



HAL
open science

Path-wise study of singular diffusions

Alexis Anagnostakis

► **To cite this version:**

Alexis Anagnostakis. Path-wise study of singular diffusions. Probability [math.PR]. Université de Lorraine, 2022. English. NNT : 2022LORR0164 . tel-03846475v3

HAL Id: tel-03846475

<https://hal.science/tel-03846475v3>

Submitted on 23 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Étude trajectorielle de diffusions singulières

THÈSE

présentée et soutenue publiquement le 18 Octobre 2022

pour l'obtention du

Doctorat de l'Université de Lorraine

(mention mathématiques)

par

Alexis Anagnostakis

Composition du jury

Encadrants : Antoine LEJAY
Denis VILLEMONTAIS

Président : Victor NISTOR

Rapporteurs : Noufel FRIKHA
Michaela SZÖLGYENYI

Examineurs : Olga ARYASOVA
Christophette BLANCHET-SCALLIET



DOCTORAT

 DOCTORAT DE L'UNIVERSITÉ DE LORRAINE
 RAPPORT DE SOUTENANCE

Nom et prénom du doctorant : ANAGNOSTAKIS ALEXIS

Titre de la thèse : Etudes trajectorielle des diffusions générales

Doctorat : Mathématiques

Date de la soutenance : 18 octobre 2022

Alexis Anagnostakis a fait une présentation bien construite et pédagogique de l'ensemble de ses travaux de thèse. Les travaux d'Alexis portent sur les diffusions singulières, les principaux résultats théoriques consistent en 1. Des critères originaux d'existence et d'unicité des équations différentielles stochastiques collantes; 2. Un résultat de convergence pour un schéma d'approximation par des chaînes de Markov en espace temps; 3. La mise au point d'un estimateur pour le temps local des équations différentielles collantes, inspiré par Jacod, et l'introduction pour la première fois d'un estimateur du paramètre de stickiness. Durant sa présentation, Alexis a illustré ses résultats théoriques par des exemples numériques très pertinent et convainquants, ce qui démontre une maîtrise certaine des aspects computationnels. De plus, Alexis a fait preuve de recul sur l'ensemble de ses travaux, et a représenté plusieurs pistes de recherches futures. Le jury a posé de nombreuses questions, auxquelles Alexis a apporté des réponses complètes et convaincantes, révélant une culture mathématique étendue au delà de son domaine de spécialité.

Président du jury : **M. VICTOR NISTOR** (A COMPLETER)

Membres du jury :

Nom	Signature	Nom	Signature
Antoine LEJAY		Denis VILLEMONAIS	
Noufel FRIKHA		Michaela SZÖLGYENYI	
Christophette BLANCHET-SCALLIET		Olga ARYASOVA	
Victor NISTOR			

Si le rapport comporte plusieurs pages ou s'il est rédigé sur un document distinct, il devra être paraphé sur chaque page et signé par le Président du jury.

Pour toutes ces raisons, le jury est convaincu qu'Alexis Anagnostakis a toutes les qualités requises pour devenir un excellent chercheur ou enseignant-chercheur, et lui désigne le titre de docteur de l'Université de Lorraine.

Résumé

Le sujet principale de cette thèse est l'étude des processus de diffusions singuliers et en particulier des diffusions collantes.

Introduites par Feller dans les années 50, les diffusions collantes sont apparues comme un cas particulier de conditions de bord dans la description analytique des diffusions générales. Leurs trajectoires passent un temps positif sur des points de l'espace d'états leur donnant l'apparence d'y coller. Quand de tels points se trouvent sur des bords atteignables de l'espace d'états on parle de réflexion collante.

La première contribution est l'approximation du temps local des diffusions d'Itô collantes. Nous définissons ce type de processus et prouvons leur description trajectoirelle. On prouve la convergence d'une classe de fonctionnelles haute-fréquence de la trajectoire du mouvement Brownien collant vers son temps local en 0. On étend avec des arguments trajectoires aux diffusions d'Itô collantes. On définit un estimateur de la *stickiness* basé sur l'approximation du temps local, puis on prouve sa consistance. On donne des résultats numériques dans le cas du mouvement Brownien collant.

La deuxième contribution de cette thèse est l'approximation de tout processus de diffusion par des marches aléatoires à valeurs dans des grilles dont les moments correspondent avec ceux du vrai processus. On appelle ces processus d'approximation *Space-Time Markov Chain Approximation* ou *STMCA* car ce sont des chaînes de Markov en espace-temps. Une particularité de ce type d'approximation est qu'on arrive à répliquer des dynamiques collantes de façon assez naturelle. On montre que avec un choix adapté de la grille on a une vitesse de convergence optimale en loi de cette approximation quand le pas de la grille tend vers 0. On appelle ce procédé *grid tuning*. On donne des résultats numériques ou on illustre la convergence en loi des processus d'approximation et la flexibilité de l'algorithme sur le problème d'approximation du temps local.

Abstract

The main object of this thesis is the study of singular diffusion processes with a focus on sticky diffusions.

Sticky diffusions were first introduced by Feller in the fifties as a case of boundary condition that can arise in the analytic description of a diffusion. Their paths spend positive amount of time at points of the state-space, giving them the appearance to “stick” on these points. When such points are located at an attainable boundary of the state-space of the process, we call it sticky reflection.

The first contribution of this thesis is to provide a suitable approximation of the local time of a sticky Itô diffusion, with statistical applications in view. We define the notion of sticky Itô diffusion and prove their path-wise descriptions. We prove that the local time of the sticky Brownian motion can be approximated by a class of high-frequency path functionals. We use the path-wise characterization to extend the result to non-explosive Itô diffusions. We prove the consistency of a stickiness estimator based on the local time approximation. We give numerical results on the stickiness estimation of a sticky Brownian motion.

The second contribution of this thesis is an approximation in law of any one-dimensional diffusion by a grid-valued conditional moment-matching random walk. The convergence occurs as the maximal grid step goes to 0. We call this type of approximation *Space-Time Markov Chain Approximation* or *STMCA*. We also show how one can achieve optimal convergence rate by suitable choices of grids. We call *grid tuning* the process of computing such a grid. One can use STMCA to set up approximation schemes for any one-dimensional diffusion process. We give various illustrated approximations examples of diffusions even in the presence of sticky behavior, discontinuous or degenerate coefficients.

MSC2020: Primary 60F17, 60J60, 60J55; Secondary 60J10

Keywords: Sticky Brownian motion, sticky boundary, Itô diffusion, local time approximation, parametric estimation, generalized diffusion, Markov chain approximation.

Acknowledgements

I experienced this thesis as an adventure and not a short one. And in any such endeavor, I believe one must find balance between creativity and consistency. So I'd like to thank everyone who has showed love, care and trust and thus helped me build up resilience and consistency. On the other hand, I'd like to thank everyone who treated me harshly, whether for constructive criticism or for other reasons. These behaviors increased my drive to succeed, pushed me even further beyond my capacities and induced these sweet creativity surges I occasionally crave.

First, I'd like to thank my PhD advisors and mentors Antoine Lejay and Denis Villemonais. Without them this work would not have been possible. Not only did they provide me with such an interesting research subject, but their attention for rigor, their knowledge and work ethic were all major contributing factors to the completion of this thesis. I can tell with confidence that they made me a better mathematician, a better scientist and most of all a better person.

Second thanks go to my family. Their constant support was the least touching and I always love their effort to understand my job.

I also thank Phuc Ngo, for her friendship, advices, positive attitude and attention.

Then, I would like to thank my old professors at the University of Crète and the recently re-branded Université Pierre et Marie Curie. In particular, Souzanna Papadopoulou, Nikos Frantzikinakis, Gilles Pagès and Markos Katsoulakis for their support, recommendations and trust in my abilities.

The last thanks go to all the lab, my home and family for the past four years. In particular, I'd like to thank: My colleague and friend Sara Mazzonetto for sharing her ideas and experience on the subject and for collaborating on the furthering of this work. Lionel Lenôtre, for his friendship and mentoring at the launch of my thesis. Pierre Etoré for exposing me to new research topics and for offering me collaboration opportunities. My 'co-bureau' Christophe Reyge for his company during all this time, I always find our little chats fun, interesting and comforting. Thomas Mazzanet and Pierre Popoli for being such easy-going and social persons, Thomas in particular for beating me out of chess for good. The enseignants-chercheurs who trusted me with teaching their students. In particular Rémi Peyre, for bringing life to the 'séminaire d'équipe' by exhibiting genuine interest and remarks on each and every presented subject. Last but not least, the lab head Anne Gégout-Petit, our Inria team leader Madalina Deaconu and the IT head Didier Gemerlé for the lab and team management and the latter for his patience and expertise.

“I dream my painting and I paint my dream.”

– Vincent van Gogh

Contents

Acknowledgements	v
Introduction (Français)	3
1 Contexte	3
2 Contributions	4
3 Organisation de cette thèse	6
Introduction (English)	9
1 Context	9
2 Contributions	10
3 Organization of the dissertation	12
1 An introduction to the analytic theory of diffusion processes	13
1.1 First notions	13
1.1.1 Markov property	13
1.1.2 Law and semi-group of a diffusion	14
1.2 Analytic characterization	16
1.2.1 Hille-Yosida theory	16
1.2.2 Adaptation to regular diffusions	18
1.2.3 Speed measure & scale function	19
1.3 Probability transition kernel	23
1.3.1 Definitions	24
1.3.2 Backward & forward equations	25
1.4 Additional useful results	26

1.4.1	Diffusions and martingales	26
1.4.2	Diffusions as time changed Brownian motions	28
1.5	The standard Brownian motion	28
1.5.1	Analytical characterization	29
2	Singular diffusions and the sticky Brownian motion	31
2.1	Diffusion process zoology	31
2.1.1	Stochastic differential equations and Itô diffusions	31
2.1.2	Boundary classification	34
2.1.3	Singular diffusions	36
2.2	The sticky Brownian motion	39
2.2.1	Definition	39
2.2.2	Properties	40
2.2.3	Probability transition kernel computation	42
2.2.4	Asymptotics	45
3	Sticky Itô diffusions	47
3.1	Definition	47
3.2	Path-wise characterization	49
3.3	Path-wise results	53
4	Local time approximation of sticky diffusions	55
4.1	Introduction and main results	55
4.2	The case of the sticky Brownian motion	58
4.2.1	Preliminary results	58
4.2.2	Proof in the sticky Brownian case	59
4.3	Proofs of the main results	65
4.3.1	Proof of Theorem 4.1.3	65
4.3.2	Proof of Corollary 4.1.4	66

5	General diffusion processes	
	as the limit of time-space Markov chains	69
5.1	Introduction	69
5.2	The Space-Time Markov Chain Approximation and its properties	70
5.2.1	The approximation scheme	70
5.2.2	Convergence rate for the general case	75
5.2.3	Grid tuning	76
5.2.4	The case of diffusions with boundary conditions	78
5.2.5	Markovian embedding	79
5.3	Moment characterization of conditional exit times	80
5.4	Convergence of the embedding times	82
5.4.1	Bounds on the conditional moments of the exit times	82
5.4.2	Convergence of the embedding times	84
5.5	Convergence rate of the Markov chain	86
5.5.1	Moment bounds	86
5.5.2	Proof of the convergence rate for a process on natural scale	88
5.6	Computations for the classical SDE case and beyond	91
6	Numerical experiments	93
6.1	Approximation in distribution	93
6.1.1	Standard Brownian motion:	93
6.1.2	Sticky Brownian motion:	97
6.1.3	Skew Brownian motion:	97
6.1.4	Ornstein-Uhlenbeck process:	97
6.1.5	Cox-Ingersoll-Ross process:	98
6.1.6	Skew and reflected Bessel process:	99
6.2	Local time approximation	99
6.2.1	Observations, conclusion	101

Appendix A	105
A.1 Strong continuity of diffusion semi-groups on $C_b(\mathbb{I})$: a counter-example	105
A.2 Semimartingale results	106
A.2.1 Change of variables in a stochastic integral	106
A.2.2 Local time scaling	106
A.2.3 Stochastic Dini's theorem	106
A.3 Joint scaling of the sticky Brownian motion	107
A.4 Conditioning on the embedded path of a diffusion	108
Bibliography	111

Introduction (Français)

1 Contexte

La découverte des processus de diffusion a été motivée par des observations de Brown [19] et d'Einstein [29] sur le mouvement de particules microscopiques en suspens dans un fluide. Le mouvement de ces particules est régi par leurs collisions avec les molécules du fluide qui engendre un comportement aléatoire. De plus, ces collisions sont si nombreuses et chaotiques qu'elles induisent un phénomène d'absence de mémoire appelé la propriété de Markov: sachant une suite d'observations de la particule à des temps différents, la dernière observation contient toute l'information prédictive sur la trajectoire future de la particule. C'est-à-dire que savoir seulement la dernière observation ou des observations passées additionnelles ne change pas nos prédictions sur la trajectoire future.

Le preuve de l'existence de tels objets a été donnée par Wiener [91] qui a donné le premier exemple de diffusion, le mouvement Brownien. Ce processus, qui a un comportement trajectorien homogène en espace et en temps, joue un rôle central dans la théorie des processus de diffusion. Feller [33] a par la suite donné une catégorisation des comportements trajectoriels de bord des diffusions en fonction des conditions latérales dans le générateur infinitésimal. Ces résultats ont permis de définir le mouvement Brownien sur le semi-espace \mathbb{R}_+ . Aussi, ils ont permis la description analytique des diffusions unidimensionnelles [34, 35, 36, 37]. Ceci a été réalisé en décrivant les diffusions par leurs actions sur l'espace des fonctions continues bornées. Actions qui peuvent être vues comme des opérations linéaires et qui, par la propriété de Markov, forment un semi-groupe. Par la théorie de Hille-Yosida des opérateurs linéaires, tout tel semi-groupe est engendré par un opérateur linéaire L appelé le générateur infinitésimal. Feller a aussi démontré la factorisation du générateur infinitésimal $L = D_m D_s$ où m est une mesure positive localement bornée et s une fonction continue et croissante. Pour une synthèse de ces résultats, voir [41].

Le motivation initiale de Feller était de voir comment des conditions de bords sur le générateur infinitésimal impactent les propriétés trajectoires d'une diffusion près du bord [33]. Dans sa tentative de les catégoriser, il a découvert que des conditions latérales de la forme

$$\frac{\rho}{2} f''(0) = f'(0+) - f'(0), \quad \frac{\rho}{2} f''(0) = f'(0+), \quad (1)$$

correspondent à un comportement collant (ou sticky) en 0. La première condition de (1) correspond à un point traversant sticky pour la diffusion et la deuxième à une réflexion sticky. Il a aussi montré que ceci correspond à un atome dans la mesure m de la factorisation $L = D_m D_s$ du générateur infinitésimal [35].

Il s'avère qu'il n'existe pas de formulation en équation différentielle stochastique classique (EDS) pour les diffusions sticky. Il est montré cependant dans [83] qu'il existe une description trajectorielle des diffusions sticky. En particulier, dans [30] il est montré que le mouvement Brownien collant X de stickiness ρ résout le système

$$\begin{aligned} dX_t &= \mathbb{1}_{X_t \neq 0} dB_t, \\ \mathbb{1}_{X_t=0} dt &= dL_t^0(X) \end{aligned}$$

ou B est un mouvement Brownien standard.

Après leur introduction, les diffusions sticky sont restés largement dans l'oubli dans la littérature probabiliste. En parallèle, les solutions d'EDS ont gagné beaucoup en popularité. Ceci a été en partie dû à l'apparition au modèle de Black-Scholes [13] conçu pour pricer et hedger (couvrir) des produits dérivés financiers. Ceci a aussi causé une ambiguïté autour du terme de diffusion. Certains auteurs utilisent le terme diffusion d'Itô pour décrire les solutions d'EDS homogènes en temps [77] et diffusion générale ou généralisée pour les processus Markov fort continus [53, 80].

Depuis peu, on est témoin d'un regain d'intérêt particulier dans les diffusions singulières et sticky. Il s'avère que ces processus répliquent des dynamiques observées dans la nature. On y trouve des applications en biologie, médecine, finance, en mécanique classique et quantique. Dans [42, 44], le mouvement de particules près d'une membrane ou une barrière sticky est étudié. Dans [20], des conditions de bord sticky sont utilisés pour modéliser la concentration d'un pathogène dans un organisme quand celle-ci est proche de 0. L'explication de ce phénomène est qu'il est difficile pour un pathogène de déclencher une infection si sa concentration est faible. Dans [76], un modèle d'Ornstein-Uhlenbeck avec réflexion sticky en 0 est proposé pour modéliser la dynamique des taux d'intérêts proche de 0. Dans [24], les auteurs utilisent un mouvement Brownien sticky pour décrire le mouvement de particules à proximité d'une source d'émission. Dans [75, 85, 12, 43, 50, 60, 60], des modèles à diffusions sticky sont proposés pour répliquer la dynamique de particules dans des colloïdes, les grosses particules dans un liquide ont tendances à coller entre-elles quand elles entrent en contact.

A part ces applications, les diffusions sticky ont un intérêt théorique. Leur étude échappe aux outils classiques utilisés pour étudier les EDS classique. De plus, elles peuvent être utilisés pour créer de nouveaux objets probabilistes, comme les couplages sticky [28].

Pour une présentation plus détaillée des diffusions sticky, voir [78]. Voir aussi [15], pour une aperçu plus large de leurs applications.

2 Contributions

Dans cette thèse, on adresse plusieurs sujets concernant les diffusions sticky: approximation du temps local, estimation du paramètre de stickiness et simulation numérique.

Le premier sujet qu'on adresse est l'approximation du temps local. Dans [55], l'auteur utilise des fonctionnelles trajectoriels haute-fréquence pour approximer le temps local de diffusions d'Itô solutions de

$$dX_t = \mu(X_t) dt + \sigma(X_t) dB_t, \tag{2}$$

ou B est un mouvement Brownien standard et (μ, σ) deux fonctions réelles qui garantissent l'existence de solution forte de (2). Cependant, ce résultat puissant ne peut pas s'appliquer à des diffusions sticky. La présence d'un atome dans la mesure de vitesse résulte en une explosion de la statistique. De plus, il est impossible de réduire le problème au mouvement Brownien standard, comme dans [55] pour les diffusions d'Itô classiques.

Dans le Chapitre 4, on prouve que si X est une diffusion d'Itô sticky en 0, g une fonction intégrable qui s'annule sur un voisinage de 0 et $\alpha \in (0, 1/2)$, alors la fonctionnelle

$$\frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g(n^\alpha X_{\frac{i-1}{n}}). \quad (3)$$

converge en probabilité au temps local de X en 0. On remarque qu'à un niveau sticky, le temps local est proportionnel au temps d'occupation et leur relation est régie par le paramètre de stickiness. On démontre que le temps d'occupation peut être approché de façon consistante par la statistique

$$\frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{X_{\frac{i-1}{n}}}. \quad (4)$$

On combine ces deux résultats pour mettre en place le premier (à notre connaissance) estimateur consistant du paramètre de stickiness qui consiste simplement à diviser la statistique (4) par la statistique (3).

Pour démontrer les résultats sur l'approximation du temps local nous utilisons une description trajectorielle des diffusions d'Itô sticky. Dans le Chapitre 3, on considère le système

$$\begin{aligned} dX_t &= \mu(X_t) \mathbb{1}_{X_t \neq 0} dt + \sigma(X_t) \mathbb{1}_{X_t \neq 0} dB_t, \\ \mathbb{1}_{X_t=0} dt &= \frac{\rho}{2} dL_t^0(X), \end{aligned} \quad (5)$$

où B est un mouvement Brownien standard. On démontre que le système admet une solution unique en loi qui est une diffusion d'Itô sticky. On démontre aussi le résultat "contraposé": que toute diffusion d'Itô sticky est solution d'un système de la forme (5)-(6). On fini par prouver des version sticky du théorème de Girsanov et du lemme d'Itô. Des résultats similaires sont formulés dans [83] dans lequel la stickiness est présentée comme un "spatial delay".

Le deuxième sujet qu'on aborde est celui de la simulation de ces processus. Le schéma numérique le plus populaire pour la simulation de diffusions est le schéma d'Euler. Il s'avère que ce schéma est mal défini pour les diffusions sticky. Pour remédier à ça, plusieurs schémas ont été proposés: Dans [4], l'auteur montre que le mouvement Brownien sticky est limite de marches aléatoires symétriques qui sont forcées en 0 à chaque passage pour un temps prédéfini. Ceci nous donne une méthode directe pour simuler ce processus. Dans [76, 73, 39, 60], les auteurs utilisent des Chaînes de Markov à temps continu (CTMC) pour simuler des processus qui résolvent une EDS et ont une barrière sticky. Une grille est défini sur l'espace d'état de la diffusion X . Le processus d'approximation est alors une CTMC à valeur dans \mathbf{g} et dont les intensités de transition sont calculées en utilisant une discrétisation du générateur infinitésimal de X sur \mathbf{g} . Dans [74], une approximation CTMC est considérée pour des

diffusions, solutions d'EDS multi-dimensionnelles qui sont réfléchies de façon sticky sur des hyperplans linéaires. Dans [7, 9], les auteurs définissent un schéma numérique basé sur une équation de différence finie où la taille du pas dépend du comportement moyen local de la mesure de vitesse. Ils peuvent de cette façon simuler des diffusions à échelle naturelle avec des points sticky qui se situent à l'intérieur de l'espace d'états.

Dans le Chapitre 5, on étend le schéma numérique défini dans [31] pour simuler tout processus de diffusion général, ainsi que ceux avec des points sticky, skew ou des conditions de bord. De plus, si on sait implémenter l'algorithme pour une diffusion non-sticky, il est direct de considérer des points sticky. On démontre que les quantités nécessaires pour l'implémentation de l'algorithme sont solution de problèmes paraboliques et admettent des représentations en intégrales définies. On peut utiliser ces représentations pour calculer ou approximer ces quantités numériquement. On prouve aussi que, si la grille est "adaptée" au processus qu'on veut approximer, alors la vitesse de convergence est optimale quand on fait tendre le pas de la grille vers 0. Cette optimalité est validée par le principe d'invariance de Donsker [26]. Dans le Chapitre 6, on donne un exemple d'application pour le problème de l'approximation du temps local du Chapitre 4. On observe que:

- les trajectoires sont approchées en distance de p -Wasserstein pour la norme $L^\infty([0, T])$,
- on peut choisir la grille sur laquelle le processus d'approximation prends ces valeurs.

Grace à ces deux faits, on peut choisir des grilles ayant une précision importante à proximité du point de stickiness pour augmenter la qualité des estimateur du temps local et de stickiness. Ce principe est illustré dans le Chapitre 6 sur un problème de benchmark de l'estimateur de stickiness. Ceci peut aussi être appliqué pour l'approximation du temps local de diffusions d'Itô non-sticky de [55].

Ces travaux ont donné lieu à deux articles [5, 6].

3 Organisation de cette thèse

La thèse est structurée de la façon suivante:

Chapitre 1: On donne la description analytique des diffusions unidimensionnelles ainsi que quelques résultats sur les martingales. On donne des caractérisations de la loi d'une diffusion: semi-groupe, résolvante, générateur infinitésimal, fonction échelle & mesure de vitesse. On présente aussi plusieurs résultats fondamentaux comme: la formule de Dynkin, le problème de martingale ou le théorème de Dubins-Schwarz. Les deux derniers résultats nous donne un lien structurel entre les diffusions et les martingales.

Chapitre 2: Ce chapitre est dédié aux diffusions singulières. On donne quels type de comportement trajectoire une diffusion peut avoir et on montre comment on peut les déduire à partir de la description analytique. On définit le mouvement Brownien collant (sticky Brownian motion) et on démontre certaines de ses propriétés. On montre les équations forward et backward satisfaites par le noyau de transition de la

diffusions. On résout ses équations pour calculer le noyau de transition du mouvement Brownien sticky.

Chapitre 3: Ce chapitre est dédié aux EDS sticky. On définit cette classe de processus et démontre leur description trajectorielle qui généralise les résultats de [30] et [76]. On utilise cette description pour démontrer des variantes sticky de quelques résultats classiques du calcul stochastique: lemme d'Itô, théorème de Girsanov.

Chapitre 4: On montre qu'on peut approcher le temps local d'une diffusion sticky, à son point de stickiness, par des fonctionnelles d'observations trajectorielles haute-fréquence de la forme (3). On donne des conditions nécessaires pour que la convergence ait lieu, puis on utilise ces résultats pour définir un estimateur consistant du paramètre de stickiness.

Chapitre 5: On prouve que l'algorithme établi dans [31], peut être utilisé pour la simulation de tout processus de diffusion unidimensionnel. En particulier, des marches aléatoires à valeurs dans des grilles, dont les probabilités et temps conditionnels de transition sont les mêmes en espérance que ceux de la diffusion X , converge en loi vers X quand la taille de la cellule maximale de la grille tend vers 0. On prouve qu'on peut borner les distances de Wasserstein de tout ordre $p \geq 1$ entre ces processus par une métrique de grilles qui dépend à la fois de la fonction échelle et de la mesure de vitesse de X . Il est donc possible d'adapter la grille pour atteindre des ordres de convergences optimaux.

Chapitre 6: Ce chapitre est dédié aux simulations numériques. On applique l'algorithme sur des diffusions qui ont différentes propriétés puis on compare les histogrammes des simulations avec les vraies densités de transitions. Quand nécessaire, nous utilisons des approximations numériques des probabilités et temps moyens de transition et on remarque que l'algorithme est toujours pertinent pour des grilles adaptées. On montre aussi les propriétés de l'approximation de temps local établie au Chapitre 4. Finalement, on montre que pour ce genre de problème, dans le contexte de simulations Monte Carlo, par un choix adapté de la grille, on peut contrôler le nombre de données observées par la statistique. Par conséquent, on peut atteindre des ordres de convergence supérieurs.

Introduction (English)

1 Context

The discovery of diffusion processes was motivated by observations, made by Brown [18] and Einstein [29], on the motion of suspended microscopic particles in a fluid. The motion of these particles is entirely induced by collisions with the fluid molecules and is hence random in nature. These collisions are so numerous and chaotic they create an absence of memory phenomena called Markov property: Given knowledge of the current position of the particle, past observations yield no additional predictive value for future positions.

It was Wiener [91] who proved the existence of such objects and exhibited the first example of diffusion, the Brownian motion. This process which has a homogeneous path-wise behavior in both space and time, plays a central role in the theory of diffusion processes. Then, Feller [33] categorized the boundary path-wise features of a diffusion in accordance with the lateral conditions in the infinitesimal generator. These results allowed in particular to define the Brownian motion defined on a semi-plane. Also, they resulted in the analytic description of one-dimensional diffusions [34, 35, 36, 37]. This was achieved by describing diffusions as actions on the space of continuous bounded functions C_b . These actions can be seen as linear operators and from the Markov property, form a semi-group. From the Hille-Yosida theory for linear operators, it turns out that each such semi-group is induced by an operator L called infinitesimal generator. Feller also proved the factorization of the infinitesimal generator $L = D_m D_s$, where m is a locally bounded positive measure and s a continuous increasing function. For a review of these results, see [41].

Feller's initial motivation was to see how boundary conditions of the infinitesimal generator affect the path-wise behavior of the underlying process near the boundary [33]. In his attempt to categorize them, he discovered that the lateral conditions of the form

$$\frac{\rho}{2} f''(0) = f'(0+) - f'(0), \quad \frac{\rho}{2} f''(0) = f'(0+),$$

correspond to stickiness and sticky reflection at 0 respectively. He also showed [35] that one can factorize the infinitesimal generators of a diffusion X as $L = D_m D_s$, where m and s are the speed measure and scale function of X (see Proposition 1.2.20). Thus, a sticky point for X corresponds to an atom in its speed measure m .

It turns out that there exists no classical SDE formulation for sticky diffusions. It is though showed independently in [11, 30] that there is a path-wise description for the sticky Brownian motion. In particular, the sticky Brownian motion X of stickiness ρ

solves the system

$$\begin{aligned} dX_t &= \mathbb{1}_{X_t \neq 0} dB_t, \\ \mathbb{1}_{X_t=0} dt &= dL_t^0(X) \end{aligned}$$

where B is a standard Brownian motion.

After their introduction, sticky diffusions were largely left unnoticed in the probabilistic literature. Moreover, stochastic differential equations or SDEs gained in popularity. This was partially due to the introduction of the Black-Scholes model [13] for pricing and hedging financial derivatives. As a result, an ambiguity was induced around the term diffusion. Some authors use terms like Itô diffusion [77] to express time-homogeneous SDE solutions and general or generalized diffusions for continuous strong Markov processes [53, 80].

Recently, we are witnessing a regain of interest in sticky processes. It turns out that they can accurately replicate dynamics found in nature. Applications range from biology, finance to quantum and classical mechanics. In [42, 44], the motion of molecules near a cell membrane or sticky wall is studied. In [20], sticky boundary conditions are used to model the concentration of a pathogen in an organism at near-zero levels. The cause of this is that it is hard for low quantities of pathogen to instigate an infection. In [76], an Ornstein-Uhlenbeck process with sticky reflection at 0 is used to describe the dynamics of interest rates near 0. In [24], the authors use a sticky Brownian motion to describe the motion of particles near a point-source of emission. In [75, 85, 12, 43, 50, 60, 60], sticky diffusions are proposed to replicate particle dynamic in colloids, coarse particles in a liquids that tend to “stick” with each-other upon contact.

Besides their applications, sticky diffusions have inherent theoretical interest. Their study escape the classical frameworks established for classical SDEs. Moreover, they are used to create new probabilistic objects like sticky couplings [28].

A more in depth presentation of the origins of sticky diffusions can be found in [78]. Also, in [15], a good overview of the applications of sticky diffusions is given.

2 Contributions

In this thesis, we address several subjects regarding sticky diffusions: local time approximation, estimation of the stickiness parameter and their numerical simulation.

The first subject we address is the approximation of the local time of a sticky diffusion at a sticky threshold. In [55], the author uses high-frequency path functionals (3) to approximate the local time of non-explosive homogeneous SDE solutions. While this result is powerful in its own, it cannot take into account sticky points. Indeed, the presence of a sticky points in the speed measure of the diffusion would result in an explosion of the statistic (3).

In Chapter 4, we prove that if X is an SDE with a sticky point at 0, if one forces the test function g to be 0 in an open interval around 0, (3) converges to the local time of X at 0. We then combine this result with an approximation of the occupation time at 0 to set up the first (to our knowing) consistent stickiness parameter estimator.

In order to prove these results, we use a path-wise description of sticky Itô diffusions. In Chapter 3, we consider the system

$$dX_t = \mu(X_t)\mathbb{1}_{X_t \neq 0} dt + \sigma(X_t)\mathbb{1}_{X_t \neq 0} dB_t, \quad (1)$$

$$\mathbb{1}_{X_t=0} dt = \frac{\rho}{2} dL_t^0(X), \quad (2)$$

where B is a standard Brownian motion. We show that uniqueness in law holds for the solution of the system (1)-(2), which is one of a sticky Itô diffusion. We also show the “inverse” result: that all sticky Itô diffusions admit a path-wise description of the form (1)-(2). We end this chapter by proving sticky versions of the Girsanov theorem and Itô’s lemma. Similar results are proven in [83], where the stickiness is called “spatial delay”.

The second subject we address is the numerical simulation of one-dimensional diffusions. The most popular numerical scheme to simulate diffusion processes is the Euler scheme. It turns out though that this scheme is not well defined for sticky diffusions. In order to address this issue, several numerical schemes were proposed for their approximation: In [4], the author showed that the sticky Brownian motion is the limit of symmetrical random walks that freeze a pre-defined amount of time every time they hit 0. This gives us a straightforward way to simulate this process. In [76, 73, 39, 60], the authors use continuous time Markov chains CTMC to simulate homogeneous SDE solutions with a reflective sticky boundary at 0. A grid \mathbf{g} is defined over the state-space of the diffusion X . The approximation process is then the \mathbf{g} -valued CTMC with jump intensities computed using a discretization of the infinitesimal generator of X over \mathbf{g} . In [74], a CTMC approximation is considered for multi-dimensional SDE diffusions that exhibit sticky reflection on a hyperplane. In [7, 9], the authors define a numerical scheme via a finite difference equation where the step magnitude depends on the mean local behavior of the speed measure. This way, they can simulate diffusions on natural scale with non-boundary sticky points.

In Chapter 5, we extend the numerical scheme defined in [31] to simulate any generalized diffusion process, including sticky ones. Moreover, if one can implement it for a non-sticky version of the diffusion, considering additional sticky points is straightforward. We prove that the involved quantities in the algorithm can be either computed by solving parabolic problems, computing definite integrals or approximate the latter numerically. We also prove that, if the grid is “adapted” to the process we want to approximate, we achieve optimal convergence rate for the approximation process. This optimality is validated by the Donsker invariance principle [26]. In Chapter 6, we give an application for the local time approximation of Chapter 4. We observe that:

- the paths of the approximation are $L^\infty([0, T])$ -close in the p -Wasserstein distance,
- we can choose the grid that is the state-space of the approximation process.

Thus, we can choose grids with increasing precision around points of stickiness to get better estimations of the local time and the stickiness. This is illustrated in Chapter 6 for the benchmarking problem of the stickiness parameter estimator. This principle can also be applied for the non-sticky local time approximation of [55].

These works gave birth to the papers [5, 6].

3 Organization of the dissertation

The thesis is organized as follows:

In Chapter 1, we give elements of the analytical theory of diffusion processes along with several results on martingales. We give different characterizations of diffusions: semi-group, resolvent family, infinitesimal generator, scale function & speed measure. We then exhibit several results like Dynkin's formula, the martingale problem and the Dubins-Schwarz theorem. The two latter results give us a structural link between diffusions and martingales.

Chapter 2 is dedicated to singular diffusions. We give various path-wise behaviors a diffusion may exhibit and show how they translate into their analytical description. We define the sticky Brownian motion and prove several of its properties. We show that the transition kernel solves the forward and backward equations. We solve these equations for the sticky Brownian motion and compute its probability transition kernel.

Chapter 3 is dedicated to sticky SDEs, processes that behave like classical SDE away from a point of stickiness. We define this class of processes analytically and prove that they always admit a path-wise description, like the ones derived in [30] and [76] for the sticky Brownian motion and the Ornstein-Uhlenbeck process with sticky reflection. We derive explicit results for these processes that rely on their path-wise formulation, similar to the ones we have for classical SDEs (Itô's lemma, Girsanov theorem).

In Chapter 4, we prove that one can approximate the local time of a sticky SDE, at its point of stickiness, by high-frequency path functionals of the form:

$$\frac{n^\alpha}{n} \sum_{i=1}^{\lfloor nt \rfloor} g(n^\alpha X_{\frac{i-1}{n}}). \quad (3)$$

We give necessary conditions for the convergence to occur and use this result to set up a consistent estimator of the stickiness parameter.

In Chapter 5, we prove that the algorithm, established in [31], can be applied for the simulation of any generalized diffusion process. In particular, grid-valued random walks whose transition probabilities and conditional transition times match the ones of a diffusion X , converge in law to X as the maximum cell size of the grid converges to 0. We prove that we can bound any p -Wasserstein distance between these processes by a grid metric that depends on both the speed measure and scale function of X . It is thus possible to adapt the grids to X in order to achieve optimal convergence rates.

Chapter 6 is dedicated to numerical experiments. We exhibit the properties of Algorithm 1 by simulating approximations of diffusions that exhibit different features and compare simulated values with the theoretical densities. When necessary, we use numerical approximations of the quantities driving the approximation process and see that the Algorithm 1 is still relevant. Last, we exhibit the properties of the local time approximation and stickiness parameter estimator established in Chapter 4. We also show that in the context of a Monte Carlo simulation, using the STMCA approximations developed in Chapter 5 by suitable choice of the grid one can control the amount of data observed by the statistic (4.1). We can thus achieve higher orders of convergence.

Chapter 1

An introduction to the analytic theory of diffusion processes

This chapter is a synthesis of known results found in [17, 27, 49, 53, 59, 71, 82].

Chapter outline: In this section, we introduce the notion of diffusion process and give analytic characterization of their law. In particular we define the scale function & speed measure characterization. Many results proved in this thesis rely on the latter. We then introduce the notion of probability transition kernel along with some elements on its computation, namely the Kolmogorov backward and forward equations. Last, we introduce the most elementary diffusion process, the Brownian motion.

1.1 First notions

1.1.1 Markov property

The main object of this thesis is the study of several aspects of singular diffusion processes. We focus in particular on the simulation and local time approximation of sticky sticky diffusions.

Diffusion processes are processes that have continuous sample paths and satisfy the strong Markov property. This property can be expressed as follows.

Definition 1.1.1. Let X be an adapted process with state-space \mathbb{I} defined on a family of filtered probability spaces $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P}_x)_{x \in \mathbb{I}}$ such that for every $x \in \mathbb{I}$, $\mathbb{P}_x(X_0 = x) = 1$. Let $(\theta_t)_{t \geq 0}$ be a family of shift operators for X . The process X is said to satisfy the strong Markov property iff for any measurable function f , stopping time τ and $s > 0$:

$$\mathbb{E}(f(X_{\tau+s}) | \mathcal{F}_\tau) \mathbb{1}_{\tau < \infty} = \mathbb{E}_{X_\tau}(f(X_s \circ \theta_\tau)) \mathbb{1}_{\tau < \infty}. \quad (1.1)$$

A process that satisfies the strong Markov property is called a strong Markov process. If (1.1) holds for all deterministic $\tau \geq 0$, then we say X satisfies the weak Markov property and we call it Markov process.

In Definition 1.1.1:

- The state-space \mathbb{I} of a process X is all the possible values taken by X .
- The family $(\mathcal{F}_t)_{t \geq 0}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P}_x)$, i.e. a family indexed by time such that for any $s < t$, $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$. A filtration expresses and expresses the observable events of the universe Ω up to time $t > 0$.
- An $(\mathcal{F}_t)_{t \geq 0}$ -adapted or adapted process is a process X such that $\{X_s \in A\} \in \mathcal{F}_t$ for all $t \geq s \geq 0$.
- A family of operators $(\theta_t)_{t \geq 0}$ on Ω is called shift operators for X iff $X_t(\theta_s \omega) = X_{t+s}(\omega)$.

We now introduce the notion of diffusion.

Definition 1.1.2 (see [59], p. 376). A diffusion on $\mathbb{I} \subset \mathbb{R}$ is a strong Markov process with state-space \mathbb{I} and continuous sample paths.

The dynamic of a diffusion is time-homogeneous and depends only on the position in space. They be seen as the motion of charged particle in space subjected to the action of a potential. Sticky diffusions are diffusion processes that spend a positive amount of time at certain point(s) of their state space (see Sections 2.1.3,2.2).

1.1.2 Law and semi-group of a diffusion

For any \mathcal{X} locally compact space with countable basis, let

- $C(\mathcal{X})$ be the space of continuous real-valued functions defined on \mathbb{I} ,
- $C_b(\mathcal{X})$ the subspace of bounded functions of $C(\mathcal{X})$,
- $C_0(\mathcal{X})$ the subspace of functions of $C(\mathcal{X})$ that vanish at infinity, i.e. $|f(x)| \rightarrow 0$ as $\|x\| \rightarrow \infty$,
- $\|\cdot\|_\infty$ the norm defined for every measurable function f by

$$\|f\|_\infty = \sup_{x \in \mathcal{X}} |f(x)|.$$

We equip the spaces $C(\mathcal{X}), C_b(\mathcal{X}), C_0(\mathcal{X})$ with the norm $\|\cdot\|_\infty$ and when there is no ambiguity $\|\cdot\| = \|\cdot\|_\infty$. These functional spaces can be used to quantify the notion of law of a process with state-space $\mathbb{I} = \mathcal{X}$.

The law of a diffusion is the full random behavior of the process. One way to quantify this behavior is via test functions. We say that two random variables Z_1 and Z_2 have the same law iff for all $f \in C_b$,

$$\mathbb{E}(f(Z_1)) = \mathbb{E}(f(Z_2)).$$

Similarly, we say that two processes X and Y with state-space $\mathbb{I} \subset \mathbb{R}^d$ have the same law iff for all $x \in \mathbb{I}$ and measurable $F : C_b(\mathbb{R}_+, \mathbb{I}) \rightarrow \mathbb{R}$,

$$\mathbb{E}_x(F(X_t; t \geq 0)) = \mathbb{E}_x(F(Y_t; t \geq 0)).$$

So an equality in law between processes (resp. random variables) means that their action on measurable path-functionals (resp. bounded continuous functions) is the same. In the case of continuous processes (like diffusions) or càdlàg ones, it is possible to factorize this action in time. Let $(P_t)_{t \geq 0}$ be the family of operators defined for every measurable bounded f , $x \in \mathbb{I}$ and $t \geq 0$ by

$$P_t f(x) = \mathbb{E}_x \left(f(X_t) \right). \quad (1.2)$$

It is possible to ditch the notion of shift operator and re-express the Markov property in terms of $(P_t)_{t \geq 0}$ as follows.

Definition 1.1.3. Let X and $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)_{x \in \mathbb{I}}$ be as in Definition 1.1.1 and $(P_t)_{t \geq 0}$ be the family of operators defined in (1.2). Then, X satisfies the strong Markov property iff

$$\mathbb{E} \left(f(X_{\tau+s}) \mid \mathcal{F}_\tau \right) \mathbb{1}_{\tau < \infty} = P_s f(X_\tau) \mathbb{1}_{\tau < \infty} \quad (1.3)$$

for all $f \in C_b(\mathbb{I})$, stopping times τ and $s \geq 0$.

If X is a Markov process and $(P_t)_{t \geq 0}$ the family of operators defined in (1.2), from (1.3),

$$\begin{aligned} P_t P_s f(x) &= \mathbb{E}_x \left(P_s f(X_t) \right) = \mathbb{E}_x \left(\mathbb{E}_{X_t} \left(f(X_{s+t}) \right) \right) \\ &= \mathbb{E}_x \left(\mathbb{E} \left(f(X_{s+t}) \mid \mathcal{F}_t \right) \right) = P_{s+t} f(x). \end{aligned}$$

Thus, the Markov property induces a semi-group structure on $(P_t)_{t \geq 0}$. As such, we call $(P_t)_{t \geq 0}$ the semi-group of X .

Proposition 1.1.4. *The law of a diffusion process is characterized by its semi-group.*

Proof. Let X be a diffusion process and $(P_t)_{t \geq 0}$ its semi-group. From Kolmogorov's extension theorem [87, p.196], it suffices to show that for any ordered finite set of times $0 < t_1 < t_2 < \dots < t_n < \infty$, the law of $(X_{t_1}, \dots, X_{t_n})$ is uniquely determined by $(P_t)_{t \geq 0}$. Equivalently, it suffices to show that for any $n \in \mathbb{N}$, $0 < t_1 < t_2 < \dots < t_n < \infty$ and bounded measurable application $F : \mathbb{I}^n \rightarrow \mathbb{R}$, $\mathbb{E} \left(F(X_{t_1}, \dots, X_{t_n}) \right)$ is uniquely determined by x and $(P_t)_{t \geq 0}$. From the functional version of the monotone class theorem [80, p.3], by considering the π -system of indicator functions on \mathbb{I} , it suffices to show it for all measurable functions $F : \mathbb{I}^n \rightarrow \mathbb{R}$ of the form

$$F(x_1, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n),$$

where $f_1, \dots, f_n : \mathbb{I} \rightarrow \mathbb{R}$ are bounded measurable functions. From (1.3),

$$\begin{aligned} &\mathbb{E}_x \left(f_1(X_{t_1}) \dots f_n(X_{t_n}) \right) \\ &= \mathbb{E}_x \left(f_1(X_{t_1}) \dots f_{n-1}(X_{t_{n-1}}) \mathbb{E} \left(f_n(X_{t_n}) \mid \mathcal{F}_{t_{n-1}} \right) \right) \\ &= \mathbb{E}_x \left(f_1(X_{t_1}) \dots f_{n-1}(X_{t_{n-1}}) P_{t_n - t_{n-1}} f_n(X_{t_{n-1}}) \right) \end{aligned}$$

and recursively

$$\mathbb{E}_x \left(f_1(X_{t_1}) \dots f_n(X_{t_n}) \right) = P_{t_1} \left(f_1 P_{t_2 - t_1} \left(f_2 \dots P_{t_n - t_{n-1}} f_n \right) \dots \right) (x),$$

which depends only on $(P_t)_{t \geq 0}$ and $x \in \mathbb{I}$. Thus, the law of $(X_{t_1}, \dots, X_{t_n})$ under \mathbb{P}_x depends only on $(P_t)_{t \geq 0}$. \square

Using the same arguments as the proof on Proposition 1.2.14, one can prove the following stronger version of the strong Markov property:

Theorem 1.1.5. *Let X be a strong Markov process defined on a family of probability spaces $(\Omega, (\mathcal{F}_t)_{t \geq 0}, P_x)_{x \in \mathbb{I}}$ with state space \mathbb{I} such that for every $x \in \mathbb{I}$, $P_x(X_0 = x) = 1$. Then for any stopping time τ and measurable functional $F : \mathcal{M}(\mathbb{R}_+, \mathbb{I}) \rightarrow \mathbb{R}$, where $\mathcal{M}(\mathbb{R}_+, \mathbb{I})$ is the space of measurable functions from \mathbb{R}_+ to \mathbb{I} ,*

- $F((X_{t+\tau})_{t \geq 0}) \mathbb{1}_{\tau < \infty}$ is \mathcal{F}_∞ -measurable.
- For any $x \in \mathbb{I}$, $E_x \left(\mathbb{1}_{\tau < \infty} F((X_{t+\tau})_{t \geq 0}) \middle| \mathcal{F}_\tau \right) = E_{X_\tau} \left(F((X_t)_{t \geq 0}) \right) \mathbb{1}_{\tau < \infty}$.

1.2 Analytic characterization

In the previous section we characterized the law of a diffusion process with its semi-group. It turns out that this is not a very convenient characterization as it cannot be written down explicitly. Moreover, it is hard to infer the path-wise properties of the diffusion by looking at its semi-group.

In this section we give two equivalent characterization of the law of a diffusion: the resolvent family and the infinitesimal generator. The link between them is the Hille-Yosida theorem.

1.2.1 Hille-Yosida theory

The Hille-Yosida theory is a series of results that establish links between classes of linear operators over Banach spaces. It was initially developed to assess the existence and uniqueness of a solution of

$$\frac{d}{dt} u(t, x) = A_x u(t, x)$$

by looking at the properties of a dissipative operator A_x . It was then applied with great success to the study of diffusion processes by Feller [33].

Let $(\mathbb{H}, \|\cdot\|)$ be a Banach space and $\mathcal{L}(\mathbb{H})$ the space of bounded linear operators from \mathbb{H} to \mathbb{H} . We define the following families of operators over \mathbb{H} :

Definition 1.2.1. A family $(P_t)_{t \geq 0} \subset \mathcal{L}(\mathbb{H})$ is called a strongly continuous contracting semi-group over \mathbb{H} iff:

1. for any $s, t \geq 0$ and $v \in \mathbb{H}$: $P_t P_s v = P_{t+s} v$,
2. for every $t > 0$ and $v \in \mathbb{H}$: $\|P_t v\| \leq \|v\|$,
3. for every $v \in \mathbb{H}$: $\|P_t v - v\| \rightarrow 0$ as $t \rightarrow 0$.

A strongly continuous contracting semi-group over $(C_0(\mathcal{X}, \mathbb{R}), \|\cdot\|_\infty)$, with \mathcal{X} a locally compact separable space, is called a Feller semi-group. Usually $\mathcal{X} \subset \mathbb{R}^d$, but this is not always the case (see [45, 65]).

Definition 1.2.2. A family $(R_\lambda)_{\lambda>0} \subset \mathcal{L}(\mathbb{H})$ is called a strongly continuous contracting resolvent family iff

1. for every $\lambda, \mu > 0$: $R_\lambda - R_\mu = (\mu - \lambda)R_\lambda R_\mu$,
2. for every $\lambda > 0$ and $v \in \mathbb{H}$: $\|\lambda R_\lambda v\| \leq \|v\|$,
3. for every $v \in \mathbb{H}$: $\|\lambda R_\lambda v - v\| \rightarrow 0$ as $\lambda \rightarrow \infty$.

Definition 1.2.3. An operator $L : \text{dom}(L) \rightarrow \mathbb{H}$, with $\text{dom}(L) \subset \mathbb{H}$ is called:

1. dissipative iff for every $\lambda > 0$ and $v \in \text{dom}(L)$: $\|\lambda v - L v\| \geq \|\lambda v\|$,
2. closed iff for every $(v_n)_{n \in \mathbb{N}} \subset \text{dom}(L)$ and $v, w \in \mathbb{H}$ such that $v_n \rightarrow v$ and $\overline{L v_n} \rightarrow w$, then $v \in \text{dom}(L)$ and $L v = w$.

To each strongly continuous contracting semi-group $(P_t)_{t \geq 0}$ one can associate a resolvent family and an operator on \mathbb{H} as follows.

Proposition 1.2.4. Let $(P_t)_{t \geq 0}$ be a strongly continuous contracting semi-group over a Banach space \mathbb{H} . The family of operators $(R_\lambda)_{\lambda > 0}$ defined for every $v \in \mathbb{H}$ and $\lambda > 0$ by

$$R_\lambda v = \int_0^\infty e^{-\lambda t} P_t v dt$$

is a strongly continuous contracting resolvent family over \mathbb{H} . We call it the resolvent of $(P_t)_{t \geq 0}$.

Definition 1.2.5. Let $(P_t)_{t \geq 0}$ be a strongly continuous contracting semi-group over a Banach space \mathbb{H} . The infinitesimal generator of $(P_t)_{t \geq 0}$ is the operator L defined for every $v \in \text{dom}(L)$ by

$$L v = \lim_{t \rightarrow 0} \frac{1}{t} (P_t v - v), \quad \text{where} \quad \text{dom}(L) := \{v \in \mathbb{H} : L v \in \mathbb{H}\}. \quad (1.12)$$

Theorem 1.2.6 (Hille-Yosida, see [32], Chapter XIII, §9 and §10). Let \mathbb{H} be a Banach space. Then,

- For each strongly continuous contracting semi-group $(P_t)_{t \geq 0}$, there exist a unique strongly continuous contracting resolvent $(R_\lambda)_{\lambda > 0}$ and a closed dissipative operator L with dense domain such that: $(R_\lambda)_{\lambda > 0}$ and L are respectively the resolvent and infinitesimal generator of $(P_t)_{t \geq 0}$.
- Every strongly continuous contracting resolvent $(R_\lambda)_{\lambda > 0}$ over \mathbb{H} is the resolvent of a strongly continuous contracting semi-group $(P_t)_{t \geq 0}$ over \mathbb{H} .
- Every closed dissipative operator L over \mathbb{H} with dense domain is the infinitesimal generator of a strongly continuous contracting semi-group $(P_t)_{t \geq 0}$ over \mathbb{H} .

Moreover, for every $\lambda > 0$,

$$R_\lambda^{-1} = (\lambda \text{Id} - L) \quad \text{and} \quad \text{dom}(L) = R_\lambda(\mathbb{H}) \subset \overline{R_\lambda(\mathbb{H})} = \mathbb{H}.$$

The following result also holds.

Proposition 1.2.7 (Chapman-Kolmogorov equations). *Let $(P_t)_{t \geq 0}$ be a strongly continuous contracting semi-group over a Banach space \mathbb{H} and L be its infinitesimal generator. Then, for any $v \in \text{dom}(L)$ and $t \geq 0$,*

- *the application $[s \rightarrow P_s v]$ is strongly differentiable in \mathbb{H} and $\frac{d}{dt} P_t v = L P_t v = P_t L v$,*
- *$P_t v - v = \int_0^t L P_s v \, ds = \int_0^t P_s L v \, ds$,*
- *$P_t(\text{dom}(L)) \subset \text{dom}(L)$ and $P_t L = L P_t$ on $\text{dom}(L)$.*

Proof. The proof is the same as the one of Proposition 1.2 of [80, Chapter VII; §1]. \square

1.2.2 Adaptation to regular diffusions

We now prove that the results of the previous sections can be adapted to regular diffusions with state-space \mathbb{I} an interval of \mathbb{R} .

Definition 1.2.8 (Definition 45.2 of [82]). A diffusion process X with state-space \mathbb{I} an interval of \mathbb{R} is called regular iff for all $x \in \text{Int}(\mathbb{I})$ and $y \in \mathbb{I}$,

$$P_x(\tau_y < \infty) > 0,$$

where $\tau_y = \inf\{t \geq 0 : X_t = y\}$.

To apply the Hille-Yosida theorem to diffusions, we need the semi-group to be strongly continuous and contracting over a proper ambient space. The semi-group $(P_t)_{t \geq 0}$ of a regular diffusion with state-space \mathbb{I} , an interval of \mathbb{R} , is not always internal on $C_0(\mathbb{I})$, (see [82, §50]), and not always strongly continuous on $C_b(\mathbb{I})$ (see Proposition A.1.1). To guarantee these properties, we first consider regular Feller diffusions. Feller diffusions are diffusions whose semi-group is a Feller semi-group (see Definition 1.2.1).

Theorem 1.2.9 (Theorem 20.13 of [59]). *Let X be a regular diffusion, with state-space \mathbb{I} an interval of \mathbb{R} . Let $\bar{\mathbb{I}}$ be the extension of \mathbb{I} by any potential entrance boundaries of \mathbb{I} (see Section 2.1.2). Then, X can be extended to a continuous Feller process on $\bar{\mathbb{I}}$.*

Corollary 1.2.10. *A regular diffusion with no entrance boundaries is a Feller diffusion.*

The aforementioned result allow us to extend the following results from regular Feller diffusions to regular diffusions. For the rest of this section, we will suppose X to be a regular diffusion with state-space \mathbb{I} an interval of \mathbb{R} and $(P_t)_{t \geq 0}$ the family of operators defined for every $f \in C_b(\mathbb{I})$ by (1.2).

Definition 1.2.11 (see [80], p. 89). Let Y be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} . We call resolvent of Y , the family of operators $(R_\lambda)_{\lambda > 0}$ defined for any $\lambda > 0$ and $f \in C_b(\mathbb{I})$ by

$$R_\lambda f(x) = \int_0^\infty E_x(f(Y_s)) e^{-\lambda s} \, ds.$$

Definition 1.2.12. Let Y be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} . We call infinitesimal generator of Y , the operator L defined for any $f \in \text{dom}(L)$ and $x \in \mathbb{I}$ by

$$L f(x) = \lim_{t \rightarrow 0} \frac{1}{t} \left(\mathbb{E}_x(f(Y_t)) - f(x) \right), \quad (1.16)$$

where

$$\text{dom}(L) = \left\{ f \in C_b(\mathbb{I}) : L f \in C_b(\mathbb{I}) \right\}. \quad (1.17)$$

Now let X be a regular Feller diffusion. The following are all corollaries of the Hille-Yosida theorem:

Corollary 1.2.13 (of Theorem 1.2.6). *Let $(R_\lambda)_{\lambda>0}, L$ be respectively the resolvent and infinitesimal generator of X . Then,*

- $(R_\lambda)_{\lambda>0}$ is a strongly continuous contracting resolvent family over $C_0(\mathbb{I})$,
- L is a closed dissipative operator on $R_\lambda(C_0(\mathbb{I}))$,
- $R_\lambda(C_0(\mathbb{I}))$ is dense in $C_0(\mathbb{I})$.

Corollary 1.2.14 (of Proposition 1.1.4 and Theorem 1.2.6). *Let $(P_t)_{t \geq 0}, (R_\lambda)_{\lambda>0}$ and L be respectively the semi-group, resolvent and infinitesimal generator of X . The law of X is characterized either by $(P_t)_{t \geq 0}, (R_\lambda)_{\lambda>0}$ or L .*

Corollary 1.2.15 (of Theorem 1.2.6). *Let X and Y be two regular diffusions with state-space \mathbb{I} an interval of \mathbb{R} and respective (semi-group, resolvent, infinitesimal generator) triplets $((P_t^X)_{t \geq 0}, (R_\lambda^X)_\lambda, L^X), ((P_t^Y)_{t \geq 0}, (R_\lambda^Y)_\lambda, L^Y)$. If either of the following holds:*

- $(P_t^X)_{t \geq 0} = (P_t^Y)_{t \geq 0}$,
- $(R_\lambda^X)_\lambda = (R_\lambda^Y)_\lambda$,
- $L^X = L^Y$,

then, $((P_t^X)_{t \geq 0}, (R_\lambda^X)_\lambda, L^X) = ((P_t^Y)_{t \geq 0}, (R_\lambda^Y)_\lambda, L^Y)$ and X, Y have the same law.

1.2.3 Speed measure & scale function

We now give an additional characterization of diffusions, the scale function & speed measure pair. This characterization yields a very convenient factorization of the infinitesimal generator of a diffusion. Moreover, it allows to also factorize the space and time behavior of a diffusion:

- the scale function mainly captures the propensity of the process to move in a particular direction,
- the speed measure expresses the speed at which the process moves.

In Section 1.4, we see that several results are expressed through these objects.

Let X be a regular diffusion process, taking values in an interval $\mathbb{I} \subset \mathbb{R}$ and defined on a family of probability spaces $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)_{x \in \mathbb{I}}$ such that for every $x \in \mathbb{I}$, $\mathbb{P}_x(X_0 = x) = 1$. Let also for any $a, b \in \mathbb{I}$,

$$\tau_a = \inf \{t \geq 0 : X_t = a\}$$

and $\tau_{ab} = \tau_a \wedge \tau_b$.

Proposition 1.2.16. *There exists a continuous, increasing function s such that for any $a < x < b \in \mathbb{I}$:*

$$\mathbb{P}_x(\tau_b < \tau_a) = \frac{s(x) - s(a)}{s(b) - s(a)}, \quad (1.18)$$

where $\tau_a = \inf\{t > 0 : X_t = a\}$. The function s is unique modulo an affine transformation and is called scale function of X . A diffusion whose scale function s is the identity function $s = [x \rightarrow x]$ is said to be on natural scale.

Proof. See [80, p. 301-302]. □

Proposition 1.2.17. *There exists a unique strictly positive¹ locally finite² measure m over $\text{int}(\mathbb{I})$ such that for every $x, a, b \in \mathbb{I}$ with $a < x < b$:*

$$\mathbb{E}_x(\tau_{ab}) = \int_{(a,b)} G_{(a,b)}(x, y) m(dy), \quad (1.19)$$

where $\tau_{ab} = \min\{\tau_a, \tau_b\}$ and $[(a, b, x, y) \rightarrow G_{(a,b)}(x, y)]$ is the function defined for any $a, b, x, y \in \mathbb{I}$ with $a < b$ by

$$G_{a,b}(x, y) = \begin{cases} \frac{(s(x) - s(a))(s(b) - s(y))}{s(b) - s(a)}, & \text{for } a < x \leq y < b, \\ \frac{(s(y) - s(a))(s(b) - s(x))}{s(b) - s(a)}, & \text{for } a < y < x < b. \end{cases} \quad (1.20)$$

The measure m is called speed measure of X .

Proof. See [80, p. 304-305]. □

Proposition 1.2.18 (Green formula). *For every measurable function $f : \mathbb{I} \rightarrow \mathbb{R}$ and $x, a, b \in \mathbb{I}$ with $a < x < b$,*

$$\mathbb{E}_x \left(\int_0^{\tau_{ab}} f(X_s) ds \right) = \int_a^b G_{(a,b)}(x, y) f(y) m(dy). \quad (1.21)$$

Proof. See [80, p.305]. □

¹A measure m over \mathbb{I} an interval of \mathbb{R} is said to be strictly position iff for any $(a, b) \subset \mathbb{I}$, $m((a, b)) > 0$.

²A measure m on a topological space \mathcal{X} is said to be locally finite iff for any compact subset K of \mathcal{X} , $m(K) < \infty$.

Proposition 1.2.19 (martingale problem). *Let L be the infinitesimal generator of X . Then, for any $f \in \text{dom}(L)$, the process $M(f)$ defined for every $t \geq 0$ by*

$$M_t(f) = f(X_t) - f(X_0) - \int_0^t L f(X_u) du, \quad (1.22)$$

is a martingale.

Proof. For any $s \leq t$, from the Markov property and (1.13),

$$\begin{aligned} \mathbb{E} \left(M_t(f) \middle| \mathcal{F}_s \right) &= M_s(f) + \mathbb{E} \left(f(X_t) - f(X_s) + \int_s^t L f(X_u) du \middle| \mathcal{F}_s \right) \\ &= M_s(f) + P_{t-s} f(X_s) - f(X_s) + \int_s^t L P_{u-s} f(X_s) du = M_s(f). \end{aligned}$$

Thus, $M(f)$ is a martingale. \square

Proposition 1.2.20. *Let L , s and m be respectively the infinitesimal generator, the scale function and the speed measure of X . Then,*

- for every $f \in \text{dom}(L)$ and $x \in \text{Int}(\mathbb{I})$,

$$L f(x) = D_m D_s f(x), \quad (1.23)$$

- $\text{dom}(L) \subset \left\{ f \in C_b(\mathbb{I}) : D_m D_s f \in C_b(\text{Int}(\mathbb{I})) \right\}$,
- if \mathbb{I} is an open interval of \mathbb{R} , $\text{dom}(L) = \left\{ f \in C_b(\mathbb{I}) : D_m D_s f \in C_b(\text{Int}(\mathbb{I})) \right\}$,

where

$$D_s g(x) = \lim_{h \rightarrow 0; h > 0} \frac{g(x+h) - g(x)}{s(x+h) - s(x)} \quad \text{and} \quad D_m g(x) = \lim_{h \rightarrow 0; h > 0} \frac{g(x+h) - g(x)}{m([x, x+h])}. \quad (1.26)$$

Proof. For $f \in \text{dom}(L)$, let $M(f)$ be the process defined for every $t \geq 0$ by

$$M_t(f) = f(X_t) - f(X_0) - \int_0^t L f(X_s) ds, \quad (1.27)$$

where $\|M_t(f)\| \leq 2\|f\| + t\|L f\|$. Since X is regular for any $x, a, b \in \text{Int}(\mathbb{I})$ such that $a < x < b$, $\mathbb{E}_x(\tau_{ab}) < \infty$ and

$$\mathbb{E}_x \left(M_{\tau_{ab}}(f) \right) = f(b) \frac{s(x) - s(a)}{s(b) - s(a)} + f(a) \frac{s(b) - s(x)}{s(b) - s(a)} - f(x) - \mathbb{E}_x \left(\int_0^{\tau_{ab}} L f(X_s) ds \right)$$

From the martingale stopping theorem, $(M_{t \wedge \tau_{ab}}(f))_{t \geq 0}$ is a martingale that vanishes at 0. Thus, from (1.21),

$$\begin{aligned} &f(b)(s(x) - s(a)) + f(a)(s(b) - s(x)) - (s(b) - s(a))f(x) \\ &= (s(b) - s(a)) \int_a^b L f(y) G_{a,b}(x, y) m(dy) \\ &= \int_a^b \left((s(x) - s(a))(s(b) - s(y)) \mathbb{1}_{x \leq y} + (s(y) - s(a))(s(b) - s(x)) \mathbb{1}_{x > y} \right) L f(y) m(dy). \end{aligned}$$

Derivating with respect to s ,

$$\begin{aligned} \frac{f(b) - f(a)}{s(b) - s(a)} - D_s f(x) &= \int_a^b \frac{(s(b) - s(y)) \mathbb{1}_{x \leq y} - (s(y) - s(a)) \mathbb{1}_{x > y}}{s(b) - s(a)} L f(y) m(dy) \\ &= \int_a^x \frac{s(y) - s(a)}{s(b) - s(a)} L f(y) m(dy) + \int_x^b \frac{s(b) - s(y)}{s(b) - s(a)} L f(y) m(dy) \end{aligned}$$

Derivating with respect to m ,

$$D_m D_s f(x) = \frac{s(y) - s(a)}{s(b) - s(a)} L f(x) + \frac{s(b) - s(y)}{s(b) - s(a)} L f(x) = L f(x),$$

proving (1.23). This also proves that

$$\text{dom}(L) \subset \text{dom}(D_m D_s) = \left\{ g \in C_b(\text{Int}(\mathbb{I})) : D_m D_s g \in C_b(\text{Int}(\mathbb{I})) \right\}.$$

If \mathbb{I} is an open interval of \mathbb{R} ,

$$\text{dom}(L) \subset \text{dom}(D_m D_s) = \left\{ g \in C_b(\text{Int}(\mathbb{I})) : D_m D_s g \in C_b(\mathbb{I}) \right\}.$$

Since X has only natural boundaries and L is a closed operator (see Theorem 1.2.6),

$$\text{dom}(D_m D_s) \subset \text{dom}(L),$$

which finishes the proof. \square

From Proposition 1.2.20, we observe that if we are given the speed measure, the scale function and the boundary conditions of the infinitesimal generator of a one dimensional diffusion process X with state-space an interval of \mathbb{R} , we have a full description of that law of X . This characterization turns out to be very convenient for singular diffusions. When we introduce new diffusions, we do it through s and m (see Section 2.2, Chapter 3 and Section 6.1).

The fact that we also need to have the conditions at the boundary of the infinitesimal generator can be seen in [14, p. 118-122]. There, Brownian motions with boundary behaviors are defined and they share the same scale function and speed measure.

Corollary 1.2.21. *The pair scale function and speed measure (s, m) characterizes the law of a diffusion process with state-space \mathbb{I} an open interval of \mathbb{R} .*

Proof. Direct from Corollary 1.2.14, (1.23) and (1.24). \square

Proposition 1.2.22. *Let s and m be the scale function and speed measure of X . Then, the process $s(X) = (s(X_t))_{t \geq 0}$ is on natural scale, has speed measure $m'(dy) = m(ds^{-1}(y))$ and state-space the interval $s(\mathbb{I})$ of \mathbb{R} .*

Proof. Let (s_0, m_0) be the scale function and speed measure of $s(X)$. Moreover, for every $\zeta \in \mathbb{I}$ and $\zeta' \in s(\mathbb{I})$,

$$\begin{aligned} \tau_\zeta &= \inf\{t > 0; X_t = \zeta\}, & \tau_{ab} &= \tau_a \wedge \tau_b, \\ \tau'_{\zeta'} &= \inf\{t > 0; s(X_t) = \zeta'\}, & \tau'_{ab} &= \tau'_a \wedge \tau'_b. \end{aligned} \tag{1.28}$$

Thus, for any $\zeta \in \mathbb{I}$, $\tau'_{s(\zeta)} = \tau_\zeta$. From (1.18), for any $a, x, b \in \mathbb{I}$ such that $a < x < b$:

$$\mathbb{P}_x \left(\tau'_b < \tau'_a \right) = \mathbb{P}_{s^{-1}(x)} \left(\tau_{s^{-1}(b)} < \tau_{s^{-1}(a)} \right) = \frac{x - a}{b - a},$$

thus $s_0(x)$ is necessarily an affine function. From (1.19) and (1.28), if $a, x, b \in \mathbb{I}$ such that $a < x < b$:

$$\begin{aligned} \mathbb{E}_{s(x)}(\tau'_{s(a)s(b)}) &= \mathbb{E}_x(\tau_{ab}) = \int_a^b \frac{(s(x) \wedge s(y) - s(a))(s(b) - s(x) \vee s(y))}{s(b) - s(a)} m(dy) \\ &= \int_{s(a)}^{s(b)} \frac{(s(x) \wedge \zeta - s(a))(s(b) - s(x) \vee \zeta)}{s(b) - s(a)} m(ds^{-1}(\zeta)). \end{aligned}$$

From (1.19), for every $x \in s(\mathbb{I})$, $s_0(x) = x$ and $m_0(dx) = m(ds^{-1}(x))$. \square

Proposition 1.2.23 (Dynkin operator). *Let X be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} and infinitesimal generator L . For any $x \in \mathbb{I}$ and $h > 0$, let $\tau_x = \inf\{t > 0 : X_t = x\}$ and $\tau_x(h) = \tau_{x-h} \wedge \tau_{x+h}$. Then, if $f \in \text{dom}(L)$, for any $x \in \mathbb{I}$,*

$$L f(x) = \lim_{h \rightarrow 0} \frac{\mathbb{E}_x \left(f(X_{\tau_x(h)}) \right) - f(x)}{\mathbb{E}_x(\tau_x(h))}. \quad (1.29)$$

Proof. Let h small enough such that $(x - h, x + h) \subset \mathbb{I}$ and $M(g)$ the martingale defined for any $g \in \text{dom}(L)$ in (1.27). As $f \in \text{dom}(L)$, the process $(M_{t \wedge \tau_x(h)}(f))_{t \geq 0}$ is a martingale. Also, since X is regular, from (1.21),

$$\mathbb{E}_x \left(f(X_{\tau_x(h)}) \right) - f(x) = \mathbb{E}_x \left(\int_0^{\tau_x(h)} L f(X_s) \right) = \int_{x-h}^{x+h} G_{x-h, x+h}(x, y) L f(y) m(dy).$$

From (1.19), and the intermediate value theorem,

$$\frac{\mathbb{E}_x \left(f(X_{\tau_x(h)}) \right) - f(x)}{\mathbb{E}_x(\tau_x(h))} = \frac{\int_{x-h}^{x+h} G_{x-h, x+h}(x, y) L f(y) m(dy)}{\int_{x-h}^{x+h} G_{x-h, x+h}(x, y) m(dy)}$$

As $f \in \text{dom}(L)$, $L f \in C_b(\mathbb{I})$, if $\delta_h(x) = \sup \{ |L f(y) - L f(x)|; |y - x| < h \}$, $\delta_h(x) \rightarrow 0$ as $h \rightarrow 0$. Thus,

$$\begin{aligned} \left| \frac{\mathbb{E}_x \left(f(X_{\tau_x(h)}) \right) - f(x)}{\mathbb{E}_x(\tau_x(h))} - L f(x) \right| &= \frac{\int_{x-h}^{x+h} G_{x-h, x+h}(x, y) |L f(y) - L f(x)| m(dy)}{\int_{x-h}^{x+h} G_{x-h, x+h}(x, y) m(dy)} \\ &\leq \frac{\int_{x-h}^{x+h} G_{x-h, x+h}(x, y) \delta_x(h) m(dy)}{\int_{x-h}^{x+h} G_{x-h, x+h}(x, y) m(dy)} = \delta_x(h) \rightarrow 0 \end{aligned}$$

as $h \rightarrow 0$, which is (1.29). \square

1.3 Probability transition kernel

In the previous sections we defined diffusion processes by means of functional analysis, via actions they induce on the space of bounded continuous functions. While this

point of view is convenient from an analytical perspective, the notion of probability transition kernel is much more relevant from a statistical perspective. Indeed, it is the transition kernel that expresses the statistical properties of the process via its marginal laws. In this section, we prove the existence of this object and give the first elements of its computation in a fully generalized setting. We also see the role played by the resolvent kernel in proving these results.

1.3.1 Definitions

Proposition 1.3.1. *Let X be a diffusion process with speed measure m , semi-group $(P_t)_{t \geq 0}$ and state-space \mathbb{I} , an interval of \mathbb{R} . There exists a function $[(t, x, y) \rightarrow p(t, x, y)] : \mathbb{R}_+ \times \mathbb{I}^2 \rightarrow \mathbb{R}_+$ such that for any $f \in L^2$, $t \geq 0$ and $x \in \mathbb{I}$,*

$$P_t f(x) = \int_{\mathbb{I}} f(y) p(t, x, y) m(dy). \quad (1.30)$$

Moreover, p satisfies the following properties:

- continuity in all arguments (t, x, y) ,
- symmetry in space, i.e. for all $x, y \in \mathbb{I}$ and $t > 0$, $p(t, x, y) = p(t, y, x)$,
- convolution, i.e. $p(t + s, x, y) = \int_{\mathbb{I}} p(t, x, \zeta) p(s, \zeta, y) m(d\zeta)$.

The relation (1.33) is also called Chapman-Kolmogorov equation.

Proof. See [53, p.149]. □

Definition 1.3.2. The function $[(t, x, y) \rightarrow p(t, x, y)]$ is called probability transition kernel of X .

Definition 1.3.3. Let X be a diffusion process with probability transition kernel $[(t, x, y) \rightarrow p(t, x, y)]$. The function $[(\lambda, x, y) \rightarrow r(\lambda, x, y)]$ defined for every $(\lambda, x, y) \in (0, \infty) \times \mathbb{I}^2$ by

$$r(\lambda, x, y) = \int_0^\infty p(t, x, y) e^{-\lambda t} dt \quad (1.34)$$

is called resolvent kernel of X . From Proposition 1.3.1 and (1.34), the resolvent kernel is continuous.

Let $(R_\lambda)_{\lambda > 0}$ be the resolvent family, $(P_t)_{t \geq 0}$ the semi-group, $[(t, x, y) \rightarrow p(t, x, y)]$ the probability transition kernel and $[(\lambda, x, y) \rightarrow r(\lambda, x, y)]$ the resolvent kernel of a diffusion X . From the positive Fubini theorem,

$$\begin{aligned} R_\lambda f(x) &= \int_0^\infty P_t f(x) e^{-\lambda t} dt = \int_0^\infty \int_{\mathbb{I}} f(y) p(t, x, y) m(dy) e^{-\lambda t} dt \\ &= \int_{\mathbb{I}} f(y) \int_0^\infty p(t, x, y) e^{-\lambda t} dt m(dy) = \int_{\mathbb{I}} f(y) r(\lambda, x, y) m(dy), \end{aligned}$$

for any measurable $f : \mathbb{I} \rightarrow \mathbb{R}$, $\lambda > 0$ and $x \in \mathbb{I}$. This justifies the usage of the term resolvent kernel for the function $[(\lambda, x, y) \rightarrow r(\lambda, x, y)]$.

1.3.2 Backward & forward equations

Let X be a diffusion process with state-space \mathbb{I} , an interval of \mathbb{R} and infinitesimal generator L . Also, let L^* be the adjoint operator of L in $L^2(\mathbb{I})$ in the sense that

$$\langle Lu, v \rangle_{L^2} = \langle u, L^*v \rangle_{L^2},$$

for any $u \in \text{dom}(L)$ and $v \in \text{dom}(L^*)$. This makes sense since $\text{dom}(L) \subset C_b(\mathbb{I}) \subset L^2(\mathbb{I})$. The partial differential equations

$$(\partial_t - L)u(t, x) = 0$$

and

$$(\partial_t - L^*)u(t, x) = 0$$

are respectively called the *backward* and *forward equations*. The reason why is made explicit by the following theorem, which is also a common point of entry for obtaining analytic expressions (when possible) of probability transition kernels (see [14, 24, 25, 67]).

Proposition 1.3.4. *Let X be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} with infinitesimal generator L . Then, the probability transition kernel $[t, x, y \rightarrow p(t, x, y)]$ of X solves*

$$(\partial_t - L_x)p(t, x, y) = 0, \tag{1.35}$$

$$(\partial_t - L_y^*)p(t, x, y) = 0, \tag{1.36}$$

$$p(t, x, y) \rightarrow \delta_x(dy) \quad \text{weakly as } t \rightarrow 0, \tag{1.37}$$

$$p(t, x, y) \rightarrow \delta_y(dx) \quad \text{weakly as } t \rightarrow 0, \tag{1.38}$$

for all $(t, x, y) \in \mathbb{R}_+ \times \mathbb{I}^2$.

Lemma 1.3.5. *The resolvent kernel $[(\lambda, x, y) \rightarrow r(\lambda, x, y)]$ of a diffusion X with state-space \mathbb{I} solves*

$$(\lambda - L_x)r(\lambda, x, y) = 0, \tag{1.39}$$

$$(\lambda - L_y^*)r(\lambda, x, y) = 0, \tag{1.40}$$

for any $(\lambda, x, y) \in (0, \infty) \times \mathbb{I}^2$, where L^* is the adjoint operator of L in $L^2(\mathbb{I})$.

Proof. Let $f \in \text{dom}(L)$ and $u(t, x) = P_t f(x)$. From (1.13), for any $t \geq 0$, $[x \rightarrow u(t, x)] \in \text{dom}(L)$ and

$$\partial_t u(t, x) = L_x u(t, x),$$

where the subscript in L_x is used to denote the scope of the operator L . Thus, by definition of the probability transition kernel p of X ,

$$\int_{\mathbb{I}} (\partial_t - L_x)p(t, x, y)f(y)m(dy) = 0.$$

Since $\text{dom}(L)$ is dense in C_b which is itself dense in L^2 and since m is a strictly positive measure,

$$(\partial_t - L_x)p(t, x, y) = 0, \tag{1.41}$$

for all $(t, x) \in \mathbb{R}_+ \times \mathbb{I}$, $m(dy)$ -almost everywhere. From the continuity of $[(t, x, y) \rightarrow p(t, x, y)]$ (1.41) holds also for all $(t, x, y) \in \mathbb{R}_+ \times \mathbb{I}^2$. For any $\lambda > 0$, by multiplying by $e^{-\lambda t}$ and integrating with respect to t on \mathbb{R}_+ ,

$$(\lambda - L_x)r(\lambda, x, y) = 0,$$

which is (1.39). Multiplying (1.39) with $f(y)r(\lambda, y, x)$ and integrating over \mathbb{I} with respect to $m(dx)$,

$$\begin{aligned} 0 &= \int_{\mathbb{I}} r(\lambda, y, x)(\lambda - L_x)r(\lambda, x, y)m(dx) \\ &= \lambda \int_{\mathbb{I}} r(\lambda, y, x)r(\lambda, x, y)m(dx) - \int_{\mathbb{I}} r(\lambda, y, x)L_x r(\lambda, x, y)m(dx) \\ &= \lambda \int_{\mathbb{I}} r(\lambda, y, x)r(\lambda, x, y)m(dx) - \int_{\mathbb{I}} r(\lambda, x, y)L_x^* r(\lambda, y, x)m(dx) \\ &= \int_{\mathbb{I}} (\lambda r(\lambda, y, x) - L_x^* r(\lambda, y, x))r(\lambda, x, y)m(dx). \end{aligned} \quad (1.42)$$

Multiplying (1.42) with a measurable function $f(y)$ and integrating over \mathbb{I} with respect to $m(dy)$,

$$\int_{\mathbb{I}} \int_{\mathbb{I}} (\lambda r(\lambda, y, x) - L_x^* r(\lambda, y, x))f(y)r(\lambda, x, y)m(dy)m(dx) = 0.$$

Since X is regular, $m(dy)$ is strictly positive measures and $p(t, x, y) > 0$ for all $(t, x, y) \in (0, \infty) \times \text{Int}(\mathbb{I}) \times \text{Int}(\mathbb{I})$. Also, from (1.34), $r(\lambda, x, y) > 0$ for all $(t, x, y) \in (0, \infty) \times \text{Int}(\mathbb{I}) \times \text{Int}(\mathbb{I})$. Thus,

$$\lambda r(\lambda, y, x) - L_x^* r(\lambda, y, x) = 0$$

for all $t \in \mathbb{R}_+ \times \mathbb{I}$, $m(dy)m(dx)$ -almost everywhere and from the continuity of $[(\lambda, x, t) \rightarrow r(\lambda, x, y)]$ for all $(t, x, y) \in \mathbb{R}_+ \times \mathbb{I}^2$. \square

Proof of Proposition 1.3.4. Equation (1.35) was proven in the proof of Proposition 1.3.5. Equation (1.36) is proven applying the inverse Laplace transform to (1.40). The strong continuity of the diffusion's semi-group $(P_t)_{t \geq 0}$ implies the weak continuity at zero, which are (1.37) and (1.38). \square

1.4 Additional useful results

1.4.1 Diffusions and martingales

Diffusion processes and martingales/semi-martingales are two distinct classes of processes that are central to the theory of continuous processes. We observe that there is no inclusive relation between these classes:

- the Brownian motion is a martingale and a diffusion process (see Section 1.5),
- the process $X = (\sqrt{|B_t|})_{t \geq 0}$, with B the standard Brownian motion, is a diffusion but not a semi-martingale (see [92]),

- strong solutions of non-homogeneous stochastic differential equations (see Section 2.1.1) are semi-martingales but not diffusions,
- the fractional Brownian motion with Hurst coefficient $H \neq 1/2$ (see [80, p. 38]) is neither a semi-martingale, nor a diffusion (see [81]).

While this is the case there these processes are linked. The object of this section is to exhibit some of these links.

Lemma 1.4.1. *Let X be a diffusion process, L its infinitesimal generator and v_0 the function defined for every $x, a, b \in \mathbb{I}$ by $v_0(x) = P_x(\tau_b < \tau_a)$. Then $v_0 \in \text{dom}(L)$ and for any $a, b \in \mathbb{I}$, it solves:*

$$\begin{cases} Lu = 0, & x \in (a, b), \\ u(a) = 0, \\ u(b) = 1. \end{cases} \quad (1.43)$$

Proof. From (1.18), (1.23) and since m is a positive measure (see Proposition 1.2.17), for any $x \in \mathbb{I}$,

$$L v_0(x) = D_m D_s \frac{s(x) - s(a)}{s(b) - s(a)} = D_m \frac{1}{s(b) - s(a)} = 0$$

as m is a strictly positive measure. The functions v_0 and $L v_0 = 0$ are both in $C_b(\mathbb{I})$. Thus,

$$v_0 \in \text{dom}(L) \quad \text{and} \quad L v_0 = 0. \quad (1.44)$$

Moreover,

$$v_0(b) = P_b(\tau_b < \tau_a) = P_b(0 < \tau_a) = 1. \quad (1.45)$$

$$v_0(a) = P_a(\tau_b < \tau_a) = P_b(\tau_b < 0) = 0. \quad (1.46)$$

From (1.44), (1.45) and (1.46), v_0 solves (1.43). \square

Proposition 1.4.2. *A diffusion process on natural scale is a local martingale.*

Proof. Let X be a diffusion process with state-space \mathbb{I} , speed measure m and scale function s the identity function, i.e. $s(x) = x$ for any $x \in \mathbb{I}$. In (1.44) we observed that $s \in \text{dom}(L)$ but only for the process stopped at the boundary $X^{\tau_{ab}} = (X_{t \wedge \tau_{ab}})_{t \geq 0}$. In general, $s \notin C_b(\mathbb{I})$ and consequently $s \notin \text{dom}(L)$. This is the case for processes on natural scale with unbounded state-space, i.e. $\mathbb{I} = \mathbb{R}$. Let $\tau_{-n,n} = \inf \{t \geq 0; X_t \notin (-n, n)\}$, $X^{\tau_{-n,n}}$ be the stopped process $X^{\tau_{-n,n}} = (X_{t \wedge \tau_{-n,n}})_{t \geq 0}$ and $M(s)$ the process defined in (1.22). From Proposition 1.2.19, the process $M^{\tau_{-n,n}}(s) = (M_{t \wedge \tau_{-n,n}}(s))_{t \geq 0}$ is a martingale. From (5.19), $L s = 0$ and

$$M_t^{\tau_{-n,n}}(s) = s(X_t^{\tau_{-n,n}}) - s(X_0^{\tau_{-n,n}}) - \int_0^t L s(X_r^{\tau_{-n,n}}) dr = X_t^{\tau_{-n,n}} - X_0^{\tau_{-n,n}}$$

Thus, $X^{\tau_{-n,n}} = (M_t^{\tau_{-n,n}}(s) + X_0^{\tau_{-n,n}})_{t \geq 0}$ is also a martingale and the result is proven. Since $\tau_{-n,n}$ is increasing $\tau_{-n,n} \rightarrow \infty$ almost surely as $n \rightarrow \infty$, the result is proven. \square

From Propositions 1.2.22 and 1.4.2, we have the following:

Corollary 1.4.3. *If X is a diffusion process with scale function s . Then, the process $s(X) = (s(X_t))_{t \geq 0}$ is a local martingale.*

1.4.2 Diffusions as time changed Brownian motions

Theorem 1.4.4. *Let X be a diffusion on natural scale with state-space \mathbb{I} an interval of \mathbb{R} and speed measure m defined on a family of probability spaces $\mathcal{P} = (\Omega, \mathcal{F}, \mathbb{P}_x)_{x \in \mathbb{I}}$ such that for every $x \in \mathbb{I}$, $\mathbb{P}_x(X_0 = x) = 1$. Then, there exists a Brownian motion B , defined on an extension of \mathcal{P} , such that $\mathbb{P}_x(B_0 = x)$ and for every $t \geq 0$,*

$$X_t = B_{\gamma(t)},$$

where $\gamma(t)$ is the right inverse of $A(t) = \frac{1}{2} \int_{\mathbb{R}} L_t^y(B) m(dy)$.

Proof. See [82, p.278-279]. □

Corollary 1.4.5 (Remark (ii) of [82], p. 277). *Let m be a locally finite strictly positive measure on \mathbb{R} and B a standard Brownian motion defined on a family of probability space $(\mathcal{P}_x)_{x \in \mathbb{R}} = (\Omega, \mathcal{F}, \mathbb{P}_x)_{x \in \mathbb{R}}$ such that for every $x \in \mathbb{R}$, $\mathbb{P}_x(B_0 = x) = 1$. Let also $L^y(B)$ be the time-continuous version of the local time of B at y , A the time-change defined for every $t \geq 0$ by*

$$A(t) = \frac{1}{2} \int_{\mathbb{R}} L_t^y(B) m(dy)$$

and γ the right-inverse of A . Then, the process $X = (B_{\gamma(t)})_{t \geq 0}$ is a diffusion process on natural scale with speed measure m and for every $x \in \mathbb{R}$, $\mathbb{P}_x(X_0 = x) = 1$.

Corollary 1.4.6. *Let X° be a diffusion on natural scale with state-space \mathbb{I} and speed measure m° . Then, if κ is a finite measure on \mathbb{I} , the process X defined through s and m , where*

$$s(x) = x, \quad m(dx) = m^\circ(dx) + \kappa(dx),$$

is a martingale with state-space \mathbb{I} .

Proof. This is a direct consequence of Proposition 1.4.2, Theorem 1.4.4 and Corollary 1.4.5. □

1.5 The standard Brownian motion

The most elementary continuous process is the Brownian motion. It also lies at the intersection of the most studied classes of processes: Lévy processes, martingales, diffusions. The Brownian motion can be defined as the process $B = (B_t)_{t \geq 0}$ with

- $B_0 = 0$,
- $[t \rightarrow B_t]$ is almost surely continuous,
- for any $t \geq s$, $B_t - B_s$ is independent of \mathcal{F}_s ,
- for any $t \geq s$, $B_t - B_s \sim \mathcal{N}(0, t - s)$.

We observe that since a Gaussian random variable can take any value in \mathbb{R} , the state-space of B is \mathbb{R} .

For sake of convenience and consistency with the previous sections, we lift the property (1.47). Instead, we consider a Brownian motion defined on a family of probability space $(\Omega, (\mathcal{F})_{t \geq 0}, \mathbb{P}_x)_{x \in \mathbb{R}}$ any process such that (1.48)-(1.50) hold and that

$$\mathbb{P}_x(B_0 = x) = 1.$$

From (1.49) and (1.50), for any measurable function $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$\mathbb{E}(f(B_t) | \mathcal{F}_s) = \mathbb{E}(f(B_t - B_s + B_s) | \mathcal{F}_s) = \mathbb{E}(f(\sqrt{t-s}Z + B_s) | \mathcal{F}_s),$$

where $Z \sim \mathcal{N}(0, 1)$ is independent of \mathcal{F}_s . Since Z is independent of \mathcal{F}_s and B_s is \mathcal{F}_s -measurable and from (1.50),

$$\mathbb{E}(f(B_t) | \mathcal{F}_s) = \mathbb{E}(f(\sqrt{t-s}Z + B_s) | \mathcal{F}_s) = \mathbb{E}_{B_s}(f(\sqrt{t-s}Z)) = \mathbb{E}_{B_s}(f(B_t)) \quad (1.51)$$

which is the Markov property. From (1.48), and since B is a strong Markov process, B is also a diffusion.

1.5.1 Analytical characterization

From (1.50) and (1.51),

$$P_t f(x) = \int_{\mathbb{R}} f(y) \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2} dy.$$

Thus, from [53, p. 149], the probability transition kernel $[(t, x, y) \rightarrow p(t, x, y)]$ and the speed measure m of the standard Brownian motion are defined for every $(t, x, y) \in \mathbb{R}_+ \times \mathbb{R}^2$ by

$$p(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2}, \quad m(dy) = 2 dy.$$

We observe that for every $(t, y) \in \mathbb{R}_+ \times \mathbb{I}$, the function $[x \rightarrow p(t, x, y)] \in C_b^2(\mathbb{R})$ and solves

$$\left(\partial_t - \frac{1}{2}\partial_{xx}\right)p(t, x, y) = 0.$$

Thus, from (1.35) and (1.17),

$$L = \frac{1}{2} D_x^2, \quad \text{dom}(L) = \{f \in C_b : Lf \in C_b\}$$

and the scale function s of X is the identity function, *i.e.* $s = [x \rightarrow x]$.

Chapter 2

Singular diffusions and the sticky Brownian motion

This chapter is a synthesis of known results found in [14, 21, 30, 52, 80].

Chapter Outline: In Section 2.1, we present all possible pathwise-features a one-dimensional diffusion can exhibit, categorize them accordingly and see how these features translate into the analytical characterization. In Section 2.2, we introduce the sticky Brownian motion, prove some of its properties and compute its probability transition kernel.

2.1 Diffusion process zoology

2.1.1 Stochastic differential equations and Itô diffusions

We call stochastic differential equation or SDE any expression of the form

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t, \quad (2.1)$$

where $\mu, \sigma : \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}$ and W is a standard Brownian motion. The stochastic differential equation (2.1) can also be alternatively formulated in the integral form

$$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s. \quad (2.2)$$

When μ and σ have no time-dependence the SDE

$$dX_t = \mu(X_t) dt + \sigma(X_t) dW_t \quad (2.3)$$

is called time-homogeneous.

Definition 2.1.1 (solution, Definition 1.2 of [80], Chapter IX, §1). *A solution of the SDE* (2.1) is a pair (X, B) of adapted processes defined on a probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ such that B is a $(\mathcal{F}_t)_{t \geq 0}$ -standard Brownian motion and (2.2) holds.

Not unlike ordinary differential equations, questions of existence and uniqueness of a solution to (2.1) arise.

Definition 2.1.2 (path-wise uniqueness, Definition 1.3-(1) of [80], Chapter IX, §1). The SDE (2.2) is said to have *path-wise uniqueness* if for two pairs $(X, W), (X', W')$, defined on the same probability space, that solve it with $W = W'$ and $X_0 = X'_0$ almost surely, then $X = X'$ almost surely.

A necessary condition for path-wise uniqueness in a one-dimensional setting is given by the following theorem.

Theorem 2.1.3 (Yamada-Watanabe). *Let μ, σ be two measurable real-valued functions such that*

- *there exists a function $\rho : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that*

$$\int_{0+} \frac{du}{\rho^2(u)} = \infty$$

and for all $x, y \in \mathbb{R}$, $|\sigma(x) - \sigma(y)| \leq \rho(|x - y|)$,

- *b is Lipschitz continuous.*

Then, path-wise uniqueness holds for (2.1).

For other versions of this result see [80, Chapter IX, Section §3]. In [10] an example of an SDE where no path-wise uniqueness holds for its solution.

Definition 2.1.4 (uniqueness in law, Definition 1.3-(2) of [80], Chapter IX, §1). There is *uniqueness in law* for the SDE (2.2) if for any two of its solutions $(X, W), (X', W')$ (not necessarily defined on the same probability space) with $\text{Law}(X_0) = \text{Law}(X'_0)$, the processes X and X' have the same law.

Proposition 2.1.5 (Proposition 1.4 of [80], Chapter IX, §1). *There is uniqueness in law for an SDE if for every $x \in \mathbb{R}$ and any two of its solutions $(X, W), (X', W')$ such that $X_0 = X'_0 = x$, then X and X' have the same law.*

To define the notion of strong solution, we introduce the natural filtration of a process. The natural filtration of a process X is the smallest filtration $(\mathcal{F}_t^X)_{t \geq 0}$ to which X is adapted.

Definition 2.1.6 (strong solution, Definition 1.5 of [80], Chapter IX, §1). A solution of an SDE is called *strong solution* iff X is adapted to $(\mathcal{F}_t^W)_{t \geq 0}$, where $(\mathcal{F}_t^W)_{t \geq 0}$ is the natural filtration of W . A solution of an SDE that is not strong will be called *weak solution*.

Proposition 2.1.7 (Corollary 3.23 of [62], Chapter 5). *Let (X, W) be a weak solution of (2.1) defined on a probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)$. Then, if path-wise uniqueness hold for (2.1). X is a strong solution of (2.1).*

The notion of strong solution be defined alternatively (see [80, Chapter IX; §1]) as:

Proposition 2.1.8 (Yamada-Watanabe). *Let (X, W) be a solution of (2.2). Then, it is a strong solution of (2.2) iff there exists a measurable function $F : C(\mathbb{R}_+, \mathbb{R}) \rightarrow C(\mathbb{R}_+, \mathbb{R})$ such that*

$$X = F(W).$$

The difference between these two notions of strong and weak solution is subtle. A strong solution is uniquely determined by the driving Brownian motion W . This is not the case for a weak solution.

For necessary conditions for the existence, path-wise uniqueness, uniqueness in law of solutions and the notions of a strong/weak solution along with relevant examples see [21] and [80, Chapter IX; §2, §3].

Proposition 2.1.9. *A process that solves (2.3) is a diffusion process. We call such processes Itô diffusions.*

Proof. Let X be a process that solves (2.3). The path-wise continuity of X is direct from the expression (2.3). For any $f \in C_b$, $t > 0$ and stopping time τ ,

$$\begin{aligned} \mathbb{E}_x \left(f(X_{\tau+t}) \middle| \mathcal{F}_\tau \right) &= \mathbb{E}_x \left(f \left(X_0 + \int_0^\tau \mu(X_s) ds + \int_0^\tau \sigma(X_s) dW_s \right) \middle| \mathcal{F}_\tau \right) \\ &= \mathbb{E}_{X_\tau} \left(f \left(X_\tau + \int_\tau^{\tau+t} \mu(X_s) ds + \int_\tau^{\tau+t} \sigma(X_s) dW_s \right) \right). \end{aligned} \quad (2.4)$$

Also,

$$\begin{aligned} \mathbb{E}_{X_\tau} \left(f(X_t \circ \theta_\tau) \right) &= \mathbb{E}_{X_\tau} \left(f \left(X_0 \circ \theta_\tau + \int_0^t \mu(X_s \circ \theta_\tau) ds + \int_0^t \sigma(X_s \circ \theta_\tau) d(W_s \circ \theta_\tau) \right) \right) \\ &= \mathbb{E}_{X_\tau} \left(f \left(X_\tau + \int_\tau^{\tau+t} \mu(X_s) ds + \int_\tau^{\tau+t} \sigma(X_s) dW_s \right) \right). \end{aligned} \quad (2.5)$$

From (2.4) and (2.5), the process X satisfies the strong Markov property (1.1). \square

We now prove the analytical characterization of Itô diffusions.

Proposition 2.1.10. *Let X be the diffusion process that solves (2.3) and L its infinitesimal generator. Then,*

$$L f = \mu f' + \frac{1}{2} \sigma^2 f'' \quad (2.6)$$

for every $f \in \text{dom}(L)$, where $\text{dom}(L) = \{f \in C_b(\mathbb{I}) : L f \in C_b(\mathbb{I})\}$. Moreover, if $\mu, \sigma \in C(\mathbb{I})$, $\text{dom}(L) = C_b^2(\mathbb{I})$.

Proof. From Itô's lemma, for any $f \in C^2(\mathbb{I})$,

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\langle X \rangle_s \\ &= f(X_0) + \int_0^t \left(f'(X_s) \mu(X_s) + \frac{1}{2} f''(X_s) \sigma^2(X_s) \right) ds \\ &\quad + \frac{1}{2} \int_0^t f'(X_s) \sigma(X_s) dW_s. \end{aligned}$$

From (1.16),

$$\begin{aligned} L f(x) &= \lim_{t \rightarrow 0} \frac{1}{t} \left(\mathbb{E}_x [f(X_t)] - f(x) \right) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{E}_x \left[\int_0^t \left(f'(X_s) \mu(X_s) + \frac{1}{2} f''(X_s) \sigma^2(X_s) \right) ds \right] \\ &= \mathbb{E}_x \left[\lim_{t \rightarrow 0} \frac{1}{t} \int_0^t \left(f'(X_s) \mu(X_s) + \frac{1}{2} f''(X_s) \sigma^2(X_s) \right) ds \right]. \end{aligned} \quad (2.7)$$

From the intermediate value theorem, P_x -almost surely

$$\frac{1}{t} \int_0^t \left(f'(X_s) \mu(X_s) + \frac{1}{2} f''(X_s) \sigma^2(X_s) \right) ds \longrightarrow f'(x) \mu(x) + \frac{1}{2} f''(x) \sigma^2(x), \quad (2.8)$$

as $t \longrightarrow 0$. From (2.7) and (2.8), (2.6) is proven. □

Proposition 2.1.11. *Let (μ, σ) be such that*

$$dX_t = \mu(X_t) dt + \sigma(X_t) dW_t$$

has a unique weak solution X with state-space \mathbb{I} an interval of \mathbb{R} and scale function & speed measure (s, m) . Then, for every $x \in \mathbb{I}$,

$$s(x) = \int_a^x e^{-\int_a^y \frac{2\mu(u)}{\sigma^2(u)} du} dy \quad \text{and} \quad m(dx) = \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} dx. \quad (2.9)$$

Proof. From (1.43) and (2.6), s solves

$$s'' + 2 \frac{\mu}{\sigma^2} s' = 0$$

and consequently

$$s'(x) = c_0 e^{-\int_a^x \frac{2\mu(u)}{\sigma^2(u)} du}, \quad (2.10)$$

where $c_0 \in \mathbb{R}$. From (1.23) and (2.6), for any $f \in \text{dom}(L)$,

$$D_m D_s f = \mu f' + \frac{1}{2} \sigma^2 f''. \quad (2.11)$$

From (1.26) and (2.10),

$$D_s f(x) = f'(x)/s'(x) = c_0^{-1} e^{\int_a^x \frac{2\mu(u)}{\sigma^2(u)} du} f'(x)$$

Let $m(dx) \ll dx$ and m' be the Radon-Nikodym derivative $m' = \frac{dm}{dx}$. From (1.26), $D_m = m' D_x$. Thus,

$$D_m D_s f(x) = m'(x) D_x \frac{f'(x)}{s'(x)} = m'(x) \left(\frac{2\mu(x)}{\sigma^2(x)} \frac{f'(x)}{s'(x)} + \frac{f''(x)}{s'(x)} \right) \quad (2.12)$$

From (2.11) and (2.12),

$$m'(x) = \frac{2}{\sigma^2(x)} \frac{1}{s'(x)}$$

which proves (2.9). □

2.1.2 Boundary classification

The paths of a diffusion process may exhibit various features in the vicinity of a boundary point of its state-space. We refer to these features as boundary behavior of a diffusion. It turns out the boundary behavior of a diffusion can be classified as follows (see Section 5.11 of Itô's book [52]):

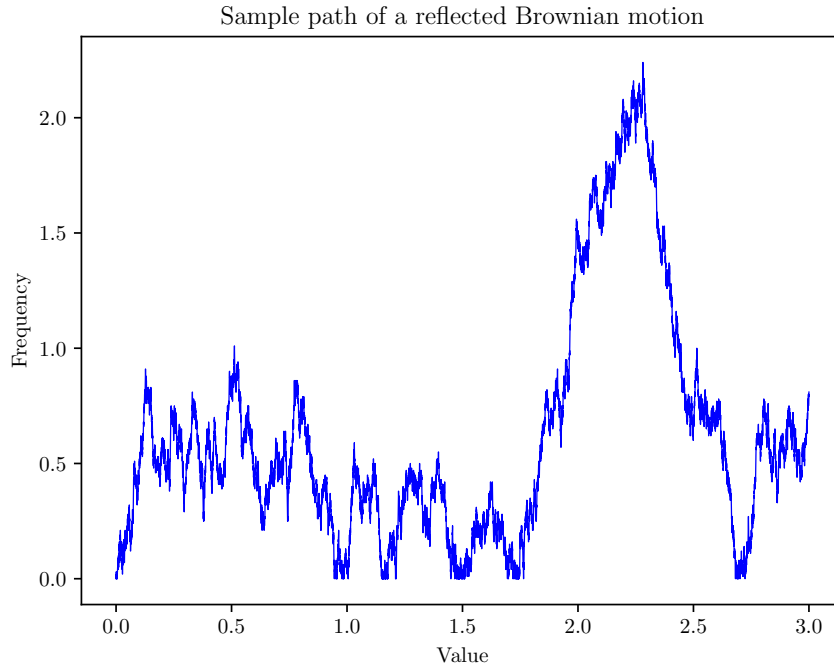


Figure 2.1: Simulated trajectory of a reflected Brownian motion using Algorithm 1.

Proposition 2.1.12 (see [63]). *Let X be a diffusion process with state-space \mathbb{I} , an interval of \mathbb{R} . A boundary point of \mathbb{I} can either be of the following types:*

- *Regular boundary: the process can both enter and leave from a regular boundary (reflection, sticky reflection, absorption),*
- *Exit boundary: the process can reach the boundary from an interior point but cannot reach an interior point from the boundary (explosion, absorption),*
- *Entrance boundary: the process can reach an interior point from the boundary but cannot reach the boundary from an interior point (repulsion),*
- *Natural boundary: the process cannot reach the boundary from an interior point and an interior point cannot be reached from the boundary.*

One can check the boundary behavior of a diffusion process by looking at the behavior of its speed measure in the vicinity of the boundary. This can be done as follows:

Proposition 2.1.13 (see Section 5.11 of [52]). *Let X be a diffusion process defined through (s, m) with state-space \mathbb{I} , an interval of \mathbb{R}_+ , and 0 be a boundary point of \mathbb{I} . For $c > 0$, let*

$$\mathcal{I} = \int \int_{0 < y < x < c} m(dx) dy, \quad \mathcal{II} = \int \int_{0 < y < x < c} m(dy) dx.$$

Then,

- *0 is a regular boundary iff $\mathcal{I} < \infty$ and $\mathcal{II} < \infty$,*

- 0 is an exit boundary iff $\mathcal{I} < \infty$ and $\mathcal{I}\mathcal{I} = \infty$,
- 0 is an entrance boundary iff $\mathcal{I} = \infty$ and $\mathcal{I}\mathcal{I} < \infty$.
- 0 is a natural boundary iff $\mathcal{I} = \infty$ and $\mathcal{I}\mathcal{I} = \infty$.

2.1.3 Singular diffusions

From (2.9), we observe that the scale function and speed measure of any Itô diffusion satisfy

$$s \in C^1(\mathbb{I}) \quad \text{and} \quad m(dx) \ll dx. \quad (2.13)$$

These processes are called *absolutely continuous*. Diffusion processes that do not satisfy (2.13) are called *singular diffusions*. Diffusion processes that do not satisfy (2.13) at a single point $\zeta \in \mathbb{I}$ of their state-space are said to have a *point-wise singularity*. These singularities correspond to specific path-wise features and can be categorized as follows.

Definition 2.1.14. Let X be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} and scale function s . The process X is said to have a *skew point* at $\zeta \in \text{Int}(\mathbb{I})$ iff $s'(\zeta-) \neq s'(\zeta+)$, where s' is the right-derivative of s .

Definition 2.1.15. A diffusion X with state-space \mathbb{I} and speed measure m is said to have a *sticky point* at $\zeta \in \mathbb{I}$ iff $m(\{\zeta\}) > 0$. The quantity $m(\{\zeta\})$ is called *stickiness* of X at ζ .

Proposition 2.1.16. Let X be a diffusion process with state-space \mathbb{I} such that $0 \in \text{Int}(\mathbb{I})$. Then, X has a skew point at 0 iff

$$\lim_{h \rightarrow 0} \left(P_0(\tau_{-h} < \tau_h) - P_0(\tau_h < \tau_{-h}) \right) \neq 0. \quad (2.14)$$

Proof. From (1.18),

$$P_0(\tau_{-h} < \tau_h) = \frac{s(h) - s(0)}{s(h) - s(-h)} \longrightarrow \frac{s'(0+)}{s'(0+) + s'(0-)}$$

and

$$P_0(\tau_{-h} < \tau_h) = \frac{s(0) - s(-h)}{s(h) - s(-h)} \longrightarrow \frac{s'(0-)}{s'(0+) + s'(0-)}$$

as $h \rightarrow 0$. Thus, (2.14) holds iff $s'(0-) \neq s'(0+)$, which concludes the proof. \square

Proposition 2.1.17. Let X be a diffusion process with state-space \mathbb{I} such that $0 \in \mathbb{I}$. Then, X has a sticky point at 0 iff it spends a positive amount of time at 0 on the event $\{\tau_0 < \infty\}$.

Proof. Let s be the scale function and m the speed measure of X . Let also X be defined on a family of probability spaces $(\Omega, (\mathcal{F}_t)_{t \geq 0}, P_x)_{x \in \mathbb{I}}$ such that for every $x \in \mathbb{I}$, $P(X_0 = x) = 1$. From Proposition 1.2.22 and Theorem 1.4.4, there exists a Brownian motion B defined on an extension of the probability space such that $s(X) = B_{\gamma(t)}$, where γ is the right-inverse of

$$A(t) = \int_{\mathbb{I}} L_t^y(B) m'(dy)$$

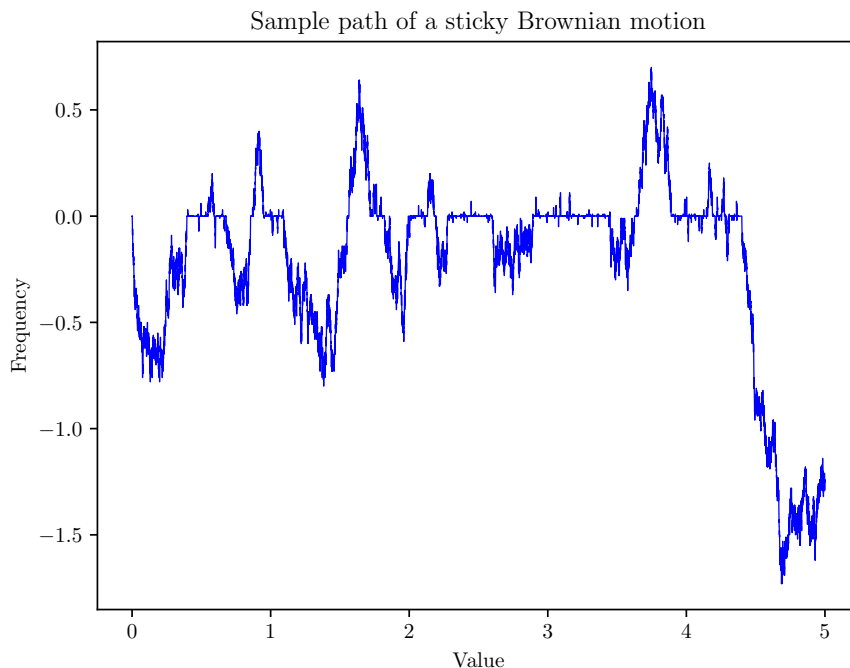


Figure 2.2: Simulated trajectory of a sticky Brownian motion of parameter $\rho = 1$ using Algorithm 1.

and $m'(dx) = m(ds^{-1}(x))$. Let $\rho = m(\{0\}) > 0$ be the stickiness at 0 of X , m_\circ the locally finite strictly positive measure defined by $m_\circ = m - m(\{0\})\delta_0$ and γ° the right-inverse of

$$A^\circ(t) = \int_{\mathbb{I}} L_t^y(B) m'_\circ(dy).$$

From Proposition 1.2.22 and Corollary 1.4.5, the process $X^\circ = (s^{-1}(B_{\gamma^\circ(t)}))_{t \geq 0}$ is a diffusion process with scale function s and speed measure m_\circ . Also,

$$A(t) = A^\circ(t) + \frac{\rho}{2} L_t^0(B)$$

and

$$A(t) = A_\rho \circ A^\circ(t),$$

where $A_\rho(t) = t + \frac{\rho}{2} L_{\gamma_\rho(t)}^0(B) = t + \frac{\rho}{2} L_t^0(X^\circ)$. Thus, $X = (X_{\gamma_\rho(t)}^\circ)_{t \geq 0}$, where γ_ρ is the right-inverse of A_ρ , and from Theorem A.2.2 and Lemma A.2.4,

$$\begin{aligned} \int_0^t \mathbb{1}_{X_s=0} ds &= \int_0^t \mathbb{1}_{X_{\gamma_\rho(s)}^\circ=0} ds = \int_0^{\gamma_\rho(t)} \mathbb{1}_{X_{\gamma_\rho(s)}^\circ=0} dA_\rho(s) \\ &= \int_0^{\gamma_\rho(t)} \mathbb{1}_{X_{\gamma_\rho(s)}^\circ=0} ds + \frac{\rho}{2} \int_0^{\gamma_\rho(t)} \mathbb{1}_{X_{\gamma_\rho(s)}^\circ=0} dL_s^0(X^\circ) = \frac{\rho}{2} L_{\gamma_\rho(t)}^0(X^\circ) = L_t^0(X) \end{aligned}$$

which proves the result. \square

We now give several examples of singular diffusions. When we want to introduce a new path-wise feature, it is common practice to first study the most elementary object having it. Among diffusion processes, the most elementary is the standard Brownian motion. We define the skew and sticky Brownian motions, which are respectively the most elementary skew and sticky diffusion processes.

Example 2.1.18 (Skew Brownian motion, see [47]). The skew Brownian motion of parameter $\beta \in (0, 1)$ is the diffusion process defined through s and m , where

$$s(x) = \begin{cases} x/\beta & x > 0 \\ x/(1-\beta) & x \leq 0 \end{cases}, \quad m(dx) = \begin{cases} 2\beta x & x > 0 \\ 2(1-\beta)x & x \leq 0 \end{cases},$$

for every $x \in \mathbb{R}$. We observe that

$$s'(0+) - s'(0-) = (1 - 2\beta)/\beta(1 - \beta).$$

Thus $s \in C^1(\mathbb{R})$ iff $\beta = 1/2$. The cases $\beta = 1/2$ and $\beta \in \{0, 1\}$ correspond to the standard and reflected Brownian motion respectively.

Example 2.1.19 (Sticky Brownian motion, see [33, 53]). The sticky Brownian motion is the diffusion process defined through s and m , where

$$s(x) = x, \quad m(dx) = 2 dx + \rho \delta_0(dx),$$

for every $x \in \mathbb{R}$. We observe that $m(\{0\}) = \rho > 0$, which is the stickiness of the process at 0. Section 2.2 is dedicated to this process. The cases $\rho = 0$ and $\rho = \infty$ correspond to the standard Brownian motion and the Brownian motion absorbed at 0 (see Section 2.2.4).

The following example illustrates that it is possible to have a point that is both sticky and skew for a diffusion process.

Example 2.1.20 (Sticky-skew Brownian motion, see [88]). The sticky-skew Brownian motion is the diffusion process defined through s and m , where

$$s(x) = \begin{cases} x/\beta & x > 0 \\ x/(1-\beta) & x \leq 0 \end{cases}, \quad m(dx) = \rho \delta_0(dx) + \begin{cases} 2\beta x & x > 0 \\ 2(1-\beta)x & x \leq 0 \end{cases},$$

for every $x \in \mathbb{R}$. We observe that the process is both sticky and skew at 0.

A process X such that there exists a countable subset \mathcal{N} of \mathbb{I} such that

$$\mathbb{1}_{\mathbb{I} \setminus \mathcal{N}} m(dx) \ll \mathbb{1}_{\mathbb{I} \setminus \mathcal{N}} dx \quad \text{and} \quad s \in C^1(\mathbb{I} \setminus \mathcal{N}) \cap C^1(\mathbb{I})$$

is called a diffusion with point-wise singularities. Then, each $x \in \text{Int}(\mathbb{I}) \cap \mathcal{N}$ is either a skew and/or a sticky point of X .

The following example illustrates that there are processes that belong to none of the aforementioned categories:

Example 2.1.21 (Brownian motion slowed on the Cantor set, see [7]). Let,

- $C_{1/3}$ be the Cantor set (see [84, p. 8, 38])

$$C_{1/3} = \left\{ \sum_{n=1}^{\infty} \frac{x^n}{3^n} : x_i \in \{0, 2\}, \forall i \in \mathbb{N} \right\},$$

- $f_{C_{1/3}}$ the Cantor-Lebesgue function (see [84]-p.38,126), i.e. a continuous increasing function, with zero-derivative on $\mathbb{R} \setminus C_{1/3}$ such that

$$f_{C_{1/3}}(0) = 0, \quad f_{C_{1/3}}(1) = 1,$$

- $m_{C_{1/3}}$ the measure defined by extension on $\mathcal{B}(\mathbb{R})$ such that for any $x < y$,

$$m_{C_{1/3}}((x, y)) = f_{C_{1/3}}(y) - f_{C_{1/3}}(x),$$

- X the diffusion process on natural scale with speed measure

$$m(dx) = 2 dx + m_{C_{1/3}}(dx).$$

We observe that X is a singular diffusion process with no point-wise singularities.

2.2 The sticky Brownian motion

2.2.1 Definition

The sticky Brownian motion is the diffusion process that behaves like a standard Brownian motion away from 0 which is a sticky point for the process. We can define it through s and m as follows.

Definition 2.2.1. The *sticky Brownian motion* is the diffusion process with state-space \mathbb{R} defined through s and m , where, for every $x \in \mathbb{R}$:

$$s(x) = x \quad \text{and} \quad m(dx) = 2 dx + \rho \delta_0(dx). \quad (2.15)$$

Proposition 2.2.2. *The infinitesimal generator of the sticky Brownian motion of stickiness $\rho > 0$ is:*

$$\begin{aligned} L &= \frac{1}{2} \frac{d^2}{dx^2}, \\ \text{dom}(L) &= \left\{ f \in C_b(\mathbb{R}) : f \in C^2(\mathbb{R}^*); \right. \\ &\quad \left. \frac{1}{\rho}(f(0+) - f(0-)) = \frac{1}{2}f''(0-) = \frac{1}{2}f''(0+) \right\}. \end{aligned}$$

Proof. It is a particular case of Proposition 3.1.3 of Chapter 3. □

Proposition 2.2.3. *Let X be a sticky Brownian motion of stickiness $\rho > 0$. Then, X solves*

$$\begin{aligned} dX_t &= \mathbb{1}_{X_t \neq 0} dB_t, \\ \mathbb{1}_{X_t = 0} dt &= \frac{\rho}{2} dL_t^0(X), \end{aligned}$$

where B is a standard Brownian motion.

Proof. It is a particular case of Proposition 3.2.3 of Chapter 3. □

Proposition 2.2.4. *Let X be a sticky Brownian motion of stickiness $\rho > 0$ defined on a probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)$ such that $\mathbb{P}_x(X_0 = x) = 1$. Then, there exists a Brownian motion W such that $\mathbb{P}_x(W_0 = x) = 1$ and for every $t \geq 0$,*

$$X_t = W_{\gamma(t)} \quad \text{and} \quad \gamma^{-1}(t) = t + \frac{\rho}{2} L_t^0(X), \quad (2.16)$$

where γ^{-1} is the right-inverse of γ . Since almost surely γ^{-1} is strictly increasing and continuous, so is γ and thus γ^{-1} is the proper inverse of γ .

Proof. It is a particular case of Theorem 1.4.4. □

2.2.2 Properties

Proposition 2.2.5. *Let X be a sticky Brownian motion of stickiness $\rho > 0$ and $\tau_0 = \inf\{s > 0 : X_s = 0\}$ the hitting time of 0 by X . Then, on the event $\{\tau_0 < t\}$, the random set $O_t = \{s < t : X_s = 0\}$ is almost surely Cantor-like, i.e. totally disconnected and of positive Lebesgue measure.*

Proof. Let $t_1, t_2 \in O_t$ such that $t_1 < t_2$ and γ the random time-change defined in (2.16). From Corollary 2.2.4, $\gamma(t_1) < \gamma(t_2)$ and $W_{\gamma(t_1)} = W_{\gamma(t_2)} = 0$.

The Brownian motion is almost surely not constant on any interval, there exists a $\gamma(t_1) < \zeta < \gamma(t_2)$ such that:

$$W_\zeta = 0 \quad \text{and} \quad X_{\gamma^{-1}(\zeta)} = 0,$$

where $t_1 < \gamma^{-1}(\zeta) < t_2$. Thus O_t is totally disconnected. The fact that the process spends a positive amount is a particular case of Proposition 2.1.17. □

Proposition 2.2.6. *Let $\{(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}_x^\rho); x \in \mathbb{R}, \rho \geq 0\}$ be a family of filtered probability spaces and $X^\rho = (X_t^\rho)_{t \geq 0}$ a process defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0})$ such that under \mathbb{P}_x^ρ it is the sticky Brownian motion of stickiness parameter ρ and $\mathbb{P}_x^\rho(X_0^\rho = x)$. Then,*

$$\text{Law}_{\mathbb{P}_x^\rho} \left(X_{ct}^\rho, L_{ct}^0(X^\rho); t \geq 0 \right) = \text{Law}_{\mathbb{P}_x^{\rho/\sqrt{c}}} \left(\sqrt{c} X_t^{\rho/\sqrt{c}}, \sqrt{c} L_t^0(X^{\rho/\sqrt{c}}); t \geq 0 \right), \quad (2.17)$$

where $L_t^0(X^\rho)$ and $L_t^0(X^{\rho/\sqrt{c}})$ are the local times of X^ρ and $X^{\rho/\sqrt{c}}$ respectively.

We prove this result by expressing the sticky Brownian motion as a time-changed Brownian motion. A simpler proof is given in the Appendix that makes use of the joint density of the process with its local time (see Lemma A.3.1).

Proof of Proposition 2.2.6. Let X^ρ be a process defined on $\mathcal{P}_x^\rho = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}_x^\rho)$ such that under \mathbb{P}_x^ρ it is a sticky Brownian motion of stickiness ρ . From Theorem 1.4.4, there exists a Brownian motion B defined on an extension of \mathcal{P}_x^ρ such that for every $t \geq 0$,

$$X_t^\rho = B_{\gamma(t)}$$

where $\gamma(t)$ is the right inverse of $A(t) = \frac{1}{2} \int_{\mathbb{R}} L_t^y(B) m(dy)$. Moreover from (A.4), \mathbb{P}_x^ρ -almost surely, for every $t \geq 0$,

$$L_t^0(X^\rho) = L_{\gamma(t)}^0(B). \quad (2.18)$$

From Proposition 1.10 of [80, Chapter I], for every $c > 0$, the process B' defined as,

$$B'_t = B_{ct}/\sqrt{c}, \quad (2.19)$$

is a standard Brownian motion. Moreover, \mathbb{P}_x^ρ -almost surely,

$$\begin{aligned} L_{ct}^0(B) &= \mathbb{P}_x\text{-}\lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_0^{ct} \mathbb{1}_{|B_s| < \epsilon} ds = \mathbb{P}_x\text{-}\lim_{\epsilon \rightarrow 0} \frac{c}{2\epsilon} \int_0^t \mathbb{1}_{|B_{cs}| < \epsilon} ds \\ &= \mathbb{P}_x\text{-}\lim_{\epsilon \rightarrow 0} \frac{c}{2\epsilon} \int_0^t \mathbb{1}_{|B'_s| < \epsilon/\sqrt{c}} ds = \sqrt{c} L_t^0(B'). \end{aligned} \quad (2.20)$$

Let $A_\rho(t) = t + \frac{\rho}{2} L_t^0(B)$ and $A'_\rho(t) = t + \frac{\rho}{2} L_t^0(B')$ with $\gamma_\rho(t)$ and $\gamma'_\rho(t)$ be their respective right inverses. From (2.20), \mathbb{P}_x^ρ -almost surely,

$$A_\rho(ct) = ct + \frac{\rho}{2} L_{ct}^0(B) = ct + \sqrt{c} \frac{\rho}{2} L_t^0(B') = c \left(t + \frac{\rho}{2\sqrt{c}} L_t^0(B') \right) = c A'_{\rho/\sqrt{c}}(t),$$

$$\gamma_\rho(ct) = \inf \{ s > 0 : A_\rho(s) > ct \} = \inf \{ s > 0 : A'_{\rho/\sqrt{c}}(s/c) > t \} = c \gamma'_{\rho/\sqrt{c}}(t). \quad (2.21)$$

From (2.19) and (2.21), \mathbb{P}_x^ρ -almost surely,

$$B_{\gamma_\rho(ct)} = \sqrt{c} B'_{\gamma'_{\rho/\sqrt{c}}(t)}.$$

Let $X^{\rho/\sqrt{c}}$ be the process such that

$$X_t^{\rho/\sqrt{c}} = B'_{\gamma'_{\rho/\sqrt{c}}(t)}, \quad (2.22)$$

which is a martingale from the martingale stopping theorem. As such its local time $L^0(X^{\rho/\sqrt{c}})$ is a well defined object. From (A.4), \mathbb{P}_x -almost surely, for every $t \geq 0$,

$$L_t^0(X^{\rho/\sqrt{c}}) = L_{\gamma'_{\rho/\sqrt{c}}(t)}^0(B'). \quad (2.23)$$

We observe that for a measure $m_\rho(dy) = 2dy + \rho\delta_0(dy)$,

$$A_\rho(t) = \int_{\mathbb{R}} L_t^y(B) m_\rho(dy) \quad \text{and} \quad A'_\rho(t) = \int_{\mathbb{R}} L_t^y(B') m_\rho(dy).$$

Thus, from (2.22) and Corollary 1.4.5, $\sqrt{c} X_t^{\rho/\sqrt{c}} = X_{ct}^\rho$ is a sticky Brownian motion of stickiness parameter ρ/\sqrt{c} and whose local time process $L^0(X^{\rho/\sqrt{c}})$, from (2.18), (2.20) and (2.23), satisfies \mathbb{P}_x^ρ -almost-surely

$$L_t^0(X^{\rho/\sqrt{c}}) = L_t^0(X^\rho)/\sqrt{c},$$

for every $t \geq 0$. Thus, setting $\mathbb{P}_x^{\rho/\sqrt{c}} = \mathbb{P}_{x\sqrt{c}}^\rho$, (A.5) is proven. \square

Corollary 2.2.7. *Let $X^\rho = (X_t^\rho)_{t \geq 0}$ be the sticky Brownian motion of stickiness parameter $\rho > 0$ and $(P_t^\rho)_{t \geq 0}$ its semi-group. Then, for every measurable function $h : \mathbb{R} \mapsto \mathbb{R}$,*

$$P_t^{\rho\sqrt{n}} h(x\sqrt{n}) = \mathbb{E}_x \left(h(\sqrt{n} X_{\frac{t}{n}}^\rho) \right). \quad (2.24)$$

2.2.3 Probability transition kernel computation

Proposition 2.2.8. *The probability transition kernel of the sticky Brownian motion of stickiness ρ with respect to m is the function $p : \mathbb{R}_+ \times \mathbb{R}^2 \mapsto \mathbb{R}_+$ defined for every $t > 0$ and $x, y \in \mathbb{R}$ by*

$$p_\rho(t, x, y) = u_1(t, x, y) - u_2(t, x, y) + v_\rho(t, x, y), \quad (2.25)$$

where

$$\begin{cases} u_1(t, x, y) &= \frac{1}{2\sqrt{2\pi t}} e^{-(x-y)^2/2t}, \\ u_2(t, x, y) &= \frac{1}{2\sqrt{2\pi t}} e^{-(|x|+|y|)^2/2t}, \\ v_\rho(t, x, y) &= \frac{1}{\rho} e^{4(|x|+|y|)/\rho+8t/\rho^2} \operatorname{erfc}\left(\frac{|x|+|y|}{\sqrt{2t}} + \frac{2\sqrt{2t}}{\rho}\right). \end{cases} \quad (2.26)$$

Lemma 2.2.9. *The resolvent kernel $r(\lambda, x, y)$ of the sticky Brownian motion solves*

$$(\lambda - L_x)r(\lambda, x, y) = 0, \quad (2.27)$$

$$r(\lambda, 0-, y) = r(\lambda, 0+, y), \quad (2.28)$$

$$r(\lambda, y-, y) = r(\lambda, y+, y), \quad (2.29)$$

$$\frac{1}{\rho}(\partial_x r(\lambda, 0+, y) - \partial_x r(\lambda, 0-, y)) = \frac{1}{2}\partial_{xx}r(\lambda, 0, y), \quad (2.30)$$

where $L_x = \frac{1}{2}D_x^2$ and $\partial_x u(x, y) = \lim_{h \rightarrow 0} \frac{1}{h}(u(x+h, y) - u(x, y))$. Moreover, for all $x \in \mathbb{R}$,

$$\lim_{|y| \rightarrow \infty} r(\lambda, x, y) = 0. \quad (2.31)$$

Proof. From Lemma 1.3.5, the resolvent kernel solves (2.27)-(2.30). From (1.30),

$$\int_{\mathbb{I}} p(t, x, y)(2 dy + \rho \delta_0(dy)) = 1.$$

Thus,

$$\int_{\mathbb{I}} r(\lambda, x, y)(2 dy + \rho \delta_0(dy)) = \int_0^\infty e^{-\lambda t} dt = \frac{1}{\lambda},$$

which proves (2.31). \square

Lemma 2.2.10 (Theorem 6.1 of [35]). *For each $\lambda > 0$ the differential equation*

$$\lambda u(\lambda, x) - L_x u(\lambda, x) = 0 \quad (2.32)$$

has two convex solutions, one increasing ψ_λ from 0 to ∞ , one decreasing ϕ_λ from ∞ to 0 such that $\psi_\lambda(0) = \phi_\lambda(0) = 1$. The pair $(\phi_\lambda, \psi_\lambda)$ is called minimal pair or minimal solutions.

Example 2.2.11. If $L = \frac{1}{2}D_x^2$, then, for any $\lambda > 0$, the minimal solutions of (2.32) are

$$\phi_\lambda(x) = e^{-\sqrt{2\lambda}x}, \quad \psi_\lambda(x) = e^{\sqrt{\lambda}x}.$$

Lemma 2.2.12. *The Green function $g(\lambda, x, y)$ of Problem (2.27)-(2.31) solves (2.27)-(2.31) along with*

$$\partial_x g(\lambda, y-, y) - \partial_x g(\lambda, y+, y) = 2 \quad (2.33)$$

and is

$$g(\lambda, x, y) = A_\lambda e^{-\sqrt{2\lambda}(|y|+|x|)} + B_\lambda e^{-\sqrt{2\lambda}|y-x|}, \quad (2.34)$$

where

$$A_\lambda = -\left(1 + \frac{2\sqrt{2}}{\rho\sqrt{\lambda}}\right)^{-1} \frac{1}{\sqrt{2\lambda}}, \quad B_\lambda = \frac{1}{\sqrt{2\lambda}}.$$

Proof. Let $(\phi_\lambda, \psi_\lambda)$ be the minimal pair solving (2.32). Let also $g(\lambda, x, y)$ be a function of the form

$$g(\lambda, x, y) = \begin{bmatrix} \phi_\lambda(x) \\ \psi_\lambda(x) \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \phi_\lambda(y) \\ \psi_\lambda(y) \end{bmatrix}^t,$$

where A^t is the transpose of the matrix A , supposed C^2 on every of the sets

$$\begin{aligned} D_1 &= \{0 < x < y\}, & D_3 &= \{y < x < 0\}, & D_5 &= \{x < 0 < y\}, \\ D_2 &= \{0 < y < x\}, & D_4 &= \{x < y < 0\}, & D_6 &= \{y < 0 < x\}, \end{aligned}$$

delimited by $\{x = y\}$, $\{x = 0\}$ and $\{y = 0\}$, where $\mathbb{R}^2 = \bigcup_{i=1}^6 \bar{D}_i$. Thus, there exists constants $\{(a_i, b_i, c_i, d_i)\}_i$ such that

$$g(\lambda, x, y) = \begin{bmatrix} \phi_\lambda(x) \\ \psi_\lambda(x) \end{bmatrix} \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \begin{bmatrix} \phi_\lambda(y) \\ \psi_\lambda(y) \end{bmatrix}^t$$

for any $(x, y) \in D_i$ and $i \in \{1, 2, \dots, 6\}$. From (2.31),

$$\begin{aligned} c_1 &= d_1 = 0, \\ b_2 &= d_2 = 0, \\ a_3 &= b_3 = 0, \\ a_4 &= c_4 = 0, \\ a_5 &= c_5 = d_5 = 0, \\ a_6 &= b_6 = d_6 = 0 \end{aligned}$$

and

$$g(\lambda, x, y) = \begin{cases} \phi_\lambda(y) (a_1 \phi_\lambda(x) + b_1 \psi_\lambda(x)), & 0 < x < y, \\ \phi_\lambda(x) (a_2 \phi_\lambda(y) + c_2 \psi_\lambda(y)), & 0 < y < x, \\ \psi_\lambda(y) (c_3 \phi_\lambda(x) + d_3 \psi_\lambda(x)), & y < x < 0, \\ \psi_\lambda(x) (b_4 \phi_\lambda(y) + d_4 \psi_\lambda(y)), & x < y < 0, \\ b_5 \phi_\lambda(y) \psi_\lambda(x), & x < 0 < y, \\ c_6 \phi_\lambda(x) \psi_\lambda(y), & y < 0 < x, \end{cases} \quad (2.35)$$

$$\partial_x g(\lambda, x, y) = \sqrt{2\lambda} \begin{cases} \phi_\lambda(y) (-a_1 \phi_\lambda(x) + b_1 \psi_\lambda(x)), & 0 < x < y, \\ -\phi_\lambda(x) (a_2 \phi_\lambda(y) + c_2 \psi_\lambda(y)), & 0 < y < x, \\ \psi_\lambda(y) (-c_3 \phi_\lambda(x) + d_3 \psi_\lambda(x)), & y < x < 0, \\ \psi_\lambda(x) (b_4 \phi_\lambda(y) + d_4 \psi_\lambda(y)), & x < y < 0, \\ b_5 \phi_\lambda(y) \psi_\lambda(x), & x < 0 < y, \\ -c_6 \phi_\lambda(x) \psi_\lambda(y), & y < 0 < x, \end{cases}$$

$$\partial_{xx}g(\lambda, x, y)g(\lambda, x, y) = 2\lambda g(\lambda, x, y).$$

From (2.28),

$$b_5 = a_1 + b_1, \quad c_6 = c_3 + d_3. \quad (2.36)$$

From (2.29),

$$\begin{aligned} a_1 &= a_2, & b_1 &= c_2, \\ c_3 &= b_4, & d_3 &= d_4. \end{aligned} \quad (2.37)$$

From (2.33),

$$b_1 + c_2 = \sqrt{2/\lambda}, \quad b_4 + c_3 = \sqrt{2/\lambda}. \quad (2.38)$$

From (2.37) and (2.38),

$$b_1 = c_2 = c_3 = b_4 = 1/\sqrt{2\lambda}. \quad (2.39)$$

From (2.30),

$$b_1 - a_1 - b_5 = \rho\sqrt{\lambda/2}b_5, \quad c_3 - c_6 - d_3 = \rho\sqrt{\lambda/2}c_6, \quad (2.40)$$

From, (2.37), (2.39) and (2.40),

$$b_1 = -a_1 \left(1 + \frac{2\sqrt{2}}{\rho\sqrt{\lambda}} \right)$$

and

$$a_1 = a_2 = d_3 = d_4 = - \left(1 + \frac{2\sqrt{2}}{\rho\sqrt{\lambda}} \right)^{-1} \frac{1}{\sqrt{2\lambda}}. \quad (2.41)$$

From (2.35), (2.36), (2.39) and (2.41),

$$g(\lambda, x, y) = \begin{cases} A_\lambda e^{-\sqrt{2\lambda}(|y|+|x|)} + B_\lambda e^{-\sqrt{2\lambda}|y-x|}, & xy \geq 0, \\ (A_\lambda + B_\lambda) e^{-\sqrt{2\lambda}(|x|+|y|)}, & xy < 0, \end{cases}$$

Observing that $|x - y| = |x| + |y|$ for $xy \leq 0$ yields (2.34). \square

Lemma 2.2.13 (see [1], p.1026). *Let $\widehat{\mathcal{L}}_\lambda$ be the inverse Laplace transform with respect to the variable λ . Then, the following Laplace inversion formulas hold,*

$$\begin{aligned} \widehat{\mathcal{L}}_\lambda \left(\frac{e^{-k\sqrt{\lambda}}}{\sqrt{\lambda}} \right) &= \frac{e^{-k^2/4t}}{\sqrt{\pi t}}, \\ \widehat{\mathcal{L}}_\lambda \left(\frac{e^{-k\sqrt{\lambda}}}{\mu + \sqrt{\lambda}} \right) &= \frac{e^{-k^2/4t}}{\sqrt{\pi t}} - \mu e^{\mu k + \mu^2 t} \operatorname{erfc} \left(\mu\sqrt{t} + \frac{k}{2\sqrt{t}} \right). \end{aligned}$$

Proof of Theorem 2.26. From Lemma 2.2.9, the resolvent kernel $[(\lambda, x, y) \mapsto r(\lambda, x, y)]$ of the sticky Brownian motion is (2.34). From (1.34), if $[(t, x, y) \mapsto p(t, x, y)]$ and $[(\lambda, x, y) \mapsto r(\lambda, x, y)]$ are respectively the probability transition and resolvent kernels of a diffusion, then,

$$p(t, x, y) = \widehat{\mathcal{L}}_\lambda r(\lambda, x, y),$$

where $\widehat{\mathcal{L}}_\lambda$ is the inverse Laplace transform. From Lemma 2.2.13,

$$\widehat{\mathcal{L}}_\lambda \left(\frac{1}{\sqrt{2\lambda}} e^{-\sqrt{2\lambda}(|y-x|)} \right) = \frac{1}{\sqrt{2\pi t}} e^{-|x-y|^2/2}$$

and

$$\begin{aligned}\widehat{\mathcal{L}}_\lambda\left(-\frac{1}{4/\rho+\sqrt{2\lambda}}e^{-\sqrt{2\lambda}(|y|+|x|)}\right) &= -\frac{1}{\sqrt{2}}\widehat{\mathcal{L}}_\lambda\left(\frac{1}{4/\rho+\sqrt{\lambda}}e^{-\sqrt{2\lambda}(|y|+|x|)}\right) \\ &= -\frac{1}{\sqrt{2\pi t}}e^{-(|x|+|y|)^2/2t} - \frac{1}{\sqrt{2}}\frac{4}{\rho}e^{4(|y|+|x|)/\rho+8t/\rho^2}\operatorname{erfc}\left(\frac{|y|+|x|}{\sqrt{2t}}+\frac{4\sqrt{t}}{\rho}\right).\end{aligned}$$

Thus, if $p_\rho(t, x, y)$ and $g(\lambda, x, y)$ are the functions defined in (4.10) and (2.34) respectively,

$$\widehat{\mathcal{L}}_\lambda(g(\lambda, x, y)) = p(t, x, y),$$

which proves (4.10). \square

Corollary 2.2.14. *The probability transition kernel of the Brownian motion sticky at $\ell \in \mathbb{R}$ of stickiness ρ with respect to $m(dx) = 2dx + \rho\delta_\ell(dx)$ is the function $p : \mathbb{R}_+ \times \mathbb{R}^2 \mapsto \mathbb{R}_+$ defined for every $t > 0$ and $x, y \in \mathbb{R}$ by*

$$p_\rho^\ell(t, x, y) = u_1(t, x, y) - u_2(t, x - \ell, y - \ell) + v_\rho(t, x - \ell, y - \ell),$$

where

$$\begin{cases} u_1(t, x, y) &= \frac{1}{2\sqrt{2\pi t}}e^{-(x-y)^2/2t}, \\ u_2(t, x, y) &= \frac{1}{2\sqrt{2\pi t}}e^{-(|x|+|y|)^2/2t}, \\ v_\rho(t, x, y) &= \frac{1}{\rho}e^{4(|x|+|y|)/\rho+8t/\rho^2}\operatorname{erfc}\left(\frac{|x|+|y|}{\sqrt{2t}}+\frac{2\sqrt{2t}}{\rho}\right).\end{cases}$$

2.2.4 Asymptotics

Proposition 2.2.15. *Let X^ρ be the sticky Brownian motion of parameter $\rho > 0$, B the standard Brownian motion and B^* the Brownian motion absorbed at 0. Then,*

$$\begin{aligned}\operatorname{Law}(X^\rho) &\longrightarrow \operatorname{Law}(B) && \text{as } \rho \longrightarrow 0, \\ \operatorname{Law}(X^\rho) &\longrightarrow \operatorname{Law}(B^*) && \text{as } \rho \longrightarrow \infty.\end{aligned}$$

Proof. From (4.10), the probability transition kernel of the sticky Brownian motion X^ρ is defined for every $(t, x, y) \in \mathbb{R}_+ \times \mathbb{R}^2$ by

$$p_\rho(t, x, y) = u_1(t, x, y) - u_2(t, x, y) + v_\rho(t, x, y), \quad (2.42)$$

where u_1 , u_2 and v_ρ are defined in (2.26). We observe that

$$\lim_{\rho \rightarrow \infty} v_\rho(t, x, y) = \lim_{\rho \rightarrow \infty} \frac{1}{\rho}e^{4(|x|+|y|)/\rho+8t/\rho^2}\operatorname{erfc}\left(\frac{|x|+|y|}{\sqrt{2t}}+\frac{2\sqrt{2t}}{\rho}\right) = 0. \quad (2.43)$$

Also, from L'Hôpital's rule,

$$\begin{aligned}
\lim_{\rho \rightarrow 0} v_\rho(t, x, y) &= \lim_{\rho \rightarrow 0} \frac{1}{\rho} e^{4(|x|+|y|)/\rho+8t/\rho^2} \operatorname{erfc} \left(\frac{|x|+|y|}{\sqrt{2t}} + \frac{2\sqrt{2t}}{\rho} \right) \\
&= \lim_{\rho \rightarrow 0} \frac{1}{\rho} \frac{\operatorname{erfc} \left(\frac{|x|+|y|}{\sqrt{2t}} + \frac{2\sqrt{2t}}{\rho} \right)}{e^{-4(|x|+|y|)/\rho-8t/\rho^2}} = \lim_{\rho \rightarrow 0} \frac{1}{\rho} \frac{\frac{2}{\sqrt{\pi}} \frac{2\sqrt{2t}}{\rho^2} \exp \left[- \left(\frac{|x|+|y|}{\sqrt{2t}} + \frac{2\sqrt{2t}}{\rho} \right)^2 \right]}{\left(4(|x|+|y|)/\rho^2 + 16t/\rho^3 \right) e^{-4(|x|+|y|)/\rho-8t/\rho^2}} \\
&= \lim_{\rho \rightarrow 0} \frac{1}{\sqrt{\pi}} \frac{4\sqrt{2t}}{\left(4\rho(|x|+|y|) + 16t \right)} \exp \left(- \frac{(|x|+|y|)^2}{2t} \right) \\
&= \frac{1}{2\sqrt{2\pi t}} \exp \left(- \frac{(|x|+|y|)^2}{2t} \right). \quad (2.44)
\end{aligned}$$

From (2.42), (2.43) and (2.44),

$$\begin{aligned}
\lim_{\rho \rightarrow \infty} p_\rho(t, x, y) &= u_1(t, x, y) - u_2(t, x, y), \\
\lim_{\rho \rightarrow 0} p_\rho(t, x, y) &= u_1(t, x, y),
\end{aligned}$$

where (see [14, p.121]) $u_1(t, x, y)$ and $u_1(t, x, y) - u_2(t, x, y)$ are respectively the probability transition kernels of the standard and absorbed at 0 Brownian motions. From Corollary 1.2.15 and Proposition 1.3.1, the result is proven. \square

Chapter 3

Sticky Itô diffusions

The content of this chapter is a part of the article [5].

Chapter Outline: In Section 3.1, we define the notions of sticky stochastic differential equation and sticky Itô diffusion through s and m . In Section 3.2, we prove that any sticky Itô diffusion solves a system of the form (3.8)-(3.9) we call path-wise representation. We also prove that this path-wise representation characterizes the law of these processes. In Section 3.3, we prove path-wise results on sticky SDEs, namely the sticky versions of Itô's lemma and Girsanov's theorem. These results are used in Chapter 4 to prove the local time approximation for sticky Itô diffusions.

For an alternative proof of these results, see [83]. Also, path-wise characterizations are proven in [30] and [76] for the sticky Brownian motion and Ornstein-Uhlenbeck process with sticky reflection respectively.

3.1 Definition

We call sticky Itô diffusions any process that have have a point of stickiness and a dynamic described by a classical stochastic differential equation away from that point.

In Section 1.2.3, we introduced the notions of scale function and speed measure and proved they characterize the law of a diffusion. This characterization has the advantage to be local in nature. As such, this characterization is very convenient for sticky diffusions and especially sticky Itô diffusion.

From (2.9) and Definition 2.1.15, we define this class of processes as follows.

Definition 3.1.1. A *time-homogeneous sticky SDE solution* or *sticky Itô diffusion* is a diffusion process of state-space \mathbb{I} , an interval of \mathbb{R} , defined through s and m , where s and m have the form

$$s(x) = \int_a^x e^{-\int_a^y \frac{2\mu(u)}{\sigma^2(u)} du} dy, \quad m(dx) = \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} dx + \rho \delta_0(dx), \quad (3.1)$$

for every $x \in \mathbb{I}$ with μ and σ being two real-valued measurable functions. For sake of convenience and in order to keep explicit the dependence on (μ, σ, ρ) , we will denote with (μ, σ, ρ) -sticky SDE solution the diffusion process described by (3.1).

We admit the following conditions on μ and σ which prevents explosive, skew and oscillating phenomena. Moreover, (3.3) allows the use of the plain Girsanov theorem.

Condition 3.1.2. *Let B a standard Brownian motion defined on a probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)$. We consider the following SDE:*

$$dX_t = \mu(X_t) dt + \sigma(X_t) dB_t, \quad (3.2)$$

where

- (μ, σ) are taken so that (3.2) has a unique non-explosive strong solution,
- if X be the solution of (3.2) such that $\mathbb{P}(X_0 = x) = 1$, for $\theta = (\sigma'(X_t) - \frac{\mu(X_t)}{\sigma(X_t)})_{t \geq 0}$ and every $x \in \mathbb{I}$,

$$\mathbb{E}_x \left[\exp \left(\int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right) \right] = 1, \quad (3.3)$$

- $\sigma \in C^1(\mathbb{I})$.

Proposition 3.1.3. *The infinitesimal generator of the (μ, σ, ρ) -sticky SDE is*

$$\mathbb{L} f = \mu f' + \frac{1}{2} \sigma^2 f'', \quad (3.4)$$

for every $f \in \text{dom}(\mathbb{L})$, where

$$\begin{aligned} \text{dom}(\mathbb{L}) &= \left\{ f \in C_b(\mathbb{I}) : f \in C_b^2(\mathbb{I} \setminus \{0\}); \right. \\ &\quad \left. \mathbb{L} f(0-) = \mathbb{L} f(0+) = \frac{1}{\rho} (f'(0+) - f'(0-)) \right\}. \end{aligned} \quad (3.5)$$

Proof. From (1.23), for any $f \in \text{dom}(\mathbb{L})$

$$\mathbb{D}_m \mathbb{D}_s f = \mathbb{L} f.$$

From (5.42), for any $x \in \text{Int}(\mathbb{I})$,

$$\mathbb{D}_s f(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{s(x+h) - s(x)} = (f'/s')(x).$$

For $x \neq 0$,

$$\begin{aligned} \mathbb{D}_m \mathbb{D}_s f(x) &= \frac{d}{m(dx)} (f'/s')(x) \\ &= \frac{1}{2} s'(x) \sigma^2(x) \left(f''(x)/s'(x) - s''(x) f'(x)/(s'(x))^2 \right) \\ &= \mu(x) f'(x) + \frac{1}{2} \sigma^2(x) f''(x). \end{aligned} \quad (3.6)$$

For $x = 0$,

$$\begin{aligned} \mathbb{D}_m \mathbb{D}_s f(0) &= \lim_{h \rightarrow 0} \frac{\mathbb{D}_s f(h) - \mathbb{D}_s f(0)}{m((0, h])} \\ &= \lim_{h \rightarrow 0} \frac{\mathbb{D}_s f(h) - \mathbb{D}_s f(0)}{\int_0^h 2/(s'(\zeta) \sigma^2(\zeta)) d\zeta + \rho} = \frac{1}{\rho} \left(\mathbb{D}_s f(0+) - \mathbb{D}_s f(0-) \right). \\ &= \frac{1}{\rho} \frac{1}{s'(0)} \left(f'(0+) - f'(0-) \right) = \frac{1}{\rho} \left(f'(0+) - f'(0-) \right), \end{aligned} \quad (3.7)$$

where the last equality results from the definition of the scale function (see Proposition 1.2.16). From (1.12), (3.6) and (3.7), we have proven (3.4) and (3.5). \square

3.2 Path-wise characterization

Theorem 3.2.1. *We consider the following system*

$$dX_t = \mu(X_t)\mathbb{1}_{X_t \neq 0} dt + \sigma(X_t)\mathbb{1}_{X_t \neq 0} dB_t, \quad (3.8)$$

$$\mathbb{1}_{X_t=0} dt = \frac{\rho}{2} dL_t^0(X), \quad (3.9)$$

where B is a standard Brownian motion, $L^0(X)$ is the local time process of X at 0 and (μ, σ) are a pair of real-valued functions over \mathbb{I} that satisfy Condition (3.1.2). Then, the system (3.8)-(3.9) has a jointly unique weak solution.

Proof. Let $(\Omega, (\mathcal{F})_{t \geq 0}, \mathbb{P}_x)$ be a probability space, B^1 a Brownian motion defined on $(\Omega, (\mathcal{F})_{t \geq 0}, \mathbb{P}_x)$ and Y be the strong solution of

$$dY_t = \mu(Y_t) dt + \sigma(Y_t) dB_t^1, \quad (3.10)$$

such that

$$\mathbb{P}_x(Y_0 = 0) = 1. \quad (3.11)$$

Let A and its right-inverse γ be the continuous and strictly increasing time-transforms defined for every $t \geq 0$ by

$$A(t) = t + \frac{\rho}{2} L_t^0(Y), \quad \gamma(t) = \inf \{s > 0 : A(s) > t\}.$$

Let X and B be the processes defined for every $t \geq 0$ by

$$X_t = Y_{\gamma(t)}, \quad (3.12)$$

$$B_t = B_{\gamma(t)}^1 + \int_0^t \mathbb{1}_{X_s=0} dB_s^0, \quad (3.13)$$

where B^0 is a Brownian motion independent of B^1 . From (A.3), for every $t \geq 0$,

$$B_{\gamma(t)}^1 = \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} dB_s^1 = \int_0^t \mathbb{1}_{X_s \neq 0} dB_{\gamma(s)}^1$$

and

$$\begin{aligned} \gamma(t) &= \langle B_{\gamma(\cdot)}^1 \rangle_t = \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} ds \\ &= \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} \left(ds + \frac{\rho}{2} dL_s^0(Y) \right) = \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} dA(s) = \int_0^t \mathbb{1}_{X_s \neq 0} ds. \end{aligned} \quad (3.14)$$

Thus,

$$\langle B \rangle_t = \left\langle B_{\gamma(\cdot)}^1 + \int_0^{\cdot} \mathbb{1}_{X_s=0} dB_s^0 \right\rangle_t = \int_0^t \mathbb{1}_{X_s \neq 0} ds + \int_0^t \mathbb{1}_{X_s=0} ds = t. \quad (3.15)$$

From Lévy's characterization (see [62, Chapter 3; Theorem 3.16]), B is a standard Brownian motion. From (A.3),

$$\begin{aligned}
X_t = Y_{\gamma(t)} &= \int_0^{\gamma(t)} dY_s = \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} dY_s \\
&= \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} \mu(Y_s) ds + \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} \sigma(Y_s) dB_s^1 \\
&= \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} \mu(Y_s) \left(ds + \frac{\rho}{2} dL_s^0(Y) \right) + \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} \sigma(Y_s) dB_s^1 \\
&= \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} \mu(Y_s) dA(s) + \int_0^{\gamma(t)} \mathbb{1}_{Y_s \neq 0} \sigma(Y_s) dB_s^1 \\
&= \int_0^t \mathbb{1}_{X_s \neq 0} \mu(X_s) ds + \int_0^t \mathbb{1}_{X_s \neq 0} \sigma(X_s) dB_{\gamma(s)}^1 \\
&= \int_0^t \mathbb{1}_{X_s \neq 0} \mu(X_s) ds + \int_0^t \mathbb{1}_{X_s \neq 0} \sigma(X_s) \left\{ dB_{\gamma(s)}^1 + \mathbb{1}_{X_s=0} dB_s^0 \right\} \\
&= \int_0^t \mathbb{1}_{X_s \neq 0} \mu(X_s) ds + \int_0^t \mathbb{1}_{X_s \neq 0} \sigma(X_s) dB_s. \quad (3.16)
\end{aligned}$$

Moreover,

$$\begin{aligned}
\int_0^t \mathbb{1}_{X_s=0} ds &= \int_0^t \mathbb{1}_{Y_{\gamma(s)}=0} dA(\gamma(s)) = \int_0^{\gamma(t)} \mathbb{1}_{Y_s=0} dA(s) \\
&= \int_0^{\gamma(t)} \mathbb{1}_{Y_s=0} \left(ds + \frac{\rho}{2} dL_s^0(Y) \right) = \frac{\rho}{2} \int_0^{\gamma(t)} \mathbb{1}_{Y_s=0} dL_s^0(Y) = \frac{\rho}{2} L_{\gamma(t)}^0(Y) = \frac{\rho}{2} L_t^0(X). \quad (3.17)
\end{aligned}$$

From (3.11) and (3.12), $P_x(X_0 = x) = 1$. From (3.16) and (3.17), the pair $(X_t, B_t)_{t \geq 0}$ solves (3.8)-(3.9), proving the existence of a solution.

For the uniqueness, we reset all notations. Let $(\Omega, (\mathcal{F}_t)_{t \geq 0}, P_x)$ be a probability space, $(\tilde{X}_t, \tilde{B}_t)_{t \geq 0}$ a solution of (3.8)-(3.9) such that $P_x(X_0 = x) = 1$ and \tilde{A} being the right-inverse of $\tilde{\gamma}$ which is the time-transform defined for every $t \geq 0$ by

$$\tilde{\gamma}(t) = \int_0^t \mathbb{1}_{\tilde{X}_s \neq 0} ds. \quad (3.18)$$

Let \tilde{Y} be the process defined for every $t \geq 0$ by

$$\tilde{Y}_t = \tilde{X}_{\tilde{A}(t)}. \quad (3.19)$$

From the definition of \tilde{Y} , $P_x(Y_0 = x) = 1$. From (3.9), if $L^0(\tilde{X})$ and $L^0(\tilde{Y})$ are the local times at 0 of \tilde{X} and \tilde{Y} respectively,

$$\begin{aligned}
t = \tilde{\gamma}(\tilde{A}(t)) &= \int_0^{\tilde{A}(t)} \mathbb{1}_{\tilde{X}_s \neq 0} ds = \tilde{A}(t) - \int_0^{\tilde{A}(t)} \mathbb{1}_{\tilde{X}_s=0} ds \\
&= \tilde{A}(t) - \frac{\rho}{2} L_{\tilde{A}(t)}^0(\tilde{X}) = \tilde{A}(t) - \frac{\rho}{2} L_t^0(\tilde{Y}). \quad (3.20)
\end{aligned}$$

The time-transform \tilde{A} is continuous and strictly increasing, thus $\tilde{\gamma}$ is its proper inverse.

From (3.8),

$$\begin{aligned}
\tilde{Y}_t = \tilde{X}_{\tilde{A}(t)} &= \tilde{X}_0 + \int_0^{\tilde{A}(t)} \mu(\tilde{X}_s) \mathbb{1}_{\tilde{X}_s \neq 0} ds + \int_0^{\tilde{A}(t)} \sigma(\tilde{X}_s) \mathbb{1}_{\tilde{X}_s \neq 0} d\tilde{B}_s \\
&= \tilde{X}_0 + \int_0^{\tilde{A}(t)} \mu(\tilde{X}_s) d\tilde{\gamma}(s) + \int_0^{\tilde{A}(t)} \sigma(\tilde{X}_s) \mathbb{1}_{\tilde{X}_s \neq 0} dB_s \\
&= \tilde{X}_0 + \int_0^t \mu(\tilde{Y}_s) ds + \int_0^{\tilde{A}(t)} \sigma(\tilde{X}_s) \mathbb{1}_{\tilde{X}_s \neq 0} d\tilde{B}_s.
\end{aligned}$$

Let $\tilde{B}_t^1 = \int_0^{\tilde{A}(t)} \mathbb{1}_{\tilde{X}_s \neq 0} d\tilde{B}_s$ where,

$$\langle \tilde{B}^1 \rangle_t = \int_0^{\tilde{A}(t)} \mathbb{1}_{\tilde{X}_s \neq 0} ds = \tilde{\gamma}(\tilde{A}(t)) = t.$$

Thus, from Levy's characterization \tilde{B}^1 is a standard Brownian motion and as $d\tilde{B}_t^1 = \mathbb{1}_{\tilde{Y}_s \neq 0} d\tilde{B}_{\tilde{A}(t)}$,

$$\tilde{Y}_t = \tilde{X}_0 + \int_0^t \mu(\tilde{Y}_s) ds + \int_0^t \sigma(\tilde{Y}_s) \mathbb{1}_{\tilde{Y}_s \neq 0} d\tilde{B}_{\tilde{A}(s)} = \tilde{X}_0 + \int_0^t \mu(\tilde{Y}_s) ds + \int_0^t \sigma(\tilde{Y}_s) d\tilde{B}_s^1.$$

Let (Y, \tilde{B}^1) be the solution of (3.10) such that $P_x(Y_0 = x) = 1$, γ the right-inverse of $A(t) = t + \frac{\rho}{2} L_t^0(Y)$ and X the process such that $X_t = Y_{\gamma(t)}$ for every $t \geq 0$. From (3.14) and (3.18)

$$\gamma(t) = \tilde{\gamma}(t) \quad \text{and} \quad A(t) = \tilde{A}(t), \quad (3.21)$$

for every $t \geq 0$. From (3.19) and (3.21) and as $P(X_0 = Y_0 = \tilde{Y}_0 = x) = 1$, almost surely

$$Y_{\gamma(t)} = X_t = \tilde{Y}_{\tilde{\gamma}(t)}, \quad (3.22)$$

for every $t \geq 0$. Hence X is uniquely determined by $(B_t^1, L_t^0(Y))_{t \geq 0}$ and (3.8)-(3.9) has a unique solution. \square

Proposition 3.2.2. *If (X, B) is the joint solution of (3.8)-(3.9), then X is the (ρ, μ, σ) -sticky SDE solution.*

Proof. Let X be the solution of (3.8)-(3.9) and γ the time transform

$$\gamma(t) = \int_0^t \mathbb{1}_{X_s \neq 0} ds.$$

Moreover, let A be the right-inverse of γ and Y the process such that for every $t \geq 0$,

$$Y_t = X_{A(t)}. \quad (3.23)$$

From (3.10), (3.18), (3.21) and (3.22), the process Y solves (3.10). From Proposition 2.6 of [80, Chapter VII] and of [14, p.17], Y is the diffusion process defined through s and m where

$$s'(x) = e^{-\int_a^x \frac{2\mu(u)}{\sigma^2(u)} du}, \quad m(dx) = \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} dx. \quad (3.24)$$

Then, $s(Y) = (s(Y_t))_{t \geq 0}$ is a diffusion process on natural scale with speed measure $m_{s(Y)}(dx) = m_Y(s^{-1}(dx))$. From Theorem 1.4.4, there exists a Brownian motion W such that,

$$s(Y_t) = W_{\gamma_{s(Y)}(t)}, \quad \text{for every } t \geq 0, \quad (3.25)$$

where $\gamma_{s(Y)}(t)$ is the right-inverse of $A_{s(Y)}(t) = \int_{\mathbb{R}} L_t^y(W) m_{s(Y)}(dy)$. From (3.23) and (3.25),

$$s(X_t) = W_{\gamma_{s(Y)}(\gamma(t))}, \quad \text{for every } t \geq 0.$$

From (3.20), the right-inverse of $\gamma_{s(Y)}(\gamma(t))$ is

$$A(A_{s(Y)}(t)) = A_{s(Y)}(t) + \frac{\rho}{2} L_{A_{s(Y)}(t)}^0(Y) = A_{s(Y)}(t) + \frac{\rho}{2} L_t^0(W) = \frac{1}{2} \int_{\mathbb{R}} L_t^y(W) \nu(dy),$$

where $\nu(dx) = m_{s(Y)}(dx) + \rho\delta_0(dx)$. Thus, from Corollary 1.4.5, $(s(X_t))_{t \geq 0}$ is a diffusion process on natural scale of speed measure $m(s^{-1}(dx)) + \rho\delta_0(dx)$. As s is continuous and invertible, X is a diffusion process with scale function s_X and speed measure m_X where

$$s_X(x) = s(x), \quad m_X(dx) = m(dx) + \rho\delta_0(dx) \quad (3.26)$$

or a (ρ, μ, σ) -sticky SDE solution. \square

Proposition 3.2.3. *Let X be a (ρ, μ, σ) -sticky SDE defined on a probability space $(\Omega, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}_x)$. Then, there exists a Brownian motion W such that under \mathbb{P}_x ,*

$$X_t = x + \int_0^t \mu(X_s) \mathbb{1}_{X_s \neq 0} ds + \int_0^t \sigma(X_s) \mathbb{1}_{X_s \neq 0} dW_s, \quad (3.27)$$

$$\int_0^t \mathbb{1}_{X_s=0} ds = \frac{\rho}{2} L_t^0(X). \quad (3.28)$$

Proof. Let A be the right-inverse of $\gamma = [t \mapsto \int_0^t \mathbb{1}_{X_s \neq 0} ds]$ and Y the process such that,

$$Y_t = X_{A(t)},$$

for every $t \geq 0$. Let (s_X, m_X) and (s_Y, m_Y) be the scale function and speed measure pairs of X and Y respectively. From (3.24) and (3.26),

$$s_Y = s_X, \quad m_Y = m_X - \rho\delta_0. \quad (3.29)$$

From (5.42) and (3.29),

$$s_Y'(x) = e^{-\int_a^x \frac{2\mu(u)}{\sigma^2(u)} du}, \quad m_Y(dx) = \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} dx.$$

Thus, from Theorem 3.12 of [80, Chapter VII], the infinitesimal generator L_Y of Y is

$$L_Y f(x) = \mu(x) f'(x) + \frac{1}{2} \sigma^2(x) f''(x),$$

for every $x \in \mathbb{I}$ and $f \in \text{dom}(L_Y)$ where $\text{dom}(L_Y) = C^2(\mathbb{I})$. From Theorem 2.7 of [80, Chapter VII], there exists a Brownian motion W^1 on an extension of $(\Omega, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}_x)$ such that Y almost surely solves,

$$dY_t = \mu(Y_t) dt + \sigma(Y_t) dW_t^1,$$

for every $t \geq 0$. Let W be the process such that

$$W_t = W_{\gamma(t)}^1 + \int_0^t \mathbb{1}_{X_s=0} dW_s^0.$$

From (3.13) and (3.15) and Lévy's characterization, W is a standard Brownian motion. From (3.10), (3.18), (3.21) and (3.22), (X, W) jointly solve (3.27)-(3.28). \square

Corollary 3.2.4. *A sticky SDE solution is a semi-martingale.*

Proof. Let X be a (μ, σ, ρ) -sticky Itô diffusion. From Proposition 3.2.3, the process X solves (3.27)-(3.28). The relation (3.27) is an explicit Doob-Meyer representation of X . \square

3.3 Path-wise results

Proposition 3.3.1 (Sticky Girsanov). *Let $(\Omega, (\mathcal{F})_{t \geq 0}, \mathbb{P})$ be a probability space and X the process that solves*

$$dX_t = \mu(X_t) \mathbb{1}_{X_t \neq 0} dt + \sigma(X_t) \mathbb{1}_{X_t \neq 0} dB_t, \quad (3.30)$$

$$\mathbb{1}_{X_t=0} dt = \frac{\rho}{2} dL_t^0(X), \quad (3.31)$$

where B is a \mathbb{P} -Brownian motion. Let θ be a processes such that $\mathbb{P}(\int_0^T \theta_s ds < \infty) = 1$, $\mathcal{E}(\theta)$ the process such that

$$\mathcal{E}_t(\theta) = \exp \left(\int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right)$$

for every $t \geq 0$ and \mathbb{Q} the probability measure such that $d\mathbb{Q} = \mathcal{E}_t(\theta) d\mathbb{P}$. Then, if $\mathbb{E}_{\mathbb{P}}$ is the expectancy under \mathbb{P} and $\mathbb{E}_{\mathbb{P}}(\mathcal{E}_t(\theta)) = 1$, the process X solves

$$d\tilde{X}_t = \left(\mu(\tilde{X}_t) + \theta_t \sigma(\tilde{X}_t) \right) \mathbb{1}_{\tilde{X}_t \neq 0} dt + \sigma(\tilde{X}_t) \mathbb{1}_{\tilde{X}_t \neq 0} d\tilde{B}_t, \quad (3.32)$$

$$\mathbb{1}_{\tilde{X}_t=0} dt = \frac{\rho}{2} dL_t^0(\tilde{X}), \quad (3.33)$$

where $\tilde{B}_t = B_t - \int_0^t \theta_s ds$ is a standard Brownian motion under \mathbb{Q} .

Proof. Let X be the the solution of (3.30)-(3.31), γ the time-change $\gamma(t) = \int_0^t \mathbb{1}_{X_s \neq 0} ds$ for every $t \geq 0$, A its right-inverse and $Y = (X_{A(t)})_{t \geq 0}$. Let \tilde{B} be the process defined by $\tilde{B}_t = B_t - \int_0^t \theta_s ds$ for every $t \geq 0$. Then, from Theorem 6.3 of [70], \tilde{B} is a standard Brownian motion under \mathbb{Q} and the probability measures \mathbb{P} and \mathbb{Q} are equivalent. By substitution,

$$dX_t = \left(\mu(X_t) + \theta_t \sigma(X_t) \right) \mathbb{1}_{X_t \neq 0} dt + \sigma(X_t) \mathbb{1}_{X_t \neq 0} d\tilde{B}_t.$$

Moreover, since: the local time $L_t^0(X)$ and the quadratic variation $\langle X \rangle_t$ are defined as limits in probability, $\mathbb{P} \sim \mathbb{Q}$ and

$$\langle X \rangle_t = \int_0^t \mathbb{1}_{X_s \neq 0} ds = t - \int_0^t \mathbb{1}_{X_s=0} ds$$

Thus, (3.31) holds also under \mathbb{Q} and X solves (3.32)-(3.33). \square

Proposition 3.3.2 (Sticky Itô formula). *Let X be a process defined on a probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)$ such that $\mathbb{P}_x(X_0 = x) = 1$ and*

$$\begin{aligned} dX_t &= \mu(X_t) \mathbb{1}_{X_t \neq 0} dt + \sigma(X_t) \mathbb{1}_{X_t \neq 0} dB_t, \\ \mathbb{1}_{X_t=0} dt &= \frac{\rho}{2} dL_t^0(X), \end{aligned} \quad (3.34)$$

where B is a $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)$ -standard Brownian motion. Then, for every real valued C^2 function f such that $f(0) = 0$ and $f'(0) \neq 0$, the process $f(X) = (f(X_t))_{t \geq 0}$ is solution of

$$\begin{aligned} df(X_t) &= \left(f'(X_t)\mu(X_t)\frac{1}{2}f''(X_t)\sigma^2(X_t) \right) \mathbb{1}_{X_t \neq 0} dt + f'(X_t)\sigma(X_t)\mathbb{1}_{X_t \neq 0} dB_t \quad (3.35) \\ \mathbb{1}_{f(X_t)=0} dt &= f'(0)\frac{\rho}{2} dL_t^0(f(X)) \quad (3.36) \end{aligned}$$

and $\mathbb{P}_x(f(X_0) = f(x)) = 1$.

Proof. The process X is a semi-martingale as,

$$X_t = X_0 + \int_0^t \mu(X_s)\mathbb{1}_{X_s \neq 0} ds + \int_0^t \sigma(X_s)\mathbb{1}_{X_s \neq 0} dB_s,$$

where $\int_0^t \mu(X_s)\mathbb{1}_{X_s \neq 0} ds$ is a process of bounded variation and $\int_0^t \sigma(X_s)\mathbb{1}_{X_s \neq 0} dB_s$ is a local martingale. Thus, we may apply the standard Itô formula for $f \in C^2(\mathbb{R})$,

$$\begin{aligned} df(X_t) &= f'(X_t) dX_t + \frac{1}{2}f''(X_t) d\langle X \rangle_t = \\ &= \left(f'(X_t)\mu(X_t) + \frac{1}{2}f''(X_t)\sigma^2(X_t) \right) \mathbb{1}_{X_t \neq 0} dt + f'(X_t)\sigma(X_t)\mathbb{1}_{X_t \neq 0} dB_t, \end{aligned}$$

thus proving (3.35). Lemma A.2.3 and (3.34) yield (3.36), thus the result is proven. \square

Chapter 4

Local time approximation of sticky diffusions

In this chapter we present the theoretical results of [5]. Numerical illustrations are given in Chapter 6.

4.1 Introduction and main results

The last decades we have seen the appearance of local time approximations and their applications in statistical estimation problems. In [55], it is proven that the high-frequency statistic of an Itô diffusion X

$$\frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g(n^\alpha X_{\frac{i-1}{n}} - \ell) \quad (4.1)$$

for a given point ℓ , converges as $n \rightarrow \infty$ uniformly in time, in probability to $\lambda(g)L_t^\ell(X)$. Since then, similar results were proven in the case of the skew and the oscillating Brownian motions [68, 72] and the fractional Brownian motion [79, 57]. These gave rise to the usage of local time statistics [67, 68, 69], *i.e.* statistics based on these approximations. In this chapter we prove that under some assumptions on the test-function g , the same results holds for Itô diffusions with a sticky point at 0. Then, we use this result to define a local time statistic which we prove to be a consistent estimator of the stickiness parameter.

Introduced by Feller in [33], one-dimensional sticky diffusions are continuous processes that satisfy the strong Markov property and spend positive amount of time at some points of their state-space \mathbb{I} . These points, called sticky points, can be located either in the interior or at an attainable boundary the boundary of \mathbb{I} (sticky reflection). A diffusion has a sticky point at 0 iff its speed measure m has an atom at 0, *i.e.* $m(\{0\}) > 0$. The mass of that atom $\rho = m(\{0\})$ is called stickiness parameter, it expresses how much time the process spends at 0 and there are no references known to us for its estimation. The infinitesimal generator of these processes are also known to have Wencel boundary conditions at the points of stickiness.

Sticky processes have been recently used to model phenomena in finance, biology, quantum and classical mechanics. In particular, they can be used to describe the

behavior of interest rates around 0 [76, 58], the behavior of molecules near a membrane [44], the concentration of pathogens in a healthy individual [20], the dynamics of mesoscale particles upon contact in colloids [86, 61] and the motion of quantum particles when they reach a source of emission [24]. From a theoretical standpoint, they are used to create new types of probabilistic couplings [28] and appear as the limit of storage processes [48]. Many papers have appeared recently that address the numerical challenges of simulating sticky diffusions [4, 8, 73, 16, 6].

In this chapter, we prove a local time approximation result for a class of sticky processes called sticky Itô diffusions. These processes solve a homogeneous SDE away from a countable and isolated set of points in the state-space where they exhibit stickiness. For simplicity, we suppose that there is one unique sticky point located at $0 \in \mathbb{I}$. Thus, the dynamic of the process is fully described by the drift and diffusivity functions μ, σ that describe the SDE the process solves away from 0, and the stickiness parameter $\rho > 0$ at 0 (see Chapter 3). To abbreviate, we call this process the (ρ, μ, σ) -SID. The most elementary sticky Itô diffusion is the $(\rho, 0, 1)$ -SID called the sticky Brownian motion (see [30]). It is the process that has a Brownian dynamic away from 0 and a sticky point at 0.

We will further suppose that the functions μ and $\sigma : \mathbb{I} \mapsto \mathbb{R}$ that govern the dynamic of the underlying SDE satisfy the following condition:

Condition 4.1.1. *Let B a standard Brownian motion defined on a probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)$. We consider the following SDE:*

$$dX_t = \mu(X_t) dt + \sigma(X_t) dB_t, \quad (4.2)$$

where

- (μ, σ) are taken so that (4.2) has a unique non-explosive strong solution,
- if X be the solution of (4.2) such that $\mathbb{P}(X_0 = x) = 1$, then, for $\theta = (\sigma'(X_t) - \frac{\mu(X_t)}{\sigma(X_t)})_{t \geq 0}$ and every $x \in \mathbb{I}$,

$$\mathbb{E}_x \left[\exp \left(\int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right) \right] = 1,$$

- $\sigma \in C^1(\mathbb{I})$.

A sticky Itô diffusion is a semimartingale (see Corollary 3.2.4). As such, if X is a sticky Itô diffusion defined on a probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, its local time at ℓ can either be defined (see [80, Chapter VI, §1])

- as the continuous, strictly-increasing process $L^\ell(X)$ such that for every $t \geq 0$,

$$|X_t - \ell| - |X_0 - \ell| = \int_0^t \text{sgn}(X_s - \ell) dX_s + L_t^\ell(X),$$

where $\text{sgn}(x) = \mathbb{1}_{x \geq 0} - \mathbb{1}_{x < 0}$,

- for every $t \geq 0$, as the limit in probability

$$L_t^\ell(X) = \mathbb{P}\text{-}\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^t \mathbb{1}_{0 \leq X_s - \ell < \epsilon} d\langle X \rangle_s, \quad (4.3)$$

- if the functions (μ, σ) satisfy Condition 3.1.2, for every $t \geq 0$, as the limit in probability

$$L_t^\ell(X) = \text{P-lim}_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_0^t \mathbb{1}_{|X_s - \ell| < \epsilon} d\langle X \rangle_s. \quad (4.4)$$

The object of this paper is to prove necessary conditions for the statistic (4.1) to converge to the local time of a sticky Itô diffusions. The convergence occurs locally uniformly in time, in probability.

Definition 4.1.2. A sequence of processes X^n is said to converge locally uniformly in time, in P-probability to a process X iff

$$\sup_{s \leq t} \{|X_s^n - X_s|\} \xrightarrow{n \rightarrow \infty} 0,$$

in P-probability, for every $t \geq 0$

In particular, we prove the following results:

Theorem 4.1.3. Let X^ρ be a (ρ, μ, σ) -SID defined on a probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \text{P}_x)$ such that $\text{P}_x(X_0 = x) = 1$ for any x in the state-space \mathbb{I} of X and (μ, σ) satisfy Condition 3.1.2. Let be $g : \mathbb{R} \mapsto \mathbb{R}$ be a bounded Lebesgue-integrable function which vanishes on an open interval around 0 and $T : \mathbb{R} \mapsto \mathbb{R}$ a continuously differentiable function such that for an $\epsilon > 0$:

$$T(0) = 0, \quad T'(0) = 1, \quad \epsilon \leq T'(x) \leq 1/\epsilon, \quad |T''(x)| \leq 1/\epsilon, \quad (4.5)$$

for every $x \in \mathbb{R}$. For every such function T , let $g_n[T]$ be the sequence of functions such that:

$$g_n[T](x) = g(n^\alpha T(n^{-\alpha}x)), \quad (4.6)$$

for every x and n . Then, for every $\alpha \in (0, 1/2)$,

$$\frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g_n[T](n^\alpha X_{\frac{i-1}{n}}^\rho) \xrightarrow{n \rightarrow \infty} \frac{\lambda(g)}{\sigma(0)} L_t^0(X^\rho). \quad (4.7)$$

locally uniformly in time, in P_x -probability.

We now apply our result to give a consistent estimator of the stickiness parameter. Numerical illustrations are given in Section 6.2.

Corollary 4.1.4 (of Theorem 4.1.3 and Lemma 4.3.1). *In the same setting as Theorem 4.1.3, if g is a bounded integrable function which vanishes on an open interval around 0, T that satisfies (4.5), $g_n[T]$ defined in (4.6) and $\alpha \in (0, 1/2)$, then,*

$$\hat{\rho}_n(X) := 2 \frac{\lambda(g)}{\sigma(0)} \frac{1}{n^\alpha} \frac{\sum_{i=1}^{[nt]} \mathbb{1}_{X_{\frac{i-1}{n}} = 0}}{\sum_{i=1}^{[nt]} g_n[T](n^\alpha X_{\frac{i-1}{n}})} \quad (4.8)$$

is a consistent estimator of ρ as $n \rightarrow \infty$.

Remark. In particular both results hold if we take T the identity function and thus replace $g_n[T]$ with g . For example, these holds for the function $g = [x \mapsto \mathbb{1}_{1 < |x| < 2}]$.

The first two results gives us the convergence properties of the statistic at and away from the threshold of stickiness. The last result results in a consistent stickiness estimator, which is the first attempt to estimate the stickiness parameter of a diffusion.

4.2 The case of the sticky Brownian motion

This section is dedicated to the proof of the following result:

Theorem 4.2.1. *Theorem 4.1.3 holds for the sticky Brownian motion.*

As the proof of Theorem 4.2.1 is rather tedious, we have isolated parts of it in Lemmas 4.2.3, 4.2.4 and 4.2.5.

4.2.1 Preliminary results

Lemma 4.2.2 (semi-group bound). *Let $(P_t^\rho)_{t \geq 0}$ be the semi-group of the sticky Brownian motion of parameter $\rho > 0$. There exists a constant $K > 0$ that does not depend on ρ such that for every real-valued function $h(x)$ such that $h(0) = 0$,*

$$|P_t^\rho h(x)| \leq K \frac{\lambda(|h|)}{\sqrt{t}}, \quad (4.9)$$

for every $t > 0$, where $\lambda(g) = \int_{\mathbb{R}} g(x) dx$.

Proof. Let $p_\rho(t, x, y)$ be the probability transition kernel of the sticky Brownian motion of parameter $\rho > 0$ with respect to its speed measure $m(dy) = 2 dy + \frac{\rho}{2} \delta_0(dy)$. From [14, p. 108],

$$p_\rho(t, x, y) = u_1(t, x, y) - u_2(t, x, y) + v_\rho(t, x, y), \quad (4.10)$$

for every $x, y \in \mathbb{R}$ and $t > 0$, where

$$\begin{cases} u_1(t, x, y) &= \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t}, \\ u_2(t, x, y) &= \frac{1}{\sqrt{2\pi t}} e^{-(|x|+|y|)^2/2t}, \\ v_\rho(t, x, y) &= \frac{2}{\rho} e^{4(|x|+|y|)/\rho + 8t/\rho^2} \operatorname{erfc}\left(\frac{|x|+|y|}{\sqrt{2t}} + \frac{2\sqrt{2t}}{\rho}\right). \end{cases}$$

We observe that for $x, y \in \mathbb{R}$ and $t \geq 0$: $\exp(-(|x|+|y|)^2/2t) < \exp(-(x-y)^2/2t)$ and

$$u_1(t, x, y) - u_2(t, x, y) \leq \frac{1}{2\sqrt{2\pi t}} e^{-(x-y)^2/2t}. \quad (4.11)$$

The Mills ratio of a Gaussian random variable (see [46, p. 98]) yields $\operatorname{erfc}(x) \sim e^{-x^2}/x$. Thus, there exists a constant $K_{\text{Mills}} > 0$ such that

$$v_\rho(t, x, y) \leq K_{\text{Mills}} \frac{2\sqrt{2t}}{\rho(|x|+|y|)+8t} e^{-(|x|+|y|)^2/2t} \leq K_{\text{Mills}} \frac{1}{2\sqrt{2t}} e^{-(x-y)^2/2t}. \quad (4.12)$$

From (4.11) and (4.12), for $K = 1 + K_{\text{Mills}}\sqrt{\pi}/2$,

$$p_\rho(t, x, y) \leq K \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t} \quad (4.13)$$

and

$$|P_t h(x)| \leq \int_{\mathbb{R}} |h(y)| p_\rho(t, x, y) dy \leq K \int_{\mathbb{R}} |h(y)| \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t} dy.$$

Observing $e^{-(x-y)^2/2t} \leq 1$ yields (4.9). \square

Lemma 4.2.3. *Let X^ρ be the sticky Brownian motion of stickiness $\rho > 0$ and g be an integrable function such that $g(0) = 0$. Then, there exists a constant $K > 0$ such that for every $x \in \mathbb{R}$, $t > 0$ and $n \in \mathbb{N}$:*

$$\mathbb{E}_x \left(\sup_{s \leq t} \left| \frac{u_n}{n} \sum_{i=1}^{\lfloor ns \rfloor} g(u_n X_{\frac{i-1}{n}}^\rho) \right| \right) \leq K \left(\frac{u_n}{n} |g(u_n x)| + \lambda(|g|) \sqrt{t} \right). \quad (4.14)$$

Proof. We observe that

$$\begin{aligned} \mathbb{E}_x \left(\sup_{s \leq t} \left| \frac{u_n}{n} \sum_{i=1}^{\lfloor ns \rfloor} g(u_n X_{\frac{i-1}{n}}^\rho) \right| \right) &\leq \mathbb{E}_x \left(\frac{u_n}{n} \sum_{i=1}^{\lfloor nt \rfloor} \left| g(u_n X_{\frac{i-1}{n}}^\rho) \right| \right) \\ &= \frac{u_n}{n} |g(u_n x)| + \frac{u_n}{n} \sum_{i=2}^{\lfloor nt \rfloor} \mathbb{E}_x \left(\left| g(u_n X_{\frac{i-1}{n}}^\rho) \right| \right). \end{aligned} \quad (4.15)$$

Then, if $h_n(x) = g(u_n x / \sqrt{n})$,

$$\sum_{i=1}^{\lfloor nt \rfloor} h_n(\sqrt{n} X_{\frac{i-1}{n}}^\rho) = \sum_{i=1}^{\lfloor nt \rfloor} g(u_n X_{\frac{i-1}{n}}^\rho), \quad \lambda(|h_n|) = \frac{\sqrt{n}}{u_n} \lambda(|g|), \quad h_n(0) = 0. \quad (4.16)$$

From (4.15) and (4.16),

$$\mathbb{E}_x \left(\sup_{s \leq t} \left| \frac{u_n}{n} \sum_{i=1}^{\lfloor ns \rfloor} g(u_n X_{\frac{i-1}{n}}^\rho) \right| \right) \leq \frac{u_n}{n} |g(u_n x)| + \frac{u_n}{n} \sum_{i=2}^{\lfloor nt \rfloor} \mathbb{E}_x \left(\left| h_n(\sqrt{n} X_{\frac{i-1}{n}}^\rho) \right| \right). \quad (4.17)$$

As $h_n(0) = 0$, from (2.24) and (4.9),

$$\mathbb{E}_x \left(\left| h_n(\sqrt{n} X_{\frac{i-1}{n}}^\rho) \right| \right) = P_{i-1}^{\rho \sqrt{n}} |h_n(x \sqrt{n})| \leq K \frac{\lambda(|h_n|)}{\sqrt{i-1}}. \quad (4.18)$$

From (4.16), (4.18) and as $\sum_{i=1}^{\lfloor nt \rfloor} \frac{1}{\sqrt{i}} \leq 2\sqrt{nt}$,

$$\sum_{i=2}^{\lfloor nt \rfloor} \mathbb{E}_x \left(\left| h_n(\sqrt{n} X_{\frac{i-1}{n}}^\rho) \right| \right) \leq 2K \lambda(|h_n|) \sqrt{nt} = 2K \frac{n}{u_n} \lambda(|g|) \sqrt{t}. \quad (4.19)$$

From (4.17) and (4.19), we get (4.14). \square

4.2.2 Proof in the sticky Brownian case

Lemma 4.2.4. *Let X be a sticky Brownian motion with local time $L^a(X)$. Moreover, let g and T be two real-valued functions such that g is bounded and integrable and T satisfies (4.5). Then, for any $t \geq 0$,*

$$\int_{\mathbb{R}} g_n(x) L_{\lfloor nt \rfloor / n}^{x/n^\alpha}(X) dx \xrightarrow{n \rightarrow \infty} \lambda(g) L_t^0(X), \quad (4.20)$$

where g_n is given by (4.6).

Proof. From Trotter's theorem [89], the local time of the standard Brownian motion $L^x(B)$ admits a version that is (t, x) -jointly continuous. As the time-change γ also admits a continuous version, from (A.4), the local time of X $L^x(X)$ admits a version that is (t, x) -jointly continuous. Thus,

$$|L_{[nt]/n}^{x/n^\alpha}(X) - L_t^0(X)| \xrightarrow{n \rightarrow \infty} 0, \quad (4.21)$$

for every $t \geq 0$ and $x \in \mathbb{R}$. Moreover, there exists a positive random variable U such that

$$|L_{[nt]/n}^{x/n^\alpha}(X) - L_t^0(X)| \leq U, \quad (4.22)$$

for every x and n . Thus, as g is bounded,

$$\begin{aligned} & \left| \int_{\mathbb{R}} g_n(x) \left(L_{[nt]/n}^{x/n^\alpha}(X) - L_t^0(X) \right) dx \right| \\ & \leq \|g\|_\infty \int_{|x| \leq q} \left| L_{[nt]/n}^{x/n^\alpha}(X) - L_t^0(X) \right| dx + U \int_{|x| > q} |g_n(x)| dx. \end{aligned} \quad (4.23)$$

From (4.21), (4.22) and Lebesgue convergence theorem,

$$\int_{|x| \leq q} \left| L_{[nt]/n}^{x/n^\alpha}(X) - L_t^0(X) \right| \xrightarrow{n \rightarrow \infty} 0. \quad (4.24)$$

With a change of variables,

$$\begin{aligned} \int_{|x| > q} |g_n(x)| dx &= \int_{|x| > q} |g(n^\alpha T(x/n^\alpha))| dx \\ &= \int_{u_n T(q/u_n)}^\infty g(y) \frac{1}{T'(y/u_n)} dy + \int_{-\infty}^{u_n T(-q/u_n)} g(y) \frac{1}{T'(y/u_n)} dy \end{aligned} \quad (4.25)$$

From (4.5) and (4.25),

$$\limsup_n \int_{|x| > q} |g_n(x)| dx \leq \frac{1}{\epsilon} \left(\int_q^\infty g(y) dy + \int_{-\infty}^{-q} g(y) dy \right)$$

which since g is integrable converges to 0 as $q \rightarrow \infty$. From (4.23), (4.24) and (4.25),

$$\left| \int_{\mathbb{R}} g_n(x) \left(L_{[nt]/n}^{x/n^\alpha}(X) - L_t^0(X) \right) dx \right| \xrightarrow{n \rightarrow \infty} 0. \quad (4.26)$$

Using again the same change of variables as in (4.25),

$$\int_{\mathbb{R}} g_n(x) dx = \int_{\mathbb{R}} g(n^\alpha T(x/n^\alpha)) dx = \int_{\mathbb{R}} g(x) \frac{1}{T'(T^{-1}(x/n^\alpha))} dx. \quad (4.27)$$

Thus, as g is integrable and $T'(x) \geq \epsilon$ for every $x \in \mathbb{R}$, from Lebesgue convergence theorem,

$$\int_{\mathbb{R}} g_n(x) dx \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}} g(x) dx. \quad (4.28)$$

Equations (4.26) and (4.28) yield (4.20). \square

Lemma 4.2.5. Let T_n be the functional define for each real-valued function h and $x \in \mathbb{R}$ by

$$T_n[h](x) = \int_0^1 (P_{n^{2\alpha}s/n}^{n^\alpha\rho} h(x) - h(x)) ds,$$

where $(P_t^{n^\alpha\rho})_{t \geq 0}$ is the semi-group of the sticky Brownian motion of stickiness parameter $n^\alpha\rho$. Then, for every bounded integrable Lipschitz function k such that k vanishes on an open interval around 0,

$$\lambda(|T_n[k_n]|) \xrightarrow[n \rightarrow \infty]{} 0, \quad (4.29)$$

where $k_n(x) = k(n^\alpha T(x/n^\alpha))$.

Proof. From Jensen's inequality,

$$\lambda(|T_n[k_n]|) \leq \int_0^1 \int_{\mathbb{R}} |P_{n^{2\alpha}s/n}^{n^\alpha\rho} k_n(x) - k_n(x)| dx ds. \quad (4.30)$$

From (4.13),

$$\begin{aligned} |P_{n^{2\alpha}s/n}^{n^\alpha\rho} k_n(x) - k_n(x)| &\leq \int_{\mathbb{R}} |k_n(y) - k_n(x)| p_{n^\alpha\rho}(n^{2\alpha}s/n, x, y) m(dy) \\ &= \int_{\mathbb{R}} |k_n(y) - k_n(x)| p_{n^\alpha\rho}(n^{2\alpha}s/n, x, y) dy + \frac{n^\alpha\rho}{2} |k_n(x)| p_{n^\alpha\rho}(n^{2\alpha}s/n, x, 0) \\ &\leq K \left[\int_{\mathbb{R}} |k_n(y) - k_n(x)| \frac{1}{n^\alpha \sqrt{2\pi s/n}} e^{-(x-y)^2 n / 2sn^{2\alpha}} dy + |k_n(x)| \frac{\sqrt{n\rho}}{2\sqrt{2\pi s}} e^{-x^2 n / 2sn^{2\alpha}} \right], \end{aligned}$$

where $\frac{1}{n^\alpha \sqrt{2\pi s/n}} e^{-(x-y)^2 n / 2sn^{2\alpha}}$ is the probability density function of a Gaussian $\mathcal{N}(x, n^{2\alpha}s/n)$.

Thus from positive Fubini,

$$\begin{aligned} \int_{\mathbb{R}} |P_{n^{2\alpha}s/n}^{n^\alpha\rho} k_n(x) - k_n(x)| dx &\leq K \left[\mathbb{E} \left[\int_{\mathbb{R}} |k_n(x + n^\alpha \sqrt{s/n} Z) - k_n(x)| dx \right] \right. \\ &\quad \left. + \int_{\mathbb{R}} |k_n(x)| \frac{\sqrt{n\rho}}{2\sqrt{2\pi s}} e^{-x^2 n / 2sn^{2\alpha}} dx \right], \quad (4.31) \end{aligned}$$

where $Z \sim \mathcal{N}(0, 1)$ under \mathbb{P}_x . For the first additive term of right-hand side of (4.31), with the same argument as (4.27),

$$\int_{\mathbb{R}} k_n(x) dx \leq \frac{1}{\epsilon} \int_{\mathbb{R}} k(x) dx.$$

Thus, for every $\omega \in \Omega$,

$$\int_{\mathbb{R}} |k_n(x + n^\alpha \sqrt{s/n} Z) - k_n(x)| dx \leq \int_{\mathbb{R}} |k_n(x + n^\alpha \sqrt{s/n} Z)| dx + \int_{\mathbb{R}} |k_n(x)| dx \leq \frac{2}{\epsilon} \lambda(|k|). \quad (4.32)$$

Moreover since $T \in C^1$ and k is Lipschitz, \mathbb{P}_x -almost surely,

$$\begin{aligned} |k_n(x + n^\alpha \sqrt{s/n} Z) - k_n(x)| &= |k(n^\alpha T(x/n^\alpha + \sqrt{s/n} Z)) - k(n^\alpha T(x/n^\alpha))| \\ &\leq |k(n^\alpha T(x/n^\alpha) + \sqrt{s/n} Z \|T'\|_\infty) - k(n^\alpha T(x/n^\alpha))| \leq |k|_{\text{Lip}} \sqrt{s/n} Z \|T'\|_\infty, \quad (4.33) \end{aligned}$$

which converges to 0 as $n \rightarrow \infty$. From (4.32), (4.33) and Lebesgue convergence theorem,

$$\mathbb{E} \left[\int_{\mathbb{R}} |k_n(x + n^\alpha \sqrt{s/n} Z) - k_n(x)| dx \right] \xrightarrow{n \rightarrow \infty} 0. \quad (4.34)$$

For the second additive term of right-hand side of (4.31): let $\delta > 0$ be a positive real number such that $k(x) = 0$ for every $x \geq 0$ such that $x \notin (\delta, 1/\delta)$. From (4.5), T is strictly increasing, thus,

$$\delta \leq u_n T(x/u_n) \leq 1/\delta,$$

is equivalent to

$$u_n T^{-1}(\delta/u_n) \leq x \leq u_n T^{-1}(1/u_n \delta).$$

From (4.5),

$$\begin{aligned} \liminf_n u_n T^{-1}(\delta/u_n) &\geq \delta \epsilon, \\ \limsup_n u_n T^{-1}(1/u_n \delta) &\leq 1/\delta \epsilon. \end{aligned}$$

Thus, there exists $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$, $\text{supp } k_n \subset (\epsilon\delta/2, 2/\epsilon\delta)$. Thus, since k is bounded,

$$\begin{aligned} \int_{\mathbb{R}} |k_n(x)| \frac{\sqrt{n}\rho}{2\sqrt{2\pi s}} e^{-x^2 n/2sn^{2\alpha}} dx &\leq 2\|k\|_\infty \int_{\epsilon\delta/2}^{2/\epsilon\delta} \frac{\sqrt{n}\rho}{2\sqrt{2\pi s}} e^{-x^2 n/2sn^{2\alpha}} dx \\ &\leq \frac{\rho\|k\|_\infty}{\sqrt{2\pi s}} \frac{2}{\epsilon\delta} \sqrt{n} e^{-n^{1-2\alpha}(\epsilon\delta)^2/8s}, \end{aligned} \quad (4.35)$$

which converges to 0 as $n \rightarrow \infty$. From (4.30), (4.31), (4.34) and (4.35), the convergence (4.29) is proven. \square

Proof (of Theorem 4.2.1). Let X be the sticky Brownian motion of parameter $\rho > 0$, $(P_t^\rho)_{t \geq 0}$ its semi-group and $L^x(X)$ its local time at x . From the occupation times formula and the characterization of $\langle X^\rho \rangle_t$ in [30],

$$\int_0^t f(X_s) \mathbb{1}_{X_s \neq 0} ds = \int_{\mathbb{R}} f(y) L_t^y(X) dy.$$

Thus, if $f(0) = 0$,

$$\int_0^t f(X_s) ds = \int_{\mathbb{R}} f(y) L_t^y(X) dy. \quad (4.36)$$

By applying consecutive change of variables and from (4.36),

$$\begin{aligned} \int_{\mathbb{R}} g_n(x) L_{[nt]/n}^{x/n^\alpha}(X) dx &= n^\alpha \int_{\mathbb{R}} g_n(n^\alpha x) L_{[nt]/n}^x(X) dx \\ &= n^\alpha \int_0^{[nt]/n} g_n(n^\alpha X_s) ds = \frac{n^\alpha}{n} \int_0^{[nt]} g_n(n^\alpha X_{s/n}) ds. \end{aligned}$$

Thus,

$$\begin{aligned}
\frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g_n(n^\alpha X_{\frac{i-1}{n}}) &= \frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g_n(n^\alpha X_{\frac{i-1}{n}}) - \frac{n^\alpha}{n} \int_0^{[nt]} g_n(n^\alpha X_{s/n}) ds + \int_{\mathbb{R}} g_n(x) L_{[nt]/n}^{x/n^\alpha}(X) dx \\
&= \sum_{i=1}^{[nt]} \frac{n^\alpha}{n} \int_0^1 \left(g_n(n^\alpha X_{\frac{i-1}{n}}) - g_n(n^\alpha X_{\frac{i-1}{n} + \frac{s}{n}}) \right) ds \\
&\quad + \frac{n^\alpha}{n} \int_0^{nt-[nt]} \left(g_n(n^\alpha X_{\frac{[nt]}{n}}) - g_n(n^\alpha X_{\frac{[nt]}{n} + \frac{s}{n}}) \right) ds \\
&\quad + \int_{\mathbb{R}} g_n(x) L_{[nt]/n}^{x/n^\alpha}(X) dx.
\end{aligned} \tag{4.37}$$

For the second additive term at the right hand side of (4.37),

$$\begin{aligned}
&\left| \frac{n^\alpha}{n} \int_0^{nt-[nt]} \left(g_n(n^\alpha X_{\frac{[nt]}{n}}) - g_n(n^\alpha X_{\frac{[nt]}{n} + \frac{s}{n}}) \right) ds \right| \\
&\leq \frac{n^\alpha}{n} \int_0^{nt-[nt]} \left| g_n(n^\alpha X_{\frac{[nt]}{n}}) - g_n(n^\alpha X_{\frac{[nt]}{n} + \frac{s}{n}}) \right| ds \leq 2 \|g\|_\infty \frac{n^\alpha}{n},
\end{aligned} \tag{4.38}$$

which converges to 0 as $n \rightarrow \infty$.

For the first additive term at the right hand side of (4.37), let

$$\begin{aligned}
A_t^n &:= \sum_{i=1}^{[nt]} \frac{n^\alpha}{n} \int_0^1 \left(g_n(n^\alpha X_{\frac{i-1}{n} + \frac{s}{n}}) - g_n(n^\alpha X_{\frac{i-1}{n}}) \right) ds, \\
B_t^n &:= \sum_{i=1}^{[nt]} \frac{n^\alpha}{n} \int_0^1 \mathbb{E}_x \left(g_n(n^\alpha X_{\frac{i-1}{n} + \frac{s}{n}}) - g_n(n^\alpha X_{\frac{i-1}{n}}) \middle| \mathcal{F}_{\frac{i-1}{n}} \right) ds.
\end{aligned}$$

As the cross terms have expectancy 0, from Minkowski's inequality,

$$\mathbb{E}_x \left(|A_t^n - B_t^n|^2 \right) \leq 2 \frac{n^{2\alpha}}{n^2} \sum_{i=1}^{[nt]} \mathbb{E}_x \left(\int_0^1 \left(g_n(n^\alpha X_{\frac{i-1}{n} + \frac{s}{n}}) - g_n(n^\alpha X_{\frac{i-1}{n}}) \right) ds \right)^2 \leq 4 \|g\|_\infty^2 \frac{n^{2\alpha}}{n} \frac{[nt]}{n}, \tag{4.39}$$

which converges to 0 as $n \rightarrow \infty$. Thus, $A_t^n - B_t^n$ converges to 0 in $L^2(\mathbb{P}_x)$ and consequently in $L^1(\mathbb{P}_x)$. As such, proving that $A_t^n \xrightarrow{n \rightarrow \infty} 0$ in $L^1(\mathbb{P}_x)$ is equivalent to proving that $B_t^n \xrightarrow{n \rightarrow \infty} 0$ in $L^1(\mathbb{P}_x)$. To prove the latter, we define for each real-valued function h the functional,

$$T_n[h](x) = \int_0^1 (P_{n^{2\alpha}s/n}^{n^\alpha \rho} h(x) - h(x)) ds.$$

From (A.5),

$$B_t^n = \frac{n^\alpha}{n} \sum_{i=1}^{[nt]} T_n[g_n](n^\alpha X_{\frac{i-1}{n}}).$$

From (4.14), there exists a constant $K' > 0$ that does not depend on n or ρ such that

$$\mathbb{E}_x \left(|B_t^n| \right) \leq K' \left(\frac{u_n}{n} |T_n[g_n](u_n x)| + \sqrt{t} \lambda(|T_n[g_n]|) \right).$$

From Jensen and (4.6), for every $x \in \mathbb{R}$, $t > 0$ and $\rho > 0$,

$$|P_t^\rho g_n(x)| \leq \|g\|_\infty,$$

By taking $K'' = K'(2\|g\|_\infty \vee 1)$,

$$\mathbb{E}_x(|B_t^n|) \leq K'' \left(\frac{n^\alpha}{n} + \sqrt{t} \lambda(|T_n[g_n]|) \right). \quad (4.40)$$

Thus, as $0 < \alpha < 1/2$, it remains to prove that $\lambda(|T_n[g_n]|) \rightarrow 0$. For this we use a Lipschitz approximation of g . In particular, as g is bounded and in $L^1(dx)$, for each p it is possible to find a Lipschitz function k_p such that, $k_p(0) = 0$ and $\lambda(|g - k_p|) < 1/p$. Let $k_p^n(x) = k_p(n^\alpha T(x/n^\alpha))$, from (4.5),

$$\lambda(|g_n - k_p^n|) < 1/p\epsilon. \quad (4.41)$$

Let $p_\rho(t, x, y)$ be the sticky Brownian motion transition kernel given in (4.10). As $p_\rho(t, x, y) = p_\rho(t, y, x)$ for every $x, y \in \mathbb{R}$,

$$\begin{aligned} \lambda(|P_t^\rho g_n - P_t^\rho k_p^n|) &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} |g_n(y) - k_p^n(y)| p_\rho(t, x, y) dy dx = \\ &\int_{\mathbb{R}} |g_n(y) - k_p^n(y)| \int_{\mathbb{R}} p_\rho(t, x, y) dx dy = \int_{\mathbb{R}} |g_n(y) - k_p^n(y)| \int_{\mathbb{R}} p_\rho(t, y, x) dx dy \\ &\leq \int_{\mathbb{R}} |g_n(y) - k_p^n(y)| dy = \lambda(|g_n - k_p^n|). \end{aligned} \quad (4.42)$$

From (4.41) and (4.42),

$$\lambda(|T_n[g_n]|) \leq \frac{2}{p\epsilon} + \lambda(|T_n[k_p^n]|).$$

Thus, from (4.29),

$$\lambda(|T_n[g_n]|) \xrightarrow{n \rightarrow \infty} 0. \quad (4.43)$$

From (4.39), (4.40) and (4.43),

$$A_t^n = \sum_{i=1}^{\lfloor nt \rfloor} \frac{n^\alpha}{n} \int_0^1 \left(g_n(n^\alpha X_{\frac{i-1}{n} + \frac{s}{n}}) - g_n(n^\alpha X_{\frac{i-1}{n}}) \right) ds \xrightarrow[n \rightarrow \infty]{P_x} 0. \quad (4.44)$$

From (4.20), (4.37), (4.38) and (4.44), for every $t \geq 0$,

$$\frac{n^\alpha}{n} \sum_{i=1}^{\lfloor nt \rfloor} g_n(n^\alpha X_{\frac{i-1}{n}}) \xrightarrow[n \rightarrow \infty]{P_x} \lambda(g) L_t^0(X).$$

If g is a positive function the processes $\frac{n^\alpha}{n} \sum_{i=1}^{\lfloor nt \rfloor} g_n(n^\alpha X_{\frac{i-1}{n}})$ are non-decreasing with P_x -almost surely, a continuous limit. Thus, from Lemma A.2.6, the convergence is locally uniform in time, in probability. For an arbitrary g satisfying the conditions of Theorem 4.2.1, let $g = g^+ - g^-$, where $g^+(x) = \max\{g(x), 0\}$ and $g^-(x) = \max\{-g(x), 0\}$. Then as g^+ and g^- are both positive function and thus,

$$\frac{n^\alpha}{n} \sum_{i=1}^{\lfloor nt \rfloor} g_n^+(n^\alpha X_{\frac{i-1}{n}}) \xrightarrow[n \rightarrow \infty]{} \lambda(g^+) L_t^0(X), \quad \frac{n^\alpha}{n} \sum_{i=1}^{\lfloor nt \rfloor} g_n^-(n^\alpha X_{\frac{i-1}{n}}) \xrightarrow[n \rightarrow \infty]{} \lambda(g^-) L_t^0(X),$$

locally uniformly in time, in P_x -probability. Using the triangle inequality for the absolute value and the $L^\infty(0, t)$ -norm, the locally uniform convergence of (4.7) in P_x -probability for X a sticky Brownian motion is proven. \square

4.3 Proofs of the main results

4.3.1 Proof of Theorem 4.1.3

The proof works by extending Theorem 4.2.1 using the path-wise tools developed in Chapter 3.

Proof (of Theorem 4.1.3). We suppose there exists a $\delta > 0$ such that

$$\delta \leq \sigma(x) \leq 1/\delta, \quad |\sigma'(x)| \leq 1/\delta, \quad (4.45)$$

for every $x \in \mathbb{I}$. From Proposition 3.2.3, there exists a Brownian motion B such that (X, B) jointly solves

$$\begin{aligned} dX_t &= \mu(X_t) \mathbb{1}_{X_t \neq 0} dt + \sigma(X_t) \mathbb{1}_{X_t \neq 0} dB_t, \\ \mathbb{1}_{X_t=0} dt &= \frac{\rho}{2} dL_t^0(X). \end{aligned}$$

Let \mathbb{Q}_x be the probability measure such that $d\mathbb{Q}_x = \mathcal{E}_t(\theta) d\mathbb{P}_x$ where

$$\begin{aligned} \mathcal{E}_t(\theta) &= \exp\left(\int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds\right), \\ \theta_t &= \sigma'(X_t) - \frac{\mu(X_t)}{\sigma(X_t)}. \end{aligned}$$

From Lemma 3.3.1, (X, \tilde{B}) jointly solve

$$\begin{aligned} dX_t &= \frac{1}{2} \sigma(X_t) \sigma'(X_t) \mathbb{1}_{X_t \neq 0} dt + \sigma(X_t) \mathbb{1}_{X_t \neq 0} d\tilde{B}_t, \\ \mathbb{1}_{X_t=0} dt &= \frac{\rho}{2} dL_t^0(X), \end{aligned}$$

where $\tilde{B} = B_t - \int_0^t \theta_s ds$ is a standard Brownian motion under \mathbb{Q}_x . Let S be the function defined for every $x \in \mathbb{R}$ by

$$S(x) = \int_0^x \frac{1}{\sigma(y)} dy.$$

We observe that S is strictly increasing and $S(0) = 0$. Thus, from (3.35)-(3.36), the process $X' = (S(X_t))_{t \geq 0}$ solves

$$dX'_t = \frac{1}{\sigma(X_t)} dX_t - \frac{1}{2} \frac{\sigma'(X_t)}{\sigma^2(X_t)} d\langle X \rangle_t = \mathbb{1}_{X'_t \neq 0} d\tilde{B}_t, \quad (4.46)$$

$$\mathbb{1}_{X'_t=0} dt = \frac{\rho}{2\sigma(0)} dL_t^0(X'). \quad (4.47)$$

From Proposition 3.2.2, X' is a sticky Brownian motion of parameter $\rho > 0$. Let U and T_0 be the functions defined for every $x \in \mathbb{R}$ by

$$U(x) = S^{-1}(x)/\sigma(0), \quad T_0(x) = T(U(x)). \quad (4.48)$$

Then,

$$T'_0(x) = T' \left(\frac{S^{-1}(x)}{\sigma(0)} \right) \frac{\sigma(x)}{\sigma(0)}, \quad T''_0(x) = T'' \left(\frac{S^{-1}(x)}{\sigma(0)} \right) \frac{\sigma^2(x)}{\sigma^2(0)} + T' \left(\frac{S^{-1}(x)}{\sigma(0)} \right) \frac{\sigma'(x)}{\sigma(0)}. \quad (4.49)$$

From (4.5), (4.45) and (4.49), for every $x \in \mathbb{R}$,

$$T_0(0) = T(0) = 0,$$

$$T'_0(0) = T'(0) = 1,$$

$$\epsilon\delta/\sigma(0) \leq T'_0(x) \leq 1/\epsilon\delta\sigma(0),$$

$$|T''_0(x)| \leq \|T''\|_\infty \|\sigma\|_\infty^2 / \sigma^2(0) + \|T'\|_\infty \|\sigma'\|_\infty / \sigma(0) \leq \frac{1}{\epsilon\delta\sigma(0)} \left(\frac{1}{\delta\sigma(0)} + 1 \right).$$

Thus, T_0 satisfies (4.5) for $\epsilon_{T_0} = \epsilon\delta(\sigma(0) \wedge 1/\sigma(0))$ and from Theorem 4.2.1 and (4.46)-(4.47),

$$\frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g_n[T_0] \left(n^\alpha X'_{\frac{i-1}{n}} \right) \xrightarrow{n \rightarrow \infty} \lambda(g)L_t^0(X'),$$

locally, uniformly in time, in \mathbb{Q}_x -probability, where $g_n[T_0](x) = g(n^\alpha T_0(x/n^\alpha))$. From (4.48) and since

$$g_n[T_0](n^\alpha x) = g(n^\alpha T_0(x)) = g(n^\alpha T(U(x))) = g_n[T](U(n^\alpha x))$$

we get

$$\frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g_n[T](n^\alpha X_{\frac{i-1}{n}}) \xrightarrow{n \rightarrow \infty} \lambda(g)L_t^0(X'), \quad (4.50)$$

locally, uniformly in time, in \mathbb{Q}_x -probability. From (4.3), the local time is defined as a limit in probability. Thus, from (4.50) and as $P_x \sim \mathbb{Q}_x$,

$$\frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g_n[T](n^\alpha X_{\frac{i-1}{n}}) \xrightarrow{n \rightarrow \infty} \lambda(g)L_t^0(X'),$$

locally, uniformly in time, in P_x -probability. From Lemma A.2.3, $L_t^0(X)/\sigma(0)$ is a version of $L_t^0(X')$. Thus, since we supposed (4.45), (4.7) is proven in the case of bounded σ , $1/\sigma$ and σ' and from section 2-5 of [55], the proof is extended to any $\sigma \in C^1$. \square

4.3.2 Proof of Corollary 4.1.4

From (3.9), the occupation/local time ratio is the stickiness parameter. Thus having an estimation of the two aforementioned quantities that does not require the knowledge of ρ allow us to use the ratio as a consistent estimator. We first show that the occupation times can be consistently approximated by Riemann sums. Then, we use it along with Theorem 4.1.3 to prove Corollary 4.1.4.

Lemma 4.3.1. *Let X be a semi-martingale and $A^0(X)$ be its occupation time of 0 defined for every $t \geq 0$ by*

$$A_t^0(X) = \int_0^t \mathbb{1}_{X_s=0} ds.$$

Then,

$$\frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{X_{\frac{i-1}{n}}=0} \xrightarrow{n \rightarrow \infty} A_t^0(X), \quad (4.51)$$

locally uniformly in time, in probability.

Proof. As both $t \mapsto \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{X_{\frac{i-1}{n}}=0}$ and $A^0(X)$ are increasing processes, with the same argument as in the proof of Theorem 4.2.1, it suffices to prove the convergence in probability for each $t > 0$. If $\delta > 0$ and $\epsilon > 0$ are two positive numbers,

$$\begin{aligned}
& \mathbb{P}_x \left(\left| \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{X_{\frac{i-1}{n}}=0} - A_t^0(X) \right| > \delta \right) \\
&= \mathbb{P}_x \left(\left| \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{|X_{\frac{i-1}{n}}| < \epsilon} - \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{0 < |X_{\frac{i-1}{n}}| < \epsilon} - A_t^0(X) \right| > \delta \right) \\
&\leq \mathbb{P}_x \left(\left| \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{|X_{\frac{i-1}{n}}| < \epsilon} - A_t^0(X) \right| + \left| \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{0 < |X_{\frac{i-1}{n}}| < \epsilon} \right| > \delta \right) \\
&\leq \mathbb{P}_x \left(\left| \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{0 < |X_{\frac{i-1}{n}}| < \epsilon} \right| > \frac{\delta}{2} \right) + \mathbb{P}_x \left(\left| \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{|X_{\frac{i-1}{n}}| < \epsilon} - A_t^0(X) \right| > \frac{\delta}{2} \right). \quad (4.52)
\end{aligned}$$

From (4.9) for $h(x) = \mathbb{1}_{0 < |x| < \epsilon}$ and as $\sum_{i=1}^{[nt]} \frac{1}{\sqrt{i}} \leq 2\sqrt{nt}$,

$$\mathbb{E}_x \left(\frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{0 < |X_{\frac{i-1}{n}}| < \epsilon} \right) = \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{E}_x \left(h(X_{\frac{i-1}{n}}) \right) \leq \frac{1}{n} 2\epsilon K \sqrt{n} \sum_{i=1}^{[nt]} \frac{1}{\sqrt{i}} \leq 4\epsilon K \sqrt{t}$$

Thus, from Markov's inequality,

$$\mathbb{P}_x \left(\left| \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{0 < |X_{\frac{i-1}{n}}| < \epsilon} \right| > \frac{\delta}{2} \right) \leq \frac{8\epsilon K \sqrt{t}}{\delta}. \quad (4.53)$$

We consider the following functions:

$$\begin{aligned}
\phi(x) &= (2 - |x|) \mathbb{1}_{1 \leq |x| < 2} + \mathbb{1}_{|x| < 1}, \\
\psi(x) &= 2(1 - |x|) \mathbb{1}_{0.5 \leq |x| < 1} + \mathbb{1}_{|x| < 0.5}.
\end{aligned}$$

The functions ϕ and ψ are both continuous and bounded with compact support and

$$\psi(x) \leq \mathbb{1}_{|x| < 1} \leq \phi(x).$$

The composed function $\phi(X_t)$ and $\psi(X_t)$ are both a.s. continuous functions of t , hence a.s. Riemann integrable. Thus,

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^{[nt]} \phi\left(\frac{1}{\epsilon} X_{\frac{i-1}{n}}\right) &\xrightarrow[n \rightarrow \infty]{} \int_0^t \phi\left(\frac{1}{\epsilon} X_s\right) ds \\
\frac{1}{n} \sum_{i=1}^{[nt]} \psi\left(\frac{1}{\epsilon} X_{\frac{i-1}{n}}\right) &\xrightarrow[n \rightarrow \infty]{} \int_0^t \psi\left(\frac{1}{\epsilon} X_s\right) ds
\end{aligned}$$

From Lebesgue convergence theorem both $\int_0^t \phi\left(\frac{1}{\epsilon} X_s\right) ds$ and $\int_0^t \psi\left(\frac{1}{\epsilon} X_s\right) ds$ converge to $A_t^0(X)$ as $\epsilon \rightarrow 0$. Thus for each $\delta > 0$ there exists an $\epsilon > 0$ such that

$$\begin{aligned}
\mathbb{P}_x\text{-}\limsup_n \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{|X_{\frac{i-1}{n}}| < \epsilon} &\leq A_t^0(X) + \delta, \\
\mathbb{P}_x\text{-}\liminf_n \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{1}_{|X_{\frac{i-1}{n}}| < \epsilon} &\geq A_t^0(X) - \delta.
\end{aligned}$$

Thus, for each $\delta > 0$ there exists an $\epsilon_0 > 0$ such that

$$P_x \left(\left| \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{1}_{|X_{\frac{i-1}{n}}| < \epsilon} - A_t^0(X) \right| > \frac{\delta}{2} \right) \xrightarrow{n \rightarrow \infty} 0. \quad (4.54)$$

From (4.52), (4.53) and (4.54), for each $\epsilon' > 0$ by choosing $\epsilon' = \epsilon\delta$ in (4.53), there exists a $\delta > 0$ such that

$$0 \leq \limsup_n P_x \left(\left| \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{1}_{X_{\frac{i-1}{n}}=0} - A_t^0(X) \right| > \delta \right) \leq 4\epsilon K \sqrt{t}.$$

Thus, $\frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{1}_{X_{\frac{i-1}{n}}=0}$ converges in probability to $A_t^0(X)$. From (3.9), $A_t^0(X)$ admits almost surely a continuous version. Thus, from Lemma A.2.6, (4.51) is proven. \square

Proof (of Corollary 4.1.4). We consider that we are on the event $\mathcal{L} = \{L_t^0(X) > 0\}$. By integrating (3.9) we observe that

$$A_t^0(X) = \frac{\rho}{2} L_t^0(X). \quad (4.55)$$

Thus,

$$\mathcal{L} = \{L_t^0(X) > 0\} = \{A_t^0(X) > 0\}.$$

From (4.7) and (4.51),

$$\frac{(1/n) \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{1}_{X_{\frac{i-1}{n}}=0}}{(n^\alpha/n) \sum_{i=1}^{\lfloor nt \rfloor} g_n[T](n^\alpha X_{\frac{i-1}{n}})} \xrightarrow[n \rightarrow \infty]{P_x} \frac{A_t^0(X)}{(\lambda(g)/\sigma(0))L_t^0(X)}. \quad (4.56)$$

On the event \mathcal{L} , both terms in the ratio on the right-hand side of (4.56) are strictly positive. As such, the ratio is a well-defined real-valued random variable. Thus, from (4.55) we get (4.8) on \mathcal{L} . \square

Chapter 5

General diffusion processes as the limit of time-space Markov chains

In this chapter, we present the theoretical results of [6]. Numerical illustrations are given in Chapter 6.

5.1 Introduction

In the diffusion process literature, the most well-studied and straightforward way to approximate diffusion processes is the Euler scheme. While such approximations works well for non-degenerate stochastic differential equations, this is not the case for more general diffusion processes [51]. The Euler scheme is also not well-defined for processes that exhibit sticky features, skew behavior [40] or slowly reflecting boundaries.

Several works aim at overcoming the shortcomings of the Euler scheme and allow us to approximate the law of more general diffusion processes. In [4], the author proposes to approximate the sticky Brownian motion with a simple random walk that stops for a fixed amount of time when it hits 0. In [76] and [73], Continuous Time Markov Chains are used to approximate slowly reflected SDE solutions, where the jumping intensities are computed using approximated discretizations of the infinitesimal generator of the diffusion. Another work where such processes are defined is [39], where the authors use a Continuous Time Markov Chain to identify events in genomics evolution. In [7, 9], the authors develop a numerical scheme to approximate diffusions on natural scale as long as a mild non-explosion condition is satisfied. They use symmetric random walks with fixed-time step whose magnitude depends on the average local behavior of the target process speed measure. Choosing the step-size this way allows the approximation process to replicate non-boundary sticky features.

In this paper, we prove the convergence in law of grid-valued random walks to any one-dimensional general diffusion process at an asymptotically optimal rate. General diffusion processes are regular one-dimensional strong Markov processes with continuous trajectories, see for instance [80, Chapter 7.3] where they are defined as linear continuous Markov processes. This convergence result allows us to set up approximation

schemes that, while make it straightforward to take into account for sticky points, can also be applied to any diffusion process that satisfy a mild non-explosion condition. This includes processes with boundary conditions like absorption, reflection or slow-reflection as well as the skew diffusions such as the Skew Brownian motion [66] and its generalizations. The values taken by the random walk correspond to values taken by the target process at random times, allowing us to classify it as an embeddable scheme along with [7] and [31]. We prove that for a grid adapted to the speed measure of the diffusion process, the laws of the random walks converges at any rate strictly inferior to $(1/2) \wedge (2/p)$ in terms of the maximum cell size for all p -Wasserstein distances. This convergence rate is optimal for $p \leq 4$ according the Donsker invariance principle, as this is the rate simple randoms walk converges to the standard Brownian motion [26].

Besides the asymptotic optimal convergence rate, the usage of such an approximation process yield several advantages. Firstly, the static character of the grid makes involved quantities good candidates for numerical approximation (see Sections 5.6 and 6.1). Moreover, this scheme makes it straightforward to take into account potential sticky points of the diffusion. Finally, its universality is further validated by the fact that the Donsker invariance principle and [26, 4, 31] are all special cases of it.

Outline. In Section 5.2, we present the approximation scheme along with its properties. In Section 5.3, we give analytical characterizations of the quantities that determine the law of the random walk defined by the algorithm, allowing us to implement it. Section 5.4 is dedicated to proving the convergence of embedding times. In Section 5.5, we prove the main convergence result in terms of the maximum cell size of the grid. The case of solution of an SDE is studied in Section 5.6.

5.2 The Space-Time Markov Chain Approximation and its properties

5.2.1 The approximation scheme

In this section we define the approximation process for a one-dimensional diffusion process on natural scale³ with state space \mathbb{I} , an open interval of \mathbb{R} . The general case is obtained by a change of scale, as detailed in Section 5.2.2.

The possible values taken by the approximation process are given as input of the scheme and must form a *covering grid* of \mathbb{I} . We introduce incrementally this notion as follows, which we illustrate by Figure 5.1. Let \mathbb{I} be an interval of \mathbb{R} .

- A *grid* over \mathbb{I} is a countable subset of \mathbb{I} with no accumulation points within \mathbb{I} .
- A *cell* \mathbf{c} of a grid \mathbf{g} is an open interval with endpoints in \mathbf{g} with a single element of \mathbf{g} in its interior, *i.e.*,

$$\text{card}(\mathbf{g} \cap \mathbf{c}) = 1.$$

We denote with $\mathcal{C}(\mathbf{g})$ the set of all cells of the grid \mathbf{g} .

³Which means that $s(x) = x$.

- We call $x \in \mathbb{I}$ the *center* of the cell $\mathbf{c} \in \mathcal{C}(\mathbf{g})$ iff

$$\mathbf{c} \cap \mathbf{g} = \{x\}.$$

- Finally, a *covering grid* of \mathbb{I} is a grid \mathbf{g} such that $\mathbb{I} = \bigcup_{\mathbf{c} \in \mathcal{C}(\mathbf{g})} \mathbf{c}$.

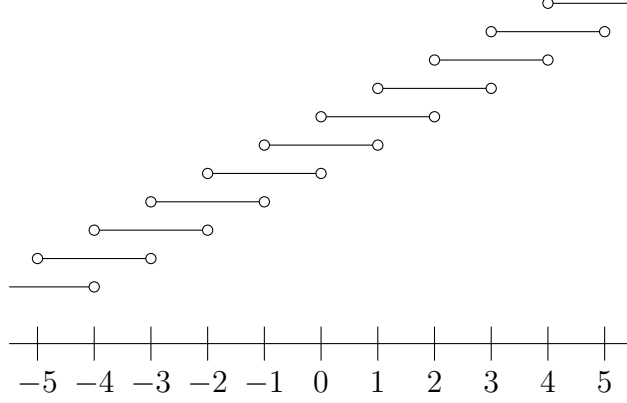


Figure 5.1: The covering grid $\mathbf{g} = \mathbb{Z}$ of $(-\infty, \infty)$ along with some of its cells $\mathcal{C}(\mathbf{g})$.

Examples of covering grids of $(0, \infty)$ and $(-\infty, \infty)$ are $\{1/n; n \in \mathbb{N}\}$ and \mathbb{Z} respectively.

For any covering grid \mathbf{g} of \mathbb{I} and diffusion process X with state-space \mathbb{I} , defined through (s, m) , let $|\mathbf{g}|$, $|\mathbf{g}|_X$ be the grid metrics:

$$|\mathbf{g}| = \sup_{\mathbf{c} \in \mathcal{C}(\mathbf{g})} |\mathbf{c}|, \quad |\mathbf{g}|_X = \sup_{\mathbf{c} \in \mathcal{C}(\mathbf{g})} \{s(\mathbf{c})m(\mathbf{c})\}, \quad (5.1)$$

where $|\mathbf{c}| = (b - a)$ and $s(\mathbf{c}) = s(b) - s(a)$ with a and b being the endpoints of \mathbf{c} . The convergence results will be expressed in the latter metric.

Let X be a diffusion process on natural scale with state-space \mathbb{I} and speed measure m . Let X be defined on the family of filtered probability spaces $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, (P_x)_{x \in \mathbb{I}})$, where P_x is the law of X such that, for any $x \in \mathbb{I}$, $P_x\{X_0 = x\} = 1$. For any covering grid \mathbf{g} of \mathbb{I} , we define the process $\widetilde{X}^{\mathbf{g}} = (\widetilde{X}_t^{\mathbf{g}})_{t \geq 0}$ as the asymmetric random walk with:

- state-space \mathbf{g} ,
- initial distribution equal to the distribution of X the first time it touches the grid,
- the same transition probabilities as X over \mathbf{g} ,
- conditional transition times that match the conditional expected transition times of X over \mathbf{g} .

Thus, under P_x , if a and b are respectively the closest lower and upper elements to x of \mathbf{g} ,

$$\widetilde{X}_0^{\mathbf{g}} = \begin{cases} a, & \text{with probability } P_x(\tau_b < \tau_a), \\ b, & \text{with probability } P_x(\tau_a < \tau_b) = 1 - P_x(\tau_b < \tau_a), \end{cases} \quad (5.2)$$

where $\tau_a := \inf\{t > 0 : X_t = a\}$. For the rest of the trajectory, we define $\tau_{ab} := \tau_a \wedge \tau_b$ and $(T^{\mathbf{g}}(n))_{n \geq 0}$ as the consecutive jumping times of $(\widetilde{X}_t^{\mathbf{g}})_{t \geq 0}$. Then, for all $k \in \mathbb{N}_0$ and $a < x < b$ adjacent points of \mathbf{g} ,

$$P\left(\widetilde{X}_{T^{\mathbf{g}}(k+1)}^{\mathbf{g}} = b \mid \widetilde{X}_{T^{\mathbf{g}}(k)}^{\mathbf{g}} = x\right) = P_x(\tau_b < \tau_a), \quad (5.3)$$

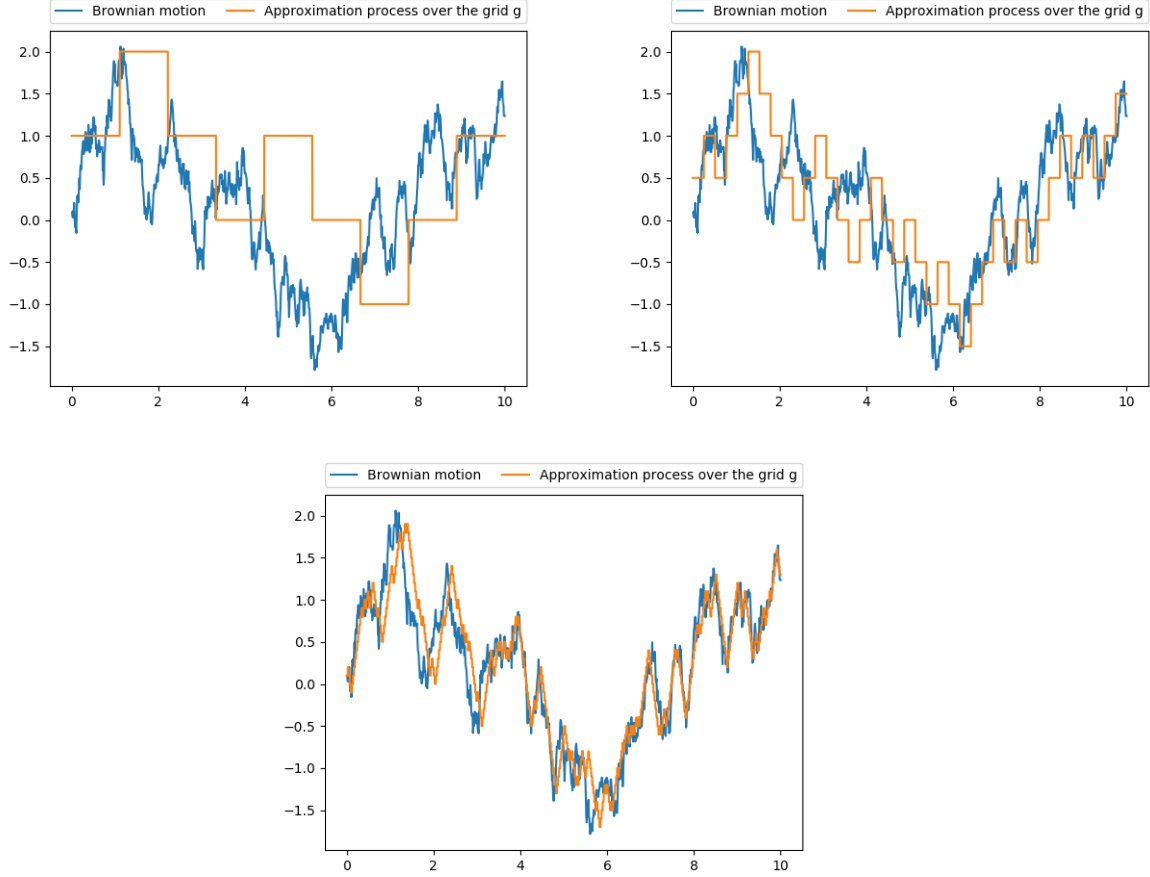


Figure 5.2: Approximations of the standard Brownian motion, Euler vs STMCA for various grid sizes $h = 1.0$, $h = 0.5$, $h = 0.1$.

and

$$T^{\mathbf{g}}(k+1) - T^{\mathbf{g}}(k) = \begin{cases} \mathbb{E}_x(\tau_{ab} | \tau_b < \tau_a), & \text{on } \{\widetilde{X}_{T^{\mathbf{g}}(k+1)}^{\mathbf{g}} = b\} \cap \{\widetilde{X}_{T^{\mathbf{g}}(k)}^{\mathbf{g}} = x\}, \\ \mathbb{E}_x(\tau_{ab} | \tau_a < \tau_b), & \text{on } \{\widetilde{X}_{T^{\mathbf{g}}(k+1)}^{\mathbf{g}} = a\} \cap \{\widetilde{X}_{T^{\mathbf{g}}(k)}^{\mathbf{g}} = x\}. \end{cases} \quad (5.4)$$

As proved in Section 5.3, the quantities that appear on the right hand side of (5.3) and (5.4) are explicit functionals of the speed measure m .

Let \mathbf{c}_x be the cell of the grid \mathbf{g} containing x , *i.e.*, $\mathbf{c}_x \in \mathcal{C}(\mathbf{g})$ and $x = \mathbf{c}_x \cap \mathbf{g}$. From (5.3), (5.4) and Bayes' rule, if $\mathbf{c}_x = (a, b)$, both $\mathbb{P}(\widetilde{X}_{T^{\mathbf{g}}(k+1)}^{\mathbf{g}} = b | \widetilde{X}_{T^{\mathbf{g}}(k)}^{\mathbf{g}} = x)$ and $T^{\mathbf{g}}(k+1) - T^{\mathbf{g}}(k)$ only depend on x . Thus, if we know the quantities

$$\begin{aligned} p^+[x, (a, b)] &= \mathbb{P}_x(\tau_b < \tau_a), & T^+[x, (a, b)] &= \mathbb{E}_x[\tau_{ab} | \tau_b < \tau_a], \\ p^-[x, (a, b)] &= \mathbb{P}_x(\tau_a < \tau_b), & T^-[x, (a, b)] &= \mathbb{E}_x[\tau_{ab} | \tau_a < \tau_b], \end{aligned} \quad (5.5)$$

for any adjacent $a < x < b$ in \mathbf{g} , we can simulate the random walk using Algorithm 1. We discuss in Section 5.3 on how to compute the quantities in (5.5). Practical examples are given in Sections 5.6 and 6.1.

This algorithm has been first introduced in [31] in the situation of SDE solutions with measurable coefficients, where the speed measure of the process satisfies

$$c \, dx \leq m(dx) \leq C \, dx. \quad (5.6)$$

Our main contribution is that we allow non-elliptic speed measures with vanishing and unbounded density (as in e.g. the Bessel process case, see Section 6.1.6), speed measures with singular part (as in e.g. the sticky Brownian motion, see Sections 6.1.2), scale functions not in C^1 (as in e.g. the skew Brownian motion and skew Bessel process, see Section 6.1.3 and 6.1.6) and non-trivial boundary behaviors (see Section 6.1.5). The probabilistic arguments we use to prove our results allow for greater flexibility, while the proofs of [31] are based on elliptic PDE theory. This allows us to handle degenerate diffusions and to perform grid tuning and achieve higher orders of convergence (see Section 5.2.3).

Remark. We observe that the process $\widetilde{X}^{\mathbf{g}}$ is not a one-dimensional Markov chain. It is though a Markov chain in space and time since the joint law of the next position of $\widetilde{X}^{\mathbf{g}}$ on the grid and the next transition time are both determined by the current position on the grid. Hence the terms: space-time Markov chains and Space Time Markov Chain Approximation (STMCA).

Remark. In the case of sticky diffusions, where the speed measure m has the form $m(dx) = m_c(dx) + \rho\delta_0(dx)$, the transition probabilities and transition times (5.5) can be directly inferred from the ones of the diffusion without the sticky term. Indeed, Proposition 5.3.2 yields

$$\mathbb{E}_x \left[\tau_{ab} \mathbb{1}_{\tau_b < \tau_a} \right] = \int_{(a,b)} G_{a,b}(x, \zeta) v_0(\zeta) m_c(d\zeta) + \rho G_{a,b}(x, 0) v_0(0),$$

where $v_0(x) = \mathbb{P}_x(\tau_b < \tau_a)$.

For the convergence, we make the further assumption that the diffusion process $(X_t)_{t \geq 0}$ satisfies the following non-explosion condition: there exists a $k_1 > 0$ such that the speed measure of the diffusion process satisfies

$$m(dx) \geq k_1 \frac{1}{1+x^2} dx, \quad (5.7)$$

for all $x \in \mathbb{I}$. Practically, this means that the process does not move faster than a log-normal process for large values of X_t . We may now express the convergence result in terms of the step-size of the grid in terms of p -Wasserstein distances. In the following result, the p -Wasserstein distance \mathcal{W}_p between two laws μ and ν of processes with càdlàg⁴ paths is defined as

$$\mathcal{W}_p[\mu, \nu] = \inf_{(\zeta, \xi) \sim \Gamma(\mu, \nu)} \left\| \|\zeta - \xi\|_{\infty} \right\|_{L^p},$$

where by $\Gamma(\mu, \nu)$ we denote the collection of all measures with marginals μ and ν .

⁴This stands for “continue à droite avec une limite à gauche”, that is right-continuous with left-limit.

Algorithm 1 Space-Time Markov Chain Approximation (STMCA) Algorithm

Input: x initial value, T time horizon, $\mathbf{g} = \{x_j\}_{j \in J}$ grid on \mathbb{I}

Output: $(\widehat{X}[k])_k$ consecutive values taken by the approximation process, $(\hat{t}[k])_k$ consecutive transition times

Initialization:

$\hat{t}[0] = 0, n = 0$

$j = \arg \min_{i \in J} \{|x_i - x|\}$

$U \sim \text{Bernoulli}(p^+[x, (x_j, x_{j+1})])$

if $U == 1$ **then**

$j = j + 1$

end if

$\widehat{X}[0] = x_j$

Main loop:

while $\hat{t}[n] < T$ **do**

$U \sim \text{Bernoulli}(p^+[x_j, (x_{j-1}, x_{j+1})])$

if $U == 1$ **then**

$j = j + 1$

$\hat{t}[n + 1] = \hat{t}[n] + T^+[x_j, (x_{j-1}, x_{j+1})]$

else

$j = j - 1$

$\hat{t}[n + 1] = \hat{t}[n] + T^-[x_j, (x_{j-1}, x_{j+1})]$

end if

$\widehat{X}[n + 1] = x_j$

$n = n + 1$

end while

Theorem 5.2.1. *Let X be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} , on natural scale, whose speed measure satisfies Condition (5.7) for some constant $k_1 > 0$. Let \mathbf{g} be a covering grid of \mathbb{I} . Then, for all $p \geq 1$, $\delta \in (0, \frac{1}{4} \wedge \frac{1}{p})$, $T > 0$ and $x \in \mathbb{I}$ there exists a constant $C > 0$ such that*

$$\mathcal{W}_p \left[\text{Law} \left((\widehat{X}_t^{\mathbf{g}})_{t \in [0, T]} \right), \text{Law} \left((X_t)_{t \in [0, T]} \right) \right] \leq C |\mathbf{g}|_X^\delta, \quad (5.8)$$

where $|\mathbf{g}|_X = \sup_{\mathbf{c} \in \mathcal{C}(\mathbf{g})} \{|\mathbf{c}|m(\mathbf{c})\}$.

Remark. *In the case where $m(dx) \geq k_1 dx$, the constant⁵ $C > 0$ in Theorem 5.2.1 does not depend on the starting point of the diffusion.*

Remark. *If X is a diffusion process on natural scale such that (5.6) holds, the bound in (5.8) can be replaced by $C|\mathbf{g}|^{2\delta}$.*

The convergence of the Wasserstein distances implies the convergence in law [90, p. 109].

Corollary 5.2.2. *For all $T > 0$, the processes $(\widehat{X}_t^{\mathbf{g}})_{t \in [0, T]}$ converges in law to $(X_t)_{t \in [0, T]}$ in the Skorokhod space $D([0, T], \mathbb{I})$ as $|\mathbf{g}|_X \rightarrow 0$.*

⁵This results in the bound of Theorem 5.2.1 not depending on the starting point of the diffusion.

In Section 5.2.5, we observe that, by suitable choice of the probability space, the values taken by the approximation process correspond to values taken by X . We call this class of approximation schemes embeddable schemes (other embeddable schemes are the ones developed in [7, 31]). Proving the convergence of an embeddable scheme usually involves: embedding the approximation process in the trajectory of the target diffusion process and controlling the embedding times, controlling the speed at which the process moves.

5.2.2 Convergence rate for the general case

The convergence results established in the previous section are proven in the case of a diffusion process on natural scale. In this section, we show how more general results can be inferred. Let X be a diffusion process with state-space \mathbb{I} an open interval of \mathbb{R} , scale function s and speed measure m . We assume that

- s belongs to the Sobolev space $W^{1,1}(\mathbb{I})$, so from Theorem 8.2 of [17], as s is continuous,

$$s(x) - s(y) = \int_y^x s'(t) dt,$$

for all y, x in \mathbb{I} .

- there exists a $k_1 > 0$ such that for all $x \in \mathbb{I}$,

$$m(dx) \geq k_1 \frac{s'(x)}{1 + (s(x))^2} dx, \quad (5.9)$$

- the inverse of s is α -Hölder continuous, *i.e.*, there exists a constant $C > 0$ such that for all $x \neq y \in \mathbb{I}$,

$$\frac{|s^{-1}(\bar{x}) - s^{-1}(\bar{y})|}{|\bar{x} - \bar{y}|^\alpha} \leq C. \quad (5.10)$$

Given a grid \mathbf{g} , we consider the random walk $\widetilde{X}_t^{\mathbf{g}}$ defined by Algorithm 1, where the transition probabilities and transition times in (5.5) can be computed using the formulas derived in Section 5.3. We obtain the following corollary of Theorem 5.2.1.

Corollary 5.2.3. *Let X be a diffusion process with scale function and speed measure satisfying the above conditions. Let also \mathbf{g} be a covering grid over the state-space \mathbb{I} of X . Then, for all $p \geq 1$, $\delta \in (0, \frac{1}{4} \wedge \frac{1}{p})$, $T > 0$ and $x \in \mathbb{I}$ there exists positive constants C_1 , and C_2 such that*

$$\mathcal{W}_p \left[\text{Law} \left((\widetilde{X}_t^{\mathbf{g}})_{t \in [0, T]} \right), \text{Law} \left((X_t)_{t \in [0, T]} \right) \right] \leq C_1 |\mathbf{g}|_X^\delta,$$

where $|\mathbf{g}|_X$ is defined in (5.1).

Proof. We define the proxy process $Y = (s(X_t))_{t \geq 0}$ which has state-space $s(\mathbb{I})$, scale function $s_Y(x) = x$ and speed measure $m_Y(dx) = m \circ s^{-1}(dx)$. From condition (5.9) and a change of variables, we get that $(Y_t)_{t \geq 0}$ satisfies condition (5.7) for the same constant k_1 . We also define $\widetilde{Y}_t^{s(\mathbf{g})}$ evolving according to Algorithm 1, with covering

grid $s(\mathbf{g}) = \{s(x); x \in \mathbf{g}\}$. It can be defined on the canonical space of X so that $s(\widetilde{X}_t^{\mathbf{g}}) = \widetilde{Y}_t^{s(\mathbf{g})}$ almost surely. Thus, Condition (5.10) implies that

$$|X_t - \widetilde{X}_t^{\mathbf{g}}| \leq C|Y_t - \widetilde{Y}_t^{s(\mathbf{g})}|^\alpha.$$

Along with the fact that

$$|\mathbf{g}|_Y = \sup_{\mathbf{c} \in \mathcal{C}(\mathbf{g})} \{|s(\mathbf{c})| m_Y(s(\mathbf{c}))\},$$

Theorem 5.2.1 implies Corollary 5.2.3. □

5.2.3 Grid tuning

We observe that for all $\epsilon > 0$, in the case of a Brownian motion, Theorem 5.2.1 yields a convergence rate of $\mathcal{O}(|\mathbf{g}|^{\frac{1}{2}-\epsilon})$ as $|\mathbf{g}| \rightarrow 0$, which is optimal from Donsker's invariance principle [26]. We would like to have this result for all diffusion processes, but the following example illustrates that this is not the case. We then show how we can remediate to this by using a custom grid and extrapolate this method to the general case via Corollary 5.2.5.

Example 5.2.4. Let X be the diffusion process with state-space \mathbb{R} , defined through s and m with

$$s(x) = x, \quad m(dx) = 2dx + \rho\delta_0(dx).$$

This process is called the *sticky Brownian motion* and is the “most elementary” sticky diffusion process. As such, it spends a positive amount of time at 0 and the Euler scheme is known to not be well defined for these processes. We observe that, for any covering grid \mathbf{g} of \mathbb{R} and $\mathbf{c} \in \mathcal{C}(\mathbf{g})$ with $\mathbf{c} \neq \mathbf{c}_0$ and $\mathbf{c}_0 \in \mathcal{C}(\mathbf{g})$ being the cell containing 0,

$$\begin{aligned} 2|\mathbf{c}|^2 = m(\mathbf{c})|\mathbf{c}| &\leq \sup_{\mathbf{c} \in \mathcal{C}(\mathbf{g})} m(\mathbf{c})|\mathbf{c}|, \\ \rho|\mathbf{c}_0| + 2|\mathbf{c}_0|^2 = m(\mathbf{c}_0)|\mathbf{c}_0| &\leq \sup_{\mathbf{c} \in \mathcal{C}(\mathbf{g})} m(\mathbf{c})|\mathbf{c}|. \end{aligned} \tag{5.11}$$

From (5.11), for a uniform grid of step size h ,

$$2h\rho < |\mathbf{g}|_X.$$

Thus, for any $\epsilon > 0$, the convergence rate given by Theorem 5.2.1 is $\mathcal{O}(h^{\frac{1}{4}-\epsilon})$ as $|\mathbf{g}| \rightarrow 0$. This means that there are functionals of the trajectory for which the convergence rate is much slower for this process in comparison with a standard Brownian motion. In order to remediate to this, we propose a preliminary step to the approximation scheme that involves finding a grid that is *adapted* to the speed measure of the process. In the case of the Brownian motion with a sticky point at 0, such a grid can be defined as one that has uniform non-adjacent cells to 0 of size h and with the cell of center 0 being $(-h^2/2\rho, h^2/2\rho)$, i.e.

$$\mathbf{g} = \left\{ \bigcup_{k \in \mathbb{Z}_+} \left\{ -\frac{h^2}{2\rho} - k\frac{h}{2} \right\} \right\} \cup \{0\} \cup \left\{ \bigcup_{k \in \mathbb{Z}_+} \left\{ \frac{h^2}{2\rho} + k\frac{h}{2} \right\} \right\}. \tag{5.12}$$

As the approximation process is a random walk, for every k steps it makes, it spends $\mathcal{O}(\sqrt{k})$ steps in the cell containing 0 (see [22]). Thus, running the algorithm on either grid yields the same algorithmic complexity, while the convergence rate is improved to $\mathcal{O}(h^{\frac{1}{2}-\epsilon})$ for the adapted grid. Numerical examples are given in Section 6.1.2.

The general case is covered by the following Corollary:

Corollary 5.2.5. *Let X be a diffusion process and $\widetilde{X}^{\mathbf{g}}$ the approximation process defined by Algorithm 1. Then, if \mathbf{g} is a grid such that*

$$|\mathbf{g}|_X \leq C|\mathbf{g}|^2, \quad (5.13)$$

we can bound the p -Wasserstein distance between the laws of $(\widetilde{X}_t^{\mathbf{g}})_{t \in [0, T]}$ and $(X_t)_{t \in [0, T]}$ in Theorem 5.2.1 by $|\mathbf{g}|^{2\delta}$ instead of $|\mathbf{g}|_X^\delta$. Thus, for all $\epsilon > 0$, the law of the random walk converges in any p -Wasserstein distance at the rate $\mathcal{O}(|\mathbf{g}|^{(\frac{1}{2} \wedge \frac{2}{p})-\epsilon})$ instead of $\mathcal{O}(|\mathbf{g}|^{(\frac{1}{4} \wedge \frac{1}{p})-\epsilon})$ as $|\mathbf{g}| \rightarrow 0$.

We now show how one can create grids such that (5.13) holds in the case of homogeneous SDEs. Considering sticky and/or skew points is straightforward.

- For a process whose speed measure satisfies (5.6) and (5.9) with one or more skew points, no grid modification is required,
- For a process whose speed measure satisfies (5.6) and (5.9) and has a sticky point at 0 of stickiness $\rho > 0$, one needs to consider the points $\{-h^2/\rho, 0, h^2/\rho\}$ to have a tuned grid,
- The case of a reflection at a boundary is treated in Section 5.2.4.

Let (μ, σ) be a pair of functions satisfying the following condition:

Condition 5.2.6. *The functions are measurable $\mathbb{R} \mapsto \mathbb{R}$ mappings and the SDE*

$$dX_t = \mu(X_t) dt + \sigma(X_t) dB_t, \quad (5.14)$$

has a unique weak solution, where B is a standard Brownian motion.

Let X be the diffusion that solves (5.14), \mathbb{I} its state-space, s its scale function and m its speed measure given by [14, p. 17]

$$s'(x) = e^{-\int_a^x \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta}, \quad \text{and} \quad m(dx) = \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} dx,$$

with s' being the right-derivative of s . Then, if $\mathbf{g} = \{x_j\}_{j \in J}$ is a covering grid such that for a constant $C > 0$ and every $j \in J$

$$\left| \left(s(x_{j+1}) - s(x_{j-1}) \right) \int_{x_{j-1}}^{x_{j+1}} \frac{2}{s'(\zeta)\sigma^2(\zeta)} d\zeta \right| \wedge (x_{j+1} - x_{j-1})^2 \leq Ch^2,$$

it satisfies (5.13) and $|\mathbf{g}| = Ch$. Thus, from Corollary 5.2.5, for any $\epsilon > 0$ using such grids give us a convergence rate of $\mathcal{O}(|\mathbf{g}|^{\frac{1}{2}-\epsilon})$ in Theorem 5.2.1.

Generating such grids numerically can be done choosing a starting point x_0 and adding points x_j to the grid iteratively as follows: given x_{j-1} , let x_j be chosen such that:

$$\left| \left(s(y) - s(x_{j-1}) \right) \int_{x_{j-1}}^y \frac{2}{s'(\zeta)\sigma^2(\zeta)} d\zeta \right| \leq h^2/2. \quad (5.15)$$

Then the next element of the grid is defined as

$$x_j = \begin{cases} x_j^{(0)} & \text{if } x_j^{(0)} - x_{j-1} \leq h, \\ x_{j-1} + h & \text{if } x_j^{(0)} - x_{j-1} > h. \end{cases} \quad (5.16)$$

The problem (5.15) can be solved numerically using a fixed point algorithm. Examples of STMCA simulations using tuned grids computed solving (5.15)-(5.16) are given in Figures 6.2 and 6.3.

5.2.4 The case of diffusions with boundary conditions

When presenting the results and the structure of the scheme, we considered only processes where \mathbb{I} is an open set, thus considering diffusion with unreachable boundaries. Our results also adapt to the situation where either ℓ and/or r are reachable, and in this case some adjustments are needed, depending on the nature of finite boundaries and on the condition at regular boundaries. In order to keep the presentation simple, we assume that the process is on natural scale and that $\mathbb{I} = [0, +\infty)$ (the adaptation to $\mathbb{I} = (\ell, r]$ or $\mathbb{I} = [\ell, r]$ or $\mathbb{I} = [\ell, r)$ with $\ell \in \mathbb{R} \cup \{-\infty\}$ and $r \in \mathbb{R} \cup \{+\infty\}$ is straightforward).

It is well known (see for instance Section 5.11 of Itô's book [52]) that the finite boundary 0 can be of four types. Setting, for some fixed $c > 0$,

$$\mathcal{I} = \int \int_{0 < y < x < c} m(dx) dy, \quad \mathcal{II} = \int \int_{0 < y < x < c} m(dy) dx,$$

then

- 0 is an exit boundary if $\mathcal{I} < \infty$ and $\mathcal{II} = \infty$,
- 0 is a regular boundary if $\mathcal{I} < \infty$ and $\mathcal{II} < \infty$,
- 0 is a natural boundary if $\mathcal{I} = \infty$ and $\mathcal{II} = \infty$,
- 0 is an entrance boundary if $\mathcal{I} = \infty$ and $\mathcal{II} < \infty$.

The entrance type been excluded for a finite boundary of a diffusion process on natural scale, and the natural type been considered in the settings of Theorem 5.2.1, this leaves us with two possible types for the boundary 0: exit or regular. If 0 is an exit boundary, then the diffusion process X is absorbed at the boundary 0. If 0 is a regular boundary, then the diffusion process can either be absorbed or reflected at 0. In these cases, the convergence result of Theorem 5.2.1 can be extended by considering a grid \mathbf{g} on \mathbb{I} containing 0 and by adapting the dynamics of $\widetilde{X}^{\mathbf{g}}$ as follows. The dynamic of $\widetilde{X}^{\mathbf{g}}$ is the same as in Algorithm 1, up to the time when it reaches 0, then:

- if 0 is an absorbing boundary (exit or regular), then the result can be immediately extended by stopping $\widetilde{X}^{\mathbf{g}}$ when it reaches 0;

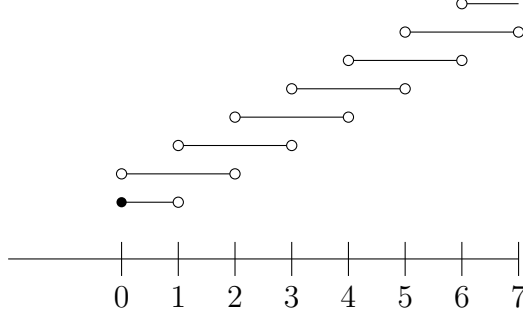


Figure 5.3: The covering grid $\mathbf{g} = \mathbb{Z}^+$ of $[0, \infty)$ along with its first cells $\mathcal{C}(\mathbf{g})$.

- if 0 is a reflecting regular boundary, then the process $\widetilde{X}^{\mathbf{g}}$ jumps from 0 to $b := \min \mathbf{g} \setminus \{0\}$ with probability 1 and after a time $\int_{[0,b)} (b - \zeta) m(d\zeta)$. We emphasize that in this configuration, 0 may be a sticky boundary (*i.e.*, with $m(0) \in (0, +\infty)$).

In both cases of reflection and absorption at 0, the boundary is attainable. Thus, 0 must be a point of any covering grid of $\mathbb{I} = [0, \infty)$. This leads to the following adaptation of the notion of grid cells. The cells $\mathcal{C}(\mathbf{g})$ of such grid \mathbf{g} are the open intervals for the induced topology of \mathbb{R} on \mathbb{I} with endpoint in \mathbf{g} . For example $\mathbf{g}_0 = \mathbb{Z}_+$ is a covering grid of $[0, \infty)$ and $\mathcal{C}(\mathbf{g}_0) = \{[0, 1), (n - 1, n + 1)_{n \in \mathbb{N}}\}$.

The proof of the convergence in these situations is omitted here, since it is a straightforward adaptation of the proof of Theorem 5.2.1, using in particular the fact that, in the case of a reflecting boundary,

$$\mathbb{E}_0 [\tau_b] = \int_{[0,b)} (b - \zeta) m(d\zeta).$$

The case of killing boundaries, and in general of a process with non-zero killing measure, leads to additional non-trivial difficulties. Devising an algorithm and a similar result as Theorem 5.2.1 for such processes remains an active area of research.

5.2.5 Markovian embedding

The consecutive values of the process $(\widetilde{X}_t)_{t \geq 0}$ defined through (5.2)-(5.4) form a Markov chain with, by construction, the same transition probabilities as $(X_t)_{t \geq 0}$ on \mathbf{g} . We define the embedding times of $(X_t)_{t \geq 0}$ in \mathbf{g} as,

$$\begin{cases} \tau_0^{\mathbf{g}} &= 0, \\ \tau_k^{\mathbf{g}} &= \inf \{t > \tau_{k-1}^{\mathbf{g}} : X_t \in \mathbf{g} \setminus \{X_{\tau_{k-1}^{\mathbf{g}}}\}\}, \quad \forall k \geq 1. \end{cases} \quad (5.17)$$

As both $\widetilde{X}_{T^{\mathbf{g}}(n)}$ and $X_{\tau_n^{\mathbf{g}}}$ are both Markov chains with the same transition probabilities with \widetilde{X}_0 forced to be equal in law to $X_{\tau_1^{\mathbf{g}}}$ (see Section 5.2.1), the following equality in law holds,

$$\text{Law}(\widetilde{X}_{T^{\mathbf{g}}(n)}; n \geq 0) = \text{Law}(X_{\tau_n^{\mathbf{g}}}; n \geq 1).$$

We define $K^{\mathbf{g}}(t)$ as the inverse of $T^{\mathbf{g}}(n)$, *i.e.*,

$$K^{\mathbf{g}}(t) = \inf \left\{ n \in \mathbb{N} : \sum_{k=1}^n \mathbb{E} \left[\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] > t \right\}. \quad (5.18)$$

Thus, we get the following Proposition.

Proposition 5.2.7. *Let $(X_t)_{t \geq 0}$ be a diffusion process, \mathbf{g} a grid defined over its state space \mathbb{I} and $\widetilde{X}_t^{\mathbf{g}}$ be the approximation process defined in (5.3) and (5.4). Then, if $(\tau_n^{\mathbf{g}})_{n \geq 0}$ are the embedding times of $(X_t)_{t \geq 0}$ in \mathbf{g} , the following equality in law holds,*

$$\text{Law}(\widetilde{X}_t; t \geq 0) = \text{Law}(X_{\tau_{K^{\mathbf{g}}(t)}}^{\mathbf{g}}; t \geq 0),$$

where $K^{\mathbf{g}}(t)$ is the random index defined in (5.18).

5.3 Moment characterization of conditional exit times

The law of the approximation process defined in the previous section was shown to be determined by the transition probabilities $P_x(\tau_b < \tau_a)$ and conditional transition times $E_x(\tau_{ab}^k | \tau_b < \tau_a)$. In this section we show that quantities of the form $v_k(x) = E_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$ yield an integral formulation with respect to the speed measure of the diffusion and involving the scale function (we do not assume that the diffusion is on natural scale in the present section). We also show that this results in them being solutions to Dirichlet problems where the differential operator is the infinitesimal generator L of the diffusion. This allows us to simulate such processes via Algorithm 1 and thus to approximate the law of the target diffusion process $(X_t)_{t \geq 0}$.

In terms of Algorithm 1, we need to compute for three adjacent points a, x, b of the grid the quantities

$$v_0(x) = P_x(\tau_b < \tau_a), \quad v_1(x) = E_x(\tau_{ab} \mathbb{1}_{\tau_b < \tau_a}) \quad \text{and} \quad \bar{v}_1(x) = E_x(\tau_{ab} \mathbb{1}_{\tau_a < \tau_b}).$$

The quantities of (5.5) are then

$$\begin{aligned} p^+[x, (a, b)] &= v_0(x), & T^+[x, (a, b)] &= \frac{v_1(x)}{v_0(x)}, \\ p^-[x, (a, b)] &= 1 - v_0(x), & T^-[x, (a, b)] &= \frac{\bar{v}_1(x)}{1 - v_0(x)}. \end{aligned}$$

Proposition 5.3.1. *The function $v_0(x) = P_x(\tau_b < \tau_a)$ is solution to the problem with Dirichlet boundary conditions*

$$\begin{cases} Lu = 0, & x \in (a, b), \\ u(a) = 0, \\ u(b) = 1, \end{cases} \quad (5.19)$$

where L is the infinitesimal generator of $(X_t)_{t \geq 0}$, which also implies that $v_0 \in \text{dom}(L)$.

Proof. Let $x \in (a, b)$, from the definition of the scale function and the factorization of the infinitesimal generator $L = D_m D_s$

$$L v_0 = D_m D_s \frac{s(\cdot) - s(a)}{s(b) - s(a)} = D_m \frac{1}{s(b) - s(a)}.$$

which equals 0 as $m(dx)$ is a positive measure. As v_0 and $L v_0 = 0$ are both functions in C_b^0 , we deduce that $v_0 \in \text{dom}(L)$ and $L v_0 = 0$. Under P_b , the stopping time τ_b equal 0 a.s. and the process has a.s. continuous trajectories, hence $\tau_a > 0$ a.s., *i.e.*,

$$v_0(b) = P_b(\tau_b < \tau_a) = P_b(0 < \tau_a) = 1.$$

This, along with the symmetrical argument, allow us to retrieve the boundary conditions of (5.19). \square

Proposition 5.3.2. *For every $k \in \mathbb{N}$, let v_k be the function defined for every $x \in (a, b)$ by $v_k(x) = E_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$. Then,*

$$v_k(x) = k \int_{(a,b)} G_{a,b}(x, \zeta) v_{k-1}(\zeta) m(d\zeta). \quad (5.20)$$

Proof. Since $\int_0^{\tau_{ab}} (\tau_{ab} - t)^{k-1} dt = \tau_{ab}^k / k$,

$$v_k(x) = k E_x \left[\mathbb{1}_{\tau_b < \tau_a} \int_0^{\tau_{ab}} (\tau_{ab} - t)^{k-1} dt \right] = k E_x \left[\mathbb{1}_{\tau_b < \tau_a} \int_0^\infty \mathbb{1}_{t \leq \tau_{ab}} (\tau_{ab} - t)^{k-1} dt \right].$$

From the Markov property, by conditioning on \mathcal{F}_t and as $\mathbb{1}_{t \leq \tau_{ab}}$ is \mathcal{F}_t -measurable,

$$v_k(x) = k E_x \left[\int_0^\infty \mathbb{1}_{t \leq \tau_{ab}} E \left[\mathbb{1}_{\tau_b < \tau_a} (\tau_{ab} - t)^{k-1} | \mathcal{F}_t \right] dt \right] = k E_x \left[\int_0^{\tau_{ab}} E_{X_t} \left[\mathbb{1}_{\tau_b < \tau_a} \tau_{ab}^{k-1} \right] dt \right].$$

The equality (5.20) results by applying directly Green's formula. \square

Proposition 5.3.3. *The function $v_k(x) = E_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$ is solution to the problem with Dirichlet boundary conditions*

$$\begin{cases} L u = -k v_{k-1}, & x \in (a, b), \\ u(a) = 0, \\ u(b) = 0. \end{cases}$$

Lemma 5.3.4. *Let $g(x) = \int_{(a,b)} G_{a,b}(x, y) f(y) m(dy)$, where $f \in C_b^0(a, b)$ and $G_{a,b}(x, y)$ is the Green function defined in (1.20). Then $g \in \text{dom}(L)$ and*

$$L g(x) = -f(x), \quad \forall x \in (a, b).$$

Proof. Let $x \in (a, b)$. Using the $D_m D_s$ factorization of L and the dominated convergence theorem we get

$$\begin{aligned} & L g(x) \\ &= D_m D_s \int_{(a,b)} \left[\mathbb{1}_{y < x} \frac{(s(y) - s(a))(s(b) - s(x))}{s(b) - s(a)} + \mathbb{1}_{y \geq x} \frac{(s(x) - s(a))(s(b) - s(y))}{s(b) - s(a)} \right] f(y) m(dy) \\ &= -D_m \int_{(a,x)} v_0(y) f(y) m(dy) + D_m \int_{[x,b)} (1 - v_0(y)) f(y) m(dy) \\ &= -v_0(x) f(x) - (1 - v_0(x)) f(x) = -f(x). \end{aligned}$$

The continuity of g is a consequence of Lebesgue's convergence theorem for integrals. Moreover as $G_{a,b}(x, y)$ is bounded by $s(b) - s(a)$, $m(dx)$ is locally finite and f is bounded, g is also bounded. So, we deduce that $f \in \text{dom}(L)$ and that on (a, b) we have $L g = -f$. \square

Proof of Proposition 5.3.3. From Proposition 5.3.2,

$$v_k(x) = \int_{(a,b)} G_{a,b}(x,y) k v_{k-1}(y) m(dy).$$

As $v_0 \in C_b^0(a,b)$, from Lemma 5.3.4 we deduce iteratively that $v_k \in C_b^0(a,b)$ for all $k \in \mathbb{N}$ and that $L v_k = -k v_{k-1}$ on (a,b) . For the boundary conditions, we observe that $\tau_{ab} = \tau_a \wedge \tau_b = 0$ a.s. under P_a , so for $k \geq 1$

$$v_k(a) = E_a(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a}) \leq E_a(\tau_{ab}^k) = 0.$$

With the same argument we show that $v_k(b) \leq 0$ and as they are obviously positive quantities $v_k(a) = v_k(b) = 0$. \square

5.4 Convergence of the embedding times

In order to prove the convergence of the process $(\widetilde{X}_t^{\mathbf{g}})_{t \geq 0}$ we need to control quantities of the form $E_x \left[\sup_{t \leq T} |t - \tau_{K\mathbf{g}(T)}^{\mathbf{g}}|^p \right]$, where $\tau_k^{\mathbf{g}}$ are the embedding times of the process X in the grid \mathbf{g} . In this section, we show the existence of such bounds in terms of the metric

$$|\mathbf{g}|_X = \sup_{\mathbf{c} \in \mathcal{C}(\mathbf{g})} \{s(\mathbf{c})m(\mathbf{c})\}$$

defined in (5.1). If there exists a constant $K > 0$ such that $m(\mathbf{c}) \leq K$ for all $\mathbf{c} \in \mathcal{C}(\mathbf{g})$, then, $|\mathbf{c}|m(\mathbf{c}) \leq K|\mathbf{c}|$ and $|\mathbf{g}|_X \leq K|\mathbf{g}|$. Thus, we can bound the quantities of interest in terms of $|\mathbf{g}|$ instead of $|\mathbf{g}|_X$. But, in doing so, we do not track correctly the convergence rates of the approximation process. For example, for the standard Brownian motion, $|\mathbf{g}|_B = |\mathbf{g}|^2$. Moreover, as shown in Section 5.2.3, such bounds give us a direct way to adapt the grid to the speed measure in order to accelerate the convergence of the scheme.

5.4.1 Bounds on the conditional moments of the exit times

Lemma 5.4.1. *Let $v_k(x) = E_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$, then for all $k \in \mathbb{N}$*

$$\left\| \frac{v_k}{v_{k-1}} \right\|_{\infty} \leq k |\mathbf{g}|_X. \quad (5.21)$$

Proof. We first observe that

$$G_{a,b}(x,y) \frac{v_0(y)}{v_0(x)} = \begin{cases} \frac{(y-a)(b-x)}{(b-a)} \frac{y-a}{x-a}, & x > y, \\ \frac{(x-a)(b-y)}{(b-a)} \frac{y-a}{x-a}, & x \leq y. \end{cases}$$

As for $x > y$ the ratio $\frac{y-a}{x-a} < 1$,

$$G_{a,b}(x,y) \frac{v_0(y)}{v_0(x)} \leq \begin{cases} \frac{(y-a)(b-x)}{(b-a)}, & x > y, \\ \frac{(y-a)(b-y)}{(b-a)}, & x \leq y, \end{cases}$$

which is bounded by $(b - a)$ in both cases as $x, y \in (a, b)$. Thus,

$$\int_a^b G_{a,b}(x, y)v_0(y)m(dy) = v_0(x) \int_a^b G_{a,b}(x, y) \frac{v_0(y)}{v_0(x)} m(dy) \leq v_0(x)(b - a)m((a, b)). \quad (5.22)$$

From Proposition 5.3.2, $v_k(x) = k \int_{(a,b)} G_{a,b}(x, y)v_{k-1}(y)m(dy)$ and

$$v_k(x) = k! \int_{(a,b)} G_{a,b}(x, x_k) \int_{(a,b)} G_{a,b}(x_k, x_{k-1}) \cdots \int_{(a,b)} G_{a,b}(x_2, x_1)v_0(x_1)m(dx_1) \dots m(dx_k). \quad (5.23)$$

From (5.22) and (5.23),

$$v_k(x) \leq k(b - a)m((a, b))v_{k-1}(x).$$

Since $(b - a)m((a, b)) \leq |\mathbf{g}|_X$, we get the desired result on v_k/v_{k-1} . \square

Corollary 5.4.2. *Let $k \in \mathbb{N}$. Then, we have the following bound for $v_k(x) = \mathbb{E}_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$,*

$$\|v_k\|_\infty \leq k!|\mathbf{g}|_X^k.$$

Lemma 5.4.3. *Let X be a diffusion process with state space \mathbb{I} an interval of \mathbb{R} , m the speed measure of X , $a, x, b \in \mathbb{I}$ such that $a < x < b$, $\mathbf{c}_x = (a, b)$, $|\mathbf{c}_x|_X = m(\mathbf{c}_x)|b - a|$ and $\lambda > 0$ such that $\lambda|\mathbf{c}_x|_X \in (0, 1)$. Then,*

$$\mathbb{E}_x(e^{\lambda\tau_{ab}}) \leq \exp\left(\frac{\lambda}{1 - \lambda|\mathbf{c}_x|_X} \mathbb{E}_x(\tau_{ab})\right). \quad (5.24)$$

Proof. Developing the exponential series,

$$\mathbb{E}_x(e^{\lambda\tau_{ab}}) = 1 + \sum_{N=1}^{\infty} \frac{\lambda^N}{N!} \mathbb{E}_x(\tau_{ab}^N) = 1 + \mathbb{E}_x(\tau_{ab}) \sum_{N=1}^{\infty} \frac{\lambda^N}{N!} \frac{\mathbb{E}_x(\tau_{ab}^N)}{\mathbb{E}_x(\tau_{ab})}.$$

From Corollary 5.4.2, we can bound the ratio of expected values by $(N - 1)!|\mathbf{c}_x|_X^{N-1}$. As $\lambda|\mathbf{c}_x|_X \in (0, 1)$,

$$\mathbb{E}_x(e^{\lambda\tau_{ab}}) \leq 1 + \mathbb{E}_x(\tau_{ab}) \sum_{N=1}^{\infty} \frac{\lambda}{N} (\lambda|\mathbf{c}_x|_X)^{N-1} \leq 1 + \mathbb{E}_x(\tau_{ab}) \frac{\lambda}{1 - \lambda|\mathbf{c}_x|_X}.$$

Thus, we only need to apply the inequality $1 + x \leq e^x$ to get (5.24). \square

Lemma 5.4.4. *Let $t, M > 0$ and $\lambda > 0$ such that $\lambda|\mathbf{g}|_X \in (0, 1)$. Then,*

$$\mathbb{P}_x(\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} > M) \leq e^{-\lambda M} e^{\lambda \frac{t + |\mathbf{g}|_X}{1 - \lambda|\mathbf{g}|_X}}. \quad (5.25)$$

Proof. From Markov's inequality,

$$\mathbb{P}_x(\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} > M) \leq e^{-\lambda M} \mathbb{E}_x \left[e^{\lambda \tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} \right] = e^{-\lambda M} \mathbb{E}_x \left[e^{\lambda \sum_{k=1}^{K^{\mathbf{g}}(t)} (\tau_k - \tau_{k-1})} \right].$$

Conditioning on the σ -algebra \mathcal{B} generated by the trajectory of X_t on the grid \mathbf{g} , *i.e.*, $\mathcal{B} = \sigma\{X_{\tau_k}; k \in \mathbb{N}_0\}$, and as $K^{\mathbf{g}}(t)$ is \mathcal{B} -measurable,

$$\mathbb{E}_x \left[e^{\lambda \sum_{k=1}^{K^{\mathbf{g}}(t)} (\tau_k - \tau_{k-1})} \right] = \mathbb{E}_x \left[\prod_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E} \left[e^{\lambda(\tau_k - \tau_{k-1})} \middle| \mathcal{B} \right] \right]. \quad (5.26)$$

From the definition of $|\mathbf{g}|_X$, $\lambda|\mathbf{c}|_X \leq \lambda|\mathbf{g}|_X \in (0, 1)$ for each cell \mathbf{c} of the grid \mathbf{g} . Thus, applying Lemma 5.4.3 on each term of the product in (5.26),

$$\mathbb{P}_x(\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} > M) \leq e^{-\lambda M} \mathbb{E}_x \left[\exp \left(\frac{\lambda}{1 - \lambda|\mathbf{g}|_X} \sum_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E}(\tau_k - \tau_{k-1} | \mathcal{B}) \right) \right] \quad (5.27)$$

From (5.18),

$$t < \sum_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E}(\tau_k - \tau_{k-1} | \mathcal{B}) \leq t + \mathbb{E}(\tau_{K^{\mathbf{g}}(t)+1} - \tau_{K^{\mathbf{g}}(t)} | \mathcal{B})$$

Thus, from (5.21),

$$\sum_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E}(\tau_k - \tau_{k-1} | \mathcal{B}) \leq t + \|v_1/v_0\|_{\infty} \leq t + |\mathbf{g}|_X, \quad (5.28)$$

From (5.27) and (5.28), we get (5.25). \square

5.4.2 Convergence of the embedding times

For this section, let X be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} , \mathbf{g} a covering grid of \mathbb{I} , $\tau_k^{\mathbf{g}}$ the embedding times of X in \mathbf{g} as defined in (5.17) and $K^{\mathbf{g}}(t)$ as defined in (5.18).

Lemma 5.4.5. *For any $T > 0$,*

$$\sum_{k=1}^{K^{\mathbf{g}}(T)} \text{Var} \left(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right) \leq 2|\mathbf{g}|_X (T + |\mathbf{g}|_X),$$

where $|\cdot|_X$ is defined in (5.1).

Proof. For all $x \in \mathbf{g}$, let \mathbf{c}_x be the cell of \mathbf{g} containing x . Then,

$$\begin{aligned} & \sum_{k=1}^{K^{\mathbf{g}}(T)} \text{Var} \left(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right) \\ &= \sum_{k=1}^{K^{\mathbf{g}}(T)} \mathbb{E} \left[(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}})^2 \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] - \left(\mathbb{E} \left[\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \right)^2 \\ &\leq \sup_{k \leq K^{\mathbf{g}}(T)} \left\{ \frac{\mathbb{E} \left[(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}})^2 \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right]}{\mathbb{E} \left[\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right]} \right\} \sum_{k=1}^{K^{\mathbf{g}}(T)} \mathbb{E} \left[(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}}) \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right]. \end{aligned}$$

So from Lemma 5.4.1 and the definition of $K^{\mathbf{g}}(t)$,

$$\begin{aligned} & \sum_{k=1}^{K^{\mathbf{g}}(T)} \text{Var} \left(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right) \\ &\leq \left\| \frac{v_2}{v_1} \right\|_{\infty} \sum_{k=1}^{K^{\mathbf{g}}(T)} \mathbb{E} \left[(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}}) \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \\ &\leq \left\| \frac{v_2}{v_1} \right\|_{\infty} \left(T + \left\| \frac{v_1}{v_0} \right\|_{\infty} \right) \leq 2|\mathbf{g}|_X (T + |\mathbf{g}|_X), \end{aligned}$$

which is the desired inequality. \square

Proposition 5.4.6. *Let $(\mathcal{F}_t)_{t \geq 0}$ be the canonical filtration of X . Let also $\mathcal{A}_n = \mathcal{F}_{\tau_n^{\mathbf{g}}}$, $\mathcal{B} = \sigma((X_{\tau_k^{\mathbf{g}}})_{k \in \mathbb{N}_0})$ and $\Delta\tau_k^{\mathbf{g}} = \tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}}$. If we define the augmented filtration $\mathcal{G}_n = \mathcal{A}_n \vee \mathcal{B}$, then the process*

$$M_n = \sum_{k=1}^n \Delta\tau_k^{\mathbf{g}} - \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right],$$

is a \mathcal{G}_n -martingale.

Proof. For $m \leq n$,

$$\begin{aligned} \mathbb{E} [M_n | \mathcal{G}_m] &= \mathbb{E} [M_n | \mathcal{A}_m, \mathcal{B}] \\ &= \mathbb{E} \left[\sum_{k=1}^m \Delta\tau_k^{\mathbf{g}} - \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] + \sum_{k=m+1}^n \Delta\tau_k^{\mathbf{g}} - \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \middle| \mathcal{A}_m, \mathcal{B} \right]. \end{aligned}$$

From Lemma A.4.1 and as $\tau_k^{\mathbf{g}}$ is \mathcal{A}_k measurable,

$$\mathbb{E} [M_n | \mathcal{G}_m] = M_m + \sum_{k=m+1}^n \mathbb{E}_{X_{\tau_m^{\mathbf{g}}}} \left[\Delta\tau_k^{\mathbf{g}} - \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \middle| \mathcal{B} \right] = M_m.$$

This proves the result. \square

Theorem 5.4.7. *For any $T > 0$,*

$$\mathbb{E} \left[\sup_{t \in [0, T]} |\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - t|^2 \right] \leq 2|\mathbf{g}|_X \left(4(T + |\mathbf{g}|_X) + 1 \right), \quad (5.29)$$

where $|\cdot|_X$ is defined in (5.1).

Proof. The convexity inequality $(a + b)^p \leq 2^{p-1}(a^p + b^p)$ yields for $p = 2$

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0, T]} |\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - t|^2 \right] &\leq 2 \mathbb{E} \left[\sup_{t \in [0, T]} \left| \tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - \sum_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \right|^2 \right] \\ &\quad + 2 \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] - t \right|^2 \right]. \quad (5.30) \end{aligned}$$

From Proposition 5.4.6, the term $\sum_{k=1}^n \Delta\tau_k^{\mathbf{g}} - \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right]$ is a \mathcal{G}_n -martingale, where $\mathcal{G}_n = \mathcal{F}_{\tau_n^{\mathbf{g}}} \vee \mathcal{B}$ and $\mathcal{B} = \sigma((X_{\tau_k^{\mathbf{g}}})_{k \in \mathbb{N}_0})$. Thus, from Doob's L^p inequality,

$$\mathbb{E} \left[\sup_{t \in [0, T]} |M_{K^{\mathbf{g}}(t)}|^2 \right] = \mathbb{E} \left[\sup_{k \leq K^{\mathbf{g}}(T)} |M_k|^2 \right] \leq 2 \mathbb{E} \left[|M_{K^{\mathbf{g}}(T)}|^2 \right].$$

By conditioning on \mathcal{B} , from Lemma A.4.1,

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{k=1}^{K^{\mathbf{g}}(t)} \Delta\tau_k^{\mathbf{g}} - \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \right|^2 \right] \\ \leq 2 \mathbb{E} \left[\left(\sum_{k=1}^{K^{\mathbf{g}}(T)} \Delta\tau_k^{\mathbf{g}} - \mathbb{E} \left[\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \right)^2 \right] \\ = 2 \mathbb{E} \left[\sum_{k=1}^{K^{\mathbf{g}}(T)} \text{Var} \left(\Delta\tau_k^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right) \right], \end{aligned}$$

which from Lemma 5.4.5 is bounded by $4|\mathbf{g}|_X(T + |\mathbf{g}|_X)$. For the second term on the right hand side of (5.30), from (5.18) and since $K^{\mathbf{g}}(t) \geq 1$ for any $t > 0$,

$$\begin{aligned} \sum_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E} \left[\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \mid X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] - t &\leq \mathbb{E} \left[\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - \tau_{K^{\mathbf{g}}(t)-1}^{\mathbf{g}} \mid X_{\tau_{K^{\mathbf{g}}(t)-1}^{\mathbf{g}}}, X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} \right] \\ &\leq \left\| \frac{v_1}{v_0} \right\|_{\infty} \leq |\mathbf{g}|_X. \end{aligned}$$

So having bounded both additive parts of the right hand side of (5.30), we get (5.29). \square

5.5 Convergence rate of the Markov chain

5.5.1 Moment bounds

In this section we prove that Theorem 3.1 of [9] holds also for reflected processes and for a sharper constant. This result, combined with the bound (5.29) is used to prove the convergence of the approximation process in Section 5.5.2.

Lemma 5.5.1. *Let X be a diffusion process on natural scale with state-space \mathbb{R} and a speed measure m_X that satisfies Condition (5.7). Then, for all $p \geq 2$, there exist two constants $C, C' > 0$ such that*

$$\mathbb{E}_x \left[\sup_{t \in [0, T]} |X_t - x|^p \right] \leq C' [1 + |x|^p] e^{CT}, \quad (5.31)$$

where k_1 is a constant such that Condition (5.7) is satisfied, $C \leq 8p(p-1)/k_1$ and $C' > 0$ is a constant that depends only on p .

Proof. Let Z be the diffusion process on natural scale with speed measure,

$$m_Z(dx) = \mathbb{1}_{|x| < 1} \frac{k_1}{2} dx + \mathbb{1}_{|x| \geq 1} \frac{k_1}{2x^2} dx.$$

We note that $\frac{k_1}{2x^2} \leq \frac{k_1}{1+x^2}$ for all $x \geq 1$. The dynamic of Z can be shown to be

$$dZ_t = \begin{cases} \frac{2}{\sqrt{k_1}} dB_t, & \text{for } |Z_t| < 1, \\ \frac{2}{\sqrt{k_1}} Z_t dB_t, & \text{for } |Z_t| \geq 1, \end{cases}$$

where B is a standard Brownian motion. We also assume that $P_x(Z_0 = x) = 1$. As X and Z are on natural scale, they can be expressed as time-changed Brownian motion [82, Theorem 47.1, p. 277], *i.e.*, for every $t \geq 0$, $X_t = B_{\gamma_X(t)}$ and $Z_t = W_{\gamma_Z(t)}$, where B and W are two standard Brownian motions with $P_x(B_0 = x) = P_x(W_0 = x) = 1$, respective local times $L^x(B), L^x(W)$ and with $\gamma_X(t), \gamma_Z(t)$ being the respective right-inverses⁶ of

$$A_X(t) = \frac{1}{2} \int_{\mathbb{I}} L_t^x(B) m_X(dx) \quad \text{and} \quad A_Z(t) = \frac{1}{2} \int_{\mathbb{I}} L_t^x(W) m_Z(dx).$$

⁶The right-inverse of a function f is given by,

$$f^{-1}(x) = \inf\{\zeta \geq 0 : f(\zeta) > x\}.$$

Using the same underlying Brownian motion in these definitions, we have $A_Z(t) \leq A_X(t)$, and hence $\gamma_X(t) \leq \gamma_Z(t)$. Thus,

$$\begin{aligned} \mathbb{E}_x \left[\sup_{t \in [0, T]} |X_t - x|^p \right] &= \mathbb{E}_x \left[\sup_{t \in [0, \gamma_X(T)]} |B_t - x|^p \right] \\ &\leq \mathbb{E}_x \left[\sup_{t \in [0, \gamma_Z(T)]} |B_t - x|^p \right] = \mathbb{E}_x \left[\sup_{t \in [0, T]} |Z_t - x|^p \right]. \end{aligned}$$

Thus, from Doob's L^p and convexity inequalities,

$$\mathbb{E}_x \left[\sup_{t \in [0, T]} |X_t - x|^p \right] \leq \frac{2^{p-1}p}{p-1} \left(\mathbb{E}_x [|Z_T|^p] + |x|^p \right). \quad (5.32)$$

Using Itô's formula, followed by a classical localization argument, Fatou's Lemma along with the standard dominated convergence theorem, one obtains that for all $q > 2/3$,

$$\begin{aligned} \mathbb{E}_x \left[(|Z_t| - 1)^{3q} \mathbb{1}_{|Z_t| \geq 1} + 1 \right] \\ \leq (|x| - 1)^{3q} \mathbb{1}_{|x| \geq 1} + 1 + \int_0^t \frac{3q}{2} (3q - 1) \mathbb{E}_x \left[(|Z_s| - 1)^{3q-2} \mathbb{1}_{|Z_s| \geq 1} \frac{4Z_s^2}{k_1} \right] ds. \end{aligned}$$

Using the inequality $|x - 1|^{3q-2}x^2 \leq 4|x - 1|^{3q} + 4$ for all $x \geq 1$,

$$\begin{aligned} \mathbb{E}_x \left[(|Z_t| - 1)^{3q} \mathbb{1}_{|Z_t| \geq 1} + 1 \right] \\ \leq (|x| - 1)^{3q} \mathbb{1}_{|x| \geq 1} + 1 + \int_0^t \frac{24q}{k_1} (3q - 1) \mathbb{E}_x \left[(|Z_s| - 1)^{3q} \mathbb{1}_{|Z_s| \geq 1} + 1 \right] ds. \end{aligned}$$

Using Gronwall's Lemma, we deduce that, for all $t \geq 0$,

$$\mathbb{E}_x \left[(|Z_t| - 1)^{3q} \mathbb{1}_{|Z_t| \geq 1} + 1 \right] \leq [(|x| - 1)^{3q} \mathbb{1}_{|x| \geq 1} + 1] e^{24q(3q-1)t/k_1}.$$

Let $C'_q > 0$ be a constant such that $(|x| - 1)^{3q} \mathbb{1}_{|x| \geq 1} + 1 \geq C'_q |x|^{3q}$ for all $x \in \mathbb{R}$. Thus,

$$\mathbb{E}_x \left[|Z_t|^{3q} \right] \leq C'_q [(|x| - 1)^{3q} \mathbb{1}_{|x| \geq 1} + 1] e^{24q(3q-1)t/k_1}, \quad \forall t \geq 0.$$

Hence, for all⁷ $p > 2$, we have

$$\mathbb{E}_x \left[|Z_t|^p \right] \leq C'_{p/3} [1 + |x|^p] e^{8p(p-1)t/k_1}, \quad \forall t \geq 0. \quad (5.33)$$

Replacing (5.33) in (5.32), we get the bound (5.31) for $C' = \frac{p}{p-1} 2^{p-1} (C'_{p/3} + 1)$ and $C = 8p(p-1)/k_1$. □

Proposition 5.5.2. *Let X be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} , on natural scale and with a speed measure m_X that satisfies Condition (5.7). Then, for each $T > 0$ and $\gamma \in (0, \frac{1}{2})$, there exists a constant $C > 0$ such that*

$$\left\| \sup_{s \neq t \leq T} \frac{|X_t - X_s|}{|t - s|^\gamma} \right\|_{L^p(\mathbb{P}_x)} \leq C(1 + |x|). \quad (5.34)$$

The result also holds in the presence of a reflecting boundary $\zeta \in \mathbb{I}$.

⁷The extension to $p = 2$ is straightforward.

Proof. For the non-reflecting case, the proof works using the same arguments as in the proof of Theorem 3.1 in [9].

For the reflecting case: Let X be a diffusion process on natural scale, with speed measure m satisfying (5.9) for a constant $k_1 > 0$ and a reflecting boundary at $\zeta \in \mathbb{I}$. We observe that

$$X = |X^\circ - \zeta| + \zeta \quad (5.35)$$

in law, where X° is the non-reflecting diffusion on natural scale with speed measure

$$m_{X^\circ}(dx) = \mathbb{1}_{x \geq \zeta} m(dx) + \mathbb{1}_{x < \zeta} m(2\zeta - dx)$$

which also satisfies (5.9) for the same constant k_1 . From (5.35) and the triangle inequality,

$$\frac{|X_t - X_s|}{|t - s|^\gamma} = \frac{||X_t^\circ - \zeta| - |X_s^\circ - \zeta||}{|t - s|^\gamma} \leq \frac{|X_t^\circ - X_s^\circ|}{|t - s|^\gamma}. \quad (5.36)$$

The diffusion process X° is non-reflecting and its speed measure m_{X° satisfies (5.9). Thus, (5.34) holds for X° for a constant $C > 0$ and from (5.36),

$$\left\| \sup_{s \neq t \leq T} \frac{|X_t - X_s|}{|t - s|^\gamma} \right\|_{L^p(\mathbb{P}_x)} \leq \left\| \sup_{s \neq t \leq T} \frac{|X_t^\circ - X_s^\circ|}{|t - s|^\gamma} \right\|_{L^p(\mathbb{P}_x)} \leq C(1 + |x|).$$

□

5.5.2 Proof of the convergence rate for a process on natural scale

Proof of Theorem 5.2.1. From Proposition 5.2.7,

$$\begin{aligned} & \mathcal{W}_p \left[\text{Law} \left((\widetilde{X}_t^{\mathbf{g}})_{t \in [0, T]} \right), \text{Law} \left((X_t)_{t \in [0, T]} \right) \right] \\ &= \inf \left\{ \left\| \|\zeta - \xi\| \right\|_{L^p}; \zeta \sim \text{Law} \left((\widetilde{X}_t^{\mathbf{g}})_{t \in [0, T]} \right), \xi \sim \text{Law} \left((X_t)_{t \in [0, T]} \right) \right\} \\ & \leq \left\| \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}}(t)}^{\mathbf{g}} - X_t| \right\|_{L^p(\mathbb{P}_x)}. \end{aligned}$$

Let $M > T$. From Minkowski inequality,

$$\begin{aligned} \left\| \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}}(t)}^{\mathbf{g}} - X_t| \right\|_{L^p(\mathbb{P}_x)} & \leq \left\| \mathbb{1}_{\tau_{K^{\mathbf{g}}}(T) \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}}(t)}^{\mathbf{g}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ & \quad + \left\| \mathbb{1}_{\tau_{K^{\mathbf{g}}}(T) > M} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}}(t)}^{\mathbf{g}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \quad (5.37) \end{aligned}$$

For the first additive term of (5.37), for any $\gamma > 0$, by multiplying and dividing by $|\tau_{K^{\mathbf{g}}}(t) - t|^\gamma$,

$$\begin{aligned} & \left\| \mathbb{1}_{\tau_{K^{\mathbf{g}}}(T) \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}}(t)}^{\mathbf{g}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ &= \left\| \mathbb{1}_{\tau_{K^{\mathbf{g}}}(T) \leq M} \sup_{t \in [0, T]} \left\{ \frac{|X_{\tau_{K^{\mathbf{g}}}(t)}^{\mathbf{g}} - X_t|}{|\tau_{K^{\mathbf{g}}}(t) - t|^\gamma} |\tau_{K^{\mathbf{g}}}(t) - t|^\gamma \right\} \right\|_{L^p(\mathbb{P}_x)} \\ & \leq \left\| \mathbb{1}_{\tau_{K^{\mathbf{g}}}(T) \leq M} \sup_{t \in [0, T]} \left\{ \frac{|X_{\tau_{K^{\mathbf{g}}}(t)}^{\mathbf{g}} - X_t|}{|\tau_{K^{\mathbf{g}}}(t) - t|^\gamma} \right\} \sup_{t \in [0, T]} |\tau_{K^{\mathbf{g}}}(t) - t|^\gamma \right\|_{L^p(\mathbb{P}_x)}. \end{aligned}$$

As $\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}$ is increasing with respect to t ,

$$\begin{aligned} & \left\| \mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ & \leq \left\| \sup_{s \neq t \leq M} \left\{ \frac{|X_s - X_t|}{|s - t|^\gamma} \right\} \sup_{t \in [0, T]} |\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - t|^\gamma \right\|_{L^p(\mathbb{P}_x)} \\ & = \left(\mathbb{E}_x \left[\sup_{s \neq t \leq M} \left\{ \frac{|X_s - X_t|}{|s - t|^\gamma} \right\} \sup_{t \in [0, T]} |\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - t|^\gamma \right]^p \right)^{1/p}. \end{aligned}$$

From Hölder's inequality for $q \geq 1$ and $q/(q-1)$ conjugates exponents and (5.34), for every $\gamma \in (0, \frac{1}{2})$, there exists a constant $C_1 = C_1(M, \gamma, p(q-1)/q, x) > 0$ such that

$$\begin{aligned} & \left\| \mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ & \leq \left(\mathbb{E}_x \left[\sup_{s \neq t \leq M} \frac{|X_s - X_t|}{|s - t|^\gamma} \right]^{p(q-1)/q} \right)^{q/p(q-1)} \left(\mathbb{E}_x \left[\sup_{t \in [0, T]} |\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - t|^\gamma \right]^{pq} \right)^{1/pq} \\ & \leq C_1^{q/p(q-1)} \left(\mathbb{E}_x \left[\sup_{t \in [0, T]} |\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - t|^{\gamma pq} \right] \right)^{1/pq}. \end{aligned}$$

Choosing $q = 2/\gamma p$, from (5.29), for any $\gamma \in (0, \frac{1}{2} \wedge \frac{2}{p})$

$$\left\| \mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \leq C_1^{1/p(1-\gamma p)} \left(2|\mathbf{g}|_X (4T + |\mathbf{g}|_X) \right)^{\gamma/2}. \quad (5.38)$$

For the second additive term of (5.37),

$$\mathbb{E}_x \left[\mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} > M} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t|^p \right] = \sum_{m=M}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} \in [m, m+1]} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t|^p \right]. \quad (5.39)$$

For each term of the sum in (5.39), from Hölder's inequality,

$$\begin{aligned} & \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} \in [m, m+1]} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t|^p \right] \\ & \leq \left[\mathbb{P}_x(\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} > m) \right]^{1/q'} \left[\mathbb{E}_x \left[\mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} \in [m, m+1]} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t|^p \right] \right]^{q'/(q'-1)} \\ & \leq \left[\mathbb{P}_x(\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} > m) \right]^{1/q'} \left[\mathbb{E}_x \left[\mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} < m+1} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t|^p \right] \right]^{q'/(q'-1)}. \end{aligned}$$

As each m in the sum in (5.39) satisfies $m+1 > M > T$, from Minkowski's inequality,

$$\begin{aligned} & \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} < m+1} \sup_{t \in [0, T]} |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - X_t|^p \right] \\ & \leq \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^{\mathbf{g}}(T)}^{\mathbf{g}} < m+1} 2^{p-1} \sup_{t \in [0, T]} \left\{ |X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} - x|^p + |X_t - x|^p \right\} \right] \\ & \leq 2^p \mathbb{E}_x \left[\sup_{t \in [0, m+1]} |X_t - x|^p \right], \end{aligned}$$

which from Lemma 5.5.1 is bounded by $C_2 [1 + |x|^p] e^{C_3(m+1)}$, where C_2 and C_3 are positive constants depending on p . This, along with Lemma 5.4.4 and Hölder's inequality gives us for $\lambda > 0$ chosen such that $\alpha = \lambda |\mathbf{g}|_X < 1$,

$$\mathbb{E}_x \left[\mathbb{1}_{\tau_{K\mathbf{g}(T)}^{\mathbf{g}} \in [m, m+1]} \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t|^p \right] \leq \left[C_2 [1 + |x|^{pq'}] e^{C_3(m+1)} \right]^{1/q'} \left[e^{-\lambda m} e^{\lambda \frac{T + |\mathbf{g}|_X}{1 - \lambda |\mathbf{g}|_X}} \right]^{(q'-1)/q'}$$

where $C_2(pq') \leq 2pq'(pq' - 1)/c$ and $C_2(pq') > 0$ a positive constant depending only on pq' . Thus, setting $C_4(pq') := \left[C_2(pq') [1 + |x|^{pq'}] \right]^{1/q'} e^{C_3/q'} > 0$,

$$\begin{aligned} \mathbb{E}_x \left[\mathbb{1}_{\tau_{K\mathbf{g}(T)}^{\mathbf{g}} \in [m, m+1]} \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t|^p \right] \\ \leq C_4 \exp \left\{ \frac{1}{q'} \left[C_3 m + (q' - 1) \left(\lambda \frac{T + |\mathbf{g}|_X}{1 - \lambda |\mathbf{g}|_X} - \lambda m \right) \right] \right\} \\ = C_4 \exp \left\{ \frac{1}{q'} \left[C_3 m + (q' - 1) \frac{\alpha}{|\mathbf{g}|_X} \left(\frac{T + |\mathbf{g}|_X}{1 - \alpha} - m \right) \right] \right\} \\ = C_5 \exp \left\{ \frac{1}{q'} \left[C_3 m + (q' - 1) \frac{\alpha}{|\mathbf{g}|_X} \left(\frac{T}{1 - \alpha} - m \right) \right] \right\}, \end{aligned}$$

where $C_5 := C_4 \exp\left(\frac{q'-1}{q'} \frac{\alpha}{1-\alpha}\right)$. If we choose $q' > 1$ such that⁸ $A := C_3 - (q' - 1)\alpha/|\mathbf{g}|_X < 0$,

$$\begin{aligned} \sum_{m=M}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\tau_{K\mathbf{g}(T)}^{\mathbf{g}} \in [m, m+1]} \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t|^p \right] \\ \leq C_5 \exp \left\{ \frac{(q' - 1)\alpha T}{(1 - \alpha)|\mathbf{g}|_X} \right\} \sum_{m=M}^{\infty} (e^A)^m = C_5 \exp \left\{ \frac{(q' - 1)\alpha T}{(1 - \alpha)|\mathbf{g}|_X} \right\} \frac{e^{AM}}{1 - e^A} \\ = C_5 \exp \left\{ \frac{(q' - 1)\alpha T}{(1 - \alpha)|\mathbf{g}|_X} \right\} \frac{\exp \left\{ \left(C_3 - \frac{(q'-1)\alpha}{|\mathbf{g}|_X} \right) M \right\}}{1 - \exp \left\{ C_3 - \frac{(q'-1)\alpha}{|\mathbf{g}|_X} \right\}}. \end{aligned}$$

If M is chosen such that $M > T/(1 - \alpha)$, then

$$\sum_{m=M}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\tau_{K\mathbf{g}(T)}^{\mathbf{g}} \in [m, m+1]} \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t|^p \right] \leq \frac{C_5 e^{C_3 M}}{1 - e^A} \exp \left\{ \frac{(q' - 1)\alpha}{|\mathbf{g}|_X} \left(\frac{T}{1 - \alpha} - M \right) \right\}, \quad (5.40)$$

where the bound is $O(e^{-1/|\mathbf{g}|_X})$, and can be rewritten as $C^{(1)} e^{-C^{(2)}/|\mathbf{g}|_X}$. From (5.37), (5.38) and (5.40),

$$\left\| \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \leq C_1^{1/p(1-\gamma p)} \left(2|\mathbf{g}|_X \left(4T + |\mathbf{g}|_X \right) \right)^{\gamma/2} + C^{(1)} e^{-C^{(2)}/|\mathbf{g}|_X}.$$

As $|\mathbf{g}|_X^\gamma$ and $e^{-1/|\mathbf{g}|_X}$ are both $\mathcal{O}(|\mathbf{g}|_X^{\gamma/2})$, there exists a constant $C > 0$, such that,

$$\left\| \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \leq C |\mathbf{g}|_X^{\gamma/2},$$

which is (5.8). □

⁸We remark that if this is satisfied for a grid \mathbf{g} , then it is satisfied for all grids \mathbf{g}' such that $|\mathbf{g}'|_X \leq |\mathbf{g}|_X$.

5.6 Computations for the classical SDE case and beyond

In order to implement Algorithm 1 for simulating paths of a diffusion X , one needs two things. First, a grid \mathbf{g} adapted to the scale function and speed measure of X . Second, good approximations of the transition probabilities and conditional transition times of X over \mathbf{g} . The first point was covered in Section 5.2.3.

In this section, we show how to compute the quantities (5.5) in the pure SDE case. The extension to SDE solutions with point-wise singularities is straightforward. This allows us via Algorithm 1 to simulate all such processes.

Let (μ, σ) two real-valued functions that satisfy Condition 5.2.6 and X be the diffusion that solves

$$dX_t = \mu(X_t) dt + \sigma(X_t) dB_t,$$

where B is a standard Brownian motion. Let also \mathbb{I} be the state-space of X . A straightforward computation using Itô's formula gives us the infinitesimal generator of X ,

$$(\mathbb{L}, \text{dom}(\mathbb{L})) = \left\{ \begin{array}{l} \mathbb{L} f(x) = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x), \quad \forall f \in \text{dom}(\mathbb{L}), \\ \text{dom}(\mathbb{L}) = \left(f \in C_b(\mathbb{I}) : \mathbb{L} f \in C(\mathbb{I}) \right). \end{array} \right. \quad (5.41)$$

In particular, if μ and σ are continuous, then, $\text{dom}(\mathbb{L}) = C^2(\mathbb{I})$. From Proposition 5.3.1 and (5.41) we can infer that,

$$s'(x) = e^{-\int_a^x \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} \quad \text{and} \quad m(dx) = \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} dx. \quad (5.42)$$

Thus, from Proposition 5.3.1,

$$v_0(x) = \frac{\int_a^x e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} dy}{\int_a^b e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} dy}. \quad (5.43)$$

From (5.20) and (5.42),

$$\begin{aligned} v_1(x) &= \int_a^b G_{a,b}(x, \zeta) v_0(\zeta) \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} d\zeta, \\ \bar{v}_1(x) &= \int_a^b G_{a,b}(x, \zeta) (1 - v_0(\zeta)) \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} d\zeta. \end{aligned} \quad (5.44)$$

The scale functions and the speed measures are defined up to a multiplicative constant: for $\alpha \in \mathbb{R}$ and $\lambda > 0$, the pairs (s, m) and $(\alpha + \lambda s, \lambda^{-1}m)$ are associated to the same diffusion. In particular, as we are only concerned with points $x \in [a, b]$, we could use v_0 for the scale function. The speed measure shall be adapted accordingly. For $x, \zeta \in [a, b]$, the Green function in (1.20) takes the simpler form,

$$G_{a,b}(x, \zeta) = v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)).$$

Thus, expressions (5.44) become

$$\begin{aligned} v_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) \frac{v_0(\zeta)}{v_0'(\zeta)} \frac{2}{\sigma^2(\zeta)} d\zeta, \\ \bar{v}_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) \frac{1 - v_0(\zeta)}{v_0'(\zeta)} \frac{2}{\sigma^2(\zeta)} d\zeta. \end{aligned} \tag{5.45}$$

The quantities (5.43) and (5.44) or (5.45) can be computed analytically as we do for the Ornstein-Uhlenbeck process in Section 6.1.4 or approximated numerically as for the Cox-Ingersoll-Ross process in Section 6.1.4.

Chapter 6

Numerical experiments

Chapter outline: In this chapter we aim to test numerically the theoretical results established in this thesis, namely the STMCA approximation of one-dimensional diffusions of Chapter 5, the local time approximation and the stickiness parameter estimation of Chapter 4. In Section 6.1, we give numerical STMCA approximation examples of diffusions that exhibit various path-wise features. In Section 6.2, we assess numerically the stickiness parameter estimation and the local time approximation. We also see how the flexibility of grid choice in Algorithm 1 make it suitable for simulations regarding the local time approximation.

6.1 Approximation in distribution

6.1.1 Standard Brownian motion:

The standard Brownian motion can be defined as the diffusion process with scale function and speed measure

$$s(x) = x \quad \text{and} \quad m(dx) = 2 dx.$$

Let $v_k(x) := E_x[\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a}]$ and $\bar{v}_k(x) := E_x[\tau_{ab}^k \mathbb{1}_{\tau_a < \tau_b}]$ for all $k \in \mathbb{N}_0$, where $\mathbf{c}_x = (a, b)$. Then, from the definition of the scale function,

$$v_0(x) = \frac{x - a}{b - a}.$$

From Proposition 5.3.2,

$$v_1^{BM}(x) = (x - a)(b - x) \left(\frac{2(x - a)^2}{3(b - a)^2} + \frac{b - x}{b - a} - \frac{2(b - x)^2}{3(b - a)^2} \right),$$

$$\bar{v}_1^{BM}(x) = (x - a)(b - x) \left(\frac{2(b - x)^2}{3(b - a)^2} + \frac{x - a}{b - a} - \frac{2(x - a)^2}{3(b - a)^2} \right),$$

where by *BM* we mean these quantities are associated with the Brownian motion. Thus from (5.5), we have all the necessary quantities we need to implement the algorithm.

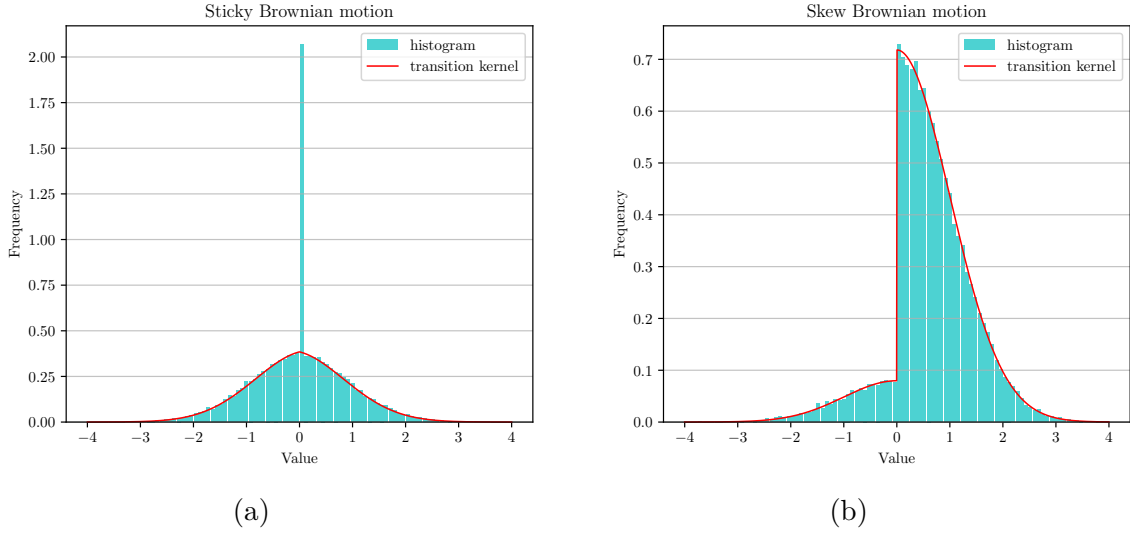


Figure 6.1: (a): histogram of simulated values at $T = 1$ of a sticky Brownian motion of parameter $\rho = 0.7$ with initial value $x_0 = 0$ using Algorithm 1 with the tuned grid (5.12) of size-criteria $h = 0.01$.
 (b): histogram of simulated values at $T = 1$ of a skew Brownian motion of parameter $\beta = 0.9$ with initial value $x_0 = 0$ using Algorithm 1 with a uniform grid of step-size $h = 0.01$.

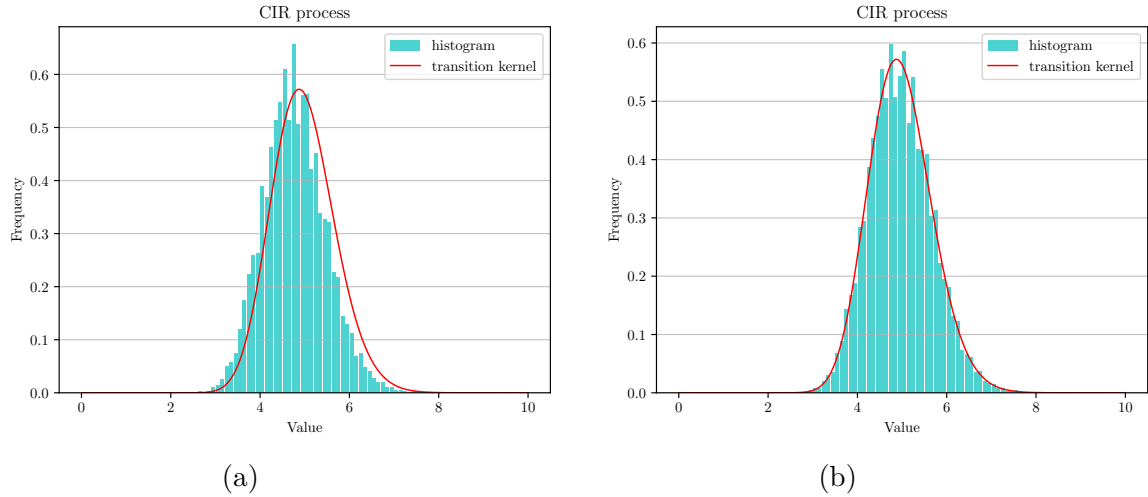


Figure 6.2: Histogram of simulated values at $T = 1$ of a CIR process of $(\theta, \mu, \sigma) = (5, 5, 1)$ with initial value $x_0 = 1$ using Algorithm 1 with:
 (a): a uniform grid of step-size $h = 0.01$ and (250, 200)-step Riemann approximation of (v_0, v_1) (simulation time: 44.5 sec).
 (b): a tuned grid of size-criteria $h = 0.01$ computed solving numerically (5.15)-(5.16) with Newton's method and (250, 200)-step Riemann approximation of (v_0, v_1) (simulation time: 46.5 sec).

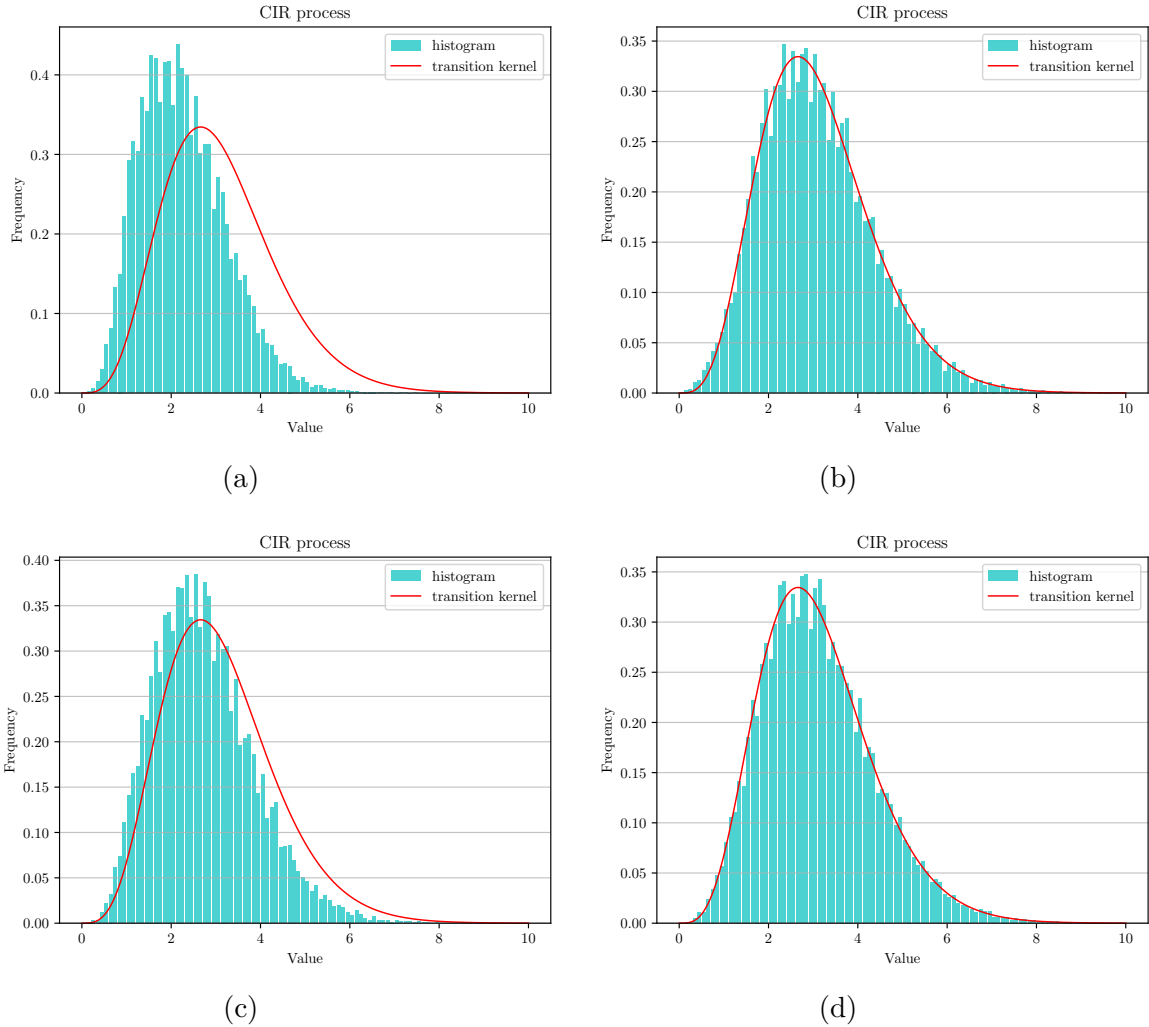


Figure 6.3: Histogram of simulated values at $T = 1$ of the CIR process of $(\theta, \mu, \sigma) = (1, 2, 1)$ with initial value $x_0 = 5$ using Algorithm 1 with:

- (a): a uniform grid of step-size $h = 0.015$ and $(100, 100)$ -step Riemann approximation of (v_0, v_1) (simulation time: 10.8 sec).
- (b): a tuned grid of size-criteria $h = 0.015$ computed solving numerically (5.15)-(5.16) with Newton's method and $(100, 100)$ -step Riemann approximation of (v_0, v_1) (simulation time: 12.8 sec).
- (c): Same as Figure 6.3a but with a $(250, 200)$ -step Riemann approximation of (v_0, v_1) (simulation time: 11.5 sec).
- (d): Same as Figure 6.3b but with a $(250, 200)$ -step Riemann approximation of (v_0, v_1) (simulation time: 12.2 sec).

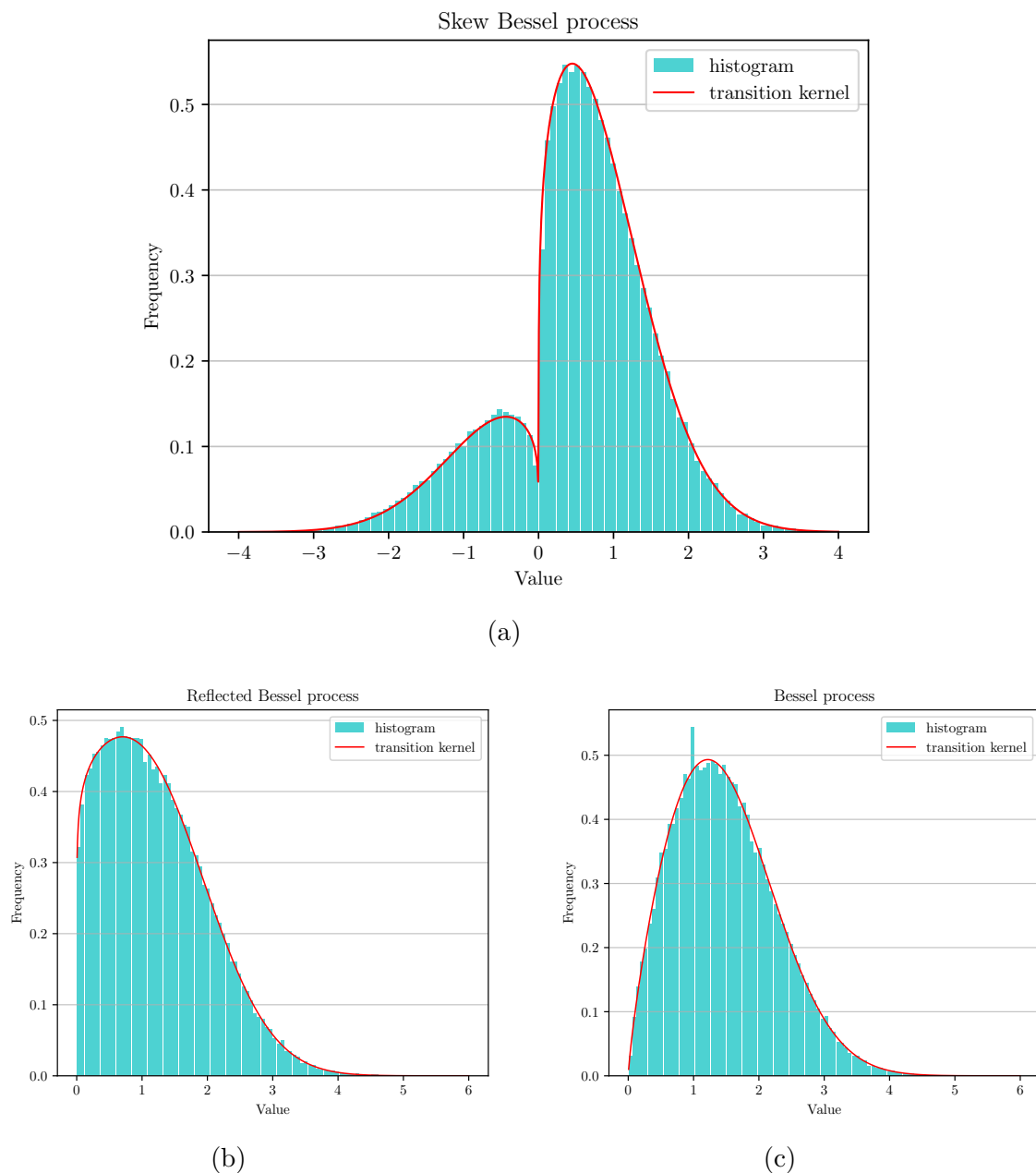


Figure 6.4: Histograms of
(a): simulated values at $T = 1$ of the Skew Bessel process of parameters $(\delta, \beta) = (1.2, 0.8)$ with initial value $x_0 = 0$ using Algorithm 1. The quantities (6.1) were approximated using 100-step Riemann approximations.
(b): histogram of simulated values at $T = 1$ of the Reflected Bessel process of parameter $\delta = 1.1$ with initial value $x_0 = 5$ using Algorithm 1 with a uniform grid of step-size $h = 0.01$.
(c): histogram of simulated values at $T = 1$ of the Bessel process of parameter $\delta = 1.8$ with initial value $x_0 = 5$ using Algorithm 1 with a uniform grid of step-size $h = 0.01$. The quantities (6.1) were approximated using 100-step Riemann approximations.

6.1.2 Sticky Brownian motion:

The Brownian motion with a sticky point at 0 where the stickiness parameter is $\rho > 0$ can be defined as the diffusion process with scale function and speed measure

$$s(x) = x \quad \text{and} \quad m(dx) = 2 dx + \rho \delta_0(dx).$$

From the definition of the scale function,

$$v_0(x) = \frac{x - a}{b - a}.$$

From Proposition 5.3.2 we may deduce the following expressions for the conditional exit times of $\mathbf{c}_x = (a, b)$,

$$\begin{aligned} v_1(x) &= v_1^{BM}(x) + \rho \mathbb{1}_{0 \in (a,b)} G_{(a,b)}(x, 0) v_0(0), \\ \bar{v}_1(x) &= \bar{v}_1^{BM}(x) + \rho \mathbb{1}_{0 \in (a,b)} G_{(a,b)}(x, 0) (1 - v_0(0)), \end{aligned}$$

where $v_1^{BM}(x)$ and $\bar{v}_1^{BM}(x)$ are the analogous quantities for the standard Brownian motion.

6.1.3 Skew Brownian motion:

The skew Brownian motion at 0 with parameter $\beta \in (0, 1)$ can be defined as the diffusion process with scale function and speed measure

$$s(x) = \begin{cases} \frac{x}{\beta}, & x \geq 0, \\ \frac{x}{1-\beta}, & x \leq 0 \end{cases} \quad \text{and} \quad m(dx) = \begin{cases} 2\beta dx, & x > 0, \\ 2(1-\beta) dx, & x < 0. \end{cases}$$

From the definition of the scale function,

$$v_0(x) = \frac{s(x) - s(a)}{s(b) - s(a)}.$$

As the β and $(1-\beta)$ terms between the speed measure and the scale function compensate themselves in the expressions giving v_k in Proposition 5.3.2,

$$v_1(x) = v_1^{BM}(x) \quad \text{and} \quad \bar{v}_1(x) = \bar{v}_1^{BM}(x).$$

6.1.4 Ornstein-Uhlenbeck process:

Let X_t be an Ornstein-Uhlenbeck with mean reversion force $\theta > 0$, long-term mean μ and diffusion parameter $\sigma > 0$, i.e.

$$dX_t = \theta(\mu - X_t) dt + \sigma dB_t,$$

where B_t is a standard Brownian motion. We will see that v_0 can be expressed in terms of the Gaussian imaginary error function as $s(x) = \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - x)\right)$,

$$v_0(x) = \frac{\int_a^x e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} dy}{\int_a^b e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} dy} = \frac{\operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - x)\right) - \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - a)\right)}{\operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - b)\right) - \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - a)\right)},$$

where $\operatorname{erfi}(x) = \frac{2}{\pi} \int_0^x e^{t^2} dt = \frac{2}{\pi} e^{x^2} D_+(x)$, with $D_+(x)$ being Dawson function⁹. Thus, for an Ornstein-Uhlenbeck process (5.44) becomes,

$$\begin{aligned} v_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) v_0(\zeta) c \exp\left(\frac{\theta(\zeta - \mu)^2}{\sigma^2}\right) \frac{2}{\sigma^2} d\zeta, \\ \bar{v}_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) (1 - v_0(\zeta)) c \exp\left(\frac{\theta(\zeta - \mu)^2}{\sigma^2}\right) \frac{2}{\sigma^2} d\zeta, \end{aligned}$$

where $c = \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - b)\right) - \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - a)\right)$.

6.1.5 Cox-Ingersoll-Ross process:

The Cox-Ingersoll-Ross process or CIR process [23], introduced first by W. Feller [38], is the diffusion that solves the SDE

$$dX_t = \theta(\mu - X_t) dt + \sigma\sqrt{X_t} dB_t,$$

where B_t is a standard Brownian motion. The parameter $\theta > 0$ expresses its mean reversion speed, μ its long term speed and $\sigma > 0$ is its diffusivity parameter. The parameter $\theta > 0$ expresses its mean reversion speed, μ its long term speed and $\sigma > 0$ is its diffusivity parameter. This equation has a diffusion coefficients that degenerates at 0. It however remains non-negative given $X_0 \geq 0$ and almost surely never hit 0 when $2\theta\mu > \sigma^2$. A large body of work have been devoted to the simulation of the CIR and related process, see e.g. [2].

From (5.43) and (5.45),

$$v_0(x) = \frac{\int_a^x y^{-\frac{2\theta\mu}{\sigma^2}} e^{\frac{2\theta}{\sigma^2}y} dy}{\int_a^b y^{-\frac{2\theta\mu}{\sigma^2}} e^{\frac{2\theta}{\sigma^2}y} dy}$$

and

$$\begin{aligned} v_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) \frac{v_0(\zeta)}{v_0'(\zeta)} \frac{2}{\sigma^2 \zeta} d\zeta, \\ \bar{v}_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) \frac{1 - v_0(\zeta)}{v_0'(\zeta)} \frac{2}{\sigma^2 \zeta} d\zeta. \end{aligned}$$

As the scale function yield no satisfactory closed formula, we perform a numerical approximation of both v_0 and the couple (v_1, \bar{v}_1) .

These functions may be computed numerically. One may choose a suitable grid when the process is close to 0 and it is noteworthy that the process may not cross 0, a problem which arise when using Euler type schemes.

⁹We use the latter representation in our numerical results.

6.1.6 Skew and reflected Bessel process:

The Skew-Bessel process (see [3]) of dimension $\delta \in (0, 2)$ and skew $\beta \in (0, 1)$ is the diffusion process with the following scale function and speed measure,

$$s(x) = \begin{cases} \frac{1}{\beta} \frac{x^{2-\delta}}{2-\delta}, & x > 0, \\ -\frac{1}{1-\beta} \frac{|x|^{2-\delta}}{2-\delta}, & x \leq 0, \end{cases} \quad m(dx) = \begin{cases} 2\beta x^{\delta-1} dx, & x > 0, \\ 2(1-\beta)|x|^{\delta-1} dx, & x \leq 0. \end{cases}$$

This yields the following expressions for the quantities we need to compute in order to implement the algorithm,

$$\begin{aligned} v_0(x) &= \frac{s(x) - s(a)}{s(b) - s(a)}, \\ v_1(x) &= \int_a^b G_{a,b}(x, \zeta) v_0(\zeta) 2|\zeta|^{\delta-1} d\zeta, \\ \bar{v}_1(x) &= \int_a^b G_{a,b}(x, \zeta) (1 - v_0(\zeta)) 2|\zeta|^{\delta-1} d\zeta, \end{aligned} \tag{6.1}$$

where $G_{a,b}(x, \zeta)$ is defined in (1.20). The probability transition kernels plotted in Figure 6.3 are computed in [3].

6.2 Local time approximation

In this section we present the results of numerical simulations regarding the local time approximation developed in Chapter 4. With numerical experiments, we assess the asymptotic properties of the local time approximation (4.7). Also, with an example we illustrate the flexibility of Space-Time Markov Chain Approximations generated via Algorithm 1. One feature of such approximation processes is that they are defined on a given grid. With a suitable choice of grid it is possible to achieve higher orders of convergence of localized path-sensitive functionals like (4.7).

We simulate trajectories of an approximation process of the sticky Brownian motion of parameter ρ . Then, we compare the stickiness parameter estimations (4.8) with the true value of ρ . For the numerical simulations we use the Space-time Markov chain approximation or STMCA Algorithm [6]. This algorithm uses grid-valued random walks to approximate the law of any one-dimensional generalized diffusion process. The STMCA Algorithm is particularly adapted to our problematic as:

- It is well suited for the simulation of sticky singular one-dimensional diffusions.
- By suitable choice of the grid, we can control the amount of path-observations of X observed through the test function g in (4.7). In particular, for grids that satisfy the condition in Corollary 2.5 of [6], the convergence speed of the STMCA algorithm is optimal. Thus, if one increases the precision of such a grid \mathbf{g} around 0, we approximate accurately X with \mathbf{g} -valued STMCA random walks and feed more path-observations to the statistic (4.7) without paying a too heavy computational cost.

The statistic: Let X be a sticky Brownian motion of parameter ρ and g the function defined for every $x \in \mathbb{R}$ by $g(x) = \mathbb{1}_{1 < |x| < 5/8}$. For every $\alpha \in (0, 1)$ and $n \in \mathbb{N}$, we define the following path-wise statistics of X :

$$T_{n,\alpha}^{(1)}(X) := \frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g(n^\alpha X_{\frac{i-1}{n}}^\rho), \quad (6.2)$$

$$T_{n,\alpha}^{(2)}(X) := \frac{2 \sum_{i=1}^{[nt]} \mathbb{1}_{X_{\frac{i-1}{n}}=0}}{n T_{n,\alpha}^{(1)}(X)}. \quad (6.3)$$

From (4.7), for any $\alpha \in (0, 1/2)$, the statistics (6.2) and (6.3) converge to $L_t^0(X)$ and ρ respectively in probability. We use (6.3) as proxy to assess the properties of the local time approximation.

The grids: For every $h > 0$, let \mathbf{g}_0 and \mathbf{g}_1 be two grids defined by:

$$\mathbf{g}_0(h) = \{0\} \cup \left\{ \pm (h^2/\rho + kh); k \in \mathbb{N} \right\}, \quad \mathbf{g}_1(h) = \left\{ \pm x_k(h); k \geq 0 \right\},$$

where $\{x_j(h)\}_{j \geq 0}$ is defined recursively by

$$x_0 = 0, \quad x_j = x_{j-1} + \left(\frac{h^2}{\rho x_{j-1} + 1} + h \left(1 - \frac{1}{x_{j-1} + 1} \right) \right) \mathbb{1}_{x_{j-1} < 1} + h \mathbb{1}_{x_{j-1} \geq 1}.$$

We observe that

$$|\mathbf{g}_1| = |\mathbf{g}_0| \quad \text{and} \quad |\mathbf{g}_1|_X = |\mathbf{g}_0|_X.$$

From Theorem 5.2.1, the asymptotic convergence of the STMCA is $\mathcal{O}(h^{1/2})$ for both grids.

α	n	$\hat{\rho}_{\text{MC}}$	\hat{S}_{MC}^2	$\hat{\sigma}_{\text{MC}}$	$\widehat{\text{acc}}$	rej/ N_{MC}
0.3	100000	1.33451	0.740592	0.860577	0.125709	0/2000
0.4	100000	1.06132	0.142955	0.378094	0.0460073	0/2000
0.5	100000	1.3148	0.162762	0.403438	0.0116839	0/2000
0.55	100000					2000/2000
0.6	100000					2000/2000
0.65	100000					2000/2000

Table 6.1: Stickiness parameter estimations using the grid \mathbf{g}_0 for $h = 0.01$. The missing values in the table corresponds to cases where the statistic (6.3) is observed to be 0. Computation time (single-core): 4 seconds.

Within each of the following tables we use the same simulated STMCA trajectories of the sticky Brownian motion of parameter $\rho = 1$:

Simulation metrics: The integer N_{MC} will be the Monte Carlo simulation size. For every $j \leq N_{\text{MC}}$ we simulate the path of an approximation process X_t^j . To assess the quality of each Monte Carlo estimation we use the following metrics:

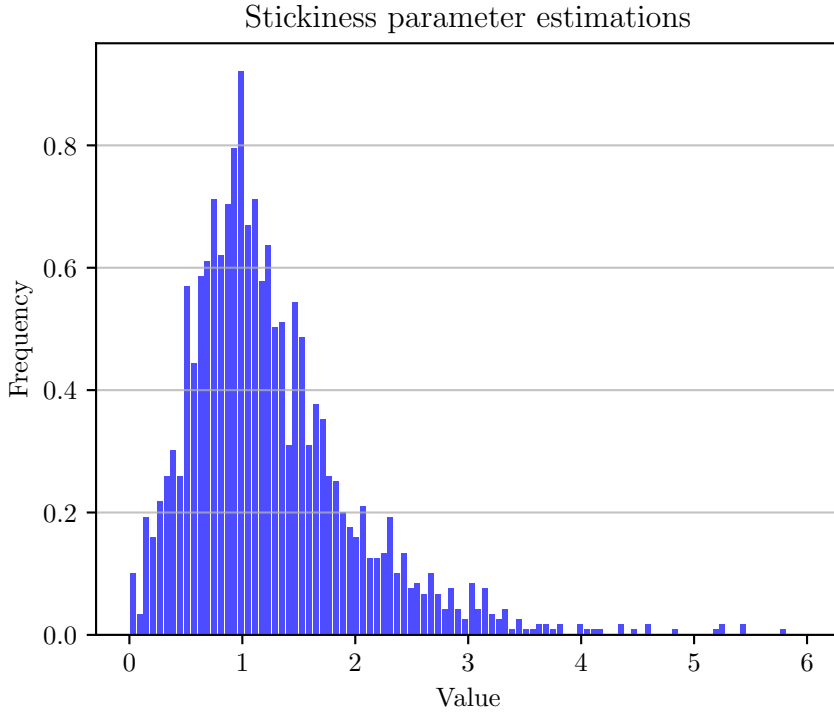


Figure 6.5: Stickiness parameter estimations histograms using (6.3) with $n = 100000$ and $\alpha = 0.3$. (true value $\rho = 1$)

- $(\widehat{\rho}_{MC}, \widehat{S}_{MC}^2, \widehat{\sigma}_{MC})$: Monte Carlo estimation, variance and standard deviation of the estimated stickiness parameters,
- \widehat{acc} : average number of path-values observed by g , i.e

$$\widehat{acc} = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{g(X_{(i-1)/n}^j) \neq 0},$$

- rej : percentage of trajectories where the local time estimation equals 0, i.e

$$rej = \#\left\{j \leq N_{MC} : \frac{n^\alpha}{n} \sum_{i=1}^{[nt]} g(n^\alpha X_{\frac{i-1}{n}}^j) = 0\right\}.$$

6.2.1 Observations, conclusion

From the numerical experiments (Tables 6.1-6.5), we observe the following:

- As long as n is high enough, the higher the value of $\alpha \in (0, 1)$, the lower the estimated variance and the better the approximation of the local time,
- Also, the higher the α , the more the trajectory of X^ρ is inflated and the less things are observed through g . Thus, having a finite set of path-wise observations of X^ρ , one must find an $\alpha \in (0, 1)$ large enough to trigger the asymptotic regime of (4.1) and low enough so we do not dump too much path-wise observations.

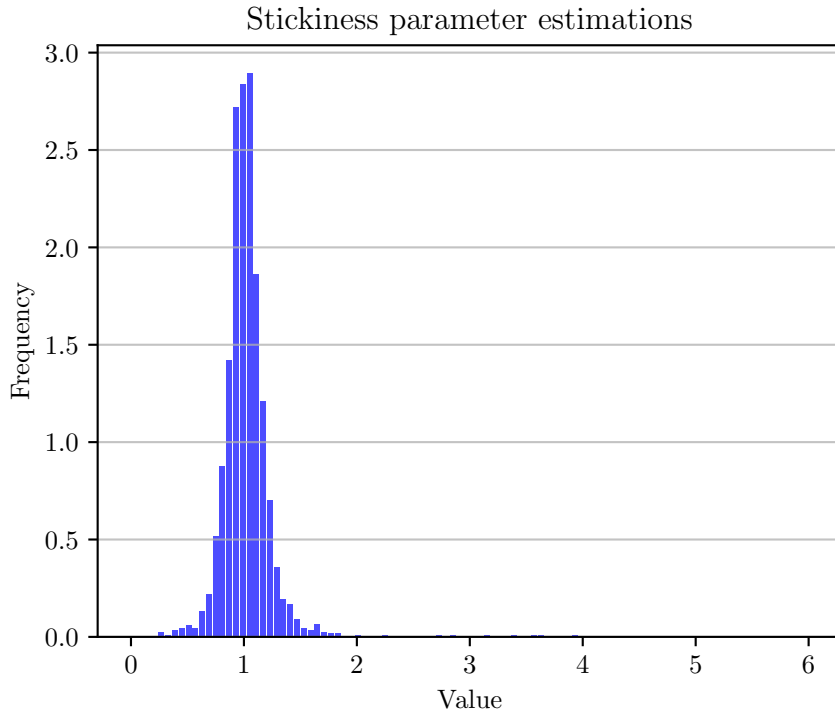


Figure 6.6: Stickiness parameter estimations histograms using (6.3) with $n = 100000$ and $\alpha = 0.55$. (true value $\rho = 1$)

- In Table 6.5 we see that for a fixed $c > 0$, every (n, α) such that $\log n = c/\alpha$ yield the same Monte Carlo variance. This relation can be guessed from (4.4) and (4.1),
- The convergence (4.7) seems to hold for $\alpha \in [1/2, 1)$, values not covered by Theorem 4.1.3. We thus conjecture the following:

Algorithm 1 gives us the flexibility to remediate to the latter by using grids of higher precision around the point of stickiness and thus achieving higher orders of convergence for (6.2) without significant increase in the numerical complexity.

We observe that the usage of grid \mathbf{g}_1 yields far superior results than \mathbf{g}_0 . Using \mathbf{g}_1 we have an abundance of simulated path-wise observations close to the point of stickiness. The statistic (6.2) remains thus relevant for large values of α and we can achieve higher orders of convergence.

Conjecture 6.2.1. *Theorem 4.1.3 holds for any $\alpha \in (0, 1)$ and the rate of convergence of both (4.8) and (4.7) are increasing in terms of α .*

α	n	$\hat{\rho}_{MC}$	\hat{S}_{MC}^2	$\hat{\sigma}_{MC}$	\widehat{acc}	rej/ N_{MC}
0.3	100000	1.25282	0.629184	0.793211	0.127209	0/2000
0.4	100000	1.06863	0.18705	0.432493	0.0437251	0/2000
0.5	100000	1.02367	0.0721315	0.268573	0.0140703	0/2000
0.55	100000	1.01813	0.0672543	0.259334	0.00796982	0/2000
0.6	100000	1.01266	0.031502	0.177488	0.00448427	0/2000
0.65	100000	1.01594	0.0268068	0.163728	0.00251455	2/2000

Table 6.2: Stickiness parameter estimations using the grid \mathbf{g}_1 for $h = 0.01$. Computation time (single-core): 67 seconds.

α	n	$\hat{\rho}_{MC}$	\hat{S}_{MC}^2	$\hat{\sigma}_{MC}$	\widehat{acc}	rej/ N_{MC}
0.6	20000	1.04379	0.14514	0.380973	0.0121671	0/1000
0.6	40000	1.02967	0.03178	0.178288	0.0079693	0/1000
0.6	100000	1.00135	0.01592	0.126199	0.0046799	0/1000
0.6	300000	1.00638	0.00850	0.092235	0.0023960	0/1000

Table 6.3: Simulation results for fixed α and different values of n .

α	n	$\hat{\rho}_{MC}$	\hat{S}_{MC}^2	$\hat{\sigma}_{MC}$	\widehat{acc}	rej/ N_{MC}
0.3	100000	1.28303	0.66712	0.81677	0.126344	0/2000
0.4	100000	1.12398	0.38812	0.62299	0.043843	0/2000
0.5	100000	1.04014	0.09153	0.30253	0.014309	0/2000
0.6	100000	1.01083*	0.04976*	0.22307*	0.004598*	1/2000
0.7	100000	1.06961*	0.14296*	0.37810*	0.001367*	13/2000
0.8	100000	1.01483*	0.12386*	0.35194*	0.000467*	23/2000

Table 6.4: Simulation results for fixed n and different values of α . Estimation with an asterisk were performed removing the trajectories with $T_n^{(1)}(X) = 0$.

α	n	$\hat{\rho}_{MC}$	\hat{S}_{MC}^2	$\hat{\sigma}_{MC}$	\widehat{acc}	rej/ N_{MC}
0.55536	20000	1.04813	0.094319	0.307115	0.0190297	0/1000
0.519033	40000	1.04893	0.095243	0.308616	0.0190053	0/1000
0.477724	100000	1.04946	0.100884	0.317623	0.0190204	0/1000
0.436109	300000	1.04931	0.096736	0.311024	0.0190139	0/1000

Table 6.5: Simulation results for (n, α) satisfying $\log n = 5.5/\alpha$.

Appendix A

A.1 Strong continuity of diffusion semi-groups on $C_b(\mathbb{I})$: a counter-example

Proposition A.1.1. *Let $(P_t)_{t \geq 0}$ be the semi-group of the standard Brownian motion B , defined for every measurable bounded f , $x \in \mathbb{R}$ and $t \geq 0$ by*

$$P_t f(x) = \mathbb{E}_x(f(B_t)).$$

Then, $(P_t)_{t \geq 0}$ is not strongly continuous on $C_b(\mathbb{R})$.

Proof. Let ϕ_n be the sequence of functions defined for every $n \in \mathbb{N}$ and $x \in \mathbb{R}$ by

$$\phi_n(x) = (1 - n|x|)\mathbb{1}_{|x| < 1/n}.$$

Let also f and g be the functions defined for every $x \in \mathbb{R}$ by

$$f(x) = \sum_{n \in \mathbb{N}} \phi_n(x - n^2)$$

and

$$g(x) = \mathbb{1}_{|x - n^2| < 1/n} + \mathbb{1}_{|x - n^2| > \Delta_n},$$

where $\Delta_n = n^2 - (n - 1)^2 - 1/n = 2n - 1 - 1/n$. We observe that for all $n \in \mathbb{N}$,

$$f(n^2) = 1. \tag{A.1}$$

Also, if $Z \sim \mathcal{N}(0, 1)$,

$$\begin{aligned} P_t f(n^2) &= \mathbb{E}_{n^2}(f(\sqrt{t}Z)) \leq \mathbb{E}(g(n^2 + \sqrt{t}Z)) \\ &= \mathbb{P}\left(Z \in \left(-\frac{1}{n\sqrt{t}}, \frac{1}{n\sqrt{t}}\right)\right) + \mathbb{P}(|Z| \geq \Delta_n/\sqrt{t}). \end{aligned}$$

Thus, for all $t \geq 0$ and $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ large enough such that for all $n \geq n_0$,

$$P_t f(n^2) < \epsilon. \tag{A.2}$$

From (A.1) and (A.2), for all $t \geq 0$,

$$\sup_{x \in \mathbb{R}} |P_t f(x) - f(x)| \geq 1$$

and thus $(P_t)_{t \geq 0}$ is not strongly continuous over $C_b(\mathbb{R})$. □

A.2 Semimartingale results

A.2.1 Change of variables in a stochastic integral

Definition A.2.1. • An $(\mathcal{F}_t)_{t \geq 0}$ -time change is any almost surely increasing, right-continuous family of $(\mathcal{F}_t)_{t \geq 0}$ -stopping times $(T_t)_{t \geq 0}$ such that $T_0 = 0$.

- A process M is said to be in synchronization with T iff M is constant on $[T_{s-}, T_s]$.
- Given an $(\mathcal{F}_t)_{t \geq 0}$ -semi-martingale M , let $L(M, (\mathcal{F}_t)_{t \geq 0})$ be the class of $(\mathcal{F}_t)_{t \geq 0}$ -progressively measurable, predictable càdlàg processes H for which the integral $\int_0^t H_s dM_s$ can be constructed.

Theorem A.2.2 (Theorem 3.1 of [64], Proposition 10.21 of [54]). *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space, M an $(\mathcal{F}_t)_{t \geq 0}$ -semimartingale in synchronization with an $(\mathcal{F}_t)_{t \geq 0}$ -time change T . If $H \in L(Z, (\mathcal{F}_t)_{t \geq 0})$, then $(H_{T_t-})_{t \geq 0} \in L(Z \circ T, (\mathcal{F}_{T_t})_{t \geq 0})$. Also,*

$$\int_0^{T_t} H_s dM_s = \int_0^t H_{T_{s-}} dM_{T_s}, \quad (\text{A.3})$$

for every $t \geq 0$.

A.2.2 Local time scaling

Lemma A.2.3 (Exercise 1.23 of [80]-Chapter VI). *Let X be a continuous semimartingale and f a strictly increasing difference of two convex functions. If $f(X) = (f(X_t))_{t \geq 0}$, then for every $a \in \mathbb{I}$ and $t \geq 0$,*

$$L_t^{f(a)}(f(X)) = f'(a)L_t^a(X)$$

almost surely, where f' is the right-derivative of f .

Lemma A.2.4 (Exercise 1.27 of [80]-Chapter VI). *Let X be a semi-martingale with state-space \mathbb{I} , T a time change and Y is the time changed process such that $Y_t = X_{T(t)}$ for every $t \geq 0$. Then for every $a \in \mathbb{I}$ and $t \geq 0$,*

$$L_t^a(Y) = L_{T(t)}^a(X) \quad (\text{A.4})$$

almost surely.

A.2.3 Stochastic Dini's theorem

Definition A.2.5. A sequence of processes X^n is said to converge locally uniformly in time, in \mathbb{P} -probability to a process X iff

$$\sup_{s \leq t} \left\{ |X_s^n - X_s| \right\} \xrightarrow{n \rightarrow \infty} 0,$$

in \mathbb{P} -probability, for every $t \geq 0$

Lemma A.2.6 ((2.2.16) of [56]). *Let $\{X^n\}_{n \in \mathbb{N}}$ be a sequence of increasing processes and X an almost surely continuous bounded process. If $\{X^n\}_{n \in \mathbb{N}}$ and X are defined on the probability space $(\Omega, (\mathcal{F})_{t \geq 0}, \mathbb{P})$ and*

$$X_t^n \xrightarrow[n \rightarrow \infty]{} X_t$$

in probability, for every $t \in D$ with D a dense subset of \mathbb{R}_+ . Then, the convergence is locally uniform in time, in probability.

A.3 Joint scaling of the sticky Brownian motion

Lemma A.3.1. *Let $\{(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x^\rho); x \in \mathbb{R}, \rho \geq 0\}$ be a family of filtered probability spaces and $X^\rho = (X_t^\rho)_{t \geq 0}$ a process defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$ such that under \mathbb{P}_x^ρ it is the sticky Brownian motion of stickiness parameter ρ and $\mathbb{P}_x^\rho(X_0^\rho = x) = 1$. Then,*

$$\begin{aligned} \text{Law}_{\mathbb{P}_x^\rho} \left(X_{ct}^\rho, L_{ct}^0(X^\rho), A_{ct}^+(X^\rho); t \geq 0 \right) \\ = \text{Law}_{\mathbb{P}_{\sqrt{c}x}^{\rho/\sqrt{c}}} \left(\sqrt{c}X_t^{\rho/\sqrt{c}}, \sqrt{c}L_t^0(X^{\rho/\sqrt{c}}), cA_t^+(X^{\rho/\sqrt{c}}); t \geq 0 \right), \end{aligned} \quad (\text{A.5})$$

where $\mathbb{P}_x^\rho(X_0^{\rho/\sqrt{c}} = \sqrt{c}x) = 1$ and $(L^0(X^\rho), A^+(X^\rho))$ and $(L^0(X^{\rho/\sqrt{c}}), A^+(X^{\rho/\sqrt{c}}))$ are the local times, occupation times¹⁰ of \mathbb{R}_+ pairs of X^ρ and $X^{\rho/\sqrt{c}}$ respectively.

Proof. From [88], the joint density of $(X_t^\rho, L_t^0(X^\rho), A_{ct}^+(X^\rho))$ is defined for every $(t, x, y, l, o) \in \mathbb{R}_+ \times \mathbb{R}^2 \times [0, \frac{2}{\rho}] \times [0, t]$ by

$$\begin{aligned} \mathbb{P}_x(X_t^\rho \in dy, L_t^0(X^\rho) \in dl, A_{ct}^+(X^\rho) \in do) \\ = q_\rho(t, x, y, l, o) dy dl do = h\left(o - \rho l, \frac{\ell}{2} + x\right) h\left(t - o, \frac{\ell}{2} - y\right) dy dl do, \end{aligned}$$

where $h(t, x)$ is the function defined for every $t \geq 0$ and $x \in \mathbb{R}$ by

$$h(t, x) = \frac{|x|}{\sqrt{2\pi t^{3/2}}} e^{-x^2/2t}.$$

We observe that

$$h(ct, x) = c^{-1}h(t, x/\sqrt{c}).$$

Thus,

$$\begin{aligned} q_\rho(ct, x, y, l, o) &= h\left(o - \rho l, \frac{\ell}{2} + x\right) h\left(ct - o, \frac{\ell}{2} - y\right) \\ &= h\left(c\left(\frac{o}{c} - \frac{\rho l}{c}\right), \frac{\ell}{2} + x\right) h\left(c\left(t - \frac{o}{c}\right), \frac{\ell}{2} - y\right) \\ &= c^{-2}h\left(\frac{o}{c} - \frac{\rho l}{c}, \frac{\ell}{2\sqrt{c}} + \frac{x}{\sqrt{c}}\right) h\left(t - \frac{o}{c}, \frac{\ell}{2\sqrt{c}} - \frac{y}{\sqrt{c}}\right) \\ &= c^{-2}q_{\rho/\sqrt{c}}\left(t, \frac{x}{\sqrt{c}}, \frac{y}{\sqrt{c}}, \frac{l}{\sqrt{c}}, \frac{o}{c}\right) \end{aligned}$$

¹⁰The occupation time by X of a measurable set $B \in \mathbb{B}(\mathbb{R})$ is the process $A^B(X)$ defined for every $t \geq 0$ by

$$A_t^B(X) = \int_0^t \mathbb{1}_{X_s \in B} ds.$$

and

$$q_\rho(ct, x, y, l, o) dy dl do = q_{\rho/\sqrt{c}}\left(t, \frac{x}{\sqrt{c}}, \frac{y}{\sqrt{c}}, \frac{l}{\sqrt{c}}, \frac{o}{c}\right) d\left(\frac{y}{\sqrt{c}}\right) d\left(\frac{l}{\sqrt{c}}\right) \left(d\frac{o}{c}\right),$$

which finishes the proof. \square

A.4 Conditioning on the embedded path of a diffusion

Let X be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} defined on a family of probability spaces $\mathcal{P} = (\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)_{x \in \mathbb{I}}$ such that for every $x \in \mathbb{I}$, $\mathbb{P}_x(X_0 = x) = 1$. Let also \mathbf{g} be a covering grid of \mathbb{I} . We define the embedding times of $(X_t)_{t \geq 0}$ in \mathbf{g} as,

$$\begin{cases} \tau_0^{\mathbf{g}} &= 0, \\ \tau_k^{\mathbf{g}} &= \inf \left\{ t > \tau_{k-1}^{\mathbf{g}} : X_t \in \mathbf{g} \setminus \{X_{\tau_{k-1}^{\mathbf{g}}}\} \right\}, \quad \forall k \geq 1. \end{cases} \quad (\text{A.6})$$

Lemma A.4.1. *Let X be a diffusion process with state-space \mathbb{I} an interval of \mathbb{R} defined on a family of probability spaces $\mathcal{P} = (\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_x)_{x \in \mathbb{I}}$ such that for every $x \in \mathbb{I}$, $\mathbb{P}_x(X_0 = x) = 1$. Let also \mathbf{g} be a covering grid of \mathbb{I} , $\{\tau_j\}_{j \in \mathbb{Z}_+}$ the sequence of embedding times of X in \mathbf{g} defined in (A.6) and \mathcal{B} the sigma-algebra defined by $\mathcal{B} = \sigma\{X_{\tau_j}; j \in \mathbb{N}_0\}$. Then, for any measurable path-functional $F : C^0(\mathbb{R}_+, \mathbb{I}) \mapsto \mathbb{R}$ and $j \geq 1$ and $x \in \mathbb{I}$,*

$$\mathbb{E}_x \left(F((X_t^{\tau_{j-1}, \tau_j})_{t \geq 0}) \middle| \mathcal{B} \right) = \mathbb{E}_x \left(F((X_t^{\tau_{j-1}, \tau_j})_{t \geq 0}) \middle| X_{\tau_{j-1}}, X_{\tau_j} \right),$$

where X^{τ_{j-1}, τ_j} is the process defined for every $t \geq 0$ by $X_t^{\tau_{j-1}, \tau_j} = X_{(\tau_{j-1} + t) \wedge \tau_j}$.

Proof. Let us fix x_1, x_2, \dots be a sequence of points in the grid. Let us define

$$Q(x; x_1, x_2, \dots) := \mathbb{P}_x \left(X_{\tau_1} = x_1, X_{\tau_2} = x_2, \dots \right)$$

By the strong Markov property,

$$\begin{aligned} \mathbb{P}_x \left(X_{\tau_{i+1}} = x_{i+1}, X_{\tau_{i+2}} = x_{i+2}, \dots \middle| \mathcal{F}_{\tau_i} \right) &= \mathbb{P}_{X_{\tau_i}} \left(X_{\tau_1} = x_{i+1}, X_{\tau_2} = x_{i+2} \right) \\ &= Q(X_{\tau_i}; x_{i+1}, x_{i+2}, \dots). \end{aligned}$$

Using the strong Markov property twice, first by conditioning first with respect to $\mathcal{F}_{\tau_{j+1}}$ and then with respect to \mathcal{F}_{τ_j} ,

$$\begin{aligned} &\mathbb{E}_x \left(\mathbb{1}_{X_{\tau_1} = x_1, \dots, X_{\tau_j} = x_j} F((X_t^{\tau_{j-1}, \tau_j})_{t \geq 0}) \mathbb{1}_{X_{\tau_j} = x_j, X_{\tau_{j+1}} = x_{j+1}, \dots} \right) \\ &= \mathbb{E}_x \left(\mathbb{1}_{X_{\tau_1} = x_1, \dots, X_{\tau_j} = x_j} F((X_t^{\tau_{j-1}, \tau_j})_{t \geq 0}) \mathbb{1}_{X_{\tau_j} = x_j} Q(X_{\tau_j}) \right) \\ &= \mathbb{E}_x \left(\mathbb{1}_{X_{\tau_1} = x_1, \dots, X_{\tau_{j-1}} = x_{j-1}} (X_t^{\tau_{j-1}, \tau_j})_{t \geq 0} \right) Q(x_j) = \mathbb{E}_x \left(\mathbb{1}_{X_{\tau_1} = x_1, \dots, X_{\tau_j} = x_{j-1}} \right) R(x_{j-1}, x_j) Q(x_j) \end{aligned}$$

with $R(x, y) := \mathbb{P}_x \left(F((X_t^{0, \tau_1})_{t \geq 0}) \mathbb{1}_{X_{\tau_1} = y} \right)$. Therefore, using the definition of the conditional expectation,

$$\begin{aligned} \mathbb{E}_x \left(F((X_t^{\tau_{j-1}, \tau_j})_{t \geq 0}) \middle| X_{\tau_1} = x_1, X_{\tau_2} = x_2, \dots \right) \\ = \frac{R(x_{j-1}, x_j)}{\mathbb{P}_{x_{j-1}}(X_{\tau_1} = x_j)} = \mathbb{E}_{x_{j-1}} \left(F((X_t^{0, \tau_1})_{t \geq 0}) \middle| X_{\tau_1} = x_j \right). \end{aligned}$$

This is sufficient to prove the result. □

Bibliography

- [1] Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables. Reprint of the 1972 ed.* English. A Wiley-Interscience Publication. Selected Government Publications. New York: John Wiley & Sons, Inc; Washington, D.C.: National Bureau of Standards. xiv, 1046 pp.; \$ 44.95 (1984). 1984.
- [2] Aurélien Alfonsi. “On the discretization schemes for the CIR (and Bessel squared) processes”. In: *Monte Carlo Methods Appl.* 11.4 (2005), pp. 355–384.
- [3] Larbi Alili and Andrew Aylwin. “On the semi-group of a scaled skew Bessel process”. In: *Statist. Probab. Lett.* 145 (2019), pp. 96–102.
- [4] Madjid Amir. “Sticky Brownian motion as the strong limit of a sequence of random walks”. In: *Stochastic Process. Appl.* 39.2 (1991), pp. 221–237.
- [5] Alexis Anagnostakis. *Functional convergence to the local time of a sticky diffusion.* 2022.
- [6] Alexis Anagnostakis, Antoine Lejay, and Denis Villemonais. *General diffusion processes as the limit of time-space Markov chains.* 2020.
- [7] Stefan Ankirchner, Thomas Kruse, and Mikhail Urusov. “A functional limit theorem for coin tossing Markov chains”. In: *Ann. Inst. Henri Poincaré Probab. Stat.* 56.4 (2020), pp. 2996–3019.
- [8] Stefan Ankirchner, Thomas Kruse, and Mikhail Urusov. “A functional limit theorem for irregular SDEs”. In: *Ann. Inst. Henri Poincaré Probab. Stat.* 53.3 (2017), pp. 1438–1457.
- [9] Stefan Ankirchner, Thomas Kruse, and Mikhail Urusov. *Wasserstein convergence rates for coin tossing approximations of continuous Markov processes.* Preprint arXiv:1903.07880. Mar. 2019.
- [10] M. T. Barlow. “Skew Brownian Motion And A One Dimensional Stochastic Differential Equation”. English. In: *Stochastics* 25.1 (1988), pp. 1–2.
- [11] Richard F. Bass. “A stochastic differential equation with a sticky point”. In: *Electron. J. Probab.* 19 (2014), no. 32, 22.
- [12] R. J. Baxter. “Percus-Yevick Equation for Hard Spheres with Surface Adhesion”. In: *Journal of Chemical Physics* 49 (1968), pp. 2770–2774.
- [13] Fischer Black and Myron Scholes. “The Pricing of Options and Corporate Liabilities”. In: *Journal of Political Economy* 81.3 (1973), pp. 637–654.
- [14] Andrei N. Borodin and Paavo Salminen. *Handbook of Brownian motion—facts and formulae.* Probability and its Applications. Birkhäuser Verlag, Basel, 1996, pp. xiv+462.

- [15] Nawaf Bou-Rabee and Miranda C. Holmes-Cerfon. “Sticky Brownian motion and its numerical solution”. In: *SIAM Rev.* 62.1 (2020), pp. 164–195.
- [16] Nawaf Bou-Rabee and Miranda C. Holmes-Cerfon. “Sticky Brownian motion and its numerical solution”. In: *SIAM Rev.* 62.1 (2020), pp. 164–195.
- [17] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011, pp. xiv+599.
- [18] Robert Brown. “A Brief Account of Microscopical Observations made in the Months of June, July, and August, 1827, on the Particles contained in the Pollen of Plants; and on the General Existence of Active Molecules in Organic and Inorganic Bodies”. In: *The Miscellaneous Botanical Works of Robert Brown*. Ed. by John Joseph Bennett. Vol. 1. Cambridge Library Collection - Botany and Horticulture. Cambridge University Press, 2015, 463–486.
- [19] Robert Harrison Brown. “XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies”. In: *Philosophical Magazine* 4 (), pp. 161–173.
- [20] Àngel Calsina and József Z. Farkas. “Steady states in a structured epidemic model with Wentzell boundary condition”. In: *J. Evol. Equ.* 12.3 (2012), pp. 495–512.
- [21] Alexander S. Cherny and Hans-Jürgen Engelbert. *Singular stochastic differential equations*. English. Vol. 1858. Lect. Notes Math. Berlin: Springer, 2005.
- [22] K. L. Chung and G. A. Hunt. “On the zero $\sum_1^n \pm 1$ ”. English. In: *Ann. Math. (2)* 50 (1949), pp. 385–400.
- [23] John C. Cox, Jonathan E. Ingersoll Jr., and Stephen A. Ross. “A theory of the term structure of interest rates”. In: *Econometrica* 53.2 (1985), pp. 385–407.
- [24] M. J. Davies and A. Truman. “Brownian motion with a sticky boundary and point sources in quantum mechanics”. In: vol. 20. 4-5. Lagrange geometry, Finsler spaces and noise applied in biology and physics. 1994, pp. 173–193.
- [25] David Dereudre, Sara Mazzonetto, and Sylvie Roelly. “An explicit representation of the transition densities of the skew Brownian motion with drift and two semipermeable barriers”. English. In: *Monte Carlo Methods Appl.* 22.1 (2016), pp. 1–23.
- [26] Monroe D. Donsker. “An invariance principle for certain probability limit theorems”. In: *Mem. Amer. Math. Soc.* 6 (1951), p. 12.
- [27] E. B. Dynkin. *Markov processes. Vols. I, II*. Vol. 122. Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone. Die Grundlehren der Mathematischen Wissenschaften, Bände 121. Academic Press Inc., Publishers, New York; Springer-Verlag, Berlin-Göttingen-Heidelberg, 1965, Vol. I: xii+365 pp.; Vol. II: viii+274.
- [28] Andreas Eberle and Raphael Zimmer. “Sticky couplings of multidimensional diffusions with different drifts”. In: *Ann. Inst. Henri Poincaré Probab. Stat.* 55.4 (2019), pp. 2370–2394.
- [29] Albert B. Einstein. “On the Motion of Small Particles Suspended in Liquids at Rest Required by the Molecular-Kinetic Theory of Heat”. In: 2004.

- [30] Hans-Jürgen Engelbert and Goran Peskir. “Stochastic differential equations for sticky Brownian motion”. In: *Stochastics* 86.6 (2014), pp. 993–1021.
- [31] Pierre Étoré and Antoine Lejay. “A Donsker theorem to simulate one-dimensional processes with measurable coefficients”. In: *ESAIM Probab. Stat.* 11 (2007), pp. 301–326.
- [32] W. Feller. *An introduction to probability theory and its applications. Vol II. 2nd ed.* English. Wiley Ser. Probab. Math. Stat. John Wiley & Sons, Hoboken, NJ, 1971.
- [33] William Feller. “The parabolic differential equations and the associated semigroups of transformation”. English. In: *Ann. Math. (2)* 55 (1952), pp. 468–519.
- [34] William Feller. “Diffusion processes in one dimension”. English. In: *Trans. Am. Math. Soc.* 77 (1954), pp. 1–31.
- [35] William Feller. “Generalized second order differential operators and their lateral conditions”. English. In: *Ill. J. Math.* 1 (1957), pp. 459–504.
- [36] William Feller. “On boundaries and lateral conditions for the Kolmogorov differential equations”. English. In: *Ann. Math. (2)* 65 (1957), pp. 527–570.
- [37] William Feller. “On the intrinsic form for second order differential operators”. English. In: *Ill. J. Math.* 2 (1958), pp. 1–18.
- [38] William Feller. “Two singular diffusion problems”. In: *Ann. of Math. (2)* 54 (1951), pp. 173–182.
- [39] Anna Ferrer-Admetlla et al. “An Approximate Markov Model for the Wright-Fisher Diffusion and Its Application to Time Series Data”. In: *Genetics* 203 (Apr. 2016).
- [40] Noufel Frikha. “On the weak approximation of a skew diffusion by an Euler-type scheme”. English. In: *Bernoulli* 24.3 (2018), pp. 1653–1691.
- [41] Masatoshi Fukushima. *Feller’s Contributions to the One-Dimensional Diffusion Theory and Beyond.* William Feller Selected Papers II. Springer, 2015, pp. 63–76.
- [42] Alberto Gandolfi, Anna Gerardi, and Federico Marchetti. “Association rates of diffusion-controlled reactions in two dimensions”. English. In: *Acta Appl. Math.* 4 (1985), pp. 139–159.
- [43] Domenico Gazzillo et al. “Multicomponent adhesive hard sphere models and short-ranged attractive interactions in colloidal or micellar solutions”. In: *Physical review. E, Statistical, nonlinear, and soft matter physics* 74 (Dec. 2006), p. 051407.
- [44] Carl Graham. “Homogenization and propagation of chaos to a nonlinear diffusion with sticky reflection”. English. In: *Probab. Theory Relat. Fields* 101.3 (1995), pp. 291–302.
- [45] Adam Gregosiewicz. *Sticky diffusions on graphs.* 2022.
- [46] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and random processes.* Third. Oxford University Press, New York, 2001, pp. xii+596.
- [47] J. M. Harrison and L. A. Shepp. “On skew Brownian motion”. English. In: *Ann. Probab.* 9 (1981), pp. 309–313.

- [48] J. Michael Harrison and Austin J. Lemoine. “Sticky Brownian motion as the limit of storage processes”. In: *J. Appl. Probab.* 18.1 (1981), pp. 216–226.
- [49] Einar Hille and Ralph S. Phillips. *Functional analysis and semi-groups*. 3rd printing of rev. ed. of 1957. English. Vol. 31. Colloq. Publ., Am. Math. Soc. American Mathematical Society (AMS), Providence, RI, 1974.
- [50] Miranda Holmes-Cerfon, Steven J. Gortler, and Michael P. Brenner. “A geometrical approach to computing free-energy landscapes from short-ranged potentials”. In: *Proceedings of the National Academy of Sciences* 110.1 (2013), E5–E14.
- [51] Martin Hutzenhaller and Arnulf Jentzen. *Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients*. English. Vol. 1112. Mem. Am. Math. Soc. Providence, RI: American Mathematical Society (AMS), 2015.
- [52] Kiyosi Itô. *Essentials of stochastic processes. Translated from the 1957 Japanese original*. English. Vol. 231. Transl. Math. Monogr. Providence, RI: American Mathematical Society (AMS), 2006.
- [53] Kiyosi Itô and Henry P. jun. McKean. *Diffusion processes and their sample paths*. English. Berlin: Springer-Verlag, 1996, pp. xviii + 326.
- [54] J. Jacod. *Calcul stochastique et problèmes de martingales*. French. Vol. 714. Lect. Notes Math. Springer, Cham, 1979.
- [55] Jean Jacod. “Rates of convergence to the local time of a diffusion”. In: *Ann. Inst. H. Poincaré Probab. Statist.* 34.4 (1998), pp. 505–544.
- [56] Jean Jacod and Philip Protter. *Discretization of processes*. Vol. 67. Stochastic Modelling and Applied Probability. Springer, Heidelberg, 2012, pp. xiv+596.
- [57] Arturo Jaramillo, Ivan Nourdin, and Giovanni Peccati. “Approximation of fractional local times: zero energy and derivatives”. English. In: *Ann. Appl. Probab.* 31.5 (2021), pp. 2143–2191.
- [58] Yuri Kabanov, Masaaki Kijima, and Sofiane Rinaz. “A positive interest rate model with sticky barrier”. In: *Quantitative Finance* 7.3 (2007), pp. 269–284.
- [59] Olav Kallenberg. *Foundations of modern probability*. Second. Probability and its Applications (New York). Springer-Verlag, New York, 2002, pp. xx+638.
- [60] Yoav Kallus and Miranda Holmes-Cerfon. “Free energy of singular sticky-sphere clusters”. In: *Phys. Rev. E* 95 (2 2017), p. 022130.
- [61] Yoav Kallus and Miranda Holmes-Cerfon. “Free energy of singular sticky-sphere clusters”. In: *Physical Review E* 95.2 (2017).
- [62] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*. Second. Vol. 113. Graduate Texts in Mathematics. Springer-Verlag, New York, 1991, pp. xxiv+470.
- [63] Samuel Karlin and Howard M. Taylor. *A second course in stochastic processes*. English. New York etc.: Academic Press, A Subsidiary of Harcourt Brace Jovanovich, Publishers. XVIII, 542 p. \$ 35.00 (1981). 1981.
- [64] Kei Kobayashi. “Stochastic calculus for a time-changed semimartingale and the associated stochastic differential equations”. In: *J. Theoret. Probab.* 24.3 (2011), pp. 789–820.

- [65] Vadim Kostykin, Jürgen Potthoff, and Robert Schrader. “Brownian motions on metric graphs”. English. In: *J. Math. Phys.* 53.9 (2012), pp. 095206, 36.
- [66] Antoine Lejay. “On the constructions of the skew Brownian motion”. In: *Probab. Surv.* 3 (2006), pp. 413–466.
- [67] Antoine Lejay, Lionel Lenôtre, and Géraldine Pichot. “Analytic expressions of the solutions of advection-diffusion problems in one dimension with discontinuous coefficients”. English. In: *SIAM J. Appl. Math.* 79.5 (2019), pp. 1823–1849.
- [68] Antoine Lejay, Ernesto Mordecki, and Soledad Torres. “Is a Brownian motion skew?”. English. In: *Scand. J. Stat.* 41.2 (2014), pp. 346–364.
- [69] Antoine Lejay and Paolo Pigato. “Statistical estimation of the oscillating Brownian motion”. In: *Bernoulli* 24.4B (2018), pp. 3568–3602.
- [70] Robert S. Liptser and Albert N. Shiryaev. *Statistics of random processes. I.* expanded. Vol. 5. Applications of Mathematics (New York). General theory, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2001, pp. xvi+427.
- [71] Petr Mandl. *Analytical treatment of one-dimensional Markov processes.* English. Vol. 151. Grundlehren Math. Wiss. Berlin-Heidelberg-New York: Springer Verlag, 1968.
- [72] Sara Mazzonetto. *Rates of convergence to the local time of Oscillating and Skew Brownian Motions (preprint).* 2019.
- [73] Christian Meier, Lingfei Li, and Zhang Gongqiu. “Markov Chain Approximation of One-Dimensional Sticky Diffusions”. In: (Oct. 2019). Preprint arXiv:1910.14282.
- [74] Christian Meier, Lingfei Li, and Gongqiu Zhang. *Simulation of Multidimensional Diffusions with Sticky Boundaries via Markov Chain Approximation.* 2021.
- [75] Guangnan Meng et al. “The Free-Energy Landscape of Clusters of Attractive Hard Spheres”. In: *Science* 327.5965 (2010), pp. 560–563.
- [76] Yutian Nie and Vadim Linetsky. “Sticky reflecting Ornstein-Uhlenbeck diffusions and the Vasicek interest rate model with the sticky zero lower bound”. In: *Stochastic Models* 0.0 (2019), pp. 1–19.
- [77] Bernt Øksendal. *Stochastic differential equations. An introduction with applications.* English. Berlin: Springer, 1998, pp. xix + 324.
- [78] Goran Peskir. “On Boundary Behaviour of One-Dimensional Diffusions: From Brown to Feller and Beyond”. In: 2014.
- [79] Mark Podolskij and Mathieu Rosenbaum. “Comment on: Limit of Random Measures Associated with the Increments of a Brownian Semimartingale: Asymptotic behavior of local times related statistics for fractional Brownian motion”. In: *Journal of Financial Econometrics* 16.4 (Dec. 2017), pp. 588–598.
- [80] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion.* Third. Vol. 293. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pp. xiv+602.
- [81] L. C. G. Rogers. “Arbitrage with fractional Brownian motion”. English. In: *Math. Finance* 7.1 (1997), pp. 95–105.

- [82] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Itô calculus, Reprint of the second (1994) edition. Cambridge University Press, Cambridge, 2000, pp. xiv+480.
- [83] Michael Salins and Konstantinos Spiliopoulos. “Markov processes with spatial delay: path space characterization, occupation time and properties”. English. In: *Stoch. Dyn.* 17.6 (2017). Id/No 1750042, p. 21.
- [84] Elias M. Stein and Rami Shakarchi. *Real analysis. Measure theory, integration, and Hilbert spaces*. English. Vol. 3. Princeton Lect. Anal. Princeton, NJ: Princeton University Press, 2005.
- [85] George Stell. “Sticky spheres and related systems”. In: *Journal of Statistical Physics* 63 (1991), pp. 1203–1221.
- [86] George Stell. “Sticky spheres and related systems”. In: *J. Statist. Phys.* 63.5-6 (1991), pp. 1203–1221.
- [87] Terence Tao. *An introduction to measure theory*. English. Vol. 126. Providence, RI: American Mathematical Society (AMS), 2011, pp. xvi + 206.
- [88] Wajdi Touhami. “On skew sticky Brownian motion”. English. In: *Stat. Probab. Lett.* 173 (2021). Id/No 109086, p. 9.
- [89] H. F. Trotter. “A property of Brownian motion paths”. In: *Illinois J. Math.* 2 (1958), pp. 425–433.
- [90] Cédric Villani. *Optimal transport*. Vol. 338. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Old and new. Springer-Verlag, Berlin, 2009, pp. xxii+973.
- [91] N. Wiener. “Differential-space.” English. In: *Bull. Am. Math. Soc.* 29 (1923), p. 105.
- [92] Marc Yor. “Un exemple de processus qui n’est pas une semi-martingale”. In: *Temps locaux. Astérisque* 52-53 (1978).