2.2 Illustrations

We illustrate the above given methodology by computing the distribution of $C(s_0)$ for representative examples of Markovian random walks with connected span, consistently used to model physical phenomena. We focus on normal random walks (symmetric and biased), but also random walks with short-range memory (persistent), or connected random walks with large jumps (resetting).

Symmetric random walk. Consider first the classical symmetric nearest neighbor random walk with transition probabilities $p(s \to s+1) = p(s \to s-1) = \frac{1}{2}$. Note that this case has been studied in the limit $N \to \infty$ (unconfined system) and in the continuous space approximation in [Dayan & Havlin 1992]. The splitting probability $\pi_{s_1,\underline{s_2}}(s_0)$, satisfies the following equation

$$\begin{aligned} &2\pi_{s_1,\underline{s_2}}(s_0) - \pi_{s_1,\underline{s_2}}(s_0-1) - \pi_{s_1,\underline{s_2}}(s_0+1) = 0 \\ &\pi_{s_1,\underline{s_2}}(s_2) = 1 \\ &\pi_{s_1,\underline{s_2}}(s_1) = 0 \end{aligned} \tag{1.9}$$

and is given by:

$$\begin{aligned} &0 \le s_0 \le s_1,\ \pi_{s_1,\underline{s_2}}(s_0) = \frac{s_0 - s_1}{s_2 - s_1} \\ &s_1 \le s_0 \le N,\ \pi_{\underline{s_1},N}(s_0) = \frac{N - s_0}{N - s_1}. \end{aligned} \tag{1.10}$$

As a result, the distribution of $C(s_0)$ is straightforwardly obtained by combining equations (1.10) and (1.6), and we display in figure 1.2(a) the exact distribution $P(C = n|s_0)$ (along numerical simulations) for given s_0 and N. Importantly, the sharp discontinuity in the distribution stems from the introduction of the confining periodic domain, in opposition to a strictly decreasing distribution in the infinite case, recovered in the $s_0 \ll N$ limit.

Biased random walk. While symmetric random walks model particles undergoing free diffusion, the biased random walk, with hopping probabilities $p(s \to s+1) = 1-p(s \to s-1) = p$, where $0 \le p \le 1$, is a paradigmatic model of particle diffusion in external force fields [Redner 2001]. In this case, the splitting probability obeys the following equation:

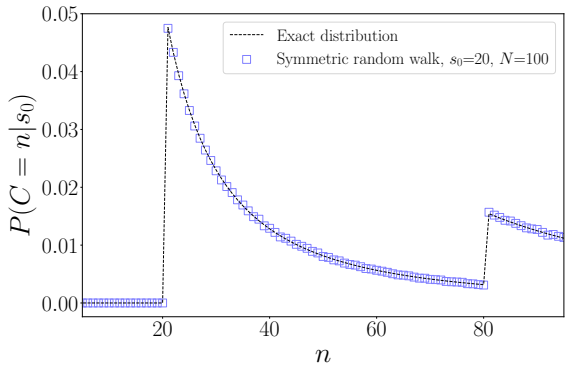
$$\begin{aligned} &\pi_{s_1,\underline{s_2}}(s_0) - p\pi_{s_1,\underline{s_2}}(s_0+1) - (1-p)\pi_{s_1,\underline{s_2}}(s_0-1) = 0 \\ &\pi_{s_1,\underline{s_2}}(s_2) = 1 \\ &\pi_{s_1,\underline{s_2}}(s_1) = 0 \end{aligned} \tag{1.11}$$

and is given by

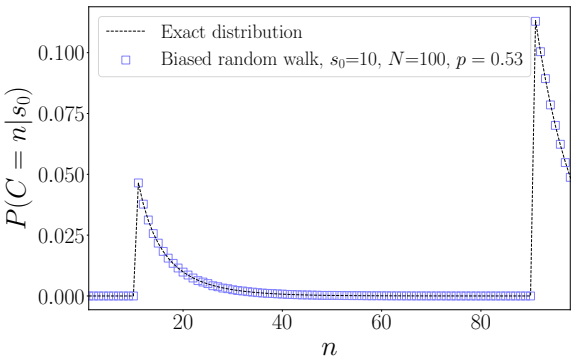
$$0 \le s_0 \le s_1,\ \pi_{0,\underline{s_1}}(s_0) = \frac{\alpha^{s_0} - 1}{\alpha^{s_1} - 1} \tag{1.12}$$

$$s_1 \le s_0 \le N,\ \pi_{\underline{s_1},N}(s_0) = \frac{\alpha^{s_0} - \alpha^N}{\alpha^{s_1} - \alpha^N} \tag{1.13}$$

with $\alpha = \frac{1-p}{p}$. Similarly to the symmetric case, the distribution displays two sharp jumps when $s_0 = O(N)$ or $N - s_0 = O(N)$ whose interpretation is unchanged; their relative weight is now controlled by the bias (see figure 1.2(b)).



(a) Distribution of $C(s_0)$ for a symmetric nearest neighbor random walk. The discontinuity at $n = 80$ corresponds to trajectories exiting clockwise, which visit at least 80 sites. In turn, for $n < 80$, the territory distribution is identical to that of the maximum in the semi-infinite space.



(b) Distribution of $C(s_0)$ for a biased nearest neighbor random walk. In that case, the walker is biased in a clockwise manner. In turn, the relative weight of trajectories exiting clockwise is increased, as shown by the high peak at $n = 80$.

Figure 1.2

Persistent nearest neighbor random walk. We now show that equation (1.6) is easily adapted to deal with random walks with short-ranged memory. As a paradigmatic example of such processes, we consider the persistent random walk [Weiss 1994, Tejedor *et al.* 2012], defined as follows: at each time step the walker performs a step identical to the previous one with probability p, and opposite with probability $1 - p$. By considering two consecutive steps of the walk, its dynamics become Markovian, and the splitting probability can be obtained explicitly. Precisely, we introduce the conditional splitting probabilities $u_{0,\underline{s}}(s_0)$ and $v_{0,\underline{s}}(s_0)$ such that

$$\begin{aligned} u_{0,\underline{s}}(s_0) &= \pi_{0,\underline{s}}(s_0|\text{previous step is taken rightwards}) \\ v_{0,\underline{s}}(s_0) &= \pi_{0,\underline{s}}(s_0|\text{previous step is taken leftwards}), \end{aligned} \tag{1.14}$$

and associated boundary conditions given by $u_{0,\underline{s}}(s) = 1$ and $v_{0,\underline{s}}(0) = 0$. Denoting a the probability of taking the first step rightwards, the splitting probability $\pi_{0,\underline{s}}(s_0, a)$ is then given by:

$$\pi_{0,\underline{s}}(s_0, a) = a u_{0,\underline{s}}(s_0 + 1) + (1 - a) v_{0,\underline{s}}(s_0 - 1), \tag{1.15}$$

and the distribution of $C(s_0)$ can be rewritten as:

$$\begin{aligned} P(C = n|s_0, a) =& 1_{s_0+1 \leq n} \pi_{0,\underline{n-1}}(s_0, a) \pi_{\underline{0},n}(n - 1, p) \; + \\ & 1_{n \geq N - s_0 + 1} \pi_{\underline{N-n+1},N}(s_0, a) \pi_{N-n,\underline{N}}(N - n + 1, 1 - p). \end{aligned} \tag{1.16}$$

Note that at the "maximum" of the trajectory, the walker needs to turn around, hence the $\pi_{.,.}(.|p)$ and $\pi_{.,.}(.|1-p)$ terms. The conditional splitting probabilities $u_{0,\underline{s}}(s_0)$ and $v_{0,\underline{s}}(s_0)$ obey the following set of equations (where we rewrite u_{s_0} and v_{s_0} to increase readability)

$$\begin{aligned} u_{s_0} &= pu_{s_0+1} + (1-p)v_{s_0-1}, \\ v_{s_0} &= pv_{s_0-1} + (1-p)u_{s_0+1}, \end{aligned} \tag{1.17}$$

which can be recast as:

$$\begin{aligned} 2u_{s_0} &= u_{s_0+1} + u_{s_0-1}, \\ v_{s_0} &= \frac{1}{1-p}(u_{s_0+1} - pu_{s_0+2}). \end{aligned} \tag{1.18}$$

Enforcing the boundary conditions leads to the following expressions of u_{s_0}, v_{s_0}:

$$\begin{aligned} u_{s_0} &= 1 + B(s_0 - s), \\ v_{s_0} &= Bs_0, \\ B &= \frac{p-1}{(1-s)(1-p)-p}, \end{aligned} \tag{1.19}$$

and we finally obtain the splitting probability for $0 < s_0 < s$:

$$\pi_{0,\underline{s}}(s_0, a) = \frac{p-1}{(1-p)(1-s)-p}s_0 + \frac{1-p-a}{1+p(s-2)-s}. \tag{1.20}$$

In turn, the exact distribution of $C(s_0)$ is computed by making use of equation (1.16). Of note:

- The continuous counterpart of the persistent random walk is the Run and Tumble particle model [Van Kampen 1992, Dhar *et al.* 2019], for which the splitting probability is known and takes a similar form: an affine function with tumbling-rate dependent slope and intercept.
- Taking $p = a = \frac{1}{2}$ yields the normal walk result.
- We display numerical agreement in figure 1.3. Again, the distribution displays two sharp jumps for $n = s_0 + 1$ and $n = N - s_0 + 1$. Here, the corresponding peaks are sharpened as the persistence time of the random walk (controlled by the parameter p) is increased.

Resetting random walk. We finally consider the resetting random walk, which has recently received a lot of attention, mostly in the continuous space and time setting. While initially introduced in [Evans & Majumdar 2011], we refer the reader to [Evans *et al.* 2020] for a comprehensive overview of advances on resetting Brownian motion. In the discrete lattice version, the walker either performs nearest neighbor symmetric hopping with probability $1-\lambda$, or resets to a given fixed position s_p with probability $\lambda \in [0, 1]$. Importantly, even if resetting events lead to large jumps whose range can cover many sites, the span of the process remains connected, because unvisited sites can only be reached by nearest neighbor hopping (all resetting jumps

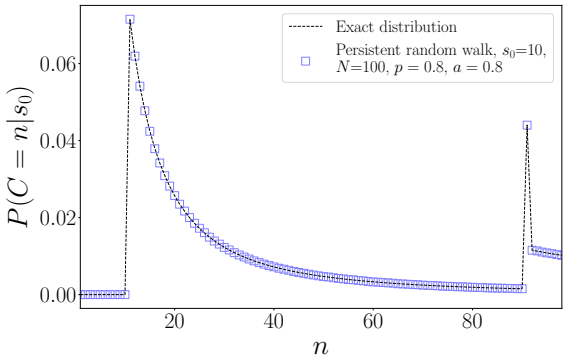


Figure 1.3: Distribution of C for a given persistent random walk with $p = 0.8$. The singular peak at $n = 90$ corresponds to trajectories taking their first step clockwise.Note that the probabilistic weight of trajectories with $n > 90$ is small. In that case, the first step needs to be taken counter-clockwise, and reversal of the walk must occur within 10 steps.

lead to s_p). As a result, the resetting random walk has connected span, and equation (1.6) is still valid.

Again, we determine the splitting probability $\pi_{0,\underline{s}}(s_0)$. We emphasize that, to the best of our knowledge, this quantity has not been given in the literature for discrete processes, even if related quantities have been studied in the continuous setup [Chechkin & Sokolov 2018, Pal & Prasad 2019]. Recalling that the resetting site is s_p, $\pi_{0,\underline{s}}(s_0)$ obeys the following equation:

$$\begin{aligned} &\pi_{0,\underline{s}}(s_0+1) - \frac{2}{1-\lambda}\pi_{0,\underline{s}}(s_0) + \pi_{0,\underline{s}}(s_0-1) = -\frac{2\lambda}{1-\lambda}\pi_{0,\underline{s}}(s_p) \\ &\pi_{0,\underline{s}}(0) = 0 \\ &\pi_{0,\underline{s}}(s) = 1, \end{aligned} \tag{1.21}$$

and the solution of equation (1.21) is given by the sum of the homogeneous solution $h(s_0)$ and a particular solution $p(s_0)$. First, the homogeneous solution reads

$$h(s_0) = \frac{r_+^{s_0} - r_-^{s_0}}{r_+^s - r_-^s}, \tag{1.22}$$

where $r_\pm = \frac{1}{1-\lambda} \pm \sqrt{\frac{1}{(1-\lambda)^2} - 1}$. Second, we construct the particular solution using the Green function [Barton 1989] of the problem defined by:

$$\begin{aligned} &G(s_0+1, s_2) - \frac{2}{1-\lambda}G(s_0, s_2) + G(s_0-1, s_2) = \delta_{s_0,s_2} \\ &G(0, s_2) = G(s, s_2) = 0. \end{aligned} \tag{1.23}$$

Following classical Green function calculations, we obtain

$$G(s_1, s_2) = 1_{s_1 \le s_2} G_-(s_1, s_2) + 1_{s_1 > s_2} G_+(s_1, s_2), \tag{1.24}$$

with

$$G_-(s_1, s_2) = A(s_2)^{-1}(r_+^{s_1} - r_-^{s_1}), \quad G_+(s_1, s_2) = A(s_2)^{-1}\frac{r_+^{s_2} - r_-^{s_2}}{r_+^{s_2} - r_-^{s_2}\frac{r_+^s}{r_-^s}}(r_+^{s_1} - r_-^{s_1}\frac{r_+^s}{r_-^s}) \tag{1.25}$$

and

$$\begin{aligned} A(s_2) =&(r_+^{s_2} - r_-^{s_2})(r_+^{s_2} - r_-^{s_2}\frac{r_+^s}{r_-^s})^{-1}(r_+^{s_2+1} - r_-^{s_2+1}\frac{r_+^s}{r_-^s}) \\ &- \frac{2}{1-\lambda}(r_+^{s_2} - r_-^{s_2}) + (r_+^{s_2-1} - r_-^{s_2-1}). \end{aligned} \tag{1.26}$$

Finally, the particular solution is given by

$$p(s_0) = -\frac{2}{1-\lambda}\pi_{0,\underline{s}}(s_p)\sum_{s_2} G(s_0, s_2). \tag{1.27}$$

In turn, taking $s_0 = s_p$ and writing $\pi_{0,\underline{s}}(s_p) = h(s_p) + p(s_p)$, we determine $\pi_{0,\underline{s}}(s_p)$ explicitly:

$$\pi_{0,\underline{s}}(s_p) = \frac{h(s_p)}{1 + \frac{2\lambda}{1-\lambda}\sum_{s_2} G(s_p, s_2)}, \tag{1.28}$$

and derive the splitting probability for arbitrary s_p and s_0:

$$\pi_{0,\underline{s}}(s_0) = h(s_0) - \frac{h(s_p)2\lambda\sum_{s_2} G(s_0, s_2)}{1 - \lambda + 2\lambda\sum_{s_2} G(s_p, s_2)}, \tag{1.29}$$

which covers in particular the case of resetting to the initial position $s_p = s_0$. Once again, the distribution of $C(s_0)$ is obtained from equation (1.6), and we conclude this section by highlighting an interesting behavior illustrated in figure 1.4: for $s_0 < N/2$, as the resetting rate increases, a spike in the distribution grows at $2s_0$. Indeed, for $\lambda \ll 1$, resetting jumps can be neglected and one recovers the case of the symmetric nearest neighbor random walk. However, for $\lambda \gg 1$, many resetting events occur, leading to a symmetric exploration of the domain around s_0 before the target is reached. The explored territory before exit is thus approximately $2s_0$, yielding the observed peak in the distribution.

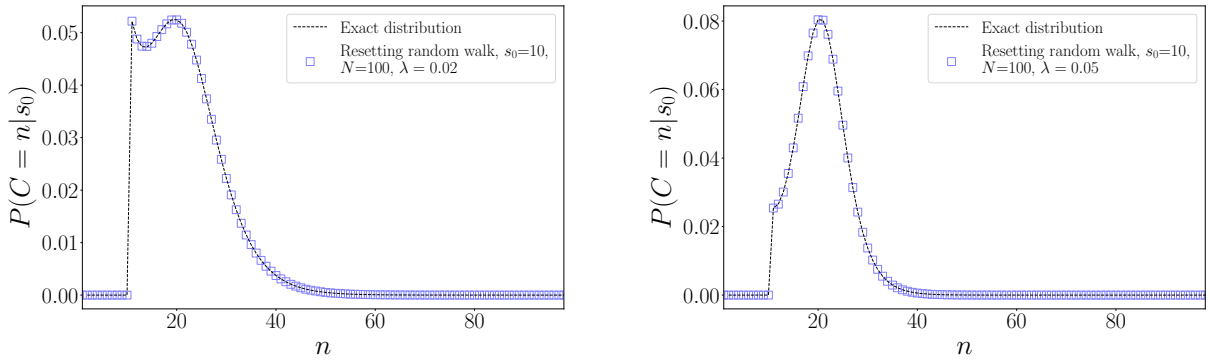


Figure 1.4: Distribution of C for a resetting random walk with fixed s_0 and N, for $\lambda = 0.02$ and $\lambda = 0.05$. As λ increases, the distribution develops a peak around $2s_0$, corresponding to a symmetric exploration of space around s_0, until the target is found.

1.2.3 Extensions

Equation (1.4) is specific to the 1D periodic lattice geometry, but can easily be extended to other one-dimensional geometries such as reflecting boundary conditions, as well as to continuous space and time stochastic processes.

Reflecting boundary conditions. Let us first focus on the case of lattice walks with reflecting boundary conditions. Consider a discrete Markovian random walk with connected span, on a one-dimensional lattice of N sites. The target site is still located at 0, and we assume that there is a reflecting boundary at site $N-1$: this effectively means that all nearest neighbor jumps from site $N-1$ lead to site $N-2$. In that case, the clockwise/counter-clockwise partition is not relevant, and the distribution of $C(s_0)$ is readily written as:

$$\begin{aligned} &P(C=n|s_0) = \pi_{0,\underline{n-1}}(s_0)\pi_{\underline{0},n}(n-1) \text{ if } s_0 < n < N \\ &P(C=n|s_0) = \pi_{0,\underline{N-1}}(s_0) \text{ if } n = N. \end{aligned} \tag{1.30}$$

Importantly, as for periodic boundary conditions, this expression only involves splitting probabilities. In particular, for all random walks considered before, we have access to the explored territory in the reflecting geometry.

Continuous space and time stochastic processes. Equation (1.4) between the maximum and the number of distinct sites visited naturally extends to continuous space and time stochastic processes. Consider a one-dimensional Markovian stochastic process starting from x_0 and evolving in the interval $[0, L]$ with absorbing boundary conditions at 0 and L. For processes with connected span, the distribution of the visited domain $C(x_0)$ is simply given by a continuous version of equation (1.4):

$$P(C=x|x_0) = 1_{x_0 \leq x}\ \mu(x|x_0) + 1_{L-x_0 \leq x}\ \mu^r(x|L-x_0), \tag{1.31}$$

such that

- $\mu(x|s_0)$ is the continuous distribution of the maximum in the semi-infinite space, with $\int_0^\infty \mu(x|x_0)\mathrm{d}x = 1$.
- $\mu^r(.|x_0)$ is the distribution of the maximum for the *reversed* process, which can also be interpreted as the absolute value of the minimum, for the initial process started at position $-x_0$. Note that for symmetric processes (*eg* Brownian motion), $\mu(.|x_0) = \mu^r(.|x_0)$.
- $P(C=x|x_0)$ should be understood as a continuous distribution, in the sense that $\int_0^L P(C=x|x_0)\mathrm{d}x = 1$.

As for lattice random walks, we express the distribution $\mu(x|x_0)$ of the maximum in terms of continuous splitting probabilities $\pi_{0,\underline{x}}(x_0)$ only. Note that since the Markovian argument $\mu(x|x_0) = \pi_{0,\underline{x}}(x_0)\pi_{\underline{0},x}(x)$ is not valid in the continuous setup ($\pi_{\underline{0},x}(x) = 0$ by definition), we rewrite the splitting probability as a partition over the farthest point reached, whose probability density is $\mu(.|x_0)$ itself:

$$\pi_{0,\underline{x}}(x_0) = 1 - \pi_{\underline{0},x}(x_0) = 1 - \int_0^x \mu(u|x_0)\mathrm{d}u, \tag{1.32}$$

such that

$$\mu(x|x_0) = -\frac{\mathrm{d}}{\mathrm{d}x}\pi_{0,\underline{x}}(x_0). \tag{1.33}$$

As a result, combining equations (1.31) and (1.33), we obtain the exact distribution of the explored territory for continuous processes with connected span, provided $\pi_{0,\underline{x}}(x_0)$ can be computed. Additionally, we emphasize that equation (1.33) is valid beyond the Markovian hypothesis. Since splitting probabilities of classical and biased Brownian motion are well known (see [Redner 2001] for instance) we illustrate the ease of use of the continuous formalism on the specific case of the resetting Brownian motion.

Consider a Brownian particle with diffusion coefficient D and subject to resetting events to x_p occurring with rate λ. The splitting probability obeys the following backward equation:

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\pi_{0,\underline{x}}(x_0) - r^2\pi_{0,\underline{x}}(x_0) = r^2\pi_{0,\underline{x}}(x_p), \tag{1.34}$$

where x_p denotes the resetting position, x_0 the current position, and $r = \sqrt{\frac{\lambda}{D}}$. The complete solution (which can also be found in [Pal & Prasad 2019]) is given by:

$$\pi_{0,\underline{x}}(x_0) = \frac{\sinh(rx_0)}{\sinh(rx)} + \frac{\pi_{0,\underline{x}}(x_p)}{\sinh(rx)}\left[\sinh(rx) - \sinh(r(x-x_0)) - \sinh(rx_0)\right]. \tag{1.35}$$

Taking $x_0 = x_p$, the splitting probability starting from x_p with resetting to x_p reads

$$\pi_{0,\underline{x}}(x_p) = \frac{\sinh(rx_p)}{\sinh(r(x-x_p)) + \sinh(rx_p)}, \tag{1.36}$$

and we finally obtain

$$\mu(x|x_p) = \frac{r\sinh(rx_p)\cosh(r(x-x_p))}{(\sinh(rx_p) + \sinh(r(x_p-x)))^2}. \tag{1.37}$$

Let us conclude with a few comments:

- Equation (1.37) is much simpler than its discrete counterpart, obtained from the splitting probability of the discrete resetting walk (1.29). In turn, it can easily be analyzed asymptotically.
- As $r \to 0$, no resetting occurs and we recover the classical Brownian result: $\mu(x|x_p) = \frac{x_p}{x^2}$
- As $r \to \infty$ resetting events occur infinitely often, and we show that $\mu(x|x_p) \to \delta(x - 2x_p)$. In turn, we obtain the analytical validation of the asymptotic concentration of $C(s_0)$ around $2s_p$ displayed in figure 1.4.

1.3 An example of disconnected random walk - the relocating random walk

As discussed at length in the previous section, the determination of the distribution of $C(s_0)$ relies heavily on both **(i)** the one-dimensional geometry and **(ii)** the connected span property. While our approach is sufficient to derive exact results for a variety of classical examples across the discrete and continuous spectrum of stochastic processes, we now wish to go beyond these hypotheses and investigate different kinds of processes.

1.3.1 Systematic relocation - or - the exact Golden Coupon problem

As a first deviation from hypothesis **(ii)**, we focus on the relocating walk, a typical example of lattice random walk with non-connected span, which belongs to the family of intermittent random walks [Benichou *et al.* 2008, Benichou *et al.* 2008, Bénichou *et al.* 2011]. At each time step, the random walker either performs a nearest neighbor hop, or relocates with probability λ to a site drawn uniformly from the set of N available sites, until 0 is found.

We first focus on systematic relocation, *ie* $\lambda = 1$. In that case, the determination of the distribution of the territory explored before exit can be recast in solving the following *coupon* problem [Holst 1986]. Assume that a coupon is drawn uniformly out of N different coupons labeled from 0 to $N-1$. The experiment is repeated until coupon 0 (the golden coupon) is drawn. Recalling that the random walk defined above starts from a given site $s_0 \neq 0$, the number of distinct sites visited is equal to the number of distinct coupons drawn before the golden one. Importantly, all coupons are strictly equivalent, so that their order of first arrival is distributed uniformly. In turn, the number of distinct coupons drawn before the golden one is uniform over $[\![0, N-1]\!]$. The distribution of the visited territory immediately follows:

$$P(C = n|s_0) = \frac{1}{N-1}, \tag{1.38}$$

where we take into account the fact that s_0 counts as a visited site. While seemingly trivial, this first result for non-connected random walks will have far-reaching consequences for stochastic processes in higher dimensions, as will be seen in chapter 2.

1.3.2 Rare relocation - perturbative results

Consider now the opposite rare relocation limit, namely $\lambda \to 0$. When $\lambda = 0$, the walker performs a symmetric random walk, and the distribution of $C(s_0)$ is given by equation (1.6). The first order correction is obtained by considering trajectories where relocation occurs only once before 0 is reached. To simplify further, we only focus on cases where the number n of visited sites is smaller than s_0, such that contributing trajectories have to relocate. The first order derivation details are given in Appendix A and we provide numerical illustration in figure 1.5.

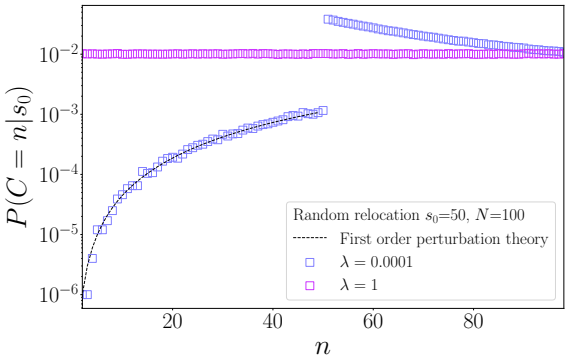


Figure 1.5: Distribution of $C(s_0)$ for the relocating random walk. In the $\lambda = 1$ systematic relocation case, the distribution is uniform. In the rare relocation case $\lambda \ll 1$, the first order perturbation theory is accurate. We emphasize that in the case $\lambda = 0$, $P(C = n|s_0 = 50) = 0$ for all $n \leq 50$, in contrast to the $\lambda > 0$ case. For $n > s_0$, the distribution of C is dominated by the symmetric random walk behavior.

As a concluding remark, we emphasize that the perturbation theory has two controlled levels of approximation:

- It is a first order expansion in the number of relocation events: we neglect all trajectories with more than one jump.
- We neglect *crossover situations*, where the set of visited sites before relocation overlaps with the set of visited sites after relocation. Note that in the limit $n \ll s_0$, this approximation becomes exact as no overlapping configurations exist. Our estimation is thus only a lower bound of the true distribution, since more trajectories contribute to the overall probabilistic weight.
- In a similar fashion, a lower bound for the second order expansion in the number of jumps can be obtained by considering trajectories where the final set of visited sites can be split into 3 non overlapping sets.

1.4 Conclusion

As a starting point in a general study of the territory $C(s_0)$ explored by a confined random walker before reaching a target, we focus on the specific case of one-dimensional random walks with connected span, *ie* that leave no holes. Due to the high constraints imposed by the geometry, we derive a systematic procedure to evaluate the distribution of the visited territory $P(C = n|s_0)$ for various bounded geometries, in terms of splitting probabilities only.

We can now answer the initial question stemming from the *Rosenstock trapping problem* [Rosenstock 1961]. For a lattice random walk with traps distributed with frequency λ, we recall that the probability $P(\text{survival}|s_0)$ to escape the system without encountering any traps reads (see equation (1.2)):

$$P(\text{survival}|s_0) = \sum_{n=1}^{N} P(C = n|s_0)(1-\lambda)^n. \tag{1.39}$$

In turn, the derivation of $P(C = n|s_0)$ provides an exact and quantitative description of such survival probability, with potential applications to chemical reaction processes in closed reservoirs.

In a first attempt to relax the connected span hypothesis, we successively investigate two limits of a specific non-connected intermittent random walk, the random relocating normal walk. In the systematic relocation regime, we obtain the exact distribution of $C(s_0)$ by reducing its evaluation to solving a combinatorial *coupon* problem. The opposite rare relocation limit is harder to deal with, and we compute the distribution of $C(s_0)$ perturbatively. Importantly, these examples indicate that, depending on the microscopic details of the random walk, a plethora of behaviors can emerge.

In the following chapter, we aim to investigate further the behavior of $C(s_0)$ across a broad range of stochastic processes, from diffusion on fractals to lattice random walks with long ranged jumps, as has been done for first passage times (FPT) [Bénichou & Voituriez 2014]. To that end, we devote chapter 2 to the study of the statistics of $C(s_0)$ in the *large confining volume limit*, in which the microscopic details of the transport rules become irrelevant with respect to more general properties.

Chapter 2

Scale-invariant processes in the large volume limit

Contents

One-dimensional systems have taught us that explicit expressions for the distribution of the explored territory $C(s_0)$ before reaching a target can be computed. While such expressions highly depend on the specific rules of the underlying random walk, we also exhibit a common pattern: the strong link between the determination of $C(s_0)$ and the splitting probability $\pi_{s_1,\underline{s_2}}(s_0)$. Based on that observation, we devote this chapter extending this relation to a broader class of stochastic processes, namely general *Markovian scale-invariant processes*, among which diffusion in higher dimension, diffusion on fractals, and diffusion in disordered media. By exploiting the universal behavior of the splitting probability emerging in the *large confining volume limit*, we further characterize the statistics of $C(s_0)$.

As a starting point, we provide a self-consistent overview of the technical ingredients needed to determine splitting probabilities for general scale-invariant processes. We introduce Pseudo Green Functions H and express $\pi_{s_1,\underline{s_2}}(s_0)$ in terms of H. Making use of the large volume limit behavior of H functions, we describe the corresponding emerging behavior of splitting probabilities.

In turn, we focus on the first moment $\langle C(s_0)\rangle$ of the territory, and exhibit the corresponding universality classes, going beyond the connected one dimensional case. Finally, we show that the full distribution of $C(s_0)$ can be recast into robust universal scaling forms.

2.1 Splitting Probabilities of general scale-invariant processes

Most of the information contained in section 2.1 can be found in [Condamin *et al.* 2005, Condamin *et al.* 2007b, Chevalier *et al.* 2011, Bénichou & Voituriez 2014, Barton 1989]. However, we find it useful to provide a self-contained manuscript, and want to emphasize the sequential approximation schemes that lead to universal scaling forms for the splitting probability.

2.1.1 Backward equations for the splitting probability

As a general starting point, we consider a particle evolving in some domain $\mathcal{D}$ of volume V, starting at position $\mathbf{x_0}$ and with reflecting boundary conditions on the surface Σ of $\mathcal{D}$. We hereafter consider *Markovian* dynamics, for which one can define a backward operator $\mathcal{L}_{\mathbf{x_0}}$ (the adjoint of the classical forward operator [Van Kampen 1992]), such that the splitting probability $\pi_{\mathbf{x_1},\underline{\mathbf{x_2}}}(\mathbf{x_0})$, defined as the probability that the particle reaches target 1 before target 2, obeys the following set of equations:

$$\begin{aligned} &\mathcal{L}_{\mathbf{x_0}}\pi_{\mathbf{x_1},\underline{\mathbf{x_2}}}(\mathbf{x_0}) = 0 && \text{for } \mathbf{x_0} \in \mathcal{D} \\ &\pi_{\mathbf{x_1},\underline{\mathbf{x_2}}}(\mathbf{x_0}) = 0 && \text{for } \mathbf{x_0} \text{ on } \Sigma_1 \\ &\pi_{\mathbf{x_1},\underline{\mathbf{x_2}}}(\mathbf{x_0}) = 1 && \text{for } \mathbf{x_0} \text{ on } \Sigma_2 \\ &\partial_{\mathbf{n}}\pi_{\mathbf{x_1},\underline{\mathbf{x_2}}}(\mathbf{x_0}) = 1 && \text{for } \mathbf{x_0} \text{ on } \Sigma, \end{aligned} \tag{2.1}$$

where Σ_1 and Σ_2 are the surfaces of respectively targets 1 and 2 located at positions $\mathbf{x_1}$ and $\mathbf{x_2}$. In the specific case of Brownian motion, the backward operator is simply given by the Laplacian $\mathcal{L}_{\mathbf{x_0}} = \Delta_{\mathbf{x_0}}$, and we illustrate the situation in figure 2.1.

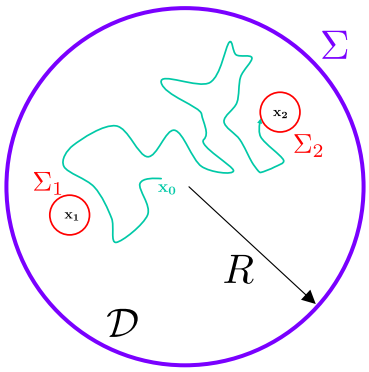


Figure 2.1: Starting from $\mathbf{x_0}$, the Brownian particle evolves until it reaches either Σ_1 or Σ_2. In this specific example, the confining domain is spherical. Note that the boundary conditions on the outer shell are taken reflecting.

Importantly, we emphasize that equation (2.1) does not usually have explicit analytical solutions. As an example, even for the simple case of Brownian motion in a disk, arbitrary target location prevents any form of analytical solution. As a first step towards simplifying the evaluation of splitting probabilities, we introduce the pseudo green functions of the problem.

2.1.2 Example: Normal walk on hypercubes

As an introductory example, consider a symmetric random walker on a periodic hypercubic lattice of N sites, with two targets at sites s_1 and s_2. Denoting Δ_{s_i} the discrete Laplacian acting on variable s_i [1], equation (2.1) naturally extends to the lattice geometry, and the splitting probability $\pi_{1,\underline{2}}(s_0)$ to s_2 rather than s_1 obeys the following equation:

$$\begin{aligned} &\Delta_{s_0}\pi_{1,\underline{2}}(s_0) = 0 \\ &\pi_{1,\underline{2}}(s_2) = 1 \\ &\pi_{1,\underline{2}}(s_1) = 0. \end{aligned} \tag{2.2}$$

Pseudo green functions. For fixed s_j, the pseudo green function of the problem $H(s_i, s_j) \equiv H_{ij}$ is defined by the following equation [Barton 1989]:

$$\Delta_i H_{ij} = \delta_{ij} - \frac{1}{N}, \tag{2.3}$$

where δ_{ij} is the Kronecker delta symbol. In fact, it has been shown [Condamin *et al.* 2007b] that $\pi_{1,\underline{2}}(s_0)$ can be expressed in terms of pseudo green functions only:

$$\pi_{1,\underline{2}}(s_0) = \frac{H_{02} - H_{10} + H_{11} - H_{12}}{H_{11} + H_{22} - 2H_{12}}. \tag{2.4}$$

Importantly, computing H is simpler than solving equation (2.2), as the targets s_1 and s_2 have been suppressed. We provide in Appendix B exact expressions of pseudo green functions for 2 and 3 dimensional periodic hypercubes. Of note, the simplification of first-passage time problems with the help of H functions was first initiated in [Noh & Rieger 2004], for mean first-passage times on complex networks.

Infinite green functions. Consider now the unbounded lattice, and define the infinite green function $G_\infty(s_i, s_j)$ by

$$G_\infty(s_i, s_j) = \sum_{n=0}^{\infty} P^\infty(s_j, n|s_i), \tag{2.5}$$

where the infinite propagator $P^\infty(s_j, n|s_i)$ is defined as the probability for the unbounded walker to be at s_j on the n^{th} step, starting from s_i. Note that depending on the dimension of the lattice, $G_\infty(s_i, s_j)$ may or may not be finite, but differences of G_∞ are always well defined. Importantly, we emphasize that G_∞ satisfies the following equation:

$$\Delta_i G_\infty(s_i, s_j) = \delta_{ij}. \tag{2.6}$$

[1] The periodic boundary conditions are directly embedded in the discrete Laplacian.

Large volume limit. Let us now fix s_1, s_2 and s_0 and focus on the large volume limit $N \to \infty$ of $\pi_{1,\underline{2}}(s_0)$. Making use of equations (2.3) and (2.6), differences of H converge towards differences of G_∞ and we obtain [Condamin *et al.* 2005, Condamin *et al.* 2007a]:

$$\pi_{1,\underline{2}}(s_0) \underset{N\to\infty}{\sim} \frac{G_\infty(s_0,s_2) - G_\infty(s_1,s_0) + G_\infty(s_1,s_1) - G_\infty(s_1,s_2)}{G_\infty(s_1,s_1) + G_\infty(s_2,s_2) - 2G_\infty(s_1,s_2)}. \tag{2.7}$$

Finally, defining r_{ij} the Euclidean distance between sites s_i and s_j, and making use of the translational invariance of the walk, the splitting probability is asymptotically given by:

$$\pi_{1,\underline{2}}(s_0) \underset{N\to\infty}{\sim} \frac{G_\infty(r_{02}) - G_\infty(r_{10}) + G_\infty(0) - G_\infty(r_{12})}{2\left[G_\infty(0) - G_\infty(r_{12})\right]}. \tag{2.8}$$

We emphasize that the evaluation of $\pi_{1,\underline{2}}(s_0)$ has been reduced to the sole evaluation of $G_\infty(r)$. In particular, the source target-distance dependence of $\pi_{1,\underline{2}}(s_0)$ is fully contained in G_∞. In the following subsection, we exploit the specific scaling form of G_∞ for scale-invariant processes to uncover universal behavior of the asymptotic splitting probability.

2.1.3 Asymptotic splitting probabilities of scale-invariant processes

Walk dimension, fractal dimension and universality classes. Ultimately, we want to quantify the explored territory $C(s_0)$ for general *Markovian scale-invariant processes*. As a first step, we assume that the space in which the stochastic process evolves is characterized by its *fractal dimension* d_f, such that the volume accessible at a distance R scales as:

$$V \sim R^{d_f}. \tag{2.9}$$

For hypercubic lattices or general Euclidean spaces, the fractal dimension is naturally given by the dimension of the embedding space (*eg* $d_f = 3$ for three dimensional systems). However, more complex structures such as fractal graphs or disordered media (see figure 2.3) allow for non integer values of the fractal dimension. For more information on the topic, we refer the reader to the classical textbook [Havlin & ben Avraham 1987].

The stochastic process $x(t)$ is *scale-invariant* if, for $\lambda \in \mathbb{R}$, $x(t)$ and $\lambda^h x(t/\lambda)$ have the same statistics, for some exponent h characterizing the process. In other words, scale-invariance can be understood as the absence of any characteristic lengthscale in a given trajectory, such that its statistical properties are preserved by zooming in or out. For such pocesses, the mean square displacement $\langle x^2(t)\rangle$ (MSD) has a known power law scaling with time[2]:

$$\langle x^2(t)\rangle \underset{t\to\infty}{\propto} t^{\frac{2}{d_w}}, \tag{2.10}$$

where $d_w = h^{-1}$ is defined as the *walk dimension*. Additionally, the infinite propagator takes the following scaling form [Havlin & ben Avraham 1987]:

$$P^\infty(\mathbf{x}, t|\mathbf{x_0}) = \frac{1}{t^{\frac{d_f}{d_w}}} \Pi\left(\frac{|\mathbf{x}-\mathbf{x_0}|}{t^{\frac{1}{d_w}}}\right) \tag{2.11}$$

[2]When the mean squared displacement is ill-defined, in the case of α-stable processes for instance, the MSD should be understood as the *typical* displacement.

with Π some process dependent function. In turn, the associated infinite Green function is given by:

$$\begin{aligned} G_\infty(\mathbf{x_1}) &= \int_0^\infty \frac{1}{t^{\frac{d_f}{d_w}}} \Pi\left(\frac{|\mathbf{x_1}|}{t^{\frac{1}{d_w}}}\right) \mathrm{d}t \\ &= |\mathbf{x_1}|^{d_w - d_f} \int_0^\infty \frac{1}{u^{\frac{d_f}{d_w}}} \Pi\left(\frac{1}{u^{\frac{1}{d_w}}}\right) \mathrm{d}u. \end{aligned} \tag{2.12}$$

It is known [Bénichou & Voituriez 2014] that the scaling function $\Pi(s)$ decays quickly enough at large s to ensure convergence near 0. Additionally, since Π has a finite limit as $s \to 0$, the behavior of the integral near $+\infty$ is dictated by the d_f/d_w ratio. In turn, 3 distinct universal regimes emerge:

- for $d_f > d_w$, the integral is well behaved near $+\infty$, and the infinite Green function G_∞ is finite. From equation (2.5), $G_\infty(\mathbf{x_1})$ appears as the mean number of returns to position $\mathbf{x_1}$, which is thus finite. The case $d_f > d_w$ is called *transient*, or *non-compact*. Pictorially, the space is too big for the walker, which has a non zero probability of never coming back to a previously visited position.

- for $d_f < d_w$, the integral diverges near $+\infty$, and G_∞ is infinite. In turn, the mean number of returns to position $\mathbf{x_1}$ is infinite, and this case is referred to as *recurrent*, or *compact*. The space is now too cramped, and the walker returns to previously visited places with probability 1.

- The case $d_f = d_w$ is called *marginal*. While analytically singular, its interpretation in terms of returns to previously visited locations is no different than in the recurrent (compact) case, since the integral diverges near $+\infty$.

Large volume approximation of splitting probabilities We now focus on identifying the universal behavior of splitting probabilities. By making use of results presented in [O'Shaughnessy & Procaccia 1985], and in the large confining volume limit $V \to \infty$ with fixed source and targets, equations (2.2), (2.4) and (2.8) can be shown to hold for general Markovian processes, such that

$$\pi_{\mathbf{x_1},\underline{\mathbf{x_2}}}(\mathbf{x_0}) \underset{V\to\infty}{\sim} \frac{G_\infty(r_{02}) - G_\infty(r_{10}) + G_\infty(0) - G_\infty(r_{12})}{2\left[G_\infty(0) - G_\infty(r_{12})\right]}. \tag{2.13}$$

In the case of scale-invariant processes, differences of G_∞ are evaluated according to equation (2.12). As a result, three distinct universal behaviors arise for the splitting probabilities, depending on the d_w/d_f ratio. Denoting r_{ij} the distance between $\mathbf{x_i}$ and $\mathbf{x_j}$, we finally obtain:

$$\pi_{\mathbf{x_2},\underline{\mathbf{x_1}}}(\mathbf{x_S}) \sim \begin{cases} \dfrac{A + B(r_{1S}^{d_w-d_f} - r_{2S}^{d_w-d_f} - r_{12}^{d_w-d_f})}{2(A - Br_{12}^{d_w-d_f})} & \text{for } d_w < d_f \\ \dfrac{A + B\log\left(\frac{r_{2S}\ r_{12}}{r_{1S}}\right)}{2\left[A + B\log(r_{12})\right]} & \text{for } d_w = d_f \\ \dfrac{r_{2S}^{d_w-d_f} + r_{12}^{d_w-d_f} - r_{1S}^{d_w-d_f}}{2r_{12}^{d_w-d_f}} & \text{for } d_w > d_f \end{cases} \tag{2.14}$$

where A and B are constants depending only on the scaling function Π in equation (2.11), and will be dealt with specifically for each example. Note that these three universal regimes entirely capture the functional dependence of the splitting probability in the geometrical disposition of the targets and source in the domain. For numerical validation, we refer the reader to [Condamin *et al.* 2007b] and [Bénichou & Voituriez 2014].

2.2 First moment of $C(s_0)$

With the previous exposition, we have shown that the large volume limit allows for the emergence of universal behavior for splitting probabilities. We now extend these results to the territory explored before reaching a target.

2.2.1 From splitting probabilities to the explored territory

Let us first consider a multi-dimensional lattice Ξ of N sites, with reflecting boundary conditions. Let s_T be a fixed target site inside the domain, and s_0 the starting point of a nearest neighbor random walk on the lattice. The visited territory before reaching s_T is then simply given by a sum of indicator random variables:

$$C(s_0) = \sum_{s\in\Xi} \mathbb{1}(s \text{ is visited before } s_T|s_0). \tag{2.15}$$

Importantly, $C(s_0)$ is evaluated along a single trajectory, such that the indicator functions $\mathbb{1}(.|s_0)$ are strongly correlated random variables. Intuitively, close-by sites have a high probability of being visited conjointly, whereas far-away sites may be considered as being independent. Taking the expectation in equation (2.15), we obtain an exact equation for the first moment $\langle C(s_0)\rangle$:

$$\langle C(s_0)\rangle = \sum_{s\in\Xi} \pi_{s_T,\underline{s}}(s_0). \tag{2.16}$$

Note that for hypercubic lattices, the H functions are known exactly (see Appendix B) and we obtain explicit expressions for the mean territory visited by making use of expression (2.4). Additionally, these exact expressions will serve as test cases against asymptotic large volume results.

2.2.2 Spherical assumption for fractal spaces

From now on, we consider general isotropic scale-invariant stochastic processes characterized by the walk dimension d_w, and whose propagator obeys the scaling relation (2.11). Following the classical treatment of diffusion in fractal spaces proposed in [O'Shaughnessy & Procaccia 1985] and [Havlin & ben Avraham 1987], we approximate fractal media by a spherically symmetric domain of radius R, solid angle Ω_{d_f} and fractal dimension d_f. In turn, the volume V of the confining domain is given by:

$$V = \Omega_{d_f} R^{d_f}, \tag{2.17}$$

and the accessible volume between two shells located at radius r and $r + \mathrm{d}r$ reads:

$$\mathrm{d}V = \Omega_{d_f} d_f r^{d_f - 1}. \tag{2.18}$$

Note on the numerical evaluation of observables in fractal spaces. In order to satisfy the isotropy requirement, the scale-invariance property, and the spherical assumption in the numerical evaluation of the explored territory, $C(s_0)$ must be averaged over all source and target pairs (s_0, s_T) such that the source-target distance is constant.

Note on the definition of distances. We systematically denote the distance between two points $\mathbf{x_i}$, $\mathbf{x_j}$ as r_{ij}. We stress that for continuous spaces, this distance is the standard euclidean distance, while for graphs and lattices it should be understood as the *chemical distance*, *ie* the smallest number of steps needed to reach one point starting from the other.

Average territory for scale-invariant processes. Under the spherical assumption, the expected visited territory for Markovian scale-invariant processes is given by:

$$\langle C(\mathbf{x_0}) \rangle = \int_V \pi_{\mathbf{x_T}, \underline{\mathbf{x}}}(\mathbf{x_0}) \mathrm{d}\mathbf{x} \tag{2.19}$$

where $\mathbf{x_T}$ is the location of the target. In the $R \to \infty$ limit, with fixed source $\mathbf{x_0}$ and target $\mathbf{x_T}$, we exploit the asymptotic behavior of the splitting probability (2.14) and exhibit universal behavior of $\langle C(\mathbf{x_0}) \rangle$.

2.2.3 Compact case

Universal scaling form. In the compact (or recurrent) case $d_w > d_f$, we recall the asymptotic expression of the splitting probability (see (2.14)):

$$\pi_{\mathbf{x_2}, \underline{\mathbf{x_1}}}(\mathbf{x_0}) \sim \frac{r_{20}^{d_w - d_f} + r_{12}^{d_w - d_f} - r_{10}^{d_w - d_f}}{2 r_{12}^{d_w - d_f}} \tag{2.20}$$

and denote $\alpha = d_w - d_f$, with $\alpha > 0$. Using equation (2.19), the mean explored territory directly reads

$$\langle C(\mathbf{x_0}) \rangle \sim \int_V \frac{r_{T0}^{\alpha} + r_{1T}^{\alpha} - r_{01}^{\alpha}}{2 r_{1T}^{\alpha}} \mathrm{d}\mathbf{x_1}. \tag{2.21}$$

Denoting the source-target distance $r_s \equiv r_{TO}$, we rewrite the integral as:

$$\langle C(r_s)\rangle \sim r_s^\alpha \int_V \frac{1+\left(\frac{r_{1T}}{r_s}\right)^\alpha - \left(\frac{r_{01}}{r_s}\right)^\alpha}{2r_{1T}^\alpha} \mathrm{d}\mathbf{x_1}. \tag{2.22}$$

Under the spherical assumption, the distance r_{01} can always be written as:

$$\begin{aligned} r_{01} &= r_s\sqrt{1+2\frac{r_{T1}}{r_s}\cos(\theta)+\left[\frac{r_{T1}}{r_s}\right]^2} \\ &\equiv r_s f_\perp\left(\frac{r_{T1}}{r_s},\theta\right) \end{aligned} \tag{2.23}$$

where θ is the angle between the vectors $\mathbf{x_1}-\mathbf{x_0}$ and $\mathbf{x_T}-\mathbf{x_0}$. Plugging into the integral, we obtain

$$\langle C(r_s)\rangle \sim r_s^\alpha \int_V \frac{1+\left(\frac{r_{T1}}{r_s}\right)^\alpha - \left(f_\perp(\frac{r_{T1}}{r_s},\theta)\right)^\alpha}{2r_{T1}^\alpha} \mathrm{d}\mathbf{x_1}, \tag{2.24}$$

and integrating out the solid angle yields

$$\langle C(r_s)\rangle \sim d_f\Omega_{d_f} r_s^\alpha \int_V \frac{1+\left(\frac{r_{T1}}{r_s}\right)^\alpha - \left(f_*(\frac{r_{T1}}{r_s})\right)^\alpha}{2r_{T1}^\alpha} r_{T1}^{d_f-1}\mathrm{d}r_{T1} \tag{2.25}$$

where f_* is the integrated version of $f_\perp$. In the large volume limit, with $\mathbf{x_0}$ and $\mathbf{x_T}$ fixed, one has $R \gg r_s$, and we make the change of variable $u = \frac{r_{T1}}{r_s}$ to finally obtain:

$$\langle C(r_s)\rangle \sim \Omega_{d_f}\ d_f\ r_s^{d_f}\int_0^{\frac{R}{r_s}} \frac{1+u^\alpha - f_*(u)^\alpha}{2u^\alpha} u^{d_f-1}\mathrm{d}u. \tag{2.26}$$

Note that the determination of the upper bound of the integral relies on the fact that the target is located far from the boundaries of the confining domain. In turn, the scaling of integral (2.26) with R is controlled by the integrand $J(u) = \frac{1+u^\alpha-f_*(u)^\alpha}{2u^\alpha}u^{d_f-1}$.

1. We first make sure that no divergences occur near 0. As $u \to 0$, $f_*(u) = 1+\beta u + o(u)$ (with β constant), such that the integrand asymptotically behaves as

   $$J(u) \underset{u\to 0}{\propto} u^{\min(d_f-1,\ 2d_f-d_w)}. \tag{2.27}$$

   In all known cases, $d_w \leq d_f+1$ and $d_f \geq 1$. As a result $J(u) \propto u^{d_f-1}$ for small u, and the integral is well-behaved near 0.

2. We now focus on divergences occurring close to $+\infty$. In that case, $f_*(u) \sim u$, such that the integrand is equivalent to

   $$J(u) \underset{u\to\infty}{\propto} u^{2df-d_w-1}. \tag{2.28}$$

   Since $d_w \leq d_f+1 \leq 2d_f$, we have $2df - d_w - 1 \geq -1$ and the integral diverges as R goes to infinity.

Making use of equation (2.26) we finally obtain the universal behavior of the mean explored territory in the large volume limit $R \gg r_s$:

$$\frac{\langle C(r_s\rangle)}{V} \underset{1\ll r_s \ll R}{\sim} \frac{d_f}{2(2d_f - d_w)} \left(\frac{r_s}{R}\right)^{d_w - d_f}. \tag{2.29}$$

Note that in the specific case $d_w = 2d_f$, the divergence is logarithmic and we obtain:

$$\frac{\langle C(r_s\rangle)}{V} \sim -\frac{d_f}{2} \frac{r_s^{d_f}}{R^{d_f}} \log\left(\frac{r_s}{R}\right). \tag{2.30}$$

Let us make a few comments:

- For a given stochastic process, d_f and d_w are easily measurable from data. In turn, the result (2.29) can be straightforwardly applied.
- In the large volume limit, the exact position of the target $\mathbf{x_T}$ becomes irrelevant. However, the source-target distance r_s is highly relevant for the average of the visited territory. In particular, for a compact system, starting close to the target ensures that it is found quickly, so that only a small fraction of the total domain is visited.
- We stress that while equation (2.29) is an asymptotic result relying heavily on the spherically symmetric description of the environment, its range of applicability is fairly wide. In the following paragraphs, we aim to highlight its relevance across examples of representative random walks.

Normal random walk on the periodic ring. Consider first the symmetric random walk on a periodic one-dimensional lattice of N sites, for which $d_f = 1$, $d_w = 2$ and $\pi_{s_1,\underline{s_2}}(s_0) = \frac{s_0 - s_1}{s_2 - s_1}$ (see chapter 1). For a target site located at 0, the mean territory explored is given by:

$$\langle C(s_0)\rangle = \sum_{s=0}^{N-1} \pi_{0,\underline{s}}(s_0). \tag{2.31}$$

Defining the nth harmonic number as $h_n = \sum_{k=1}^{n} \frac{1}{k}$, we obtain for $s_0 > 0$:

$$\langle C(s_0)\rangle = 2 + (N - s_0)(h_{N-1} - h_{N-s_0+1}) + s_0(H_{N-1} - h_{s_0}), \tag{2.32}$$

which is asymptotically equivalent to

$$\langle C(s_0)\rangle \underset{1\ll s_0 \ll N}{\sim} -N \frac{s_0}{N} \log\left(\frac{s_0}{N}\right). \tag{2.33}$$

The spherical assumption corresponds to $R = N/2$, with R the farthest possible distance from the absorbing target located at site 0. Rewriting (2.33), we obtain:

$$\frac{\langle C(s_0)\rangle}{N} \underset{1\ll s_0 \ll R}{\sim} -\frac{s_0}{2R} \log\left(\frac{s_0}{R}\right), \tag{2.34}$$

in agreement with equation (2.30). As announced in chapter 1, we now focus on an example of one-dimensional lattice walk with non-connected span.

Riemann walks with $\mu > 1$. The Riemann walk [Hughes 1995, Mariz *et al.* 2001] - or discrete Levy Flight - is a classical example of discrete random walk with long ranged jumps. More precisely, the transition kernel is given by $p(s_1 \to s_2) = \frac{1}{2\zeta(1+\mu)}|s_1 - s_2|^{-(1+\mu)}$, where ζ is the Riemann Zeta function, needed for normalization. In the specific case $1 < \mu < 2$, the jump distribution has a well defined mean but no variance. As a result, the Riemann walk displays a super-diffusive behavior [Bouchaud & Georges 1990], such that its typical position $\tilde{x}_n$ after n steps scales as:

$$\tilde{x}_n \underset{n\gg 1}{\propto} n^{\frac{1}{\mu}}, \tag{2.35}$$

in contrast to classical diffusive systems, for which $\tilde{x}_n \underset{n\gg 1}{\propto} \sqrt{n}$ (*eg* the normal walk). In that case $d_w = \mu$, and choosing $1 < \mu < 2$ on a one-dimensional periodic lattice, we obtain a paradigmatic example of compact random walk with non-connected span. We display in figure 2.2 the agreement between numerical simulations and the universal formula (2.29).

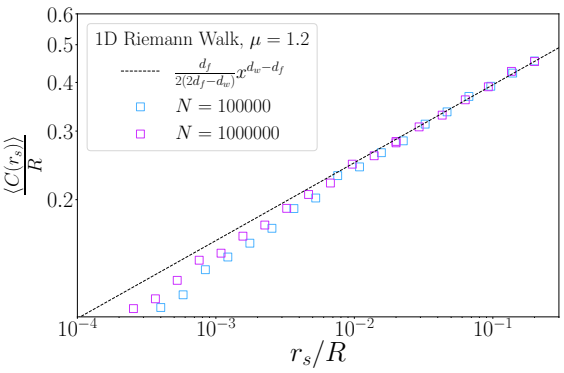
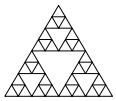


Figure 2.2: Mean explored territory for a compact Riemann Walk with $\mu = 1.2$. The r_s dependence follows prediction (2.29). Importantly, the behavior we aim to describe is only true in the $1 \ll r_s \ll R$ limit, as indicated by the deviation from the theory at small r_s/R values.

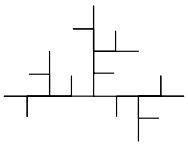
We point out that there exists two conventions for discovering new sites: the *arrival* or *crossing* convention [Dybiec *et al.* 2016]. In the arrival convention, sites are visited only when the walker effectively lands on the site. In the crossing convention, any site that is flown over is considered as visited. To enforce the non-connected span property, we follow the arrival convention. In particular, the walker can fly over the target 0 a number of times before actually landing on it.

Fractal graphs. We now consider nearest neighbor random walks on spaces with non integer fractal dimension d_f, and focus on two classical examples of deterministic fractal graphs, the *T-graph* and the *Sierpinski Gasket* [Havlin & ben Avraham 1987] (see figure 2.3), for which the pairs (d_w, d_f) are respectively equal to $\left(\frac{\log(6)}{\log(2)}, \frac{\log(3)}{\log(2)}\right)$ and $\left(\frac{\log(5)}{\log(2)}, \frac{\log(3)}{\log(2)}\right)$. Note that we consider diffusion on the finite graphs of order n, such that the walker evolves in a bounded domain. The large volume limit is obtained by increasing n. We emphasize the accuracy of the prefactor predicted in equation (2.29) and display numerical agreement in figure 2.3(d-e). In turn, our results highlight the relevancy of the *spherically symmetric* description of fractal structures.

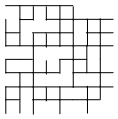
Disordered medium. To conclude, we investigate a paradigmatic example of transport in disordered media, namely diffusion in a critical two-dimensional percolation cluster [Havlin & ben Avraham 1987, Bouchaud & Georges 1990]. The critical percolation cluster is obtained from a square lattice of N sites by deleting each bond with a probability $p_c = 1/2$ and keeping the largest connected component. In this case, $d_f = 91/48$ is known exactly, and the walk dimension is approximately equal to $d_w \simeq 2.878$ [Havlin & ben Avraham 1987]. We display agreement between simulations and theory in figure 2.3(f). Of note, each percolation cluster is an independent realization of the percolating procedure described above, and $\langle C(r_s)\rangle$ is averaged over different clusters.



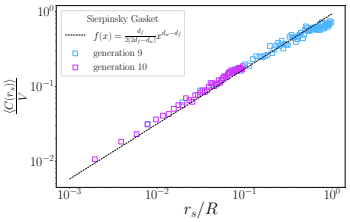
(a) The Sierpinski gasket is a deterministic fractal constructed by iteratively assembling the triangular structure depicted above. In its finite version, generation n contains $\frac{3}{2}(1+3^{n-1})$ sites.



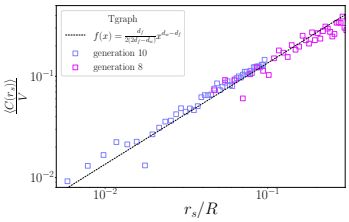
(b) The Tgraph is a deterministic fractal constructed by iteratively assembling the tree structure depicted above. In its finite version, generation n contains $3^n - 2$ sites.



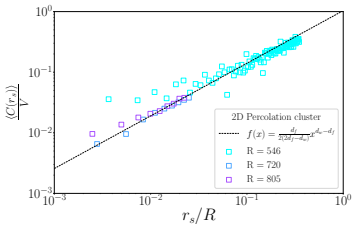
(c) The critical percolation cluster is obtained from a $2D$ lattice by deleting each bond with a probability $p = 0.5$ and keeping the largest connected component.



(d) Mean explored territory on Sierpinski gaskets of increasing size. The r_s dependence follows prediction (2.29). Importantly, the definition of R is left to interpretation. We define it as the longest possible distance between two nodes of the gasket.



(e) Mean explored territory on Tgraphs of increasing size. The r_s dependence follows prediction (2.29). The choice of R is made similarly to the Sierpinski gasket.



(f) Mean explored territory on critical percolation clusters. The r_s dependence follows prediction (2.29). All r_s pairs are averaged over different critical percolation clusters. The R parameter is also random and depends on each realization.

Figure 2.3

2.2.4 Non-compact case

Consider now the non-compact case, and define $\alpha = d_w - d_f$, with $\alpha < 0$. Plugging the large volume limit of the splitting probability (2.14)

$$\pi_{\mathbf{x_2},\underline{\mathbf{x_1}}}(\mathbf{x_0}) \sim \frac{A + B(r_{10}^{d_w-d_f} - r_{20}^{d_w-d_f} - r_{12}^{d_w-d_f})}{2(A - Br_{12}^{d_w-d_f})} \tag{2.36}$$

into equation (2.19) yields

$$\langle C(\mathbf{x_0})\rangle \sim \int_V \frac{A + B(r_{10}^{\alpha} - r_{T0}^{\alpha} - r_{1T}^{\alpha})}{2(A - Br_{1T}^{\alpha})} \mathrm{d}\mathbf{x_1}, \tag{2.37}$$

and we obtain:

$$\langle C(\mathbf{x_0})\rangle - \frac{V}{2} \sim \int_V \frac{B(r_{10}^\alpha - r_{T0}^\alpha)}{2(A - Br_{1T}^\alpha)} \mathrm{d}\mathbf{x_1}. \tag{2.38}$$

Similarly to the compact case, we denote $r_s \equiv r_{TO}$, integrate out the angular dependence and make the change of variable $u = r_{T1}/r_s$ to obtain:

$$\langle C(r_s)\rangle - \frac{V}{2} \sim B d_f \Omega_{d_f} r_s^\alpha r_s^{d_f} \int_0^{\frac{R}{r_s}} \frac{f_*(u)^\alpha - 1}{2\left[A - B(ur_s)^\alpha\right]} u^{d_f - 1} \mathrm{d}u \tag{2.39}$$

where the definition of $f_*(u)$ is identical to the compact case, and we denote the integrand $J(u) = \frac{f_*(u)^\alpha - 1}{2[A - B(ur_s)^\alpha]} u^{d_f - 1}$.

1. As $u \to 0$, recalling that $\alpha < 0$, $J(u) \propto u^{2d_f - d_w}$, and the integral is well-behaved since $d_f > d_w$.

2. For large u, $J(u) \propto u^{d_f - 1}$ and the integral diverges as R grows to infinity.

As a result, for large R and fixed source target distance r_s, the mean explored territory takes the following universal form:

$$\frac{\langle C(r_s)\rangle}{V} \underset{1 \ll r_s \ll R}{\sim} \frac{1}{2} - \frac{C_1}{r_s^{d_f - d_w}} \quad \text{with} \quad C_1 = \frac{B}{2A}. \tag{2.40}$$

Importantly:

- A and B are constant coefficients that depend on the microscopic details of the random walk. However, the plateauing feature displayed by equation (2.40) is universal, and the algebraic convergence speed depends on d_w and d_f only.

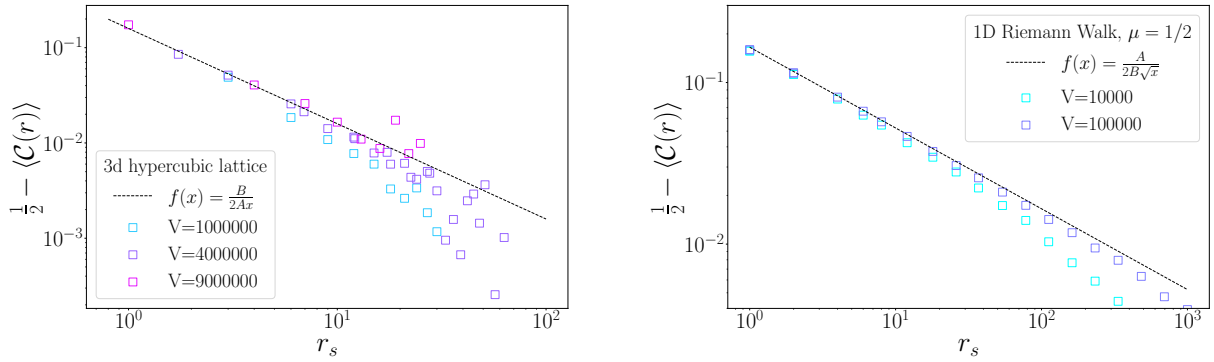
- The r_s dependence of $\langle C(r_s)\rangle$ is strikingly different from the compact case. As the source-target distance increases, its influence on the explored territory vanishes, and the memory of the starting point is lost.

- The limiting value $\langle C(r_s)\rangle \to V/2$ has a physical interpretation; in a non-compact exploration process, every site can be thought of as *equivalent*. In turn, the average behavior is simply to visit half the space before finding the target.

Similarly to the compact universal class, we illustrate the universal result (2.40) across two representative examples of non-compact processes: classical diffusion on 3 dimensional hypercubes and non-compact Riemann Walks in one-dimension.

Normal walk on a 3D hypercube. Since the walk dimension of the Normal walk is always $d_w = 2$, embedding it in a 3 dimensional space makes it non-compact. The A and B constants can be found in [Hughes 1995] and we display numerical agreement with formula (2.40) in figure 2.4(a). We emphasize that there is not fit parameter.

Riemann walk with $\mu < 1$. We follow the same procedure as for the compact Riemann walk and choose $0 < \mu < 1$ on a one-dimensional periodic lattice to obtain non-compact exploration. To determine A and B, we consider the mean first-passage time (MFPT) t_{0s} to 0 starting from s at distance r_s. For compact Riemann walks, the MFPT is known both exactly [Tejedor *et al.* 2011] and asymptotically reads $t_{0s} \sim N(A - Br_s^{\mu-1})$ for large $1 \ll r_s \ll N$ [Bénichou & Voituriez 2014]. As a result, we identify A and B and display agreement between (2.40) and numerical simulations in figure 2.4(b).



(a) Normal random walk on a 3 dimensional hypercube. As the volume increases, the mean visited territory converges towards the asymptotic limit (2.40). Note that convergence is slower than for the one-dimensional system. As the mean FPT scales with the confining volume V, simulations quickly become time consuming.

(b) Non-compact Riemann Walk. As the volume increases, the mean visited territory converges towards the asymptotic limit (2.40). We emphasize the absence of a fit parameter.

Figure 2.4

2.2.5 Marginally compact case

Having extensively covered the compact and non-compact cases, we conclude with the marginal case $d_f = d_w$. The large volume limit splitting probability is given by (2.14):

$$\pi_{\mathbf{x_2},\underline{\mathbf{x_1}}}(\mathbf{x_0}) \sim \frac{A + B\log\left(\frac{r_{20}\ r_{12}}{r_{10}}\right)}{2\left[A + B\log(r_{12})\right]} \tag{2.41}$$

and the mean explored territory reads:

$$\langle C(\mathbf{x_0})\rangle \sim \int_V \frac{A + B\log\left(\frac{r_{T0}\ r_{1T}}{r_{10}}\right)}{2\left[A + B\log(r_{1T})\right]} \mathrm{d}\mathbf{x_1}. \tag{2.42}$$

Similarly to the compact and non-compact cases, we denote $r_s \equiv r_{OT}$ and $r_{10} = r_s f_\perp(\frac{r_{1T}}{r_s}, \theta)$, integrate the angular dependence and obtain:

$$\langle C(\mathbf{r_s})\rangle \sim d_f \Omega_{d_f}\left[A + B\log(r_s)\right] \int_0^R \frac{1}{2\left[A + B\log(r_{1T})\right]} r_{T1}^{d_f-1} \mathrm{d}r_{T1}. \tag{2.43}$$

The integral is convergent near 0, and integration by parts yields:

$$\langle C(r_s)\rangle \sim d_f \Omega_{d_f} \left[A + B\log(r_s)\right]\left(\left[\frac{r_{T1}^{d_f}}{2d_f\left[A + B\log(r_{1T})\right]}\right]_0^R + \int_0^R \frac{B}{2\left[A + B\log(r_{1T})\right]^2} r_{T1}^{d_f-1} \mathrm{d}r_{T1}\right), \tag{2.44}$$

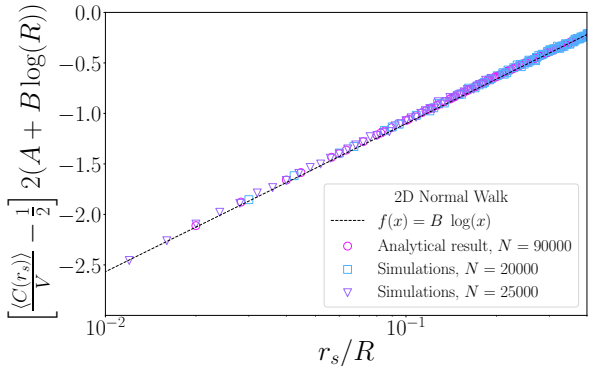
from which we finally obtain the universal behavior of $\langle C(r_s)\rangle$ in the large volume limit:

$$\frac{\langle C(r_s)\rangle}{V} \underset{1\ll r_s \ll R}{\sim} \frac{1}{2} + \frac{B\log\left(\frac{r_s}{R}\right)}{2\left[A + B\log(R)\right]}. \tag{2.45}$$

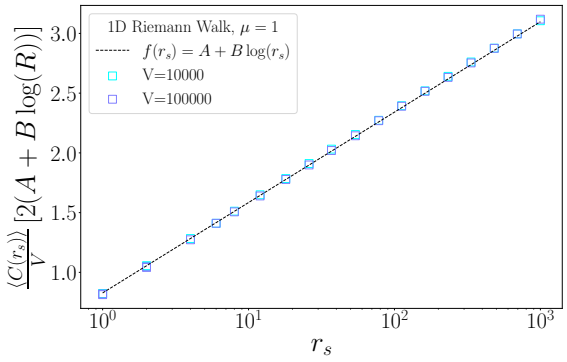
- As in the non-compact case, A and B are constant coefficients that depend on the microscopic details of the random walk. However, the logarithmic r_s dependence of $\langle C(r_s)\rangle$ is universal and process independent.
- The marginal case shares many similarities with the non-compact case. Indeed the mean explored territory converges toward $1/2$ as r_s increases. However, the convergence speed is logarithmic, much slower than the corresponding algebraic decay for the non-compact case.
- Marginal systems are not as widespread as compact or non-compact ones. We here focus on two specific examples for which both A and B can be computed: the normal walk on a two dimensional hypercube, and the marginal one-dimensional Riemann Walk.

Normal walk on a 2D lattice. The symmetric random walk on a 2D square lattice is the canonical example of a marginally compact system, and both $A \simeq 1.0293374$ and $B = 2/\pi$ can be found in [Hughes 1995]. Importantly, the H functions are known exactly in this case, such that combining equations (2.16) and (2.4) yields explicit values of $\langle C(r_s)\rangle$ for any lattice size. We display the agreement between the large volume asymptotic result (2.45), numerical simulations and exact analytical values in figure 2.5(a).

Riemann walk with $\mu = 1$ on 1D lattice We finally consider a Riemann Walk with $\mu = 1$ on the 1D periodic lattice of N sites. As in the compact case, the MFPT t_{0s} is known exactly, and is given by $t_{0s} \sim N(A + B\log(r_s))$ for $1 \ll rs \ll N$. In turn, we obtain explicit expressions of A and B, and display numerical agreement with simulations in figure 2.5(b).



(a) Symmetric random walk on a 2D lattice. The r_s dependence of the mean visited territory rescaled according to equation (2.45) collapses for various lattice sizes N. Additionally, the values of $C(r_s)$ obtained from the exact H functions also validate the asymptotic result (2.45) Note that the definition of R is chosen such that the volume of the spherical disk and 2D lattice match.



(b) 1D Riemann walk with $\mu = 1$. The r_s dependence of the mean visited territory rescaled according to equation (2.45) collapses for various lattice sizes V. Note however that while the asymptotic normalizing factor $2(A+B\log(R))$ is correct in the $V \to \infty$ limit, we here normalize by $\sum_{n=1}^{V/2} 2(A+B\log(n))$ (2.43) to obtain good convergence.

Figure 2.5

2.3 Complete distribution of $C(s_0)$

Let us briefly summarize the results obtained for the first moment of $C(s_0)$. In the large volume limit $V \to \infty$ and with fixed source and target such that $V \gg r_s^{d_f} \gg 1$, we identified three universality classes for the behavior of $\langle C(r_s)\rangle$, which fully capture the functional dependence on the geometrical parameters of the problem, independently of the microscopic details of the walk:

$$\frac{\langle C(r_s)\rangle}{V} \sim \begin{cases} \dfrac{1}{2} - \dfrac{B}{2Ar_s^{d_f-d_w}} & \text{for } d_w < d_f \\ \dfrac{1}{2} + \dfrac{B\log\left(\frac{r_s}{R}\right)}{2\left[A + B\log(R)\right]} & \text{for } d_w = d_f \\ \dfrac{d_f}{2(2d_f - d_w)}\left[\dfrac{r_s}{R}\right]^{d_w-d_f} & \text{for } d_w > d_f \end{cases} \tag{2.46}$$

Importantly, it appears from (2.46) that the rescaled variable $\mathcal{C}(r_s) = C(r_s)/V$, defined as the *fraction* of the domain visited before reaching the target, is of particular relevance. In the following sections, we uncover the universal behavior of the full distribution of $\mathcal{C}$ in the large volume limit.

2.3.1 Non-compact and marginal cases

As already noted, the marginal and non-compact cases display strong similarities with regards to the scaling form of $\langle \mathcal{C}(r_s)\rangle$. In particular $\langle \mathcal{C}(r_s)\rangle$ converges towards $1/2$ as r_s increases. Based on that remark, we handle both cases conjointly.

Second moment. Consider first a lattice Ξ of N sites with a target at site s_T and a source-target distance r_s. In that case, $\mathcal{C}(r_s)^2$ is given by:

$$\mathcal{C}^2(r_s) = \frac{1}{N^2} \sum_{s_1,s_2\in\Xi} \mathbb{1}(s_1 \text{ is visited before } s_T \mid s_0)\mathbb{1}(s_2 \text{ is visited before } s_T \mid s_0), \tag{2.47}$$

from which the second moment is obtained as:

$$\langle \mathcal{C}^2(r_s)\rangle = \frac{1}{N^2} \sum_{s_1,s_2\in\Xi} \pi_{s_T,\underline{s_1},\underline{s_2}}(s_0), \tag{2.48}$$

where $\pi_{s_T,\underline{s_1},\underline{s_2}}(s_0)$ is defined as the three point splitting probability of reaching both s_1 and s_2 before s_T. We now focus on the large volume limit of equation (2.48).

1. Specifying the order in which the sites are visited yields the simplified exact result

$$\langle \mathcal{C}^2(r_s)\rangle = \frac{2}{N^2} \sum_{s_1,s_2\in\Xi} \pi_{s_T,s_1,\underline{s_2}}(s_0)\pi_{s_T,\underline{s_1}}(s_2). \tag{2.49}$$

2. On the one hand, the asymptotic results obtained for $\langle \mathcal{C}(r_s)\rangle$ indicate that sums of splitting probabilities are controlled by sites located far from s_T and s_0. Assuming further that the divergences come from sites s_1 and s_2 far apart from each other, we first rewrite $\pi_{s_T,\underline{s_1}}(s_2)$ as:

$$\pi_{s_T,\underline{s_1}}(s_2) \sim \frac{1}{2}. \tag{2.50}$$

   On the other hand, the three point splitting probability can be exactly expressed in terms of H functions. We show in Appendix C that when sites s_1 and s_2 are far from s_T, the three point splitting probability asymptotically reads:

$$\pi_{s_T,s_1,\underline{s_2}}(s_0) \sim \pi_{s_T,\underline{s_2}}(s_0) \left[\frac{2(H_0 - H_{T2})}{H_{1T} - H_{T2} + 3(H_0 - H_{T2})}\right] \tag{2.51}$$

   where $H_0 \equiv H_{ii}$. Additionally, for s_1, s_2 far from s_T, $H_{1T} \sim H_{2T}$, such that:

$$\pi_{s_T,s_1,\underline{s_2}}(s_0) \sim \frac{2}{3}\pi_{s_T,\underline{s_2}}(s_0). \tag{2.52}$$

3. In the large volume limit, equation (2.49) is finally recast as:

$$\begin{aligned}\langle \mathcal{C}^2(r_s)\rangle &\sim \frac{2}{N^2} \sum_{s_1,s_2\in\Xi} \frac{2}{3}\pi_{s_T,\underline{s_2}}(s_0) \times \frac{1}{2} \\ &\sim \frac{2}{3N} \sum_{s_2\in\Xi} \pi_{s_T,\underline{s_2}}(s_0) \\ &\sim \frac{2}{3}\langle \mathcal{C}(r_s)\rangle.\end{aligned} \tag{2.53}$$

Importantly both the first and second moment of $\mathcal{C}(r_s)$ have the same asymptotic scaling behavior, stemming mostly from observations (2.50) and (2.52). To proceed further, we extend this argument to higher moments of $\mathcal{C}(r_s)$, and give a *geometrical interpretation* to the form of the n-point splitting probabilities.

***n*-point splitting probabilities and further moments.** Consider a given set $(s_1, ..., s_n)$ of n targets far from each other, and far from s_T, with s_0 at a distance r_s from s_T. Define $\pi^*(r_s)$ the probability that the walker moves far enough from s_T so that all $n+1$ targets $(s_1, .., s_n, s_T)$ become equivalent. As a result, we rewrite the large volume limit of the n-point splitting probabilities as:

$$\pi_{s_T,s_1,...\underline{s_n}}(r_s) \sim \frac{\pi^*(r_s)}{n+1} \quad \text{and} \quad \pi_{s_T,s_1,...\underline{s_{n-1}}}(s_n) \sim \frac{1}{n} \tag{2.54}$$

Importantly, by making use of the definition (2.16) of $\langle \mathcal{C}(r_s) \rangle$, we obtain

$$\pi^*(r_s) = 2\langle \mathcal{C}(r_s) \rangle. \tag{2.55}$$

Consequently, the asymptotic behavior of the n^{th} moment is given by

$$\begin{aligned} \langle \mathcal{C}^n(r_s) \rangle &= \frac{1}{N^n} \sum_{s_1,...,s_n \in \Xi} \pi_{s_T,\underline{s_1},...,\underline{s_n}}(s_0) \\ &\sim \frac{n!}{N^n} \sum_{s_1,...,s_n \in \Xi} \pi_{s_T,s_1...,\underline{s_n}}(s_0) \dots \pi_{s_T,\underline{s_1}}(s_2) \\ &\sim \frac{n!}{N^n} \sum_{s_1,...,s_n \in \Xi} \frac{\pi^*(r_s)}{n+1} \frac{1}{n} \cdots \frac{1}{2} \\ &\sim \frac{\pi^*(r_s)}{n+1}. \end{aligned} \tag{2.56}$$

Complete distribution of $\mathcal{C}(r_s)$ for non-compact and marginal processes. Since the asymptotic large volume behavior of all moments is known, we derive the full distribution of $\mathcal{C}(r_s)$. To that end, we consider two types of trajectories: those that travel far enough from s_T, with associated probabilistic weight $\pi^*(r_s)$ and those that stay in the vicinity of s_T. In turn, the distribution of $\mathcal{C}$ is given as a sum of two conditional distributions:

$$P(\mathcal{C} = x|r_s) = \pi^*(r_s)\mu_1(x|r_s) + [1 - \pi^*(r_s)]\, \mu_2(x|r_s). \tag{2.57}$$

- For fixed r_s, as $V \to \infty$, the explored territory for trajectories staying close to s_T shrinks to 0, and $\mu_2(x|r_s) \to \delta(x)$.
- In turn , the only possible form for $\mu_1(x|r_s)$ to satisfy the scaling (2.56) of the moments of $\mathcal{C}(r_s)$ is $\mu_1(x|r_s) = 1$, hence a *uniform* distribution.

Finally, for non-compact and marginal processes, we uncover the universal behavior of the full distribution of $\mathcal{C}(r_s)$ in the large volume limit:

$$P(\mathcal{C}=x|r_s) \underset{1\ll r_s\ll R}{\sim} 2\langle \mathcal{C}(r_s)\rangle + [1-2\langle \mathcal{C}(r_s)\rangle]\,\delta(x), \tag{2.58}$$

and we display agreement with simulations in figure 2.6 for a variety of non-compact and marginal cases.

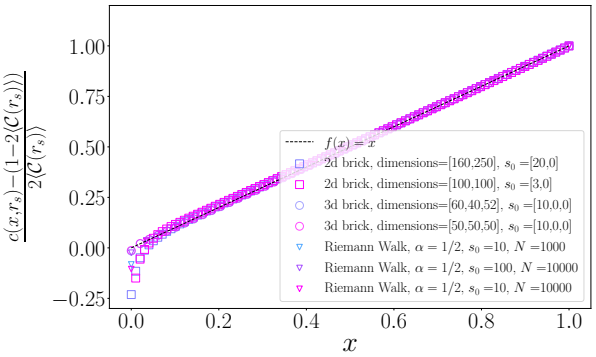


Figure 2.6: Cumulative distribution $c(x,r_s)=\int_0^x P(\mathcal{C}=u|r_s)\mathrm{d}u$ for a variety of non-compact and marginal processes. The cumulative distributions collapse on a single r_s independent function when rescaled according to (2.58). Note that our theory applies for domains of arbitrary shape. As an example we focus here on random walks in paralelepipedic domains.

A short comment on the golden coupon. The universal behavior (2.57) of $\mathcal{C}(r_s)$ directly parallels the *golden coupon* solution exhibited in chapter 1 in the case of the systematic relocation random walk. In the large volume limit, if the walker moves far enough from the target, all N sites of the domain become equivalent. In turn, the order in which they are discovered is uniformly distributed over all possible permutations of $[\![1,N]\!]$, and the conditional distribution of $\mathcal{C}(r_s)$ is uniform.

2.3.2 Compact case

The non-compact universal form (2.58) stems mainly from assumption (2.54), which is not expected to hold in the compact case. However, partitioning trajectories in a fashion similar to (2.57) is still exact. In the spherically symmetric picture, we introduce the *shell* splitting probability $\pi^{shell}(r_s,R_{out})$, defined as the probability to reach an outer shell of radius R_{out}, starting at a distance r_s from the absorbing target, which is asymptotically given by [Levernier *et al.* 2018]:

$$\pi^{shell}(r_s,R_{out}) \underset{1\ll r_s\ll R_{out}}{\propto} \left[\frac{r_s}{R_{out}}\right]^{d_w-d_f}. \tag{2.59}$$

For a bounding domain of radius R, we partition trajectories on whether or not the outer boundary of the confining domain is reached before s_T and obtain:

$$P(\mathcal{C}=x|r_s)=\pi^{shell}(r_s,R)\mu_1(x|r_s)+\left[1-\pi^{shell}(r_s,R)\right]\mu_2(x|r_s). \tag{2.60}$$

On trajectories conditioned to reach the boundary. We first focus on the distribution of $\mathcal{C}(r_s)$ conditioned to reach the boundary. For fixed r_s and $R \to \infty$, $\mu_1(x|r_s)$ becomes independent of r_s. Indeed, the domain visited prior to reaching the boundary is negligible compared to the domain visited after the boundary has been found. Additionally, $\mu_1(x) \to 0$ as $x \to 0$, since trajectories reaching the boundary must visit a substantial fraction of the domain.

On trajectories conditioned to never reach the boundary. Trajectories that never reach the boundary are statistically equivalent to unbounded trajectories. Consequently, we focus on the distribution of the number of sites visited in the infinite medium, before reaching the target[3]. For unbounded processes, the distribution of the farthest shell reached before finding s_T is obtained from the shell splitting probability:

$$P(\text{farthest reached shell is at distance } R_{out}|r_s) = -\left.\frac{\mathrm{d}}{\mathrm{d}R}\pi^{shell}(r_s,R)\right|_{R=R_{out}} \underset{1\ll r_s\ll R_{out}}{\propto} \frac{r_s^{d_w-d_f}}{R_{out}^{d_w-d_f+1}} \tag{2.61}$$

Note that equation (2.61) is the spherical equivalent of the one-dimensional equation (1.33) relating the distribution of the maximum and the splitting probability. Furthermore, for compact exploration processes, the number of visited sites in a ball of radius R is known to scale as R^{d_f}. In turn, the distribution of the number $C^\infty(s_0)$ of distinct sites visited before reaching s_T in the infinite medium is asymptotically given by

$$P(C^\infty = n|r_s) \propto \frac{r_s^{d_w-d_f}}{n^{\frac{d_w}{d_f}}}. \tag{2.62}$$

As a result, the conditional distribution of the visited fraction reads:

$$\mu_2(x|r_s) \underset{1\ll r_s\ll R}{\propto} \left[\frac{r_s}{R}\right]^{d_w-d_f} \frac{1}{x^{\frac{d_w}{d_f}}} \tag{2.63}$$

Plugging equation (2.63) into (2.57), we finally obtain a universal scaling form for the distribution of $\mathcal{C}(r_s)$, valid in the large volume limit:

$$P(\mathcal{C} = x|r_s) \underset{1\ll r_s\ll R}{\propto} \left[\frac{r_s}{R}\right]^{d_w-d_f} \left[\frac{1}{x^{\frac{d_w}{d_f}}} + \tilde{\mu}_1(x)\right], \tag{2.64}$$

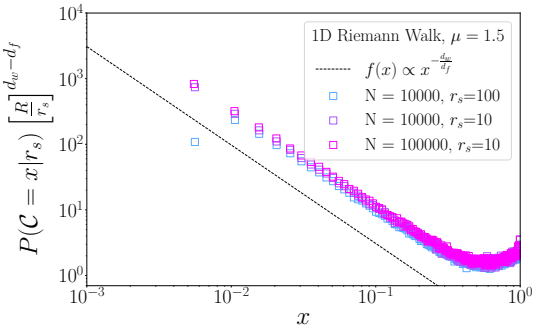
where $\tilde{\mu}_1$ is a process dependent but r_s and R independent function. We conclude this part with a few important remarks:

- The dependence of $\mathcal{C}(r_s)$ on the geometrical parameters r_s and R is fully explicit in equation (2.64). In turn, the scaling of all moments of $\mathcal{C}(r_s)$ with r_s and R is explicit, and important quantities, such as the fluctuations of $\mathcal{C}(r_s)$, can be analyzed.

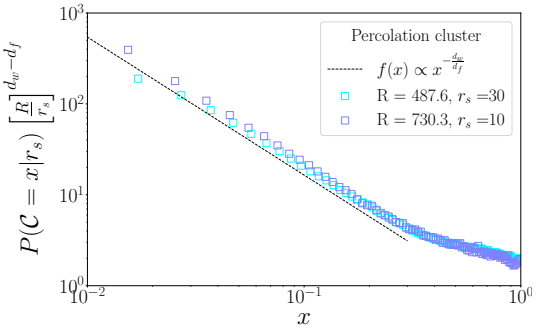
[3] In the *compact* case, the target is reached with probability 1.

- For $x \ll 1$, $\tilde{\mu}_1(x) = o(x^{-d_w/d_f})$, and the distribution of $\mathcal{C}(r_s)$ displays an algebraic behavior, characterized by d_w and d_f only. However, as $x \to 1$, the functional form of $\tilde{\mu}_1(x)$ is process dependent, and left undetermined in our approach.

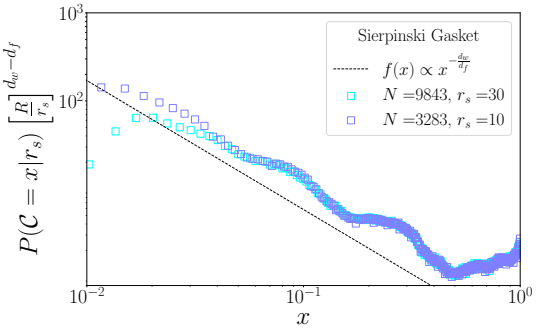
- Finally, we illustrate the relevance and broad range of applicability of equation (2.64) in figure 2.7, across the various compact processes studied in the previous section.



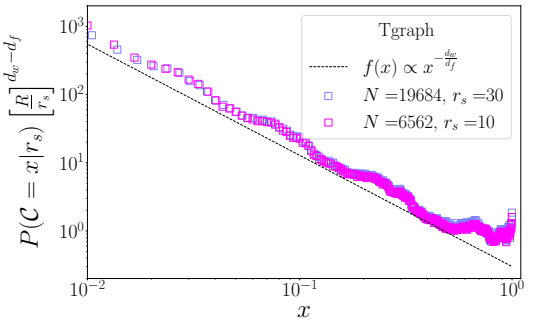
(a) Distribution of $\mathcal{C}(r_s)$ for a compact Riemann walk with $\mu = 3/2$ on a ring of N sites. Upon rescaling according to equation (2.64), the distributions collapse to a R and r_s independent function, whose small x behavior is algebraic. R is taken equal to $N/2$.



(b) Distribution of $\mathcal{C}(r_s)$ for 2D critical percolation clusters. The distributions are averaged over different source-target pairs, and different realizations of percolation clusters, starting from a $2D$ lattice with fixed number of sites 150^2 (blue) and 200^2 (purple). R is the mean maximal node distance over these realizations.



(c) Distribution of $\mathcal{C}(r_s)$ for the Sierpinski gasket. The distributions are averaged over different source-target pairs. N is the number of nodes for a given generation of the Sierpinski gasket, and R is the maximal node distance.



(d) Distribution of $\mathcal{C}(r_s)$ for the Tgraph. The distributions are averaged over different source-target pairs. N is the number of nodes for a given generation of the Tgraph, and R is the maximal node distance.

Figure 2.7

2.4 Conclusion

To go further than the one-dimensional random walks with connected span studied in chapter 1, we focus on general Markovian scale-invariant processes, characterized by their fractal dimension d_f and walk dimension d_w. Guided by earlier results and emergent universal behavior for first-passage observables in large confining domains, such as the first-passage time or splitting probabilities [Bénichou & Voituriez 2014], we investigate the *large volume limit* of the explored territory before reaching a target. In that case, both the source s_0 and target s_T are inside the bulk of the domain, and the source target distance r_s is small with respect the lengthscale $R \propto V^{1/d_f}$ of the confining domain.

Considering first the average $\langle C(r_s)\rangle$ of the explored territory, we identify three universality classes, namely *compact*, *marginal* and *non-compact*, for which the functional dependence of $\langle C(r_s)\rangle$ on the geometrical parameters r_s and R is independent of the microscopic transport details. We stress that the resulting scaling forms provide physical insight on the thoroughness of the exploration. In particular, marginal and non-compact processes tend to dodge the target, even for small r_s.

In fact, we show that these universality classes hold for the full distribution of the visited fraction $\mathcal{C}(r_s) = C(r_s)/V$ of the domain, which displays the following universal asymptotic behavior:

$$\begin{aligned} P(\mathcal{C}=x|r_s) &\underset{1\ll r_s\ll R}{\sim} 2\langle \mathcal{C}(r_s)\rangle + [1-2\langle \mathcal{C}(r_s)\rangle]\,\delta(x) && d_f \geq d_w \\ P(\mathcal{C}=x|r_s) &\underset{1\ll r_s\ll R}{\propto} \left[\frac{r_s}{R}\right]^{d_w-d_f} \left[\frac{1}{x^{\frac{d_w}{d_f}}} + \tilde{\mu}_1(x)\right] && d_f < d_w. \end{aligned} \tag{2.65}$$

The marginal and non-compact cases ($d_w \leq d_f$) are most striking. Indeed, the distribution of $\mathcal{C}(r_s)$ is quasi-uniform, and the Rosenstock survival probability discussed at length in chapter 1 is straightforwardly obtained, with potential applications in realistic systems, such as chemical reaction processes in three-dimensional gases or liquids. In the compact case, we stress that although the $\tilde{\mu}_1(x)$ function is process dependent, we expect its behavior near $x = 1$ to display some universal features, which still need to be investigated.

Taken together, chapter 1 and 2 draw a comprehensive picture of the territory explored before reaching a target in a confining domain. In particular, we provide a framework to evaluate its distribution, either exactly, or asymptotically in the large volume limit, which is valid across paradigmatic examples of stochastic processes. However, focusing on the geometrical description of the domain exploration, we completely neglect the intrinsic dynamical aspects. In particular, the time needed to reach the target conveys crucial information on the explored territory. For instance, shorter explorations hint at smaller explored territory.

In the following chapter, we unveil connections between the geometrical and kinetic aspects of exploration, by investigating the joint distribution of the explored territory and first-passage time, which fully quantifies their coupling.

Chapter 3

Joint statistics of space and time

Contents

The statistics of first-passage times (FPTs) to a target site of interest have proved to play a key role in determining the *kinetics* of space exploration [Redner 2001, Metzler *et al.* 2014, Bray *et al.* 2013]. In particular, the case of first-passage times in confined domains appears to be highly relevant to assess the efficiency of target search processes, and has lead to an important activity [Condamin *et al.* 2007b, Bénichou & Voituriez 2014, Cheviakov *et al.* 2010, Schuss *et al.* 2007]. On the other hand, the territory explored before reaching a given target in a confining domain, which characterizes the *geometry* of exploration, has been thoroughly investigated in chapters 1 and 2. Intuitively, the FPT and territory are highly correlated random variables: with high probability, longer exploration times lead to larger explored territories. However, even for one-dimensional processes, the precise quantification of the coupling between these two variables has been largely unexplored. The aim of this chapter is to fully elucidate the interplay between the FPT and the territory explored before reaching the target, for general one-dimensional processes.

To that end, we introduce the joint distribution $\sigma(s, n|s_0)$ of the number of sites visited s, and the FPT n to 0, for one-dimensional lattice random walks starting from s_0. Importantly, this joint distribution completely encompasses the coupling between *space and time* observables. Similarly to chapter 1, we first present a systematic methodology to derive σ in the case of one-dimensional Markovian walks with connected span. In turn, we illustrate our method by computing σ for representative examples of one-dimensional lattice random walks.

We then naturally extend this methodology to continuous and reflecting one-dimensional geometries. Additionally, in the large n and s limit, we show that the joint distribution takes a

universal scaling form, which holds for general scale-invariant processes and applies in particular to classical non-Markovian processes, such as the Fractional Brownian Motion (FBM) or the Random Acceleration Process (RAP) [Bicout & Burkhardt 2000].

We conclude this chapter with two important applications. First, the derivation of the joint law grants access to conditional distributions of either the FPT or the territory, which we suggest can be of key importance for experimental setups where only partial information is available. In particular, one can answer questions such as: knowing that a particle came back after n steps, how much of the space did it visit? Second, we show that the joint distribution is an essential tool for the derivation of the persistence properties of a strongly correlated random walk, the Self Avoiding True Walk (SATW) [Sapozhnikov 1994], for which the explicit form of the FPT distribution has yet to be derived. Of note, the results presented here are mostly contained in [Klinger *et al.* 2022a].

3.1 Joint distribution - systematic derivation

We first consider the case of Markovian lattice walks with connected span [1]. For a Markovian random walk, the propagator $P(s,n|s_0)$ defined as the probability for the walker to be at site s on the n^{th} step, obeys the forward equation [Van Kampen 1992]

$$P(s,n+1|s_0) = \mathcal{L}_s^\dagger P(s,n|s_0), \tag{3.1}$$

where the forward operator $\mathcal{L}^\dagger$ acting on s is defined by the microscopic transport rules. In turn, denoting the adjoint backward operator $\mathcal{L}$ acting on s_0, the propagator obeys the backward equation:

$$P(s,n+1|s_0) = \mathcal{L}_{s_0} P(s,n|s_0). \tag{3.2}$$

Importantly, for one-dimensional random walks with connected span, the joint distribution $\sigma(s,n|s_0)$ of the number of visited sites s and the FPT n to 0 is strictly equivalent to the joint distribution of the maximum and the FPT (see figure 3.1 for a schematic illustration), and our derivations rely heavily on this observation. Additionally, we emphasize that $\sigma(s,n|s_0)$ should not be confused with other extreme value statistics observables, *eg.* the distribution of the maximum and the time at which this maximum is reached [Randon-Furling & Majumdar 2007, Mori *et al.* 2021].

3.1.1 Leftward exit-time probability

To determine $\sigma(s,n|s_0)$, we introduce the *leftward exit-time probability* (LETP) $F_{\underline{0},s}(n|s_0)$, defined as the probability that the random walker reaches 0 before s, at time n exactly. Partitioning trajectories that contribute to that event on the value of the maximum reached before the FPT to 0, we rewrite $F_{\underline{0},s}$ (shorthand for $F_{\underline{0},s}(n|s_0)$) as:

$$F_{\underline{0},s}(n|s_0) = \sum_{s'=s_0}^{s-1} \sigma(s',n|s_0), \tag{3.3}$$

[1] We here recall the definition: let s_1 and s_2 be visited sites, for any site s such that $s \in [\![s_1, s_2]\!]$, s is visited.

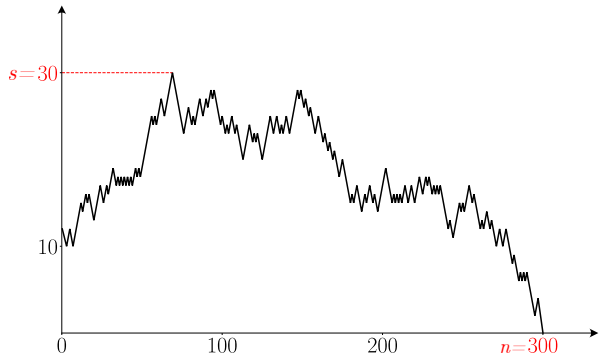


Figure 3.1: Example trajectory contributing to $\sigma(30, 300|12)$ for a symmetric $1D$ random walk. Starting from $s_0 = 12$, the walker reaches 0 for the first time at step $n = 300$. The running maximum up to n is $s = 30$.

and express $\sigma(s, n|s_0)$ in terms of the LETP only:

$$\sigma(s, n|s_0) = F_{\underline{0},s+1}(n|s_0) - F_{\underline{0},s}(n|s_0) \equiv D_s F_{\underline{0},s}(n|s_0). \tag{3.4}$$

- Equation (3.4) has clear implications: the determination of $F_{\underline{0},s}$ is sufficient to obtain fully explicit forms for the joint distribution, and characterize completely the coupling between the FPT and the explored territory.

- Note that $F_{\underline{0},s}$ is not a probability distribution since $\sum_{n=0}^{\infty} F_{\underline{0},s}(n|s_0) = \pi_{\underline{0},s}(s_0)$. In fact, equation (3.4) is a refined version of equation (1.8), which can be recovered by summing over n.

By partitioning over the first step of the walk, we show that the LETP obeys the same backward equation as the propagator, for all $0 < s_0 < s$:

$$F_{\underline{0},s}(n+1|s_0) = \mathcal{L}_{s_0} F_{\underline{0},s}(n|s_0), \tag{3.5}$$

with boundary conditions $F_{\underline{0},s}(n|s) = 0$ and $F_{\underline{0},s}(n|0) = \delta_{n,0}$, for all $n \geq 0$. Introducing the generating function $\tilde{F}_{\underline{0},s}(\xi|s_0) = \sum_{n=0}^{\infty} \xi^n F_{\underline{0},s}(n|s_0)$, we obtain:

$$\begin{aligned} &\tilde{F}_{\underline{0},s}(\xi|s_0) = \xi \mathcal{L}_{s_0} \tilde{F}_{\underline{0},s}(\xi|s_0) \quad \text{with} \quad 0 < \xi < 1 \\ &\tilde{F}_{\underline{0},s}(\xi|s) = 0 \\ &\tilde{F}_{\underline{0},s}(\xi|0) = 1. \end{aligned} \tag{3.6}$$

Equations (3.4) together with (3.6) draw a clear road-map towards the determination of *space and time* correlations for one-dimensional Markovian walks with connected span, provided that the system (3.6) can be solved. In the following subsection, we compute $\sigma(s, n|s_0)$ for classical examples of random walks already encountered in Chapter 1, to which we refer the reader for definitions of specific transport rules.

3.1.2 Examples

Symmetric and Biased random walk. We begin with the nearest neighbor symmetric random walk, for which the backward equation for $\tilde{F}_{\underline{0},s}$ reads:

$$\tilde{F}_{\underline{0},s}(\xi|s_0+1) - \frac{2}{\xi}\tilde{F}_{\underline{0},s}(\xi|s_0) + \tilde{F}_{\underline{0},s}(\xi|s_0-1) = 0. \tag{3.7}$$

Enforcing the boundary conditions and denoting $r_\pm = \frac{1}{\xi}(1 \pm \sqrt{1-\xi^2})$ yields $\tilde{F}_{\underline{0},s}(\xi|s_0) = (r_+^{s-s_0} - r_-^{s-s_0})(r_+^s - r_-^s)^{-1}$. Plugging into equation (3.4), we explicitly compute the generating function of the joint distribution:

$$\begin{aligned}\tilde{\sigma}(s,\xi|s_0) &= \tilde{F}_{\underline{0},s+1}(\xi|s_0) - \tilde{F}_{\underline{0},s}(\xi|s_0) \\ &= \frac{r_+ - r_-}{r_+^s - r_-^s}\frac{r_+^{s_0} - r_-^{s_0}}{r_+^{s+1} - r_-^{s+1}}(r_+ r_-)^{s-s_0} \\ &= \frac{r_+ - r_-}{r_+^s - r_-^s}\frac{r_+^{s_0} - r_-^{s_0}}{r_+^{s+1} - r_-^{s+1}}.\end{aligned} \tag{3.8}$$

In turn, series expansion in powers of ξ yields $\sigma(s,n|s_0)$ for all values of n.

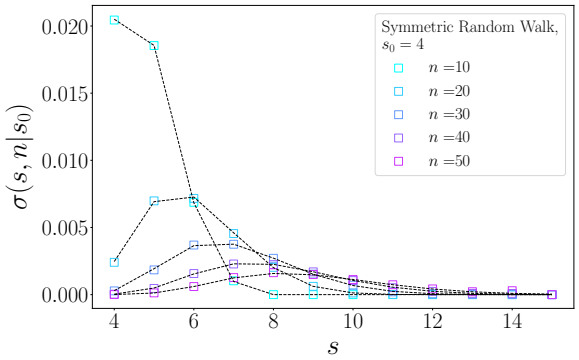
The biased nearest neighbor random walk is dealt with similarly. Denoting p the probability to step rightwards, the backward equation is given by:

$$p\tilde{F}_{\underline{0},s}(\xi|s_0+1) - \frac{1}{\xi}\tilde{F}_{\underline{0},s}(\xi|s_0) + q\tilde{F}_{\underline{0},s}(\xi|s_0-1) = 0, \tag{3.9}$$

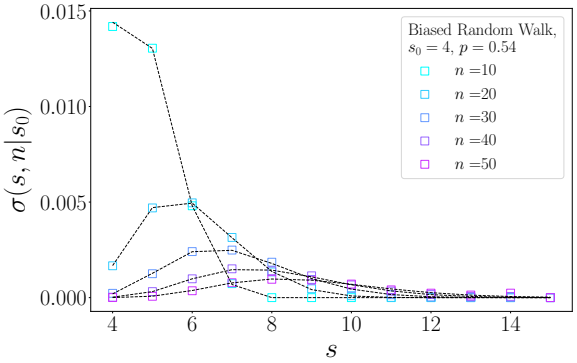
with $q = 1-p$. Denoting now $r_\pm = \frac{1}{2p\xi}(1 \pm \sqrt{1-4pq\xi^2})$, the LETP reads $\tilde{F}_{\underline{0},s}(\xi|s_0) = r_-^{s_0} - r_-^s(r_+^{s_0} - r_-^{s_0})(r_+^s - r_-^s)^{-1}$, and the generating function of the joint distribution straightforwardly follows.

Resetting and Persistent random walks. The resetting and persistent random walks have both been studied in Chapter 1. Consequently, we relegate the full derivation of $\tilde{F}_{\underline{0},s}$ to Appendix D, as the technical tools involved are redundant with earlier results. Making use of equation (3.4), we then obtain explicit expressions for $\sigma(s,n|s_0)$ and provide numerical evidence of the validity of our approach across these 4 examples in figure 3.2.

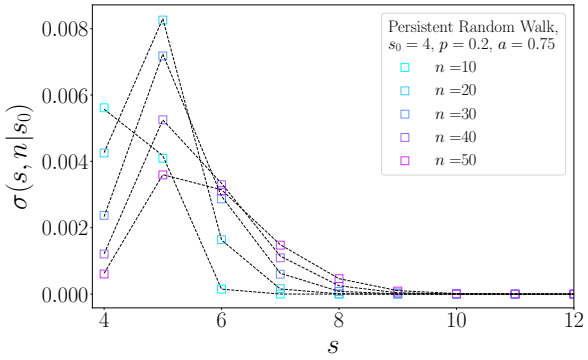
- It is clear from figure 3.2 that with high probability, a longer FPT leads to a larger explored territory. In other words, giving more time to the random walk allows it to explore more territory.
- We emphasize that the key information about $\sigma(s,n|s_0)$ is contained in the leftward exit-time probabilities $F_{\underline{0},s}$. Their determination is thus necessary and sufficient to fully characterize the interplay between the explored territory and the time needed to do so.
- We finally stress that while the technical tools needed to derive $\tilde{F}_{\underline{0},s}$ appear fairly standard, to the best of our knowledge the LETPs for lattice systems have not yet been investigated.



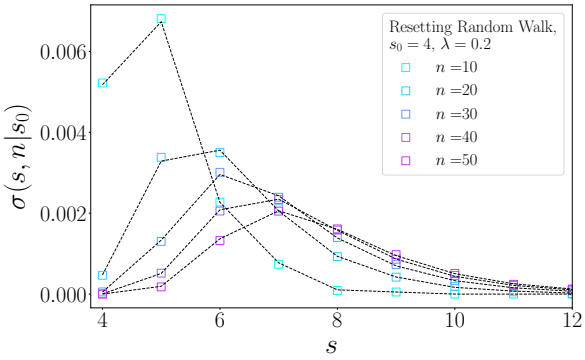
(a) Joint distribution $\sigma(s, n|s_0)$ for a symmetric nearest neighbor random walk. As n increases, the walker can go farther, and the most probable maximum drifts to infinity.



(b) Joint distribution $\sigma(s, n|s_0)$ for a biased nearest neighbor random walk. Note the shape similarly with the unbiased case. In fact, in the continuous setting, for fixed time t, the conditioned maximum distribution can be shown to be proportional to the unbiased case.



(c) Joint distribution $\sigma(s, n|s_0)$ for a persistent nearest neighbor random walk. The dashed lines are obtained from the small ξ expansion of $\sigma(s, \xi|s_0)$ computed in Appendix D.



(d) Joint distribution $\sigma(s, n|s_0)$ for a resetting nearest neighbor random walk. The dashed lines are obtained from the small ξ expansion of $\sigma(s, \xi|s_0)$ computed in Appendix D.

Figure 3.2

3.1.3 Extensions

We now show that the methodology for the derivation of $\sigma(s, n|s_0)$ in the case of Markovian one-dimensional lattice walks with connected span naturally extends to both bounded lattices as well as continuous space and time Markovian processes.

Bounded domains. Consider first the case of a $1D$ bounded lattice, with an absorbing site at 0, and a reflecting wall at site N. For all values of s such that $s < N + 1$, the trajectories contributing to $\sigma(s, n|s_0)$ are strictly identical to those in the unbounded medium. As a result, the bounded and unbounded joint distributions are equal. We thus focus on the specific case $s = N$. Defining $F_{\underline{0}}^{bounded}(n|s_0)$ the FPT probability to 0 in the bounded geometry, and partitioning over whether or not the site $N - 1$ is reached yields:

$$\sigma^{bounded}(N, n|s_0) + F_{\underline{0},N}(n|s_0) = F_{\underline{0}}^{bounded}(n|s_0). \tag{3.10}$$

In turn, the joint distribution is known for $s = N$, provided $F_{\underline{0}}^{bounded}(n|s_0)$ can be computed (usually by solving a backward equation).

Continuous space and time. We now turn to stochastic processes defined in continuous space and time, and adapt equation (3.4) to meet the continuous formalism:

$$\sigma(x,t|x_0) = \frac{\partial}{\partial x} F_{\underline{0},x}(t|x_0), \tag{3.11}$$

where $\sigma(x,t|x_0)$ and $F_{\underline{0},x}(t|x_0)$ are the continuous counterparts of $\sigma(s,n|s_0)$ and $F_{\underline{0},s}(n|s_0)$. As an illustration, we briefly discuss the Brownian case. Introducing the Laplace transform [2] of the LETP $\tilde{F}_{\underline{0},x}(p|x_0) = \int_0^\infty e^{-pt} F_{\underline{0},x}(t|x_0)\mathrm{d}t$, the classical diffusion equation reads:

$$p\tilde{F}_{\underline{0},x}(p|x_0) = D\frac{\partial^2}{\partial^2 x_0}\tilde{F}_{\underline{0},x}(p|x_0) \tag{3.12}$$

with D the diffusion coefficient. Solving for $\tilde{F}_{\underline{0},x}$ with $\tilde{F}_{\underline{0},x}(p|x) = 0$ and $\tilde{F}_{\underline{0},x}(p|0) = 1$, we successively obtain:

$$\begin{aligned}
\tilde{F}_{\underline{0},x}(p|x_0) &= \frac{\sinh\left[\sqrt{\frac{p}{D}}(x-x_0)\right]}{\sinh\left[\sqrt{\frac{p}{D}}x\right]} \qquad \tilde{\sigma}(x,p|x_0) = \sqrt{\frac{p}{D}}\frac{\sinh\left[\sqrt{\frac{p}{D}}x_0\right]}{\sinh^2\left[\sqrt{\frac{p}{D}}x\right]} \\
\sigma(x,t|x_0) &= \frac{2D\pi}{x^3}\sum_{k=1}^{\infty} e^{-(k\pi)^2 D\tau} k\sin(k\pi\tilde{x}_0)\left[2(k\pi)^2 D\tau - 2 - \frac{k\pi\tilde{x}_0}{\tan(k\pi\tilde{x}_0)}\right],
\end{aligned} \tag{3.13}$$

with $\tilde{x}_0 = x_0/x$ and $\tau = t/x^2$. Importantly:

- Taking $p \to 0$ in the Laplace Transform is equivalent to considering the time integrated quantity. In turn, we recover from equation (3.13) the distribution of the maximum of the Brownian motion $\mu(x|x_0) = x_0/x^2$.
- The explicit determination (3.13) of $\sigma(x,t|x_0)$ in the real variables x and t is obtained from residue analysis; the resulting sum of decaying exponentials is the hallmark of confined systems.
- Note that for Brownian motion, an alternative determination of $\sigma(x,t|x_0)$ can be found in [Borodin & Salminen 1996].

Additionally, we provide explicit expressions of $\tilde{\sigma}(x,p|x_0)$ for Biased Brownian motion and resetting Brownian motion in Appendix D. We finally consider the case of bounded one-dimensional continuous processes which are treated similarly to bounded lattice walks. For a reflecting boundary located at position L, and defining the bounded distribution $F_{\underline{0}}^{bounded}(t|x_0)$ of the FPT to 0, the joint distribution in the specific case $x = L$ is given by:

[2] The Laplace transform is simply the continuous counterpart of the generating function transform in the discrete time case

$$\sigma^{bounded}(L,t|x_0) = F_{\underline{0}}^{bounded}(t|x_0) - F_{\underline{0},L}(t|x_0). \tag{3.14}$$

In turn, for Markovian processes, the bounded FPT distribution is derived from the corresponding backward equation. As a concluding example, the joint distribution for a bounded Brownian particle is given by:

$$\sigma^{bounded}(L,t|x_0) = \frac{D\pi}{L^2}\sum_{k=1}^{\infty} e^{-(k\pi)^2 D\tau}(-1)^{k+1} k \sin(k\pi\tilde{x}_0), \tag{3.15}$$

with $\tau = t/4L^2$ and $\tilde{x}_0 = x_0/2L$.

3.2 Joint distribution in the large space and time limit

By exploiting the specifics of one-dimensional stochastic processes with connected span, we derived in section 3.1 a general methodology to evaluate exactly the joint distribution of the explored territory and the FPT to 0. In turn, we completely determined the intuitive correlations between these two random observables. We focus now on more general one-dimensional scale-invariant stochastic processes for which the previous methodology might not apply,as is the case for non-Markovian processes. To that end, extending an approach presented in [Levernier *et al.* 2018], we identify emergent universal scaling forms of the joint distribution $\sigma(x,t|x_0)$, in the large x and t limit. Without loss of generality, we now exclusively use the continuous formalism.

3.2.1 Universal scaling form

As a starting point, consider an unbounded stochastic process $x(t)$ whose FPT distribution $F_{\underline{0}}(t|x_0)$ to 0 has an algebraic large t decay, and define the persistence exponent θ accordingly [Bray *et al.* 2013]:

$$F_{\underline{0}}(t|x_0) \underset{t\to\infty}{\sim} \frac{k(x_0)}{t^{1+\theta}}, \tag{3.16}$$

where $k(x_0)$ is a process dependent function. As an illustration, the FPT distribution for a Brownian particle is given by [Redner 2001] $F_{\underline{0}}(t|x_0) = x_0(2\pi t^3)^{-1/2}e^{-x_0^2/(2t)}$ such that $F_{\underline{0}}(t|x_0) \propto x_0 t^{-3/2}$ at large t, and $\theta = 1/2$.

We now consider a secondary target at fixed position x, and introduce $T_{typ} \propto x^{d_w}$ the typical time needed for the process to reach x, where the walk dimension d_w has been defined in chapter 2. To evaluate the LETP, we distinguish two cases:

- At short times $t \ll T_{typ}$, trajectories contributing to $F_{\underline{0},x}(t|x_0)$ never come close to x, and are thus identical to unbounded trajectories. In turn, $F_{\underline{0},x}(t|x_0) \underset{t\ll T_{typ}}{\sim} F_{\underline{0}}(t|x_0)$.

- For times $t \gg T_{typ}$, trajectories that reach x do not contribute to $F_{\underline{0},x}(t|x_0)$, and we rewrite the leftward exit-time probability as

$$F_{\underline{0},x}(t|x_0) \sim F_{\underline{0}}(t|x_0) g_1\left(\frac{t}{x^{d_w}}\right), \tag{3.17}$$

where $g_1(u)$ is a smooth cut off function that vanishes as $u \to \infty$, and $g_1(0) = 1$ to recover the short times regime. Importantly, g_1 must be an dimensionless function, and cannot depend on x_0 since the starting point dependence is captured in $F_{\underline{0}}(t|x_0)$. As a result, g_1 can only be a function of the rescaled variable t/x^{d_w}.

As a result, defining the rescaled variable $\tau = t/x^{d_w}$ and combining equations (3.11), (3.16) and (3.17), we obtain the following scaling form for the joint distribution:

$$\sigma(x,t|x_0) \underset{\substack{x\to\infty \\ t\to\infty \\ \tau \text{ fixed}}}{\sim} \frac{h(x_0)}{x^{d_w\theta+1}} \frac{1}{x^{d_w}} f(\tau), \tag{3.18}$$

where $h(x_0)$ and $f(\tau)$ are both process-dependent functions, and f has been normalized. Equation (3.18) is the main result of this section, and deserves a few comments:

- While both h and f are process-dependent functions, equation (3.18) has important practical applications. In particular, for fixed x_0, empirical measurement of the joint distribution for a given x provides the functional form of σ for all (x,t) pairs.

- Since $f(\tau)$ is normalized, integrating equation (3.18) over t yields:

$$\int_0^\infty \sigma(x,t|x_0)\mathrm{d}t = \mu(x|x_0) = \frac{h(x_0)}{x^{d_w\theta+1}}, \tag{3.19}$$

  with $\mu(x|x_0)$ the marginal distribution of the maximum of the process before reaching zero. In turn, we recover the algebraic decay of $\mu(x|x_0)$ as a function of d_w and θ only, in agreement with known results [Majumdar *et al.* 2010b]. Additionally, this identification yields the small x_0 behavior of $h(x_0)$:

$$h(x_0) \underset{x_0\to 0}{\propto} x_0^{d_w\theta}. \tag{3.20}$$

- The previous observation confers a clear physical interpretation to the f function: $f(\tau)$ is the distribution of the rescaled variable $\tau = t/x^{d_w}$, conditioned on the value of the maximum x. Importantly, we stress that f is independent of x_0.

3.2.2 Illustrations

We now illustrate the result (3.18) on representative examples of one-dimensional stochastic processes. Importantly, we stress that the collapse of the conditional distribution of $\tau = t/x^{d_w}$, for various values of x and x_0 is enough to guarantee the validity of equation (3.18).

Brownian motion. Taking the $x \to \infty$, $t \to \infty$ limit with τ fixed in the explicit joint distribution (3.13), we obtain:

$$\sigma(x,t|x_0) \sim \frac{x_0}{x^2}\frac{f_{BM}(\tau)}{x^2}$$
$$f_{BM}(\tau) = 2D\pi^2 \sum_{k=1}^{\infty} e^{-(k\pi)^2 D\tau} k^2 \left[2(k\pi)^2 D\tau - 3\right], \tag{3.21}$$

in agreement with equation (3.18), and display numerical evidence in figure 3.3(a). Importantly, we emphasize that the asymptotic conditional distribution f_{BM} holds for any symmetric Markovian random walk whose $i.i.d$ increments satisfy the central limit theorem[3]. We now turn to processes for which the $f(\tau)$ function cannot be computed exactly.

Riemann walks. Already introduced in Chapter 2, Riemann Walks offer a continuously varying walk dimension $d_w = \mu$, where μ characterizes the algebraic decay of the transition kernel $p(s_1 \to s_2) \propto |s_1 - s_2|^{-(1+\mu)}$. We consider the *crossing* convention, such that all flown-over sites are considered visited, and the process is terminated upon first crossing of 0. In that case, the persistence exponent is known and given by $\theta = 1/2$ [Hughes 1995]. Of note, the persistence exponent for Riemann walks is predicted by the more general Sparre Andersen theorem [Andersen 1954, Bray *et al.* 2013]. We display numerical agreement with our theory in figure 3.3(b), for $\mu = 3/2$.

Fractional Brownian Motion. Both Brownian motion and Riemann walks are Markovian stochastic processes. As a first representative example of non-Markovian random process, we consider the Fractional Brownian Motion (FBM). The FBM is a Gaussian process defined by its auto-correlation function:

$$\langle (x(t_1) - x_0)(x(t_2) - x_0)\rangle = 2\left[t_1^{2H} + t_2^{2H} + |t_1 - t_2|^{2H}\right] \tag{3.22}$$

where $0 < H < 1$ is the *Hurst* exponent. Importantly, for $H \neq 1/2$, the FBM is non-Markovian and displays anomalous behavior. Its walk dimension $d_w = 1/H$ is directly read off from (3.22), and its persistence exponent has been shown to satisfy $\theta = 1-H$ [Molchan 1999, Bray *et al.* 2013]. We provide simulation details in Appendix D and display numerical agreement with our theory in figure 3.3(c).

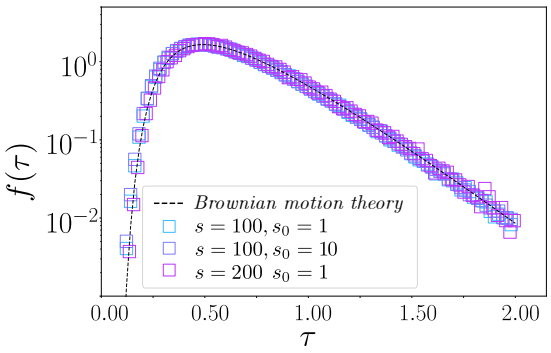
Random Acceleration Process. Lastly, we consider the Random Acceleration process (RAP), whose time evolution is given by the following stochastic differential equation:

$$\ddot{x}(t) = \eta(t), \tag{3.23}$$

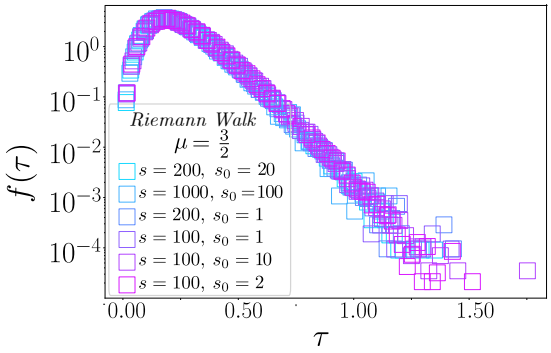
where η(t) is a standard white noise with zero mean and $\langle \eta(t_1)\eta(t_2)\rangle = 2D\delta(t_1 - t_2)$. Importantly, the RAP is a non-Markovian process and the pair (d_w, θ) is equal to $(2/3, 1/4)$. Note that the RAP has been used to model the dynamics of friction-less particles whose acceleration is continuously given random kicks, in contrast to the classical over-damped Langevin picture

[3] In that case the value of D must be extracted from the MSD of the process, which is asymptotically given by $\langle x^2(t)\rangle \underset{t\to\infty}{\sim} 2Dt$.

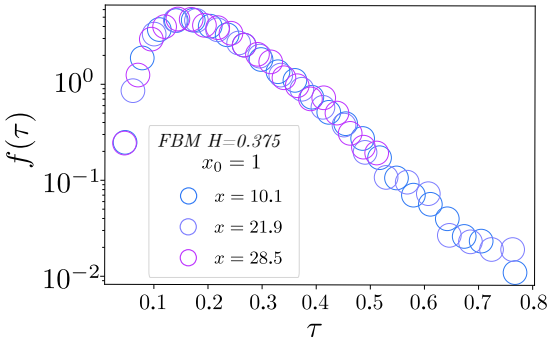
$\dot{x}(t) = \eta(t)$. We refer the reader to [Bicout & Burkhardt 2000, Burkhardt 2007] for an overview of RAP results, provide simulation details in Appendix D and display numerical agreement with our theory in figure 3.3(d).



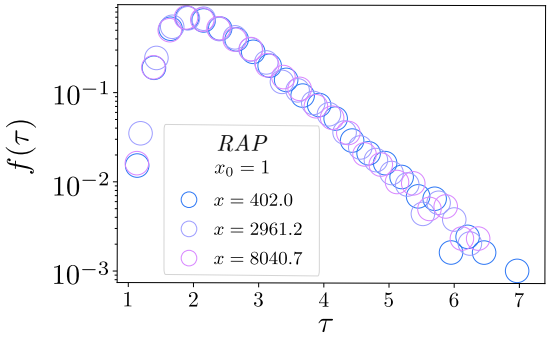
(a) Symmetric normal walk: conditional distribution of the rescaled variable $\tau = n/s^2$ for various maximums s and initial positions s_0. For large s, the distribution converges to the universal Brownian motion result $f_{BM}(\tau)$ (3.21).



(b) Riemann walk: conditional distribution of the rescaled variable $\tau = n/s^\mu$ for various maximums s and initial positions s_0. All distributions collapse to an unknown universal function.



(c) Fractional Brownian Motion: conditional distribution of the rescaled variable τ for various maximums x. All distributions collapse to an unknown universal function.



(d) Random Acceleration Process: conditional distribution of the rescaled variable τ for various maximums x. All distributions collapse to an unknown universal function.

Figure 3.3

Supported by the numerical evidence accumulated in figure 3.3, we have shown that equation (3.18) holds for general scale-invariant one-dimensional processes, and is even valid beyond the Markovian hypothesis. Consequently, we emphasize that the coupling between the space explored before reaching a target and the time needed to do so is fully governed by the scaling variable $\tau = t/x^{d_w}$.

3.3 Applications

In this section, we investigate the implications of this emergent scaling behavior on two specific applications. First, we focus on the derivation of conditional space and time distributions, and their application to determining conditional survival probabilities. Second, we show that the joint distribution appears as key tool for the evaluation of the persistence properties of a paradigmatic example of self-interacting random walk: the Self Avoiding True Walk (SATW).

3.3.1 Conditional distributions and the Rosenstock trapping problem

As an important side product, the determination of the joint distribution $\sigma(x,t|x_0)$ allows for the evaluation of two conditional distributions:

$$G_{sp}(x|t,x_0) = \frac{\sigma(x,t|x_0)}{F_{\underline{0}}(t|x_0)} \qquad G_{tm}(t|x,x_0) = \frac{\sigma(x,t|x_0)}{\mu(x|x_0)} \tag{3.24}$$

with G_{sp} the distribution of the explored territory before reaching a target conditioned on the value of the FPT, and G_{tm} the distribution of the FPT conditioned on the value of the explored territory [4]. In turn, G_{sp} provides an explicit solution to a conditional version of the Rosenstock trapping problem [Rosenstock 1961, Hughes 1995], which initially motivated the study of the explored territory $C(s_0)$ in chapter 1.

For the sake of clarity, we rephrase the Rosenstock problem to fit with the continuous space and time notations, and provide a schematic of the situation in figure 3.4(a). Consider a catalytic particle that enters a one-dimensional chemical reactor at x_0 and leaves it at 0, and assume that the time t spent in the reactor is observed, which is equivalent to conditioning with respect to the FPT. The reactor contains Poisson distributed point-like reactive targets of density ρ [5], which trigger a reaction upon encounter with the catalytic particle. Furthermore, we assume that the dynamics of the particle remain unchanged upon reaction. The efficiency of such schematic catalytic reaction can be quantified by the Rosenstock reaction probability P_t that the catalytic particle has activated at least one reactive site before exiting the domain, given the exit-time t, and reads

$$P_t = \int_0^\infty (1 - e^{-\rho x}) G_{sp}(x|t,x_0)\mathrm{d}x. \tag{3.25}$$

While an explicit expression of $\sigma(x,t|x_0)$ is required to derive G_{sp} and P_t exactly, the emergent scaling structure for large time t is sufficient to obtain quantitative results.

Making use of (3.18), the asymptotic conditional distribution of the FPT is given by $G_{tm}(t|x,x_0) \sim x^{-d_w} f(\tau)$ with f a process-dependent normalized function. Similarly, the conditional distribution of the explored territory can be recast in a scaling form of the rescaled variable $x/t^{1/d_w}$, and is asymptotically given by:

[4] $F_{\underline{0}}(.|x_0)$ and $\mu(.|x_0)$ are the marginal distributions of respectively the FPT and the maximum.
[5] In probabilistic terms, the probability of observing a target in some interval of size dx is given by ρdx.

$$G_{sp}(x|t,s_0) \sim \frac{1}{t^{\frac{1}{d_w}}}\phi\left(\frac{x}{t^{\frac{1}{d_w}}}\right)$$
$$\phi(\chi) = \frac{\chi^{-d_w(\theta+1)-1} f(\chi^{-d_w})}{\int_0^\infty u^{-d_w(\theta+1)-1} f(u^{-d_w})du}. \tag{3.26}$$

Numerical evidence of the scaling behavior (3.26) is provided in figure D.1 of Appendix D. Plugging equation (3.26) into (3.25), we show that P_t is a function of the reduced variable $\rho t^{1/d_w}$ only:

$$P_t \underset{\substack{t\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} \int_0^\infty (1 - e^{-\rho t^{1/d_w} u})\phi(u)\mathrm{d}u. \tag{3.27}$$

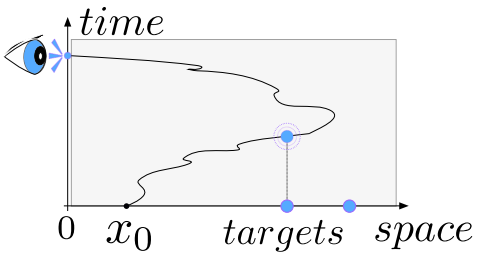
Furthermore, in the limit $\rho t^{1/d_w} \ll 1$, the conditional reaction probability displays an algebraic growth

$$P_t \propto \rho t^{1/d_w}. \tag{3.28}$$

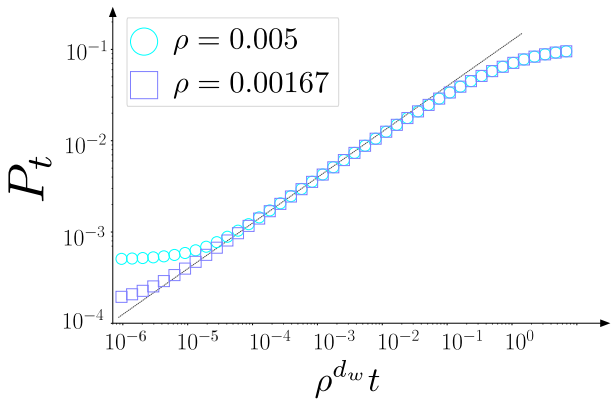
Importantly, equation (3.28) quantitatively describes the intuitive increase of P_t with t. In the specific case of Brownian motion, $f_{BM}(\tau)$ is given in equation (3.21) such that ϕ can be exactly computed. In turn, the reaction probability in the regime $x_0^2/D \ll t \ll 1/(D\rho^2)$ is given by:

$$P_t \sim \sqrt{\pi}\rho(Dt)^{\frac{1}{2}}, \tag{3.29}$$

and we conclude with numerical illustration in figure 3.4(b).



(a) Starting from x_0, the particle evolves in the hidden grayed-out domain containing randomly distributed traps (blue dots), and is observed at some time t (eye) upon reaching 0 for the first time. The conditional reaction probability quantifies the likelihood of trap encounter given the extra FPT to 0 information.



(b) Rosenstock reaction probability for Brownian motion with varying trap density ρ. For large t, P_t is a function of $\rho^{d_w} t$ only. The dashed line corresponds to the $\rho^{d_w} t \ll 1$ limit (3.29). Note that for t too small, the scaling regime is not reached and the collapse fails.

Figure 3.4

3.3.2 Persistence properties of the Self Avoiding True Walk

Transport rules. The one-dimensional Self Avoiding True Walk (SATW) [Sapozhnikov 1994] is a lattice random walk pertaining to the family of the self-interacting random walks and reinforcement walks, whose transport rules at step n depend on the full set of visited sites up to step n. More precisely, at each time step, if both neighboring sites have already been visited, the random walker hops on either of them with probability 1/2. However, if one of them is virgin, *ie* has never been visited, it is chosen with probability $0 \leq \beta \leq 1$. Note that this can either be an attractive effect ($\beta < 1/2$) or a repulsive one ($\beta > 1/2$) (see figure 3.5). Historically, the SATW has been studied as a prototypical example of process with long-range memory and proved more recently to be relevant for the description of the dynamics of motile cells [d'Alessandro *et al.* 2021].

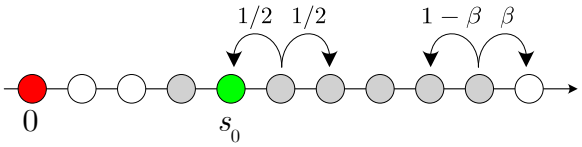


Figure 3.5: One-dimensional SATW dynamics. All grayed-out sites have been previously visited, and all white sites are virgin sites. Upon hopping, if both neighboring sites have already been visited, the jump is performed symmetrically. If one of the sites is virgin, it is chosen with probability β.

FPT distribution. We focus here on the specific case of the SATW starting from $s_0 = 1$, with absorbing site at 0, and aim to explicitly derive the asymptotic large n FPT distribution $F_{\underline{0}}(n|s_0 = 1)$. While the persistence exponent has recently been derived in [Barbier-Chebbah *et al.* 2020], and is given by $\theta = \frac{1-\beta}{2\beta}$, the exact prefactor of $F_{\underline{0}}(n|s_0 = 1)$ has yet to be computed. We fill this gap by considering $F_{\underline{0}}(n|s_0 = 1)$ as a marginal of the joint distribution $\sigma(s, n|1)$:

$$F_{\underline{0}}(n|1) = \sum_{s=1}^{\infty} \sigma(s, n|1) = \sum_{s=1}^{\infty} \mu(s|1) G_{tm}(n|s, 1). \tag{3.30}$$

with $\mu(s|1)$ the distribution of the maximum before reaching 0, and $G_{tm}(n|s, 1)$ the conditional distribution of the FPT to 0 with s fixed.

1. First, $\mu(s|1)$ is entirely determined by the splitting probability $\pi_{0,\underline{s}}(1)$, and both quantities are asymptotically found to be given by [Barbier-Chebbah *et al.* 2020]:

$$\begin{aligned} \pi_{0,\underline{s}}(1) &\underset{s\to\infty}{\sim} \frac{\Gamma(-2+2/\beta)}{\Gamma(-1+1/\beta)} s^{-\frac{1-\beta}{\beta}} \\ \mu(s|1) &\underset{s\to\infty}{\sim} \frac{\Gamma(-2+2/\beta)}{\Gamma(-1+1/\beta)} \frac{1-\beta}{\beta} s^{-\frac{1-\beta}{\beta}-1} \end{aligned} \tag{3.31}$$

2. Next, by partitioning over times at which virgin sites are discovered, we show that the generating function of the joint distribution $\tilde{\sigma}(s, \xi|1)$ can be expressed in terms of the LETP and RETP only:

$$\tilde{\sigma}(s,\xi|1) = \frac{\xi}{2}\left(\prod_{s'=3}^{s}\tilde{F}_{0,\underline{s'}}(\xi|s'-1)\right)\tilde{F}_{\underline{0},s+1}(\xi|s). \tag{3.32}$$

As for the splitting probability, $\tilde{F}_{0,\underline{s}}(\xi|s-1)$ and $\tilde{F}_{\underline{0},s+1}(\xi|s)$ can be exactly computed, yielding an explicit expression for the joint distribution. In turn, in the large n and s limit, we rewrite the joint distribution in the scaling form (3.18) with $d_w = 2$ [Barbier-Chebbah *et al.* 2020] and obtain:

$$\sigma(s,n|1) \underset{\substack{s\to\infty\\ n\to\infty\\ n/s^2\text{fixed}}}{\sim} \mu(s|1) f_{SATW}(n/s^2)s^{-2}, \tag{3.33}$$

where the conditional distribution $f_{SATW}(\tau)$ of $\tau = n/s^2$ is defined by its strikingly simple Laplace transform:

$$\tilde{f}_{SATW}(p) = \int_0^\infty e^{-pu} f_{SATW}(u)\mathrm{d}u = \left(\frac{\sqrt{2p}}{\sinh(\sqrt{2p})}\right)^{\frac{1}{\beta}}. \tag{3.34}$$

Note that the detailed calculations leading to equations (3.31), (3.32) and (3.34) are reproduced in Appendix E.

3. Finally, plugging equation (3.33) into equation (3.30) yields the asymptotic FPT distribution:

$$\begin{aligned} F_{\underline{0}}(n|s_0=1) &\underset{n\to\infty}{\sim} \frac{\Gamma(-2+2/\beta)}{\Gamma(-1+1/\beta)}\frac{(1-\beta)}{2\beta}A(\beta)n^{-\frac{1-\beta}{2\beta}-1} \\ A(\beta) &= \int_0^\infty f_{SATW}(\tau)\tau^{\frac{1-\beta}{2\beta}}d\tau. \end{aligned} \tag{3.35}$$

Exact inversion of $\tilde{f}_{SATW}$. For $\beta = 1/n$ with $n \in \mathbb{N}^*$, the poles of $\tilde{f}_{SATW}(z)$ are given by $z_k = -k^2\pi^2/2$ and $\tilde{f}_{SATW}$ can be analytically inverted. For concreteness, we provide two exact formulas:

$$\begin{aligned} \beta = \frac{1}{3} : f_{SATW}(\tau) &= \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2\tau}{2}}\left[\frac{1}{2}(-1)^{k+1}k^6\pi^6\tau^2\right. \\ &\qquad \left. -\frac{9}{2}(-1)^{k+1}k^4\pi^4\tau + \frac{1}{2}(-1)^{k+1}k^4\pi^4 + 6(-1)^{k+1}k^2\pi^2\right] \\ \beta = \frac{1}{4} : f_{SATW}(\tau) &= \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2\tau}{2}}\left[\frac{k^8\pi^8\tau^3}{6} - 3k^6\pi^6\tau^2\right. \\ &\qquad \left. +\frac{25k^4\pi^4\tau}{2} + \frac{2k^6\pi^6\tau}{3} - \frac{10k^4\pi^4}{3} - 10k^2\pi^2\right], \end{aligned} \tag{3.36}$$

and display numerical agreement in figure 3.6(a). Note that, as β decreases, the polynomial part in equation (3.36) accumulates more terms. However, careful residue analysis yields the leading large τ behavior, valid for any $\beta = 1/n$:

$$f_{SATW}(\tau) \underset{\tau\to\infty}{\sim} e^{-\frac{\pi^2}{2}\tau} \frac{\pi^{\frac{2}{\beta}}}{(\frac{1}{\beta}-1)!} \tau^{\frac{1}{\beta}-1}. \tag{3.37}$$

FPT prefactor - derivation of $A(\beta)$. We now show that $A(\beta)$ can be computed exactly. Consider first the case where $n = 1/\beta$ is an *odd* integer, and rewrite $A(\beta) = \int_0^\infty f_{SATW}(\tau)\tau^m \mathrm{d}\tau$ with $m = (n-1)/2$. For a given function $g(\tau)$ whose Laplace transform $\mathcal{L}[g]$ is defined as $\mathcal{L}[g](s) = \int_0^\infty e^{-\tau s} g(\tau)\mathrm{d}\tau$, one has

$$\int_0^\infty g(\tau)\tau^m \mathrm{d}\tau = (-1)^m \frac{\mathrm{d}^m}{\mathrm{d}^m s} \mathcal{L}[g](s)\Big|_{s=0}, \tag{3.38}$$

and choosing $g(\tau) = \mathcal{L}^{-1}[\tilde{f}_{SATW}](\tau)$ yields

$$\begin{aligned} A(\beta) &= (-1)^m \frac{\mathrm{d}^m}{\mathrm{d}^m s} \mathcal{L}\left[\mathcal{L}^{-1}[\tilde{f}_{SATW}]\right](s)\Big|_{s=0} \\ &= (-1)^m \frac{\mathrm{d}^m}{\mathrm{d}^m s} \left(\frac{\sqrt{2s}}{\sinh(\sqrt{2s})}\right)^n\Big|_{s=0} \end{aligned}. \tag{3.39}$$

We introduce the *higher order Bernoulli polynomials* $B_n^l(x)$ [Adelberg 1992] defined by their generating function:

$$\left[\frac{t}{e^t-1}\right]^l e^{xt} = \sum_{n=0}^\infty B_n^l(x)\frac{t^n}{n!}, \tag{3.40}$$

and combine equations (3.39) and (3.40) to obtain:

$$A(\beta) = B_{2m}^{2m+1}\left(\frac{2m+1}{2}\right)\frac{2^{3n}m!}{2m!}. \tag{3.41}$$

Importantly, with the help of the generating function (3.40), the $B_n^l(x)$ can be shown to satisfy a number of recurrence relations. Using in particular the fact that

$$\begin{aligned} B_n^l(x+1) &= B_n^l(x) + nB_{n-1}^{l-1}(x), \\ B_n^{l+1}(x) &= \left(1-\frac{n}{l}\right)B_n^l(x) + \frac{n}{l}(x-l)B_{n-1}^l(x), \end{aligned} \tag{3.42}$$

we finally obtain:

$$A(\beta) = (2m-1)(2m-3)(2m-5)... \equiv (2m-1)!! \tag{3.43}$$

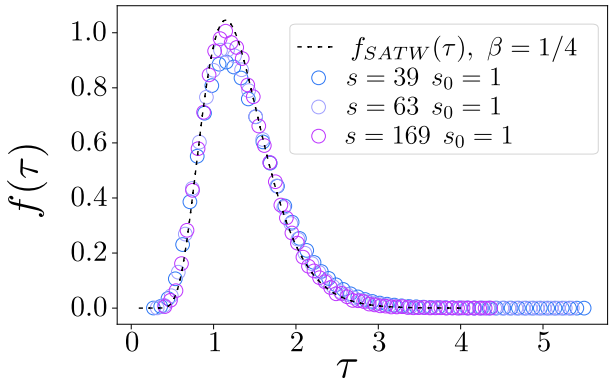
As a concluding step, we focus on general β values. While equation (3.43) has only been obtained in the case $\beta = 1/(2m+1)$, we analytically extend it to all β with the help of the Γ function:

$$A(\beta) = \frac{2^{\frac{1-\beta}{2\beta}}\Gamma\left(\frac{1}{2\beta}\right)}{\sqrt{\pi}}, \tag{3.44}$$

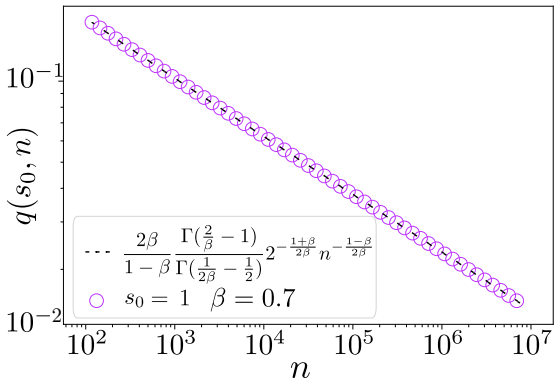
and provide an explicit expression for the asymptotic FPT distribution:

$$F_{\underline{0}}(n|s_0=1) \underset{n\to\infty}{\sim} \frac{\Gamma(\frac{2}{\beta}-1)}{\Gamma(\frac{1}{2\beta}-\frac{1}{2})} 2^{-\frac{1+\beta}{2\beta}} n^{-\frac{1-\beta}{2\beta}-1}. \tag{3.45}$$

Numerical agreement between equation (3.45) and simulations is provided in figure 3.6(b).



(a) Conditional distribution of the rescaled variable τ for various maxima s. All distributions collapse to the single universal curve (3.36). Note that there is an apparent convergence speed: the collapse improves as s increases.



(b) Survival probability $q(s_0, n) = \sum_{k=n+1}^{\infty} F_{\underline{0}}(k|s_0)$ of a SATW with $\beta = 0.7$. Both the persistence exponent θ and prefactor obtained from equation (3.45) are correctly recovered.

Figure 3.6

3.4 Conclusion

To investigate further the territory explored before reaching a target, we focus in this chapter on its coupling with the random stopping time at which it is observed, namely the FPT to the target. To quantify the correlations between these *space* and *time* observables, we introduce their joint distribution $\sigma(x,t|x_0)$, derive analytical results across general one-dimensional stochastic processes and address underlying physical implications.

We first consider one-dimensional Markovian random walks with connected span, and develop a systematic framework to explicitly derive the joint distribution, relying on the determination of the *leftward exit-time probability* (LETP). In turn, we compute a variety of novel exact expressions for $\sigma(x,t|x_0)$, spanning discrete and continuous processes, as well as unbounded and bounded geometries.

For general one-dimensional scale-invariant processes, we focus on the large space and time limit, and uncover emergent scaling forms for the LETP, leading to corresponding universal behavior of the joint distribution. In fact, we show that the coupling between *space* and *time* statistics is entirely governed by a single rescaled parameter $\tau = t/x^{d_w}$, and provide illustrations across a broad range of one-dimensional processes, including representative examples of Non-Markovian processes and self-interacting random walks.

Exploiting the emergent universal form of the joint distribution, we unveil similar behavior for the conditional distributions of either the explored territory knowing the FPT, or vice-versa, and illustrate their relevance on two specific applications. Focusing first on a conditional version of the Rosenstock trapping problem [Rosenstock 1961], we determine the likelihood of reacting with a trap given the FPT through 0 of a catalytic particle in a closed reactor. Second, we make use of the joint distribution approach to explicitly compute the prefactor of the FPT distribution for a paradigmatic example of self-interacting walk, the Self Avoiding True Walk.

We emphasize that the next natural step would be to consider the *space* and *time* coupling for more general scale-invariant processes, in higher dimensions for instance, which is left for future work. One can only guess, but the *compact* universality class could lend itself to a treatment similar to the simple one-dimensional geometry, while in the *non-compact* case, a one-to-one correspondence between the FPT and number of visited sites would seem very appealing.

Part II

Jump processes in confinement

Chapter 4

Splitting probabilities of jump processes

Contents

The following four chapters deal exclusively with a specific type of stochastic process, namely discrete-time and continuous space random walks, or *jump processes*. In turn, we take this introductory chapter as an opportunity to briefly review and contextualize the results presented in this manuscript.

In a first part, we provide a precise definition of one-dimensional symmetric jump processes, and review known results for unbounded jump processes (which we also refer to as *unconstrained*), and jump processes killed upon entering the negative half axis (*ie* whose movement is stopped upon becoming negative). We discuss in depth the survival probability of the latter and highlight the technical difficulties arising for *bounded* jump processes, *ie* killed upon exiting an interval.

In a second part, we present the main result of this chapter, namely the determination of the splitting probability $\pi_{0,\underline{x}}(x_0)$, defined as the probability for a jump process starting from x_0 to escape a confining interval $[0, x]$ through x rather than through 0, in the large x limit. We illustrate our result on a variety of examples, and exhibit experimental relevance in the context of particle scattering and so-called *transmission probabilities*. Of note, this chapter builds up on results presented in [Klinger *et al.* 2022b].

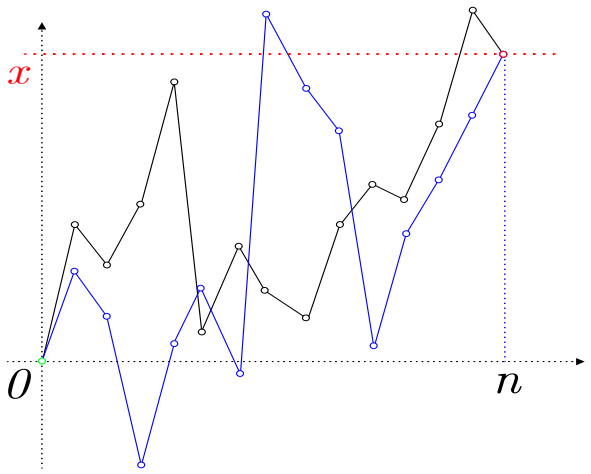


Figure 4.1: Example trajectories contributing to the infinite propagator $G_\infty(x,n|0)$ (blue) and semi-infinite propagator $G_0(x,n)$ (black). In the semi-infinite case, the whole trajectory stays above 0. In both cases the walker starts from 0 and is located at position x after n steps.

4.1 Propagators of general jump processes

Let us consider a particle starting from $x_0 \geq 0$, whose successive positions are given by $x_{n+1} = x_n + \xi_n$, where the ξ_i's are independent and identically distributed (*i.i.d.*), with common distribution $p(\ell)$. In the following, we focus on *symmetric continuous jump processes*, such that the single jump distribution is taken to be symmetric $p(\ell) = p(-\ell)$, and the cumulative distribution is continuous.

Of note, jump processes have been used to model a variety of physical phenomena, from polymer bead location [Rouse 1953] to financial asset values and queuing theory [Asmussen 2003]. Additionally, we emphasize that jump processes are relevant for the description of discrete time series, which naturally arise in empirical data series. Indeed, continuous models are not well suited to properly capture probabilistic effects arising from the discreetness of the data. As an illustration, consider the probability of a given time series $(x_0, x_1, ..x_n)$ with $x_0 = 0$ to be positive at all times. Since $x_0 = 0$, any continuous model fitting the data yields 0 as an answer. In contrast, a jump process fitting the data has a strictly non zero probability of being positive at all times, as we will see shortly.

4.1.1 Unbounded jump processes

We first focus on jump processes evolving freely on $\mathbb{R}$, and characterize their statistics at all times n. We point out that the analytical results presented here are classical, and can be found in [Van Kampen 1992, Hughes 1995] for instance.

Infinite propagator. The infinite propagator $G_\infty(x,n|x_0)$, defined as the probability for the particle to be located at position x after n steps (see figure 4.1), fully captures the statistics of the jump process after n steps. Importantly, since the jumps are *i.i.d.* and symmetric, the propagator depends only on $x - x_0$, and we denote $G_\infty(x,n) \equiv G_\infty(x,n|x_0 = 0)$. Partitioning

over the last step of the walk, one obtains the *forward* equation:

$$G_\infty(x,n) = \delta_{0,n}\delta(x) + [1-\delta_{0,n}]\int_{-\infty}^{\infty} p(x-y)G_\infty(y,n-1)\mathrm{d}y, \tag{4.1}$$

and introducing the Fourier transformed generating function $G_\infty(k,\xi) = \sum_{n=0}^{\infty}\xi^n\int_{-\infty}^{\infty}e^{ikx}G_\infty(x,n)\mathrm{d}x$ of the infinite propagator, it can be shown that (see Appendix G)

$$G_\infty(k,\xi) = \frac{1}{1-\xi\tilde{p}(k)}, \tag{4.2}$$

where $\tilde{p}(k) = \int_{-\infty}^{\infty}e^{ik\ell}p(\ell)\mathrm{d}\ell$ is the Fourier Transform of the single jump distribution. In turn, the real space and time propagator $G_\infty(x-x_0,n)$ is given by:

$$G_\infty(x-x_0,n) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[e^{-ik(x-x_0)}\frac{1}{2\pi i}\oint\frac{1}{\xi^{n+1}}G_\infty(k,\xi)\mathrm{d}z\right]\mathrm{d}k. \tag{4.3}$$

Importantly, equation (4.3) depends only on $\tilde{p}(k)$. Note that the determination of explicit expressions for $G_\infty(x-x_0,n)$ from (4.3) is only possible for a few specific jump distributions $\tilde{p}(k)$. However, numerical evaluation of equation (4.3) is always feasible.

Limit laws for large n and x. In the large n and x limit, it is known [Gnedenko & Kolmogorov 1955, Stone 1967, Feller 1971, Bouchaud & Georges 1990] that the infinite propagator converges towards a limit continuous propagator $G_\infty^{(c)}(x,t|x_0)$, defined only by the large ℓ behavior of the single jump distribution $p(\ell)$, or equivalently by the small k behavior of $\tilde{p}(k)$. More precisely, in the small k limit, one generally has [1]

$$\tilde{p}(k) \underset{k\to 0}{=} 1-(a_\mu|k|)^\mu + o(k^\mu) \tag{4.4}$$

where $\mu\in]0,2]$ is the Levy exponent, characterizing the tails of the jump distribution, and $a_\mu > 0$ its characteristic length scale. Two situations arise:

- For $\mu = 2$, the jump distribution has a finite second moment [2] and the limit continuous process is a Brownian motion with diffusion coefficient $D = a_2^2$ and propagator

$$G_\infty^{(c)}(x,t|x_0) = \frac{1}{\sqrt{4\pi Dt}}e^{-\frac{(x-x_0)^2}{4Dt}}. \tag{4.5}$$

  where $t = n$ is now the continuous time.

- For $\mu < 2$, the jump process is dubbed *heavy-tailed*, and $p(\ell)$ has a diverging second moment. The tails of its jump distribution behave as power laws:

[1] We will take this as a definition of jump processes discussed hereafter. Other cases exist but we will consider them marginal. See [Bouchaud & Georges 1990] for a more in-depth review.

[2] Again, this is not required. For instance, distributions $p(\ell)\propto l^{-3}$ have $\mu = 2$ and no second moment. However, we also consider this case to be marginal.

$$\int_x^\infty p(l)\mathrm{d}l \propto \frac{1}{x^\mu}, \tag{4.6}$$

and the limit continuous process is an α-stable process, whose propagator is given via its Fourier Transform:

$$G_\infty^{(c)}(k,t) = e^{-t|ka_\mu|^\mu}. \tag{4.7}$$

Importantly, the asymptotic behavior of other observables, such as first-passage times for instance, can be extracted from these limit processes. For more details on the continuous processes, we refer the reader to [Redner 2001] for the Brownian case, and [Kyprianou & Pardo 2022] for the α-stable case.

4.1.2 Jump processes on the semi-infinite line

We now turn to jump processes killed (meaning stopped) upon entering the negative half-axis, and denote their propagator $G(x,n|x_0)$. In other words, trajectories contributing to $G(x,n|x_0)$ are n-step long trajectories, starting from x_0 and ending at x, such that all intermediate positions of the particle are positive (see figure 4.1).

Semi-infinite propagator. Similarly to the unconstrained case, G obeys a *forward* equation obtained by partitioning over the last step of the walk

$$G(x,n|x_0) = \delta_{0,n}\delta(x) + [1-\delta_{0,n}]\int_0^\infty p(x-y)G(y,n-1|x_0)\mathrm{d}y, \tag{4.8}$$

but the modified bounds of the integral do not allow for a direct Fourier Transform. However, equation (4.8) belongs to the more general type of Wiener-Hopf integral equations [Hopf 1935]. As a result, by making use of specific associated technical tools, an explicit expression of the semi-infinite propagator has been derived as a function of $\tilde{p}(k)$ only in [Ivanov 1994]. For consistency we give its expression here, and present calculation details in Appendix G, where we provide a clear exposition of the original derivation. The semi-infinite propagator is given explicitly as a double Laplace transform and generating function:

$$\begin{aligned}
&\sum_{n=0}^\infty \xi^n \left[\int_0^\infty\int_0^\infty e^{-sx-s_1x_0}G(x,n|x_0)\mathrm{d}x\mathrm{d}x_0\right] = \frac{G_0(s,\xi)G_0(s_1,\xi)}{s+s_1}\\
&G_0(s,\xi) = \exp\left[-\frac{s}{2\pi}\int_{-\infty}^\infty \frac{\log\left[1-\xi\tilde{p}(k)\right]}{s^2+k^2}\mathrm{d}k\right],
\end{aligned} \tag{4.9}$$

where $G_0(x,n)$ is defined as the semi-infinite propagator starting from $x_0 = 0$. Finally, depending on the specific form of $\tilde{p}(k)$, the real space and time $G(x,n|x_0)$ can either be obtained explicitly, or numerically evaluated.

Survival probability. A quantity of prime importance in the semi-infinite geometry is the so-called survival probability $q(x_0, n)$, defined as the probability for the walker starting from x_0 to stay positive during its first n steps. In terms of semi-infinite propagator $G(x, n|x_0)$, the survival probability is given by integrating over the position of the particle on its n^{th} step

$$q(x_0, n) = \int_0^\infty G(x, n|x_0)\mathrm{d}x, \tag{4.10}$$

and reads, in the Laplace transformed space [3]:

$$\sum_{n=0}^{\infty} \xi^n \left[\int_0^\infty e^{-sx_0} q(x_0, n)\mathrm{d}x_0\right] = \frac{1}{s\sqrt{1-\xi}} \exp\left[-\frac{s}{2\pi}\int_{-\infty}^{\infty} \frac{\log\left[1-\xi\tilde{p}(k)\right]}{s^2+k^2}\mathrm{d}k\right]. \tag{4.11}$$

We now make a few important remarks on the survival probability stemming from equation (4.11).

- The small x_0 behavior of $q(x_0, \xi)$ is obtained by identifying the leading large s behavior in the LHS and RHS of equation (4.11). In turn, we recover the celebrated Sparre Andersen result [Andersen 1954] of the survival probability starting from 0, initially derived using combinatorial techniques:

$$\sum_{n=0}^{\infty} \xi^n q(0, n) = \frac{1}{\sqrt{1-\xi}}, \tag{4.12}$$

  such that

$$q(0, n) = \binom{2n}{n}\frac{1}{2^{2n}}. \tag{4.13}$$

  We emphasize that $q(0, n)$ is a purely combinatorial number, and is independent of the single jump distribution $p(\ell)$. Importantly, due to the discrete nature of the jump process, $q(0, n)$ is *non-vanishing* for all values of n. In contrast, for continuous non-smooth processes, $q(x_0, t)$, defined as the probability to stay positive up to time t, necessarily vanishes when $x_0 \to 0$.

- The large n behavior of $q(x_0, n)$ is obtained by considering the $\xi \to 1$ limit in equation (4.11), and is given by [Doney 2012, Majumdar *et al.* 2017]

$$q(x_0, n) \underset{n\to\infty}{\sim} \frac{1}{\sqrt{n}}\left[\frac{1}{\sqrt{\pi}} + V(x_0)\right], \tag{4.14}$$

  where $V(x_0)$ is defined by its Laplace transform:

$$\int_0^\infty e^{-sx_0} V(x_0)\mathrm{d}x_0 = \frac{1}{s\sqrt{\pi}}\left(\exp\left[-\frac{s}{\pi}\int_0^\infty \frac{\mathrm{d}k}{s^2+k^2}\ln(1-\tilde{p}(k))\right] - 1\right). \tag{4.15}$$

[3] This is obtained from the limit $G_0(s \to 0, \xi)$ in equation (4.9).

- The function $V(x_0)$ can be further analyzed in the small x_0 limit [Majumdar *et al.* 2017, Klinger *et al.* 2022b], and displays universal behavior depending only on the small scale details of $p(\ell)$[4]:

$$V(x_0) = \begin{cases} -\left[\pi^{-\frac{3}{2}} \int_0^\infty \mathrm{d}k \log(1-\tilde{p}(k))\right] x_0 + o(x_0) & \text{if } \tilde{p}(k) \underset{k\to\infty}{=} o(k^{-1}) \\ \dfrac{\beta}{2\sqrt{\pi}\Gamma(1+\nu)\cos(\pi\nu/2)} x_0^\nu + o(x_0^\nu) & \text{if } \tilde{p}(k) \underset{k\to\infty}{\sim} \beta k^{-\nu} \text{ with } \nu < 1 \\ -\dfrac{\beta}{\pi^{3/2}} x_0 \ln(x_0) + o\left(x_0 \ln(x_0)\right) & \text{if } \tilde{p}(k) \underset{k\to\infty}{\sim} \beta k^{-1} \end{cases} \tag{4.16}$$

Limit laws for large n and x. Similarly to the unconstrained case, jump processes on the semi-infinite line also converge towards limit continuous processes for large n and x values. More precisely, one needs to consider jump processes *conditioned* to stay positive, and introduce a normalized propagator $G(x,n|x_0)/q(x_0,n)$. A few recent papers from the mathematical literature, among which [Caravenna 2005, Caravenna & Chaumont 2008] focus on proving convergence theorems for *conditioned* jump processes towards corresponding *conditioned* continuous stable processes. In turn, asymptotic results for FPT distributions of jump processes can be obtained, as we will see shortly.

4.1.3 Jump processes in an interval

We finally consider jump processes killed upon exiting an interval $[0,x]$. As in the infinite and semi-infinite cases, the *bounded* propagator $G_{[0,x]}$ obeys a *forward* equation

$$G_{[0,x]}(u,n|x_0) = \delta_{0,n}\delta(u-x_0) + [1-\delta_{0,n}] \int_0^x p(u-y) G_{[0,x]}(y,n-1|x_0)\mathrm{d}y. \tag{4.17}$$

Importantly, the bounds of the integral are now 0 and x, preventing any kind of Fourier or Laplace transforms to simplify equation (4.17).

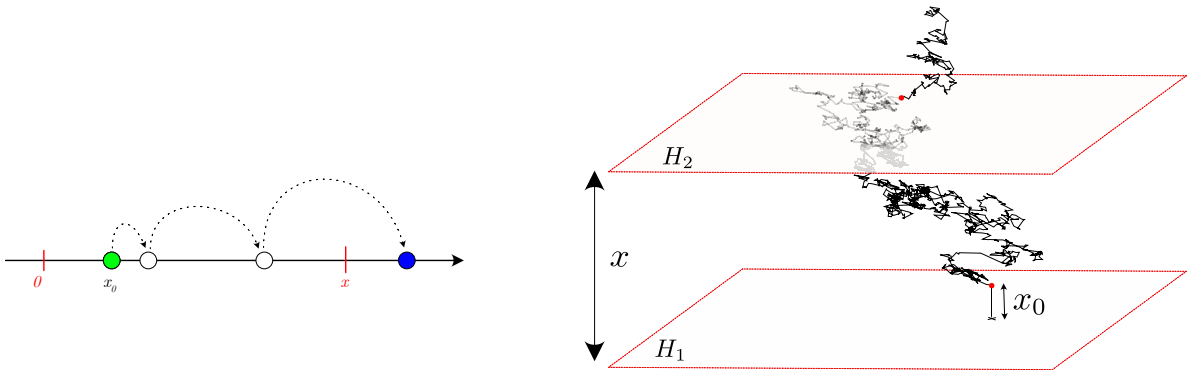
We emphasize that, in contrast to the infinite and semi-infinite cases, *much less* is known in the bounded case. In fact, for arbitrary jump processes, no explicit form of $G_{[0,x]}$ has been derived up to now [5], nor of any other observables, such as the survival probability in the interval for instance, defined as the probability to stay inside the interval during the first n steps. In the following chapters, we investigate some natural observables associated to bounded jump processes and begin with the splitting probability, for which there is to date no general explicit expression.

[4] The specific case of the linear behavior was already accounted for in [Majumdar *et al.* 2017].

[5] For some $p(\ell)$, the bounded propagator can be derived by exploiting the specific form of the single jump distribution. For instance, the exponential jump distribution $p(\ell) \propto e^{-|\ell|}$ is tractable, see [Mori *et al.* 2020b].

4.2 Determination of the splitting probability

The splitting probability $\pi_{0,\underline{x}}(x_0)$ is defined as the probability to strictly exit the interval through x rather than through 0 (hence the underlined x) starting from x_0. In other words, the splitting probability conveys information relative to the direction of exit of the interval: either rightwards or leftwards (see schematic 4.2(a)).



(a) Example trajectory contributing to $\pi_{0,\underline{x}}(x_0)$. Starting from x_0 the walker takes two steps inside the interval, before jumping over x.

(b) 3 dimensional hyperplane splitting probability: starting from a distance x_0 from H_1, the isotropic jump process crosses H_2 before H_1. Importantly this is strictly equivalent to the splitting probability $\pi^{\perp}_{0,\underline{x}}(x_0)$ of the effective one-dimensional projected jump process.

Figure 4.2

By partitioning over the first step of the walk, we first show that the splitting probability obeys a *backward* equation:

$$\pi_{0,\underline{x}}(x_0) = \int_x^\infty \mathrm{d}y p(y - x_0) + \int_0^x \pi_{0,\underline{x}}(y) p(y - x_0) \mathrm{d}y, \tag{4.18}$$

and make a few important remarks.

- Because of the integral bounds appearing in equation (4.18), the splitting probability $\pi_{0,\underline{x}}(x_0)$ has only been explicitly derived for a few specific single jump distributions, such as the exponential jump distribution $p(\ell) = \frac{1}{2}e^{-|\ell|}$ [Van Kampen 1992].

- It is clearly apparent from equation (4.18) that $\pi_{0,\underline{x}}(x_0)$ is non vanishing as $x_0 \to 0$. In other words, even when starting from $x_0 = 0$, the jump process has a non 0 probability of escaping the interval through x rather than through 0. In contrast, splitting probabilities of continuous processes obey a backward differential equation, and $\pi_{0,\underline{x}}(0) = 0$ by definition.

To obtain explicit results, we focus in the following on the large x limit of the splitting probability, and uncover asymptotic universal behavior of $\pi_{0,\underline{x}}(x_0)$, valid for arbitrary symmetric continuous $p(\ell)$.

4.2.1 Splitting and survival probabilities

The core ingredient of our asymptotic result is to exhibit a link between the splitting probability $\pi_{0,\underline{x}}(x_0)$ in the bounded interval $[0, x]$, and the survival probability $q(x_0, n)$ in the semi-infinite space $[0, +\infty]$. We here take the opportunity to thank Quentin Berger from the LPSM, Sorbonne Université, for providing help and guidance in establishing a clean proof of this relation.

We first introduce τ^x and τ^0, the two stopping times associated with the strict crossing of respectively x and 0, and rewrite $\pi_{0,\underline{x}}(x_0) = P(\tau^x < \tau^0|x_0)$. Next, partitioning over the step at which first exit occurs, we have

$$P(\tau^x < \tau^0|x_0) = \sum_{k=1}^{\infty} P(\tau^x = k, \tau^0 > k|x_0). \tag{4.19}$$

We now show that in the large x limit, the above given sum is controlled by steps $k \propto n_x \equiv \left[\frac{x}{a_\mu}\right]^\mu$, with μ and a_μ defined in (4.4). Introducing an arbitrary parameter ε and splitting the sum in 3 parts yields [6]:

$$\begin{aligned} P(\tau^x < \tau^0|x_0) = &\sum_{k=1}^{\varepsilon n_x} P(\tau^x = k, \tau^0 > k|x_0) \qquad (A) \\ &+ \sum_{k=\varepsilon n_x}^{\varepsilon^{-1} n_x} P(\tau^x = k, \tau^0 > k|x_0) \qquad (B) \\ &+ \sum_{k=\varepsilon^{-1} n_x}^{\infty} P(\tau^x = k, \tau^0 > k|x_0) \qquad (C). \end{aligned} \tag{4.20}$$

With x large and fixed, let us show that both terms (A) and (C) vanish as ε goes to 0. Starting with term (C), we have:

$$(C) \leq P\Big(\min(\tau^x, \tau^0) \geq n_x \varepsilon^{-1} | x_0 \Big). \tag{4.21}$$

For a given μ, it is known that the large time leading order of the survival probability in the interval $[0, x]$ is given by a decaying exponential $P(\min(\tau^x, \tau^0) \geq n|x_0) \propto e^{-c_1 \frac{n}{x^\mu}}$, where c_1 is a process dependent but an x and n independent constant [Kwaśnicki 2012, Dybiec *et al.* 2016]. In turn, we obtain

$$(C) \leq c_2 e^{-\frac{c_3}{\varepsilon}}, \tag{4.22}$$

with c_2 and c_3 some x independent constants. As a result, (C) can be arbitrarily small as $\varepsilon \to 0$, independently of the value of x . Focusing now on the (A) term, we obtain an upper bound

$$\begin{aligned} (A) &\leq P\left(\tau^x \leq n_x \varepsilon | x_0\right) \\ &\leq 1 - P\left(\tau^x \geq n_x \varepsilon | x_0\right) \end{aligned} \tag{4.23}$$

[6] the bounds of the sums should be understood as integer parts.

which involves the survival probability of the jump process on the semi-infinite line. Using results from [Doney 2012], one has

$$P\left(\tau^x \geq n_x\varepsilon|x_0\right) \underset{x\to\infty}{\sim} \int_0^\infty h_\varepsilon(t)\mathrm{d}t, \tag{4.24}$$

where h_ε is the first-passage time distribution to 0 of the semi-infinite limit process defined in the previous section, starting from $\varepsilon > 0$. As $\varepsilon \to 0$, $h_\varepsilon(t) \to \delta(t)$, such that (A) can also be made arbitrarily small as $\varepsilon \to 0$, independently of the value of x. We now rewrite the only non-vanishing term (B) by conditioning on the time at which x is crossed

$$P(\tau^x < \tau^0|x_0) = \sum_{k=\varepsilon n_x}^{\varepsilon^{-1}n_x} P(\tau^x = k|\tau^0 > k, x_0)q(x_0, k), \tag{4.25}$$

where $q(x_0, k) = P(\tau^0 > k|x_0)$ is the semi-infinite survival probability discussed above. We now make a few remarks:

- First, for large x, k is large and $x \gg x_0$, such that $q(x_0, k)$ takes the form (4.14)

$$P(\tau^x < \tau^0|x_0) = \sum_{k=\varepsilon n_x}^{\varepsilon^{-1}n_x} P(\tau^x = k|\tau^0 > k, x_0)\frac{1}{\sqrt{k}}(1 + V(x_0)). \tag{4.26}$$

- Second, using results from [Caravenna & Chaumont 2008] and [Doney 2012], the conditional FPT distribution $P(\tau^x = k|\tau^0 > k, x_0)$ converges towards an x_0 independent and scale-invariant limit distribution

$$P(\tau^x = k|\tau^0 > k, x_0) \underset{x\to\infty}{\sim} \frac{1}{n_x}h^+\left(\frac{k}{n_x}\right). \tag{4.27}$$

  We emphasize that h^+ depends only on μ, as the rescaling factor a_μ has been absorbed in the definition of n_x.

Plugging equation (4.27) into (4.26) and taking ε to 0 finally yields

$$\pi_{0,\underline{x}}(x_0) \underset{x\to\infty}{\sim} q(x_0, n_x)\int_0^\infty \frac{h^+(t)}{t}\mathrm{d}t. \tag{4.28}$$

Importantly, we successfully related the splitting probability, which is an unknown observable associated to the bounded jump process, to the semi-infinite survival probability, which has been extensively studied in the literature. However, equation (4.28) still contains an unknown process dependent proportionality constant $\int_0^\infty h^+(t)t^{-1}\mathrm{d}t$, which remains to be evaluated.

4.2.2 Prefactor identification

We emphasize that the unknown coefficient in the asymptotic result (4.28) depends only on the value of μ, which characterizes the limit continuous process. In particular, it is both independent of x_0 and a_μ.

Continuous limit. Since equation (4.28) is valid for all $x_0 \ll x$, we place ourselves in the *continuous limit*, defined for initial x_0 condition such that $a_\mu \ll x_0 \ll x$. Importantly, in this limit, both the splitting probability and survival probability are asymptotically equivalent to those of the limit continuous processes, denoted by $\pi^{(c)}_{0,\underline{x}}(x_0)$ and $q^{(c)}(x_0, n)$. In turn, we identify the sought-after prefactor:

$$\int_0^\infty \frac{h^+(t)}{t}\mathrm{d}t \underset{x\to\infty}{\sim} \frac{\pi_{0,\underline{x}}(x_0)}{q(x_0,n_x)} \underset{a_\mu \ll x_0 \ll x}{\sim} \frac{\pi^{(c)}_{0,\underline{x}}(x_0)}{q^{(c)}(x_0,n_x)} \tag{4.29}$$

Let us first focus on the splitting probability, whose explicit expression is given by [Widom 1961, Blumenthal *et al.* 1961, Majumdar *et al.* 2010b]:

$$\pi^{(c)}_{0,\underline{x}}(x_0) = \frac{\Gamma(\mu)}{\Gamma^2(\frac{\mu}{2})}\int_0^{\frac{x_0}{x}} [u(1-u)]^{\mu/2-1}\,\mathrm{d}u \tag{4.30}$$

and is valid for all μ and $0 \le x_0 \le x$. Note that expression (4.30) vanishes as $x_0 \to 0$, as expected of a continuous description. Considering further the small x_0 regime, one obtains:

$$\pi^{(c)}_{0,\underline{x}}(x_0) \underset{a_\mu \ll x_0 \ll x}{\sim} \frac{2\Gamma(\mu)}{\mu\Gamma^2(\frac{\mu}{2})}\left(\frac{x_0}{x}\right)^{\frac{\mu}{2}}. \tag{4.31}$$

We next turn to the survival probability, and make use of the asymptotic behavior of $q(x_0, n)$ obtained in [Majumdar *et al.* 2017], which yields for large n and $a_\mu \ll x_0$

$$q^{(c)}(x_0,n) \sim \frac{1}{\sqrt{n}}\frac{a_\mu^{-\frac{\mu}{2}}}{\sqrt{\pi}\Gamma(1+\frac{\mu}{2})}x_0^{\frac{\mu}{2}}. \tag{4.32}$$

Combining equations (4.29), (4.31) and (4.32), we finally obtain a universal exact asymptotic formula for the splitting probability of arbitrary jump processes:

$$\begin{aligned}\lim_{x\to\infty}\left[\frac{\pi_{0,\underline{x}}(x_0)}{A_\mu(x)}\right] &= \frac{1}{\sqrt{\pi}} + V(x_0)\\ A_\mu(x) &= \left(\frac{a_\mu}{x}\right)^{\mu/2} 2^{\mu-1}\Gamma\left(\frac{1+\mu}{2}\right).\end{aligned} \tag{4.33}$$

- Equation (4.33) is the main result of this chapter. It is valid for any symmetric continuous jump process, and for all x_0 regimes such that $0 \le x_0 \ll x$. In particular, taking $x_0 \gg a_\mu$, one recovers the continuous limit regime.

- Of prime importance is the regime $0 \le x_0 \lesssim a_\mu$ in which the discrete nature of the process is highly relevant, and is not captured by a continuous description. In this regime, the x_0 behavior of $\pi_{0,\underline{x}}(x_0)$ is described by the universality classes defined in equation (4.16), in strong contrast with the single behavior predicted by the continuous limit (4.31).

- Finally, we underline a striking result: the splitting probability starting from 0 takes the following universal asymptotic form:

$$\pi_{0,\underline{x}}(0) \underset{x\to\infty}{\sim} \frac{2^{\mu-1}}{\sqrt{\pi}}\Gamma\left(\frac{1+\mu}{2}\right)\left(\frac{a_\mu}{x}\right)^{\frac{\mu}{2}}. \tag{4.34}$$

Importantly, while $\pi_{0,\underline{x}}(x_0)$ from equation (4.33) depends on the full jump distribution $p(\ell)$ via $V(x_0)$, the splitting probability starting from zero depends on $p(\ell)$ via a_μ and μ only. In turn, analytical or numerical evaluation of μ and a_μ is sufficient to determine $\pi_{0,\underline{x}}(0)$.

4.3 Illustrations and applications of the asymptotic splitting probability

While equation (4.33) is an exact asymptotic result, we provide a few exactly solvable cases for illustration.

4.3.1 Exactly solvable cases and numerical simulations

Exponential jump distribution. We first consider the jump distribution $p(\ell) = \frac{\gamma}{2}e^{-\gamma|\ell|}$, for which the splitting probability is known exactly and is given by [Van Kampen 1992]:

$$\begin{aligned}\pi_{0,\underline{x}}(x_0) &= \frac{1}{x\gamma+2}x_0\gamma + \frac{1}{x\gamma+2}\\ \pi_{0,\underline{x}}(x_0) &\underset{x\to\infty}{\sim} \frac{1}{x}\left[x_0+\frac{1}{\gamma}\right].\end{aligned} \tag{4.35}$$

In that case, the Fourier transform of $p(\ell)$ reads $\tilde{p}(k) = \frac{1}{(k/\gamma)^2+1}$, such that $\mu = 2$ and $a_2 = \gamma^{-1}$. Furthermore, the specific form of $\tilde{p}(k)$ allows for the explicit derivation of $V(x_0)$:

$$\begin{aligned}\int_0^\infty e^{-sx_0}V(x_0)\mathrm{d}x_0 &= \frac{1}{s\sqrt{\pi}}\left(\exp\left[-\frac{s}{\pi}\int_0^\infty \frac{\mathrm{d}k}{s^2+k^2}\ln(1-\tilde{p}(k))\right]-1\right)\\ &= \frac{1}{s\sqrt{\pi}}\left(\exp\left[-\frac{s}{\pi}\int_0^\infty \frac{\mathrm{d}k}{s^2+k^2}\left(\ln(k^2/\gamma^2)-\ln(1+k^2/\gamma^2)\right)\right]-1\right).\end{aligned} \tag{4.36}$$

By making use of the identity $\int_0^\infty \frac{\ln(\alpha+\beta k^2)}{\lambda^2+k^2}\mathrm{d}k = \frac{\pi}{\lambda}\ln(\sqrt{\alpha}+\sqrt{\beta}\lambda)$ [Majumdar *et al.* 2010a], we obtain:

$$\begin{aligned}\int_0^\infty e^{-sx_0}V(x_0)\mathrm{d}x_0 &= \frac{\gamma}{\sqrt{\pi}s^2},\\ V(x_0) &= \frac{\gamma x_0}{\sqrt{\pi}}.\end{aligned} \tag{4.37}$$

In turn, plugging into equation (4.33) with $A_2(x) = \sqrt{\pi}/(\gamma x)$, we compute the asymptotic splitting probability:

$$\begin{aligned}\pi_{0,\underline{x}}(x_0) &\underset{x\to\infty}{\sim} A_2(x)\left[\frac{1}{\sqrt{\pi}}+\frac{\gamma x_0}{\sqrt{\pi}}\right]\\ &\underset{x\to\infty}{\sim} \frac{1}{x}\left[x_0+\frac{1}{\gamma}\right],\end{aligned} \tag{4.38}$$

in agreement with the direct computation (4.35).

Gamma jump process. We derive explicitly the splitting probability for the gamma jump process defined by $p(\ell) = \frac{\gamma^2}{2}|\ell|e^{-\gamma|\ell|}$. We stress that, to the best of our knowledge, this quantity has not been reported in the literature. Exploiting the singular nature of $p(\ell)$, we show that the jump distribution obeys the following differential equation

$$\left[\frac{\partial^2}{\partial^2 y} - \gamma^2\right]^2 p(y-x) = 2\gamma^4\delta(y-x) + 2\gamma^2\delta^{(2)}(y-x), \tag{4.39}$$

with $\delta^{(2)}$ the second derivative of the Dirac delta function. Recalling the backward equation for the splitting probability

$$\pi_{0,\underline{x}}(x_0) = \int_0^x p(y-x_0)\pi_{0,\underline{x}}(y)\mathrm{d}y + \int_x^\infty p(y-x_0)\mathrm{d}y \tag{4.40}$$

and applying the differential operator $[d_2 - \gamma^2]^2$ with respect to x_0 (where $d_2 = \frac{\partial^2}{\partial^2 x_0}$), we obtain a differential equation for the splitting probability:

$$\frac{\partial^4}{\partial^4 x_0}\pi_{0,\underline{x}}(x_0) - 3\gamma^2\frac{\partial^2}{\partial^2 x_0}\pi_{0,\underline{x}}(x_0) = 0. \tag{4.41}$$

In turn, decomposing onto orthogonal solutions $Ae^{-\sqrt{3}\gamma x_0}, Be^{+\sqrt{3}\gamma x_0}, Cx_0, D$, and inserting the solution into the integral equation (4.40) is sufficient to identify coefficients A, B, C and D, which are given in Appendix F. For the sake of simplicity, we focus on the splitting probability starting from 0. On the one hand, developing the exact solution for large x and taking x_0 to 0 yields

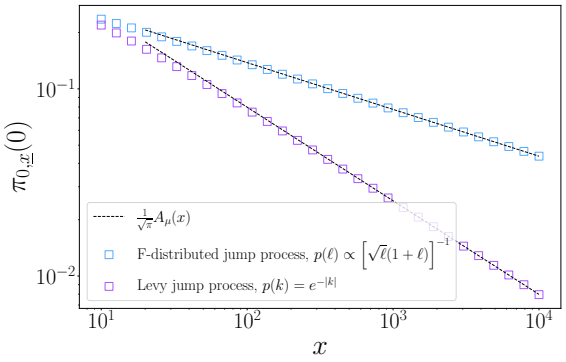
$$\pi_{0,\underline{x}}(0) \underset{x\to\infty}{\sim} \frac{\sqrt{3}}{\gamma x}. \tag{4.42}$$

On the other hand, the Fourier Transform of $p(\ell)$ is given by $\tilde{p}(k) = \gamma^2\frac{\gamma^2-k^2}{(\gamma^2+k^2)^2}$, such that $\mu = 2$ and $a_2 = \sqrt{3}/\gamma$. Plugging μ and a_2 into equation (4.34), we recover the asymptotic result (4.42) obtained from the direct calculation of $\pi_{0,\underline{x}}(x_0)$.

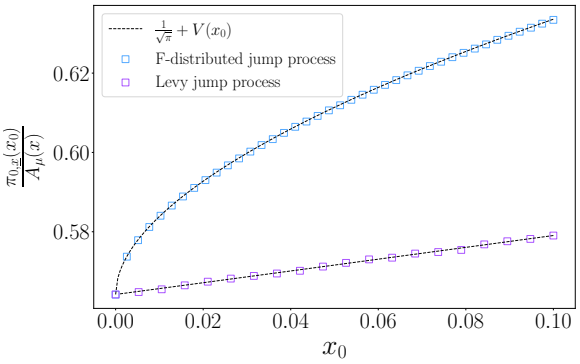
Numerical simulations. For other jump processes, the exact determination of $\pi_{0,\underline{x}}(x_0)$ does not seem possible, and we resort to numerical simulations. We choose to focus on processes with $\mu < 2$ as both the exponential and gamma jump processes belong to the $\mu = 2$ universality class. We display in figure 4.3 results for two heavy-tailed processes: one *Levy Flight* with $\tilde{p}(k) = e^{-|k|}$, such that $\mu = 1$ and $a_\mu = 1$, and the F-distributed jump process, defined by $p(\ell) = (2\pi\sqrt{|\ell|}(1+|\ell|))^{-1}$ and with $\mu = 1/2$ and $a_\mu = 2/\pi$.

4.3.2 Transmission probability in scattering experiments

The determination of splitting probabilities for jump processes is not only of theoretical interest, but also has direct implications in experimental situations. As an illustration, we consider hereafter the case of photon scattering in hot atomic vapors, recently studied experimentally in [Baudouin *et al.* 2014, Araújo *et al.* 2021, Lopez *et al.* 2023]. The trajectories of photons in



(a) Algebraic decay of the large x splitting probability from 0, computed from equation (4.34). The F-distributed jump process has $\mu = 1/2$ and $a_\mu = 2/\pi$. Both exponent and prefactor are correctly captured.



(b) x_0 dependence of $\pi_{0,\underline{x}}(x_0)$ in the large x limit, with $x = 10^6$ fixed. The dashed lines are obtained from numerical inversion of the Laplace transform of $V(x_0)$ (4.15). Importantly, the limiting value at $x_0 = 0$ is clearly non 0.

Figure 4.3

three dimensional hot atomic vapors have been shown to be correctly modeled by isotropic *Levy Flights*, with reorientation events occurring upon scattering, and jump length distribution $\tilde{p}(k) = e^{-|a_\mu k|^\mu}$ [7]. In the specific case of hot vapors, the Levy exponent μ has been experimentally measured and is such that $\mu \simeq \frac{1}{2}$.

Since measuring single jump distributions for scattered photons can be difficult, recent experimental setups have focused on the *transmission probability* across three dimensional slabs of width x containing the scattering medium (see 4.2(b)). By shining a beam of photons through one side and measuring the transmitted intensity on the other side, one can obtain the corresponding transmission probability, defined as the probability for a photon to exit the slab through the side opposite to its entry point. Importantly, incoming photons hit the slab from the exterior, and the motion parallel to the two delimiting hyperplanes H_1 and H_2 (see 4.2(b)) is irrelevant to the direction of exit. In turn, the transmission probability is strictly equivalent to the splitting probability from 0 of the jump process describing the motion perpendicular to H_1 and H_2. Consequently, making use of equation (4.34) and identifying the effective 1D perpendicular jump process is enough to determine the transmission probability, in the limit of large slab width x.

Effective jump process. We first focus on the specific case of 3D isotropic Levy Flights, and derive the single jump distribution of the one dimensional jump process describing the motion perpendicular to H_1 and H_2, which we refer to as the projected process. We recall that, in the 3D general case, the projected distribution $p_\perp(\ell)$ is given by $p_\perp(\ell) = \frac{1}{2}\int_{|\ell|}^{\infty} \frac{p(r)+p(-r)}{r} \mathrm{d}r$ [Mori *et al.* 2020a]. In the Fourier formalism, we obtain:

[7] Note that to ease calculations the jump length can be negative. As a result, the distribution $p_1(\ell)$ of the true jump length is given by $p_1(\ell) = p(\ell) + p(-\ell)$.

$$\begin{aligned}\tilde{p}_\perp(k) &= \frac{1}{2}\int_{-\infty}^{\infty} e^{ik\eta}\left[\int_0^\infty \frac{1}{l}\theta\left(1-\left|\frac{\eta}{l}\right|\right)(p(l)+p(-l))\mathrm{d}l\right]\mathrm{d}\eta \\ &= \int_0^\infty \frac{\sin(kl)}{kl}(p(l)+p(-l))\mathrm{d}l \\ &= \frac{1}{k}\int_{-\infty}^{\infty}\frac{\sin(kl)}{l}p(l)\mathrm{d}l\end{aligned} \tag{4.43}$$

We now use the specific form of the 3D jump distribution $\tilde{p}(k) = e^{-|a_\mu k|^\mu}$ and write $\sin(kl) = \mathrm{Im}(e^{ikl})$:

$$\begin{aligned}\tilde{p}_\perp(k) &= \frac{1}{k}\int_{-\infty}^{\infty}\frac{e^{ikl}}{l}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ik'l-|a_\mu k'|^\mu}\mathrm{d}k'\right]\mathrm{d}l \\ &= \frac{1}{k}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{e^{il(k-k')}}{l}\mathrm{d}l\frac{1}{2\pi}e^{-|a_\mu k'|^\mu}tdk' \\ &= \frac{i}{k}\int_{-\infty}^{\infty}\frac{1}{2}\mathrm{sign}(k-k')e^{-|a_\mu k'|^\mu}\mathrm{d}k'.\end{aligned} \tag{4.44}$$

Taking the imaginary part finally yields Fourier transform of the projected jump distribution:

$$\tilde{p}_\perp(k) = \frac{\Gamma\left(\frac{1}{\mu}\right) - \Gamma\left(\frac{1}{\mu}, (a_\mu k)^\mu\right)}{a_\mu \mu k} \tag{4.45}$$

with $\Gamma(s,x)$ the incomplete gamma function. The small k and large k asymptotic analysis leads to

$$\begin{aligned}\tilde{p}_\perp(k) &\underset{k\to 0}{\sim} 1 - \frac{(a_\mu k)^\mu}{1+\mu} + o(k^\mu) \\ \tilde{p}_\perp(k) &\underset{k\to\infty}{\sim} \frac{1}{k}\frac{\Gamma\left(\frac{1}{\mu}\right)}{c\mu} + o\left(\frac{1}{k}\right)\end{aligned} \tag{4.46}$$

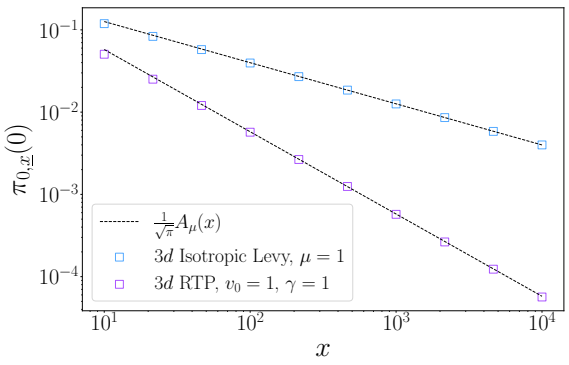
from which the transmission probability is obtained by extracting μ and a_μ, and making use of equation (4.34). Importantly:

- The $x^{-\frac{\mu}{2}}$ decay of the transmission probability observed experimentally [Araújo *et al.* 2021] is in agreement with the x decay of $\pi_{0,\underline{x}}(0)$. By fitting the prefactor of the experimentally observed algebraic decay and comparing with the analytical prediction (4.34) one could extract a numerical value for a_μ, which characterizes the underlying scattering process. We emphasize that measuring transmission probabilities is experimentally easier than measuring the single jump distribution of a scattered photon. In turn our result provides offers a simplified access to the characteristic lengthscale of the photon scattering process.
- Note that the k^{-1} behavior of the projected jump distribution for large k leads to a non-linear small x_0 behavior (see equation (4.16)) of the transmission probability for photons starting inside the slab. This analytical result is in agreement with recent experiments [Lopez *et al.* 2023] focusing on the x_0 dependence of the transmission probability, and where such non-linear behavior is reported.

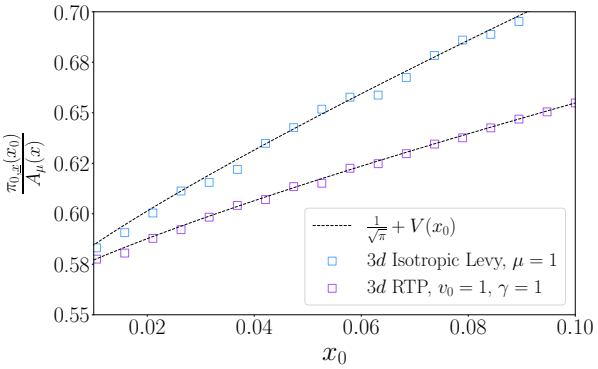
Isotropic run and tumble particles. We next focus on a second photon scattering model: the isotropic *Run and Tumble* model (RTP) which we already discussed in chapter 1. In this model photons - or more generally particles - perform exponentially distributed run phases followed by isotropic reorientation. Importantly, in the original model, the particles travel at constant speed v, such that the process is not discrete in time and runs have a duration $t = \ell v^{-1}$. However, the splitting probability is a strictly geometrical quantity, and does not depend on time, such that this distinction is irrelevant. Of note, the isotropic Run and Tumble model has recently received a lot of attention [Tailleur & Cates 2008, Rupprecht *et al.* 2016, Mori *et al.* 2020a]. For the RTP, the run length distribution is given by $p(\ell) = \frac{\gamma}{2v} e^{-\frac{\gamma}{v}|\ell|}$, with γ the reorientation (tumbling) rate, and v the particle speed. Note that, by convention, the run length can be negative. The projected jump distribution reads:

$$\begin{aligned} p_\perp(\ell) &= \frac{\gamma}{2v}\Gamma\left(0, \frac{\gamma}{v}|\ell|\right) \\ \tilde{p}_\perp(k) &= \frac{\gamma \arctan\left(\frac{vk}{\gamma}\right)}{vk}, \end{aligned} \tag{4.47}$$

from which the asymptotic transmission probability is obtained with the help of equation (4.34). We display agreement with numerical simulations of transmission probabilities through 3D slabs, for both RTP and Levy Flights, in figure 4.4.



(a) Transmission probability for a 3d isotropic Levy jump process and a Run and Tumble particle. In both cases the Levy exponent μ of the projected process is identical to that of the initial process, and the prefactor is correctly captured.



(b) Transmission probability for isotropic 3d processes starting inside the slab. The non-linear x_0 dependence stems from the distributions (4.46) and (4.47) of the projected processes.

Figure 4.4

4.4 Conclusion

Jump processes are discrete time and continuous space stochastic processes, whose trajectories can be decomposed into *i.i.d.* increments with a symmetric and continuous distribution $p(\ell)$. Introduced in the early 20^{th} century, they have been used to model a variety of physical phenomena, from one-dimensional particle diffusion to three dimensional photon scattering in heterogeneous media.

One-dimensional unbounded jump processes are well understood, and characterized by their propagator $G_\infty(x, n|x_0)$, defined as the probability to be at x on the n^{th} step, starting from x_0, which can be explicitly written as a function of $p(\ell)$ only. For jump processes killed (*ie* stopped) upon entering the negative half axis, the corresponding semi-infinite propagator $G(x, n|x_0)$ can also be explicitly computed in terms of $p(\ell)$ only, and we reproduce the original derivation [Ivanov 1994] in Appendix G. In turn, the determination of G yields the corresponding survival probability $q(x_0, n)$, defined as the probability to stay positive during the first n steps, for which we exhibit unreported universal x_0 behavior in the small $x_0 \ll 1$ regime.

However, it appears from studying infinite and semi-infinite jump processes that *very few* results exist for *bounded* jump processes, killed upon strictly exiting an interval $[0, x]$. While asymptotic results can be extracted from limit continuous processes, defined only by the tails of the jump distribution $p(\ell)$, these cannot properly capture the specific behavior arising from the discrete nature of the jump process, and in particular the non vanishing value of observables for processes starting from 0 or x.

In an effort to bridge that gap, we study the most natural observable associated to jump processes in an interval, the splitting probability $\pi_{0,\underline{x}}(x_0)$ defined as the probability to exit the interval through x rather than through 0. Our results are valid for general symmetric continuous jump processes. We derive a universal, exact large x asymptotic form of the splitting probability, and obtain a fully explicit determination of the transmission probability ($x_0 = 0$), in striking contrast with the trivial prediction $\pi^{(c)}_{0,\underline{x}}(0) = 0$ obtained from the continuous limit. We illustrate these results on paradigmatic models of jump processes, with applications to photon scattering in heterogeneous media in realistic 3D slab geometries.

We emphasize that the splitting probability is only a geometrical observable, and does not carry information about the dynamical aspects of interval exit events. In particular, the rightward exit-time probability $F_{0,\underline{x}}(n|x_0)$, defined as the probability to escape through x on the n^{th} step exactly, contains refined information on the exit events. To further characterize the statistics of *bounded* jump processes, we focus in the next chapter on both $F_{0,\underline{x}}(n|x_0)$ and its leftward counterpart $F_{\underline{0},x}(n|x_0)$.

Chapter 5

Exit-time probabilities of jump processes

Contents

The aim of this chapter is to obtain refined information on the escape from an interval for general symmetric continuous jump processes. While chapter 4 focused on the sole direction of exit, quantified by the splitting probability $\pi_{0,\underline{x}}(x_0)$, we now investigate the dynamics of escape, by considering jointly the direction of exit and the step n at which exit occurs.

To that end, we introduce and study the leftward and rightward exit-time probabilities $F_{\underline{0},x}(n|x_0)$ (LETP), $F_{0,\underline{x}}(n|x_0)$ (RETP), defined as the probability that the walker escapes the interval through 0 (respectively x) on the n^{th} step exactly. In the large n and x limit, we first consider the continuous limit $x_0 \gg a_\mu$ where a_μ is the characteristic lengthscale of the jump process. In this regime, we show that the LETP and RETP are equivalent to the continuous leftward and rightward exit-time distributions of the limit continuous process, and we review existing results for Brownian motion and α-stable processes.

We then focus specifically on the rightward exit-time probability in the large n and x regime with $0 \leq x_0 \lesssim a_\mu$, and show that the RETP takes an asymptotic universal scaling form, valid for general jump processes. The leftward exit-time probability is dealt with similarly. Importantly, while asymptotic in x and n, our results properly capture the discrete nature of jump processes. In particular, we obtain exact non-vanishing LETP and RETP values for the *edge* initial condition $x_0 = 0$, which, as for $\pi_{0,\underline{x}}(0)$, is not properly described by continuous models. Note that most of the results presented in this chapter can be found in [Klinger *et al.* 2023].

5.1 Asymptotic behavior of exit-time probabilities in the continuous limit

In an effort to characterize further the dynamics of *bounded* jump processes, we first investigate the complete exit-time probability $F_{\underline{0},\underline{x}}(n|x_0)$, defined as the probability for the walker to exit the interval $[0,x]$ on the n^{th} step exactly, whether it be through x or 0. Mimicking the treatment of the bounded propagator, $F_{\underline{0},\underline{x}}(n|x_0)$ obeys a backward equation obtained by partitioning over the first step of the walk:

$$F_{\underline{0},\underline{x}}(n|x_0) = (1-\delta_{n,1})\int_0^x p(y-x_0)F_{\underline{0},\underline{x}}(n-1|y)\mathrm{d}y + \delta_{n,1}\left[1-\int_0^x p(y-x_0)\mathrm{d}y\right]. \tag{5.1}$$

As we have seen extensively, this type of integral equation cannot be solved exactly for arbitrary jump distributions. To proceed further and uncover general results, we consider in the following the asymptotic limits $n\to\infty$ and $x\to\infty$. Note that, even in this limit, the x_0 dependence of $F_{\underline{0},\underline{x}}(n|x_0)$ in the regime $0\leq x_0 \lesssim a_\mu$ is still controlled by the discrete nature of the jump process. In particular, we emphasize that for all n and x values, $F_{\underline{0},\underline{x}}(n|0)$ is strictly non-vanishing. We now briefly review existing results for limit continuous processes.

5.1.1 Brownian case

Recall that for $\mu=2$, jump processes converge at large n towards Brownian motion with diffusion coefficient $D=a_2^2$. In turn, in the continuous limit $x_0 \gg a_\mu$, the complete exit-time probability converges to the corresponding continuous first exit-time distribution $F^{(c)}_{\underline{0},\underline{x}}(n|x_0)$ of a Brownian motion, which can be exactly computed (see [Borodin & Salminen 1996] for instance) and is given by:

$$F^{(c)}_{\underline{0},\underline{x}}(n|x_0) = \frac{2D\pi}{x^2}\sum_{k=1}^{\infty} k\sin\left(\frac{k\pi x_0}{x}\right)\left[1+(-1)^{k+1}\right]e^{-\frac{Dk^2\pi^2 n}{x^2}}. \tag{5.2}$$

It is clearly seen from equation (5.2), that $F^{(c)}_{\underline{0},\underline{x}}(n|x_0)$ vanishes as $x_0\to 0$, in agreement with the fact that $F_{\underline{0},\underline{x}}(n|x_0)\sim F^{(c)}_{\underline{0},\underline{x}}(n|x_0)$ only for $x_0 \gg a_\mu$.

5.1.2 α-stable case

Let us now focus on the $\mu<2$ case, for which jump processes converge to an α-stable process of exponent μ. Denoting $G^{(c)}_{[0,x]}(u,t|x_0)$ the propagator of an α-stable process killed upon exiting the interval $[0,x]$, $G^{(c)}_{[0,x]}$ is known to satisfy the *fractional diffusion equation* [Zoia *et al.* 2007, Kwasnicki 2017]

$$\frac{\partial}{\partial t}G^{(c)}_{[0,x]}(u,t|x_0) = \left[\frac{\partial^2}{\partial^2 u}\right]^{\frac{\mu}{2}} G^{(c)}_{[0,x]}(u,t|x_0) \tag{5.3}$$

with absorbing boundary conditions at 0 and x. Denoting ψ_k and λ_k the eigenfunctions associated to the fractional Laplacian $\left[\frac{\partial^2}{\partial^2 u}\right]^{\frac{\mu}{2}}$ in the interval $[0,2]$[1] [Kwaśnicki 2012], the propagator is given by

$$G^{(c)}_{[0,x]}(u,t|x_0) = \frac{2}{x}\sum_{k=1}^{\infty}\psi_k\left(\frac{2x_0}{x}\right)\psi_k\left(\frac{2u}{x}\right)e^{-\lambda_k\left[\frac{2a_\mu}{x}\right]^\mu t}. \tag{5.4}$$

from which the first exit-time distribution is obtained as

$$\begin{aligned} F^{(c)}_{\underline{0},\underline{x}}(t|x_0) &= -\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^x G^{(c)}_{0,x}(u,t|x_0)\mathrm{d}u\right] \\ &= \sum_{k=1}^{\infty} C_k(x_0)\lambda_k\left[\frac{2a_\mu}{x}\right]^\mu e^{-\lambda_k\left[\frac{2a_\mu}{x}\right]^\mu t} \\ C_k(x_0) &= \psi_k\left(\frac{2x_0}{x}\right)\int_0^2 \psi_k(u)\mathrm{d}u \end{aligned} \tag{5.5}$$

Let us now make a few comments on the fractional diffusion eigenvalue problem.

- Equation (5.5) is formally exact. In the case $\mu = 2$, the ψ_k and λ_k are known exactly, and one recovers equation (5.2). However, for $\mu < 2$, no exact forms of ψ_k and λ_k are known.
- Controlled approximations of the eigenvalues and eigenfunctions have been constructed in [Kwaśnicki 2012], by combining eigenfunctions of the semi-infinite and infinite fractional diffusion problem. We provide in Appendix H a short overview of this construction, and refer the reader to the article for more details.
- In particular, it was shown in [Kwaśnicki 2012] that the eigenvalues of the limit α-stable process on the interval $[0,2]$ are approximately given by:

$$\lambda_k \simeq \left[\frac{k\pi}{2} - \frac{(2-\mu)\pi}{8}\right]^\mu. \tag{5.6}$$

- Additionally, we emphasize that the small x_0 behavior of the approximated eigenfunctions $\psi_k(x_0)$ is known and given by [Kwaśnicki 2011]

$$\psi_k(x_0) \underset{x_0\to 0}{\sim} \left[\sqrt{\frac{\mu}{2}}\Gamma\left(\frac{\mu}{2}\right)\right]^{-1}\left[\left(\frac{k\pi}{2} - \frac{(2-\mu)\pi}{8}\right)x_0\right]^{\frac{\mu}{2}}. \tag{5.7}$$

  Importantly, since $\psi_k(x_0) \propto x_0^{\mu/2}$ for small x_0, all first-passage observables defined for bounded or semi-infinite α-stable processes vanish as $x_0^{\mu/2}$. In particular, this is the case for the first exit-time distribution, $F^{(c)}_{\underline{0},\underline{x}}(n|x_0)$, the splitting probability $\pi^{(c)}_{\underline{0},\underline{x}}(x_0)$ and the semi-infinite FPT distribution through zero $F^{(c)}_{\underline{0}}(n|x_0)$.

[1] We here consider the interval $[0,2]$ to match closely the results presented in [Kwaśnicki 2012]. In turn, the eigenvalues and eigenfunctions for an arbitrary interval $[0,x]$ are given by $\lambda_k^{[0,x]} = \lambda_k\left[\frac{2a_\mu}{x}\right]^\mu$ and $\psi_k^{[0,x]}(x_0) = \psi_k\left(\frac{2x_0}{x}\right)\sqrt{\frac{2}{x}}$.

As in the Brownian case, in the continuous limit $x_0 \gg a_\mu$ and for large x and n, the complete exit-time probability of heavy-tailed jump processes converges towards its continuous counterpart $F_{\underline{0},\underline{x}}(n|x_0) \sim F^{(c)}_{\underline{0},\underline{x}}(n|x_0)$.

We are now left to deal with the $0 \leq x_0 \lesssim a_\mu$ regime, for which the limit process approach is not sufficient. To that end, partitioning trajectories on the exit direction (see figure 5.1), we introduce the leftward and rightward exit-time probabilities $F_{\underline{0},x}(n|x_0)$ and $F_{0,\underline{x}}(n|x_0)$ (LETP/RETP) such that:

$$F_{\underline{0},\underline{x}}(n|x_0) = F_{\underline{0},x}(n|x_0) + F_{0,\underline{x}}(n|x_0) \tag{5.8}$$

Note that the LETP has already been encountered in chapter 3 in the context of general continuous scale-invariant processes. Addressing each case separately, we exhibit universal asymptotic behavior for $F_{\underline{0},\underline{x}}(n|x_0)$.

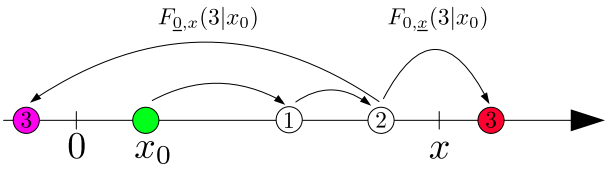


Figure 5.1: Two example trajectories contributing to $F_{\underline{0},x}(3|x_0)$ and $F_{0,\underline{x}}(3|x_0)$. After taking two steps inside the interval, the third step takes the walker outside, either left (purple) or right (red).

5.2 Rightward exit-time probability

5.2.1 General scaling form in the large n and x limit

We first consider the case of the RETP $F_{0,\underline{x}}(n|x_0)$, defined as the probability that the particle escapes the interval through x on the n^{th} step exactly. Importantly, we emphasize that the RETP is not a normalized quantity, and that

$$\sum_{n=1}^{\infty} F_{0,\underline{x}}(n|x_0) = \pi_{0,\underline{x}}(x_0). \tag{5.9}$$

Let us introduce the normalized conditional rightward exit-time distribution $h(x,n|x_0)$ defined by

$$h(x,n|x_0) = \frac{F_{0,\underline{x}}(n|x_0)}{\pi_{0,\underline{x}}(x_0)} \tag{5.10}$$

and proceed in three steps:

1. In the large x and n limit, based on results from [Caravenna & Chaumont 2008], $h(x,n|x_0)$ converges towards the first-passage time distribution through x of a limit continuous process conditioned to stay positive. In turn, for large x, this first-passage time distribution becomes independent of x_0: $h(x,n|x_0) \equiv h(x,n)$. In particular, $h(x,n)$ can be evaluated for large x_0 values, for which the continuous limit applies:

$$h(x,n) = \frac{F_{0,\underline{x}}(n|x_0)}{\pi_{0,\underline{x}}(x_0)} \underset{\substack{x\to\infty \\ n\to\infty}}{\sim} \lim_{u\to 0} \left[\frac{F^{(c)}_{0,\underline{x}}(n|u)}{\pi^{(c)}_{0,\underline{x}}(u)} \right]. \tag{5.11}$$

Importantly, the limit in the RHS ratio is well defined when $u \to 0$. Indeed both the RETP and splitting probability can be projected on the eigenfunction basis ψ_k, and hence have the same leading behavior as $u \to 0$.

2. Since the limit continuous processes conditioned to stay positive are scale-invariant (see for instance [Chaumont 1997]), the distribution of the first-passage time through x takes the following asymptotic scaling form

$$h(x,n) \underset{\substack{x\to\infty \\ n\to\infty \\ \tau \text{ fixed}}}{\sim} h^+(\tau) \left[\frac{a_\mu}{x} \right]^\mu, \tag{5.12}$$

where $\tau = n/n_x$ and $n_x = (x/a_\mu)^\mu$ is the typical number of steps needed to escape the interval. Note that the scaling function h^+ has already been encountered in equation (4.27) of chapter 4.

3. Finally, In the large n and x limit, the RETP is given by:

$$F_{0,\underline{x}}(n|x_0) \underset{\substack{x\to\infty \\ n\to\infty \\ \tau \text{ fixed}}}{\sim} \pi_{0,\underline{x}}(x_0) h_\mu(\tau) n^{-1}, \tag{5.13}$$

where $h_\mu(\tau) = h^+(\tau)\tau$ is a universal scaling function depending only on μ. Importantly, the specific x_0 dependence of $F_{0,\underline{x}}(n|x_0)$ arising from the discrete nature of jump processes is fully contained in the splitting probability $\pi_{0,\underline{x}}(x_0)$, which we extensively studied in the previous chapter. In particular, $F_{0,\underline{x}}(n|0)$ is clearly non-vanishing.

To characterize further the RETP, we derive explicit expressions of the universal function h_μ, for all $\mu \leq 2$.

5.2.2 Brownian case

In the case $\mu = 2$, similarly to $F^{(c)}_{\underline{0},\underline{x}}(n|x_0)$, $F^{(c)}_{0,\underline{x}}(n|x_0)$ can be computed exactly [Borodin & Salminen 1996] and is given by:

$$F^{(c)}_{0,\underline{x}}(n|x_0) = \frac{2D\pi}{x^2} \sum_{k=1}^{\infty} e^{-\frac{Dk^2\pi^2 n}{x^2}} k(-1)^{k+1} \sin\left(\frac{k\pi x_0}{x} \right). \tag{5.14}$$

Making use of equations (5.11) and (5.12), $h_2(\tau)$ is then obtained explicitly:

$$h_2(\tau) = 2\tau\pi^2 \sum_{k=1}^{\infty} k^2(-1)^{k+1} e^{-k^2\pi^2\tau}. \tag{5.15}$$

As a result, for jump processes with $\mu = 2$, we have completely characterized the large n and x asymptotic RETP, for all x_0 regimes. For the representative example of the exponential jump process with $p(\ell) = \frac{\gamma}{2} e^{-\gamma|\ell|}$, we display numerical agreement for $F_{0,\underline{x}}(n|0)$ in figure 5.2(a).

5.2.3 α-stable case

For $\mu < 2$, $F^{(c)}_{0,\underline{x}}(n|x_0)$ is not known explicitly, and $h_\mu(\tau)$ cannot be determined for all values of τ. However, its large and small τ behavior can be analyzed.

Large τ behavior. Recalling the definition of $n_x \equiv (x/a_\mu)^\mu$, we first consider the limit $n \gg n_x$. After a large number of steps, the jump process loses the memory of its starting point and delocalizes itself inside the interval. In turn, rightward and leftward exit events become equally likely, and the RETP converges towards *half* of the CETP:

$$F^{(c)}_{0,\underline{x}}(n|x_0) \underset{n \gg n_x}{\sim} \frac{F^{(c)}_{\underline{0},\underline{x}}(n|x_0)}{2}. \tag{5.16}$$

We emphasize that this argument is valid for all μ values, and can be explicitly checked in the case $\mu = 2$ by comparing equations (5.14) and (5.2). In turn, making use of equation (5.5) and evaluating the continuous splitting probability $\pi^{(c)}_{0,\underline{x}}(x_0)$ given in the previous chapter (equation (4.31)), we show that the leading τ behavior of $h_\mu(\tau)$ is given by a single decaying exponential, with an explicit prefactor:

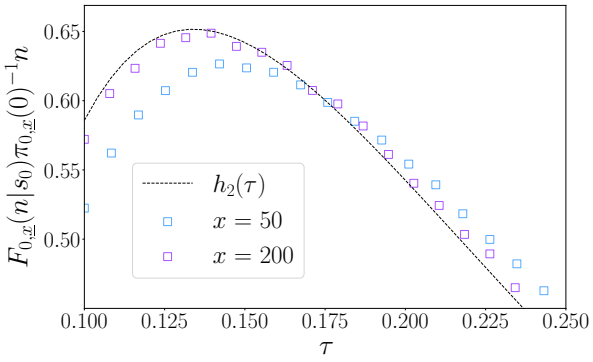
$$\begin{aligned} h_\mu(\tau) &\underset{\tau \gg 1}{\sim} \gamma_\mu \left[\lambda_1 2^\mu \tau\right] e^{-\lambda_1 2^\mu \tau} \\ \gamma_\mu &\equiv \sqrt{\mu} 2^{\frac{\mu}{2} - \frac{3}{2}} \frac{\Gamma\left(\frac{\mu}{2}\right)}{\Gamma(\mu)} \left[\frac{\pi}{2} - \frac{(2-\mu)\pi}{8}\right]^{\frac{\mu}{2}} \int_0^2 \psi_1(u) \mathrm{d}u. \end{aligned} \tag{5.17}$$

Of note, taking $\mu = 2$ in equation (5.17) yields the correct leading behavior of $h_2(\tau)$. Finally, for $\mu < 2$, we emphasize that while equation (5.17) is exact, neither λ_1 nor ψ_1 are known explicitly, and we revert to the the approximations constructed in [Kwaśnicki 2012] to compare our result to numerical simulations. In particular, we illustrate the large τ behavior of the RETP from 0 in figure 5.2(b).

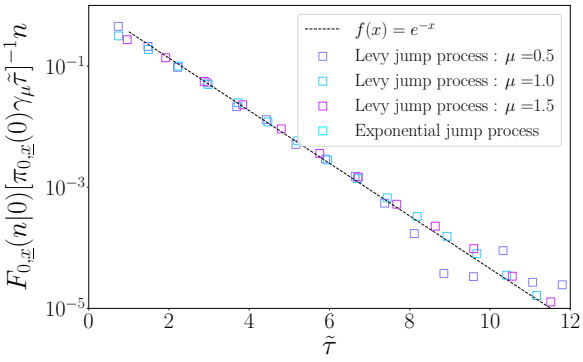
Small τ behavior. Consider now the opposite limit $n \ll n_x$, *ie* $\tau \ll 1$. While no simple link exists between $F^{(c)}_{0,\underline{x}}(n|x_0)$ and $F^{(c)}_{\underline{0},\underline{x}}(n|x_0)$ in this limit, the small τ behavior of $h_\mu(\tau)$ can be extracted by exploiting the power-law behavior of $p(\ell)$. Partitioning rightward exit events over the position of the last step yields

$$F_{0,\underline{x}}(n|x_0) = \int_0^x G_{[0,x]}(u, n-1|x_0) \left[\int_{x-u}^\infty p(\ell) \mathrm{d}\ell\right] \mathrm{d}u, \tag{5.18}$$

with $G_{[0,x]}(u, n-1|x_0)$ the *bounded* propagator. For the sake of simplicity, we consider hereafter that $x_0 = 0$. Fixing both n and u, and taking $x \to \infty$, we have $G_{[0,x]}(u, n-1|0) \sim G(u, n-1|0)$ where we recall that $G(u, n-1|0)$ is the semi infinite propagator, which is known exactly (see



(a) Large x and n asymptotic RETP for an exponential jump process with $p(\ell) \propto e^{-\gamma|\ell|}$. As x increases, the RETP rescaled according to (5.13) converges towards the universal $h_2(\tau)$ function (5.15).



(b) Large τ behavior of the rescaled RETP for Levy jump processes with $\tilde{p}(k) = e^{-|k|^\mu}$ and exponential jump process. With $\tilde{\tau} = \lambda_1 2^\mu \tau$, all curves collapse on a single exponential predicted by (5.17).

Figure 5.2

chapter 4). In turn, the asymptotic RETP is given by the following convolution

$$F_{0,\underline{x}}(n|0) \underset{x\to\infty}{\sim} \int_0^x G(u, n-1|0)U(x-u)\mathrm{d}u, \tag{5.19}$$

where $U(x) = \int_x^\infty p(\ell)\mathrm{d}\ell$ is the cumulative distribution of $p(\ell)$. By introducing the generating function and Laplace transform, the asymptotic RETP finally reads:

$$\sum_{n=1}^{\infty} \xi^n \left[\int_0^\infty e^{-sx} F_{0,\underline{x}}(n|0)\mathrm{d}x\right] \underset{s\to 0}{\sim} \xi\ G(s,\xi|0)U(s). \tag{5.20}$$

Both $G(s,\xi|0)$ and $U(s)$ can then be analyzed in the $s \to 0$ limit to extract the leading large x behavior of $F_{0,\underline{x}}(n|0)$. Recalling that in the small k limit, one has $\tilde{p}(k) = 1 - (a_\mu|k|)^\mu + o(|k|)^\mu$, we obtain a comprehensive list of small s behaviors:

$$0 < \mu < 1 : \begin{cases} G(s,\xi|0) = \dfrac{1}{\sqrt{1-\xi}}\left[1 - \dfrac{\xi}{1-\xi} c_\mu (a_\mu s)^\mu + o(s^\mu)\right] \\ U(s) = c_\mu\ (a_\mu)^\mu\ s^{\mu-1} + o(s^{\mu-1}) \end{cases}$$

$$\mu = 1 : \quad \begin{cases} G(s,\xi|0) = \dfrac{1}{\sqrt{1-\xi}}\left[1 + \left(\dfrac{\xi}{1-\xi}\right)\dfrac{a_1}{\pi} s\log(s) + O(s)\right] \\ U(s) = -\dfrac{a_1}{\pi}\log(s) + O(1) \end{cases} \tag{5.21}$$

$$1 < \mu < 2 : \begin{cases} G(s,\xi|0) = \dfrac{1}{\sqrt{1-\xi}}\left[1 + \left(\dfrac{\xi}{1-\xi}\right)^{\frac{1}{\mu}} \dfrac{a_\mu\ s}{\sin(\pi/\mu)} - \dfrac{\xi}{1-\xi} c_\mu (a_\mu s)^\mu + o(s^\mu)\right] \\ U(s) = \langle p\rangle + c_\mu\ (a_\mu)^\mu\ s^{\mu-1} + o(s^{\mu-1}) \end{cases}$$

where $c_\mu = \sec(\frac{\pi\mu}{2})/2$ and $\langle p \rangle = \int_0^\infty p(\ell)\mathrm{d}\ell$. To leading order in $s \to 0$, and upon Laplace inversion, we finally obtain the following exact asymptotic form of $F_{0,\underline{x}}(n|0)$ valid for all values of $\mu < 2$:

$$F_{0,\underline{x}}(n|0) \underset{x\to\infty}{\sim} q(0,n-1)\frac{\Gamma(\mu)}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^\mu. \tag{5.22}$$

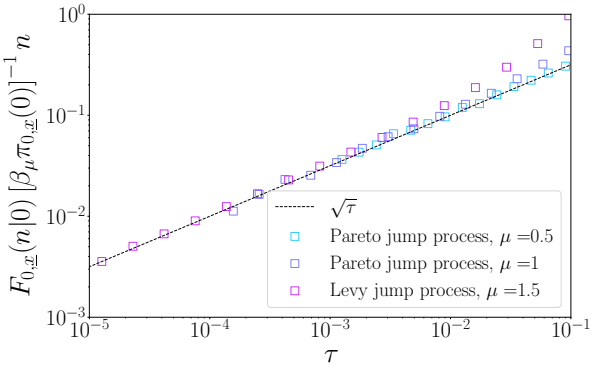
where we recall that $q(0,n)$ is the survival probability in the semi-infinite geometry, starting from $x_0 = 0$. Importantly, equation (5.22) has a clear physical interpretation. For a heavy-tailed jump process, the probability to stay in a large interval for $n-1$ steps before exiting rightwards is simply given by the $n-1$ step survival probability on the semi-infinite line, times the probability to perform a single jump large enough to cross the full interval. Of note, this *single big jump* interpretation has also been used to evaluate other observables for general unbounded Levy Flights [Foss S. 2015, Vezzani *et al.* 2019, Vezzani *et al.* 2020].

Finally, making use of the large n asymptotic behavior (4.14) of $q(0,n)$, we derive the universal small τ behavior of $h_\mu(\tau)$

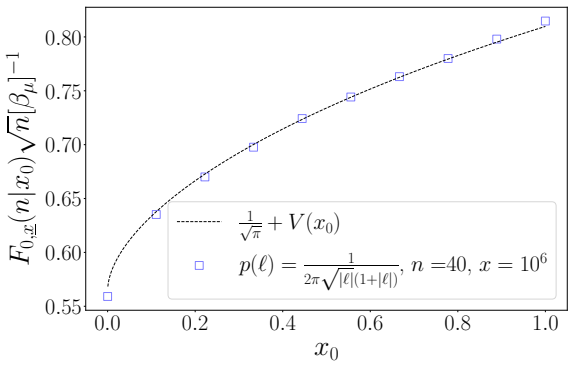
$$h_\mu(\tau) \underset{\tau\ll 1}{\sim} \Gamma(\mu/2)\sin(\pi\mu/2)\pi^{-\frac{3}{2}}\sqrt{\tau}, \tag{5.23}$$

and conclude with a few remarks:

- The small τ behavior predicted by equation (5.23) is displayed in figure 5.3(a), where we numerically estimate $F_{0,\underline{x}}(n|0)$ for a variety of heavy-tailed jump processes.
- As announced, the small τ behavior of h_μ holds independently of the initial condition x_0. In turn, we illustrate the agreement between simulations and $F_{0,\underline{x}}(n|x_0)$ for various initial positions x_0 in figure 5.3(b).
- As $\mu \to 2$, the right-hand-side of equation (5.23) vanishes, in agreement with the stronger small τ decay of $h_2(\tau)$ given by equation (5.15).



(a) Small τ behavior of the RETP for Pareto jump processes with $p(\ell) \propto \mathbb{1}_{|\ell>1|}|\ell|^{-(1+\mu)}$ and Levy jump process with $\tilde{p}(k) = e^{-|k|^\mu}$. Properly rescaled according to equations (5.13) and (5.23), all curves collapse to a single function. Here $\beta_\mu = \Gamma(\mu/2)\sin(\pi\mu/2)\pi^{-\frac{3}{2}}$.



(b) x_0 dependence of the RETP for a F-distributed jump process. In the regime $n \ll n_x$, the x_0 dependence is fully contained in the survival probability $q(x_0,n)$ (4.14). The rescaled RETP according to (5.22) and the numerically inverted $\frac{1}{\sqrt{\pi}} + V(x_0)$ collapse perfectly.

Figure 5.3

5.3 Leftward exit-time probability

Having extensively discussed the RETP, we now focus on the leftward exit-time probability $F_{\underline{0},x}(n|x_0)$ (LETP) for which we also derive universal asymptotic results in the large n and x limit, valid in all x_0 regimes.

5.3.1 General scaling form in the large n and x limit

We emphasize that the argument developed here closely follows section 3.2.1 of chapter 3, in which we derived an asymptotic scaling form of $F^{(c)}_{\underline{0},x}(t|x_0)$ for general continuous scale-invariant processes. Indeed, for large n and x, the limit continuous processes are scale-invariant, and we can directly use equation (3.17) for jump processes. Denoting $\tau = n/n_x$ with $n_x \equiv (x/a_\mu)^\mu$, we obtain the general scaling form

$$F_{\underline{0},x}(n|x_0) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} F_{\underline{0}}(n|x_0) g_\mu(\tau). \tag{5.24}$$

where g_μ is a universal x_0 and a_μ independent function which depends only on μ. Importantly, the complete x_0 dependence is contained the first-passage time distribution through 0 in the semi infinite geometry: $F_{\underline{0}}(n|x_0) = q(x_0, n-1) - q(x_0, n)$. In turn, $F_{\underline{0},x}(n|0)$ is clearly non-vanishing. Finally, we stress that equation (5.24) is the direct analog of equation (5.13) for the RETP, and proceed to characterize further the g_μ function.

5.3.2 Brownian case

Since $g_\mu(\tau)$ is independent of x_0, we rewrite equation (5.24) in the continuous limit $x_0 \gg a_\mu$, where both the LETP and FPT through 0 are asymptotically given by their continuous counterparts:

$$g_\mu(\tau) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} \lim_{u\to 0} \left[\frac{F^{(c)}_{\underline{0},x}(n|u)}{F^{(c)}_{\underline{0}}(n|u)} \right] \tag{5.25}$$

In turn, in the case $\mu = 2$, both continuous LETP and FPT distribution are known [Borodin & Salminen 1996] and given by

$$\begin{aligned} F^{(c)}_{\underline{0},x}(n|u) &= \frac{2D\pi}{x^2} \sum_{k=1}^{\infty} e^{-\frac{Dk^2\pi^2 n}{x^2}} k(-1)^{k+1} \sin\left(\frac{k\pi(x-u)}{x}\right) \\ F^{(c)}_{\underline{0}}(n|u) &= \frac{u}{\sqrt{4\pi D t n^3}} e^{-\frac{u^2}{4Dn}}, \end{aligned} \tag{5.26}$$

with $D = a_2^2$. In turn, we obtain an explicit expression for $g_2(\tau)$

$$g_2(\tau) = 4\pi^{\frac{5}{2}} \tau^{\frac{3}{2}} \sum_{k=1}^{\infty} e^{-k^2\pi^2\tau} k^2. \tag{5.27}$$

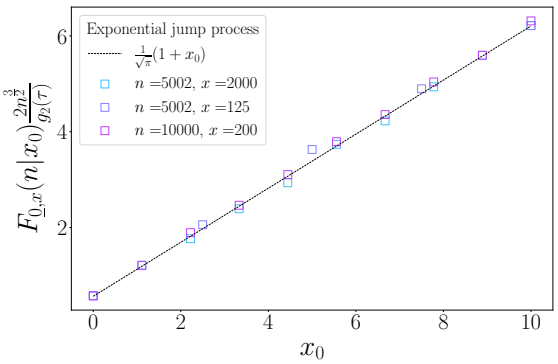
Taken together, equations (5.27) and (5.24) completely characterize the asymptotic LETP for general jump processes with $\mu = 2$. As an illustration, we provide numerical evidence of the x_0 behavior of $F_{\underline{0},x}(n|x_0)$ for an exponential jump process in figure 5.4(a).

5.3.3 α-stable case

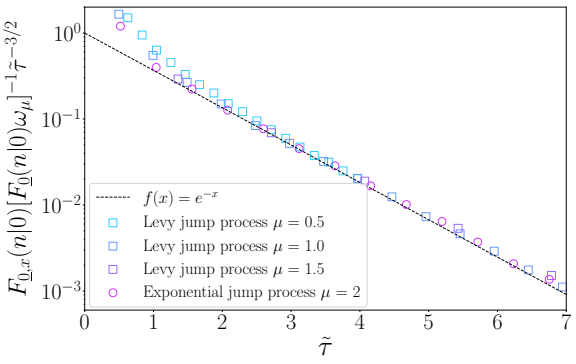
For jump processes with $\mu < 2$, no explicit expression of g_μ can be obtained for general τ. However, following the RETP treatment, we compute the small and large τ asymptotic behavior of $g_\mu(\tau)$:

$$\begin{aligned} g_\mu(\tau) &\underset{\tau \ll 1}{\sim} 1, \\ g_\mu(\tau) &\underset{\tau \gg 1}{\sim} \omega_\mu \left[\lambda_1 2^\mu \tau\right]^{3/2} e^{-\lambda_1 2^\mu \tau}, \\ \omega_\mu &\equiv \sqrt{\frac{2\pi}{\mu}} \frac{\Gamma\left(1+\frac{\mu}{2}\right)}{\Gamma\left(\frac{\mu}{2}\right)} \int_0^2 \psi_1(u) \mathrm{d}u. \end{aligned} \tag{5.28}$$

As a concluding remark, we point out that in both $\tau \gg 1$ and $\tau \ll 1$ regimes, taking $\mu \to 2$ in (5.28) yields the correct asymptotic behavior of $g_2(\tau)$ extracted from equation (5.27), and we illustrate the large τ behavior in figure 5.4(b), for both heavy-tailed and narrow-tailed jump processes.



(a) x_0 dependency of the LETP for exponential jump processes with $p(\ell) \propto e^{-|\ell|}$ and fixed n and x. For the exponential jump process, $F_{\underline{0}}(n|x_0) \underset{n\to\infty}{\sim} n^{-3/2} \frac{1}{2\sqrt{\pi}} [1+x_0]$ (4.14). Correctly rescaled by $g_2(\tau)$ (5.24), the LETP displays the same affine x_0 behavior.



(b) Large τ behavior of the rescaled LETP for Levy jump processes with $\tilde{p}(k) = e^{-|k|^\mu}$ and exponential jump process. With $\tilde{\tau} = \lambda_1 2^\mu \tau$, all curves collapse on a single exponential predicted by (5.28). Note that the dashed line does not contain any fitting parameter.

Figure 5.4

5.4 Conclusion

Going beyond the geometrical information contained in the splitting probability $\pi_{0,\underline{x}}(x_0)$, the complete exit-time probability $F_{\underline{0},\underline{x}}(n|x_0)$ quantifies the dynamical aspects of interval exit for bounded jump processes. While in the *continuous limit* $x_0 \gg a_\mu$, the CETP converges towards the continuous exit-time distribution $F^{(c)}_{\underline{0},\underline{x}}(n|x_0)$ of the associated continuous limit process, this approach is not relevant to address the $x_0 \lesssim a_\mu$ regime.

To fully characterize the CETP in all x_0 regimes, we find it necessary to introduce the rightward and leftward exit-time probabilities $F_{0,\underline{x}}(n|x_0)$ and $F_{\underline{0},x}(n|x_0)$, such that $F_{\underline{0},\underline{x}}(n|x_0) = F_{0,\underline{x}}(n|x_0) + F_{\underline{0},x}(n|x_0)$. In the large n and x limit, defining $n_x = (x/a_\mu)^\mu$ and keeping $\tau = n/n_x$ fixed, we uncover universal scaling behavior for both quantities, summarized in table 5.1.

	$F_{0,\underline{x}}(n\|x_0) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} \pi_{0,\underline{x}}(x_0) h_\mu(\tau) n^{-1}$	$F_{\underline{0},x}(n\|x_0) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} F_{\underline{0}}(n\|x_0) g_\mu(\tau)$
$\mu = 2$	$h_2(\tau) = 2\tau\pi^2 \sum_{k=1}^{\infty} k^2 (-1)^{k+1} e^{-k^2\pi^2\tau}$	$g_2(\tau) = 4\pi^{\frac{5}{2}} \tau^{\frac{3}{2}} \sum_{k=1}^{\infty} e^{-k^2\pi^2\tau} k^2$
$\mu < 2$	$h_\mu(\tau) \underset{\tau\ll 1}{\sim} \Gamma(\mu/2) \sin(\pi\mu/2) \pi^{-\frac{3}{2}} \sqrt{\tau}$	$g_\mu(\tau) \underset{\tau\ll 1}{\sim} 1$
	$h_\mu(\tau) \underset{\tau\gg 1}{\sim} \gamma_\mu \left[\lambda_1 2^\mu \tau\right] e^{-\lambda_1 2^\mu \tau}$	$g_\mu(\tau) \underset{\tau\gg 1}{\sim} \omega_\mu \left[\lambda_1 2^\mu \tau\right]^{3/2} e^{-\lambda_1 2^\mu \tau}$

Table 5.1

Importantly, the small x_0 dependence of the RETP and LETP - and in particular the non-vanishing values for the initial position $x_0 = 0$ - only appears through the splitting probability $\pi_{0,\underline{x}}(x_0)$ and the semi infinite FPT distribution $F_{\underline{0}}(n|x_0)$, which we extensively discussed in chapter 4. As a side product, we obtain the sought-after asymptotic behavior of the CETP, valid for all x_0 regimes.

Recalling equation (3.11) from chapter 3, we stress that the RETP and LETP appear as essential tools to investigate *extreme value statistics*. In particular, the joint distribution $\sigma(x, n|x_0)$ of the maximum x and FPT n through 0 straightforwardly adapts to jump processes

$$\sigma(x, n|x_0) = \frac{\mathrm{d}}{\mathrm{d}x} F_{\underline{0},x}(n|x_0), \tag{5.29}$$

with the $x_0 \lesssim a_\mu$ regime correctly captured by $F_{\underline{0},x}(n|x_0)$. In turn, we focus in the next chapter on deriving essential building blocks to address general *extreme value* problems for jump processes.

Chapter 6

Extreme value statistics of jump processes

Contents

In a broad sense, *extreme value* problems focus on the determination of the statistics of the extremums of a set of random variables $(x_1, \dots x_n)$. Studied early on for sets of independent random variables [Fréchet 1927, Gumbel 1935], extreme value statistics (EVS) have led to fruitful applications across statistical physics [Bouchaud & Mézard 1997, Biroli *et al.* 2007]. The case of extreme value statistics of random walks has been studied in depth for continuous time stochastic processes [Chung 1976, Randon-Furling & Majumdar 2007, Majumdar *et al.* 2008]. However, with the exception of the first moments of the running maximum [Spitzer 1956, de Bruyne *et al.* 2021], extreme value statistics of jump processes have yet to be investigated thoroughly.

We first review existing results regarding EVS of Markovian continuous time stochastic processes, and analyze the technical tools involved in their computation, which we broadly refer to as the *Markovian path decomposition* technique. In turn, we show how this technique is straightforwardly adapted to deal with EVS of jump processes. In particular, we emphasize that the survival probability and semi-infinite propagator constitute essential building blocks to determine the distributions of the running maximum and time at which it is reached for unbounded jump processes.

For jump processes stopped upon crossing 0 for the first time (which we refer to as semi-infinite jump processes), we introduce an additional essential building block to address extremums distributions, the *strip probability* $\mu_{0,\underline{x}}(n)$, defined as the probability that the walker starting from 0 reaches its maximum x on its n^{th} step exactly, without crossing 0. We first provide an exact characterization of $\mu_{0,\underline{x}}(n)$ in terms of the LETP and RETP, valid for general jump processes. Next, defining the typical number of steps $n_x \equiv (x/a_\mu)^\mu$ needed to reach x, we analyze the asymptotic behavior of the strip probability in the $1 \ll n \ll n_x$ regime for processes with $\mu < 2$, and in the scaling regime n/n_x fixed for processes with $\mu = 2$. In turn, we illustrate the range of applicability of the strip probability $\mu_{0,\underline{x}}(n)$ by investigating various extreme value observables of semi-infinite jump processes, as well as their *span*. Finally, we make use of $\mu_{0,\underline{x}}(n)$ to provide a refined description of the LETP in the $1 \ll n \ll n_x$ limit.

6.1 Extreme value statistics of unbounded jump processes

6.1.1 EVS of continuous space and time stochastic processes

The running maximum $M(t)$ is the most natural extreme value observable for a continuous stochastic process and, as such, has been extensively studied since the introduction of both Brownian motion and α-stable processes. Its cumulative distribution is easily shown to be given by

$$P(M(t) \leq x|x_0) = q^{(c)}(x - x_0, t), \tag{6.1}$$

where we recall that $q^{(c)}(x_0, t)$ is the continuous survival probability up to time t for a process issued from x_0 and killed upon crossing 0 for the first time. In the case of Brownian motion, the survival probability has been derived early on (see [Levy 1937] for instance) such that the full distribution of $M(t)$ is known. For general α-stable processes however, the derivation of $q^{(c)}(x_0, t)$ is much harder, and explicit results for the distribution of $M(t)$ are only available in specific cases [Darling 1956, Doney 1987, Kuznetsov 2011].

We emphasize that many other important EVS results have been obtained in the context of Brownian motion. For instance, the distribution of the time t_m at which $M(t)$ is reached, known as the arcsine-law, has been obtained by Levy [Levy 1937]

$$P(t_m = u|t) = \frac{1}{\pi\sqrt{u(t-u)}}. \tag{6.2}$$

Additionally, a variety of joint distributions of extremums and related times have been derived for a Brownian motion B_t. In particular, the joint distribution of the time t_m and running maximum $M(t)$ has been investigated in [Majumdar *et al.* 2008, Mori *et al.* 2021] for free Brownian motion and constrained versions such as Brownian bridges, as well as general stationary processes; the distribution of the time at which the maximum is reached before the FPT to 0 has been studied in [Randon-Furling & Majumdar 2007]; the distribution of the time between the running maximum $M(t)$ and minimum $m(t)$ has been derived in [Mori *et al.* 2020b] and, very recently, the joint distribution of the time T at which $M(t)$ first crosses a fixed level m and the position B_T has been investigated in [Randon-Furling *et al.* 2022].

In most of these cases, the computations rely on the Markovian nature of Brownian motion, and the fact that Brownian trajectories can be decomposed into statistically independent parts, which we broadly refer to as the *Markovian path decomposition* technique. As an example, we consider the joint distribution $\rho_1^{(c)}(x, t_m|t, x_0 = 0)$ of the running maximum $M(t)$ and the time t_m at which it is reached, for an unbounded Brownian motion issued from $x_0 = 0$. By decomposing the trajectory around t_m, and making use of the symmetry of the trajectory, it was shown in [Majumdar *et al.* 2008] that

$$\rho_1^{(c)}(x, t_m|t, x_0 = 0) = \lim_{\varepsilon \to 0} \mathcal{N}(\varepsilon) G^{(c)}(x, t_m|\varepsilon) q^{(c)}(\varepsilon, t - t_m), \tag{6.3}$$

where $G^{(c)}(x, t_m|\varepsilon)$ is the propagator of the Brownian motion issued from $x_0 = \varepsilon$ and killed upon reaching 0 for the first time, and $\mathcal{N}(\varepsilon)$ is a normalization factor so that

$$\int_0^\infty \int_0^t \rho_1^{(c)}(x, t_m|t, x_0 = 0) \mathrm{d}x \mathrm{d}t_m = 1. \tag{6.4}$$

Equation (6.3) deserves a few important comments:

- The introduction of the ε cut-off parameter is necessary to ensure that the weights $G^{(c)}(x, t_m|\varepsilon)$ and $q^{(c)}(\varepsilon, t - t_m)$ of the independent parts of the trajectory do not vanish. Indeed, for Brownian motion, $q^{(c)}(x_0 = 0, t) = 0$ by definition.
- The determination of $\rho_1^{(c)}(x, t_m|t, x_0 = 0)$ necessarily implies the exact derivation of the normalization factor $\mathcal{N}(\varepsilon)$ and may, at times, require tedious computations.

Importantly, we emphasize that these two difficulties are intrinsically linked to the continuous time nature of Brownian motion, and arise in most extreme value computations involving a Markovian decomposition of the trajectory. However, in the case of general discrete time jump processes, we circumvent these difficulties, and show that EVS can be directly obtained via the Markovian path decomposition technique.

6.1.2 First results for jump processes

Running maximum. Following the previous exposition for continuous processes, we first consider the distribution $\mu(x|n, x_0)$ of the running maximum x reached by an n step long general jump process. It is clear that equation (6.1) is easily adapted in the case of symmetric continuous jump processes, such that

$$\mu(x|n, x_0) = \frac{\mathrm{d}}{\mathrm{d}x} q(x - x_0, n), \tag{6.5}$$

where we recall that $q(x - x_0, n)$ is the survival probability in the semi-infinite geometry, and whose expression for a general jump process is given in equation (4.11) of chapter 4 and reminded here

$$\sum_{n=0}^{\infty} \xi^n \left[\int_0^\infty e^{-sx_0} q(x_0, n) \mathrm{d}x_0 \right] = \frac{1}{s\sqrt{1-\xi}} \exp\left[-\frac{s}{2\pi} \int_{-\infty}^{\infty} \frac{\log\left[1 - \xi \tilde{p}(k)\right]}{s^2 + k^2} \mathrm{d}k \right]. \tag{6.6}$$

We emphasize that while equation (6.5) specifically depends on the details of the jump distribution $p(\ell)$, it clearly captures the discrete nature of the jump process. In particular, $\mu(x|n,x_0)$ is non vanishing for $x = x_0$, in contrast to its continuous counterpart (6.1).

Importantly, we stress that the first moment of the maximum has been studied for jump processes with $\mu = 2$ [Comtet & Majumdar 2005] and $\mu < 2$ [de Bruyne *et al.* 2021] in the large n limit. However, the complete distribution $\mu(x|n,x_0)$ remains to be investigated. Focusing on the asymptotic regime $n \to \infty$ and $x \to \infty$, we obtain the leading behavior of $\mu(x|n,x_0)$ for general jump processes. Throughout this chapter, we take $x_0 = 0$ without loss of generality.

- We first focus on the case $\mu = 2$. In the limit $n \to \infty$, the jump process converges towards a Brownian motion with $D = a_2^2$. In turn, in the scaling regime $n \to \infty$, $x \to \infty$ and $\tau = n/(x/a_2)^2$ fixed, the distribution of the maximum converges towards the distribution of the running maximum of the Brownian motion [Redner 2001]:

$$\mu(x|n,x_0=0) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} \frac{1}{a_2\sqrt{\pi n}} e^{-\left[\frac{x}{a_2}\right]^2 \frac{1}{4n}}. \tag{6.7}$$

- For heavy-tailed jump processes with $\mu < 2$, the first large x vanishing order of the survival probability (6.6) reads

$$1 - q(x,n) \underset{1\ll n\ll n_x}{\sim} \frac{n}{\pi} \sin\left(\frac{\pi\mu}{2}\right) \Gamma(\mu) \left[\frac{a_\mu}{x}\right]^\mu, \tag{6.8}$$

  where the notation $n_x \equiv (x/a_\mu)^\mu$ will now be used extensively. In turn, we obtain the asymptotic behavior of the distribution of the maximum $\mu(x|n,0)$:

$$\mu(x|n,x_0=0) \underset{1\ll n\ll n_x}{\sim} \frac{\mu n}{\pi} \sin\left(\frac{\pi\mu}{2}\right) \Gamma(\mu) \left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}, \tag{6.9}$$

  valid for all $\mu < 2$. Note that the linear n dependence of the distribution of the maximum can be interpreted in terms of the single big jump principle [Vezzani *et al.* 2019]. Indeed, for the running maximum to be equal to x, the walker has exactly n chances to perform one big jump bringing him close to x.

Time at which the running maximum is reached. We next focus on the distribution $\rho(n_m|n,x_0=0)$ of the time n_m at which the running maximum is reached. While the results presented in this paragraph are not new, and can be found in *eg* [Majumdar *et al.* 2020], we take the computation of $\rho(n_m|n,x_0=0)$ as an opportunity to showcase the ease of use of the Markovian path decomposition technique for jump processes.

Consider a trajectory contributing to $\rho(n_m|n,x_0=0)$ illustrated in figure 6.1(a), and decompose the trajectory in two parts around n_m. Since the jump process is Markovian, the two parts can be evaluated independently. In turn, by making use of the symmetry of the jump process, the weights of the left and right parts are respectively given by the survival probabilities $q(0,n_m)$ and $q(0,n-n_m)$. Explicit expressions for the survival probabilities from 0 are obtained

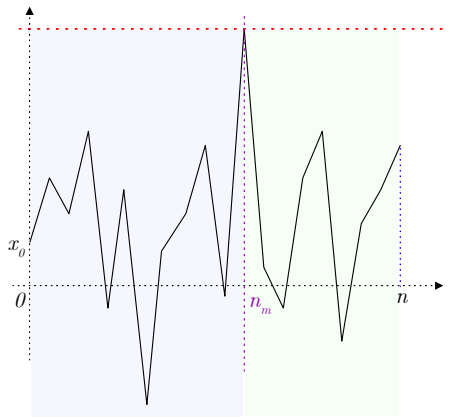
from the Sparre Andersen theorem (see equation (4.13) of chapter 4), and the distribution of n_m is finally given by

$$\begin{aligned}\rho(n_m|n, x_0 = 0) &= q(0, nm)q(0, n - n_m) \\ &= \binom{2n_m}{n_m}\frac{1}{2^{2n_m}}\binom{2(n-n_m)}{n-n_m}\frac{1}{2^{2(n-n_m)}}.\end{aligned} \tag{6.10}$$

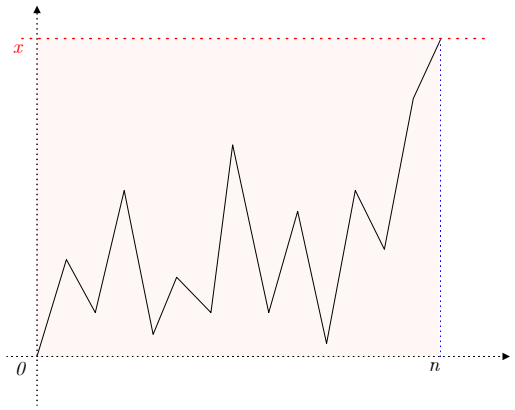
Note that $\rho(n_m|n, x_0 = 0)$ is strikingly independent of the jump distribution $p(\ell)$, whether it be heavy-tailed or not for instance. This important result stems from the underlying combinatorial properties of jump processes, which were already illustrated in the sole Sparre Andersen theorem. In the large n_m and n limit, with $n \gg n_m \gg 1$, the distribution of n_m asymptotically reads

$$\rho(n_m|n, x_0 = 0) \underset{n\gg n_m\gg 1}{\sim} \frac{1}{\pi}\frac{1}{\sqrt{n_m(n-n_m)}}, \tag{6.11}$$

which can be seen as a discrete version of the continuous arcsine-law (6.2). As a concluding remark, we emphasize that in contrast to continuous time processes such as Brownian motion, no difficulties arise from decomposing a trajectory into various independent parts. As a result, we now show that the Markovian path decomposition is an essential technical tool to derive a variety of EVS observables for general jump processes.



(a) Example trajectory of an unbounded jump process with maximum reached at step n_m. The blue and green parts are statistically independent, and with respective probabilistic weight $q(0, n_m)$ and $q(0, n - n_m)$.



(b) Example trajectory contributing to $\mu_{0,\underline{x}}(n)$. In exactly n steps the jump process crosses the red strip of width x.

Figure 6.1

6.1.3 Joint distributions for infinite jump processes from Markovian path decomposition

We illustrate the range of applicability of the previous decomposition by providing new and explicit expressions for two important joint distributions belonging to the EVS of infinite jump processes. Without loss of generality we hereafter choose $x_0 = 0$.

Maximum and time of maximum for free jump processes. Consider first the joint distribution $\rho_1(x, n_m|n, 0)$ of the maximum x and time at which it is reached n_m, for an n step long jump process starting from 0. Recall that the semi-infinite propagator $G_0(x, n)$, defined as the probability that the particle issued from 0 stays positive during n steps and is at x on the n^{th} step, is given for all $\tilde{p}(k)$ by

$$\sum_{n=0}^{\infty} \xi^n \left[\int_0^{\infty} e^{-sx} G_0(x, n)\mathrm{d}x\right] = \exp\left[-\frac{s}{2\pi}\int_{-\infty}^{\infty} \frac{\log\left[1-\xi\tilde{p}(k)\right]}{s^2+k^2}\mathrm{d}k\right]. \tag{6.12}$$

By decomposing a trajectory contributing to $\rho_1(x, n_m|n, 0)$ into two parts around n_m, and making use of the symmetry of the process, the joint distribution reads

$$\rho_1(x, n_m|n, 0) = G_0(x, n_m)q(0, n - n_m). \tag{6.13}$$

While equation (6.14) is exact, the joint distribution of x and n_m depends explicitly on the full jump distribution $\tilde{p}(k)$. In particular, in the specific case of the exponential jump process with $p(\ell) = \frac{\gamma}{2}e^{-\gamma|\ell|}$, $G_0(x, n)$ can be computed exactly(see [Majumdar *et al.* 2017] for instance) and reads in the generating function formalism:

$$\sum_{n=0}^{\infty} G_0(x, n)\xi^n = \gamma e^{-\sqrt{1-\xi}\gamma x}(1 - \sqrt{1-\xi}), \tag{6.14}$$

so that $\rho_1^{\text{exp}}(x, n_m|n, 0)$ can be evaluated for arbitrary n, n_m and x. As an illustration, we numerically evaluate the cumulative distribution of $\rho_1^{\text{exp}}(x, n_m|n, 0)$ for fixed n and n_m, and compare with our result in figure 6.2(a). However, for most jump processes the semi-infinite propagator $G_0(x, n)$ cannot be written explicitly. In turn, we derive universal asymptotic results in the $n \gg n_m \gg 1$ and $x \to \infty$, valid for general jump processes.

- Again, we first consider the case jump processes with $\mu = 2$. In the $n \gg n_m \gg 1$ limit, the jump process converges toward a Brownian motion, and the joint distribution ρ_1 converges towards the joint distribution of the running maximum and time at which it is reached of a Brownian motion, given by [Majumdar *et al.* 2008]:

$$\rho_1(x, n_m|n, 0) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed} \\ n_m \gg n}}{\sim} \frac{x}{a_2}\frac{1}{2a_2\pi n_m^{3/2}\sqrt{n-n_m}} e^{-\left[\frac{x}{a_2}\right]^2 \frac{1}{4n_m}}. \tag{6.15}$$

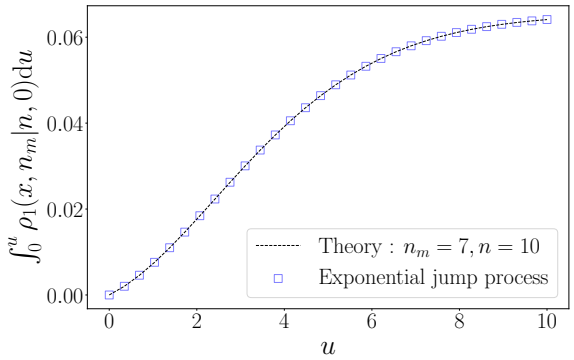
- For heavy-tailed processes with $\mu < 2$ we perform the asymptotic analysis of $G_0(x, n)$ in the $n \gg 1$ and $x \to \infty$ and obtain the following leading x behavior:

$$G_0(x, n) \underset{1\ll n\ll n_x}{\sim} \sqrt{\frac{n}{\pi}}\frac{2\mu}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\Gamma(\mu)\left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}. \tag{6.16}$$

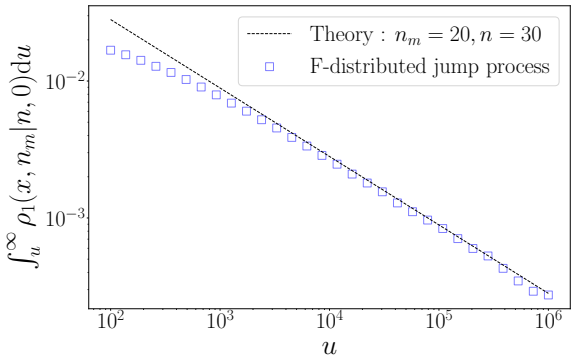
  In turn, in the limit $n \gg n_m$, the joint distribution ρ_1 takes the following universal asymptotic form:

$$\rho_1(x, n_m|n, 0) \underset{n_x \gg n \gg n_m \gg 1}{\sim} \frac{1}{\pi}\sqrt{\frac{n_m}{n-n_m}}\frac{2\mu}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\Gamma(\mu)\left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}, \tag{6.17}$$

valid for all values of $\mu < 2$. Finally, we illustrate the asymptotic result (6.17) in figure 6.2(b), where we numerically evaluate the cumulative distribution of ρ_1 for a a jump process with $p(\ell) \propto (\sqrt{|\ell|}(1+|\ell|))^{-1}$ and a given set of parameters $n = 20$ and $n_m = 30$. We emphasize the quick convergence of equation (6.17) even for small n and n_m values.



(a) ρ_1 cumulative distribution for an exponential jump process with $p(\ell) \propto e^{-|\ell|}$. In this case, ρ_1 is known exactly from equation (6.14).



(b) ρ_1 cumulative distribution for a F-distributed jump process with $p(\ell) \propto (\sqrt{|\ell|}(1+|\ell|))^{-1}$ and corresponding $\mu = 1/2$. The dashed line is obtained from (6.17), with no fit parameter. As x increases, the $1 \ll n_m \ll n_x$ regime is reached, and both theory and simulations converge.

Figure 6.2

Trivariate distribution. As a second illustration of the Markovian path decomposition technique to evaluate EVS observables for infinite jump processes, we consider the joint distribution $\rho_2(x, n_m, x_f|n, 0)$ of the maximum x, the time at which it is reached n_m and the final position x_f for a n step long jump process. The derivation of ρ_2 is straightforward: by decomposing the trajectory around n_m and making use of the symmetry and translational invariance of the jump process, we obtain

$$\rho_2(x, n_m, x_f|n, 0) = G_0(x, n_m)G_0(x - x_f, n - n_m). \tag{6.18}$$

We emphasize that the weight of each independent part of the trajectory is well-defined, showcasing the ease of use of the Markovian path decomposition.

Similarly to the previous joint distribution, we derive universal asymptotic behavior in the large n_m, x, $n-n_m$ and $x-x_f$ regime. First, for jump processes with $\mu = 2$, the joint distribution in the scaling regime $\tau_1 = n_m/(x/a_2)^2$ and $\tau_2 = (n-n_m)/((x-x_f)/a_2)^2$ fixed is simply given by the corresponding distribution for the standard Brownian motion [Borodin & Salminen 1996]:

$$\rho_2(x, n_m, x_f|n, 0) \underset{\substack{\tau_1 \text{ fixed} \\ \tau_2 \text{ fixed}}}{\sim} \frac{x}{a_2}\frac{x-x_f}{a_2}\frac{1}{4(a_2)^2\pi(n_m(n-n_m))^{3/2}} e^{-\left[\frac{x}{a_2}\right]^2\frac{1}{4n_m}-\left[\frac{x-x_f}{a_2}\right]^2\frac{1}{4(n-n_m)}}. \tag{6.19}$$

Next, for jump processes with $\mu < 2$, the asymptotic behavior of ρ_2 is simply obtained by plugging the asymptotic expression of $G_0(x,n)$ into equation (6.18), and we obtain

$$\rho_2(x, n_m, x_f|n, 0) \sim \sqrt{n_m(n-n_m)}\frac{4\mu^2}{\pi^3}\sin^2\left(\frac{\pi\mu}{2}\right)\Gamma^2(\mu)\left[\frac{a_\mu^2}{x(x-x_f)}\right]^\mu \frac{1}{x(x-x_f)}. \tag{6.20}$$

As a concluding remark, we emphasize that expressions for ρ, ρ_1 and ρ_2 hold only because the jump process can freely become negative, so that all distributions can be expressed in terms of semi-infinite quantities, such as the survival probability and semi-infinite propagator. For jump processes killed (*ie* stopped) upon the first crossing of 0, the determination of extreme value statistics requires a new, additional building block, which we introduce in the next section.

6.2 Strip probability of jump processes

The *strip* probability $\mu_{0,\underline{x}}(n)$ of a jump process is defined as the probability for a process starting from $x_0 = 0$ to stay positive during its first n steps, and to reach its maximum x on its n^{th} step exactly. In other words, trajectories that contribute to $\mu_{0,\underline{x}}(n)$ go from one end of a strip of width x to the other end in exactly n steps, as depicted in figure 6.1(b). Importantly, $\mu_{0,\underline{x}}(n)$ is *not* a semi-infinite observable, since the jump process cannot exit the interval $[0,x]$ during its n steps.

Sparre-Andersen-like behavior of the integrated strip probability. Remarkably, the integrated strip probability $p_n = \int_0^\infty \mu_{0,\underline{x}}(n)\mathrm{d}x$, defined as the probability that the jump process stays positive during n steps and reaches its maximum on its last step takes a universal form, independently of the jump distribution $p(\ell)$:

$$p_n = \frac{1}{2n}. \tag{6.21}$$

We stress that this striking result was first brought to our attention in a conjecture from [Mori *et al.* 2020b]. Although we expect that this combinatorial result has been proved in the mathematical literature, we provide in Appendix I a custom proof of equation (6.21).

6.2.1 Exact expression of the strip probability

To derive an exact expression of the strip probability, we introduce the joint distribution $\sigma(x,n|0)$ of the running maximum x and FPT n through 0, already encountered in chapter 3. By partitioning over the step k at which x is reached, and decomposing the trajectory around k, the joint distribution σ is given by

$$\sigma(x,n|0) = \sum_{k=1}^{n-1} \mu_{0,\underline{x}}(k) F_{0,\underline{x}}(n-k|0). \tag{6.22}$$

Additionally, we know from equation (5.29) of chapter 5 that $\sigma(x,n|0) = \frac{\mathrm{d}}{\mathrm{d}x}F_{\underline{0},x}(n|0)$. As a result, the strip probability is given in the generating function formalism:

$$\sum_{n=2}^{\infty} \xi^n \mu_{0,\underline{x}}(n) \equiv \mu_{0,\underline{x}}(\xi) = \frac{\frac{\mathrm{d}}{\mathrm{d}x} F_{\underline{0},x}(\xi|0)}{F_{0,\underline{x}}(\xi|0)}. \tag{6.23}$$

Importantly, equation (6.23) is exact, and reduces the evaluation of the strip probability to that of the LETP $F_{\underline{0},x}(\xi|0)$ and RETP $F_{0,\underline{x}}(\xi|0)$, which have been extensively studied in chapter 5. As a first explicit result, we compute the strip probability for one exactly solvable case.

Exponential jump process. In the specific case of the exponential jump process with $p(\ell) = \frac{\gamma}{2} e^{-\gamma|\ell|}$, the LETP can be exactly computed from the integral equation obtained by partitioning over the position of the walker after one step:

$$F_{\underline{0},x}(n|x_0) = (1-\delta_{n,1}) \int_0^x F_{\underline{0},x}(n-1|u) p(u-x_0) \mathrm{d}u + \delta_{n,1} \int_{-\infty}^0 p(u-x_0) \mathrm{d}u. \tag{6.24}$$

By making use of the fact that $\left[\frac{\mathrm{d}^2}{\mathrm{d}^2 u} - \gamma^2\right] p(u-x) = \gamma^2 \delta(u-x)$, we explicitly solve equation (6.24) and obtain the LETP in the generating function formalism:

$$F_{\underline{0},x}(\xi|x_0) = \frac{\xi\left(\sqrt{1-\xi}\cosh((x-x_0)\gamma\sqrt{1-\xi}) + \sinh((x-x_0)\gamma\sqrt{1-\xi})\right)}{2\sqrt{1-\xi}\cosh(x\gamma\sqrt{1-\xi}) - (\xi-2)\sinh(x\gamma\sqrt{1-\xi})}. \tag{6.25}$$

In turn, since the process is symmetric, the RETP is simply given by $F_{0,\underline{x}}(\xi|x_0) = F_{\underline{0},x}(\xi|x-x_0)$, and we obtain the explicit strip probability:

$$\mu_{0,\underline{x}}(\xi) = \frac{\gamma(1-\xi)\xi}{(2-\xi)\sqrt{1-\xi}\sinh\left(\gamma\sqrt{1-\xi}x\right) + 2(1-\xi)\cosh\left(\gamma\sqrt{1-\xi}x\right)} \tag{6.26}$$

which we illustrate in figure 6.3(a), for a fixed number of steps n.

Of note, the derivation of an explicit expression of $\mu_{0,\underline{x}}(\xi)$ is only possible because of the specific form of the jump distribution $p(\ell)$. However, for general jump processes with arbitrary symmetric $p(\ell)$, we expect the strip probability to display emerging universal behavior in the large n and x limit. In the following we consider separately the case $\mu = 2$ and $\mu < 2$.

6.2.2 Asymptotic behavior of the strip probability - $\mu = 2$

Consider first the case of a jump process with $\mu = 2$. Again, in the large n and x limit and in the scaling regime $\tau = n/(x/a_2)^2$ fixed, the process converges towards a Brownian motion of diffusion coefficient $D = a_2^2$. Additionally, in the limit $x \to \infty$, no overshoot occurs when the jump process crosses level x for the first time. In turn, the strip probability is asymptotically proportional to the RETP: $\mu_{0,\underline{x}}(n) \propto F_{0,\underline{x}}(n|0)$, and we find the proportionality constant to be equal to $1/a_2$. By making use of the asymptotic behavior of the RETP given in table 5.1 of chapter 5, we obtain:

$$\mu_{0,\underline{x}}(n) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} 2\left[\frac{a_2}{x}\right]^2 \frac{1}{x}\pi^2 \sum_{k=1}^{\infty} k^2 (-1)^{k+1} e^{-k^2\pi^2\tau}. \tag{6.27}$$

6.2.3 Asymptotic behavior of the strip probability - $\mu < 2$

For heavy-tailed jump processes with $\mu < 2$, overshoots occur even in the limit $x \to \infty$, such that the identification of the strip probability and the RETP is no longer valid. Additionally, the exact expression (6.23) cannot be used to asymptotically analyze $\mu_{0,\underline{x}}(n)$; indeed, the large x leading order of the LETP $F_{\underline{0},x}(\xi|0)$ obtained in chapter 4 is given by $F_{\underline{0},x}(\xi|0) \sim F_{\underline{0}}(\xi|0)$ and independent of x. To circumvent these difficulties, we introduce the cumulative strip probability $\mu_{0,\underline{\geq x}}(n)$

$$\mu_{0,\underline{\geq x}}(n) = \int_x^\infty \mu_{0,\underline{u}}(n)\mathrm{d}u, \tag{6.28}$$

defined as the probability that the walker stays positive for n steps, reaches its maximum on its last step, and that this maximum is larger than x. We now partition trajectories on the first step k at which x is crossed, and the corresponding position u of the walker upon crossing x.

1. The probabilistic weight of the first part of the trajectory is given by the *hitting* distribution $\hat{F}_{0,\underline{x}}(u,k|0)$ defined as the distribution of the walker's position right after exiting the interval. In particular, we stress that the RETP is a marginal of the hitting distribution.

$$\int_x^\infty \hat{F}_{0,\underline{x}}(u,k|0)\mathrm{d}u = F_{0,\underline{x}}(k|0). \tag{6.29}$$

2. During the remaining $n-k$ steps, we require the walker to stay positive,bib and reach a maximum $y \geq u$ on its last step. Considering the trajectory backwards,*ie* from step n to step k, the associated probabilistic weight is given by $G_{[0,y]}(y-u, n-k|0)$, where we recall that $G_{[0,y]}$ is the bounded propagator of the jump process.

In turn, $\mu_{0,\underline{\geq x}}(n)$ is given by

$$\begin{aligned}\mu_{0,\underline{\geq x}}(n) &= \sum_{k=1}^n \int_x^\infty \hat{F}_{0,\underline{x}}(u,k|0) \int_u^\infty G_{[0,y]}(y-u, n-k|0)\mathrm{d}y\mathrm{d}u \\ &= \sum_{k=1}^n \int_x^\infty \hat{F}_{0,\underline{x}}(u,k|0) \int_0^\infty G_{[0,y+u]}(y, n-k|0)\mathrm{d}y\mathrm{d}u.\end{aligned} \tag{6.30}$$

Importantly, in the large x limit, since $u > x$, we have $G_{[0,y+u]}(y, n-k|0) \sim G_0(y, n-k|0)$. Note that this argument is identical to the one used for the asymptotic analysis of $F_{0,\underline{x}}(n|0)$ in equation (5.19) of chapter 5. As a result, in the limit $n_x \gg n \gg 1$, the cumulative strip probability asymptotically reads

$$\begin{aligned}\mu_{0,\underline{\geq x}}(n) &\sim \sum_{k=1}^n \int_x^\infty \hat{F}_{0,\underline{x}}(u,k|0) \int_0^\infty G_0(y, n-k)\mathrm{d}y\mathrm{d}u \\ &\sim \sum_{k=1}^n \int_x^\infty \hat{F}_{0,\underline{x}}(u,k|0)\mathrm{d}u\ q(0,n-k) \\ &\sim \sum_{k=1}^n F_{0,\underline{x}}(k|0) q(0,n-k).\end{aligned} \tag{6.31}$$

The probabilistic interpretation is clear: the walker stays in the interval $[0, x]$ for $k-1$ steps before performing a big jump beyond x. The boundary at 0 is then too far to be seen and the remaining probabilistic weight is simply the semi-infinite survival probability $q(0, n-k)$. Recalling the asymptotic form (5.22) of the RETP $F_{0,\underline{x}}(n|0)$

$$F_{0,\underline{x}}(n|0) \underset{1\ll n\ll n_x}{\sim} q(0,n-1)\frac{\Gamma(\mu)}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^\mu, \tag{6.32}$$

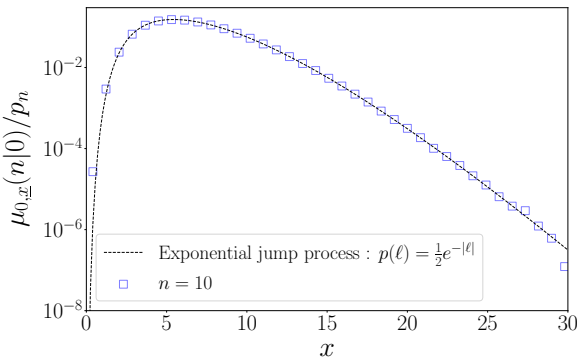
we finally obtain the cumulative strip probability in the generating function formalism

$$\mu_{0,\underline{>x}}(\xi) \underset{1\ll n\ll n_x}{\sim} \frac{1}{\pi}\Gamma(\mu)\sin\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^\mu \frac{\xi}{1-\xi}. \tag{6.33}$$

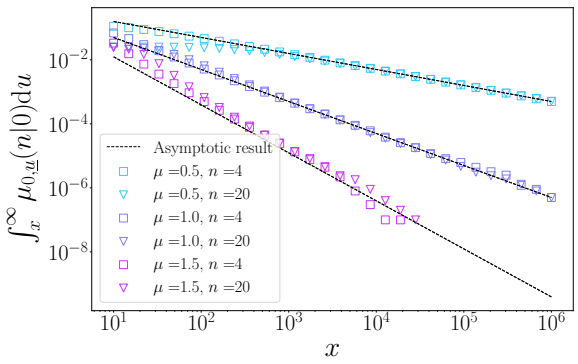
In turn, in the regime $1 \ll n \ll n_x$, the strip probability displays the following universal behavior, valid for all $0 < \mu < 2$:

$$\mu_{0,\underline{x}}(n) \underset{1\ll n\ll n_x}{\sim} \frac{\mu}{\pi}\,\Gamma(\mu)\sin\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}. \tag{6.34}$$

Equation (6.34) constitutes the main result of this section, and we provide numerical illustration in figure 6.3. Note that, surprisingly, the strip probability is independent of n in the $(x/a_\mu)^\mu \gg n \gg 1$ limit. We emphasize that this specific behavior only arises when the walker is constrained to stay positive. Indeed, the unconstrained counterpart of $\mu_{0,\underline{x}}(n)$ is the probability that the walker reaches its maximum x on its last step, without having to stay positive. In turn, this probability is given by $G_0(x, n)$, which clearly depends on n (see equation (6.16)).



(a) Distribution of the maximum for a 10 step long exponential jump processes conditioned to stay positive and to reach its maximum on its last step. The dashed line is computed from series expansion of the explicit result (6.26).



(b) Cumulative strip probability for Pareto jump processes with $p(\ell) \propto \mathbb{1}_{|\ell|>1}|\ell|^{-(1+\mu)}$. The dashed lines correspond to the n independent large x asymptotic result (6.34).

Figure 6.3

6.3 Extreme value statistics of semi-infinite jump processes

We devote this last section to the determination of EVS observables for semi-infinite jump processes killed (*ie* stopped) upon crossing 0 for the first time. We emphasize that all the results in this section are new and that the quantities derived throughout chapters 4,5 and 6, namely the splitting probability $\pi_{0,\underline{x}}(x_0)$, the RETP $F_{0,\underline{x}}(n|x_0)$ and the strip probability $\mu_{0,\underline{x}}(n)$, constitute necessary and sufficient building blocks to analyze a variety of important extreme value observables, for which we provide exact expressions and uncover universal asymptotic behavior.

As an introductory example, we consider the distribution $\mu(x|x_0)$ of the maximum reached before crossing 0, which is given by the splitting probability:

$$\mu(x|x_0) = -\frac{\mathrm{d}}{\mathrm{d}x}\pi_{0,\underline{x}}(x_0). \tag{6.35}$$

Equation (6.35) is exact, but $\mu(x|x_0)$ depends on the specific shape of the single jump distribution $p(\ell)$. However, by making use of the large x asymptotic behavior (4.33) of $\pi_{0,\underline{x}}(x_0)$, we obtain:

$$\mu(x|x_0) \underset{x\to\infty}{\sim} \mu 2^{\mu-2}\Gamma\left(\frac{1+\mu}{2}\right)\left[\frac{1}{\sqrt{\pi}} + V(x_0)\right]\left[\frac{a_\mu}{x}\right]^{\frac{\mu}{2}}\frac{1}{x}. \tag{6.36}$$

Importantly, equation (6.36) is valid for general symmetric jump processes with $0 < \mu \leq 2$. Additionally, we emphasize that the x_0 dependence of $\mu(x|x_0)$ is fully contained in the function $V(x_0)$, which has been extensively characterized in chapter 4. Of note, we hereafter systematically consider $x_0 = 0$.

6.3.1 Joint space and time distributions

As announced, we now investigate joint distributions of extremums and times, for processes killed upon becoming strictly negative.

Maximum and time at which it is reached. Consider first the joint distribution $\rho_3(x, n|0)$ of the maximum x and the step n at which it is reached before the first-passage across 0. By using the Markovian path decomposition technique around step n, ρ_3 is simply given by:

$$\rho_3(x, n|0) = \mu_{0,\underline{x}}(n)\pi_{0,\underline{x}}(0). \tag{6.37}$$

Again, we stress that the path decomposition technique is well adapted to deal with jump processes; the derivation of the corresponding joint distribution for Brownian motion, given in [Randon-Furling & Majumdar 2007] is more tedious. We now focus on the emergent universal asymptotic behavior of ρ_3 in the large n and x limit, which is easily extracted from the previous analysis of both $\mu_{0,\underline{x}}(n)$ and $\pi_{0,\underline{x}}(0)$.

- For jump processes in the Brownian basin of attraction, *ie* with $\mu = 2$, ρ_3 displays a universal asymptotic behavior in the scaling limit $n \to \infty$, $x \to \infty$ and $\tau = n/(x/a_2)^2$ fixed:

$$\rho_3(x,n|0) \underset{\substack{n\to\infty\\ x\to\infty\\ \tau \text{ fixed}}}{\sim} 2\left[\frac{a_2}{x}\right]^3 \frac{1}{x}\pi^2 \sum_{k=1}^{\infty} k^2(-1)^{k+1} e^{-k^2\pi^2\tau}. \tag{6.38}$$

- For heavy-tailed processes with $\mu < 2$, we determined both asymptotic behaviors of $\mu_{0,\underline{x}}(n)$ and $\pi_{0,\underline{x}}(0)$ in the $n_x \gg n \gg 1$ limit. In turn, we obtain the following asymptotic behavior:

$$\rho_3(x,n|0) \underset{1\ll n\ll n_x}{\sim} \frac{2^{\mu-1}\ \mu\ \Gamma(\frac{1+\mu}{2})\ \Gamma(\mu)\ \sin(\frac{\pi\mu}{2})}{\pi^{\frac{3}{2}}} \left[\frac{a_\mu}{x}\right]^{\frac{3\mu}{2}} \frac{1}{x}. \tag{6.39}$$

  Note that, because of the asymptotic behavior of $\mu_{0,\underline{x}}(n)$, ρ_3 is also independent of n. However, we recall that this only holds in the $1 \ll n \ll (x/a_\mu)^\mu$ limit.

Maximum, time at which it is reached, and FPT through 0. As a second example, we consider the multivariate distribution $\rho_4(x,n,n_f|0)$ of the maximum x, step at which it is reached n and FPT n_f through 0. Decomposing the trajectory yields the following exact expression:

$$\rho_4(x,n,n_f|0) = \mu_{0,\underline{x}}(n) F_{0,\underline{x}}(n_f - n|0), \tag{6.40}$$

from which we extract the large x and n asymptotic behavior of ρ_4 in the usual regimes:

$$\begin{aligned}
&\mu = 2 \quad \rho_4(x,n,n_f|0) \underset{\substack{n_f\to\infty\\ n\to\infty\\ x\to\infty\\ \tau \text{ fixed}}}{\sim} 4\left[\frac{a_2}{x}\right]^5 \frac{1}{x}\pi^4 \sum_{k,l=1}^{\infty} k^2 l^2 (-1)^{k+l} e^{-k^2\pi^2 n\left[\frac{a_2}{x}\right]^2 - l^2\pi^2(n_f-n)\left[\frac{a_2}{x}\right]^2} \\
&\mu < 2 \quad \rho_4(x,n,n_f|0) \underset{1\ll n\ll n_f\ll n_x}{\sim} \frac{\mu}{\sqrt{\pi(n_f-n)}} \left[\frac{\Gamma(\mu)}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\right]^2 \left[\frac{a_\mu}{x}\right]^{2\mu} \frac{1}{x}.
\end{aligned} \tag{6.41}$$

Maximum and first-passage time through 0. With the help of the strip probability, we can finally evaluate the joint distribution $\sigma(x,n|0)$ of the maximum x and the FPT n through 0 for general jump processes.

- When $\mu = 2$, the asymptotic behavior of the joint distribution σ in the usual scaling regime is obtained from the corresponding joint distribution for Brownian motion (3.13), which we computed in chapter 3. As a result, $\sigma(x,n|0)$ is asymptotically given by:

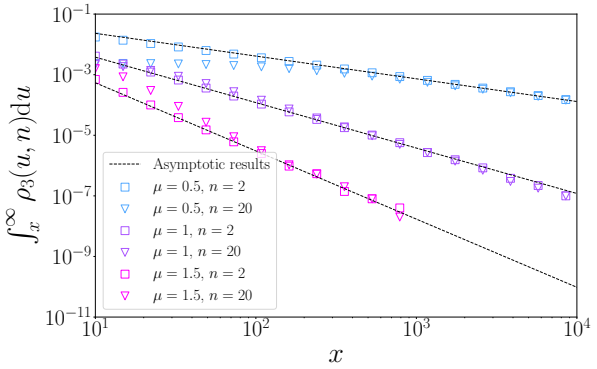
$$\sigma(x,n|0) \underset{\substack{n\to\infty\\ x\to\infty\\ \tau \text{ fixed}}}{\sim} 2\pi^2 \left[\frac{a_2}{x}\right]^3 \frac{1}{x} \sum_{k=1}^{\infty} e^{-k^2\pi^2 n\left[\frac{a_2}{x}\right]^2} k^2 \left[2k^2\pi^2 n\left[\frac{a_2}{x}\right]^2 - 3\right], \tag{6.42}$$

- For heavy-tailed jump processes with $\mu < 2$, we revert to the Markovian path decomposition, and rewrite equation (6.22) in the generating function formalism to obtain:

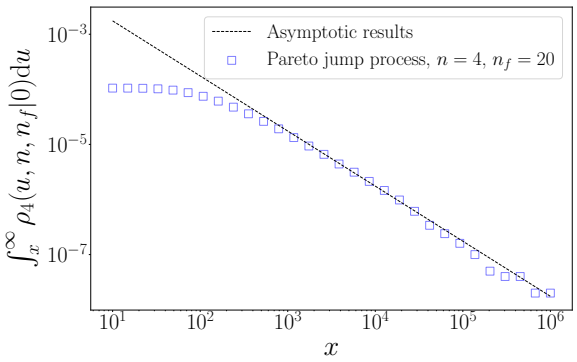
$$\sigma(x,\xi|0) = \mu_{0,\underline{x}}(\xi) F_{0,\underline{x}}(\xi|0). \tag{6.43}$$

Equation (6.43) is exact, and its asymptotic behavior in the limit $1 \ll n \ll (x/a_\mu)^\mu$ is finally given by

$$\sigma(x,n|0) \underset{1\ll n\ll n_x}{\sim} \frac{2\mu\sqrt{n}}{\pi^{\frac{5}{2}}}\Gamma^2(\mu)\sin^2\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^{2\mu}\frac{1}{x}. \tag{6.44}$$



(a) Asymptotic cumulative distribution of ρ_3 obtained from equation (6.39) for processes with $\tilde{p}(k) = e^{-|k|^\mu}$. We exhibit the n independence of ρ_3 for large x values.



(b) Asymptotic cumulative distribution of ρ_4 for a Pareto jump process with $p(\ell) = \mathbb{1}_{|\ell|>1}\frac{1}{4}|\ell|^{-3/2}$. We exhibit the convergence towards result (6.41) as $x \to \infty$.

Figure 6.4

The asymptotic results (6.39) and (6.41) are numerically illustrated in figure 6.4. As a concluding remark, these three examples show that the strip probabilities and splitting probabilities are essential quantities to characterize EVS observables for semi-infinite jump processes. We now turn to a final application example, and investigate the *span* of general jump processes.

6.3.2 Span of jump processes

We consider the distribution $s(x,n)$ of the span of an n step long unbounded jump process issued from 0, defined as the difference between its maximum M and minimum m up to step n. In view of applying the Markovian path decomposition technique, we split trajectories contributing to $s(x,n)$ into three parts of respective length n_1, n_2 and n_3.

- Without loss of generality we consider that the maximum M of the trajectory is reached before the minimum m. During the first n_1 steps, the walker travels from its initial position to M.
- In a second part, of duration n_2, the walker travels from M to m, crossing a distance of x exactly. Importantly, the associated probabilistic weight is equal to $\mu_{0,\underline{x}}(n_2)$.
- The remaining n_3 steps take the walker from m to some final position inside $[m, M]$.

By making use of the translational invariance of the jump process and the symmetry of the single jump distribution, the span distribution is then given by:

$$s(x,n) = 2 \sum_{\substack{n_1,n_2,n_3 \\ n_1+n_2+n_3=n}} \left[\int_0^x G_{[0,x]}(u,n_1|0)\mathrm{d}u\right] \mu_{0,\underline{x}}(n_2) \left[\int_0^x G_{[0,x]}(u,n_3|0)\mathrm{d}u\right], \tag{6.45}$$

where the factor 2 arises from the interchangeability of m and M. In the generating function formalism, we rewrite equation (6.45) as

$$s(x,\xi) = 2 \left[\int_0^x G_{[0,x]}(u,\xi|0)\mathrm{d}u\right] \mu_{0,\underline{x}}(\xi) \left[\int_0^x G_{[0,x]}(u,\xi|0)\mathrm{d}u\right]. \tag{6.46}$$

Equation (6.46) is exact, but involves the bounded propagators $G_{[0,x]}$ (defined as in chapter 4), which can only be computed for specific jump distributions.

Exponential jump process. When $p(\ell) = \frac{1}{2}e^{-|\ell|}$, the bounded propagator can be computed (see [Mori *et al.* 2021] for instance) and is given by

$$G_{[0,x]}(u,\xi|x_0=0) = -\frac{\xi\left(\sinh\left(\sqrt{1-\xi}(x-y)\right) + \sqrt{1-\xi}\cosh\left(\sqrt{1-\xi}(x-y)\right)\right)}{(\xi-2)\sinh\left(\sqrt{1-\xi}x\right) - 2\sqrt{1-\xi}\cosh\left(\sqrt{1-\xi}x\right)}. \tag{6.47}$$

Combining the expression (6.26) of $\mu_{0,\underline{x}}(\xi)$ and (6.47), we obtain the explicit span distribution of the exponential jump process:

$$s^{\exp}(x,\xi=1-\alpha) = -\frac{4(1-\alpha)\sqrt{\alpha}e^{\sqrt{\alpha}x}\left(-\alpha-2\sqrt{\alpha}\sinh\left(\sqrt{\alpha}x\right)+(-\alpha-1)\cosh\left(\sqrt{\alpha}x\right)+1\right)}{\left(-\alpha+\sqrt{\alpha}+(\alpha+\sqrt{\alpha})\,e^{\sqrt{\alpha}x}\right)^2\left(2\sqrt{\alpha}\cosh\left(\sqrt{\alpha}x\right)-(-\alpha-1)\sinh\left(\sqrt{\alpha}x\right)\right)}. \tag{6.48}$$

Asymptotic behavior of the span distribution. For general jump processes for which the bounded propagator cannot be computed exactly, we uncover universal asymptotic behavior in the usual large n and large x limit.

- When $\mu = 2$, we expect the span distribution to converge towards the corresponding continuous span distribution of Brownian motion, which can be found in [Borodin & Salminen 1996], or, more recently [Régnier *et al.* 2022]. In turn, defining again $\tau = n/(x/a_2)^2$ we obtain

$$s(x,n) \underset{\substack{n\to\infty \\ x\to\infty \\ \tau \text{ fixed}}}{\sim} \frac{4}{a_2\sqrt{\pi n}} \sum_{k=1}^{\infty} (-1)^{k+1} k^2 e^{-\frac{k^2}{4\tau}}. \tag{6.49}$$

- When $\mu < 2$ and in the limit $1 \ll n \ll n_x$ we have $G_{[0,x]}(u,\xi|0) \sim G_0(u,\xi)$. Note that we have now used this argument three times: for the derivation of the RETP, the

strip probability and the span distribution. As a result, the leading order behavior of $\int_0^x G_{[0,x]}(u,\xi|0)\mathrm{d}u$ is given by

$$\begin{aligned}\int_0^x G_{[0,x]}(u,\xi|0)\mathrm{d}u &\underset{1\ll n\ll n_x}{\sim} \int_0^x G_0(u,\xi)\mathrm{d}u \\ &\underset{1\ll n\ll n_x}{\sim} q(0,\xi) = \frac{1}{\sqrt{1-\xi}},\end{aligned} \tag{6.50}$$

and we obtain the leading order of $s(x,\xi)$, valid for general heavy-tailed jump processes:

$$s(x,\xi) \underset{1\ll n\ll n_x}{\sim} 2\left(\frac{1}{\sqrt{1-\xi}}\right)^2 \frac{\mu}{\pi}\,\Gamma(\mu)\sin\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}\frac{\xi}{1-\xi}. \tag{6.51}$$

Finally, the asymptotic span distribution in real space and time is given by

$$s(x,n) \underset{1\ll n\ll n_x}{\sim} \frac{2n\mu}{\pi}\,\Gamma(\mu)\sin\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^\mu \frac{1}{x}, \tag{6.52}$$

and we display agreement with simulations in figure 6.5(a).

6.3.3 Refined information on $F_{\underline{0},x}(n|0)$ for heavy-tailed jump processes

We conclude this chapter by deriving the first vanishing order of $F_{\underline{0},x}(n|0)$ in the $1 \ll n \ll n_x$ regime, which was not contained in the asymptotic analysis of the LETP presented in chapter 5. Let us define the first vanishing correction $h(x,\xi) \equiv F_{\underline{0}}(\xi|0) - F_{\underline{0},x}(\xi|0)$, and recall that $\sigma(x,\xi|0) = \frac{\mathrm{d}}{\mathrm{d}x}F_{\underline{0},x}(\xi|0)$, such that

$$\frac{\mathrm{d}}{\mathrm{d}x}h(x,\xi) = \sigma(x,\xi|0). \tag{6.53}$$

In turn, making use of the asymptotic behavior (6.44) of $\sigma(x,n|0)$, we obtain

$$h(x,\xi) \underset{x\to\infty}{\sim} \frac{\Gamma(\mu)^2\sin^2\left(\frac{\pi\mu}{2}\right)}{2\pi^2}\left[\frac{a_\mu}{x}\right]^{2\mu}\frac{\xi^2}{(1-\xi)^{\frac{3}{2}}}, \tag{6.54}$$

so that the LETP to first vanishing order is simply given by

$$F_{\underline{0},x}(\xi|0) \underset{x\to\infty}{=} F_{\underline{0}}(\xi|0) - \frac{\Gamma(\mu)^2\sin^2\left(\frac{\pi\mu}{2}\right)}{2\pi^2}\left[\frac{a_\mu}{x}\right]^{2\mu}\frac{\xi^2}{(1-\xi)^{\frac{3}{2}}} + o(x^{-2\mu}). \tag{6.55}$$

Importantly:

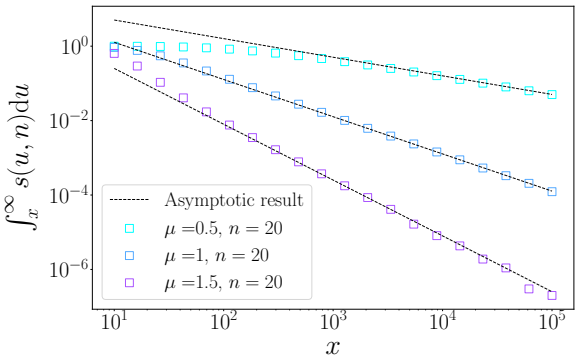
- The derivation of this first vanishing correction requires the knowledge of the asymptotic behavior of $\sigma(x,n|0)$. In turn, we obtained this behavior from the analysis of the strip probability, showcasing once more the essential role of this quantity in addressing EVS of jump processes.

- The algebraic x decay of the LETP, proportional to $x^{-2\mu}$, is stronger than the corresponding one for the RETP, proportional to $x^{-\mu}$. Indeed trajectories that contribute to this first correction need to approach the x boundary significantly, and then come back to 0, in a *back and forth* motion. In contrast, trajectories contributing to the RETP simply need to approach the x boundary.

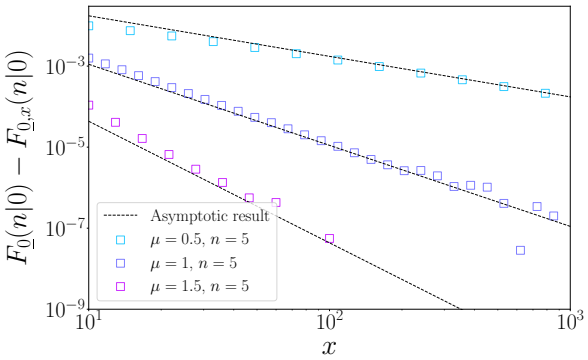
- Applying this argument further, we expect that both RETP and LETP can be written as series expansions in powers of a_μ/x:

$$\begin{aligned} F_{0,\underline{x}}(\xi|0) &= \sum_{k=0}^{\infty} a_k(\xi) \left[\frac{a_\mu}{x}\right]^{(2k+1)\mu} \\ F_{\underline{0},x}(\xi|0) &= \sum_{k=0}^{\infty} b_k(\xi) \left[\frac{a_\mu}{x}\right]^{2k\mu} \end{aligned} \tag{6.56}$$

  where the explicit determination of the $a_k(\xi)$ and $b_k(\xi)$ would require further investigation.



(a) Cumulative distribution of the span for three heavy-tailed processes with $\tilde{p}(k) = e^{-|k|^\mu}$ and fixed $n = 20$. The dashed lines come from the asymptotic result (6.52).



(b) First vanishing order of the LETP for three heavy-tailed processes with $\tilde{p}(k) = e^{-|k|^\mu}$ and fixed n. The dashed lines come from equation (6.55). Of note, since n is small here, we exactly invert the generating function in equation (6.55) by doing a series expansion in ξ instead of using the asymptotic large $\sqrt{n}$ behavior.

Figure 6.5

6.4 Conclusion

The discrete time nature of jump processes facilitates Markovian path decomposition, which is a standard tool for the analysis of *extreme value* statistics. In turn, prior knowledge on the semi-infinite survival probability $q(0,n)$ and the semi-infinite propagator $G_0(x,n)$ allows to derive exact results for a variety of extreme value observables of unbounded jump processes.

However, for semi-infinite jump processes killed (*ie* stopped) upon entering the negative half-line, the trajectories contributing to the EVS cannot become negative anymore, and both $q(0,n)$ and $G_0(x,n)$ become irrelevant in this context. As a workaround, we introduce the semi-infinite counterpart of $G_0(x,n)$, the strip probability $\mu_{0,\underline{x}}(n)$ defined as the probability for a walker starting from 0 to stay positive during n steps, and to reach its maximum x on its last step exactly.

We provide an explicit expression of the strip probability as a function of the LETP and RETP introduced in chapter 5, and uncover universal asymptotic behavior of $\mu_{0,\underline{x}}(n)$ for general jump processes. With the help of this new building block, we derive explicit and asymptotic results for extreme value observables of semi-infinite jump processes. In particular, we investigate a variety of space and time joint distributions summarized in table 6.1, characterize the *span* of jump processes, and provide a finer description of the asymptotic LETP than the one derived in the previous chapter.

EVS observable	Exact expression	$\mu<2$ - Asymptotic behavior in the regime $1 \ll n \ll n_x$	$\mu=2$ - Asymptotic behavior in the scaling regime n/n_x fixed
Infinite jump processes			
$\mu(x\|n,x_0=0)$	$\frac{\mathrm{d}}{\mathrm{d}x}q(x,n)$	$\frac{n\mu}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\Gamma(\mu)\left[\frac{a_\mu}{x}\right]^\mu\frac{1}{x}$	$\frac{1}{a_2\sqrt{\pi n}}e^{-\left[\frac{x}{a_2}\right]^2\frac{1}{4n}}$ o
$\rho(n_m\|n,x_0=0)$	$q(0,n_m)q(0,n-n_m)$	$\frac{1}{\pi}\frac{1}{\sqrt{n_m(n-n_m)}}$ o	
$\rho_1(x,n_m\|n,x_0=0)$	$G_0(x,n_m)q(0,n-n_m)$	$\frac{1}{\pi}\sqrt{\frac{n_m}{n-n_m}}\frac{2\mu}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\Gamma(\mu)\left[\frac{a_\mu}{x}\right]^\mu\frac{1}{x}$	$\frac{x}{a_2}\frac{1}{2a_2\pi n_m^{3/2}\sqrt{n-n_m}}e^{-\left[\frac{x}{a_2}\right]^2\frac{1}{4n_m}}$
$\rho_2(x,n_m,x_f\|n,0)$	$G_0(x,n_m)G_0(x-x_f,n-n_m)$	$\sqrt{n_m(n-n_m)}\frac{4\mu^2}{\pi^3}\sin^2\left(\frac{\pi\mu}{2}\right)\Gamma^2(\mu)\left[\frac{a_\mu^2}{x(x-x_f)}\right]^\mu\frac{1}{x(x-x_f)}$	$\frac{x}{a_2}\frac{x-x_f}{a_2}\frac{1}{4(a_2)^2\pi(n_m(n-n_m))^{3/2}}e^{-\left[\frac{x}{a_2}\right]^2\frac{1}{4n_m}-\left[\frac{x-x_f}{a_2}\right]^2\frac{1}{4(n-n_m)}}$
Semi infinite jump processes			
$\mu(x\|x_0=0)$	$-\frac{\mathrm{d}}{\mathrm{d}x}\pi_{0,\underline{x}}(0)$	$\mu 2^{\mu-2}\Gamma\left(\frac{1+\mu}{2}\right)\frac{1}{\sqrt{\pi}}\left[\frac{a_\mu}{x}\right]^{\frac{\mu}{2}}\frac{1}{x}$	
$\rho_3(x,n_m\|x_0=0)$	$\mu_{0,\underline{x}}(n_m)\pi_{0,\underline{x}}(0)$	$\frac{2^{\mu-1}\,\mu\,\Gamma(\frac{1+\mu}{2})\,\Gamma(\mu)\,\sin(\frac{\pi\mu}{2})}{\pi^{\frac{3}{2}}}\left[\frac{a_\mu}{x}\right]^{\frac{3\mu}{2}}\frac{1}{x}$	$2\left[\frac{a_2}{x}\right]^3\frac{1}{x}\pi^2\sum_{k=1}^\infty k^2(-1)^{k+1}e^{-k^2\pi^2 n_m\left[\frac{a_2}{x}\right]^2}$
$\rho_4(x,n_m,n_f\|x_0=0)$	$\mu_{0,\underline{x}}(n_m)F_{0,\underline{x}}(n_f-n_m\|0)$	$\frac{\mu}{\sqrt{\pi(n_f-n)}}\left[\frac{\Gamma(\mu)}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\right]^2\left[\frac{a_\mu}{x}\right]^{2\mu}\frac{1}{x}$	$4\left[\frac{a_2}{x}\right]^5\frac{1}{x}\pi^4\sum_{k,l=1}^\infty k^2l^2(-1)^{k+l}e^{-k^2\pi^2 n_m\left[\frac{a_2}{x}\right]^2-l^2\pi^2(n_f-n_m)\left[\frac{a_2}{x}\right]^2}$
$\sigma(x,n_f\|x_0=0)$	$-\frac{\mathrm{d}}{\mathrm{d}x}F_{\underline{0},x}(n_f\|0)$	$\frac{2\mu\sqrt{n_f}}{\pi^{\frac{5}{2}}}\Gamma^2(\mu)\sin^2\left(\frac{\pi\mu}{2}\right)\left[\frac{a_\mu}{x}\right]^{2\mu}\frac{1}{x}$	$2\pi^2\left[\frac{a_2}{x}\right]^3\frac{1}{x}\sum_{k=1}^\infty e^{-k^2\pi^2 n_f\left[\frac{a_2}{x}\right]^2}k^2\left[2k^2\pi^2 n_f\left[\frac{a_2}{x}\right]^2-3\right]$

Table 6.1: EVS observables for general jump processes. For the asymptotic behavior, n_x is defined as $n_x=(x/a_\mu)^\mu$ for all μ values, and x, n_m and n_f correspond respectively to the maximum, time at which the maximum is reached, and first-passage time across 0. Entries with a o are already given in the literature.

We emphasize that most of the results derived in chapters 4, 5 and 6 stem from the asymptotic derivation of the splitting probability $\pi_{0,\underline{x}}(x_0)$. In particular, the characterization of $\pi_{0,\underline{x}}(x_0)$ in the $x_0 \lesssim a_\mu$ limit naturally extends to the majority of first-passage observables, such as the RETP $F_{0,\underline{x}}(n|x_0)$ for instance. While we mostly focused on symmetric one-dimensional jump processes, we devote the concluding chapter to general isotropic jump processes, and show that the knowledge of the splitting probability can be used systematically to evaluate general first-passage observables.

Chapter 7

Matching method for isotropic jump processes

Contents

In this chapter, we introduce a general methodology to investigate *first-passage* observables for general isotropic jump processes in the large confining volume limit, which we will refer to as the *matching method*. As in the one-dimensional case, we aim to capture the specific behavior of jump processes starting close to absorbing boundaries, for which the continuous limit does not apply. We first provide a formal presentation of the method for a general first-passage observable $\mathcal{B}$, and discuss the asymptotic limits in which it holds. In turn, we illustrate the range of applicability of the methodology on a representative list of examples, spanning various jump processes across different geometries.

We first focus on geometrical observables. As a natural extension of the splitting probability $\pi_{0,\underline{x}}(x_0)$ for one-dimensional heavy-tailed jump processes, we consider the *hitting* distribution of the landing position upon exiting the interval. In higher dimensions, we focus on jump processes with $\mu = 2$, and derive two important quantities: the harmonic measure in a disk, and the splitting probability in an eccentric annulus.

The methodology equally applies to first-passage times. As an example, we explicitly compute mean exit-times from intervals, which cannot be derived from the asymptotic results obtained for exit-time probabilities in chapter 5. For jump processes with $\mu = 2$, we compute mean absorption times in disks with fully absorbing boundary, and then show that the matching

method is straightforwardly adapted to compute the complete absorption time distribution. Finally we consider the FPT to small absorbing targets in arbitrary confining domains with reflecting boundary conditions. For jump processes starting near the absorbing target, we show that the matching method can be combined with existing continuous large volume frameworks [Bénichou & Voituriez 2014] to compute the complete FPT distribution.

7.1 Matching method

7.1.1 First-passage observables for confined isotropic jump processes

Jump processes in dimension $d \geq 1$ describe the stochastic evolution in $\mathbb{R}^d$ of the position $\mathbf{x_n}$ of a particle undergoing *i.i.d.* random jumps with common distribution $p(\mathbf{l})$ at each discrete time step. As we have now seen extensively, quantifying the statistics of such processes is key in understanding real life experimental data. For example, the scattering of photons in heterogeneous media can be correctly modeled by jump processes with isotropic heavy-tailed jump distribution [Savo *et al.* 2017], and the experimentally observed Run and Tumble motion of E. Coli [Patteson *et al.* 2015] can be seen as an isotropic jump process with $p(|\mathbf{l}|) = \frac{\gamma}{v} e^{-\frac{\gamma}{v}|\mathbf{l}|}$, where γ is the tumbling rate, and v the speed of the particle.

We hereafter consider strictly isotropic jump processes in d dimensions, whose single jump distribution $p(\mathbf{l})$ depends only on the norm $|\mathbf{l}| \equiv \ell$. Additionally, the large ℓ behavior of $p(\mathbf{l})$ is equivalently given by the small $\mathbf{k}$ behavior of the Fourier transform $\tilde{p}(\mathbf{k})$ of the jump distribution [1]:

$$\tilde{p}(\mathbf{k}) \underset{\mathbf{k}\to 0}{=} 1 - |a_\mu \mathbf{k}|^\mu + o(|\mathbf{k}|^\mu). \tag{7.1}$$

We recall that a_μ is the characteristic lengthscale of the single jump distribution, and the Levy exponent μ characterizes the tails of the distribution. In particular, when $0 < \mu < 2$, the jump distribution displays a power-law behavior $p(\ell) \propto \ell^{-(1+\mu)}$.

Unbounded isotropic jump processes have been well understood early on. For instance, the infinite propagator $G_\infty(\mathbf{x}, n|\mathbf{x_0})$, defined as the distribution of the position after n steps, is known exactly in the Fourier transformed and generating function formalism:

$$\sum_{n=0}^{\infty} \xi^n \left[\int_{\mathbb{R}^d} e^{i\mathbf{k}(\mathbf{x}-\mathbf{x_0})} G_\infty(\mathbf{x}, n|\mathbf{x_0}) \mathrm{d}\mathbf{k} \right] = \frac{1}{1 - \xi\tilde{p}(\mathbf{k})}. \tag{7.2}$$

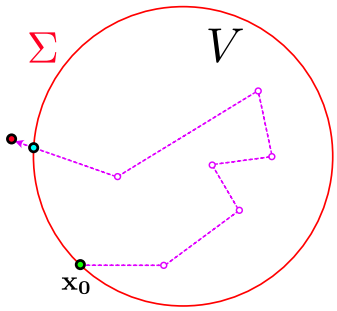
In turn, knowledge of the infinite propagator grants access to a variety of dynamical and first-passage observables, such as the FPT distribution to a given target for example. However, in a bounded geometry, there exists no general results valid for arbitrary isotropic jump processes.

In this chapter, we provide a methodology to systematically evaluate first-passage observables in the large confining volume limit. More precisely, we consider an isotropic jump process starting from $\mathbf{x_0}$ in a confining domain of volume V, whose boundary is split into a reflecting part, and an absorbing part denoted Σ, as depicted in figure 7.1. We now introduce an arbitrary first-passage observable $\mathcal{B}(\mathbf{x_0})$, evaluated upon first crossing of the absorbing part Σ of the boundary. To be more explicit, $\mathcal{B}(\mathbf{x_0})$ could equally correspond to:

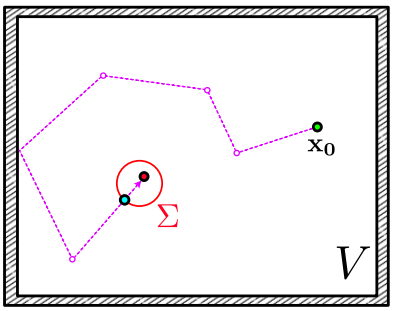
[1] Equation 7.1 should be understood as a definition of the general class of processes we are considering.

- The side through which a one-dimensional jump process escapes an absorbing interval $[0, x]$, namely the splitting probability $\pi_{0,\underline{x}}(x_0)$.
- The harmonic measure of a $2D$ jump process upon escaping a completely absorbing disk (see figure 7.4(a)).
- The FPT to an inner absorbing target in the interior of a confining domain with reflecting boundary conditions (see figure 7.1(b)).

In the following, we assume that the confining domain has a unique characteristic lengthscale R, and place ourselves in the large volume limit $R \gg a_\mu$. Importantly, $\mathcal{B}(\mathbf{x_0})$ is well defined for all initial positions $\mathbf{x_0}$ such that $\mathbf{x_0}$ belongs to the interior or the boundary of the domain. Indeed, as we have seen in chapters 4 to 6, the discreetness of the jump process allows for particles to start on the absorbing surface. To evaluate the distribution of $\mathcal{B}(\mathbf{x_0})$ for all initial positions, we investigate separately the case where $\mathbf{x_0}$ is far or close to the absorbing boundary.



(a) Two-dimensional isotropic jump process in a fully absorbing disk, starting from $\mathbf{x_0}$. Here the absorbing boundary Σ corresponds to the complete disk boundary, and the process is stopped as soon as it strictly exits the disk. Because of its discrete nature, the random walker is allowed to start *on* the boundary of the absorbing disk. In general, we refer to starting positions close to the absorbing boundary as *edge* initial condition.



(b) Two-dimensional jump process in a reflecting/absorbing confining domain, starting from $\mathbf{x_0}$, with an interior absorbing target (red circle). Here Σ only corresponds to the boundary of the interior target and the process is stopped as soon as it strictly penetrates the target. Since the process starts far from the absorbing target, this sample trajectory has a *bulk* initial condition.

Figure 7.1

7.1.2 Bulk initial conditions and the continuous limit

As for one-dimensional jump processes, d-dimensional isotropic jump processes are known to converge at large times to an associated continuous process, characterized only by the large ℓ behavior of the jump distribution, or equivalently the small $\mathbf{k}$ behavior of the Fourier transform $\tilde{p}(\mathbf{k})$:

- In the case $\mu = 2$, the jump process converges at large times to a d-dimensional Brownian motion with diffusion coefficient $D = a_2^2/d$. Note that more details on this stochastic process can be found in [Redner 2001] for instance.
- In the case $\mu < 2$, the jump process converges at large times to an isotropic α-stable process [Kyprianou & Pardo 2022], whose continuous space and time propagator is simply given by its Fourier Transform:

$$G_\infty(\mathbf{k}, n|\mathbf{0}) = e^{-n|a_\mu \mathbf{k}|^\mu}. \tag{7.3}$$

We now introduce the orthogonal projection $\mathbf{x_h}$ of $\mathbf{x_0}$ on the absorbing surface Σ, and the corresponding source-surface distance $d_e = |\mathbf{x_h} - \mathbf{x_0}|$ (see figure 7.2 for a schematic of the situation). In turn, we define the *continuous limit* as the regime $d_e \gg a_\mu$ and $R \gg a_\mu$. In this limit, the starting position is considered to be in the bulk of the domain with respect to the absorbing surface, and the observable $\mathcal{B}(\mathbf{x_0})$ is simply given by

$$\mathcal{B}(\mathbf{x_0}) \sim \mathcal{B}^{(c)}(\mathbf{x_0}), \tag{7.4}$$

where $\mathcal{B}^{(c)}(\mathbf{x_0})$ corresponds to the same observable, evaluated for the continuous limit process. As an illustration, we showed in chapter 4 equation (4.29), that in the regime $x_0 \gg a_\mu$ and $x \gg a_\mu$ the splitting probability $\pi_{0,\underline{x}}(x_0)$ is equivalent to its continuous counterpart $\pi^{(c)}_{0,\underline{x}}(x_0)$. Before shifting our focus to *edge* initial conditions, we emphasize that equation (7.4) is only relevant if the distribution of $\mathcal{B}^{(c)}(\mathbf{x_0})$ can be evaluated. Note however that it is the case for a large number of first-passage observables, as we will see shortly.

7.1.3 Edge initial conditions

We now consider the *edge* regime, defined as $d_e \lesssim a_\mu$, for which the continuous limit does not apply, and equation (7.4) is not valid. This is in particular the case when $\mathbf{x_0} \in \Sigma$.

Projected jump process. As an important tool for the matching method, we introduce the projected jump process associated to the isotropic d-dimensional one. Let us choose an arbitrary d-dimensional normalized vector $\mathbf{v}$, and consider the successive positions $(\mathbf{x_0}, \mathbf{x_1}, \dots \mathbf{x_n})$ of the particle. The successive positions of the projected jump process are given by $(x_0^\perp, x_1^\perp, \dots, x_n^\perp) = (\mathbf{x_0}.\mathbf{v}, \mathbf{x_1}.\mathbf{v}, \dots \mathbf{x_n}.\mathbf{v})$. In turn, the projected jump process is a one-dimensional, symmetric and continuous jump process, whose single jump distribution is given by

$$p_\perp(\ell) = \int \mathrm{d}\mathbf{u} p(\mathbf{u}) \delta(\mathbf{u}.\mathbf{v} - \ell). \tag{7.5}$$

Note that since the initial jump process is isotropic, $p_\perp(\ell)$ is strictly independent of the choice of $\mathbf{v}$. Importantly,the large ℓ decay of $p_\perp(\ell)$ is identical to that of $p(\mathbf{l})$; in the Fourier formalism, one has

$$\tilde{p}_\perp(k) \underset{k\to 0}{=} 1 - (a_\mu^\perp |k|)^\mu + o(|k|^\mu), \tag{7.6}$$

where μ is unchanged. Additionally, in the specific $\mu = 2$ case, the projected characteristic length $a_2^\perp$ is explicit, and given by

$$a_2^\perp = \frac{a_2}{\sqrt{d}}, \tag{7.7}$$

where d is the space dimension. We emphasize that the crucial role of the projected jump processes has already been underlined in chapter 4. Indeed, we showed that the splitting probability in $3D$

slabs is given by the one-dimensional splitting probability of the projected jump process upon a vector $\mathbf{v}$ chosen perpendicularly to the slab's delimiting hyperplanes H_1 and H_2 (see figure 4.2(b) of chapter 4 for a reminder of the geometrical setting).

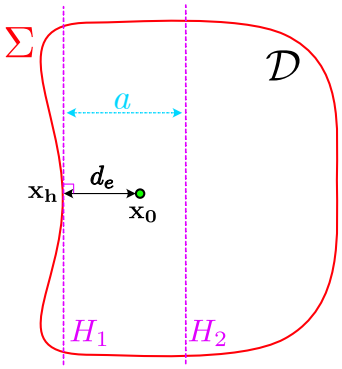


Figure 7.2: **Schematics of the matching method.** The full boundary Σ is absorbing, and the starting position $\mathbf{x_0}$ is close to Σ. The two delimiting hyperplanes H_1 and H_2 are parallel, and H_1 is tangent to Σ. Importantly, we consider the *edge* regime $d_e \lesssim a_\mu \ll a \ll R$. In this regime, trajectories conditioned to cross H_2 before H_1 are asymptotically equivalent to the trajectories of the continuous limit process.

Matching method. In the case $d_e \lesssim a_\mu$, the discrete nature of jump processes cannot be neglected, even in the large confining volume limit. To account for such discrete effects, we introduce two parallel hyperplanes H_1 and H_2 distant from an arbitrary distance a, such that H_1 is tangent to Σ at $\mathbf{x_h}$ (see figure 7.2). Note that we consider Σ to be a smooth surface.

To evaluate $P(\mathcal{B}(\mathbf{x_0}) = b)$, we partition the contributing trajectories according to whether or not H_2 is crossed before H_1. Denoting $\pi_{H_1,\underline{H_2}}(\mathbf{x_0})$ the hyperplane splitting probability that the process crosses H_2 before H_1, we obtain the following exact expression:

$$\begin{aligned} P(\mathcal{B}(\mathbf{x_0}) = b) =& \pi_{H_1,\underline{H_2}}(\mathbf{x_0}) P(\mathcal{B}(\mathbf{x_0}) = b | H_2 \text{ crossed before } H_1) \\ &+ \pi_{\underline{H_1},H_2}(\mathbf{x_0}) P(\mathcal{B}(\mathbf{x_0}) = b | H_1 \text{ crossed before } H_2). \end{aligned} \tag{7.8}$$

In turn, the hyperplane splitting probability is strictly equal to the one-dimensional splitting probability $\pi^{\perp}_{0,\underline{a}}(d_e)$ to reach a before 0 for the jump process projected perpendicularly to H_1 and H_2. Consequently, we rewrite equation (7.8) as:

$$\begin{aligned} P(\mathcal{B}(\mathbf{x_0}) = b) =& \pi^{\perp}_{0,\underline{a}}(d_e) P(\mathcal{B}(\mathbf{x_0}) = b | H_2 \text{ crossed before } H_1) \\ &+ \pi^{\perp}_{\underline{0},a}(d_e) P(\mathcal{B}(\mathbf{x_0}) = b | H_1 \text{ crossed before } H_2). \end{aligned} \tag{7.9}$$

Since the hyperplane distance a can be chosen arbitrarily, we place ourselves in the limit $a_\mu \ll a \ll R$ and proceed in 4 steps to investigate the asymptotic behavior of $P(\mathcal{B}(\mathbf{x_0}) = b)$.

1. First, in the limit $a_\mu \ll a$, we showed in chapter 4 that $\pi^{\perp}_{0,\underline{a}}(d_e)$ is known exactly, and all our results are directly applicable. In particular, $\pi^{\perp}_{0,\underline{a}}(0)$ is strictly non vanishing, and the d_e dependence of $\pi^{\perp}_{0,\underline{a}}(d_e)$ depends only on the small scale details of the projected jump distribution.

2. Second, in the limit $a_\mu \ll a$, the statistics of the jump process conditioned to cross H_2 before H_1 converge towards the statistics of the continuous limit process conditioned to cross H_2 before H_1, and we have:

$$P(\mathcal{B}(\mathbf{x_0}) = b|H_2 \text{ crossed before } H_1) \underset{a_\mu \ll a}{\sim} P(\mathcal{B}^{(c)}(\mathbf{x_0}) = b|H_2 \text{ crossed before } H_1). \quad (7.10)$$

3. To evaluate the conditional distribution of the continuous limit process, we introduce the splitting probability $\pi^{(c)}_{0,\underline{a}}(\mathbf{x_0})$ that the continuous limit process crosses H_2 before H_1, such that

$$P(\mathcal{B}^c(\mathbf{x_0}) = b|H_2 \text{ crossed before } H_1) = \frac{P(\mathcal{B}^{(c)}(\mathbf{x_0}) = b)}{\pi^{(c)}_{0,\underline{a}}(\mathbf{x_0})}. \quad (7.11)$$

Importantly, in the large volume limit $a \ll R$, the conditional distribution of $\mathcal{B}(\mathbf{x_0})$ becomes independent of the initial position $\mathbf{x_0}$, and is given by:

$$P(\mathcal{B}^c(\mathbf{x_0}) = b|H_2 \text{ crossed before } H_1) \underset{a \ll R}{\sim} \lim_{\mathbf{u} \to \mathbf{x_h}} \left[\frac{P(\mathcal{B}^{(c)}(\mathbf{u}) = b)}{\pi^{(c)}_{0,\underline{a}}(\mathbf{u})} \right]. \quad (7.12)$$

Note that we have already used this argument in equation (5.11) of chapter 5 to determine the conditional rightward first exit-time probability $F_{0,\underline{x}}(n|x_0)/\pi_{0,\underline{x}}(x_0)$ in the $x \to \infty$ limit.

4. Lastly, we need to take into account the second term in equation (7.8), corresponding to trajectories that cross H_1 before H_2. Defining the radius or curvature R_c of Σ at $\mathbf{x_h}$, we consider the limit $R_c \gg a_\mu$. In this limit, Σ is locally approximated by H_1; as a result, trajectories that cross H_1 before H_2 are immediately absorbed in the vicinity of $\mathbf{x_h}$. In turn, in the large volume limit $R \gg a_\mu$, the contribution of such trajectories to the statistics of $\mathcal{B}(\mathbf{x_0})$ is sub-leading [2] compared to trajectories that can be absorbed on the whole surface Σ, and can thus be neglected.

Finally, we obtain the large volume asymptotic behavior of arbitrary first-passage observables for general d-dimensional isotropic jump processes:

$$P(\mathcal{B}(\mathbf{x_0}) = b) \underset{\substack{a_\mu \ll a \ll R \\ a_\mu \ll R_c}}{\sim} \pi^{\perp}_{0,\underline{a}}(d_e) \lim_{\mathbf{u} \to \mathbf{x_h}} \left[\frac{P(\mathcal{B}^{(c)}(\mathbf{u}) = b)}{\pi^{(c)}_{0,\underline{a}}(\mathbf{u})} \right]. \quad (7.13)$$

Together with equation (7.4), equation (7.13) is the main result of this chapter, and deserves a few comments.

[2] As we will see in the following examples, this statement is observable-dependent, and will be verified systematically.

- The limit on the RHS of equation (7.13) is well defined. Indeed, we expect that for a given confining domain, all the distributions of continuous first-passage observables can be projected on an eigenfunction basis $\{\psi_k\}$

$$P(\mathcal{B}^{(c)}(\mathbf{u}) = b) = \sum_{k=1}^{\infty} c_k \psi_k(\mathbf{u}), \tag{7.14}$$

  where the c_k depend on both $\mathcal{B}^{(c)}$ and b. In turn, the leading $\mathbf{u}$ behavior of $P(\mathcal{B}^{(c)}(\mathbf{u}) = b)$ as $\mathbf{u} \to \mathbf{x_h}$ is identical for all first-passage observables.

- We emphasize that the key ingredient of the matching method is the one-dimensional splitting probability of the projected process, which we first derived in [Klinger *et al.* 2022b], and discussed at length in chapter 4. Importantly, the microscopic details and the discrete nature of the initial jump process are embedded in the d_e dependence of $\pi^{\perp}_{0,\underline{a}}(d_e)$.

- Equation (7.13) has the status of a ready-to-use method, requiring a few essential ingredients. More precisely, to evaluate the statistics of an arbitrary first-passage observable $\mathcal{B}(\mathbf{x_0})$ for a given isotropic jump process, one needs to be able to identify both the limit process and the projected process, and to evaluate $P(\mathcal{B}^{(c)}(\mathbf{u}) = b)$ and $\pi^{(c)}_{0,\underline{a}}(\mathbf{u})$.

In conclusion, we have provided a general and broadly applicable methodology to asymptotically evaluate first-passage observables for arbitrary confined jump processes. Our results go beyond the classical continuous approach (7.4), which fails to consider the discrete nature of jump processes and thus cannot capture the specificity of *edge* initial conditions. In this context, we derive a variety of new important results for representative examples of confined jump processes, spanning numerous observables and geometries. In the next two sections, we demonstrate the versatility and flexibility of the matching methodology by computing geometrical and temporal first-passage observables in one or more dimensions, relevant to the description of real physical systems.

7.2 Geometrical observables

In this section, we focus exclusively on geometrical first-passage observables, which only depend on the position of the particle upon exiting the confining domain and do not provide any information on the time needed to exit. As an example, the one-dimensional splitting probability is a geometrical observable.

7.2.1 Hitting distributions of one-dimensional jump processes

We first consider the most natural extension of the splitting probability: the hitting distribution $f_{[0,x]}(y|x_0)$ of the landing position y upon exiting the interval $[0, x]$. Without loss of generality, we consider hereafter that x_0=0, *ie* the particle starts on the absorbing boundary, and focus on $y > x$ values. In the large interval limit, no overshoot occurs for jump processes in the Brownian basin of attraction ($\mu = 2$), and the distribution of y concentrates around x, such that $f_{[0,x]}(y|0) \sim \delta(y - x)\pi_{0,\underline{x}}(0)$. Conversely, the hitting distribution for arbitrary symmetric

heavy-tailed jump processes with $\mu < 2$ does not concentrate, and can be obtained from the matching method.

1. We first determine the limit continuous process and the associated splitting probability. For heavy-tailed processes with $\mu < 2$, the limit process is an α-stable process of parameter μ, and the continuous splitting probability is given by [Widom 1961, Blumenthal *et al.* 1961, Majumdar *et al.* 2010b]

$$\pi^{(c)}_{0,\underline{a}}(x_0) \underset{x_0 \ll a}{\sim} \frac{2\Gamma(\mu)}{\mu\Gamma^2(\frac{\mu}{2})} \left(\frac{x_0}{a}\right)^{\frac{\mu}{2}}. \tag{7.15}$$

2. Next, hitting distributions for α-stable processes are known [Blumenthal *et al.* 1961], and read, for $y > x$

$$\begin{aligned} f^{(c)}_{[0,x]}(y|x_0) &= f^{(c)}_{[-1,1]}\left(\frac{2y}{x} - 1 | \frac{2x_0}{x} - 1\right)\frac{2}{x}, \\ f^{(c)}_{[-1,1]}(y|x_0) &= \frac{1}{\pi}\sin\left(\frac{\pi\mu}{2}\right)\left(\frac{|1-|x_0|^2|}{|1-|y|^2|}\right)^{\mu/2}|x_0 - y|^{-1}. \end{aligned} \tag{7.16}$$

   Note that the continuous hitting distribution is indeed vanishing when $x_0 \to 0$.

3. Lastly, we identify the projected jump process. In the one-dimensional case, the projected process is simply the initial jump process, and the splitting probability from 0 is given by

$$\pi^{\perp}_{0,\underline{a}}(0) \underset{a\to\infty}{\sim} \frac{2^{\mu-1}}{\sqrt{\pi}}\Gamma\left(\frac{1+\mu}{2}\right)\left(\frac{a_\mu}{a}\right)^{\frac{\mu}{2}}. \tag{7.17}$$

Making use of equations (7.15),(7.16) and (7.17), we finally obtain the hitting distribution for jump processes with $0 < \mu < 2$

$$f_{[0,x]}(y|0) \underset{\substack{x\to\infty \\ y>x}}{\sim} \frac{2^\mu\Gamma(1+\frac{\mu}{2})\sin(\frac{\pi\mu}{2})}{\pi}\left[\left(\frac{2y}{x} - 1\right)^2 - 1\right]^{-\frac{\mu}{2}}\frac{1}{y}\left[\frac{a_\mu}{x}\right]^{\frac{\mu}{2}}, \tag{7.18}$$

and we provide numerical illustration in figure 7.3. We stress that for $y > x$, trajectories that cross 0 before a do not contribute to $f_{[0,x]}(y|0)$, and the 4^{th} step of the matching method is exact.

7.2.2 $\mu = 2$ - Geometrical observables in $2D$ disks

We now turn our attention to higher dimensions, and consider the case of jump processes with $\mu = 2$ confined in a disk of radius R with fully absorbing boundary, as depicted in figure 7.1(a).

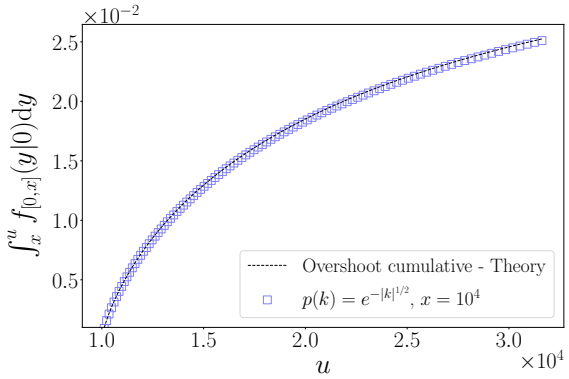


Figure 7.3: Cumulative hitting distribution for a Levy distributed heavy-tailed jump process starting from $x_0 = 0$, and fixed interval size x. The dashed line is obtained from equation (7.18).

Harmonic measure. In the large volume limit, the position of a particle issued from $(r_0, \theta_0 = \pi)$ upon crossing the absorbing disk surface is uniquely determined by the angular position θ (see figure 7.4(a)), whose distribution $\omega(r_0, \theta)$ is referred to as the harmonic measure. In the continuous limit $R - r_0 \gg a_\mu$ (*ie* bulk initial conditions), the distribution of θ is given by the corresponding Brownian harmonic measure $\omega^{(c)}(r_0, R, \theta)$ in a disk [Redner 2001]:

$$\omega^{(c)}(r_0, R, \theta) = \frac{1}{2\pi}\left[\frac{1 - \frac{r_0^2}{R^2}}{1 + \frac{2r_0}{R}\cos(\theta) + \frac{r^2}{R^2}}\right]. \tag{7.19}$$

Note that since $\theta_0 = \pi$, the harmonic measure is minimal at $\theta = 0$. However, as $r_0 \to R$, equation (7.19) vanishes, and the continuous description fails to capture the harmonic measure for initial positions close to the absorbing boundary. Following the matching method (7.13), the asymptotic harmonic measure is given by

$$\omega(r_0, \theta) \underset{a_2 \ll R}{\sim} \pi^{\perp}_{0,\underline{a}}(R - r_0) \lim_{u \to 0}\left[\frac{\omega^{(c)}(R - u, R, \theta)}{\pi^{(c)}_{0,\underline{a}}(u)}\right], \tag{7.20}$$

where a is an arbitrary lengthscale chosen such that $a_2 \ll a \ll R$. From equation (7.19), we obtain the small u behavior of the continuous harmonic measure, and the continuous Brownian splitting probability is known:

$$\begin{aligned} \omega^{(c)}(R - u, R, \theta) &\underset{u \to 0}{\sim} \frac{1}{2\pi}\frac{u}{R}\left[\frac{1}{1 + \cos(\theta)}\right], \\ \pi^{(c)}_{0,\underline{a}}(u) &\underset{u \to 0}{\sim} \frac{u}{a}. \end{aligned} \tag{7.21}$$

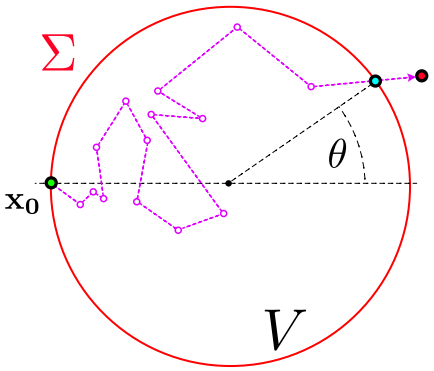
Focusing on the specific case $r_0 = R$, the splitting probability of the projected jump process reads

$$\pi^{\perp}_{0,\underline{a}}(0) \underset{a \gg a_2}{\sim} \frac{a_2^{\perp}}{a}, \tag{7.22}$$

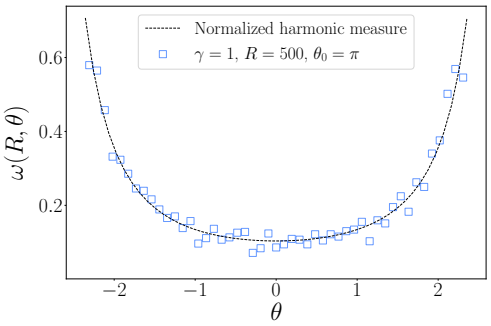
where we recall that $a_2^\perp = a_2/\sqrt{2}$. Finally, we obtain the asymptotic harmonic measure of the jump process starting from the boundary of the domain:

$$\omega(R,\theta) \underset{R\to\infty}{\sim} \frac{1}{2\pi}\frac{a_2^\perp}{R}\left[\frac{1}{1+\cos(\theta)}\right], \tag{7.23}$$

and provide numerical illustration in figure 7.4(b). Importantly, the harmonic measure (7.23) is still minimal for $\theta = 0$, but diverges for $\theta = \theta_0 = \pi$. Indeed, in the large volume limit, most trajectories escaping in the vicinity of the starting point $(R,\, \theta_0 = \pi)$ do not penetrate the bulk of the domain. In turn, the 4$^{\text{th}}$ hypothesis of the matching method is violated, and (7.20) does not hold in the case $\theta_0 = \pi$.



(a) Isotropic jump process in a disk with a fully absorbing boundary. Starting from $(R, \theta_0 = \pi)$, this sample trajectory intersects the disk at an angle θ upon its first escape, contributing to $\omega(R,\theta)$.



(b) Harmonic measure of an RTP starting from the boundary of a disk. The dashed line corresponds to equation (7.23). Note that the harmonic measure is normalized between $[-\frac{3\pi}{4}, \frac{3\pi}{4}]$.

Figure 7.4

Splitting probabilities in eccentric disks. In the context of preferential target search, geometrical observables quantify the probability that among a group of targets, a specific one is reached first. We here consider the case of a jump process with $\mu = 2$ evolving between two eccentric disks $\mathcal{D}_1$ and $\mathcal{D}_2$ of respective centers $z_i \in \mathbb{C}$ and radii R_i (see figure 7.5(a) for a schematic of the situation), and compute the probability to cross the outer rim first. We emphasize that this geometry is particularly relevant in a biophysical context, to describe transport from and to the nuclear and cellular surfaces [Stana & Lythe 2022] for instance.

Without loss of generality, we consider that the particle starts on the smaller disk of radius R_1, at position $z_0 = z_1 + R_1 e^{i\theta_0}$, and we denote $\pi_{R_1,\underline{R_2}}(z_0)$ the splitting probability to strictly cross the outer disk before the inner one. We now straightforwardly apply the matching method (7.13) to obtain:

$$\pi_{R_1,\underline{R_2}}(z_0) \underset{R_2\gg R_1\gg a\gg a_2}{\sim} \pi^\perp_{0,\underline{a}}(0) \lim_{u\to z_0}\left[\frac{\pi^{(c)}_{R_1,\underline{R_2}}(u)}{\pi^{(c)}_{0,\underline{a}}(u)}\right] \tag{7.24}$$

with $u \in \mathbb{C}$. Importantly, the limit $u \to z_0$ has to be taken perpendicularly to the H_1 plane, *ie* with fixed θ. Rewriting $u = z_0 + \varepsilon e^{i\theta_0}$ with $\varepsilon \in \mathbb{R}$, equation (7.24) is recast as

$$\pi_{R_1,\underline{R_2}}(z_0) \underset{R_2\gg R_1\gg a\gg a_2}{\sim} \pi^{\perp}_{0,\underline{a}}(0)\lim_{\varepsilon\to 0}\left[\frac{\pi^{(c)}_{R_1,\underline{R_2}}(z_0+\varepsilon e^{i\theta_0})}{\pi^{(c)}_{0,\underline{a}}(z_0+\varepsilon e^{i\theta_0})}\right]. \tag{7.25}$$

In the small ε limit, the continuous hyperplane splitting probability and projected splitting probability are given by

$$\begin{aligned}\pi^{\perp}_{0,\underline{a}}(0) &\underset{a\gg a_2}{\sim} \frac{a_2^{\perp}}{a},\\ \pi^{(c)}_{0,\underline{a}}(z_0+\varepsilon e^{i\theta}) &= \frac{\varepsilon}{a},\end{aligned} \tag{7.26}$$

such that $\pi^{\perp}_{0,\underline{a}}(0)/\pi^{(c)}_{0,\underline{a}}(z_0+\varepsilon e^{i\theta_0}) \sim a_2^{\perp}/\varepsilon$. Additionally, the continuous splitting probability to the outer rim vanishes linearly with ε

$$\pi^{(c)}_{R_1,\underline{R_2}}(z_0+\varepsilon e^{i\theta_0}) \underset{\varepsilon\to 0}{\sim} \varepsilon f_{\theta_0}, \tag{7.27}$$

where f_{θ_0} is a θ_0, R_1 and R_2 dependent constant. Finally, making use of equation (7.26) and (7.27), we find that the splitting probability starting from the inner disk is given by

$$\pi_{R_1,\underline{R_2}}(z_0) \underset{R_2\gg R_1\gg a\gg a_2}{\sim} \pi^{(c)}_{R_1,\underline{R_2}}(z_0+a_2^{\perp}e^{i\theta_0}). \tag{7.28}$$

As a result, the evaluation of the Brownian eccentric splitting probability $\pi^{(c)}_{R_1,\underline{R_2}}(z)$ for arbitrary initial position z is sufficient to determine $\pi_{R_1,\underline{R_2}}(z_0)$.

1. To compute $\pi^{(c)}_{R_1,\underline{R_2}}(z)$ explicitly, we first introduce the inter-disk distance $l = |z_2 - z_1|$ and define the constant $A = l/\sqrt{(R_2^2-R_1^2)^2 - 2l^2(R_1^2+R_2^2)+l^4}$. In turn, the conformal transformation

$$w = \frac{z}{1+Az} \tag{7.29}$$

maps the eccentric region to a concentric region between two disks of radii $\tilde{R}_2 > \tilde{R}_1$ (see [Chen *et al.* 2009] for more details), given by

$$\begin{aligned}\tilde{R}_2 &= \frac{\sqrt{1+4R_2^2a^2}-1}{2R_2a^2}\\ \tilde{R}_1 &= \frac{\sqrt{1+4R_1^2a^2}-1}{2R_1a^2}.\end{aligned} \tag{7.30}$$

2. Next, since the Laplace equation satisfied by $\pi^{(c)}$ is invariant by conformal transformation, if $w \in \mathbb{C}$ is the image of z through the conformal mapping, $\pi^{(c)}_{R_1,\underline{R_2}}(z) = \pi^{(c)}_{\tilde{R}_1,\underline{\tilde{R}_2}}(w)$.

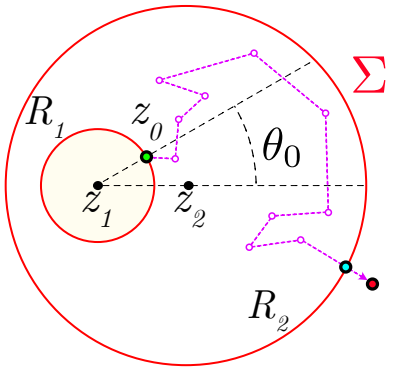
3. Finally, in the concentric geometry, the splitting probability only depends on the initial radius $r \in [\tilde{R}_1, \tilde{R}_2]$, and the splitting probability is known [Redner 2001]:

$$\pi^{(c)}_{\tilde{R}_1,\underline{\tilde{R}_2}}(r) = \frac{\log(r) - \log(\tilde{R}_1)}{\log(\tilde{R}_2) - \log(\tilde{R}_1)}. \tag{7.31}$$

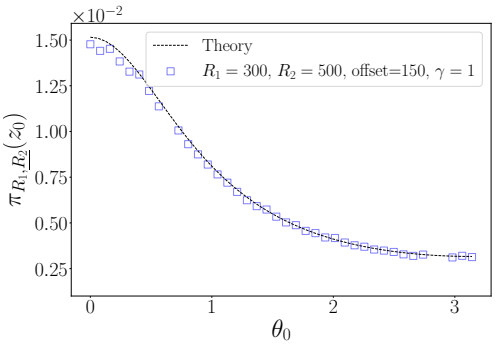
Applying this procedure, we obtain the following exact expression for the splitting probability of general jump processes starting from the inner disk at position $z_0 = z_1 + R_1 e^{i\theta_0}$:

$$\pi_{R_1,\underline{R_2}}(z_0) \underset{R_2,R_1 \gg a \gg a_2}{\sim} \frac{\log\left(\left\| \frac{z_0 + a_2^{\perp} e^{i\theta_0}}{1 + A(z_0 + a_2^{\perp} e^{i\theta_0})} \right\|\right) - \log(\tilde{R}_1)}{\log(\tilde{R}_2) - \log(\tilde{R}_1)}, \tag{7.32}$$

and display numerical agreement in figure (7.5).



(a) Eccentric disks $\mathcal{D}_1$ and $\mathcal{D}_2$ of respective radii and centers R_i and z_i. Starting on the boundary of the inner disk at position $z_0 = z_1 + R_1 e^{i\theta_0}$, the sample trajectory crosses the outer rim before the inner one, contributing to $\pi_{R_1,\underline{R_2}}(z_0)$.



(b) Splitting probability of a RTP in an eccentric annulus, starting from the inner rim at position $z_0 = z_1 + R_1 e^{i\theta_0}$. The two disk centers are offset by a distance 50. The dashed line is obtained by following the procedure outlined above.

Figure 7.5

Constant-speed tumbling processes. As a concluding remark, we emphasize that all the results obtained for geometrical observables of jump processes are directly applicable to isotropic continuous time processes with a constant speed v and random reorientation events.

As an example, consider the harmonic measure of a 2D Run and Tumble particle with tumbling rate γ and constant speed v. The length of each flight is given by $l = v\tau$, with τ the random flight duration, and it is easily seen that the characteristic flight length a_2 is given by $a_2 = v/\gamma$. As a result, we obtain from equation (7.23) the asymptotic harmonic measure of a Run and Tumble particle starting from $(R, \theta_0 = \pi)$:

$$\omega(R,\theta) \underset{R \gg \frac{v}{\gamma}}{\sim} \frac{1}{2\pi} \frac{v}{\sqrt{2}\gamma R} \left[\frac{1}{1+\cos(\theta)}\right]. \tag{7.33}$$

7.3 First-passage times

So far we only considered first-passage observables that depend solely on the position of the particle upon exit. However, the matching method is naturally extended to deal with first-passage times through the absorbing surface. We first focus on exit-times from one-dimensional intervals,

studied at length in chapter 5, and show that the matching method allows to go beyond the asymptotic results obtained for the RETP and LETP.

7.3.1 Mean exit-times from intervals

In this subsection, we consider the asymptotic mean exit-time $t(x|x_0)$ from the interval $[0, x]$, for an arbitrary jump process starting from x_0. Note that the asymptotic rightward and leftward exit-time probabilities have been investigated extensively in chapter 5. However, we have shown in table 5.1 that for heavy-tailed jump processes, the RETP and LETP are expressed in terms of unknown scaling functions h_μ and g_μ, such that one cannot explicitly evaluate $t(x|x_0)$.

Universal asymptotic behavior. To derive the asymptotic behavior of the mean exit-time, we resort to the matching method, and consider hereafter $x_0 = 0$. As we have seen in the previous section, the one-dimensional continuous splitting probability for limit processes $\pi^{(c)}_{0,\underline{a}}(u)$ is known for all $\mu \leq 2$, and we have derived the splitting probability from zero $\pi^{\perp}_{0,\underline{a}}(0)$ for all jump processes in chapter 4. Additionally, the mean exit-time from an interval for α-stable processes is known [Getoor 1961], and given by

$$t^{(c)}(x|x_0) = \sqrt{\pi}\left[2^\mu \Gamma\left(1+\frac{\mu}{2}\right)\Gamma\left(\frac{1+\mu}{2}\right)\right]^{-1}(x_0(x-x_0))^{\frac{\mu}{2}}\,. \tag{7.34}$$

Of note, this result is also true for Brownian motion with diffusion coefficient $D = 1$. In turn, the asymptotic mean exit-time for arbitrary jump processes is straightforwardly obtained from equation (7.13) and reads

$$\begin{aligned} t(x|0) &\underset{x\to\infty}{\sim} \pi^{\perp}_{0,\underline{a}}(0)\lim_{u\to 0}\left[\frac{t^{(c)}(x|u)}{\pi^{(c)}_{0,\underline{a}}(u)}\right], \\ &\underset{x\to\infty}{\sim} \frac{2^{-\mu}\sqrt{\pi}}{\Gamma\left(\frac{1+\mu}{2}\right)}\left[\frac{x}{a_\mu}\right]^{\frac{\mu}{2}}. \end{aligned} \tag{7.35}$$

We stress that equation (7.35) is valid for all jump processes with $\mu \leq 2$. Although the x scaling was expected, we identify the exact prefactor, and uncover a fully explicit asymptotic universal behavior. Importantly, we emphasize that $t(x|0)$ scales as $x^{\mu/2}$, much slower than the typical time $n_x \propto x^\mu$ needed to approach x, which we introduced in the previous chapters. This effect is mostly due to the fact that many trajectories contributing to $t(x|0)$ never approach x, and exit leftwards through 0. Finally, note that mean first exit-times should be understood as mean number of steps, as indicated by the dimensionality of equation (7.35).

A short comment on the matching method. In the 4^{th} hypothesis of the matching method, we neglect all trajectories that cross 0 before the arbitrary lengthscale a. This hypothesis is here validated a posteriori; in the limit $x \gg a \gg a_\mu$, it can be seen from equation (7.35) that the scaling with a of the mean exit-time of trajectories that exit through 0 without crossing a is bounded above by $a^{\mu/2}$. In turn, these trajectories have a sub-leading contribution to $t(x|0)$, which scales as $x^{\mu/2}$.

Illustrations. We now illustrate the asymptotic result (7.35) on both analytical and numerical examples. First, we consider the exactly solvable case of the exponential jump process with $p(\ell) = \frac{\gamma}{2} e^{-\gamma|\ell|}$, for which $\mu = 2$ and $a_2 = \gamma^{-1}$. The associated mean exit-time $t_1(x|x_0)$ from $[0, x]$ obeys the following backward equation:

$$t_1(x|x_0) = 1 + \int_0^x t_1(x|y) \frac{\gamma}{2} e^{-\gamma|y-x_0|} \mathrm{d}y, \tag{7.36}$$

which can be solved (see [Van Kampen 1992] for instance) to obtain the exact mean exit-time, valid for all x and x_0 values:

$$t_1(x|x_0) = -\frac{x_0^2\gamma^2}{2} + \frac{x x_0 \gamma^2}{2} + 1 + \frac{x\gamma}{2}. \tag{7.37}$$

As a result, in the large x limit, the mean exit-time $t_1(x|0)$ starting from 0 is asymptotically given by

$$t_1(x|0) \underset{x\to\infty}{\sim} \frac{x\gamma}{2}, \tag{7.38}$$

in agreement with the asymptotic result (7.35). In the case $\mu < 2$, there does not seem to be a process for which the integral equation (7.36) is exactly solvable. Consequently, we provide numerical estimates of $t(x|0)$ for a variety of heavy-tailed jump processes in figure 7.6, and display agreement with equation (7.35).

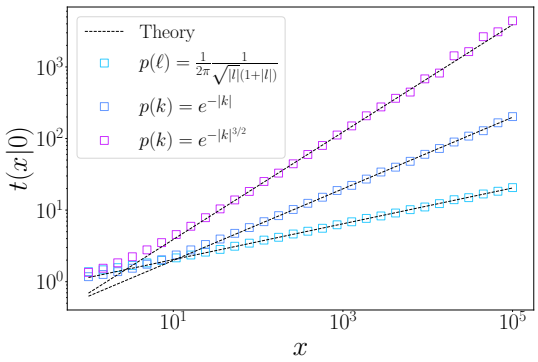


Figure 7.6: Mean exit-time for three heavy-tailed jump processes. For large x, the mean exit-time converges towards the asymptotic behavior (7.35). Of note, the F-distributed process with $p(\ell) \propto (\sqrt{|l|}(1 + |\ell|))$ has $\mu = 1/2$ and $a_\mu = 2/\pi$.

7.3.2 Mean exit-times from disks - recovering the Kac formula

By computing the harmonic measure and eccentric disk splitting probabilities, we demonstrated that the matching method is relevant to investigate geometrical observables in dimensions greater than one. In this subsection, we show that the matching method is equally applicable to address higher dimensional first-passage time observables, and focus on mean exit-times from disks with fully absorbing boundaries. Additionally, we showcase the flexibility of the method in dealing with jump processes whose first jump is not drawn from the same distribution as the subsequent jumps.

Kac formula. Specifically, we consider a Run and Tumble particle with speed v_0 and tumbling rate γ in a 2D disk of radius R. The particle starts on the boundary of the disk at position $(R, \theta_0 = \pi)$ (see figure 7.4(b)) and evolves until it strictly escapes the disk. In turn, the Kac mean exit-time $t_{KAC}(R)$ is known to be γ independent [Blanco & Fournier 2003, Bénichou *et al.* 2005, Savo *et al.* 2017], and given exactly by the so-called Kac formula:

$$t_{KAC}(R) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \mathrm{d}\theta \cos(\theta) t(R|\theta) = \frac{R}{2v_0}. \tag{7.39}$$

Importantly, $t(R|\theta)$ is the conditional mean exit-time of the Run and Tumble particle with initial speed $\mathbf{v_0} = v_0\left[\cos(\theta)\mathbf{e_x} + \sin(\theta)\mathbf{e_y}\right]$.

Matching method with specific first jump distribution. We now determine the asymptotic large R conditional mean exit-time $t(R|\theta)$ for arbitrary jump processes with $\mu = 2$, and show that applying our result to the Run and Tumble particle we recover the Kac Formula.

In the large volume limit $R \gg a_2$, only the discrete part $\pi^{\perp}_{0,\underline{a}}(0)$ of equation (7.13) is impacted by the distribution of the first jump, and we rewrite the conditional mean exit-time as:

$$t(R|\theta) \underset{R\gg a_2}{\sim} \pi^1_{H_1,\underline{H_2}}(\theta) \lim_{\varepsilon\to 0}\left[\frac{t^{(c)}(R|R-\varepsilon)}{\pi^{(c)}_{0,\underline{a}}(\varepsilon)}\right], \tag{7.40}$$

where $\pi^1_{H_1,\underline{H_2}}(\theta)$ is the splitting probability to cross the hyperplane H_2 before H_1, conditioned on the first jump being taken with an angle θ (see figure 7.7 for a schematic of the situation). As usual, $t^{(c)}(R|r)$ is the Brownian mean exit-time from the disk starting from radius r, and $\pi^{(c)}_{0,\underline{a}}(\varepsilon)$ is the Brownian one-dimensional splitting probability to a.

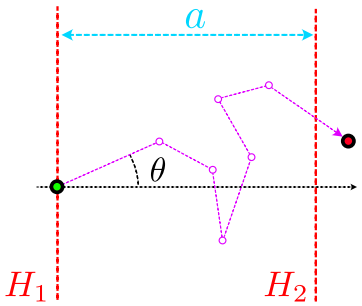


Figure 7.7: Sample trajectory contributing to the conditional hyperplane splitting probability $\pi^1_{H_1,\underline{H_2}}(\theta)$. The direction of the first jump is fixed at angle θ, and subsequent jumps are taken isotropically until H_2 is crossed before H_1.

Both $t^{(c)}(R|r)$ and $\pi^{(c)}_{0,\underline{a}}(\varepsilon)$ are known and can be found in [Redner 2001], and we obtain

$$\lim_{\varepsilon\to 0}\left[\frac{t^{(c)}(R|R-\varepsilon)}{\pi^{(c)}_{0,\underline{a}}(\varepsilon)}\right] = \frac{Ra}{2D}, \tag{7.41}$$

In turn, we compute $\pi^1_{H_1,\underline{H_2}}(\theta)$ by partitioning over the length of the first jump, which we recall is distributed according to $p(\ell)$:

$$\pi^1_{H_1,\underline{\underline{H_2}}}(\theta) = \int_0^\infty p(\ell)\pi^\perp_{0,\underline{a}}(\ell\cos(\theta))\mathrm{d}\ell, \tag{7.42}$$

such that the conditional mean exit-time is finally given by

$$t(R|\theta) \underset{R\gg a_2}{\sim} \frac{Ra}{2D}\int_0^\infty p(\ell)\pi^\perp_{0,\underline{a}}(\ell\cos(\theta))\mathrm{d}\ell. \tag{7.43}$$

We now show that in the case of the Run and Tumble particle, making use of equation (7.43) obtained from the matching method correctly yields the Kac Formula. In this specific case, the flight length is given by $\ell = v_0\tau$, where the flight duration τ is distributed according to $p(\tau) = \gamma e^{-\gamma\tau}$. In turn, the corresponding diffusion coefficient and characteristic lengthscale a_2 are given by $D = v_0^2/(2\gamma)$ and $a_2 = v_0/\gamma$. Plugging equation (7.43) into (7.39), we obtain the asymptotic the Kac mean exit-time:

$$\begin{aligned}
\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \mathrm{d}\theta\cos(\theta)t(R|\theta) &\underset{R\to\infty}{\sim} \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \mathrm{d}\theta\cos(\theta)\frac{Ra}{2D}\int_0^\infty \gamma e^{-\gamma\tau}\pi^\perp_{0,\underline{a}}(\cos(\theta)v_0\tau)\mathrm{d}\tau \\
&\underset{R\to\infty}{\sim} \frac{Ra\gamma}{2Dv_0\pi}\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^\infty e^{-\frac{\gamma\tau}{v_0\cos(\theta)}}\pi^\perp_{0,\underline{a}}(\ell)\mathrm{d}\ell \\
&\underset{R\to\infty}{\sim} \frac{R}{2}\frac{\sqrt{2}\gamma}{v_0^2\pi}\int_0^{\frac{\pi}{2}}\tilde{U}\left(\frac{\gamma}{v_0\cos(\theta)}\right)\mathrm{d}\theta,
\end{aligned} \tag{7.44}$$

where $\tilde{U}(s)$ is defined as

$$\tilde{U}(s) = \frac{v_0}{\sqrt{2}\gamma a}\int_0^\infty e^{-s\ell}\pi^\perp_{0,\underline{a}}(\ell)\mathrm{d}\ell. \tag{7.45}$$

For Run and Tumble particles, the Fourier transform of the projected jump distribution is given by $\tilde{p}_\perp(k) = 1/\sqrt{1+[kv_0\gamma^{-1}]^2}$, and we showed in equation (4.36) of chapter 4 that in the limit $a \gg a_2$

$$\tilde{U}\left(\frac{\gamma}{v_0\cos(\theta)}\right) = \frac{v_0\cos(\theta)}{\gamma}\exp\left[-\frac{1}{\cos(\theta)\pi}\int_0^\infty \frac{\mathrm{d}k}{\frac{1}{\cos^2(\theta)}+k^2}\log\left(1-\frac{1}{\sqrt{1+k^2}}\right)\right]. \tag{7.46}$$

Consequently, the Kac mean exit-time is asymptotically given by:

$$\frac{1}{\pi}\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\cos(\theta)t(R|\theta) \underset{R\to\infty}{\sim} \frac{R}{2v_0}\frac{\sqrt{2}}{\pi}\int_0^{\frac{\pi}{2}}\cos(\theta)\exp\left[-\frac{1}{\cos(\theta)\pi}\int_0^\infty \frac{\mathrm{d}k}{\frac{1}{\cos^2(\theta)}+k^2}\log\left(1-\frac{1}{\sqrt{1+k^2}}\right)\right]\mathrm{d}\theta \tag{7.47}$$

and we finally obtain

$$\frac{1}{\pi}\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\cos(\theta)t(R|\theta) \underset{R\to\infty}{\sim} \frac{R}{2v_0}, \tag{7.48}$$

in agreement with the Kac formula (7.39), which is valid for all R and v_0. In conclusion, we demonstrated that the matching method is well suited to address higher dimensional first-passage time observables and is easily adapted to take into account the specific initial conditions of a given jump process. In particular, we emphasize that the key ingredient to evaluate first-passage observables of jump processes is the determination of the hyperplane splitting probability $\pi_{H_1,\underline{H_2}}$, which relies heavily on the one-dimensional splitting probability. Moreover, it is worth stressing that the Kac mean first-passage time only serves as an example of specific initial conditions, highlighting the broad range of applicability of the matching method.

7.3.3 $\mu = 2$ - Complete distributions of first-passage times in the large volume limit

As a final application example, we consider asymptotic first-passage time distributions of jump processes with $\mu = 2$.

2D disk with fully absorbing surface. While the Kac formula only focuses on the mean first absorption time of a Run and Tumble particle after desorbtion from the surface of a disk, we show that the matching method allows for the derivation of the complete first-passage time distribution $f_{\underline{R}}(n)$ for arbitrary jump processes with $\mu = 2$, starting from the disk boundary. We emphasize that such distributions cannot be obtained by a direct continuous approximation. Introducing the FPT distribution $f^{(c)}_{\underline{R}}(t,r)$ of a Brownian particle starting from radius r and making use of the matching formula (7.13), we obtain

$$f_{\underline{R}}(n) \underset{\substack{R\gg a_2\\ n\gg 1}}{\sim} \pi^{\perp}_{0,\underline{a}}(0) \lim_{\varepsilon\to 0}\left[\frac{f^{(c)}_{\underline{R}}(n, R-\varepsilon)}{\pi^{(c)}_{0,\underline{a}}(\varepsilon)}\right]. \tag{7.49}$$

The continuous Brownian FPT distribution $f^{(c)}_{\underline{R}}(t,r)$ satisfies the following diffusion equation

$$\frac{\partial}{\partial t} f^{(c)}_{\underline{R}}(t,r) = D\Delta_r \frac{\partial}{\partial t} f^{(c)}_{\underline{R}}(t,r) \tag{7.50}$$

with boundary conditions $f^{(c)}_{\underline{R}}(t, r=R) = \delta(t)$, and can be explicitly computed in the Laplace transformed space:

$$\int_0^\infty e^{-st} f^{(c)}_{\underline{R}}(t,r)\mathrm{d}t = \tilde{f}^{(c)}_{\underline{R}}(s,r)\mathrm{d}t = \frac{I_0\left(\sqrt{\frac{sr^2}{D}}\right)}{I_0\left(\sqrt{\frac{sR^2}{D}}\right)}, \tag{7.51}$$

with I_k the modified Bessel function of order k. To obtain the first non vanishing order of $f^{(c)}_{\underline{R}}(t, R-\varepsilon)$ in the large R and t limit, we consider the distribution of the rescaled variable $\tilde{t} = t/R^2$. Rewriting $r = R(1-\varepsilon)$, we obtain the small ε behavior of the Laplace transformed distribution of $\tilde{t}$:

$$\int_0^\infty e^{-\tilde{s}\tilde{t}} f_{\underline{R}}^{(c)}(\tilde{t}, R(1-\varepsilon))\mathrm{d}\tilde{t} = \tilde{f}_{\underline{R}}^{(c)}(\tilde{s},\varepsilon) \underset{\substack{R\to\infty\\ \varepsilon\to 0}}{\sim} 1 - \varepsilon \frac{\sqrt{\tilde{s}} I_1\left(\sqrt{\frac{\tilde{s}}{D}}\right)}{\sqrt{D} I_0\left(\sqrt{\frac{\tilde{s}}{D}}\right)} + o(\varepsilon). \tag{7.52}$$

The constant term in the RHS of equation (7.52) corresponds to the Dirac delta boundary condition, such that we only consider the linearly vanishing term. In turn, making use of equation (7.49), we obtain the asymptotic large R Laplace transformed distribution $\tilde{f}_{\underline{R}}(\tilde{s})$ of the rescaled variable $\tilde{n} = n/R^2$:

$$\tilde{f}_{\underline{R}}(\tilde{s}) \underset{R\gg a_2}{\sim} \frac{a_2^\perp \sqrt{\tilde{s}} I_1\left(\sqrt{\frac{\tilde{s}}{D}}\right)}{R\sqrt{D} I_0\left(\sqrt{\frac{\tilde{s}}{D}}\right)}, \tag{7.53}$$

valid for any jump process with $\mu = 2$. In particular, we emphasize that our result holds for arbitrary constant-speed jump processes with speed v_0 and tumbling times distributed according to $p(\tau)$, for which the projected lengthscale and diffusion coefficient are respectively given by

$$a_2^\perp = \frac{v_0\sqrt{\langle\tau^2\rangle}}{2} \qquad D = \frac{v_0^2\langle\tau^2\rangle}{4\langle\tau\rangle}. \tag{7.54}$$

As a concluding example, we consider the case of the Run and Tumble particle starting from the boundary, with $p(\tau) = \gamma e^{-\gamma\tau}$, and display agreement between numerical simulations and numerical inversion of equation (7.53) in figure 7.8(a).

FPT distributions to interior targets. In a large confining domain $\mathcal{D}$ of volume V, characteristic lengthscale R, and reflecting boundary conditions (see figure 7.1(b) for a schematic of the situation), the FPT distribution of Brownian particles starting from $\mathbf{x_0}$ to a small spherical target of radius r has been shown to display universal behavior in the large volume limit $R \gg r$ [Bénichou *et al.* 2010b, Meyer *et al.* 2011]. More precisely, defining the mean first-passage time $t^{(c)}(\mathbf{x_0})$ from $\mathbf{x_0}$ to the target, and the global mean first-passage time $\langle\overline{T}\rangle = V^{-1}\int_{\mathcal{D}} t^{(c)}(\mathbf{u})\mathrm{d}\mathbf{u}$, the distribution $f^{(c)}(\theta|\mathbf{x_0})$ of the rescaled variable $\theta = t/\langle\overline{T}\rangle$ has been shown to be given by

$$f^{(c)}(\theta|\mathbf{x_0}) \underset{R\to\infty}{\sim} \left(1 - \frac{t^{(c)}(\mathbf{x_0})}{\langle\overline{T}\rangle}\right)\delta(\theta) + \frac{t^{(c)}(\mathbf{x_0})}{\langle\overline{T}\rangle} e^{-\theta}. \tag{7.55}$$

We emphasize that $t^{(c)}(\mathbf{x_0})$ vanishes as the source target distance goes to 0, such that equation (7.55) cannot be used directly to evaluate the distribution $f(n|0)$ of the FPT n of jump processes starting from the surface of the interior target. However, making use of the matching method, we extend this large volume framework to jump processes with $\mu = 2$, and compute the asymptotic distribution $f(\tilde{n}|0)$ of the rescaled variable $\tilde{n} = n/\langle\overline{T}\rangle$, in the limit $a_2 \ll r \ll R$.

Following the derivation of the FPT distribution to the fully absorbing disk, we straightforwardly apply the main equation (7.13) and obtain:

$$f(\tilde{n}|0) \underset{\substack{R\gg a_2\\ n\gg 1}}{\sim} \pi^{\perp}_{0,\underline{a}}(0) \lim_{\mathbf{x_0}\to\Sigma} \left[\frac{f^{(c)}(\theta|\mathbf{x_0})}{\pi^{(c)}_{0,\underline{a}}(\mathbf{x_0})}\right], \tag{7.56}$$

where Σ denotes the surface of the absorbing target. Importantly, to first order in the large volume limit, the direction in which we take $\mathbf{x_0} \to \Sigma$ is irrelevant. In turn, the evaluation of $t^{(c)}$ and $\langle\overline{T}\rangle$ as $\mathbf{x_0} \to \Sigma$ is sufficient to obtain the complete FPT distribution $f(\tilde{n}|0)$.

As an illustration, we consider a three-dimensional domain $\mathcal{D}$ of arbitrary shape. For Brownian motion, the large volume behavior of $t^{(c)}(\mathbf{x_0})$ and $\langle\overline{T}\rangle$ is given by [Condamin *et al.* 2007a]

$$t^{(c)}(\mathbf{x_0}) \underset{R\to\infty}{\sim} \frac{V}{4\pi D}\left(\frac{1}{r} - \frac{1}{r+\varepsilon}\right), \qquad \langle\overline{T}\rangle \underset{R\to\infty}{\sim} \frac{V}{D}\frac{1}{4\pi r}, \tag{7.57}$$

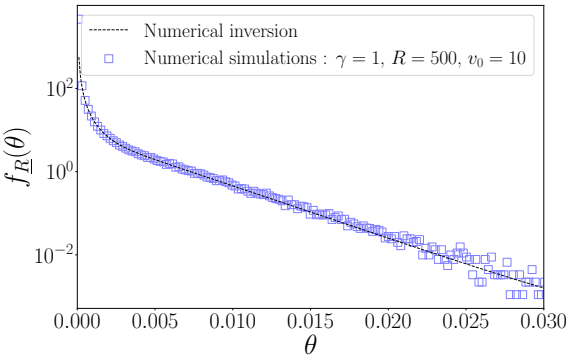
where r is the target radius, and ε the distance of the Brownian particle to the target surface. Importantly, these expressions only depend on the volume V, and not on the specific shape of the confining volume. First, note that for a jump process starting from the surface of the interior target, the asymptotic expression of $\langle\overline{T}\rangle$ is unchanged, and only the vanishing mean FPT $t^{(c)}(\mathbf{x_0})$ needs to be dealt with. In the limit $r \gg a_2$, the matching method applies, and the mean first-passage time vanishes linearly as $\varepsilon \to 0$

$$t^{(c)}(r+\varepsilon) \underset{\varepsilon\to 0}{\sim} \frac{\varepsilon V}{4\pi D r^2}. \tag{7.58}$$

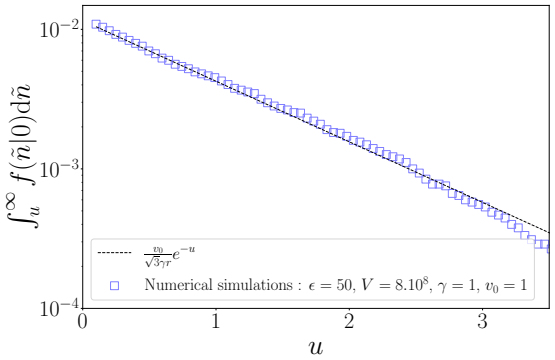
As a result, combining equation (7.56), (7.57) and (7.58), we obtain the distribution of the rescaled variable $\tilde{n}$ in the large R limit:

$$f(\tilde{n}|0) \underset{R\to\infty}{\sim} \left(1 - \frac{a_2^{\perp}}{r}\right)\delta(\tilde{n}) + \frac{a_2^{\perp}}{r}e^{-\tilde{n}}. \tag{7.59}$$

In conclusion, the matching method is fully compatible with existing large volume frameworks addressing first-passage observables for continuous processes, and allows to extend numerous asymptotic results to jump processes, while retaining the specificity arising from their discreetness. In this context, we have computed FPT distributions to interior targets for all initial position regimes, including the important *edge* initial conditions where the particle starts close to the absorbing target, and illustrate the general result (7.59) in figure 7.8(b), where we numerically estimate the FPT to a small target for a Run and Tumble particle in a large confining cubic domain.



(a) Distribution of the rescaled FPT $\theta = n/R^2$ to the boundary of a disk of radius R, for a Run and Tumble particle with $p(\tau) = \gamma e^{-\gamma\tau}$ and constant sped v_0. The particle starts from the absorbing boundary, and the theoretical dashed line comes from numerical inversion of equation (7.53).



(b) Cumulative distribution of the rescaled FPT $\theta = t/\langle\overline{T}\rangle$ to a centered spherical target in a three-dimensional cubic domain for a Run and Tumble particle starting from the target, with $a_2^\perp = v_0/(\sqrt{3}\gamma)$. The dashed is computed from equation (7.59).

Figure 7.8

7.4 Conclusion

Building upon results developed in chapter 4, the matching method allows to properly bind together known results for continuous stochastic processes and the particular behavior arising from the microscopic details of discrete time jump processes. As a result, we obtain a variety of new asymptotic results in the large confining volume limit.

We first investigate geometrical observables. We consider one-dimensional hitting distributions for heavy-tailed processes, and turn to disk geometries in the $\mu = 2$ case. Importantly, our two dimensional results for the harmonic measure and splitting probabilities of Run and Tumble particles go beyond the simple Brownian description, capturing the specifics of partly ballistic trajectories.

The matching method naturally extends to first-passage time observables. We compute mean exit-times from intervals, complementing the results obtained for exit-time probabilities in chapter 5. Additionally, we showcase the flexibility of the method in dealing with varying initial conditions by recovering the Kac formula for Run and Tumble particles in disks. Finally, we exhibit the compatibility of the matching method with existing large volume results for continuous processes, by computing the complete FPT distribution to a small target in a large confining domain of arbitrary shape.

We emphasize that the results presented in this chapter are only representative application examples of the matching method, spanning various geometries and jump processes, with experimental relevance to biology [Tailleur & Cates 2008] or optics [Baudouin *et al.* 2014, Savo *et al.* 2017]. In turn, we hope that the comprehensive framework of the matching method will be beneficial for researchers in various fields to explain and interpret non-trivial phenomena observed in empirical time series.

Conclusion

The description and interpretation of a physical phenomenon by a random transport problem involves two main aspects: the determination of a stochastic process accounting for the empirical random dynamics, and the choice of a statistical observable that addresses a specific physical question. For example, the spread of an epidemic can be modeled by a random walk on a graph (stochastic process), and the time for the complete contamination of the population corresponds to the time taken by the walker to visit all the nodes (observable). Similarly, the length of a queue can be described by a Poisson process (stochastic process); in turn, the probability that the cashier ever goes home is given by the probability of this process eventually reaching 0 (observable).

This thesis is divided into two parts, with each part corresponding to one of these two aspects. In the first part, we introduced a new observable, the territory visited by a confined random walker before reaching a given target, which quantifies the efficiency of the space exploration process, and quantitatively characterized its statistics for the broadest possible class of stochastic processes. In a second part, we focused on a single type of stochastic process, one-dimensional jump processes in confinement, which have been studied for some time but with limited analytical results. In this regard, we proposed a method to provide a variety of general results in the large confining volume limit.

The number of distinct sites visited by a random walker after n steps, which quantifies the ability of the walker to explore an unknown territory, has been mostly studied in the infinite geometry. Indeed, in a bounded geometry, the confining domain is eventually fully visited. However, the question of what portion of a closed domain is visited by a walker before escaping through a fixed exit is of particular interest in chemical and biological applications. To answer this question, we introduced the observable $C(s_0)$, defined as the territory visited by a random walker starting from s_0 before reaching a fixed target s_T.

We first considered the class of 1D Markovian processes with connected span, which includes many classical lattice random walks: the symmetric and biased normal walk, the persistent walk; but also continuous processes such as Brownian Motion with and without resetting. By establishing an exact relation between the territory visited before reaching the target located at 0, and the splitting probability $\pi_{s_1,\underline{s_2}}(s_0)$, defined as the probability that the walker reaches s_2 before s_1, we determined the exact distribution of the territory visited for numerous processes.

To go beyond the one-dimensional and connected span hypothesis, we then considered the more general class of scale-invariant processes, in the large confining volume limit. In particular, this includes nearest neighbor diffusion on deterministic and random fractal networks, but also processes with long range jumps, such as Riemann walks. By exploiting the asymptotic behavior of splitting probabilities for scale invariant processes in the large volume limit, we identified a universal scaling behavior of the first moment $\langle C(s_0)\rangle$ of the explored territory, but also of its complete distribution. In particular, we fully characterized the dependence of the visited territory on the geometrical parameters of the system, namely the source-target distance, as well as the size of the confining domain.

Finally, it is clear that the territory visited before finding the target s_T is intrinsically linked

to the time needed to reach s_T. To quantify precisely the intuitive correlations between territory and time, we considered the joint distribution $\sigma(s, n|s_0)$ of the visited territory s and the FPT n to s_T. In the case of 1D Markovian processes with connected span, we developed a methodology to explicitly compute σ, and applied it to the above given processes. Moreover, in the limit s and n large, we showed that the joint law admits a universal scaling form, valid for all one-dimensional processes, whether they be Markovian and with connected span, or not. In particular, our results are valid for the emblematic fractional Brownian motion.

These initial results offer many exciting possibilities for future investigations. For example, does the universal behavior of the joint distribution of the territory and FPT still hold in dimensions greater than 1? If one expects to find universality classes akin to those governing the behavior of $C(s_0)$, the answer to this question is currently unknown and represents a challenging problem for future research. Similarly, can we say anything if the target is not fixed but randomly distributed in the system, or if multiple targets are allowed? These natural extensions are not only exciting from a theoretical point of view, but are also part of the broader perspective of modeling realistic phenomena and systems.

Our second area of interest involved investigating the behavior of a specific stochastic process, namely one-dimensional jump processes, which describe the position x_n of a particle undergoing random jumps distributed according to a symmetric distribution $p(\ell)$ at discrete time intervals. Importantly, these models, first introduced by Pearson in 1905, are especially useful for analyzing experimental or numerical data, which is discrete by essence, as continuous descriptions such as Brownian motion are unable to capture the effects of this discretization. Jump processes in infinite and semi-infinite geometry (ie stopped upon becoming negative) have been extensively studied: for instance the distribution of the position x_n can be expressed explicitly in terms of $p(\ell)$ in both cases. However, for bounded (or confined) jump processes, *ie* those that stop at the first exit of a finite interval $[0, x]$, there are no general results valid for any $p(\ell)$.

To gain a better understanding of these confined processes, we first focused on observables related to the exit of the interval, specifically the splitting probability $\pi_{0,\underline{x}}(x_0)$ which quantifies the likelihood of the process escaping through x rather than through 0. In the limit $x \to \infty$, we derived an explicit universal behavior for the splitting probability, which holds for any symmetric jump distribution. Our result provides a complete characterization of the splitting probability's dependence on the initial position of the particle and captures the specific effects of the process's discreteness. Notably, we showed that the splitting probability starting from $x_0 = 0$ depends only on the tails of the jump distribution, and is strictly non-zero, which is not possible for a continuous non-smooth process.

To obtain more fine-grained information on interval exit events, we investigated the leftward and rightward exit-times of the jump process within the confined domain $[0, x]$. To that end, we introduced the leftward exit-time probability (LETP) $F_{\underline{0},x}(n|x_0)$, defined as the probability that the particle exits the interval through 0 on the n^{th} step, and its rightward counterpart, the RETP $F_{0,\underline{x}}(n|x_0)$. In the large interval limit, we demonstrated that these quantities exhibit universal asymptotic behavior that depends solely on the tails of the distribution. We emphasize that our findings account for the discrete effects inherent in jump processes and hold for all symmetric jump distributions $p(\ell)$ and initial positions x_0, including in particular $x_0 = 0$.

Finally, the investigation of the splitting probability, LETP, and RETP allowed us to extend

our findings in two directions. Firstly, we examined extreme value statistics (EVS) of jump processes, which are typically associated with the distribution of extremums and the time taken to reach them. We demonstrated that the LETP, RETP, and splitting probability are essential components in determining EVS observables. In turn, we computed a variety of exact and asymptotic joint distributions of extremums and first-passage times for arbitrary $p(\ell)$. Secondly, we adapted our results to the case of isotropic jump processes in dimensions larger than one. To achieve this, we developed a methodology suitable for any jump distribution that allowed us to evaluate general first-passage observables in the large volume limit. For example, we computed the distribution of the time to reach the edge of a disk for a Run and Tumble particle initially located on the disk's edge, going beyond the classical Brownian motion approximation.

In summary, we have developed a methodology to evaluate the asymptotic behavior of observables associated with the exit of a confining domain for arbitrary symmetric jump processes in dimension 1 and higher. Our results are comprehensive, and we emphasize that the analysis of the tails of $p(\ell)$ is mostly sufficient to evaluate the asymptotic behavior of the chosen observable for any process. However, there are still many unanswered questions that we hope to explore in future research. One such question is the evaluation of first asymptotic corrections in the large confining volume limit, which might be achieved using the *big jump principle* [Vezzani *et al.* 2019], and would allow for a more detailed characterization of confined jump processes. Another direction that we are eager to explore is that of the asymmetry of the process. Is it possible to identify universal behaviors in the case where $p(\ell)$ is not symmetric anymore? We know that this is the case for the survival probability of semi-infinite processes [Majumdar *et al.* 2012], and hope to extend our methodological basis to the confined asymmetric case.

Overall, there is still much to be explored in the field of confined jump processes, and we hope our work can inspire further research in this area.

Appendices

Appendix A

Relocating random walk - first order perturbation theory

We devote this Appendix to the derivation of the distribution of $C(s_0)$ for relocating random walks with $\lambda \ll 1$. We compute a first order perturbation approximation by considering trajectories that relocate exactly once before reaching 0. Additionally, we neglect all trajectories where the *pre* and *post* relocation sets of visited sites overlap. In turn, we obtain a lower bound for $P(C = n|s_0)$.

For fixed n and s_0, the contributing trajectories contain two distinct parts: a first short diffusive motion where n_1 sites are visited, followed by a relocation event that takes the walker close to 0, where the last $n - n_1$ sites are visited before the target site is reached. The probability of such trajectories can be rewritten as follows:

$$\begin{aligned} P(C = n+1|s_0) = \sum_{n_1=1}^{n} & P(\text{span} = n_1 \text{ when the first jump occurs}|s_0)\times \\ & \frac{2}{N}\sum_{s=1}^{n-n_1} \mu(n-n_1 \quad \text{and no jump has occurred}|s) \end{aligned} \tag{A.1}$$

where $\mu(n_2$ and no jump has occurred$|s)$ is the probability that the maximum before reaching 0 is equal to n_2, and no jump occurred during the whole trajectory. Of note, the $\frac{2}{N}$ factor simply takes into account the fact that after relocating, the walker is either left or right of 0. We now introduce the various quantities needed to evaluate each term in equation (A.1).

- Let the propagator $G_{[0,s]}(s_1, n|s_0)$ denote the probability to be at position s_1 after n steps, starting from s_0, with absorbing sites at 0 and s. We additionally define its generating function, with $\xi \in [0, 1]$: $G_{[0,s]}(\xi, s_1|s_0) = \sum_{n=0}^{\infty} G_{[0,s]}(n, s_1|s_0)\xi^n$. Importantly, the propagator is a known quantity that can be found in [Hughes 1995]:

$$\begin{aligned} G_{[0,s]}(s_1, \xi|s_0) &= G_{[0,s]}(s_0, \xi|s_0)\frac{\alpha^{s_1-s} - \alpha^{s-s_1}}{\alpha^{-(s-s_0)} - \alpha^{s-s_0}} \\ G_{[0,s]}(s_0, \xi|s_0) &= \left[1 - \frac{\xi\alpha}{2}\left[\frac{1-\alpha^{2(s_0-1)}}{1-\alpha^{2s_0}} + \frac{1-\alpha^{2(s-s_0-1)}}{1-\alpha^{2(s-s_0)}}\right]\right]^{-1} \\ \alpha &= \xi^{-1}(1-\sqrt{1-\xi^2}) \end{aligned} \tag{A.2}$$

- Let $\pi^{nj}_{0,\underline{s}}(s_0)$ denote the probability to reach site s before site 0, without any relocating event along the trajectory. By partitioning over the time at which s is reached, $\pi^{nj}_{0,\underline{s}}(s_0)$ is given by:

$$\begin{aligned}\pi^{nj}_{0,\underline{s}}(s_0) &= \sum_{k=1}^{\infty} P(s \text{ is reached for the first time at time } k) P(\text{no jump occurs in } k \text{ steps}) \\ &= \frac{1-\lambda}{2} G_{[0,s]}(s-1, 1-\lambda|s_0).\end{aligned} \tag{A.3}$$

- Let $q_{\underline{0},\underline{s}}(n|s_0)$ denote the survival probability after n steps in the interval $[\![0, s]\!]$ with absorbing boundary conditions. $q_{\underline{0},\underline{s}}(n|s_0)$ is given in [Hughes 1995]:

$$q_{\underline{0},\underline{s}}(n|s_0) = \frac{1}{s} \sum_{|2j+1|<s} \cos^n\left(\frac{(2j+1)\pi}{s}\right) \sin\left(\frac{(2j+1)\pi s_0}{s}\right) \cot\left(\frac{(2j+1)\pi}{2s}\right). \tag{A.4}$$

- Let $\pi^{j}_{\underline{0},\underline{s}}(s_0)$ denote the probability to relocate before reaching either 0 or s:

$$\begin{aligned}\pi^{j}_{\underline{0},\underline{s}}(s_0) &= \sum_{k=1}^{\infty} q_{\underline{0},\underline{s}}(k-1|s_0) P(\text{a jump occurs on the } k^{\text{th}} \text{ step}) \\ &= \frac{\lambda}{s} \sum_{|2j+1|<s} \frac{\sin\left(\frac{(2j+1)\pi s_0}{s}\right) \cot\left(\frac{(2j+1)\pi}{2s}\right)}{1-(1-\lambda)\cos\left(\frac{(2j+1)\pi}{s}\right)}.\end{aligned} \tag{A.5}$$

We now rewrite all terms in equation (A.1) as functions of $G_{[0,s]}(s_1, n|s_0)$, $\pi^{nj}_{0,\underline{s}}(s_0)$, $q_{\underline{0},\underline{s}}(n|s_0)$ and $\pi^{j}_{\underline{0},\underline{s}}(s_0)$. Making use of the translational invariance of the normal walker we obtain:

$$\begin{aligned}&P(\text{span} = n_1 \text{ when the first jump occurs}|s_0) = \\ &\qquad \sum_{s=0}^{n_1-1} \pi^{nj}_{0,\underline{n_1}}(1) \pi^{j}_{0,\underline{n_1+1}}(1) \left(\pi^{nj}_{0,\underline{n_1-1}}(s) + \pi^{nj}_{0,\underline{n_1-1}}(n_1-1-s)\right),\end{aligned} \tag{A.6}$$

and

$$\mu(n-n_1 \text{ and no jump has occurred}|s) = \pi^{nj}_{0,\underline{n-n_1-1}}(s) \pi^{nj}_{0,\underline{n-n_1}}(1). \tag{A.7}$$

Equations (A.6) and (A.7) can easily be evaluated. In turn, we obtain a first order approximation for the distribution of the number of distinct sites visited in the rare relocation setup, and display numerical agreement on figure 1.5 of chapter 1.

Appendix B

Exact Pseudo Green functions

We devote this Appendix to exact expressions of Pseudo green functions for random walks on periodic lattices. These results can be found in [Bénichou & Voituriez 2014]. We recall that the pseudo green function $H(s_i, s_j) \equiv H_{ij}$ satisfies the following equation

$$\Delta_i H_{ij} = \delta_{ij} - \frac{1}{N}, \tag{B.1}$$

where Δ_i is the discrete Laplacian taking the periodicity into account, δ_{ij} the Kronecker delta, and N the total number of sites in the domain.

2D paralelepipedic domain The pseudo green function for a two dimensional periodic paralelepipedic domain of dimensions (X, Y) and N sites is given by

$$H_{ij} = \frac{1}{N} \sum_{m=0}^{X-1} \sum_{n=\delta_{m0}}^{Y-1} \frac{\exp\left[\frac{2im\pi(x_i-x_j)}{X} + \frac{2in\pi(y_i-y_j)}{Y}\right]}{1 - \frac{1}{2}\left[\cos\left(\frac{2m\pi}{X}\right) + \cos\left(\frac{2n\pi}{Y}\right)\right]}. \tag{B.2}$$

3D paralelepipedic domain The pseudo green function for three dimensional periodic paralelepipedic domain of dimensions (X, Y, Z) and N sites is given by:

$$H_{ij} = \frac{1}{N} \sum_{m=0}^{X-1} \sum_{n=0}^{Y-1} \sum_{p=\delta_{(m,n)(0,0)}}^{Z-1} \frac{\exp\left[\frac{2im\pi(x_i-x_j)}{X} + \frac{2in\pi(y_i-y_j)}{Y} + \frac{2ip\pi(z_i-z_j)}{Z}\right]}{1 - \frac{1}{3}\left[\cos\left(\frac{2m\pi}{X}\right) + \cos\left(\frac{2n\pi}{Y}\right) + \cos\left(\frac{2p\pi}{Z}\right)\right]}. \tag{B.3}$$

Appendix C

Three point splitting probabilities

We devote this Appendix to the derivation of 3 point splitting probabilities in the large volume limit for non-compact processes. Let s_1, s_2 and s_3 be three target sites in a confining domain Ξ, and denote $\pi_{\underline{s_1},s_2,s_3}(s_0)$ the splitting probability to reach s_1 before s_2 and s_3 starting from s_0. Denoting Δ_i the discrete Laplacian acting on variable i, $\pi_{\underline{s_1},s_2,s_3}(s_0)$ obeys the following equation:

$$\begin{aligned}
&\Delta_0 \pi_{\underline{s_1},s_2,s_3}(s_0) = 0 \\
&\pi_{\underline{s_1},s_2,s_3}(s_1) = 1 \\
&\pi_{\underline{s_1},s_2,s_3}(s_2) = 0 \\
&\pi_{\underline{s_1},s_2,s_3}(s_3) = 0
\end{aligned} \tag{C.1}$$

In terms of pseudo green functions H_{ij}, the 3 point splitting probability is given exactly by

$$\pi_{\underline{s_1},s_2,s_3}(s_0) = \frac{[H_{10}-H_{20}+H_{23}-H_{13}]\,[H_{12}-H_{32}+H_{33}-H_{13}]-[H_{10}-H_{30}+H_{33}-H_{13}]\,[H_{12}-H_{22}+H_{23}-H_{13}]}{[H_{11}-H_{21}+H_{23}-H_{13}]\,[H_{12}-H_{32}+H_{33}-H_{13}]-[H_{11}-H_{31}+H_{33}-H_{13}]\,[H_{12}-H_{22}+H_{23}-H_{13}]}. \tag{C.2}$$

Note that this expression can be directly checked by plugging into (C.1) and using the fact that $\Delta_i H_{ij} = \delta_{ij} - \frac{1}{N}$. Denoting r_{ij} the distance between s_i and s_j, we consider the large volume limit for which $r_{ij} \gg r_{01} \equiv r_s$ for all pairs $(i,j) \neq (0,1)$. In turn, for all pairs of sites far away from each other, we have $H_{ij} \sim H_{kj}$. As a result we simplify equation (C.2) and obtain:

$$\pi_{\underline{s_1},s_2,s_3}(s_0) \sim \frac{H_{10}-H_{30}+H_{11}-H_{13}}{H_{12}-H_{31}+3(H_{11}-H_{13})}, \tag{C.3}$$

where $\forall i,\ H_{ii} = H_{11}$. Recalling the exact expression for the two point splitting probability

$$\pi_{\underline{s_1},s_3}(s_0) = \frac{H_{10}-H_{30}+H_{11}-H_{13}}{2(H_{11}-H_{31})}, \tag{C.4}$$

we rewrite the asymptotic 3 point splitting probability as

$$\pi_{\underline{s_1},s_2,s_3}(s_0) \sim \pi_{\underline{s_1},s_3}(s_0)\frac{2(H_{11}-H_{31})}{H_{12}-H_{31}+3(H_{11}-H_{13})}. \tag{C.5}$$

Finally, for s_1, s_2 and s_3 far from each other, we asymptotically have $H_{12} \sim H_{13}$, so that (C.5) is simplified further:

$$\pi_{\underline{s_1},s_2,s_3}(s_0) \sim \frac{2}{3}\pi_{\underline{s_1},s_3}(s_0). \tag{C.6}$$

Appendix D

Joint statistics of space and time

We devote this Appendix to additional information regarding the derivations and simulations presented in chapter 3.

D.1 Joint distribution derivations

In the following, all generating functions and Laplace transforms are identified by the variable name (respectively ξ and p), to make notations less cumbersome.

D.1.1 Persistent random walk

We first derive the RETP $F_{0,\underline{s}}(n|s_0)$ for the persistent random walk. Similarly to the derivation of the splitting probability, we introduce $u_{0,\underline{s}}(n|s_0)$ and $v_{0,\underline{s}}(n|s_0)$ the RETP conditioned on the direction of the previous step (respectively rightward and leftward.) Denoting a the probability of taking the first step towards the right yields:

$$F_{0,\underline{s}}(n+1,s_0) = a u_{0,\underline{s}}(n|s_0+1) + (1-a) v_{0,\underline{s}}(n|s_0-1). \tag{D.1}$$

In the generating function formalism we obtain:

$$\frac{F_{0,\underline{s}}(\xi|s_0)}{\xi} = a u_{0,\underline{s}}(\xi|s_0+1) + (1-a) v_{0,\underline{s}}(\xi|s_0-1). \tag{D.2}$$

Dropping the ξ dependence for brevity, we also obtain a set of equations for u and v:

$$\begin{aligned} u_{0,\underline{s}}(s_0+2) - \frac{\frac{1}{\xi} - \xi + 2p\xi}{p} u_{0,\underline{s}}(s_0+1) + u_{0,\underline{s}}(s_0) &= 0 \\ v_{0,\underline{s}}(s_0) &= \frac{1}{1-p}\left[\frac{1}{\xi} u_{0,\underline{s}}(s_0+1) - p u_{0,\underline{s}}(s_0+2)\right]. \end{aligned} \tag{D.3}$$

Note that taking $\xi \to 1$ yields back the governing equations for the splitting probability to reach s before 0, obtained in chapter 1. Simplifying equation (D.3), we show that $u_{0,\underline{s}}$ obeys a second order difference equation whose roots read:

$$r_\pm = \frac{\frac{1}{\xi} - \xi + 2p\xi}{2p} \pm \frac{\sqrt{(\frac{1}{\xi} - \xi + 2p\xi)^2 - 4p^2}}{2p}, \tag{D.4}$$

such that

$$u_{0,\underline{s}}(s_0) = A r_+^{s_0} + B r_-^{s_0} \tag{D.5}$$

Finally, imposing the two boundary conditions $u_{0,\underline{s}}(s) = 1$ and $v_{0,\underline{s}}(0) = 0$, we obtain the following result:

$$\begin{aligned} A &= \frac{(p\xi r_- - 1)r_-}{r_-^s r_+ - p\xi r_-^s r_+^2 - r_- r_+^s + p\xi r_-^2 r_+^s} \\ B &= -\frac{(p\xi r_+ - 1)r_+}{r_-^s r_+ - p\xi r_-^s r_+^2 - r_- r_+^s + p\xi r_-^2 r_+^s}, \end{aligned} \tag{D.6}$$

and $F_{0,\underline{s}}(\xi|s_0)$ is obtained from equation (D.2). In turn, the LETP $F_{\underline{0},s}(\xi|s_0)$ is obtained by making use of the symmetry $s_0 \to s - s_0$ and $a \to 1 - a$, and the joint distribution $\sigma(s, n|s_0)$ can be computed by performing a small ξ expansion.

D.1.2 Resetting random walk

We derive $F_{0,\underline{s}}(n|s_0)$ for the discrete resetting process. Recall that, at each step, the walker either jumps back to its initial position s_p with probability λ or hops on one of its neighbors with probability $\frac{1-\lambda}{2}$. In the generating function formalism, $F_{0,\underline{s}}(\xi|s_0)$ obeys the following backward equation:

$$F_{0,\underline{s}}(\xi|s_0+1) - \frac{2}{\xi(1-\lambda)}F_{0,\underline{s}}(\xi|s_0) + F_{0,\underline{s}}(\xi|s_0-1) = \frac{2\lambda}{(1-\lambda)}F_{0,\underline{s}}(\xi|s_p) \tag{D.7}$$

with associated boundary conditions $F_{0,\underline{s}}(\xi|0) = 0$ and $F_{0,\underline{s}}(\xi|s) = 1$. Denoting $G(s_1, s_2)$ the Green function such that

$$G(s_1+1, s_2) - \frac{2}{\xi(1-\lambda)}G(s_1, s_2) + G(s_1-1, s_2) = \delta_{s_1,s_2}, \tag{D.8}$$

$G(s_1, s_2)$ is given by

$$G(s_1, s_2) = 1_{s_1 \le s_2} G_-(s_1, s_2) + 1_{s_1 > s_2} G_+(s_1, s_2) \tag{D.9}$$

where

$$\begin{aligned} G_-(s_1, s_2) &= A(s_2)^{-1}(r_+^{s_1} - r_-^{s_1}) \\ G_+(s_1, s_2) &= A(s_2)^{-1}\frac{r_+^{s_2} - r_-^{s_2}}{r_+^{s_2} - r_-^{s_2}\frac{r_+^s}{r_-^s}}\left(r_+^{s_1} - r_-^{s_1}\frac{r_+^s}{r_-^s}\right) \end{aligned} \tag{D.10}$$

with

$$r_\pm = \frac{1}{\xi(1-\lambda)} \pm \sqrt{\left[\frac{1}{\xi(1-\lambda)}\right]^2 - 1},$$

$$A(s_2) = \left(r_+^{s_2} - r_-^{s_2}\right)\left(r_+^{s_2} - r_-^{s_2}\frac{r_+^s}{r_-^s}\right)^{-1}\left(r_+^{s_2+1} - r_-^{s_2+1}\frac{r_+^s}{r_-^s}\right) - \frac{2}{\xi(1-\lambda)}\left(r_+^{s_2} - r_-^{s_2}\right) + \left(r_+^{s_2-1} - r_-^{s_2-1}\right). \tag{D.11}$$

Introducing the homogeneous solution

$$h_{0,\underline{s}}(s_0) = \frac{r_+^{s_0} - r_-^{s_0}}{r_+^{s} - r_-^{s}}, \tag{D.12}$$

we finally obtain the RETP for arbitrary s_p

$$F_{0,\underline{s}}(\xi|s_0) = h_{0,\underline{s}}(s_0) - \frac{h(s_p)2\lambda\sum_{s_2} G(s_0,s_2)}{1-\lambda+2\lambda\sum_{s_2} G(s_p,s_2)}, \tag{D.13}$$

which covers in particular the case $s_p = s_0$. Again, taking $\xi \to 1$ yields back the splitting probability derived in chapter 1. In turn, the LETP $F_{\underline{0},s}(\xi|s_0)$ is obtained by making use of the symmetry $s_0 \to s - s_0$, and the joint distribution $\sigma(s,n|s_0)$ can be computed by performing a small ξ expansion.

D.1.3 Brownian Motion with drift

Consider biased diffusion with diffusive coefficient D and rightward drift coefficient v. The continuous LETP $F_{\underline{0},x}(t|x_0)$ obeys the following backward equation:

$$\frac{\partial}{\partial t}F_{\underline{0},x}(t|x_0) - v\frac{\partial}{\partial x_0}F_{\underline{0},x}(t|x_0) = D\frac{\partial^2}{\partial x_0^2}F_{\underline{0},x}(t|x_0). \tag{D.14}$$

Introducing the time Laplace transform $F_{\underline{0},x}(p|x_0) = \int_0^\infty e^{-pt}F_{\underline{0},x}(t|x_0)\mathrm{d}t$ yields

$$pF_{\underline{0},x}(p|x_0) - v\frac{\partial}{\partial y}F_{\underline{0},x}(p|x_0) = D\frac{\partial^2}{\partial y^2}F_{\underline{0},x}(p|x_0),$$

and imposing the prescribed boundary conditions on $F_{\underline{0},x}(p|x_0)$ we obtain

$$F_{\underline{0},x}(p|x_0) = e^{\frac{v}{2D}(x-x_0)}\frac{\sinh\left((x-x_0)\frac{|v|}{2D}\sqrt{1+\frac{4pD}{v^2}}\right)}{\sinh\left(x\frac{|v|}{2D}\sqrt{1+\frac{4pD}{v^2}}\right)}.$$

In turn, making use of equation (3.11) from chapter 3, the joint distribution is given by

$$\sigma^v(x,p|x_0) = e^{\frac{-vx_0}{2D}}\frac{|v|}{2D}\sqrt{1+\frac{4pD}{v^2}}\frac{\sinh\left(\frac{x_0|v|}{2D}\sqrt{1+\frac{4sD}{v^2}}\right)}{\sinh\left(\frac{x|v|}{2D}\sqrt{1+\frac{4pD}{v^2}}\right)^2}. \tag{D.15}$$

In real space and time variables we obtain:

$$\sigma^v(x,t|x_0) = e^{\frac{-vx_0}{2D}}e^{\frac{-v^2t}{4D}}\frac{2D\pi}{x^3}\sum_{k=1}^{\infty}e^{-(k\pi)^2\tau}k\left[2(k\pi)^2\tau - 2 - \frac{k\pi\tilde{x_0}}{\tan(k\pi\tilde{x_0})}\right]\sin(k\pi\tilde{x_0}), \tag{D.16}$$

with $\tau = Dt/x^2$ and $\tilde{x}_0 = x_0/x$. Importantly, the joint distribution can be recast as

$$\sigma^v(x,t|x_0) = e^{-\frac{vx_0}{2D}}e^{-\frac{v^2t}{4D}}\sigma^0(x,t|x_0), \tag{D.17}$$

where σ^0 denotes the joint law for symmetric diffusion. We emphasize that this proportionality relation for fixed t is apparent on figure 3.2 of chapter 3, in the discrete lattice geometry.

D.1.4 Resetting Brownian Motion

Consider finally the continuous resetting diffusion process, which diffuses with coefficient D and resets to its initial position x_p with resetting rate λ. The Laplace transformed RETP $F_{0,\underline{x}}(p|x_0)$ obeys the following backward equation

$$\frac{\partial^2}{\partial^2 x_0} F_{0,\underline{x}}(p|x_0) - \omega^2 F_{0,\underline{x}}(p|x_0) = -\frac{\lambda}{D} F_{0,\underline{x}}(p|x_p), \tag{D.18}$$

where $\omega^2 = \frac{\lambda+p}{D}$ and the boundary conditions are given by $F_{0,\underline{x}}(p|0) = 0$ and $F_{0,\underline{x}}(p|x) = 1$. Denoting $G(x_1, x_2)$ the green function such that

$$\frac{\partial^2}{\partial^2 x_1} G(x_1, x_2) - \omega^2 G(x_1, x_2) = \delta(x_1 - x_2) \tag{D.19}$$

and $G(0, x_2) = G(x, x_2) = 0$, $F_{0,\underline{x}}(p|x_0)$ is given by

$$F_{0,\underline{x}}(p|x_0) = \frac{\sinh(\omega x_0)}{\sinh(\omega x)} - \frac{\lambda}{D} F_{0,\underline{x}}(p|x_p) \int_0^x G(u, x_0) \mathrm{d}u. \tag{D.20}$$

Taking $x_0 = x_p$, we obtain

$$F_{0,\underline{x}}(p|x_0) = \frac{\sinh(\omega x_0)}{\sinh(\omega x)} \frac{1}{1 + \frac{\lambda}{D} \int_0^x G(u, x_0) \mathrm{d}u}. \tag{D.21}$$

Finally, the green function is easily shown to be equal to

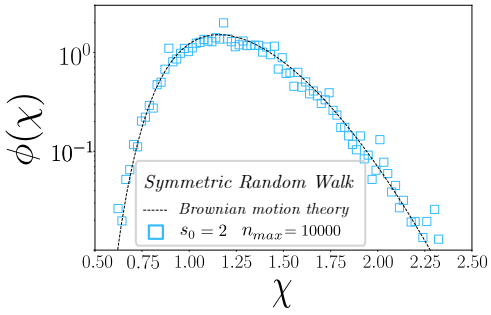
$$\begin{aligned} G(y, x_0) =& H(x_0 - y) \left[\frac{e^{-(x_0+y)\omega} (e^{2x\omega} - e^{2x_0\omega})(1 - e^{2y\omega})}{(e^{2x\omega} - 1) 2\omega} \right] \\ &+ H(y - x_0) \left[\frac{e^{-(x_0+y)\omega} (e^{2x\omega} - e^{2y\omega})(1 - e^{2x_0\omega})}{(e^{2x\omega} - 1) 2\omega} \right]. \end{aligned} \tag{D.22}$$

Making use of the symmetry $x_0 \to x - x_0$ to obtain the LETP, the joint distribution is finally obtained from equation (3.11) of chapter 3, and reads

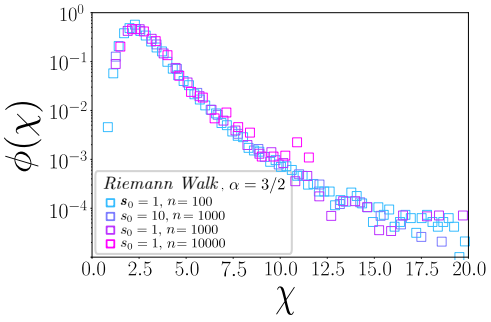
$$\sigma(x, p|x_0) = \frac{D\omega^3 [p + \lambda \cosh(\omega(x - x_0))] \sinh(\omega x_0)}{\sinh^2(\omega x)[p + \lambda \cosh(\omega x)][\sinh(\omega(x - x_0)) + \sinh(\omega x_0)]^2}. \tag{D.23}$$

D.2 Conditional maximum distribution

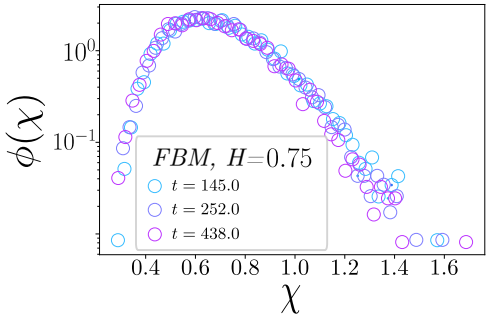
We numerically estimate the conditional distribution $\phi(\chi)$ of the rescaled variable $\chi = x/t^{\frac{1}{d_w}}$ for the scale-invariant processes studied in chapter 3. For various fixed t values, all curves collapse to t and s_0 independent universal functions.



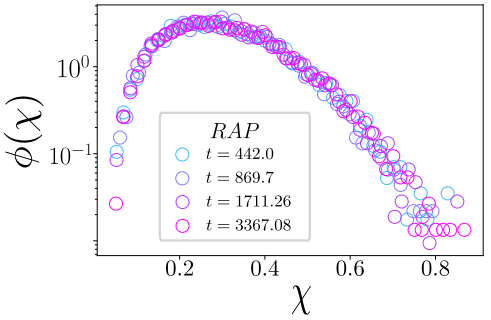
(a) Symmetric random walk: distribution of the rescaled variable χ for a fixed FPT n_{max}. The dashed line is obtained from equation (3.26) of chapter 3 and the exact $f_{BM}(\tau)$ function.



(b) Riemann walk: distribution of the rescaled variable χ for various fixed FPT n. All distributions collapse on a n and s_0 independent curve.



(c) Fractional Brownian Motion with Hurst exponent H=0.75: distribution of the rescaled variable χ for various fixed FPT t. Details on the simulation scheme can be found in the next section.



(d) Random Acceleration Process: distribution of the rescaled variable χ for various fixed FPT t. Details on the simulation scheme can be found in the next section.

Figure D.1

D.3 Simulation details

D.3.1 FBM

The algorithm used to sample the 1D FBM trajectories is based on the circulant matrix method and is detailed in [Davies & Harte 1987, Wood & Chan 1994, Dietrich & Newsam 1997]. This method allows generating exact trajectories with a constant time step $\Delta t \sim 1.10^{-4}$, until a maximal time equal to $t_{max} = 4000$. Since the process is scale-invariant, we fix $x_0 = 1$ without loss of generality. In order to evaluate $f(\tau)$ at fixed x, we keep all trajectories for which the maximum belongs to the interval $[x - dx, x + dx]$, with dx as small as possible. Similarly for $\phi(\chi)$, we only keep trajectories for which $t \in [t - dt, t + dt]$. We here recap the intervals chosen for each numerical experiment:

- Figure 3.3 of chapter 3: $H = 0.375$, $(x = 10.1,\ dx = 0.674)$, $(x = 21.9,\ dx = 1.46)$, $(x = 28.5,\ dx = 1.9)$
- Figure D.1: $H = 0.75$, $(t = 145,\ dt = 4)$, $(t = 252,\ dt = 12)$, $(t = 438,\ dt = 16)$

D.3.2 RAP

The stochastic trajectories of the random acceleration processes are generated by means of the algorithm introduced in [Bicout & Burkhardt 2000], that computes the exact probability function with a discrete time step. Since the velocity grows with time t as $\sqrt{t}$, we reduce the time step at each iteration in order to keep the discrete space interval covered small and use $\Delta t = \frac{0.05}{n^{1/3}}$ where n is the number of steps. The process being scale-invariant, we fix $x_0 = 1$ without loss of generality. To evaluate conditional distributions, we proceed similarly to the FBM and bin trajectories according to a pair (t, dt) or (x, dx). We recap the intervals chosen for each numerical experiment:

- Figure 3.3 of chapter 3: $(x = 402, dx = 27.7)$, $(x = 2961.2, dx = 197)$,$(x = 8040.7, dx = 536)$
- Figure D.1: $(t = 442,\ dt = 5)$, $(t = 869.7,\ dt = 10)$, $(t = 1711.26,\ dt = 50)$,$(t = 3367.08,\ dt = 80)$

Appendix E

Self Avoiding True Walk - detailed calculations

We reproduce in this Appendix the derivation of the splitting probability and asymptotic distribution of the maximum for the SATW process introduced in chapter 3, as well as the derivation of the joint distribution $\sigma(s, n|s_0)$. Note that these results can be found in [Klinger *et al.* 2022a].

E.1 Splitting probability of the SATW

We first derive the splitting probability $\pi_{0,\underline{s}}$ of the SATW process starting from site $s_0 = 1$, defined as the probability to reach site s before 0. It is convenient to parameterize the SATW dynamics in terms of the number of distinct sites already visited. To do so, we introduce $\pi_{0,\underline{s}}(s')$, defined as the splitting probability of a walker currently at s' and having already visited the sites $\mathcal{D}_s = [1, s-1]$. In turn, $\pi_{0,\underline{s}}$ verifies the following recurrence relation:

$$\pi_{0,\underline{s+1}} = \pi_{0,\underline{s}}\pi_{0,\underline{s+1}}(s). \tag{E.1}$$

Inside the visited territory D_s, the walker performs a classical symmetric nearest neighbor random walk such that, for $0 < s' < s$:

$$\pi_{0,\underline{s}}(s') = \frac{1}{2}\pi_{0,\underline{s}}(s'-1) + \frac{1}{2}\pi_{0,\underline{s}}(s'+1). \tag{E.2}$$

yielding

$$\pi_{0,\underline{s}}(s') = \lambda + \mu s, \tag{E.3}$$

where λ and μ can be deduced from the boundary conditions at the extremities of the visited area:

$$\begin{cases} \pi_{0,\underline{s}}(1) = (1-\beta)\,\pi_{0,\underline{s}}(2) \\ \pi_{0,\underline{s}}(s-1) = \beta + (1-\beta)\,\pi_{0,\underline{s}}(s-2). \end{cases} \tag{E.4}$$

such that

$$\pi_{0,\underline{s+1}}(s) = 1 - \frac{1-\beta}{2+\beta(s-3)}. \tag{E.5}$$

Finally, making use of the recurrence relation (E.1), the exact spitting probability starting from $s_0 = 1$ is given by:

$$\pi_{0,\underline{s}} = \prod_{s'=1}^{s-1} \left(1 - \frac{1-\beta}{2+\beta(s'-3)}\right) \underset{s\to\infty}{\sim} \frac{\Gamma(-2+2/\beta)}{\Gamma(-1+1/\beta)} s^{-\frac{1-\beta}{\beta}}. \tag{E.6}$$

Consequently, the asymptotic distribution $\mu(s|1)$ of the maximum before reaching 0 is deduced from the splitting probability

$$\mu(s|1) = \pi_{0,\underline{s}} \left(1 - \pi_{0,\underline{s+1}}(s)\right), \tag{E.7}$$

and reads:

$$\mu(s|1) \underset{s\to\infty}{\sim} \frac{\Gamma(-2+2/\beta)}{\Gamma(-1+1/\beta)} \frac{1-\beta}{\beta} s^{-\frac{1-\beta}{\beta}-1}. \tag{E.8}$$

E.2 Generating function of the joint distribution

We now consider the generating function of the joint distribution $\sigma(s,n|1)$ of the maximum s and the FPT n to 0, starting from $s_0 = 1$. Let us denote $F_{0,\underline{s}}(n|s_0)$ the probability to reach site s before 0 for the first time at step n starting from s_0 knowing that the sites $\{1,..,s-1\}$ have already been visited. By partitioning over times at which new sites are discovered, we obtain in the generating function formalism:

$$\tilde{\sigma}(s,\xi|1) = \frac{\xi}{2} \left(\prod_{s'=3}^{s} \tilde{F}_{0,\underline{s'}}(\xi|s'-1) \right) \tilde{F}_{\underline{0},s+1}(\xi|s). \tag{E.9}$$

We first compute $\tilde{F}_{0,\underline{s}}(\xi|s_0)$. For $1 < s_0 < n$, the SATW performs a simple random walk such that:

$$\tilde{F}_{0,\underline{s}}(\xi|s_0) = \frac{\xi}{2}\tilde{F}_{0,\underline{s}}(\xi|s_0+1) + \frac{\xi}{2}\tilde{F}_{0,\underline{s}}(\xi|s_0-1) \tag{E.10}$$

and

$$\tilde{F}_{0,\underline{s}}(\xi|s_0) = \lambda r_1^{s_0} + \mu r_2^{s_0} \tag{E.11}$$

with

$$\begin{cases} r_1 &= \frac{1}{\xi} - \frac{\sqrt{1-\xi^2}}{\xi} \\ r_2 &= \frac{1}{\xi} + \frac{\sqrt{1-\xi^2}}{\xi}. \end{cases} \tag{E.12}$$

We now deduce λ and μ using the boundary conditions

$$\begin{cases} \tilde{F}_{0,\underline{s}}(\xi|1) = \xi(1-\beta)\tilde{F}_{0,\underline{s}}(\xi|2) \\ \tilde{F}_{0,\underline{s}}(\xi|s-1) = \beta\xi + (1-\beta)\xi\tilde{F}_{0,\underline{s}}(\xi|s-2), \end{cases} \tag{E.13}$$

and obtain

$$\tilde{F}_{0,\underline{s}}(\xi|s-1) = \beta\xi \frac{r_1^{s-3}(r_1-(1-\beta)\xi) - r_2^{s-3}(r_2-(1-\beta)\xi)}{r_1^{s-4}(r_1-(1-\beta)\xi)^2 - r_2^{s-4}(r_2-(1-\beta)\xi)^2}. \tag{E.14}$$

Of note, the specific cases of $s=2$ and $s=3$ are given by:

$$\begin{cases} \tilde{F}_{0,\underline{2}}(\xi|1) = \frac{\xi}{2} \\ \tilde{F}_{0,\underline{3}}(\xi|2) = \frac{\beta\xi}{1-(1-\beta)^2\xi^2}. \end{cases} \tag{E.15}$$

Finally, the probabilistic weight of the last term in equation (E.9) is given by

$$\tilde{F}_{\underline{0},s+1}(\xi|s) = \beta\xi \frac{(1-(1-\beta)r_2\xi) - (1-(1-\beta)r_1\xi)}{r_1^{s-3}(r_1-(1-\beta)\xi)^2 - r_2^{s-3}(r_2-(1-\beta)\xi)^2}, \tag{E.16}$$

and plugging equation (E.16) and (E.14) into equation (E.9) yields the generating function of the joint distribution:

$$\begin{aligned} \tilde{\sigma}(\xi, s|s_0=1) = &\frac{\beta^{s-1}\xi^s}{2(1-(1-\beta)^2\xi^2)} \left(\prod_{i=0}^{s-4} \frac{r_1^{i+1}(r_1-(1-\beta)\xi) - r_2^{i+1}(r_2-(1-\beta)\xi)}{r_1^i(r_1-(1-\beta)\xi)^2 - r_2^i(r_2-(1-\beta)\xi)^2} \right) \\ &\times \frac{(1-(1-\beta)r_2\xi) - (1-(1-\beta)r_1\xi)}{r_1^{s-3}(r_1-(1-\beta)\xi)^2 - r_2^{s-3}(r_2-(1-\beta)\xi)^2}. \end{aligned} \tag{E.17}$$

E.3 Distribution of the FPT to 0 conditioned on the value of the maximum s

In this section, we consider the asymptotic behavior of $\sigma(n|s, s_0=1)$ in the large n and s limit. To derive $\sigma(n|s, s_0=1)$ we analyze independently each of the terms in the convolution (E.9), divided by the corresponding normalizing factor. Indeed, the conditional distribution can be written as

$$\tilde{\sigma}(\xi|s,1) = \xi \left(\prod_{s'=3}^{s} \left[\frac{\tilde{F}_{0,\underline{s'}}(\xi|s'-1)}{\pi_{0,\underline{s'}}(s'-1)} \right] \right) \left[\frac{\tilde{F}_{\underline{0},s+1}(\xi|s)}{\pi_{\underline{0},s+1}(s)} \right]. \tag{E.18}$$

Consider first a single term inside the product:

$$\begin{aligned} \pi_{0,\underline{i+4}}(i+3)^{-1}\tilde{F}_{0,\underline{i+4}}(\xi|i+3) &= \frac{1}{1-\frac{1-\beta}{\beta i+2}} \beta\xi \frac{r_1^{i+1}(r_1-(1-\beta)\xi) - r_2^{i+1}(r_2-(1-\beta)\xi)}{r_1^i(r_1-(1-\beta)\xi)^2 - r_2^i(r_2-(1-\beta)\xi)^2} \\ &= \frac{1}{1-\frac{1-\beta}{\beta i+2}} \left(\beta\xi \frac{r_1(r_1-(1-\beta)\xi) - \frac{r_2^i}{r_1^i} r_2(r_2-(1-\beta)\xi)}{(r_1-(1-\beta)\xi)^2 - \frac{r_2^i}{r_1^i}(r_2-(1-\beta)\xi)^2} \right). \end{aligned} \tag{E.19}$$

To obtain the leading order behavior for large n and s, we consider equation (E.19) in the limit $\xi \to 1$ and $i \to \infty$. In particular, since we are looking for a scaling form in terms of the rescaled

variable $\tau = n/s^2$, we rewrite $\xi = e^{-u}$, and consider the joint limit $u \to 0$ and $s \to \infty$ with ui^2 fixed. In this limit, the first vanishing order is given by

$$\pi_{0,\underline{i+4}}(i+3)^{-1}\tilde{F}_{0,\underline{i+4}}(\xi|i+3) \underset{\substack{u\to 0\\ i\to\infty\\ ui^2 \text{ fixed}}}{=} 1 + \frac{1-\beta}{\beta i}\left(1 - \frac{\sqrt{2ui^2}}{\tanh(\sqrt{2ui^2})}\right) + R(u,i) \tag{E.20}$$

where $R(u,i)$ is a subleading term. We now plug this asymptotic behavior in the full product and obtain

$$\begin{aligned}\prod_{s'=3}^{s}\left[\frac{\tilde{F}_{0,\underline{s'}}(\xi|s'-1)}{\pi_{0,\underline{s'}}(s'-1)}\right] &\underset{\substack{u\to 0\\ us^2 \text{ fixed}}}{\sim} \exp\left[\sum_{i=3}^{s}\ln\left(1+\frac{1-\beta}{i\beta}\left(1-\frac{\sqrt{2ui}}{\tanh(i\sqrt{2u})}\right)\right)\right] \\ &\underset{\substack{u\to 0\\ us^2 \text{ fixed}}}{\sim} \exp\left[\int_{i=0}^{1}\frac{1-\beta}{i\beta}\left(1-\frac{\sqrt{2us^2}i}{\tanh(i\sqrt{2us^2})}\right)\mathrm{d}i\right] \\ &\underset{\substack{u\to 0\\ us^2 \text{ fixed}}}{\sim} \exp\left[\frac{1-\beta}{\beta}\ln\left(\frac{\sqrt{2us^2}}{\sinh(\sqrt{2us^2})}\right)\right].\end{aligned} \tag{E.21}$$

Similarly, we obtain the leading order behavior of the last term

$$\left[\frac{\tilde{F}_{\underline{0},s+1}(\xi|s)}{\pi_{\underline{0},s+1}(s)}\right] \underset{\substack{u\to 0\\ us^2 \text{ fixed}}}{\sim} \frac{\sqrt{2us^2}}{\sinh(\sqrt{2us^2})}. \tag{E.22}$$

Combining equations (E.21) and (E.22), we finally compute the small u and large s behavior of the conditional distribution:

$$\tilde{\sigma}(\xi = e^{-u}, s|s_0 = 1) \underset{\substack{u\to 0\\ us^2 \text{ fixed}}}{\sim} \left(\frac{\sqrt{2us^2}}{\sinh(\sqrt{2us^2})}\right)^{\frac{1}{\beta}}, \tag{E.23}$$

from which we obtain the asymptotic distribution $f_{SATW}(\tau)$ of the rescaled variable $\tau = n/s^2$, given by its strikingly simple Laplace Transform:

$$\tilde{f}_{SATW}(p) = \int_0^\infty e^{-p\tau} f_{SATW}(\tau)\mathrm{d}\tau = \left(\frac{\sqrt{2p}}{\sinh(\sqrt{2p})}\right)^{1/\beta}. \tag{E.24}$$

Appendix F

Splitting probability of Gamma jump processes

We provide in this Appendix the exact formula for the splitting probability $\pi_{0,\underline{x}}(x_0)$ for a gamma jump process with jump distribution $p(\ell) = \frac{\gamma^2}{2}|\ell|e^{-\gamma|\ell|}$:

$$\pi_{0,\underline{x}}(x_0) = Ae^{-\sqrt{3}\gamma x_0} + Be^{-\sqrt{3}\gamma x_0} + Cx_0 + D$$

$$A = \frac{2e^{\sqrt{3}\gamma x}}{e^{\sqrt{3}\gamma x}\left(\left(2\sqrt{3}+3\right)\gamma x + 6\sqrt{3}+8\right) + \left(2\sqrt{3}-3\right)\gamma x + 6\sqrt{3} - 8}$$

$$B = -\frac{\left(5\sqrt{3}+9\right)^2}{3\left(e^{\sqrt{3}\gamma x}\left(\left(97\sqrt{3}+168\right)\gamma x + 276\sqrt{3}+478\right) + \left(7\sqrt{3}+12\right)\gamma x + 36\sqrt{3}+62\right)}$$

$$C = \frac{3\gamma\left(e^{\sqrt{3}\gamma x}\left(e^{\sqrt{3}\gamma x}\left(\left(2\sqrt{3}+3\right)\gamma x + \left(11\sqrt{3}+19\right)e^{\sqrt{3}\gamma x} + 7\sqrt{3}+11\right) - 3\sqrt{3}+5\right) + \left(45-26\sqrt{3}\right)\gamma x - 63\sqrt{3}+109\right)}{\left(\left(\sqrt{3}+1\right)e^{\sqrt{3}\gamma x} + \left(2\sqrt{3}-3\right)\gamma x + 5\sqrt{3}-7\right)\left(e^{2\sqrt{3}\gamma x}\left(3\left(4\sqrt{3}+7\right)\gamma x + 34\sqrt{3}+60\right) + 4\sqrt{3}e^{\sqrt{3}\gamma x} + 3\left(4\sqrt{3}-7\right)\gamma x + 34\sqrt{3}-60\right)}$$

$$D = \frac{-3e^{\sqrt{3}\gamma x}\left(2\left(\sqrt{3}-2\right)\gamma x + 9\sqrt{3}-17\right) + 3e^{2\sqrt{3}\gamma x}\left(\left(3\sqrt{3}+4\right)\gamma x + 11\sqrt{3}+17\right) + \left(47\sqrt{3}+81\right)e^{3\sqrt{3}\gamma x} + 3\left(64-37\sqrt{3}\right)\gamma x - 269\sqrt{3}+465}{\left(\left(\sqrt{3}+1\right)e^{\sqrt{3}\gamma x} + \left(2\sqrt{3}-3\right)\gamma x + 5\sqrt{3}-7\right)\left(e^{2\sqrt{3}\gamma x}\left(3\left(4\sqrt{3}+7\right)\gamma x + 34\sqrt{3}+60\right) + 4\sqrt{3}e^{\sqrt{3}\gamma x} + 3\left(4\sqrt{3}-7\right)\gamma x + 34\sqrt{3}-60\right)} \tag{F.1}$$

Appendix G

Infinite and semi-infinite propagators for jump processes

We devote this Appendix to the derivation of infinite and semi-infinite propagators of symmetric jump processes. All material presented here can be found across various classical references among which [Ivanov 1994, Hughes 1995, Majumdar *et al.* 2006] and references therein. However, we believe a comprehensive presentation is relevant to the present thesis, and choose to include some calculations in this Appendix. Importantly, most of the material presented here is a reformulation of calculations extracted from [Ivanov 1994].

G.1 Infinite space propagator

As a reminder, we consider here a discrete-time, one-dimensional jump process starting from x_0, whose symmetric single jump distribution is denoted $p(\ell) = p(-\ell)$. We first compute the infinite space propagator of the jump process $G_\infty(x, n|x_0)$, defined as the probability to be located at position x after n steps. In the following, we will be considering generating functions, Fourier transforms and Laplace transforms of the various probabilistic quantities involved. To increase readability, such transforms will be indicated respectively by the corresponding variables, ξ, k and s. As an example, we rewrite $G_\infty(x, \xi|x_0) = \sum_{n=0}^{\infty} G_\infty(x, n|x_0)\xi^n$ and $G_\infty(k, \xi|x_0) = \int_{-\infty}^{\infty} e^{ikx} G_\infty(x, \xi|x_0)\mathrm{d}x$.

The infinite space propagator is straightforwardly obtained from the single jump distribution $p(\ell)$. Indeed, for a fixed number of steps n, since the walk is unbounded, G_∞ is given by convoluting $p(\ell)$ over $\mathbb{R}$ n times:

$$G_\infty(x, n|x_0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x_1 \dots \mathrm{d}x_n p(x_1)\dots p(x_n)\delta(x_0 + x_1 + \dots + x_n = x) \tag{G.1}$$

In the generating function formalism, and using well-known properties of the convolution, one obtains:

$$\begin{aligned} &G_\infty(x, \xi|x_0) = \delta(x - x_0) + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\xi\ p(k)}{1 - \xi\ p(k)} e^{-ik(x-x_0)}\mathrm{d}k, \\ &\text{or} \\ &G_\infty(k, \xi) = \frac{1}{1 - \xi p(k)}. \end{aligned} \tag{G.2}$$

Importantly, while equation (G.2) is given in the generating function space, it is fully explicit in terms of $p(k)$, and can be inverted (at least numerically), to obtain $G_\infty(x, n|x_0)$ for any values of n.

G.2 Semi-infinite space propagator

We now consider the same jump process, but killed (*ie* stopped) upon entering $\mathbb{R}^*_-$, and define the corresponding propagator $G(x, n|x_0)$. Of note, while $\int_{\mathbb{R}} G_\infty(x, n|x_0) = 1$ for the infinite (or unbounded) process, $\int_{\mathbb{R}} G(x, n|x_0) = q(x_0, n)$, where $q(x_0, n) < 1$ is the survival probability of the process, defined as the probability to stay positive during the first n steps. We now proceed to show that similarly to G_∞, G can also be expressed solely as a function of the single jump distribution $p(\ell)$.

G as the green function of the semi-infinite problem. Let us first write the *forward* equation for G, by partitioning trajectories over the position of the particle on its $n-1$ step. In the generating function formalism, the *forward* equation is directly given by

$$G(x, \xi|x_0) = \delta(x - x_0) + \xi \int_0^\infty p(x - y)G(y, \xi|x_0)\mathrm{d}y, \tag{G.3}$$

where the $\delta(x - x_0)$ function corresponds to $G(x, n = 0|x_0)$. While equation (G.3) is not directly solvable for arbitrary $p(\ell)$ distributions, it defines G as the green function of the general in-homogeneous problem

$$f(x, \xi) = \xi \int_0^\infty p(x - y)f(y, \xi)\mathrm{d}y + S(x) \tag{G.4}$$

with $S(x)$ some arbitrary source term, such that

$$f(x, \xi) = \int_0^\infty G(x, \xi|x_0)S(x_0)\mathrm{d}x_0. \tag{G.5}$$

Equation (G.5) will prove useful in the following.

Expressing G as a function of the boundary propagator G_0. Since $G(x, \xi|x_0)$ cannot be determined directly, our first objective is to reduce its evaluation to that of $G(x, \xi|0) \equiv G_0(x)$, where the initial condition is taken to be exactly at the origin. Note that $G_0(x)$ depends on ξ but we drop that dependence to increase readability. Consider first all n-step trajectories starting from x_0 and ending at x. Necessarily, there exists a minimum $\hat{x}$ of the trajectory such that $0 \leq \hat{x} \leq \min(x_0, x)$ (see figure G.1). Using the symmetry of the single jump distribution, and partitioning over the step k at which this minimum is reached, we obtain, in the generating function formalism:

$$G(x, \xi|x_0) = \int_0^{\min(x_0,x)} G_0(x_0 - x')G_0(x - x')\mathrm{d}x'. \tag{G.6}$$

Let us now show that in the Laplace transformed space, equation (G.6) takes a simple product-like form. Taking the Laplace transform with respect to the variable x_0, we obtain:

$$\begin{aligned} \int_0^\infty e^{-s_1 x_0} G(x, \xi|x_0)\mathrm{d}x_0 = &\int_0^x \mathrm{d}x_0 e^{-s_1 x_0} \int_0^{x_0} G_0(x_0 - x')G_0(x - x')\mathrm{d}x' \quad (A) \\ &+ \int_x^\infty \mathrm{d}x_0 e^{-s_1 x_0} \int_0^x G_0(x_0 - x')G_0(x - x')\mathrm{d}x' \quad (B). \end{aligned} \tag{G.7}$$

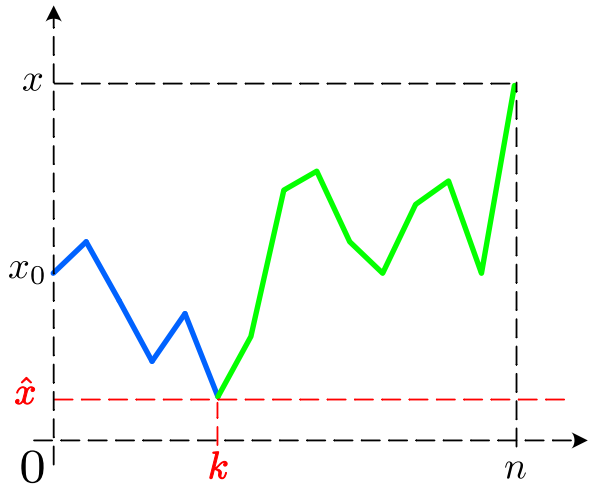


Figure G.1: Example trajectory contributing to $G(x,n|x_0)$. The minimum $\hat{x} \in [0, \min(x_0, x)]$ is reached on the k^{th} step. By decomposing the trajectory in two bits around step k and making use of the symmetry of the walk, each part has a respective probabilistic weight $G(x_0 - \hat{x}, k|0)$ (blue) and $G(x - \hat{x}, n - k|0)$ (green).

In turn, we analyze terms (A) and (B) separately. Starting with (A), inverting integration order and making a change of variables yields:

$$\begin{aligned}\int_0^x \mathrm{d}x_0 e^{-s_1 x_0} \int_0^{x_0} G_0(x_0 - x')G_0(x - x')\mathrm{d}x' &= \int_0^x \mathrm{d}x' \int_{x'}^x \mathrm{d}x_0 e^{-s_1 x_0} G_0(x_0 - x')G_0(x - x')\mathrm{d}x' \\ &= \int_0^x G_0(x - x')e^{-s_1 x'}\mathrm{d}x' \int_0^{x-x'} \mathrm{d}x_0 e^{-s_1 x_0} G_0(x_0)\mathrm{d}x'.\end{aligned} \tag{G.8}$$

Term (B) is taken care of similarly

$$\begin{aligned}\int_x^\infty \mathrm{d}x_0 e^{-s_1 x_0} \int_0^{x} G_0(x_0 - x')G_0(x - x')\mathrm{d}x' &= \int_0^x \mathrm{d}x' \int_x^\infty \mathrm{d}x_0 e^{-s_1 x_0} G_0(x_0 - x')G_0(x - x') \\ &= \int_0^x \mathrm{d}x' G_0(x - x')e^{-s_1 x'} \int_{x-x'}^\infty \mathrm{d}x_0 e^{-s_1 x_0} G_0(x_0),\end{aligned} \tag{G.9}$$

and summing the two terms yields:

$$G(x,\xi|s_1) = G_0(s_1) \int_0^x \mathrm{d}x' G_0(x - x')e^{-s_1 x'} \tag{G.10}$$

where $G_0(s) = \int_0^\infty e^{-sx'} G_0(x')\mathrm{d}x'$ is the Laplace transform of G_0. Finally, Laplace transforming equation (G.10) with respect to x yields following factorization:

$$G(s,\xi|s_1) = \frac{G_0(s)G_0(s_1)}{s + s_1} \tag{G.11}$$

Importantly, the interchangeability of x and x_0 in $G(x,\xi|x_0)$ is fully apparent in the Laplace transformed space. Additionally, equation (G.11) indicates that G_0 alone is enough to fully

determine $G(x, \xi|x_0)$.

Determination of G_0. We now determine G_0 in terms of $p(k)$ only. Consider first the associated forward equation, obtained by partitioning over the position of the process before taking its last step. Doing so directly in the generating function formalism we obtain:

$$G_0(x) = \delta(x) + \xi \int_0^\infty p(x-y) G_0(y) \mathrm{d}y, \tag{G.12}$$

where we recall that $G_0(x)$ is a shorthand for $G_0(x, \xi) = \sum_{n=0}^\infty \xi^n G(x, n|x_0 = 0)$, and the $\delta(x)$ function corresponds to $G(x, n = 0|x_0 = 0)$.

We now aim to find a self-consistent equation for G_0. Recalling that G can be both expressed as a function of G_0 and as the green function of the semi-infinite problem, we artificially generate an in-homogeneous integral problem by differentiating (G.12) with respect to ξ:

$$\begin{aligned} \partial_\xi G_0(x) &= \xi \int_0^\infty p(x-y) \partial_\xi G_0(y) \mathrm{d}y + \int_0^\infty p(x-y) G_0(y) \mathrm{d}y \\ &= \xi \int_0^\infty p(x-y) \partial_\xi G_0(y) \mathrm{d}y + \frac{1}{\xi} \left[G_0(x) - \delta(x) \right]. \end{aligned} \tag{G.13}$$

Equation (G.13) is now of the general form (G.4) such that we obtain:

$$\begin{aligned} \xi \partial_\xi G_0(x) &= \int_0^\infty G(x, \xi|x_0) \left[G_0(x_0) - \delta(x_0) \right] \mathrm{d}x_0 \\ &= \int_0^\infty G(x, \xi|x_0) G_0(x_0) \mathrm{d}x_0 - G_0(x). \end{aligned} \tag{G.14}$$

Taking the Laplace transform with respect to the x variable, and replacing G according to equation (G.10) yields

$$\begin{aligned} \xi \partial_\xi G_0(s) &= \int_0^\infty G(s, \xi|x_0) G_0(x_0) \mathrm{d}x_0 - G_0(s) \\ &= G_0(s) \int_0^\infty \mathrm{d}x_0 \int_0^{x_0} \mathrm{d}x' G_0(x_0 - x') G_0(x_0) e^{-s x'} - G_0(s), \end{aligned} \tag{G.15}$$

such that

$$\xi \partial_\xi \log \left[G_0(s) \right] + 1 = \int_0^\infty \mathrm{d}x_0 \int_0^{x_0} \mathrm{d}x' G_0(x_0 - x') G_0(x_0) e^{-s x'}. \tag{G.16}$$

Finally, inverting once again the integration order and making a change of variable, we obtain:

$$\xi \partial_\xi \log \left[G_0(s) \right] + 1 = \int_0^\infty \mathrm{d}x' e^{-s x'} \int_0^\infty \mathrm{d}x_0 G_0(x_0) G_0(x_0 + x'). \tag{G.17}$$

Equation (G.17) is close to self consistent, but is still not solvable as such. However, by focusing on the right-hand side, once can see a strong parallel with equation (G.6). Indeed, the integrated product of G_0 functions can be recast as the unbounded infinite propagator G_∞, with a very similar argument. Considering all unbounded trajectories starting from 0 and ending at x' on

their n^{th} step, and partitioning over the value of the minimum and the step at which it is reached, we obtain, in the generating function formalism:

$$G_\infty(x', \xi|0) = \int_0^\infty \mathrm{d}x_0 G_0(x_0) G_0(x_0 + x'). \tag{G.18}$$

Note that the physical interpretation of this relation is the same as equation (G.6), but the value of the minimum is now unbounded. Plugging into equation (G.17), and making use of equation (G.2) for G_∞, we obtain

$$\xi \partial_\xi \log\left[G_0(s)\right] = \frac{s}{2\pi} \int_{-\infty}^{\infty} \frac{\xi p(k)}{1 - \xi p(k)} \frac{1}{s^2 + k^2} \mathrm{d}k, \tag{G.19}$$

and a final integration with respect to ξ yields the sought-after closed form expression for $G_0(s, \xi)$:

$$G_0(s, \xi) = \exp\left[-\frac{s}{2\pi} \int_{-\infty}^{\infty} \frac{\log\left[1 - \xi p(k)\right]}{s^2 + k^2} \mathrm{d}k \right]. \tag{G.20}$$

Let us make a few concluding remarks:

- Together with equation (G.11), equation (G.20) yields an explicit expression of the semi-infinite propagator as a function of the single jump distribution $p(\ell)$, similarly to the unbounded case.
- Equation (G.20) lends itself to asymptotic analysis with respect to the parameters s and ξ. In particular, taking s to 0 yields the survival probability starting from 0:

$$\int_{\mathbb{R}} G_0(x, \xi) \mathrm{d}x = q(x_0 = 0, \xi) = \frac{1}{\sqrt{1 - \xi}} \tag{G.21}$$

  and one recovers the Sparre Andersen Theorem [Andersen 1954].

Appendix H

Approximate eigenfunctions and eigenvalues for the bounded fractional diffusion equation

In this Appendix, we reproduce results from [Kwaśnicki 2012] regarding the eigenvalues and eigenfunctions of the fractional Laplacian. We emphasize that the Fractional Laplacian is an important technical tool to describe Levy-like processes, but that the eigenfunctions ψ_k and eigenvalues λ_k of the operator in a bounded one-dimensional interval are not known analytically. Here we present approximates of these quantities; we closely follow the notations of [Kwaśnicki 2012], and consider the fractional diffusion eigenvalue problem in the domain $D = (-1, 1)$. Essentially, the objective is to find solutions to the following fractional diffusion equation:

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}\right)^{\frac{\mu}{2}} \phi(x) = \lambda\phi(x) \quad \text{for } x \in (-1, 1) \tag{H.1}$$

where $\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}\right)^{\frac{\mu}{2}}$ is the fractional Laplacian (see [Zoia *et al.* 2007, Kwasnicki 2017] for a detailed presentation of this operator), and with $\phi(-1) = \phi(1) = 0$. Denoting λ_k the k^{th} eigenvalue, sorted in increasing order, the λ_k are approximated by:

$$\lambda_k \underset{k\to\infty}{=} \left(\frac{k\pi}{2} - \frac{(2-\mu)\pi}{8}\right)^{\mu} + O\left(\frac{1}{k}\right). \tag{H.2}$$

Next, approximate expressions of the corresponding eigenfunctions ψ_k can be obtained by combining infinite and semi-infinite eigenfunctions of the fractional Laplacian, which are known explicitly. It is found that

$$\psi_k(x) = q(-x)F_{\mu_k}(1+x) - (-1)^k q(x)F_{\mu_k}(1-x), \tag{H.3}$$

with

$$\mu_k = \frac{k\pi}{2} - \frac{(2-\mu)\pi}{8} \tag{H.4}$$

and F_λ defined in the following way:

$$F_\lambda(x) = \sin\left(\lambda x + \frac{(2-\mu)\pi}{8}\right) - G(\lambda x). \tag{H.5}$$

Here G is the Laplace transform of a positive function $\gamma(s)$:

$$\begin{aligned}\gamma(s) =& \frac{\sqrt{2\mu}\sin\left(\frac{\mu\pi}{2}\right)}{2\pi}\frac{s^\mu}{1+s^{2\mu}-2s^\mu\cos\left(\frac{\mu\pi}{2}\right)} \\ & \times\exp\left[\frac{1}{\pi}\int_0^\infty\frac{\mathrm{d}r}{1+r^2}\ln\left(\frac{1-r^\mu s^\mu}{1-r^2s^2}\right)\right],\end{aligned} \tag{H.6}$$

and q is an interpolating function:

$$q(x) = \begin{cases} 0 & \text{for } x\in(-\infty,-\frac{1}{3}) \\ \frac{9}{2}\left(x+\frac{1}{3}\right)^2 & \text{for } x\in(-\frac{1}{3},0) \\ 1-\frac{9}{2}\left(x-\frac{1}{3}\right)^2 & \text{for } x\in(0,\frac{1}{3}) \\ 1 & \text{for } x\in(\frac{1}{3},\infty). \end{cases} \tag{H.7}$$

From these combined expressions, one can numerically evaluate the approximate eigenfunctions ψ_k. In particular, this is used to determinate numerical values for λ_1, γ_μ and ω_μ in table 5.1 of chapter 5.

Appendix I

Sparre-Andersen-like theorem for the integrated strip probability

We devote this Appendix to a proof of a Sparre-Andersen-like theorem arising in the context of strip probabilities studied in chapter 6. More precisely, we denote p_n the probability that a n step long jump process starting from 0 stays positive, and reaches its maximum on its last step. In terms of the strip probability $\mu_{0,\underline{x}}(n)$, one has

$$\int_0^\infty \mu_{0,\underline{u}}(n)\mathrm{d}u = p_n. \tag{I.1}$$

Remarkably, for any symmetric continuous jump process, p_n is independent of the single jump distribution $p(\ell)$, and is given by

$$p_n = \frac{1}{2n}. \tag{I.2}$$

We here provide a custom combinatorial proof of equation (I.2).

I.1 Notations

We provide a few notations useful for the proof.

Notation I.1.1. Let $(X_1, \dots, X_n)$ be some *i.i.d.* random variables. We call a trajectory $\mathcal{T}$ associated to $(X_1, \dots, X_n)$ the path generated by the partial sums $(S_1, \dots, S_n)$. We will freely use the abusive notation $\mathcal{T} = (X_1, \dots, X_n)$.

Notation I.1.2. We call a *fixed sign trajectory* the trajectory $\mathcal{T}$ associated to $(\varepsilon_1 X_1, \dots, \varepsilon_n X_n)$ where the X_i are *i.i.d.* drawn from a continuous **positive** distribution, and $(\varepsilon_1, \dots \varepsilon_n) \in \{-1, 1\}^n$ is fixed.

Notation I.1.3. We denote $\mathcal{S}$ the subset of trajectories that stay positive and attain their maximum on their last step.

With the help of these notations, the theorem we aim to prove can be rephrased as follows: let n be a fixed integer and $(X_1, \dots, X_n)$ a set of n *i.i.d.* random variables such that X_1 follows a **symmetric continuous** law. Let $(S_1, \dots, S_n)$ be their partial sums. Then:

$$p_n = P\left(\bigcap_{i=1}^n S_i > 0, \bigcap_{i=1}^n S_n \geq S_i\right) = \frac{1}{2n} \tag{I.3}$$

In other words, for a given trajectory $\mathcal{T}$ associated to $(X_1, \ldots, X_n)$, the probability that $\mathcal{T} \in \mathcal{S}$ is equal to $1/(2n)$. This applies to the case of the jump processes studied in the main text, whose increments are *i.i.d.*, and follow a symmetric and continuous law. We point out that the proof is essentially based on the decomposition of trajectories upon *fixed sign trajectories*.

I.2 Specific case of jagged trajectories

We first prove an intermediate result. Consider $(X_1, \ldots, X_n)$ a set of *i.i.d.* random variables following a **continuous positive** law. Let $\mathcal{T}$ be the fixed sign trajectory associated to $(X_1, \ldots, X_n)$ and $(1, -1, \ldots, (-1)^n)$. We call such a trajectory a *jagged* trajectory. Then:

$$P(\mathcal{T} \in \mathcal{S}) = \frac{1}{n} \qquad \text{if } n \text{ is odd and 0 otherwise} \tag{I.4}$$

Proof. Consider that $(X_1, \ldots, X_n)$ is given. Since the X_i's are equidistributed, any permutation σ of $[\![1, n]\!]$ into itself $(X_{\sigma(1)}, \ldots, X_{\sigma(n)})$, has the same probabilistic weight. In turn, these $n!$ permutations can be decomposed into $(n-1)!$ disjoint cyclic groups, containing n permutations each. We now show that among a given cyclic group, one and only one jagged trajectory $\mathcal{T}$ belongs to $\mathcal{S}$.

For clarity we introduce one more subset $\mathcal{P}$: the subset of trajectories staying strictly positive. We now propose an algorithm to transform any given trajectory $\mathcal{T}_1$ into a trajectory $\mathcal{T}_2 \in \mathcal{S}$ while staying in its cyclic group. We will examine 3 distinct cases, which cover all possible trajectory configurations. For each case, we provide the cycled trajectory, and check if it stays positive, and reaches its maximum on its last step.

1. Suppose $\mathcal{T}_1 = (X_1, \ldots, X_n) \in \mathcal{P} \cap \overline{\mathcal{S}}$. Let n_{max} be such that $\forall i \leq n, S_{n_{max}} \geq S_i$. Then $\tilde{\mathcal{T}_1} = (X_{n_{max}+1}, \ldots, X_n, X_1, \ldots, X_{n_{max}}) \in \mathcal{S}$. See below for figure and proof:

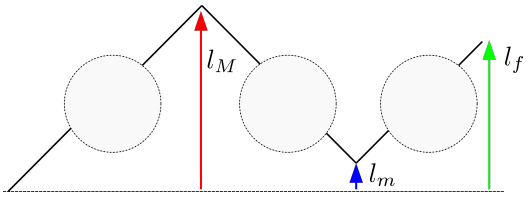


Figure I.1: sample trajectory in $\mathcal{P} \cap \overline{\mathcal{S}}$

Positivity condition: satisfied since $l_M > l_f$.
Maximum condition: $l_M - l_m < l_M - l_f + l_M \iff l_M > l_f - l_m$ which is true.

2. Suppose $\mathcal{T}_1 = (X_1, \ldots, X_n) \in \overline{\mathcal{P}}$. Let n_{min} be the index of the global minimum of $\mathcal{T}_1$ and n_{max} be the index of the local maximum of $\mathcal{T}_1$ between 1 and n_{min}. Then, if $S_{n_{max}} < S_n - S_{n_{min}}$, $\tilde{\mathcal{T}_1} = (X_{n_{min}+1}, \ldots, X_n, X_1, \ldots, X_{n_{min}}) \in \mathcal{P}$ and we repeat the above given procedure. See below for figure and proof:

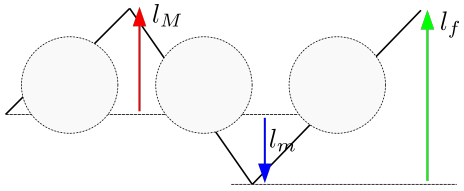


Figure I.2: sample trajectory in $\overline{\mathcal{P}}$

Positivity condition: $l_f > l_M$ true by hypothesis, so the cycled trajectory is in $\mathcal{P}$.

3. Else, let n_{max} be the index of the global maximum of $\mathcal{T}_1$, then $\tilde{\mathcal{T}_1} = (X_{n_{max}+1}, \dots, X_n, X_1, \dots, X_{n_{max}}) \in \mathcal{S}$.There are two cases depending on the relative position of the global maximum and minimum. See below for figure and proof

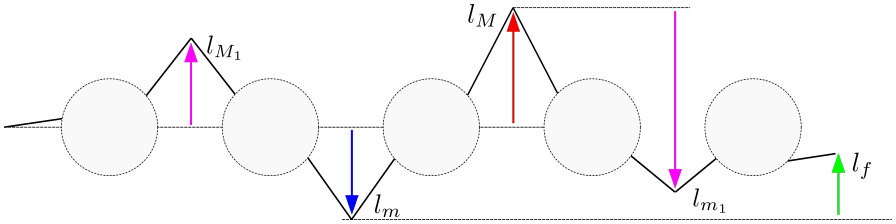


Figure I.3: case 1: the global minimum comes before the global maximum.

Positivity condition: $l_M + l_m - l_f > l_m \iff l_M - l_f > 0$. However, by hypothesis, $l_{M_1} - l_f > 0$ and $l_M > l_{M_1}$ hence the result.
Maximum condition: $l_{m_1} < l_M + l_m - l_f + l_M$. Notice that $l_{m_1} < l_M + l_m$ and $l_M < l_f$ by hypothesis, hence the result.

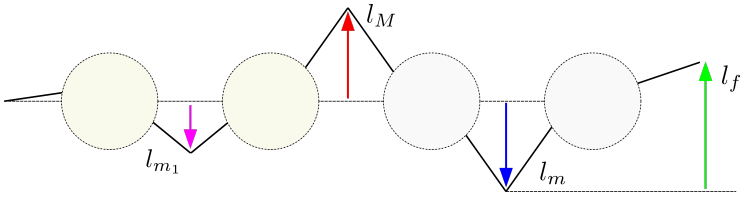


Figure I.4: case 2: the global minimum comes after the global maximum.

Positivity condition: $l_M + l_m - l_f > l_{m_1}$. By hypothesis, $l_M - l_f > 0$ and $l_m - l_{m_1} > 0$.
Maximum condition: $l_M + l_m < l_M + l_m - l_f + l_M$ immediate by hypothesis. Note that l_{m_1} may or may not exist, but it does not't change the result.

Consequently, for any given $\mathcal{T}$, we can build $\tilde{\mathcal{T}} \in \mathcal{S}$ such that $\mathcal{T}$ and $\tilde{\mathcal{T}}$ belong to the same cyclic group. Additionally, any cyclic transformation of $\mathcal{T} \in \mathcal{S}$ leads to $\tilde{\mathcal{T}} \in \overline{\mathcal{S}}$. As a result, there is a unique viable trajectory in each cyclic group, yielding the desired result.

I.3 Partitioning upon fixed sign trajectories

Let us turn back to unconstrained trajectories of length n and decompose a given trajectory upon fixed sign trajectories. We first start with a few examples.

Case $n = 3$. Consider a trajectory of length 3. With probability 2^{-3}, it belongs to fixed sign trajectories of type:

$$
\begin{array}{c}
" + + + " \\
" + - + " \\
" - + + " \\
" - + - " \\
" + - - " \\
" - + + " \\
" - - + " \\
" - - - "
\end{array}
$$

Out of these, it is clear that only trajectories of the first two kinds can belong to $\mathcal{S}$, and they do so with probability 1 and $\frac{1}{3}$ respectively, yielding a total probability $p_3 = \frac{1}{8}(1 + \frac{1}{3}) = \frac{1}{6}$ as claimed initially.

Case $n = 5$. Let us do the same work with trajectories of length 5 and list only the fixed sign trajectories contributing to p_5:

$$
\begin{array}{l}
\{' + + + + + ' \} \\
\{" + - - - - + ", " + + + - - + ", " + - + + + + "\} \\
\{" + + - + + + ", " + - - + + + ", " + + - - - + "\} \\
\{" + - + - + + "\}
\end{array}
$$

One notices a pattern: trajectories can be assembled into disjoints sets in which the number of interfaces between '+' and '-' is fixed. Equivalently, for each $i \in \mathbb{N}$ all trajectories inside a set have the same number of *consecutive sign segments* of length i. In turn, we label each trajectory with a n-tuple $(k_1, \ldots, k_n)$, such that k_i is the number of segments of size i. Finally, we denote $C(k_1, \ldots, k_n)$ the set of fixed sign trajectories that are labeled exactly by $(k_1, \ldots, k_n)$. Let us be explicit on the $n = 5$ example:

$$
\begin{array}{l}
\{" + + + + + + "\} = C(0, 0, 0, 0, 1) \\
\{" + - - - - + ", " + + + - - + ", " + - + + + + "\} = C(2, 0, 1, 0, 0) \\
\{" + + - + + + ", " + - - + + + ", " + + - - - + "\} == C(1, 2, 0, 0, 0) \\
\{" + - + - + + "\} = C(5, 0, 0, 0, 0)
\end{array}
$$

Notice that$\sum_{i=1}^{n} k_i$ must be an odd number.

General case. Let $(k_1, \ldots, k_n)$ be fixed. Let us show that

$$
\sum_{\mathcal{T} \in C(k_1, .., k_n)} P(\mathcal{T} \in \mathcal{S}) = \frac{(k_1 + \cdots + k_n)!}{k_1! ... k_n!} \frac{1}{k_1 + \cdots + k_n} \tag{I.5}
$$

Consider a *jagged* trajectory $\mathcal{U}$ of length n, with n odd, whose increments U_i are drawn from the following probability distribution:

$$U_i \sim \begin{cases} |X_1| \text{with probability } q_1 \\ |X_1| + |X_2| \text{with probability } q_2 \\ \dots \\ \sum_i^n |X_i| \text{with probability } q_n \end{cases}$$

where the X_i are the initial random variables and $\sum_i q_i = 1$. Since $\mathcal{U}$ is a *jagged* trajectory, we have from earlier results:

$$P(\mathcal{U} \in \mathcal{S}) = \frac{1}{n} \tag{I.6}$$

which can be rewritten as

$$\begin{aligned} P(\mathcal{U} \in \mathcal{S}) &= \frac{1}{n}(q_1 + \dots + q_n)^n \\ &= \sum_{\substack{(k_1,\dots,k_n),\\ k_1+\dots+k_n=n}} q_1^{k_1} \dots q_n^{k_n} \frac{(k_1 + \dots + k_n)!}{k_1!...k_n!} \frac{1}{k_1 + \dots + k_n}. \end{aligned} \tag{I.7}$$

In turn, we rewrite $P(\mathcal{U} \in \mathcal{S})$ by partitioning over the distributions of the $\mathcal{U}_i$'s, such that:

$$P(\mathcal{U} \in \mathcal{S}) = \sum_{\substack{(k_1,\dots,k_n)\\ k_1+\dots+k_n=n}} q_1^{k_1} \dots q_n^{k_n} \left[\sum_{\mathcal{T} \in C(k_1,..,k_n)} P(\mathcal{T} \in \mathcal{S}) \right] \tag{I.8}$$

Finally, identifying of the two polynomial forms yields equation (I.5).

I.4 Final summation

We are indebted to Muriel Livernet for finding a way to carry out the final summation.
We consider a trajectory $\tilde{\mathcal{T}}$ of fixed length n, and decompose $\tilde{\mathcal{T}}$ as a sequence $(k_1, \dots, k_n)$ with $\sum_{i=1}^{n} k_i i = n$ and $\sum_{i=1}^{n} k_i$ is odd. Let A_n be the set of all those sequences, and rewrite p_n as:

$$\begin{aligned} p_n &= \frac{1}{2^n} \sum_{(k_1,\dots,k_n) \in A_n} \left[\sum_{\mathcal{T} \in C(k_1,..,k_n)} P(\mathcal{T} \in \mathcal{S}) \right] \\ &= \frac{1}{2^n} \left(\sum_{(k_1,\dots,k_n) \in A_n} \frac{(k_1 + \dots + k_n)!}{k_1! \dots k_n!} \frac{1}{k_1 + \dots + k_n} \right) \end{aligned} \tag{I.9}$$

We now proceed to show that

$$\frac{1}{2^n}\left(\sum_{(k_1,\ldots,k_n)\in A_n}\frac{(k_1+\ldots+k_n)!}{k_1!\ldots k_n!}\frac{1}{k_1+\ldots+k_n}\right)=\frac{1}{2n}. \tag{I.10}$$

Defining $g_r(z)=\frac{1}{r}\left(\frac{z}{1-z}\right)^r=\frac{1}{r}(z+z^2+..+z^s+...)^r$ and developing the polynomial form yields

$$g_r(z)=\sum_{k_1+\cdots+k_s+\cdots=r}\frac{1}{r}\frac{(k_1+\cdots+k_s+\ldots)!}{k_1!\ldots k_s!\ldots}z^{k_1+2k_2+\cdots+sk_s+\cdots}. \tag{I.11}$$

To recover the 2^{-n} factor, it is convenient to introduce $h_r(z)=g_r(\frac{z}{2})$ such that

$$\begin{aligned}h_r(z)&=\sum_n b_{n,r}z^n\\ b_{n,r}&=\sum_{\substack{k_1+\cdots+k_s+\cdots=r\\ k_1+\cdots+sk_s+\cdots=n}}\frac{1}{r}\frac{(k_1+\cdots+k_s+\ldots)!}{k_1!\ldots k_s!\ldots}\frac{1}{2^n}\end{aligned} \tag{I.12}$$

which reads, keeping only plausible terms regarding the second condition under the sum:

$$b_{n,r}=\sum_{\substack{k_1+\cdots+k_n=r\\ k_1+\cdots+nk_n=n}}\frac{1}{r}\frac{(k_1+\cdots+k_n)!}{k_1!\ldots k_n!}\frac{1}{2^n} \tag{I.13}$$

We now compute the $b_{n,r}$ by defining the associated generating function:

$$F(x,y)=\sum_{n,r}b_{n,r}x^ry^n=\sum_n\left(\sum_r b_{n,r}x^r\right)y^n. \tag{I.14}$$

It is clear from equation (I.10) that $p_n=\sum\limits_{r\ \text{odd}}b_{n,r}$, such that:

$$\sum_n p_n y^n=\frac{F(1,y)-F(-1,y)}{2} \tag{I.15}$$

We finally compute $F(x,y)$:

$$\begin{aligned}F(x,y)&=\sum_{n,r}b_{n,r}x^ry^n=\sum_r\left(\sum_n b_{n,r}y^n\right)x^r=\sum_r h_r(y)x^r\\ &=\sum_r\frac{1}{r}\left(\frac{y}{2-y}\right)^r x^r=\sum_r\frac{1}{r}\left(\frac{xy}{2-y}\right)^r\\ &=-\log(1-\frac{xy}{2-y}).\end{aligned} \tag{I.16}$$

In turn

$$\frac{F(1,y)-F(-1,y)}{2}=\frac{1}{2}\left(-\ln(2-2y)+\ln(2)\right)=-\frac{1}{2}\ln(1-y)=\sum_n\frac{1}{2n}y^n \tag{I.17}$$

yielding the desired final result:

$$p_n=\frac{1}{2n} \tag{I.18}$$

Bibliography

[Adelberg 1992] Arnold Adelberg. *On the degrees of irreducible factors of higher order Bernoulli polynomials.* Acta Arith., vol. 62, no. 4, pages 329–342, 1992. (Cited on page 83.)

[Andersen 1954] E. S. Andersen. *On the fluctuations of sums of random variables II.* Mathematica Scandinavica, vol. 2, pages 195–223, 1954. (Cited on pages 5, 14, 77, 93 and 183.)

[Araújo *et al.* 2021] Michelle O. Araújo, Thierry Passerat De Silans and Robin Kaiser. *Lévy flights of photons with infinite mean free path.* Physical Review E, vol. 103, no. 1, pages 1–6, 2021. (Cited on pages 16, 100 and 102.)

[Asmussen 2003] Soren Asmussen. Applied Probability and Queues, volume 51 of *Stochastic Modelling and Applied Probability.* Springer New York, New York, NY, 2003. (Cited on pages 3, 13, 26 and 90.)

[Bachelier 1900] Louis Bachelier. *La théorie de la spéculation.* Ann. Sci. Ec. Norm. Sup., 1900. (Cited on pages 1 and 25.)

[Barbier-Chebbah *et al.* 2020] A. Barbier-Chebbah, O. Benichou and R. Voituriez. *Anomalous persistence exponents for normal yet aging diffusion.* Physical Review E, vol. 102, no. 6, pages 062115–, 12 2020. (Cited on pages 12, 81 and 82.)

[Barton 1989] G. Barton. Elements of green's functions and propagation. Oxford Science Publications, 1989. (Cited on pages 40, 48 and 49.)

[Baudouin *et al.* 2014] Q. Baudouin, R. Pierrat, A. Eloy, E. J. Nunes-Pereira, P. A. Cuniasse, N. Mercadier and R. Kaiser. *Signatures of Lévy flights with annealed disorder.* Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, vol. 90, no. 5, pages 1–11, 2014. (Cited on pages 16, 19, 100 and 154.)

[Bénichou & Voituriez 2014] O. Bénichou and R. Voituriez. *From first-passage times of random walks in confinement to geometry-controlled kinetics.* Physics Reports, vol. 539, no. 4, pages 225–284, 6 2014. (Cited on pages 9, 21, 46, 48, 51, 52, 59, 67, 69, 136 and 163.)

[Bénichou *et al.* 2005] O. Bénichou, M. Coppey, M. Moreau, P. H. Suet and R. Voituriez. *Averaged residence times of stochastic motions in bounded domains.* EPL (Europhysics Letters), vol. 70, no. 1, pages 42–48, 2005. (Cited on page 149.)

[Benichou *et al.* 2008] O. Benichou, C. Loverdo, M. Moreau and R. Voituriez. *Optimizing intermittent reaction paths.* Physical Chemistry Chemical Physics, vol. 10, no. 47, pages 7059–7072, 2008. (Cited on page 44.)

[Bénichou *et al.* 2010a] O Bénichou, C Chevalier, J Klafter, B Meyer and R Voituriez. *Geometry-controlled kinetics.* Nat Chem, vol. 2, no. 6, pages 472–477, jun 2010. (Cited on pages 3 and 26.)

[Bénichou *et al.* 2010b] O. Bénichou, C. Chevalier, J. Klafter, B. Meyer and R. Voituriez. *Geometry-controlled kinetics.* Nat Chem, vol. 2, no. 6, pages 472–477, 06 2010. (Cited on page 152.)

[Bénichou *et al.* 2011] O. Bénichou, C. Loverdo, M. Moreau and R. Voituriez. *Intermittent search strategies.* Reviews of Modern Physics, vol. 83, no. 1, pages 81–129, 03 2011. (Cited on page 44.)

[Berg *et al.* 1981] O. G. Berg, R. B. Winter and P. H. von Hippel. *Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory.* Biochemistry, vol. 20, page 6929, 1981. (Cited on pages 1, 3, 25 and 26.)

[Bicout & Burkhardt 2000] Dominique J. Bicout and Theodore W. Burkhardt. *Absorption of a randomly accelerated particle: gambler's ruin in a different game.* Journal of Physics A: Mathematical and General, vol. 33, no. 39, pages 6835–6841, 2000. (Cited on pages 12, 70, 78 and 172.)

[Biroli *et al.* 2007] Giulio Biroli, Jean-Philippe Bouchaud and Marc Potters. *Extreme value problems in random matrix theory and other disordered systems.* Journal of Statistical Mechanics: Theory and Experiment, vol. 2007, no. 07, pages P07019–P07019, jul 2007. (Cited on page 117.)

[Black & Scholes 1973] Fischer Black and Myron Scholes. *The Pricing of Options and Corporate Liabilities.* Journal of Political Economy, vol. 81, no. 3, pages 637–654, 1973. (Cited on pages 1 and 25.)

[Blanco & Fournier 2003] S. Blanco and R. Fournier. *An invariance property of diffusive random walks.* EPL (Europhysics Letters), vol. 61, no. 2, pages 168–173, 2003. (Cited on page 149.)

[Blumenthal *et al.* 1961] R. M. Blumenthal, R K Getoor and D B Ray. *On the distribution of first hits for the symmetric stable processes.* Transactions of the American Mathematical Society, vol. 99, no. 3, pages 540–554, 1961. (Cited on pages 98 and 142.)

[Borodin & Salminen 1996] Andrei N. Borodin and Paavo Salminen. Handbook of Brownian Motion — Facts and Formulae. Birkhäuser Basel, 1996. (Cited on pages 74, 106, 109, 113, 123 and 131.)

[Bouchaud & Georges 1990] Jean-Philippe Bouchaud and Antoine Georges. *Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications.* Physics Reports, vol. 195, no. 4-5, pages 127–293, 1990. (Cited on pages 2, 8, 25, 56, 57 and 91.)

[Bouchaud & Mézard 1997] Jean-Philippe Bouchaud and Marc Mézard. *Universality classes for extreme-value statistics.* Journal of Physics A: Mathematical and General, vol. 30, no. 23, pages 7997–8015, dec 1997. (Cited on page 117.)

[Bray *et al.* 2013] Alan J. Bray, Satya N. Majumdar and Grégory Schehr. *Persistence and first-passage properties in nonequilibrium systems.* Advances in Physics, vol. 62, no. 3, pages 225–361, 2013/07/04 2013. (Cited on pages 12, 69, 75 and 77.)

[Burkhardt 2007] Theodore W Burkhardt. *The random acceleration process in bounded geometries.* Journal of Statistical Mechanics: Theory and Experiment, vol. 2007, no. 07, 2007. (Cited on page 78.)

[Caravenna & Chaumont 2008] Francesco Caravenna and Loïc Chaumont. *Invariance principles for random walks conditioned to stay positive.* Annales de l'institut Henri Poincare (B) Probability and Statistics, vol. 44, no. 1, pages 170–190, 2008. (Cited on pages 94, 97 and 108.)

[Caravenna 2005] Francesco Caravenna. *A local limit theorem for random walks conditioned to stay positive.* Probability Theory and Related Fields, vol. 133, no. 4, pages 508–530, 2005. (Cited on page 94.)

[Cavagna *et al.* 2010] Andrea Cavagna, Alessio Cimarelli, Irene Giardina, Giorgio Parisi, Raffaele Santagati, Fabio Stefanini and Massimiliano Viale. *Scale-free correlations in starling flocks.* Proceedings of the National Academy of Sciences, vol. 107, no. 26, pages 11865–11870, jun 2010. (Cited on pages 1 and 25.)

[Chaumont 1997] Loïc Chaumont. *Excursion normalisée, méandre et pont pour les processus de Lévy stables.* Bulletin Des Sciences Mathematiques, vol. 121, pages 377–403, 1997. (Cited on page 109.)

[Chechkin & Sokolov 2018] A. Chechkin and I. M. Sokolov. *Random Search with Resetting: A Unified Renewal Approach.* Physical Review Letters, vol. 121, no. 5, pages 050601–, 08 2018. (Cited on page 40.)

[Chen *et al.* 2009] Jeng Tzong Chen, Ming Hong Tsai and Chein Shan Liu. *Conformal mapping and bipolar coordinate for eccentric Laplace problems.* Computer Applications in Engineering Education, vol. 17, no. 3, pages 314–322, 2009. (Cited on page 145.)

[Chevalier *et al.* 2011] C. Chevalier, O. Bénichou, B. Meyer and R. Voituriez. *First-passage quantities of Brownian motion in a bounded domain with multiple targets: A unified approach.* Journal of Physics A: Mathematical and Theoretical, vol. 44, no. 2, page 25002, 2011. (Cited on pages 9 and 48.)

[Cheviakov *et al.* 2010] Alexei F. Cheviakov, Michael J. Ward and Ronny Straube. *An Asymptotic Analysis of the Mean First Passage Time for Narrow Escape Problems: Part II: The Sphere.* Multiscale Modeling & Simulation, vol. 8, no. 3, pages 836–870, 01 2010. (Cited on page 69.)

[Chung 1976] Kai Lai Chung. *Excursions in Brownian motion.* Arkiv för matematik, vol. 14, no. 1, 1976. (Cited on page 117.)

[Comtet & Majumdar 2005] Alain Comtet and Satya N. Majumdar. *Precise asymptotics for a random walker's maximum.* Journal of Statistical Mechanics: Theory and Experiment, no. 6, 2005. (Cited on page 120.)

[Condamin *et al.* 2005] S Condamin, O Benichou and M Moreau. *First-passage times for random walks in bounded domains.* Phys Rev Lett, vol. 95, no. 26, page 260601, 2005. (Cited on pages 9, 48 and 50.)

[Condamin *et al.* 2007a] S Condamin, O Benichou and M Moreau. *Random walks and Brownian motion: a method of computation for first-passage times and related quantities in confined geometries.* Phys Rev E, vol. 75, no. 2 Pt 1, page 021111, 2007. (Cited on pages 50 and 153.)

[Condamin *et al.* 2007b] S Condamin, O Bénichou, V Tejedor, R Voituriez and J Klafter. *First-passage times in complex scale-invariant media.* Nature, vol. 450, no. 7166, pages 77–80, Nov 2007. (Cited on pages 9, 48, 49, 52 and 69.)

[d'Alessandro *et al.* 2021] Joseph d'Alessandro, Alex Barbier-Chebbah, Victor Cellerin, Olivier Benichou, RenéMarc Mège, Raphaël Voituriez and Benoît Ladoux. *Cell migration guided by long-lived spatial memory.* Nature Communications, vol. 12, no. 1, page 4118, 2021. (Cited on pages 12 and 81.)

[Darling 1956] D. A. Darling. *The Maximum of Sums of Stable Random Variables.* Transactions of the American Mathematical Society, vol. 83, no. 1, pages 164–169, 1956. (Cited on page 118.)

[Davies & Harte 1987] R. B. Davies and D. S. Harte. *Tests for Hurst effect.* Biometrika, vol. 74, no. 1, pages 95–101, 1987. (Cited on page 171.)

[Dayan & Havlin 1992] I Dayan and S Havlin. *Number of distinct sites visited by a random walker in the presence of a trap.* Journal of Physics A: Mathematical and General, vol. 25, no. 9, page L549, may 1992. (Cited on page 37.)

[de Bruyne *et al.* 2021] Benjamin de Bruyne, Satya N. Majumdar and Grégory Schehr. *Expected maximum of bridge random walks and Levy flights.* Journal of Statistical Mechanics: Theory and Experiment, vol. 2021, no. 8, 2021. (Cited on pages 117 and 120.)

[Dhar *et al.* 2019] Abhishek Dhar, Anupam Kundu, Satya N. Majumdar, Sanjib Sabhapandit and Grégory Schehr. *Run-and-tumble particle in one-dimensional confining potentials: Steady-state, relaxation, and first-passage properties.* Physical Review E, vol. 99, no. 3, page 032132, mar 2019. (Cited on page 39.)

[Dietrich & Newsam 1997] C. R. Dietrich and G. N. Newsam. *Fast and Exact Simulation of Stationary Gaussian Processes through Circulant Embedding of the Covariance Matrix.* SIAM Journal on Scientific Computing, vol. 18, no. 4, pages 1088–1107, jul 1997. (Cited on page 171.)

[Doney 1987] R. A. Doney. *On Wiener-Hopf Factorisation and the Distribution of Extrema for Certain Stable Processes.* The Annals of Probability, vol. 15, no. 4, pages 1352 – 1362, 1987. (Cited on page 118.)

[Doney 2012] R. A. Doney. *Local behaviour of first passage probabilities.* Probability Theory and Related Fields, vol. 152, no. 3-4, pages 559–588, 2012. (Cited on pages 5, 28, 93 and 97.)

[Dvoretzky & Erdös 1951] A. Dvoretzky and P. Erdös. *Some problems on random walk in space.* In Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, pages 353–367. University of California Press, Berkeley-Los Angeles, Calif., 1951. (Cited on pages 2, 3, 6, 26, 27 and 34.)

[Dybiec *et al.* 2016] Bartłomiej Dybiec, Ewa Gudowska-Nowak and Aleksei Chechkin. *To hit or to pass it over - Remarkable transient behavior of first arrivals and passages for Lévy flights in finite domains.* Journal of Physics A: Mathematical and Theoretical, vol. 49, no. 50, 2016. (Cited on pages 56 and 96.)

[Einstein 1905] Albert Einstein. *Uber die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen.* Ann. d. Physik, 1905. (Cited on pages 1 and 25.)

[Evans & Majumdar 2011] Martin R. Evans and Satya N. Majumdar. *Diffusion with Stochastic Resetting.* Physical Review Letters, vol. 106, no. 16, page 160601, 04 2011. (Cited on pages 7 and 39.)

[Evans *et al.* 2020] Martin R Evans, Satya N Majumdar and Grégory Schehr. *Stochastic resetting and applications.* Journal of Physics A: Mathematical and Theoretical, vol. 53, no. 19, page 193001, apr 2020. (Cited on page 39.)

[Feller 1971] William Feller. An introduction to probability theory and its applications. Vol. II. Second edition. John Wiley & Sons Inc., New York, 1971. (Cited on page 91.)

[Foss S. 2015] Zachary S. Foss S. Korshunov D. An introduction to heavy-tailed and subexponential distributions. Springer New York, NY, 2015. (Cited on page 112.)

[Fréchet 1927] M. Fréchet. *Sur la loi de probabilité de l'écart maximum.* In Annales de la societe Polonaise de Mathematique, volume 6, pages 93–116, 1927. (Cited on page 117.)

[Getoor 1961] R. K. Getoor. *First Passage Times for Symmetric Stable Processes in Space.* Transactions of the American Mathematical Society, vol. 101, no. 1, 1961. (Cited on pages 21 and 147.)

[Gillis & Weiss 1970] Joseph E. Gillis and George H. Weiss. *Expected Number of Distinct Sites Visited by a Random Walk with an Infinite Variance.* Journal of Mathematical Physics, vol. 11, no. 4, pages 1307–1312, apr 1970. (Cited on pages 3, 6, 27 and 34.)

[Gnedenko & Kolmogorov 1955] Boris Gnedenko and Academician A. N. Kolmogorov. Limit distributions for sums of independent random variables. 1955. (Cited on page 91.)

[Gooden *et al.* 1998] Jonathon K. Gooden, Michael L. Gross, Anja Mueller, Andrei D. Stefanescu and Karen L. Wooley. *Cyclization in Hyperbranched Polymer Syntheses: Characterization by MALDI-TOF Mass Spectrometry.* Journal of the American Chemical Society, vol. 120, no. 39, pages 10180–10186, oct 1998. (Cited on pages 3 and 26.)

[Goss *et al.* 1989] S. Goss, S. Aron, J. L. Deneubourg and J. M. Pasteels. *Self-organized shortcuts in the Argentine ant.* Naturwissenschaften, vol. 76, no. 12, pages 579–581, dec 1989. (Cited on pages 2 and 26.)

[Grassberger 2017] Peter Grassberger. *Self-Trapping Self-Repelling Random Walks.* Phys. Rev. Lett., vol. 119, page 140601, Oct 2017. (Cited on pages 2 and 26.)

[Gumbel 1935] E.J. Gumbel. *Les valeurs extrêmes des distributions statistiques.* Annales de l'institut Henri Poincaré, vol. 5, no. 2, pages 115–158, 1935. (Cited on page 117.)

[Havlin & ben Avraham 1987] S. Havlin and D. ben Avraham. *Diffusion in disordered media.* Adv.in Phys., vol. 36, no. 6, page 695, 1987. (Cited on pages 3, 8, 27, 34, 50, 53, 56 and 57.)

[Haynes & Roberts 2008] C. P. Haynes and A. P. Roberts. *Global first-passage times of fractal lattices.* Phys. Rev. E, vol. 78, no. 4, page 041111, Oct 2008. (Cited on pages 3 and 26.)

[Holst 1986] Lars Holst. *On Birthday, Collectors', Occupancy and Other Classical Urn Problems.* International Statistical Review / Revue Internationale de Statistique, vol. 54, no. 1, page 15, apr 1986. (Cited on page 44.)

[Hopf 1935] Eberhard Hopf. *Mathematical Problems of Radiative Equilibrium.* Nature, vol. 135, no. 3402, pages 51–51, jan 1935. (Cited on page 92.)

[Hughes 1995] B.D. Hughes. Random walks and random environments. Oxford University Press, New York, 1995. (Cited on pages 3, 13, 27, 34, 36, 56, 58, 60, 77, 79, 90, 161, 162 and 179.)

[Ivanov 1994] V. V. Ivanov. *Resolvent method: exact solutions of half-space transport problems by elementary means.* Astronomy and Astrophysics, vol. 286, no. 1, 1994. (Cited on pages 5, 14, 28, 92, 104 and 179.)

[Jain & Pruitt 1971] Naresh C. Jain and William E. Pruitt. *The range of transient random walk.* J. Analyse Math., vol. 24, pages 369–393, 1971. (Cited on pages 3, 6, 27 and 34.)

[Jain & Pruitt 1974] Naresh C. Jain and William E. Pruitt. *Further limit theorems for the range of random walk.* J. Analyse Math., vol. 27, pages 94–117, 1974. (Cited on pages 3, 6, 27 and 34.)

[Klinger *et al.* 2021] J. Klinger, R. Voituriez and O. Bénichou. *Distribution of the span of one-dimensional confined random processes before hitting a target.* Physical Review E, vol. 103, no. 3, pages 032107–, 03 2021. (Cited on page 33.)

[Klinger *et al.* 2022a] J. Klinger, A. Barbier-Chebbah, R. Voituriez and O. Bénichou. *Joint statistics of space and time exploration of one-dimensional random walks.* Physical Review E, vol. 105, no. 3, page 034116, mar 2022. (Cited on pages 70 and 173.)

[Klinger *et al.* 2022b] J. Klinger, R. Voituriez and O. Bénichou. *Splitting Probabilities of Symmetric Jump Processes.* Physical Review Letters, vol. 129, no. 14, page 140603, sep 2022. (Cited on pages 89, 94 and 141.)

[Klinger *et al.* 2023] J. Klinger, R. Voituriez and O. Bénichou. *Leftward, rightward, and complete exit-time distributions of jump processes.* Phys. Rev. E, vol. 107, page 054109, May 2023. (Cited on page 105.)

[Koshland 1980] D. E. Koshland. Bacterial chemotaxis as a model behavioral system. Raven, 1980. (Cited on pages 4, 13 and 28.)

[Kuznetsov 2011] Alexey Kuznetsov. *On extrema of stable processes.* The Annals of Probability, vol. 39, no. 3, may 2011. (Cited on page 118.)

[Kwaśnicki 2011] Mateusz Kwaśnicki. *Spectral analysis of subordinate Brownian motions on the half-line.* Studia Mathematica, vol. 206, no. 3, pages 211–271, 2011. (Cited on page 107.)

[Kwaśnicki 2012] Mateusz Kwaśnicki. *Eigenvalues of the fractional Laplace operator in the interval.* Journal of Functional Analysis, vol. 262, no. 5, pages 2379–2402, 2012. (Cited on pages 96, 107, 110 and 185.)

[Kwasnicki 2017] Mateusz Kwasnicki. *Ten equivalent definitions of the fractional laplace operator.* Fractional Calculus and Applied Analysis, vol. 20, no. 1, pages 7–51, 2017. (Cited on pages 106 and 185.)

[Kyprianou & Pardo 2022] Andreas E. Kyprianou and Juan Carlos Pardo. Stable Lévy Processes via Lamperti-Type Representations. Cambridge University Press, mar 2022. (Cited on pages 19, 92 and 137.)

[Kyprianou *et al.* 2014] Andreas E. Kyprianou, Juan Carlos Pardo and Alexander R. Watson. *Hitting distributions of α-stable processes via path censoring and self-similarity.* Ann. Probab., vol. 42, no. 1, pages 398–430, 2014. (Cited on page 21.)

[Larralde *et al.* 1992] Hernan Larralde, Paul Trunfio, Shlomo Havlin, H. Eugene Stanley and George H. Weiss. *Number of distinct sites visited by N random walkers.* Phys. Rev. A, vol. 45, pages 7128–7138, May 1992. (Cited on pages 6 and 34.)

[Levernier *et al.* 2018] N. Levernier, O. Bénichou, T. Guérin and R. Voituriez. *Universal first-passage statistics in aging media.* Physical Review E, vol. 98, no. 2, pages 022125–, 08 2018. (Cited on pages 3, 12, 26, 64 and 75.)

[Levy 1937] P. Levy. Théorie de l'addition des variables aléatoires. Gauthier-Villars, 1937. (Cited on pages 2, 17, 25 and 118.)

[Lifshitz *et al.* 1978] I. M. Lifshitz, A. Yu. Grosberg and A. R. Khokhlov. *Some problems of the statistical physics of polymer chains with volume interaction.* Rev. Mod. Phys., vol. 50, pages 683–713, Jul 1978. (Cited on pages 1 and 25.)

[Lopez *et al.* 2023] J. P. Lopez, A. S. M. Macedo, M. O. Araújo and T. Passerat de Silans. *Dependence of Lévy-flight transmission on the starting point for photons propagating in atomic vapors.* Phys. Rev. A, vol. 107, page 013501, Jan 2023. (Cited on pages 100 and 102.)

[Majumdar *et al.* 2006] Satya N. Majumdar, Alain Comtet and Robert M. Ziff. *Unified Solution of the Expected Maximum of a Discrete Time Random Walk and the Discrete Flux to a Spherical Trap.* Journal of Statistical Physics, vol. 122, no. 5, pages 833–856, 2006. (Cited on page 179.)

[Majumdar *et al.* 2008] Satya N. Majumdar, Julien Randon-Furling, Michael J. Kearney and Marc Yor. *On the time to reach maximum for a variety of constrained Brownian motions.* Journal of Physics A: Mathematical and Theoretical, vol. 41, no. 36, page 365005, sep 2008. (Cited on pages 17, 117, 118, 119 and 122.)

[Majumdar *et al.* 2010a] Satya N. Majumdar, Alain Comtet and Julien Randon-Furling. *Random Convex Hulls and Extreme Value Statistics.* Journal of Statistical Physics, vol. 138, no. 6, pages 955–1009, 2010. (Cited on pages 3, 26 and 99.)

[Majumdar *et al.* 2010b] Satya N. Majumdar, Alberto Rosso and Andrea Zoia. *Hitting Probability for Anomalous Diffusion Processes.* Physical Review Letters, vol. 104, no. 2, 01 2010. (Cited on pages 76, 98 and 142.)

[Majumdar *et al.* 2012] Satya N. Majumdar, Grégory Schehr and Gregor Wergen. *Record statistics and persistence for a random walk with a drift.* Journal of Physics A: Mathematical and Theoretical, vol. 45, no. 35, 2012. (Cited on pages 24 and 157.)

[Majumdar *et al.* 2017] Satya N. Majumdar, Philippe Mounaix and Grégory Schehr. *Survival probability of random walks and Lévy flights on a semi-infinite line.* Journal of Physics A: Mathematical and Theoretical, vol. 50, no. 46, page 465002, nov 2017. (Cited on pages 5, 28, 93, 94, 98 and 122.)

[Majumdar *et al.* 2020] Satya N. Majumdar, Arnab Pal and Grégory Schehr. *Extreme value statistics of correlated random variables: A pedagogical review.* Physics Reports, vol. 840, pages 1–32, 2020. (Cited on page 120.)

[Majumdar 2010] Satya N. Majumdar. *Universal first-passage properties of discrete-time random walks and Lévy flights on a line: Statistics of the global maximum and records.* Physica A: Statistical Mechanics and its Applications, vol. 389, no. 20, pages 4299–4316, 2010. Proceedings of the 12th International Summer School on Fundamental Problems in Statistical Physics. (Cited on page 17.)

[Mandelbrot & Van Ness 1968] Benoit B. Mandelbrot and John W. Van Ness. *Fractional Brownian Motions, Fractional Noises and Applications.* SIAM Review, vol. 10, no. 4, pages 422–437, oct 1968. (Cited on pages 2, 12 and 26.)

[Mariz *et al.* 2001] A. M. Mariz, F. Van Wijland, H. J. Hilhorst, S. R. Gomes and C. Tsallis. *Statistics of the one-dimensional Riemann walk.* Journal of Statistical Physics, vol. 102, no. 1-2, pages 259–283, 2001. (Cited on pages 3, 6, 27, 34 and 56.)

[Matthews 2002] M. Matthews. *A Brownian Model for Recurrent Earthquakes.* Bulletin of The Seismological Society of America - BULL SEISMOL SOC AMER, vol. 92, pages 2233–2250, 08 2002. (Cited on pages 1 and 25.)

[Metzler *et al.* 2014] R. Metzler, G. Oshanin and S. Redner. First passage problems: recent advances. World Scientific, Singapore, 2014. (Cited on page 69.)

[Meyer *et al.* 2011] B. Meyer, C. Chevalier, R. Voituriez and O. Bénichou. *Universality classes of first-passage-time distribution in confined media.* Physical Review E, vol. 83, no. 5, page 051116, 2011. (Cited on page 152.)

[Milne 1921] E. A. Milne. *Radiative Equilibrium in the Outer Layers of a Star: the Temperature Distribution and the Law of Darkening.* Monthly Notices of the Royal Astronomical Society, vol. 81, no. 5, pages 361–375, mar 1921. (Cited on pages 4 and 28.)

[Molchan 1999] GM Molchan. *Maximum of a fractional Brownian motion: Probabilities of small values.* Communications In Mathematical Physics, vol. 205, no. 1, pages 97–111, August 1999. (Cited on pages 12 and 77.)

[Monthus & Bouchaud 1996] Cécile Monthus and Jean-Philippe Bouchaud. *Models of traps and glass phenomenology.* Journal of Physics A: Mathematical and General, vol. 29, no. 14, pages 3847–3869, jul 1996. (Cited on pages 2 and 26.)

[Montroll & Weiss 1965] Elliott W. Montroll and George H. Weiss. *Random Walks on Lattices. II.* Journal of Mathematical Physics, vol. 6, no. 2, pages 167–181, 1965. (Cited on pages 2 and 26.)

[Montroll 1969] Elliott W. Montroll. *Random Walks on Lattices. III. Calculation of First-Passage Times with Application to Exciton Trapping on Photosynthetic Units.* Journal of Mathematical Physics, vol. 10, no. 4, pages 753–765, 1969. (Cited on pages 3 and 26.)

[Mori *et al.* 2020a] Francesco Mori, Pierre Le Doussal, Satya N. Majumdar and Grégory Schehr. *Universal Survival Probability for a d-Dimensional Run-and-Tumble Particle.* Physical Review Letters, vol. 124, no. 9, pages 1–6, 2020. (Cited on pages 101 and 103.)

[Mori *et al.* 2020b] Francesco Mori, Satya N. Majumdar and Grégory Schehr. *Distribution of the time between maximum and minimum of random walks.* Physical Review E, vol. 101, no. 5, page 52111, 2020. (Cited on pages 94, 118 and 124.)

[Mori *et al.* 2021] Francesco Mori, Satya N. Majumdar and Grégory Schehr. *Distribution of the time of the maximum for stationary processes.* Epl, vol. 135, no. 3, pages 1–13, 2021. (Cited on pages 17, 70, 118 and 131.)

[Noh & Rieger 2004] Jae Dong Noh and Heiko Rieger. *Random Walks on Complex Networks.* Physical Review Letters, vol. 92, no. 11, pages 118701–4, 2004. (Cited on page 49.)

[O'Shaughnessy & Procaccia 1985] O'Shaughnessy and I Procaccia. *Analytical solutions for diffusion on fractal objects.* Phys Rev Lett, vol. 54, no. 5, pages 455–458, Feb 1985. (Cited on pages 51 and 53.)

[Pal & Prasad 2019] Arnab Pal and V. V. Prasad. *First passage under stochastic resetting in an interval.* Physical Review E, vol. 99, no. 3, pages 032123–, 03 2019. (Cited on pages 40 and 43.)

[Patteson *et al.* 2015] A. E. Patteson, A. Gopinath, M. Goulian and P. E. Arratia. *Running and tumbling with E. coli in polymeric solutions.* Scientific Reports, vol. 5, no. 1, page 15761, oct 2015. (Cited on pages 19 and 136.)

[Pearson 1905] K Pearson. *The Problem of the Random Walk.* Nature, vol. 72, no. 1865, pages 294–294, jul 1905. (Cited on pages 1 and 25.)

[Pólya 1921] Georg Pólya. *Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz.* Mathematische Annalen, vol. 84, no. 1-2, pages 149–160, mar 1921. (Cited on pages 1 and 25.)

[Randon-Furling & Majumdar 2007] Julien Randon-Furling and Satya N Majumdar. *Distribution of the time at which the deviation of a Brownian motion is maximum before its first-passage time.* Journal of Statistical Mechanics: Theory and Experiment, vol. 2007, no. 10, page P10008, 2007. (Cited on pages 17, 70, 117, 118 and 128.)

[Randon-Furling *et al.* 2022] Julien Randon-Furling, Paavo Salminen and Pierre Vallois. *On a first hit distribution of the running maximum of Brownian motion.* Stochastic Processes and their Applications, vol. 150, pages 1204–1221, 2022. (Cited on page 118.)

[Redner 2001] S. Redner. A guide to first- passage processes. Cambridge University Press, Cambridge, England, 2001. (Cited on pages 19, 37, 43, 69, 75, 92, 120, 137, 143, 145 and 149.)

[Régnier *et al.* 2022] Léo Régnier, Maxim Dolgushev, S. Redner and Olivier Bénichou. *Complete visitation statistics of one-dimensional random walks.* Phys. Rev. E, vol. 105, page 064104, Jun 2022. (Cited on page 131.)

[Reynolds 2006] A. M. Reynolds. *On the intermittent behaviour of foraging animals.* Europhysics Letters (EPL), vol. 75, no. 4, pages 517–520, 2006. (Cited on pages 2 and 25.)

[Rosenstock 1961] Herbert B. Rosenstock. *Random Walks with Spontaneous Emission.* Journal of the Society for Industrial and Applied Mathematics, vol. 9, no. 2, pages 169–188, jun 1961. (Cited on pages 1, 3, 6, 25, 26, 34, 46, 79 and 85.)

[Rouse 1953] Prince E. Rouse Jr. *A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers.* The Journal of Chemical Physics, vol. 21, no. 7, pages 1272–1280, 1953. (Cited on page 90.)

[Rupprecht *et al.* 2016] Jean François Rupprecht, Olivier Bénichou and Raphael Voituriez. *Optimal search strategies of run-and-tumble walks.* Physical Review E, vol. 94, no. 1, pages 1–9, 2016. (Cited on page 103.)

[Sapozhnikov 1994] V B Sapozhnikov. *Self-attracting walk with nu <1/2.* Journal of Physics A: Mathematical and General, vol. 27, no. 6, pages L151–L153, mar 1994. (Cited on pages 12, 70 and 81.)

[Savo *et al.* 2017] Romolo Savo, Romain Pierrat, Ulysse Najar, Rémi Carminati, Stefan Rotter and Sylvain Gigan. *Observation of mean path length invariance in light-scattering media.* Science, vol. 358, no. 6364, pages 765–768, nov 2017. (Cited on pages 1, 19, 25, 136, 149 and 154.)

[Schuss *et al.* 2007] Z Schuss, A Singer and D Holcman. *The narrow escape problem for diffusion in cellular microdomains.* Proc Natl Acad Sci U S A, vol. 104, no. 41, pages 16098–103, Oct 2007. (Cited on page 69.)

[Spitzer 1956] Frank Spitzer. *A Combinatorial Lemma and Its Application to Probability Theory.* Transactions of the American Mathematical Society, vol. 82, no. 2, page 323, 1956. (Cited on page 117.)

[Stana & Lythe 2022] Remus Stana and Grant Lythe. *Diffusion in a disk with inclusion: Evaluating Green's functions.* PLOS ONE, vol. 17, no. 4, pages 1–9, 04 2022. (Cited on page 144.)

[Stone 1967] Charles Stone. *On local and ratio limit theorems.* Proc. 5th Berkeley Sympos. math. Statist. Probab., Univ. Calif. 1965/1966, 2, Part 2, 217-224 (1967)., 1967. (Cited on page 91.)

[Tailleur & Cates 2008] J. Tailleur and M. E. Cates. *Statistical Mechanics of Interacting Run-and-Tumble Bacteria.* Physical Review Letters, vol. 100, no. 21, page 218103, may 2008. (Cited on pages 103 and 154.)

[Tejedor *et al.* 2011] V Tejedor, O Bénichou, Ralf Metzler and R Voituriez. *Residual mean first-passage time for jump processes: theory and applications to Lévy flights and fractional Brownian motion.* Journal of Physics A: Mathematical and Theoretical, vol. 44, no. 25, page 255003, 2011. (Cited on page 59.)

[Tejedor *et al.* 2012] Vincent Tejedor, Raphael Voituriez and Olivier Bénichou. *Optimizing Persistent Random Searches.* Physical Review Letters, vol. 108, no. 8, pages 088103–, 02 2012. (Cited on page 38.)

[Van Kampen 1992] N.G. Van Kampen. Stochastic processes in physics and chemistry, third edition. North-Holland personnal library, 1992. (Cited on pages 4, 19, 28, 36, 39, 48, 70, 90, 95, 99 and 148.)

[Vezzani *et al.* 2019] Alessandro Vezzani, Eli Barkai and Raffaella Burioni. *Single-big-jump principle in physical modeling.* Physical Review E, vol. 100, no. 1, pages 1–20, 2019. (Cited on pages 24, 112, 120 and 157.)

[Vezzani *et al.* 2020] Alessandro Vezzani, Eli Barkai and Raffaella Burioni. *Rare events in generalized Lévy Walks and the Big Jump principle.* Scientific Reports, vol. 10, no. 1, page 2732, 2020. (Cited on page 112.)

[Weiss *et al.* 1992] George H. Weiss, Ido Dayan, Shlomo Havlin, James E. Kiefer, Hernan Larralde, H.Eugene Stanley and Paul Trunfio. *Some recent variations on the expected number of*

distinct sites visited by an n-step random walk. Physica A: Statistical Mechanics and its Applications, vol. 191, no. 1-4, pages 479–490, dec 1992. (Cited on page 34.)

[Weiss 1994] G.H. Weiss. Aspects and applications of the random walk. Amsterdam, Netherlands: North-Holland, 1994. (Cited on pages 7, 34 and 38.)

[Widom 1961] Harold Widom. *Stable Processes and Integral Equations.* Transactions of the American Mathematical Society, vol. 98, no. 3, page 430, mar 1961. (Cited on pages 98 and 142.)

[Wijland *et al.* 1997] F. van Wijland, S. Caser and H. J. Hilhorst. *Statistical properties of the set of sites visited by the two-dimensional random walk.* Journal of Physics A: Mathematical and General, vol. 30, no. 2, pages 507–531, 1997. (Cited on pages 2 and 26.)

[Wood & Chan 1994] Andrew T. A. Wood and Grace Chan. *Simulation of Stationary Gaussian Processes in [0,1] d.* Journal of Computational and Graphical Statistics, vol. 3, no. 4, page 409, dec 1994. (Cited on page 171.)

[Wright 1931] S Wright. *Evolution in Mendelian Populations.* Genetics, vol. 16, no. 2, pages 97–159, mar 1931. (Cited on pages 3 and 27.)

[Yuste & Acedo 1999] S. B. Yuste and L. Acedo. *Territory covered by N random walkers.* Physical Review E, vol. 60, no. 4, pages R3459–R3462, oct 1999. (Cited on page 34.)

[Yuste *et al.* 2013] S. B. Yuste, E. Abad and Katja Lindenberg. *Exploration and Trapping of Mortal Random Walkers.* Physical Review Letters, vol. 110, no. 22, page 220603, may 2013. (Cited on page 34.)

[Zoia *et al.* 2007] A. Zoia, A. Rosso and M. Kardar. *Fractional Laplacian in bounded domains.* Physical Review E, vol. 76, no. 2, 08 2007. (Cited on pages 106 and 185.)

[Zoia *et al.* 2009] Andrea Zoia, Alberto Rosso and Satya N. Majumdar. *Asymptotic Behavior of Self-Affine Processes in Semi-Infinite Domains.* Physical Review Letters, vol. 102, no. 12, pages 120602–4, 03 2009. (Cited on page 3.)