THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'École Normale Supérieure
Dans le cadre d'une cotutelle avec l'Université Paris-Saclay

Homologie persistante des processus stochastiques et leurs fonctions zêta

Persistent homology of stochastic processes and their zeta functions

Soutenue par
Daniel PEREZ
Le 11 juillet 2022

École doctorale n°574
École Doctorale de Mathématiques Hadamard

Spécialité
Mathématiques

Préparée au
Département de Mathématiques et Applications

Composition du jury :

Élisabeth GASSIAT Professeure, Université Paris-Saclay	*Présidente du jury* *Examinatrice*
Shmuel WEINBERGER Professeur, University of Chicago	*Rapporteur*
Gérard BEN AROUS Professeur, New York University	*Rapporteur*
Grégory GINOT Professeur, Université Paris 13	*Examinateur*
Dmitry CHELKAK Professeur, École normale supérieure	*Examinateur*
Claude VITERBO Professeur, Université Paris-Saclay	*Directeur de thèse*
Pierre PANSU Professeur, Université Paris-Saclay	*Invité* *Codirecteur de thèse*

Remerciements

J'aimerais remercier mes deux directeurs de thèse, Pierre Pansu et Claude Viterbo, sans qui ce travail n'aurait pas été possible. Leur bienveillance et l'aide qu'ils m'ont fournie au cours des trois dernières années a été clé dans l'aboutissement de cette thèse et sa réussite.

I would also like to thank Shmuel Weinberger and Gérard Ben Arous for refereeing this manuscript, for the careful attention they have given it and for the numerous discussions I've had the pleasure to have with them. By the same token, I wish to thank the jury of this thesis for their time and presence. Many thanks are also owed to Yuliy Baryshnikov for taking the time to look at and discuss my work. J'aimerais aussi remercier Jean-François Le Gall, Nicolas Curien, Frédéric Chazal et Pascal Massart avec qui j'ai eu la chance d'échanger plusieurs fois et dont les nombreuses discussions ont orienté plusieurs de axes de recherche que j'ai pu explorer, dont certains figurent dans cette thèse. Enfin, je remercie Vincent Divol et Théo Lacombe pour les quelques discussions que l'on a pu avoir au cours de ces dernières années, qui ont aussi eu une grande influence sur ma recherche.

Les trois dernières années ont été tumultueuses, la présence des équipes du DMA et du LMO ont fait de ma thèse une expérience constructive tant sur le plan professionnel que personnel. Grâce à eux, je garde un souvenir très positif des dernières années et de ma thèse en général. En particulier, j'aimerais remercier Matija Srećković (hvala brate !), Ricardo Grande, Maxime Berger et Corentin Gentil (que je dois aussi remercier pour une relecture), Rémy Mahfouf, Mickaël Latocca, Pierre Gervais et Geoffrey Beck pour leur amitié, leur support, le temps qu'on a passé ensemble et les nombreuses discussions que nous avons pu avoir au cours des derniers trois ans. J'aimerais aussi exprimer ma plus sincère gratitude à Zaina Elmir et Fabienne Renia pour leur soutient tant administratif que moral ainsi qu'au reste de l'équipe du DMA, en particulier Laurent Moonens, qui a toujours su me remonter le moral avec du thé ou du chocolat ! Du côté d'Orsay, j'aimerais remercier Nicolas Camps, Wojciech Reise et le reste de l'équipe DataShape que j'ai aussi eu le plaisir de côtoyer.

Il est impératif que je remercie Alban, Boris, Jadwiga, Gabin, Silviu, Myrto, Ana et Guillaume que j'ai eu la chance de me lier d'amitié dans le cadre de mes activités dans le monde associatif et qui ont été un support important. Évidemment, je n'oublie pas non plus Laure Monnoyer, que je dois remercier pour la lecture d'une partie de ce manuscrit et avec qui on ne s'ennuie jamais ainsi que Philippe Meyer, qui a aussi relu une partie de cette thèse et avec qui les échanges sont toujours intéressants. I also want to thank Daniel Kroff and Gabriela Fernandes for their support and in the case of Daniel for his invaluable LaTeX help whenever I needed it.

También deseo darle las gracias a mi familia, en particular a mi mamá y mi papá, quienes siempre ma han apoyado, a pesar de que no siempre ha sido facil de conciliar distancia con tiempo pasado juntos, en particular durante estos últimos años de pandemia. De igual forma quisiera darle las gracias a todos mis primos y especialmente a Diana, Juliana y Jesús, quienes siempre me han tanto soportado como apoyado. Aprovecho el parrafo en español para darle las gracias a mis amigos Jaime, Ana, Manuel, Carolina y Angelica por su amistad, y también por todos los momentos que hemos pasado juntos estos últimos años.

Diese Danksagung wäre nicht vollständig, wenn ich nicht auch Jan Maelger erwähnen würde, dessen Unterstützung während des letzten Jahres von unschätzbarem Wert war, und Georg Grützner für die zahlreichen interessanten Diskussionen, die wir in den letzten Jahren hatten.

Finally, to everyone else, whose names this brief acknowledgement section is too narrow to contain.

Résumé

Cette thèse étudie l'homologie persistante des fonctions continues à valeurs réelles f sur des espaces topologiques compacts X. L'introduction des *indices homologiques* et des *dimensions homologiques* permet de lier la théorie de la persistance à des quantités métriques de l'espace compact X, telles que sa dimension. L'étude de ces quantités permet de d'étendre les résultats de stabilité des distances Wasserstein p sur l'espace des diagrammes de persistance aux fonctions höldériennes sur des espaces métriques plus généraux que ceux précédemment établis par la littérature (qui incluent en particulier toutes les variétés riemanniennes compactes) avec des constantes explicites. En degré d'homologie zéro, une étude plus approfondie peut être réalisée à l'aide d'arbres associés à f, qui généralisent les *merge trees* définissables lorsque f est de Morse. Il est possible de lier la dimension de ces arbres à l'indice de persistance de f et à son code-barres. Nous appliquons ces résultats déterministes au cadre stochastique pour en tirer des conséquences sur les code-barres de fonctions aléatoires de régularité prescrite. Ces conséquences permettent en outre d'élaborer des tests de discrimination de distribution des processus, dont nous présentons un exemple particulier. Enfin, nous définissons les fonctions ζ associés à un processus stochastique et nous calculons ces fonctions d'autres quantités annexes pour plusieurs processus en dimension 1, ainsi que le mouvement Brownien et les processus de Lévy α-stables.

Mots clés : homologie, persistance, code-barres, topologie, probabilité

Abstract

This thesis studies the persistent homology of $\mathbb{R}$-valued continuous functions f on compact topological spaces X. The introduction of *homological indices* and *homological dimensions* allows us to link persistence theory to metric quantities of the compact space X, such as its upper-box dimension. These quantities give a precise framework to the Wasserstein p-stability results known in the literature, but also extend them to Hölder functions on more general spaces (including all compact Riemannian manifolds) with explicit constants and whose regime for p is optimal. In degree zero of homology, a more in-depth study can be made using trees associated to f, which generalize the *merge trees* definable when f is Morse. It is possible to link the dimension of these trees to the persistence index of f and to its barcode. We apply these deterministic results to the stochastic setting to draw consequences about the barcodes of random functions of prescribed regularity. These consequences also allow us to develop distributional discrimination tests for the processes, of which we present a particular example. Finally, we define the ζ-functions associated with a stochastic process and compute these functions and other related quantities for several processes in dimension one, including the Brownian motion and the α-stable Lévy processes.

Keywords : homology, persistence, barcodes, topology, probability

Contents

Chapter 1

Introduction

Dans cette thèse, nous tenterons de comprendre l'homologie des ensembles de sur-niveau d'une fonction $f : X \to \mathbb{R}$ sur certains espaces topologiques X et sous diverses hypothèses de régularité de f. Nous considérerons aussi le cas où la fonction f est aléatoire.

Cette question est particulièrement bien étudiée en degré 0 d'homologie, où il suffit de compter le nombre de composantes connexes des ensembles de sur-niveau. De multiples ouvrages dans certains cadres spécifiques traitent de cette question, dont en particulier un qui a motivé cette thèse : l'étude par Nabutovsky du nombre de minima dits « très profonds » de certaines fonctionnelles sur des variétés riemanniennes. Les travaux de Nabutovsky ne sont pas les seuls exemples intéressants d'étude du H_0 des ensembles de sur-niveau (ou sous-niveau). En théorie des probabilités, la recherche sur les propriétés fines du mouvement brownien a motivé l'introduction d'arbres définis à partir d'une certaine fonction $f : [0,1] \to \mathbb{R}$ qui encodent l'information relative aux composantes connexes des ensembles de sur-niveau [29, 37, 38, 41, 80].

Depuis les années 1990, une approche systématique a été dévéloppée pour comprendre l'homologie des ensembles de sur-niveau (et sa variation à mesure que nous changeons le niveau) pour certains espaces métriques et des fonctions assez régulières : l'homologie persistante. Cette branche de la topologie algébrique a été introduite par Barannikov dans le contexte de la théorie Morse en 1994 [8], puis dans un tout autre contexte par la communauté de l'analyse topologique de données (dite, communauté TDA pour *topological data analysis*) (*cf.* [24, 76] et les références qui y figurent).

Le reste de cette introduction se présente comme suit. D'abord, nous présenterons brièvement l'homologie persistante, de sorte à familiariser le lecteur aux concepts de base de cette théorie s'il ne l'est pas déjà. Puis nous discuterons des résultats pré-existants et du contexte bibliographique dans lequel s'inscrit cette thèse avant de présenter nos résultats en deux volets : les résultats déterministes et les résultats obtenus dans un cadre probabiliste.

1 Une introduction à l'homologie persistante

Tout au long de cette section, nous allons définir et tenter d'appréhender la notion d'homologie persistante. Pour ce faire, il est commode de segmenter cette section selon son nom. D'abord, nous rappellerons brièvement ce qu'est l'*homologie* et comment elle peut être définie, puis nous expliquerons le caractère *persistant* de l'homologie persistante. Pour une introduction plus approfondie, nous encourageons le lecteur à consulter les références classiques suivantes sur ce sujet : [24, 49, 65, 76]. Un lecteur familier avec ces concepts pourra commencer sa lecture par la section suivante.

1.1 Homologie

La motivation générale derrière la plupart des concepts en topologie algébrique est d'associer un objet algébrique (typiquement plus facile à étudier, tel qu'un module, un groupe, *etc.*) à un espace topologique, de telle sorte que, au sens large, cet objet algébrique reste invariant pour des espaces topologiques homéomorphes. De plus, nous demandons que ces invariants se comportent bien par rapport aux fonctions continues. Si nous introduisons un invariant algébrique $A(X)$, construit à partir d'un espace topologique X, nous demandons que la condition suivante soit satisfaite : si $f : X \to Y$ est une fonction continue, f induit un morphisme $A(f) : A(X) \to A(Y)$ entre les invariants algébriques de X et Y. Un exemple notable et sans doute le plus célèbre d'un tel invariant est le groupe fondamental $\pi_1(X)$ originellement introduit par Poincaré [83].

Ces concepts s'expriment de manière plus commode dans le language de la théorie des catégories, nous parlons alors de *foncteurs* entre la catégorie des espaces topologiques, **Top**, et une catégorie d'objets algébriques (comme la catégorie des groupes, **Grp**, ou celle des modules sur un anneau A, $\mathbf{Mod}_A$). En ce sens, l'homologie est un foncteur $H_* : \mathbf{Top} \to \mathbf{Mod}_A$.

Dans la suite de cette thèse, nous ferons toujours l'hypothèse que l'anneau A est en fait un corps k, de sorte que nous ayons $H_* : \mathbf{Top} \to \mathbf{Vect}_k$, où $\mathbf{Vect}_k$ est la catégorie des k-espaces vectoriels. Cette restriction permet d'éviter des complications algébriques importantes qui surviennent lorsque nous travaillons sur un anneau (notamment des problèmes d'extension et de torsion) et permet un exposé plus simple de la théorie. Nous noterons cependant que l'homologie persistante peut quand même être définie à coefficients dans un anneau ($\mathbb{Z}$ par exemple) [90], au prix d'une complication de la théorie. Nous ne détaillerons pas ici les définitions de ces notions relatives à la théorie des catégories, mais nous renvoyons le lecteur au livre introductif de Mac Lane [60] et vers les ouvrages de Hatcher et MacLane sur la topologie algébrique [49, 65] pour une discussion plus approfondie sur les foncteurs et leur apparition en topologie algébrique.

Pour définir l'homologie, concentrons-nous d'abord sur les complexes simpliciaux (un ensemble de simplexes *orientés* collés les uns aux autres le long de points, d'arêtes ou de n-faces). Cette définition permettra aussi de définir l'homologie pour les espaces topologiques dits *triangulables*, *i.e.* homéomorphes à un complexe simplicial.

Définition 1.1. Soit X un complexe simplicial. L'**espace de chaînes** est l'espace vectoriel gradué noté $C_*(X, k)$ (l'étoile désigne le **degré**), où $C_n(X, k)$ est le k-espace vectoriel libre

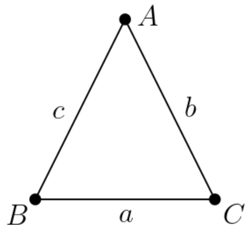


Figure 1.1: Un exemple de complexe simplicial X.

engendré par l'ensemble des n-faces du complexe simplicial. Lorsqu'il n'y aura pas d'ambiguité, nous noterons simplement $C_*(X,k) = C_*$. Nous garderons cette notation tout le long de cette section.

Définition 1.2. L'**application bord** $\partial : C_* \to C_*$ est une application linéaire définie degré par degré comme suit: les générateurs des n-chaînes sont envoyés sur la somme (signée) des générateurs de leurs bords, qui sont des éléments de C_{n-1}. Le signe devant chaque générateur du bord d'une n-face est déterminé par la compatibilité de l'orientation du générateur du bord avec l'orientation de la n-face.

Remarque 1.3. Afin d'éviter les complications relatives au signe provenant de l'orientation du complexe simplicial, il est commode de travailler sur $k = \mathbb{Z}_2$, ce que nous ferons dorénavant.

Avec cette définition, $\partial^2 = 0$, ce qui traduit le fait que le bord d'un bord est toujours vide. Ainsi,

Définition 1.4. Un **complexe de chaînes** est un espace vectoriel gradué C_* muni d'une application bord ∂ de degré -1 satisfaisant $\partial^2 = 0$.

Nous pouvons écrire ∂ comme une suite de morphismes

$$\cdots \to C_2 \xrightarrow{\partial} C_1 \xrightarrow{\partial} C_0 \xrightarrow{\partial} 0\,, \tag{1.1}$$

ayant la propriété que $\partial^2 = 0$. Cette propriété implique en particulier que $\mathrm{Im}(\partial) \subset \ker(\partial)$.

Définition 1.5. Nous appelons $\ker(\partial)$ l'**espace des cycles**.

L'homologie quantifie exactement quels cycles ne sont pas des bords. En termes algébriques, l'homologie mesure le défaut d'exactitude du complexe de chaînes. Plus précisément,

Définition 1.6. En fixant un degré n, nous pouvons définir le **n-ème groupe d'homologie sur un corps** k comme étant

$$H_n(X,k) := \ker(\partial|_{C_n}) / \,\mathrm{Im}(\partial|_{C_{n+1}})\,, \tag{1.2}$$

où $\partial|_{C_n}$ désigne la restriction de ∂ à C_n. Nous noterons $H_*(X,k)$ l'espace vectoriel gradué engendré par les $H_n(X,k)$.

L'homologie permet en outre de définir ce que nous entendons par un trou n-dimensionnel. Il est plus facile d'illustrer ceci par le biais d'un exemple. Afin de simplifier cette illustration, fixons $k = \mathbb{Z}_2$ et fixons le complexe simplicial X de la figure 1.1. Dans ce cas, $C_0 = \langle A, B, C\rangle_{\mathbb{Z}_2}$ et $C_1 = \langle a, b, c\rangle_{\mathbb{Z}_2}$ et tous les espaces de chaînes de degré supérieurs sont l'espace vectoriel $\{0\}$. L'application bord vérifie

$$\partial : a \mapsto B + C\ ,\ b \mapsto A + C\ ,\ c \mapsto A + B\,, \tag{1.3}$$

et tous les autres générateurs sont envoyés vers zéro. L'espace des cycles restreint à C_1, $\ker(\partial|_{C_1}) = (a + b + c)\mathbb{Z}_2$, est ici entièrement engendré par le cycle qui fait un tour autour du complexe simplicial. De même, il est facile de voir que $\ker(\partial|_{C_0}) = C_0$. Il s'ensuit que $H_1(X) = (a + b + c)\mathbb{Z}_2$, de sorte que nous détectons effectivement un trou « 1-dimensionnel » à l'intérieur du complexe.

Ajoutons maintenant à X une 2-face que nous noterons Δ, dont le bord est $a + b + c$. Nous avons alors $C_2 = \langle\Delta\rangle_{\mathbb{Z}_2}$ et $\partial(\Delta) = a+b+c$. En remplissant le « trou », nous avons effectivement tué le H_1 du nouveau complexe $X \cup \Delta$. Dans cet exemple rudimentaire, nous voyons en quoi les « trous » d'un espace sont bien caractérisés par le fait d'être des cycles qui ne sont pas des bords.

L'homologie en degré 0, H_0, nous intéressera particulièrement tout le long de cette thèse. Ce degré d'homologie quantifie toujours le nombre de composantes connexes d'un espace. Dans le cas des complexes simpliciaux (finis), ces derniers sont connexes si et seulement s'ils sont connexes par arcs. Ces arcs engendrent des éléments de C_1 ayant pour bord n'importe quelle paire d'éléments de C_0 (pourvu que la paire soit dans la même composante connexe de X). Ainsi, $\mathrm{Im}(\partial|_{C_1})$ est toujours engendré par des sommes de paires de sommets dans C_0 vivant dans la même composante connexe, d'où une correspondance entre les générateurs de H_0 et les composantes connexes de X.

À ce stade, étant donné un espace triangulable Y, le lecteur pourrait se demander si cette définition dépend de la triangulation choisie pour Y. Il s'avère que non et que l'homologie ainsi définie est un invariant bien défini (*cf.* [49]). En fait, le point de vue que nous avons adopté a l'avantage d'être clair, mais il est plutôt restrictif, car nous pouvons donner un sens à l'homologie dans des cadres bien plus généraux [40, 65]. D'ailleurs, nous aurons besoin de ce degré de généralité supplémentaire (plus précisément, de la définition de l'homologie de Čech, *cf.* [40, Chapitre IX]), étant donné que nous considérerons l'homologie d'espaces topologiques qui ne seront plus nécessairement triangulables.

1.2 Homologie persistante

Ayant grossièrement esquissé ce qu'est l'homologie, du moins dans le cadre des espaces triangulables, attardons-nous sur l'homologie persistante. Nous renvoyons à nouveau le lecteur aux références classiques pour une description plus détaillée de la théorie [24, 76].

Dans ce contexte, nous munissons le complexe de chaînes $C_*(X)$ d'une filtration, c'est-à-dire que $C_*(X)$ est maintenant filtré par des sous-espaces $(C_*^r)_r$ indexés par un ensemble totalement ordonné de telle sorte que pour tout $r \geq s$, $C_*^r \subset C_*^s$. Cette filtration est typiquement prescrite par les ensembles de sur-niveau d'une fonction $f : X \to \mathbb{R}$, auquel cas les inclusions ci-dessus sont induites par l'injection des sur-niveaux les uns dans les autres.

Définition 1.7. La filtration ainsi induite est appelée la **filtration par sur-niveaux** (une définition analogue peut être donnée pour la **filtration par sous-niveaux**)

Par exemple, si X est une variété riemannienne (et en particulier, donc, un espace triangulable) et $f : X \to \mathbb{R}$ est une fonction de Morse, nous pouvons filtrer le complexe $C_*(X)$ par les ensembles de sur-niveau $X_r = \{f \geq r\}$. La filtration par sur-niveaux est alors donnée par $(C_*(X_r))_{r\in\mathbb{R}}$. Pour chaque $r \in \mathbb{R}$, nous pouvons calculer l'homologie de X_r. Nous obtenons ainsi une famille d'espaces vectoriels indexés par r. Du fait du caractère fonctoriel de H_*, pour chaque $r > s$, l'inclusion (continue) $i_{r,s} : X_r \hookrightarrow X_s$ induit une application linéaire $H_*(i_{r,s}) : H_*(X_r) \to H_*(X_s)$. L'ensemble de ces morphismes nous renseigne sur la manière dont l'homologie change à mesure que nous varions r. L'homologie persistante est la famille d'espaces vectoriels $(H_*(X_r))_r$ *et* la famille de morphismes $(H_*(i_{r,s}))_{r>s}$. Ceci s'exprime encore une fois plus confortablement en termes catégoriques.

Définition 1.8. L'**homologie persistante** est un foncteur $H_*(X,f) : \mathbb{R} \to \mathbf{Vect}_k$ défini par

$$H_*(X,f)(r) := H_*(X_r) \quad \text{et} \quad H_*(X,f)(r \to s) := H_*(i_{r,s})\,, \tag{1.4}$$

où $\mathbb{R}$ est ici vu comme la catégorie induite par la relation d'ordre partiel sur $\mathbb{R}$, à savoir la catégorie dont les objets sont des éléments de $\mathbb{R}$ telle qu'il existe un morphisme $r \to s$ qui est unique si et seulement si $r > s$.

Si la fonction f est suffisamment régulière, par exemple de Morse, et que X est compact, l'homologie persistante provenant de la filtration par sur-niveaux de f peut être décomposée en modules dits d'*intervalles*. Ces derniers sont eux-mêmes des foncteurs définis comme suit.

Définition 1.9. Soit k un corps et $A \subset \mathbb{R}$ un intervalle, alors le **module d'intervalle** k_A est défini par

$$k_A(r) := \begin{cases} k & \text{si } r \in A \\ 0 & \text{sinon} \end{cases} \quad \text{et} \quad k_A(r \to s) = \begin{cases} \mathrm{id} & \text{si } r,s \in A \\ 0 & \text{sinon} \end{cases} \tag{1.5}$$

Remarque 1.10. Ces modules peuvent être vus comme des « indicatrices » dans un contexte fonctoriel.

Le théorème de décomposition énonce que sous certaines hypothèses, comprendre les modules d'intervalle suffit, car nous pouvons décomposer les modules de persistance en termes de ces derniers (voir le livre d'Oudot [76] pour une description plus complète).

Théorème 1.1 (Théorème de décomposition, Auslander, Ringel, Tachikawa, Gabriel, Azumaya)**.** *Sous certaines conditions sur X et f, si $H_*(X,f)$ désigne l'homologie persistante à valeurs dans $\mathbf{Vect}_k$, alors $H_*(X,f)$ est isomorphe à une somme directe (éventuellement infinie) de modules d'intervalles. De plus, cette décomposition est unique à isomorphisme et permutation des termes près.*

Autrement dit, si f et X sont suffisamment réguliers, en fixant un degré d'homologie n,

$$H_n(X, f) = \bigoplus_i k_{A_i} \, , \tag{1.6}$$

où les A_i sont des intervalles de $\mathbb{R}$. Nous retiendrons alors que l'information contenue dans le module de persistance est encodée dans les A_i. La collection de ces intervalles est ce que nous appelerons le **code-barres de degré** n associé à $f : X \to \mathbb{R}$, usuellement noté $\mathcal{B}_n(f)$. Nous noterons $\mathcal{B}(f)$ la somme directe sur les degrés et parfois, par abus de langage nous noterons $\mathcal{B}(f) = \mathcal{B}_0(f)$ en fonction du contexte. Une manière équivalente, mais parfois plus commode, de représenter cette information est de ne retenir que les extrémités de l'intervalle. Ainsi, nous pouvons représenter les intervalles comme un multi-ensemble de points dans le demi-plan

$$\mathcal{X} := \{(x, y) \in (\mathbb{R} \cup \{\pm\infty\})^2 \,|\, x < y\} \, . \tag{1.7}$$

Ce multi-ensemble de points de $\mathcal{X}$ est appelée le **diagramme de persistance** associé à f, et est noté $\mathrm{Dgm}_n(f)$ (ou $\mathrm{Dgm}(f)$ si le degré est implicite ou s'il s'agit d'une somme directe sur les degrés).

En réalité il y a quelques difficultés techniques provenant de l'énoncé du théorème de décomposition, qui requiert d'avoir X et f « assez réguliers ». Ces difficultés proviennent du fait que les espaces $H_*(X_r)$ peuvent *a priori* être de dimension infinie. En revanche, si le rang des morphismes $H_*(i_{r,s})$ demeure toujours fini, cela suffit pour garantir l'applicabilité du théorème de décomposition (*cf.* [23, 76] pour les détails).

Définition 1.11. Un couple (X, f) satisfaisant la condition de rang fini est dit **q-tame**.

Remarque 1.12. Si X est un espace compact et f est continue, alors le couple (X, f) est q-tame.

En effet, l'homologie persistante de tout couple q-tame (X, f) est toujours décomposable, au prix d'introduire une catégorie dite *observable* des modules de persistance [23]. Cette catégorie est définie comme la catégorie quotient des modules q-tame par des modules dits éphémères, *i.e.* de la forme $k_{\{a\}}$, au sens de la théorie de la localisation de Serre [23]. En particulier, cela signifie que nous pouvons restreindre les intervalles de la décomposition à être de la forme $[a, b[$ (avec $a < b$). Modulo ces points techniques, nous pourrons parler de diagrammes de persistance et de code-barres pour tous les couples (X, f) considérés dans cette thèse, puisqu'ils satisferont toujours la propriété q-tame.

Illustrons ces propos dans le cas de H_0. Dans ce cas, $\dim H_0(X_r)$ est le nombre de composantes connexes de X_r. Le rang de $H_0(i_{r,s})$ correspond au nombre de composantes connexes de X_s qui contiennent les composantes connexes de X_r. Enfin, le théorème de décomposition peut aussi être facilement compris : les barres dans le code-barres indiquent le moment où des composantes connexes sont « nées » et les moments où elles fusionnent avec une autre composante connexe, avec la règle que la composante connexe « la plus ancienne » (celle qui est née le plus tôt) est celle qui « survit ».

2 État de l'art

2.1 Notions de distance sur les diagrammes

Distances provenant du transport optimal partiel

Suivant les travaux de Divol et Lacombe [33], nous introduisons le transport optimal *partiel*, qui étend la théorie du transport optimal à des mesures de masses différentes (et qui peuvent être potentiellement infinies). Un exposé détaillé de la théorie est donné dans l'article cité, mais aussi dans les travaux de différents auteurs [27, 43, 58]. Divol et Lacombe s'appuient sur les travaux de Figalli et Gigli [44] et étendent les distances de Wasserstein aux mesures de Radon supportées sur des sous-ensembles ouverts propres $\mathcal{X} \subset \mathbb{R}^n$, dont le bord est noté $\partial\mathcal{X}$ (et $\overline{\mathcal{X}} := \mathcal{X} \sqcup \partial\mathcal{X}$). Pour le reste de cette section, nous noterons toujours $\mathcal{X}$ un tel ensemble.

Pour définir une notion de transport, l'idée générale est de voir $\partial\mathcal{X}$ comme un réservoir de masse infinie, capable d'absorber toute disparité de masse entre les mesures. Ainsi, si deux mesures de Radon μ et ν ont une masse différente, nous pouvons toujours définir un plan de transport d'une mesure à l'autre en envoyant le surplus de masse sur $\partial\mathcal{X}$. Symboliquement,

Définition 2.1. [44, Problème 1.1] Soit $p \in [1, +\infty)$ et soient μ, ν deux mesures de Radon supportées sur $\mathcal{X}$ telles que

$$\int_{\mathcal{X}} d(x, \partial\mathcal{X})^p \, d\mu(x) < +\infty, \quad \int_{\mathcal{X}} d(x, \partial\mathcal{X})^p \, d\nu(x) < +\infty \,. \tag{1.8}$$

L'ensemble des **plans de transport admissibles** $\Gamma(\mu, \nu)$ est défini comme l'ensemble des mesures de Radon π sur $\overline{\mathcal{X}} \times \overline{\mathcal{X}}$ satisfaisant

$$\pi(A \times \overline{\mathcal{X}}) = \mu(A) \quad \text{et} \quad \pi(\overline{\mathcal{X}} \times B) = \nu(B) \tag{1.9}$$

pour tous les ensembles de Borel $A, B \subset \mathcal{X}$. Le **coût d'un plan de transport** $\pi \in \Gamma(\mu, \nu)$ est

$$C_p(\pi) := \int_{\overline{\mathcal{X}} \times \overline{\mathcal{X}}} d(x, y)^p \, d\pi(x, y) \tag{1.10}$$

et la **distance de transport optimale (dite de Wasserstein)** $d_p(\mu, \nu)$ est

$$d_p(\mu, \nu) := \left(\inf_{\pi \in \Gamma(\mu, \nu)} C_p(\pi) \right)^{1/p} . \tag{1.11}$$

Les plans $\pi \in \Gamma(\mu, \nu)$ réalisant l'infimum dans l'équation ci-dessus sont dits **optimaux**.

Remarque 2.2. Le coût C_∞ d'un plan de transport $\pi \in \Gamma(\mu, \nu)$ est défini par

$$C_\infty(\pi) := \|d\|_{L^\infty(\pi)} \,, \tag{1.12}$$

d'où une notion de distance d_∞ induite par ce coût de transport.

Définition 2.3. L'espace des mesures de Radon sur $\mathcal{X}$ sera noté $\mathcal{D}(\mathcal{X})$ (ou simplement $\mathcal{D}$ si $\mathcal{X}$

est sous-entendu). Nous introduisons également les espaces suivants

$$\mathcal{D}_p := \left\{ \mu \in \mathcal{D} \,\middle|\, \int_{\mathcal{X}} d^p(x, \partial\mathcal{X}) \, d\mu(x) < \infty \right\} \tag{1.13}$$

et définissons $\mathcal{D}_\infty$ comme l'espace des mesures de Radon à support compact.

Pour en revenir à la théorie de la persistance, rappelons qu'il est possible de voir les diagrammes de persistance comme des mesures sur

$$\mathcal{X} := \{(x, y) \in \mathbb{R}^2 \,|\, y > x\} \,. \tag{1.14}$$

Dorénavant, $\mathcal{X}$ fera toujours référence à ce demi-espace. Considérés comme des mesures, les diagrammes de persistance ne sont rien d'autre que des sommes de mesures de Dirac. En clôturant cet espace par rapport à la topologie de la convergence vague, nous retrouvons l'ensemble des mesures de Radon sur $\mathcal{X}$.

Définition 2.4. L'ensemble des **mesures persistantes** $\mathcal{D}$ est l'ensemble des mesures de Radon (de masse potentiellement infinie) sur $\mathcal{X} := \{(x, y) \in \mathbb{R}^2 \,|\, y > x\}$.

En munissant $\mathcal{X}$ de la distance ℓ^∞ sur $\mathbb{R}^2$ définie par

$$d((p, q), (r, s)) = \max\{|p - r| \,, |q - s|\} \,, \tag{1.15}$$

les distances de transport partiel optimal d_p entre les mesures persistantes ont un sens. Les répercussions de cette approche ont été explorées par Divol et Lacombe dans [33].

Remarque 2.5. En particulier, d_∞ coïncide exactement avec la distance bottleneck usuelle sur les diagrammes (*cf.* les livres de Chazal *et al.* et d'Oudot pour une définition de la distance *bottleneck* [24, 76]).

Les notions de distance ainsi décrites correspondent à des notions de distance sur les diagrammes de persistance. *A priori*, vu que nous nous intéresserons à la topologie des ensembles de sur-niveau, il faut s'assurer que ces notions de distance ont une signification au niveau homologique. Pour d_∞ (la distance bottleneck), ceci est établi par le théorème d'isométrie, qui garantit que la distance bottleneck mesure à quel point deux diagrammes sont loin d'être isomorphes dans la catégorie des modules de persistance. Nous renvoyons le lecteur vers le livre de Chazal *et al.* [24, Chapitre 5] pour une explication plus détaillée. La nature algébrique des distances Wasserstein d_p sur l'espace des diagrammes a été discutée par Skraba et Turner dans [93, §8].

L'extension de l'espace des diagrammes de persistance à l'espace des mesures persistantes présente trois avantages principaux. Premièrement, comme montré dans [33], il est possible d'utiliser la machinerie du transport optimal pour aborder les problèmes de la théorie de la persistance. Deuxièmement, $\mathcal{D}$ est un espace linéaire, ce qui rend possible et facile la prise de moyennes et de combinaisons de diagrammes. Enfin, l'espace ainsi étendu est bien adapté au cadre stochastique en raison de la linéarité des mesures et du théorème de Tonelli : deux propriétés clés que nous exploiterons à plusieurs reprises.

2.2 Théorèmes de stabilité

Une des propriétés les plus importantes de d_∞, en plus du théorème d'isométrie, est le théorème de stabilité, qui garantit que deux filtrations provenant de fonctions L^∞-proches ont des diagrammes qui sont d_∞ proches, et que cette continuité est en fait Lipschitzienne. Plus précisement,

Théorème 2.1 (Stabilité bottleneck, Corollaire 3.6 [76])**.** *Soient $f, g : X \to \mathbb{R}$ deux fonctions continues, alors*

$$d_\infty(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq \|f - g\|_\infty \tag{1.16}$$

où $\mathrm{Dgm}(f)$ *et* $\mathrm{Dgm}(g)$ *désignent les diagrammes de f et g respectivement.*

Ce théorème établit que les code-barres sont des invariants robustes par perturbations L^∞. Par contre, d_∞ néglige ce qui se passe au niveau des petites barres, qui sont traditionnellement perçues comme du « bruit topologique ». Or, ces petites barres peuvent contenir de l'information non-triviale et importante comme la texture, d'où l'avantage des distances d_p avec $p < \infty$ (qui tiennent bien compte de *toutes* les barres). Il est classique dans ce cadre d'introduire la fonctionnelle suivante, qui quantifie une partie de l'information relative aux petites barres.

Définition 2.6. Soit X un espace topologique compact, connexe, localement connexe par arcs et soit $f : X \to \mathbb{R}$ une fonction continue. La k-ième **fonctionnelle** Pers_p **de** f est

$$\mathrm{Pers}_p(H_k(X, f)) := \left(\sum_{b \in H_k(X,f)} \ell(b \cap [\inf(f), \sup(f)])^p \right)^{1/p}, \tag{1.17}$$

où $\ell(b)$ désigne la longueur de la barre b et $H_k(X, f)$ désigne le code-barres (ou diagramme) du H_k issu de la filtration par les sur-niveaux de f. Nous noterons souvent $\mathrm{Pers}_p(f) := \mathrm{Pers}_p(H_0(X, f))$.

De plus, un résultat de stabilité pour d_p ($p < \infty$) est souhaitable pour s'assurer que la théorie demeure robuste dans le cadre Wasserstein, comme c'est le cas pour d_∞. Il est cependant impossible d'espérer un résultat aussi général que le théorème de stabilité classique, puisque les boules L^∞ arbitrairement petites autour d'une fonction f contiennent toujours des fonctions dont les diagrammes sont à distance d_p infinie de $\mathrm{Dgm}(f)$. Néanmoins, pourvu que nous fixions la régularité des fonctions et des espaces X considérés, un tel résultat est envisageable. Pour le démontrer, les résultats classiques s'appuient sur la condition suivante sur l'espace métrique X.

Définition 2.7. [28] Un espace métrique (triangulable) X **a q-persistance totale bornée** si, pour tout $k \in \mathbb{N}$, il existe une constante C_X qui ne dépend que de X telle que

$$\mathrm{Pers}_q^q(\mathrm{Dgm}_k(f)) < C_X \tag{1.18}$$

pour toute fonction 1-Lipschitzienne f.

Sous réserve de cette condition, il est possible de démontrer un théorème de stabilité Wasserstein pour les fonctions Lipschitz.

Théorème 2.2 (Cohen-Steiner, Edelsbrunner, Harer, Mileyko [28])**.** *Soit X un espace triangulable impliquant une q-persistance totale bornée et soient f et g deux fonctions Lipschitziennes sur X à valeurs dans $\mathbb{R}$. Alors, pour tout $p > q$,*

$$d_p(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq C_X(\mathrm{Lip}(f)^q \vee \mathrm{Lip}(g)^q) \, \|f - g\|_\infty^{1-\frac{q}{p}} \,, \tag{1.19}$$

où $\mathrm{Lip}(f)$ *désigne la constante de Lipschitz de f et où $a \vee b := \max\{a, b\}$.*

Ce théorème est conceptuellement important, mais quantitativement défaillant en raison de la condition *ad hoc* sur la q-persistance bornée. Plus précisement, la borne du théorème dépend d'une constante C_X associée à l'espace métrique qui pourrait donc en principe être arbitrairement grande. De plus, des conditions sur q en fonction de l'espace métrique X sont désirables afin de s'assurer du régime de validité du théorème. Dans certains cas particuliers, une borne sur q peut être trouvée.

Lemme 2.8 (Skraba, Turner, [93])**.** *Soit X une variété riemannienne compacte de dimension d. Si X a une q-persistance totale bornée alors $q \geq d$.*

Nous verrons plus tard que cette notion de q-persistance bornée est intimement liée à une notion de *dimension homologique*. Une partie de notre contribution consiste à généraliser le théorème de stabilité Wasserstein et à le rendre quantitatif. En particulier, nous donnerons des bornes explicites sur q (que nous démontrerons être optimales dans certains cas) et sur C_X, qui dépendent uniquement de quantités explicites liées à une notion de dimension topologique sur X.

2.3 Notions de dimensions homologiques

Plusieurs notions de dimension définies en termes des diagrammes de persistance existent dans la litérature : dans la thèse de Vanessa Robins [88], dans un article de Schweinhart et MacPherson [66], de Adams *et al.* [2] et plus récemment de Schweinhart [91] (dans lequel l'auteur modifie la définition donnée précédemment dans l'article écrit en collaboration avec MacPherson). Parmi ces définitions, celle qui nous intéressera le plus est celle donnée par Schweinhart dans [91], sur laquelle plusieurs résultats sont établis par le même auteur et ses collaborateurs dans [52, 91, 92].

Définition 2.9 (Définition de Schweinhart de $\dim_{\mathrm{PH}}^k$, [91])**.** Soit X un sous-ensemble borné d'un espace métrique. La **k-ième dimension homologique de X au sens de Schweinhart** est

$$\dim_{\mathrm{PH}}^k(X) := \inf_p \{p \mid \sup_{\mathbf{x}} \mathrm{Pers}_p(\mathrm{Dgm}_k(d(-, \mathbf{x}))) < \infty\} \,, \tag{1.20}$$

où le supremum est pris sur tous les ensembles finis de points $\mathbf{x} \subset X$.

Dans le cadre déterministe, Schweinhart établit plusieurs résultats concernant des sous-ensembles de $\mathbb{R}^d$, notamment concernant les dimensions homologiques de degré 1.

Théorème 2.3 (Schweinhart, [91])**.** *Soit X un sous-ensemble borné de $\mathbb{R}^2$. Si la dimension upper-box $\overline{\dim}(X)$ est strictement supérieure à 1.5, alors*

$$\dim^1_{\mathrm{PH}}(X) = \overline{\dim}(X)\,. \tag{1.21}$$

Si $X \subset \mathbb{R}^d$ est borné et $\overline{\dim}(X) > d - \frac{1}{2}$, alors

$$\overline{\dim}(X) \leq \dim^1_{\mathrm{PH}}(X) \leq d\,. \tag{1.22}$$

De manière générale, l'auteur tente de rapprocher sa notion de dimension homologique et la dimension upper-box de X. Rappelons brièvement la définition de cette notion.

Définition 2.10. Soit (X, d) un espace métrique. Sa dimension upper-box, $\overline{\dim}(X)$ (resp. lower-box, $\underline{\dim}(X)$) est la limite

$$\overline{\dim}(X) := \limsup_{\varepsilon \to 0} \frac{\log(\mathcal{N}_X(\varepsilon))}{\log(1/\varepsilon)}\,, \qquad \text{resp. } \underline{\dim}(X) := \liminf_{\varepsilon \to 0} \frac{\log(\mathcal{N}_X(\varepsilon))}{\log(1/\varepsilon)}\,, \tag{1.23}$$

où $\mathcal{N}_X(\varepsilon)$ est le nombre minimal de boules de rayon ε nécessaires pour couvrir X.

Ce faisant, Schweinhart souligne une question intéressante à laquelle nous donnerons un début de réponse.

Question 2.11 (Question 5, [91])**.** Y a-t-il des hypothèses sur X sous lesquelles nous ayons $\dim^k_{\mathrm{PH}}(X) = \overline{\dim}(X)$?

Une première réponse en degré d'homologie zéro avait déjà été établie par Kozma, Lotker et Stupp [59], qui ont démontré que la dimension homologique ainsi définie coïncide avec $\overline{\dim}(X)$ pour n'importe quel espace métrique (les auteurs démontrent en fait cette notion d'égalité pour une dimension définie en termes de l'arbre couvrant minimal, mais cette dernière coïncide elle-même avec $\dim^0_{\mathrm{PH}}$ au sens de Schweinhart). Plus précisement,

Théorème 2.4 (Kozma, Lotker, Stupp, [59])**.** *Pour un espace métrique quelconque X,*

$$\dim^0_{\mathrm{PH}}(X) = \overline{\dim}(X)\,. \tag{1.24}$$

Nous reviendrons plus tard sur ce résultat sur la dimension et l'indice homologique en degré 0 dans le cadre de nos propres résultats.

Enfin, Schweinhart démontre quelques résultats sur la dimension homologique, mais où l'homologie des nuages de points dans X est calculée via leur complexe de Rips. Nous noterons cette notion de dimension $\dim^k_{\widetilde{\mathrm{PH}}}(X)$. Bien que ces résultats ne soient *a priori* valables que pour l'homologie de Rips des nuages de points, nous verrons plus tard que nous rencontrerons des conditions similaires dans nos propres résultats.

Théorème 2.5 (Schweinhart, [91])**.** *Il existe un espace métrique X tel que*

$$\dim^1_{\widetilde{\mathrm{PH}}}(X) = 2 > 1 = \overline{\dim}(X)\,. \tag{1.25}$$

De plus, si $X \subset \mathbb{R}^d$ est borné, alors

$$\dim^1_{\widetilde{\mathrm{PH}}}(X) \leq \overline{\dim}(X) \,. \tag{1.26}$$

2.4 Résultats sur les arbres

Comme remarqué précédemment, la communauté probabiliste s'est aussi beaucoup intéressée à la question de l'étude du H_0 des ensembles de sur-niveau de fonctions $f : X \to \mathbb{R}$ dans le cas particulier $X = [0, 1]$. Le Gall [37] introduit un arbre encodant l'information des composantes connexes de sur-niveau avec une approche analytique que nous détaillerons. Ces arbres ont ensuite été utilisés par Le Gall et d'autres [37, 38] pour étudier des propriétés fines des processus de Lévy et du mouvement brownien.

Définition 2.12. Soit X un espace topologique compact, connexe et localement connexe par arcs et soit $f : X \to \mathbb{R}$ une fonction continue. La **H_0-distance**, d_f, est la pseudo-distance

$$d_f(x, y) := f(x) + f(y) - 2 \sup_{\gamma : x \mapsto y} \inf_{t \in [0,1]} f(\gamma(t)) \,, \tag{1.27}$$

où le supremum porte sur tous les chemins γ reliant x et y.

Proposition 2.13. *L'espace métrique*

$$(T_f, d_f) := (X/\{d_f = 0\}, d_f) \,, \tag{1.28}$$

où $X/\{d_f = 0\}$ est le quotient de X où nous identifions tous les points x et y satisfaisant $d_f(x, y) = 0$ est un arbre réel. Par abus de notation, nous notons d_f la distance induite sur T_f par la pseudo-distance d_f sur X.

Remarque 2.14. La définition originellement donnée par Le Gall concerne les fonctions $f : [0, 1] \to \mathbb{R}$ et ne contient donc pas le sup sur tous les chemins. La formule pour cette notion de pseudo-distance a aussi été considérée par Curien, Le Gall et Miermont [29]. Notre contribution consiste donc simplement à avoir assoupli les hypothèses sur X produisant un arbre en quotientant X par le lieu des zéros de d_f.

Picard [80] reprend une procédure similaire afin de construire un arbre associé à une fonction $f : [0, 1] \to \mathbb{R}$ maintenant simplement supposée continue à droite ayant des limites à gauche (càdlàg), mais nous ne détaillerons pas sa construction explicitement renvoyant le lecteur vers l'article cité (des constructions similaires sont aussi faites par Le Gall et Duquesne dans [38]). La figure 1.2 aide néanmoins à en comprendre l'essence.

Les branches de T_f correspondent exactement aux composantes connexes des ensembles de sur-niveaux de f. Il est pratique d'introduire des *ε-taillages* ou *ε-simplications* de T_f comme suit. Considérons la fonction $h : T_f \to \mathbb{R}$ qui à un point $\tau \in T_f$ associe la distance de τ à la plus haute feuille au-dessus de τ par rapport à la filtration sur T_f induite par f, alors

Définition 2.15. Soit $\varepsilon \geq 0$. Un **ε-taillage** ou une **ε-simplification** de T_f est le sous-espace

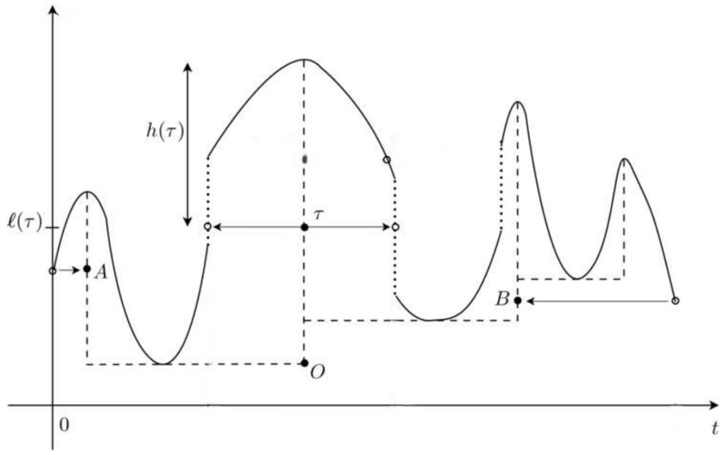


Figure 1.2: Une représentation de la construction d'un arbre associé à une fonction càdlàg. La figure est extraite de [80]

.

métrique de T_f défini par

$$T_f^\varepsilon := \{\tau \in T_f \,|\, h(\tau) \geq \varepsilon\} \tag{1.29}$$

Remarque 2.16. L'ε-taillage de T_f est obtenu en taillant l'arbre en partant du haut de ses branches d'une longueur ε. Nous noterons N^ε le nombre de feuilles de T_f^ε.

En suivant encore le courant de pensée probabiliste, il est possible en dimension un d'utiliser l'ordre total de $\mathbb{R}$ et de compter N^ε en comptant le nombre de fois où nous nous éloignons d'au moins ε d'un minimum local et d'au moins ε d'un maximum local. Cette idée peut être formalisée par la suite de temps d'arrêts suivante introduite par Neveu *et al.* [74].

Définition 2.17. En fixant $S_0^\varepsilon = T_0^\varepsilon = 0$, nous définissons une suite de temps d'arrêt de manière récursive

$$T_{i+1}^\varepsilon := \inf\left\{ t \geq S_i^\varepsilon \,\middle|\, \sup_{[S_i^\varepsilon, t]} f - f(t) > \varepsilon \right\}$$

$$S_{i+1}^\varepsilon := \inf\left\{ t \geq T_{i+1}^\varepsilon \,\middle|\, f(t) - \inf_{[T_{i+1}^\varepsilon, t]} f > \varepsilon \right\}$$

Compter le nombre de barres de longueur $\geq \varepsilon$ revient donc à compter le nombre d'oscillations de taille au moins ε. Plus précisément,

$$N^\varepsilon = \inf\{i \,|\, T_i^\varepsilon \text{ ou } S_i^\varepsilon = \inf \emptyset\}\,, \tag{1.30}$$

i.e. le plus petit i tel que l'ensemble sur lequel porte l'infimum qui définit T_i^ε et S_i^ε est vide. Cette formalisation du comptage de N^ε en dimension un nous sera utile plus tard, lorsque nous chercherons à calculer explicitement N^ε et certaines autres quantités relatives au code-barres de divers processus.

Ce point de vue laisse entrevoir le fait que la décroissance de N^ε devrait fortement dépendre de la régularité du processus. En un sens précis, la régularité détermine presque entièrement

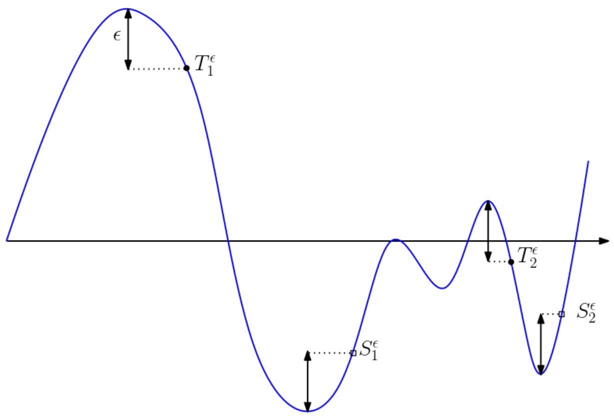


Figure 1.3: Une fonction f avec les temps T_i^ε et S_i^ε indiqués. Cette fonction a exactement 3 barres de longueur $\geq \varepsilon$ et pas seulement 2 en raison des effets de bord de l'intervalle.

l'asymptotique de N^ε lorsque $\varepsilon \to 0$, comme démontré par Picard (et interprété dans le cadre des code-barres dans [77]).

Théorème 2.6 (Picard, §3[80] et [77]). *Soit $f : [0,1] \to \mathbb{R}$ une fonction continue, alors*

$$\mathcal{V}(f) = \mathcal{L}_0(f) = \overline{\dim}\, T_f = \limsup_{\varepsilon \to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 \tag{1.31}$$

où $a \vee b = \max\{a, b\}$,

$$\mathcal{V}(f) := \inf\{p \,|\, \|f\|_{p-var} < \infty\} \quad et \quad \mathcal{L}_0(f) := \inf\{p \geq 1 \,|\, \mathrm{Pers}_p(f) < \infty\}\,. \tag{1.32}$$

Nous démontrerons que ce théorème reste vrai pour des fonctions sur des espaces métriques beaucoup plus généraux, au prix de ne plus avoir de lien aussi précis avec une notion de régularité analogue. Nous verrons aussi que pour une certaine classe d'espace métriques, il reste possible génériquement (au sens de Baire) de lier régularité et $\mathcal{L}_0(f)$.

2.5 Résultats probabilistes

Dans cette thèse, nous nous appuyerons sur de nombreux résultats probabilistes couvrant plusieurs branches différentes de la théorie des probabilités. D'abord, nous examinerons certains résultats proches des considérations que nous avons eues lorsque nous avons introduit la notion de dimension homologique de Schweinhart. Nous montrerons en outre comment se comportent les diagrammes de nuages de points aléatoires tirés selon certains types de lois. Ensuite, nous rappelerons une notion de « diagramme de persistance moyen » que nous pouvons définir dans un cadre probabiliste qui est dans un certain sens compatible avec les fonctionnelles linéaires des diagrammes de persistance. Nous discuterons au passage quelques résultats annexes concernant ces objets. Enfin, nous nous attarderons sur quelques résultats établis dans la littérature concernant les diagrammes de processus stochastiques.

Autour de la dimension homologique

Dans la suite de cette thèse, nous verrons qu'il peut être intéressant de quantifier la persistance de la fonction distance à un nuage de points et plus précisement de quantifier la divergence de la fonctionnelle Pers_p^p à mesure que nous augmentons le nombre de points dans le nuage. Le cadre probabiliste permet d'établir plusieurs résultats, dont les suivants.

Théorème 2.7 (Schweinhart, [92]). *Soit μ une mesure de probabilité d-Ahlfors régulière sur un espace métrique X et soit $\mathbf{X}_n := (X_1, \cdots, X_n)$ un vecteur d'échantillons i.i.d. de μ. Si $0 < p < d$,*

$$\mathrm{Pers}_p^p(H_0(X, d(-, \mathbf{X}_n)) \asymp n^{1-\frac{p}{d}} \tag{1.33}$$

avec une forte probabilité lorsque $n \to \infty$, où le symbole $\asymp$ indique que le rapport entre les deux quantités est borné entre des constantes positives qui ne dépendent pas de n.

Ce théorème de Schweinhart fait écho (mais est valable dans un cadre beaucoup plus général) à un théorème originellement établi par Steele en 1988 concernant les mesures de probabilité à support compact dans $\mathbb{R}^d$.

Théorème 2.8 (Steele, [95]). *Soit μ une mesure de probabilité à support compact dans $\mathbb{R}^d$ avec $d \geq 2$ et soit $\mathbf{X}_n := (X_1, \cdots, X_n)$ des échantillons i.i.d. de μ. Si $0 < p < d$, alors presque sûrement,*

$$\lim_{n \to \infty} n^{-1+\frac{p}{d}} \, \mathrm{Pers}_p^p(H_0(X, d(-, \mathbf{X}_n))) = c(p, d) \int_{\mathbb{R}^m} \rho(x)^{1-\frac{p}{d}} \, dx \tag{1.34}$$

où $\rho(x)$ est la densité de probabilité de la partie absolument continue de μ et $c(p, d)$ est une constante positive qui ne depend que de p et d.

Divol et Polonik ont récemment établi une loi forte des grands nombres similaire au résultat de Steele ci-dessus concernant les degrés d'homologie supérieurs, dans le cas de mesures de probabilité absolument continues et bornées sur le cube $[0, 1]^d$.

Théorème 2.9 (Divol et Polonik, [35]). *Soit μ une mesure de probabilité bornée sur $[0, 1]^d$ et soit $\mathbf{X}_n := (X_1, \cdots, X_n)$ un vecteur d'échantillons i.i.d. de μ, alors pour $0 < p < d$ et $0 \leq k < d$, presque sûrement,*

$$\lim_{n \to \infty} n^{-1+\frac{p}{d}} \, \mathrm{Pers}_p^p(H_k([0, 1]^d, d(-, \mathbf{X}_n))) \to \mathrm{Pers}_p^p(\nu_p^\mu)\,. \tag{1.35}$$

pour une certaine mesure de Radon non-dégénérée dépendant de p et de la measure de probabilité μ, ν_p^μ sur $\mathcal{X}$.

Nous revisiterons et mettrons en contexte le résultat de Schweinhart dans le cadre d'un théorème de généricité que nous démontrerons plus tard. Quant à lui, le résultat de Divol et de Polonik nous permettra de partiellement répondre à une question de Schweinhart sur la dimension homologique. Dans une approche à la Erdős, l'idée sera que cette construction probabiliste démontre l'existence de fonctions saturant certaines bornes.

Diagrammes de persistance moyens

Le diagramme moyen d'un processus stochastique f peut être défini par dualité pourvu que nous voyons les diagrammes de persistance comme mesures.

Définition 2.18 (Chazal, Divol, [20]). Pour tout $B \subset \mathcal{X}$ mesurable, nous notons

$$\mathbb{E}[\mathrm{Dgm}(f)]\,(B) := \mathbb{E}[\mathrm{Dgm}(f)(B)] \ . \tag{1.36}$$

Cette approche contraste celle de la définition de « moyenne » par le biais de moyennes de Fréchet. Au prix de devoir élargir l'espace des diagrammes à l'espace de toutes les mesures de Radon sur $\mathcal{X}$ (dont nous avons déjà discuté), cette approche présente plusieurs avantages. Entre autres, l'existence et l'unicité de cette moyenne sont évidentes. De plus, elle est plus facilement calculable dans un cadre théorique comme pratique. Enfin, cette définition est compatible avec les fonctionnelles linéaires des diagrammes [20]. Plus précisément, si $\mathcal{B}$ est un espace de Banach, $\psi : \mathcal{X} \to \mathcal{B}$ est une fonction continue et μ est une mesure de Radon sur $\mathcal{X}$, toute fonctionnelle $\Psi : \mathcal{D} \to \mathcal{B}$ de la forme

$$\Psi : \mu \mapsto \int \psi \, d\mu \tag{1.37}$$

est compatible avec cette notion. C'est-à-dire que la moyenne de la fonctionnelle $\mathbb{E}[\Psi]$ correspond à l'intégrale de ψ sur le diagramme moyen ainsi défini.

La continuité de ces fonctionnelles a été explorée par Divol et Lacombe dans [33].

Proposition 2.19 (Divol, Lacombe, Proposition 5.1 [33]). *En gardant les mêmes notations, une fonctionnelle $\Psi : \mathcal{D}_p \to \mathcal{B}$ est continue par rapport à d_p si et seulement si*

$$\psi \in \left\{ \varphi \in C^0(\mathcal{X}, \mathcal{B}) \;\middle|\; x \mapsto \frac{\varphi(x)}{d(x, \partial\mathcal{X})^p} \text{ est bornée} \right\} . \tag{1.38}$$

De cette caractérisation de la continuité de ces fonctionnelles, nous déduisons qu'un contrôle sur les distances entre les diagrammes moyens de deux processus stochastiques différents est souhaitable, car il permet en principe de déduire l'existence d'une borne pour toute fonctionnelle Ψ ainsi définie. Nous prouverons un tel contrôle en termes de distances Wasserstein sur les distributions des processus stochastiques sous-jacents.

Des résultats sur les diagrammes moyens pour certaines classes de processus ont été obtenus par Chazal et Divol dans [20]. Les auteurs traitent dans le même article du cas des nuages de points et de l'existence et la régularité de la densité par rapport à la mesure de Lebesgue de la mesure de Radon $\mathbb{E}\Big[\mathrm{Dgm}(H_k(\mathbb{R}^d, d(-, \mathbf{X}_n)))\Big]$ pour un certain nuage de points aléatoire $\mathbf{X}_n$. Ce qui nous intéressera ici c'est plutôt quelques résultats obtenus pour le mouvement brownien.

Théorème 2.10 (Chazal, Divol, [20]). *Soit B un mouvement brownien standard sur $[0, t]$. Le diagramme moyen $\mathbb{E}[\mathrm{Dgm}(B)]$ de l'homologie persistante en degré 0 des ensembles de sous-niveau est bien défini et admet une densité par rapport à la mesure de Lebesgue.*

Baryshnikov réussit même à calculer explicitement la densité du diagramme moyen associé au mouvement brownien avec drift sur le rayon $[0, \infty[$ sur un sous-ensemble de $\mathcal{X}$.

Proposition 2.20 (Baryshnikov, Propositions 4.1 et 4.3 [9])**.** *Soit B^μ un mouvement brownien standard avec drift $\mu > 0$ sur le rayon $[0, \infty[$. Si $x > 0$, le nombre $N^{x,x+\varepsilon}$ de points dans le diagramme de persistance dans le rectangle $]-\infty, x] \times [x+\varepsilon, \infty[$ est distribué comme une variable aléatoire géométrique de paramètre $1 - e^{2\mu\varepsilon}$. De plus, la densité de $\mathbb{E}[\mathrm{Dgm}(B^\mu)]$ par rapport à la mesure de Lebesgue (sur $\mathcal{X} \cap \{x > 0\}$) est donnée par*

$$\frac{4\mu^2 e^{2\mu\varepsilon}(1 + e^{2\mu\varepsilon})}{(e^{2\mu\varepsilon} - 1)^3}\,. \tag{1.39}$$

et diverge comme $\frac{1}{\mu\varepsilon^3}$ lorsque $\varepsilon \to 0$.

La nature de cette divergence n'est *a posteriori* pas surprenante, car elle provient de l'auto-similarité du mouvement brownien. Dans cette thèse, nous donnerons une expression explicite pour la densité du diagramme moyen du mouvement brownien sur $[0, t]$. Nous nous intéresserons aussi à ces divergences qui, comme le montre le théorème de Picard déjà mentionné, ont un lien avec la régularité et, pour les processus autosimilaires, avec leur auto-similarité.

Suivant cette logique, Picard a obtenu des résultats relatifs à la divergence de N^ε lorsque $\varepsilon \to 0$ pour les processus de Lévy (et même pour le mouvement brownien fractionnaire, *cf.* [80]).

Proposition 2.21 (Picard, §3 [80])**.** *Soit f un processus de Lévy et supposons que, presque sûrement, f soit nulle part monotone. Notons*

$$\xi(\varepsilon) := \mathbb{E}[S^\varepsilon + T^\varepsilon] \tag{1.40}$$

avec

$$S^\varepsilon := \inf\{t \mid f(t) - \inf_{[0,t]} f > \varepsilon\} \quad \textit{et} \quad T^\varepsilon := \inf\{t \mid \sup_{[0,t]} f - f(t) > \varepsilon\}\,, \tag{1.41}$$

alors, en probabilité, $\xi(\varepsilon)N^\varepsilon \to 1$ lorsque $\varepsilon \to 0$. Si $\xi(\varepsilon) = O(\varepsilon^\alpha)$ pour un certain α, cette convergence est en fait presque sûre.

Dans la suite, nous donnerons une asymptotique encore plus fine pour les processus de Lévy.

2.6 Transformées de Mellin et développements asymptotiques

La fonctionnelle $\mathrm{Pers}_p(H_k(X, f))$ qui nous a occupés jusqu'à présent est duale (vue comme fonction de p) à N_k^ε en un sens précis et contient strictement la même information que cette dernière.

Proposition 2.22 (P, Section 2.5 [79])**.** *Nous avons l'égalité suivante*

$$\mathrm{Pers}_p^p(H_k(X, f)) = p \int_0^\infty \varepsilon^{p-1} N_k^\varepsilon \, d\varepsilon\,. \tag{1.42}$$

De plus, si nous considérons $\mathrm{Pers}_p^p(H_k(X, f))$ comme fonction complexe de p, pour un certain c

$$N^\varepsilon = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \mathrm{Pers}_p^p(H_k(X, f)) \varepsilon^{-p} \, \frac{dp}{p}\,. \tag{1.43}$$

Cette proposition justifie un détour dans le monde de l'analyse complexe et des développements asymptotiques. Pers_p^p n'est en fait rien d'autre qu'une transformée de Mellin (à un facteur p près) de N^ε, ce qui entraîne des conséquences intéressantes.

Définition 2.23. Soit f une fonction localement intégrable sur le rayon $]0,\infty[$. La **transformée de Mellin** de f est définie par

$$\mathcal{M}[f(x)](s) := \int_0^\infty x^{s-1} f(x)\, dx\,. \tag{1.44}$$

Remarque 2.24. La transformée de Mellin reflète la dualité de Pontryagin du groupe abélien localement compact $(\mathbb{R}_+, \times)$. En effet, $d\log(x) = \frac{dx}{x}$ est la mesure de Haar de $(\mathbb{R}_+, \times)$. La théorie de cette transformée est donc complètement analogue à celle de la transformée de Laplace bilatérale, car l'application $\log : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$ est un isomorphisme de groupes abéliens.

Notation 2.25. Nous utiliserons parfois la notation $\mathcal{M}[f](s) = f^*(s)$.

Définition 2.26. Le **domaine fondamental** ou **bande fondamentale** de f, $\langle \alpha, \beta \rangle$ est l'ensemble maximal

$$\langle \alpha, \beta \rangle := \{z \in \mathbb{C} \,|\, \alpha < \mathrm{Re}(z) < \beta\} \tag{1.45}$$

où $f^*(s)$ est définie.

En vertu du théorème d'inversion de Laplace, nous avons le théorème suivant.

Théorème 2.11 (Inversion de Mellin, [36, 75])**.** *Soit $\langle \alpha, \beta \rangle$ le domaine fondamental (que nous supposerons non-vide) de f et soit $c \in]\alpha, \beta[$. Alors*

1. *Si f est intégrable et $f^*(c + it)$ est intégrable alors, pour presque tout x,*

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f^*(s) x^{-s}\, ds\,. \tag{1.46}$$

   Si f est continue, l'égalité est vraie partout.

2. *Si f est localement intégrable et de variation bornée dans un voisinage de x, alors*

$$\frac{f(x^+) + f(x^-)}{2} = \lim_{T\to\infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} f^*(s) x^{-s}\, ds \tag{1.47}$$

Une condition suffisante pour que la transformée de Mellin de f soit bien définie sur $\langle \alpha, \beta \rangle$ est que f vérifie

$$f(x) = O(x^{-\alpha}) \text{ lorsque } x \to 0 \quad \text{et} \quad f(x) = O(x^{-\beta}) \text{ lorsque } x \to \infty\,. \tag{1.48}$$

Le théorème suivant montre que la transformation de Mellin est un bon outil pour étudier les développements asymptotiques.

Théorème 2.12 (Correspondance fondamentale, [45])**.** *Soit $f :]0, \infty[\to \mathbb{C}$ est une fonction continue avec une bande fondamentale non vide $\langle \alpha, \beta \rangle$. Alors,*

- *Supposons que* $f^*(s)$ *admette une extension méromorphe à la bande* $\langle \gamma, \beta \rangle$ *pour* $\gamma < \alpha$, *qu'elle n'y ait qu'un nombre fini de pôles et qu'elle soit analytique sur* $\mathrm{Re}(s) = \gamma$. *Supposons également qu'il existe* $\eta \in]\alpha, \beta[$ *tel que le long d'un ensemble dénombrable de segments horizontaux de partie imaginaire* $|\mathrm{Im}(s)| = T_i$ *(avec* $T_i \to \infty$*), nous ayons, pour* $r > 1$ *et* $s \in \langle \gamma, \eta \rangle$
$$f^*(s) = O(|s|^{-r}) \quad \textit{lorsque } |s| \to \infty\,. \tag{1.49}$$
Indexons les pôles sur $\langle \gamma, \beta \rangle$ *par leur emplacement* ξ *et par leur ordre* k *et notons* $c_{\xi,k}$ *le* k*-ième coefficient dans la série de Laurent de* $f^*(s)$ *dans une couronne autour de* ξ. *Alors, nous avons un développement asymptotique de* f *autour de* 0 *de la forme*
$$f(x) \sim \sum_{(\xi,k)} c_{\xi,k} \frac{(-1)^{k-1}}{(k-1)!} x^{-\xi} \log^k(x) + O(x^{-\gamma}) \quad \textit{lorsque } x \to 0\,. \tag{1.50}$$
- *Réciproquement, si la fonction* f *a un tel développement asymptotique autour de* 0, *alors* $f^*(s)$ *admet une extension méromorphe à la bande* $\langle \gamma, \beta \rangle$.

De plus, un énoncé analogue est vrai pour les développements asymptotiques autour de ∞ *et les extensions méromorphes au-delà de* β.

Pour en revenir à l'homologie persistante, le théorème d'inversion de Mellin établit que N^ε et $\mathrm{Pers}_p^p(f)$ (maintenant vue comme fonction complexe de p) contiennent exactement la même information. La correspondance fondamentale nous donne une information plus précise : le comportement asymptotique de N^ε peut être retrouvé en étudiant les propriétés analytiques de $\mathrm{Pers}_p^p(f)$ et vice-versa. Nous exploiterons cette dualité afin d'en déduire quelques faits sur les code-barres des processus stochastiques, notamment dans le chapitre 4.

3 Résultats déterministes

Ayant fait l'état des lieux concernant les résultats existants, nous sommes maintenant prêts à exposer les résultats déterministes démontrés dans cette thèse. Ceux-ci concernent surtout une classe d'espace métriques particuliers, à savoir, les espaces *LLC*.

Définition 3.1. Un **espace métrique** (X, d) **localement linéairement connexe (LLC)**, est un espace métrique connexe tel que pour tout $r > 0$ et pour tout $z \in X$, pour tout $x, y \in B(z, r)$, il existe un arc reliant x et y tel que le diamètre de cet arc soit contrôlé linéairement en $d(x, y)$.

Notre contribution gravite autour de la notion de dimension homologique et de stabilité Wasserstein, et du lien entre ces deux concepts. Ils sont intimiment liés par le fait intuitif que les fonctions de classe de régularité C^α sont contraintes dans leur variation, et donc dans leur p-persistance, par leur régularité en un sens qui peut être rendu précis.

Accessoirement, nous avons aussi démontré des résultats sur les distances de Wasserstein sur les espaces de mesures de persistance généralisées, au sens de Divol-Lacombe [33]. La correspondance fondamentale provenant de la théorie de la transformée de Mellin montre déjà que la bande fondamentale de $\mathrm{Pers}_p^p(f)$, lorsqu'elle est non-vide, est une *bande* dont les extrémités ne dépendent que du comportement asymptotique de N^ε en 0 et en l'infini. De ce fait, si $\mathrm{Dgm}(f) \in \mathcal{D}_p \cap \mathcal{D}_q$, avec $p < q$, $\mathrm{Dgm}(f)$ est *a fortiori* dans tous les $\mathcal{D}_r$ pour $r \in [p, q]$. Une question naturelle est de savoir si ces propriétés analytiques valables pour les diagrammes d'une fonction peuvent s'étendre d'abord aux mesures de persistance généralisées, et ensuite à un théorème d'interpolation pour les distances Wasserstein entre deux mesures de persistance généralisées. Nous répondons positivement à ces deux questions.

Proposition 3.2 (Interpolation Wasserstein, Chapitre 4 Proposition 2.40, [79])**.** *Soit $1 \leq p < q \leq \infty$ et $\theta \in]0, 1[$. Définissons p_θ par*

$$\frac{1}{p_\theta} = \frac{\theta}{p} + \frac{1-\theta}{q}\,. \tag{1.51}$$

Pour tout $\mu, \nu \in \mathcal{D}_p \cap \mathcal{D}_q$,

$$d_{p_\theta}(\mu, \nu) \leq 2^{1-\theta}\, d_p^\theta(\mu, \nu)\, (\mathrm{Pers}_q(\mu) + \mathrm{Pers}_q(\nu))^{1-\theta}\,. \tag{1.52}$$

En particulier, si $p \leq r \leq q$, alors $\mathcal{D}_p \cap \mathcal{D}_q \subset \mathcal{D}_r$.

3.1 Indices homologiques, dimensions homologiques et stabilité Wasserstein

Dans ce qui suit, il sera commode d'adopter la notation et convention suivantes.

Notation 3.3. Nous noterons N_k^ε le nombre de barres de $H_k(X, f)$ de longueur $\geq \varepsilon$ avec la convention $N^\varepsilon := N_0^\varepsilon$. De même, nous noterons $N_k^{x,x+\varepsilon}$ le nombre de points du diagramme de $H_k(X, f)$ dans le rectangle $]-\infty, x] \times [x+\varepsilon, \infty[$ avec la convention $N^{x,x+\varepsilon} := N_0^{x,x+\varepsilon}$.

Convention 3.4. Nous considérerons toujours que les barres infinies du code-barres sont tuées à la valeur $\inf f$, de sorte à ne pas avoir de barres de longueur infinie.

Pour une fonction f à support compact, $\mathrm{Dgm}(f)$ est toujours à support borné dans $\mathcal{X}$. De ce fait, le dévéloppemement asymptotique de N^ε à l'infini est identiquement nul et donc $\mathrm{Pers}_p^p(H_k(X, f))$ a toujours une bande fondamentale de la forme $\langle c, \infty\rangle$ (on prendra $\langle \infty, \infty\rangle := \emptyset$). Comme nous l'avons vu, en degré 0, l'importance de la valeur de c et son lien avec la régularité de f a déjà été établi en dimension 1 par Picard [80]. Cela motive la définition suivante pour les degrés d'homologie supérieurs.

Définition 3.5 (Indices de persistance)**.** Soit $f : X \to \mathbb{R}$ une fonction continue. Le **k-ième-indice de persistance de** f est

$$\mathcal{L}_k(f) := \inf\{p \geq 1 \mid \mathrm{Pers}_p(H_k(X, f)) < \infty\}\,. \tag{1.53}$$

Nous écrirons parfois $\mathcal{L}(f) := \mathcal{L}_0(f)$. Pour autant que les degrés supérieurs d'homologie s'annulent identiquement à partir d'un certain rang, nous pouvons également définir l'**indice**

de persistance totale de f

$$\mathcal{L}_{Tot}(f) := \inf\{p \geq 1 \mid \sum_k \mathrm{Pers}_p(H_k(X,f)) < \infty\}\,. \tag{1.54}$$

Dans la suite, nous aurons besoin des définitions suivantes

Définition 3.6. Un espace métrique (X, d)

- est **géodésique** si pour toute paire $x, y \in X$ il existe un chemin γ reliant x à y dont la longueur est égale à $d(x, y)$;
- est M**-doublant** s'il existe une constante $M > 0$ telle que pour tout $x \in X$ et $r > 0$ il soit possible de couvrir la boule $B(x, r)$ par au plus M boules de rayon $\frac{r}{2}$;
- est **polonais** si X est complet et séparable;
- admet un ensemble de dimension upper-box homogène s'il existe un sous-ensemble ouvert $U \subset X$ tel que toute boule $B \subset U$ satisfait $\overline{\dim}(B) = \overline{\dim}(X)$.

Une première question naturelle est donc de savoir si en dimension supérieure nous pouvons encore lier $\mathcal{L}_k(f)$ à la régularité de f et si oui, quelle notion de régularité est appropriée. Le théorème suivant répond en partie à cette problématique.

Théorème 3.1 (Chapitre 2 Théorème 3.23 et remarque 3.28, [77]). *Soit X un espace de $\overline{\dim}(X) = d$ géodésique, compact, doublant et connexe dont toutes les boules suffisamment petites sont géodésiquement convexes. Pour $k \in \mathbb{N}$ et $f \in C^{\alpha}(X, \mathbb{R})$, $\mathcal{L}_k(f) \leq \frac{d}{\alpha}$. Sans l'hypothèse doublante, cette borne devient $\mathcal{L}_k(f) \leq \frac{d(k+1)}{\alpha}$.*

Remarque 3.7. Les hypothèses sur la plus petite borne sur $\mathcal{L}_k(f)$ sont vérifiées par toute variété riemannienne compacte. En revanche, les conditions sur X ne sont probablement pas optimales.

En prenant $\alpha = 1$, C^{α} est la classe des fonctions Lipschitz, dont les fonctions distance à un nuage de points font partie. Ce théorème permet donc d'ores et déjà de répondre partiellement à la Question 5 de Schweinhart (question 2.11). Cependant, cette réponse n'est pas complète, car il faudrait s'assurer qu'il existe des fonctions sur X appartenant à cette classe d'espaces métriques telles que l'inégalité $\mathcal{L}_k(f) \leq \frac{d}{\alpha}$ soit saturée, ou saturée à δ près pour tout $\delta > 0$. Nous reviendrons sur la saturation de cette inégalité plus tard dans un cadre plus restreint. Sans l'hypothèse que X soit doublant, une question intéressante est de savoir si la borne trouvée est optimale : la preuve du théorème laisse croire que si de tels espaces métriques existent, ils ne peuvent pas être « à géométrie bornée » et sont relativement pathologiques. Néanmoins, le résultat de Schweinhart concernant l'homologie de Rips (théorème 2.5) laisse espérer que des tels exemples existent.

Pour des raisons qui deviendront évidentes dans la suite, nous modifions la définition de Schweinhart de la k-ième dimension homologique de X de la manière suivante.

Définition 3.8 (k-ième dimension homologique de X)**.** Soit X un sous-ensemble borné d'un espace métrique. La k-ième dimension homologique de X est

$$\dim_{\mathrm{PH}}^k(X) := \sup_{f \in \mathrm{Lip}_1(X)} \mathcal{L}_k(f)\,, \tag{1.55}$$

où $\mathrm{Lip}_1(X)$ désigne l'ensemble des fonctions 1-Lipschitziennes.

Cette définition correspond exactement à la notion de dimension qui convient pour définir le régime de validité de la stabilité Wasserstein, puisqu'elle garantit que pour tout $q > \dim_{\mathrm{PH}}^k(X)$, X implique une q-persistance de degré k bornée.

Remarque 3.9. Notre définition de $\dim_{\mathrm{PH}}^k$ domine toujours celle de Schweinhart, avec quelques cas d'égalité connus.

Un corollaire immédiat du théorème est une borne sur la dimension homologique de degré supérieur pour une certaine classe d'espaces métriques.

Corollaire 3.10. *Soit X un espace géodésique, compact, doublant et connexe dont toutes les boules suffisamment petites sont géodésiquement convexes. Alors*

$$\dim_{\mathrm{PH}}^k(X) \leq \overline{\dim}(X)\,. \tag{1.56}$$

Ce résultat présente une borne analogue à celle trouvée par Schweinhart pour la dimension homologique de Rips des sous-ensembles bornés de $\mathbb{R}^d$. Le contexte n'est pas le même, mais nous avons au moins une condition suffisante pour que la même borne soit valable pour l'homologie de Čech, et ce dans n'importe quel degré.

Du théorème précédemment énoncé découle aussi l'estimation quantitative recherchée de la condition de q-persistance bornée utilisée dans les preuves de stabilité Wasserstein.

Corollaire 3.11 (Chapitre 2 Corollaire 4.16, [77])**.** *Soit X un espace métrique LLC compact de $\overline{\dim}(X) = d$. Pour tout $f \in C^\alpha_\Lambda(X)$ et $q > \frac{d}{\alpha}$,*

$$\mathrm{Pers}_q^q(H_0(X,f)) \leq (2C\Lambda)^q q \int_0^{\frac{diam(X)}{2C}} \varepsilon^{q-1} \mathcal{N}_X(\varepsilon)\, d\varepsilon\,. \tag{1.57}$$

Si l'on suppose de plus que X est géodésique et que les boules suffisamment petites de X sont géodésiquement convexes, alors pour chaque $k \in \mathbb{N}^*$, et tout $p > q > \frac{d(k+1)}{\alpha}$

$$\mathrm{Pers}_q^q(H_k(X,f)) \leq (2C\Lambda)^q q \int_0^{\frac{diam(X)}{2C}} \varepsilon^{q-1} (\mathcal{N}_X(\varepsilon) \vee K_X)^k\, d\varepsilon\,, \tag{1.58}$$

où $K_X = \mathcal{N}_X(\varepsilon^*)$ et ε^* est la valeur à partir de laquelle les boules de rayon $\geq \varepsilon$ de X ne sont plus géodésiquement convexes. Enfin, si l'on suppose en plus que X est M-doublant, alors pour tout $p > q > \frac{d}{\alpha}$,

$$\mathrm{Pers}_q^q(H_k(X,f)) \leq (2C\Lambda)^q q (M^{k+1} - M^k) \int_0^{\frac{diam(X)}{2C}} \varepsilon^{q-1} (\mathcal{N}_X(\varepsilon) \vee K_X)\, d\varepsilon\,. \tag{1.59}$$

Les bornes ainsi obtenues sur la q-persistance de toute fonction $f \in C^{\alpha}_{\Lambda}(X, \mathbb{R})$ ne dépendent que de quantités de nature topologique de X, α et Λ.

Une fois ce résultat établi, le théorème de stabilité Wasserstein peut s'étendre aux fonctions de classe $C^{\alpha}_{\Lambda}(X, \mathbb{R})$ et devient quantitatif relativement aux constantes et à la plage de p admissibles.

Théorème 3.2 (Chapitre 2 Stabilité Wasserstein, Théorème 4.13, [77])**.** *Soit X un espace métrique LLC compact de $\overline{\dim}(X) = d$ et soient $f, g \in C^{\alpha}_{\Lambda}(X, \mathbb{R})$. Alors, pour tout $p > q > \frac{d}{\alpha}$,*

$$d_p^p(H_0(X, f), H_0(X, g)) \leq C_{X,\Lambda,\alpha} \, \|f - g\|_{\infty}^{p-q} \, . \tag{1.60}$$

Si l'on suppose en plus que X est géodésique et est tel que toutes les boules suffisamment petites de X sont géodésiquement convexes, alors pour chaque $k \in \mathbb{N}^*$, et tous les $p > q > \frac{d(k+1)}{\alpha}$

$$d_p^p(H_k(X, f), H_k(X, g)) \leq C_{X,\Lambda,\alpha,k} \, \|f - g\|_{\infty}^{p-q} \, . \tag{1.61}$$

Enfin, si l'on suppose que X est doublant, alors l'inégalité ci-dessus est valable pour tout $p > q > \frac{d}{\alpha}$.

3.2 Arbres et cas particulier du H_0

Arbre associé à une fonction continue

Nous avons d'ores et déjà vu que la question du H_0 est intimement liée aux arbres métriques, et ce, de plusieurs manières. D'abord, par l'approche probabiliste qui à une fonction continue $f : X \to \mathbb{R}$ sur un espace assez régulier associe un arbre T_f qui contient l'information relative aux composantes connexes de surniveau, mais aussi dans un cadre métrique plus abstrait, en s'intéressant aux arbres couvrants minimaux d'un espace métrique quelconque. Nous adopterons ici le point de vue probabiliste, car nous nous intéresserons aux processus stochastiques de dimension 1 dans la suite.

Avec ce degré de généralité, il est possible d'extraire le code-barres de f de l'arbre qui lui est associé T_f par le biais d'un algorithme explicite. Le résultat suivant généralise un théorème précédemment établi par Curry [30, Theoreme 3.10] pour les fonctions de Morse.

Théorème 3.3 (Chapitre 2 Théorème 2.20, [77])**.** *Soit X un espace topologique compact, connexe, localement connexe par arcs et soit $f : X \to \mathbb{R}$ une fonction continue. Alors $\mathrm{Alg}(T_f) = H_0(X, f)$. Où Alg est défini par l'algorithme 1.*

D'après le théorème de Picard (théorème 2.6), il est clair que pour toute fonction f ayant indice de p-variation $\mathcal{V}(f) = d$, nous pouvons associer un arbre T_f de dimension $\overline{\dim}(T_f) = d$. Nous pouvons nous demander si la correspondance $f \mapsto T_f$ est surjective. Notre prochain résultat répond positivement à cette question.

Théorème 3.4 (Problème inverse, Chapitre 2 Théorème 2.23, [77])**.** *Soit T un arbre $\mathbb{R}$ compact tel que $\overline{\dim}\, T < \infty$. Pour tout $\delta > 0$, il existe une fonction continue $f : [0, 1] \to \mathbb{R}$ de $(\overline{\dim}\, T + \delta)$-*

Algorithm 1: L'algorithme calculant $\mathrm{Alg}(T)$

Result: $\mathbb{V}$
$\mathcal{F} \leftarrow T$;
$\mathbb{V} \leftarrow 0$;
$i \leftarrow 0$;
while $\mathcal{F} \neq \varnothing$ **do**
  Trouver γ, le chemin le plus long dans $\mathcal{F}$ en partant de α et finissant en β ;
  if $i = 0$ **then**
    $\mathbb{V} \leftarrow \mathbb{V} \oplus k[\ell(\alpha), \infty[$;
  else
    $\mathbb{V} \leftarrow \mathbb{V} \oplus k[\ell(\alpha), \ell(\beta)[$;
  end
  $\mathcal{F} \leftarrow \overline{\mathcal{F} \setminus \mathrm{Im}(\gamma)}$;
  $i \leftarrow i + 1$;
end
return $\mathbb{V}$

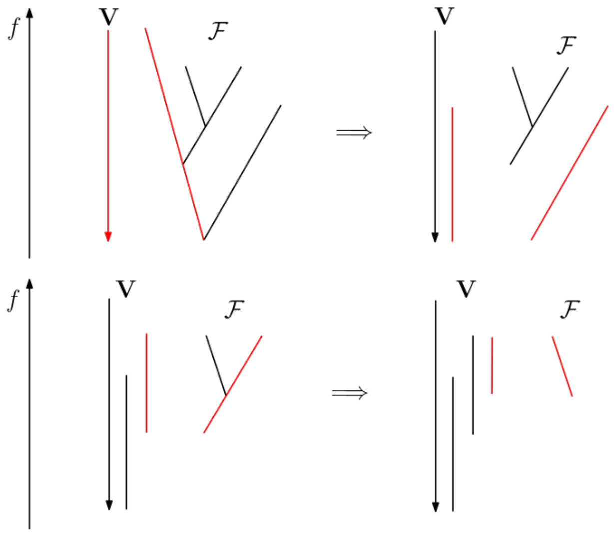


Figure 1.4: Les quatre premières itérations de l'algorithme. À chaque étape, la branche en rouge correspond à la plus longue de la forêt restante. Nous construisons progressivement le module de persistance $\mathbb{V}$ en associant à la k-ième étape un module d'intervalle à l'intervalle associée à la branche en rouge.

variation finie telle que $T = T_f$. De plus, à reparamétrisation près, f peut être prise $\frac{1}{\overline{\dim}\, T+\delta}$-Hölder continue.

Théorèmes de dimension

Étant donné l'utilité du théorème de Picard en dimension 1 (théorème 2.6), il est souhaitable de l'étendre à des espaces topologiques plus généraux.

Théorème 3.5 (Chapitre 2 Théorème 3.9, [77])**.** *Soit X un espace topologique compact, connexe, localement connexe par arcs et soit $f : X \to \mathbb{R}$ une fonction continue. Si $\overline{\dim}\, T_f < \infty$, nous* *avons la chaîne d'égalités suivante*

$$\mathcal{L}_0(f) = \limsup_{\varepsilon \to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 = \overline{\dim}\, T_f \,. \tag{1.62}$$

De plus,

$$\liminf_{\varepsilon \to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 \leq \underline{\dim}\, T_f \,, \tag{1.63}$$

où $\underline{\dim}$ *est la dimension lower-box. Pour* $\underline{\dim}\, T_f > 1$*, ces inégalités deviennent des égalités aussitôt que*

$$\limsup_{\varepsilon \to 0} \frac{N^{2\varepsilon}}{N^\varepsilon} < 1 \,. \tag{1.64}$$

Notons l'absence dans ce théorème d'une notion d'un indice de régularité de f. Contrairement au cas de $[0,1]$, où l'indice de la p-variation $\mathcal{V}(f)$ permettait un lien parfait entre $\mathcal{V}(f)$ et son indice de persistance, à défaut de pouvoir construire une notion de régularité *ad hoc*, il est possible de montrer une inégalité dans le cadre plus abstrait.

Lemme 3.12 (Regularité-dimension, Chapitre 2 Lemme 3.13, [77])**.** *Soit X un espace métrique LLC compact. En gardant les mêmes notations que précédemment,*

$$\mathcal{L}_0(f) = \overline{\dim}\, T_f \leq \frac{1}{\alpha^*} \overline{\dim}\, X \,, \tag{1.65}$$

où

$$\alpha^* := \sup \left\{ 0 \leq \alpha \leq 1 \,\middle|\, \exists \lambda \in \mathrm{Homeo}(X) \,,\ \|f \circ \lambda\|_{C^\alpha} < \infty \right\} \,. \tag{1.66}$$

Cette inégalité est parfois stricte. Par exemple, pour toute fonction de Morse sur une variété riemannienne compacte de dimension > 1, cette inégalité est stricte. Néanmoins, il est possible de montrer que le cas d'égalité est toujours vrai génériquement au sens de Baire dans une certaine classe de régularité. Ce résultat était déjà connu de S. Weinberger et de Y. Baryshnikov dans des travaux jamais publiés, mais dans un cadre beaucoup plus restreint. Je remercie S. Weinberger et Y. Baryshnikov qui m'ont communiqué leur manuscrit.

Théorème 3.6 (Chapitre 2 Théorème 3.19, [77])**.** *Soit X un espace LLC compact admettant un ensemble de dimension upper-box homogène. Pour tout $0 < \alpha \leq 1$, $\mathcal{L}_0(f) = \frac{1}{\alpha}\overline{\dim}(X)$ génériquement au sens de Baire dans $C^\alpha(X,\mathbb{R})$, i.e. l'ensemble sur lequel $\alpha\mathcal{L}_0(f) < \overline{\dim}(X)$ est maigre dans $C^\alpha(X,\mathbb{R})$.*

Ainsi, en utilisant le lemme de régularité-dimension et ce théorème, nous retrouvons le résultat de Kozma, Lotker et Stupp [59] dans un cadre plus restreint. Plus précisément, nous avons montré que

Corollaire 3.13. *Pour notre notion de $\dim^0_{\mathrm{PH}}$, pour tout espace LLC compact admettant un ensemble de dimension upper-box homogène,*

$$\dim^0_{\mathrm{PH}}(X) = \overline{\dim}(X) \,. \tag{1.67}$$

En particulier, en degré d'homologie zéro et pour cette classe d'espaces topologiques, la notion de dimension de Schweinhart et la nôtre coïncident.

Théorème 3.7 (Chapitre 2 Théorème 3.23, [77])**.** *Soit X une variété riemannienne compacte de dimension d. Alors, pour tout $0 \leq k < d$,*

$$\dim_{\mathrm{PH}}^k(X) = d \tag{1.68}$$

et génériquement au sens de Baire dans $C^\alpha(X, \mathbb{R})$, $\mathcal{L}_k(f) = \frac{d}{\alpha}$. De plus, les notions de dimensions homologiques de Schweinhart et les nôtres coïncident.

Ce dernier résultat s'appuie sur le résultat probabiliste de Divol et Polonik précédemment cité (théorème 2.9), qui montre que sur le cube $[0,1]^d$ les fonctions distance aux nuages de points aléatoires admettent une persistance croissante (et divergente) lorsqu'on augmente le nombre de points du nuage.

Les conditions du théorème ci-dessus peuvent probablement être considérablement assouplies, pourvu que l'on puisse exhiber des fonctions dont la k-ième fonctionnelle Pers_p est arbitrairement grande sur des ensembles arbitrairement petits de X. Le résultat de Divol et Polonik suggère qu'une approche à la Erdős est peut-être plus simple, puisqu'elle ne requiert pas des constructions explicites. Des bons candidats seraient alors les fonctions distance à des nuages de points aléatoires dans des petits sous-ensembles de X.

4 Résultats probabilistes

Les résultats précédents ont établi l'importance de la classe de régularité de f relativement à son indice de persistance. Dans un cadre probabiliste, il est commode d'introduire la classe de régularité suivante.

Définition 4.1. La classe des **fonctions sur X à valeurs réelles presque α-Hölder**, notée $E^\alpha(X, \mathbb{R})$, est la classe des fonctions définies par

$$E^\alpha(X, \mathbb{R}) := \bigcap_{0 \leq \beta < \alpha} C^\beta(X, \mathbb{R}) \tag{1.69}$$

En effet, les processus stochastiques ont souvent une telle régularité. Des exemples notoires incluent le mouvement brownien qui est $E^{\frac{1}{2}}([0,1], \mathbb{R})$, mais aussi le mouvement brownien fractionnaire de paramètre de Hurst H, qui est $E^H([0,1], \mathbb{R})$. En dimension supérieure, une vaste classe d'exemples est donnée par les séries de Fourier sous-gaussiennes sur les tores de dimension arbitraire (*cf.* le livre de Kahane sur les séries aléatoires pour les détails [54]).

Notation 4.2. Soit (X, δ) un espace métrique polonais, nous noterons $\mathcal{P}(X)$ (ou simplement $\mathcal{P}$ si X est sous-entendu) l'ensemble des **mesures de probabilité sur X**. De plus, pour une mesure μ et une application g, nous noterons $g_\sharp\mu$ le poussé en avant de μ par g.

Comme nous l'avons vu, la spécification de la classe de régularité d'un processus nous permet d'inférer des informations sur les diagrammes des trajectoires. Il est aussi possible d'en déduire quelques éléments sur le comportement du diagramme moyen, c'est ce que nous dit le résultat suivant.

Proposition 4.3 (Chapitre 2 Proposition 5.7, [77])**.** *Soit* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espace probabilisé et* f *un processus stochastique à valeurs dans* $\mathbb{R}$ *p.s. dans* $E^\alpha(X, \mathbb{R})$ *sur une variété riemannienne compacte* X *de dimension* d*. Pour tout* $\varepsilon > 0$, $(\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P} \in \mathcal{P}(\mathcal{D}_{\frac{d}{\alpha}+\varepsilon} \cap \mathcal{D}_\infty)$ *et a fortiori dans* $\mathcal{P}(\mathcal{D}_r)$ *pour chaque* $\frac{d}{\alpha} < r < \infty$*. De plus, si* $\frac{d}{\alpha} < q < \infty$ *et pour tout* $\beta < \alpha$, $\mathbb{E}\left[\|f\|^q_{C^\beta(X,\mathbb{R})}\right] < \infty$, *alors* $\mathbb{E}[\mathrm{Dgm}_k(f)] \in \bigcap_{\frac{d}{\alpha} < p \leq q} \mathcal{D}_p$.

4.1 Stabilité par perturbations Wasserstein en distribution

De même, le théorème de stabilité Wasserstein déterministe induit un analogue stochastique de la forme suivante.

Théorème 4.1 (Stabilité des champs aléatoires sous perturbations Wasserstein, Théorème 5.9, [77])**.** *Supposons que* f *et* g *soient deux processus stochastiques sur un espace probabilisé* $(\Omega, \mathcal{F}, \mathbb{P})$ *à valeurs dans* $\mathbb{R}$ *p.s.* $E^\alpha(X, \mathbb{R})$ *sur une variété riemannienne compacte* X *de dimension* d*. Pour tout* $k \in \mathbb{N}$ *et tout* $1 \leq p \leq \infty$,

$$W_{p,d_\infty}((\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm}_k \circ g)_\sharp \mathbb{P}) \leq W_{p,L^\infty}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P}) . \tag{1.70}$$

De plus, si les supports de $f_\sharp \mathbb{P}$ *et* $g_\sharp \mathbb{P}$ *sont compacts dans* $E^\alpha(X, \mathbb{R})$*, alors*

$$d_\infty(\mathbb{E}[\mathrm{Dgm}_k(f)] , \mathbb{E}[\mathrm{Dgm}_k(g)]) \leq W_{\infty,d_\infty}((\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm}_k \circ g)_\sharp \mathbb{P}) , \tag{1.71}$$

et pour chaque $\frac{d}{\alpha} < p \leq q \leq \infty$*, il existe une constante* $C_{X,p,\eta}$ *dépendant des supports de* $f_\sharp \mathbb{P}$ *et* $g_\sharp \mathbb{P}$ *telle que*

$$d_p(\mathbb{E}[\mathrm{Dgm}_k(f)] , \mathbb{E}[\mathrm{Dgm}_k(g)]) \leq W_{q,d_p}((\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm}_k \circ g)_\sharp \mathbb{P}) \leq C_{X,p,\eta} W^\eta_{q\eta,L^\infty}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P}) . \tag{1.72}$$

où $\eta < 1 - \frac{d}{\alpha p}$.

Ce résultat généralise les travaux de Chazal *et al.* [22, Lemme 15]. Il est d'autant plus pratique lorsqu'il est utilisé en conjonction avec le résultat de continuité établi par Divol et Lacombe sur la continuité des fonctionnelles linéaires des diagrammes moyens (proposition 2.19). En particulier, nous avons des résultats quantitatifs lorsque le module de continuité de ces fonctionnelles peut être établi, par exemple pour les paysages de persistance (*persistence landscapes* en anglais), et plus généralement pour des fonctionnelles linéaires satisfaisant les conditions de [35, Théorème 3.1].

Le résultat suivant donne une majoration facile du membre de droite dans les inégalités du théorème précédent.

Proposition 4.4 (Contrôle de W_{p,L^∞}, Chapitre 2 Proposition 5.12, [77])**.** *Soit* f *et* g *deux processus stochastiques p.s.* $E^\alpha(X, \mathbb{R})$ *sur une variété riemannienne compacte* X *de dimension* *d sur un espace probabilisé* $(\Omega, \mathcal{F}, \mathbb{P})$. *Alors,*

$$W_{p,L^\infty}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P}) \leq \|f - g\|_{L^p(\Omega, L^\infty(X,\mathbb{R}))} . \tag{1.73}$$

Enfin, le théorème de stabilité Wasserstein des champs aléatoires devrait être compris comme une borne inférieure sur la distance entre les distributions des diagrammes : il s'agit d'un résultat de semi-continuité inférieure et donc d'un théorème permettant de discriminer des processus. La raison est que si nous ne nous fixons pas une régularité E^α donnée, les boules Wasserstein arbitrairement petites autour de la distribution d'un processus E^α-régulier contiennent en général des distributions de processus dont les diagrammes ne sont dans aucun $\mathcal{D}_p$ pour $p < \infty$. Cet argument suggère qu'un contrôle des boules Wasserstein W_{p,L^∞} par des boules W_{p,d_p} ne peut être obtenu sans faire des hypothèses fortes. Autrement dit, une distance W_{p,d_p} entre deux lois de diagrammes petite n'implique pas que les processus sous-jacents sont proches.

4.2 Fonctions ζ associées à des processus stochastiques

La fonctionnelle Pers_p^p a joué un rôle central dans nos résultats jusqu'à présent. Nous introduisons donc une version de cette fonctionnelle définissable sur n'importe quelle mesure de persistance généralisée μ, puis utilisons cette définition pour définir les fonctions ζ associée à un processus stochastique.

Notation 4.5. Pour $\varepsilon > 0$ nous notons $\Delta_\varepsilon \subset \overline{\mathcal{X}}$ un voisinage tubulaire ouvert de rayon ε autour de la diagonale Δ dans $\overline{\mathcal{X}}$ et Δ_ε^c son complément dans $\mathcal{X}$. De plus, nous notons

$$R_{x,\varepsilon} :=]-\infty, x] \times [x + \varepsilon, \infty[\,. \tag{1.74}$$

Enfin, nous notons

$$R_t := \sup_{[0,t]} f - \inf_{[0,t]} f\,. \tag{1.75}$$

Définition 4.6. Soit $\mu \in \mathcal{D}$ une mesure de Radon sur $\mathcal{X}$. Lorsqu'elle est définie, la **fonction ζ associée à μ** est

$$\zeta_\mu(p) := \mathrm{Pers}_p^p(\mu) = p\,\mathcal{M}[\mu(\Delta_\varepsilon^c)]\,(p)\,. \tag{1.76}$$

De manière analogue, la **fonction ζ locale en x associée à μ** est

$$\zeta_\mu^x(p) := p\,\mathcal{M}[\mu(R_{x,\varepsilon})](p)\,. \tag{1.77}$$

Ces notations prendront plus tard tout leur sens dans l'étude de fonctions ζ associées à des processus stochastiques. La transformée de Mellin laisse déjà entrevoir l'origine de cette terminologie : les fonctions complexes $\zeta(p)$ ainsi obtenues pour certains processus partagent des propriétés analytiques similaires à celles de la fonction ζ de Riemann.

Définition 4.7. Soit f un processus stochastique sur un espace topologique compact X. Sa **fonction ζ (locale en x) associée**, notée ζ_f (ζ_f^x), est la fonction ζ (locale en x) associée à la mesure de Radon $\mathbb{E}[\mathrm{Dgm}(f)]$.

Proposition 4.8 (Chapitre 4 Proposition 2.25, [79]). *Les fonctions ζ et les fonctions ζ locales sont liées par la formule suivante*

$$\zeta_f(p) = p \int_{\mathbb{R}} \zeta_f^x(p-1)\, dx\,. \tag{1.78}$$

Un premier théorème concernant les fonctions ζ est une caractérisation de la convergence Wasserstein entre les mesures de Radon et la convergence de leurs fonctions ζ associées en tant que fonctions complexes de p.

Théorème 4.2 (Caractérisation ζ de la convergence Wasserstein p, Théorème 2.44, [79]). *Soit $(\mu_n)_n \subset \mathcal{D}_p \cap \mathcal{D}_q$ une suite de mesures q-tame et $\mu \in \mathcal{D}_p \cap \mathcal{D}_q$. Supposons que la suite $(\mu_n(\Delta_\varepsilon^c))_n$ est uniformément bornée par une fonction $g :]0,\infty[\to \mathbb{R}_+$ telle que pour $\varepsilon \in]0,1]$, $g(\varepsilon) = O(\varepsilon^{-p})$ lorsque $\varepsilon \to 0$ et sur $[1,\infty[$, $g(\varepsilon) = O(\varepsilon^{-q})$ lorsque $\varepsilon \to \infty$. Alors, les conditions suivantes sont équivalentes :*

1. *Il existe $p < r < q$ tel que $d_r(\mu_n, \mu) \xrightarrow{n\to\infty} 0$.*
2. *Il existe $p < r < q$ tel que $\mu_n \xrightarrow[n\to\infty]{v} \mu$ et $\zeta_{\mu_n}(r) \to \zeta_\mu(r)$.*
3. *Pour presque tout $x \in \mathbb{R}$ et $\varepsilon > 0$, $\mu_n(\Delta_\varepsilon^c) \to \mu(\Delta_\varepsilon^c)$ et $\mu_n(R_{x,\varepsilon}) \to \mu(R_{x,\varepsilon})$.*
4. *Pour presque tout $x \in \mathbb{R}$, $\zeta_{\mu_n}^x \to \zeta_\mu^x$ et $\zeta_{\mu_n} \to \zeta_\mu$ uniformément sur tout compact de $\langle p, q\rangle$.*
5. *Pour tout $p < r < q$, $d_r(\mu_n, \mu) \xrightarrow{n\to\infty} 0$.*

L'utilité de cette caractérisation devient apparente lorsqu'on la couple avec le résultat ci-dessous. Soit $(S_n f)_n$ une suite de processus ayant pour limite L^∞ p.s. f alors

Théorème 4.3 (Chapitre 3 Théorème **??**, [78]). *Soit $\delta_n := \|f - S_n f\|_{L^\infty}$, alors il existe un δ_n-matching entre les codes-barres de $S_n f$ et de f. En particulier, pour tout $\varepsilon \geq 2\delta_n$.*

$$N_f^{\varepsilon+\delta_n} \leq N_{S_n f}^\varepsilon \leq N_f^{\varepsilon-\delta_n} \tag{1.79}$$

Par ailleurs, si $\mathbb{E}\Big[N_f^\varepsilon\Big]$ est continue par rapport à ε, alors

$$N_{S_n f}^\varepsilon \xrightarrow[n\to\infty]{L^1} N_f^\varepsilon \quad \textit{and} \quad N_{S_n f}^\varepsilon \xrightarrow[n\to\infty]{\mathbb{P}} N_f^\varepsilon\,,$$

ce qui, à n fixé, se traduit quantitativement par

$$\mathbb{E}\Big[\Big|N_f^\varepsilon - N_{S_n f}^\varepsilon\Big|\Big] \leq 2\omega_\varepsilon(\delta_n) \quad \textit{and} \quad \mathbb{P}(\Big|N_f^\varepsilon - N_{S_n f}^\varepsilon\Big| \geq k) \leq \frac{2\omega_\varepsilon(\delta_n)}{k} \tag{1.80}$$

où ω_ε est le module de continuité de $\mathbb{E}\Big[N_f^\varepsilon\Big]$ sur l'intervalle $[\varepsilon-\delta_n, \varepsilon+\delta_n]$. Et les mêmes énoncés sont valables en remplaçant N^ε par $N^{x,x+\varepsilon}$ en n'importe quel degré d'homologie.

Ce théorème établit la validité du troisième critère de la caractérisation ζ de la convergence Wasserstein. Pourvu que les hypothèses de la caractérisation soient satisfaites pour un certain

p et q, nous avons alors $d_r(\mathbb{E}[\mathrm{Dgm}(S_n f)], \mathbb{E}[\mathrm{Dgm}(f)]) \xrightarrow{n\to\infty} 0$ pour tout $r \in]p,q[$ établissant ainsi un résultat de continuité séquentielle pour les distances de Wasserstein sur l'espace des diagrammes qui ne dépend plus de la compacité du support des distributions de $S_n f_\sharp \mathbb{P}$ et $f_\sharp \mathbb{P}$. En revanche, ce résultat de continuité séquentielle n'est plus quantitatif.

Remarque 4.9. La continuité de $\mathbb{E}[N^\varepsilon]$ et de $\mathbb{E}[N^{x,x+\varepsilon}]$ par rapport à ε est immédiate pourvu que $\mathbb{E}[\mathrm{Dgm}(f)]$ admette une densité par rapport à la mesure de Lebesgue sur $\mathcal{X}$. Des critères établissant l'absolue continuité de ces mesures de Radon ont été obtenus par Chazal and Divol dans [20] pour les nuages de points aléatoires ainsi que pour le mouvement brownien. Pour le mouvement brownien, nous donnerons plus tard une expression explicite de la densité de son diagramme moyen.

Processus en dimension 1

De la théorie déjà bien connue des semimartingales, nous pouvons tirer les deux conditions suivantes sur leurs fonctions ζ et les développements asymptotiques de N^ε.

Proposition 4.10 (Chapitre 4 Proposition **??**, [79])**.** *Soit f une semimartingale continue $f = M + A$ sur l'intervalle $[0,t]$ telle que, pour $s \geq 1$,*

$$\mathbb{E}\Big[[M]_t^{s/2} + \Big(\int_0^t |dA|_s\Big)^s\Big] < \infty\,. \tag{1.81}$$

Alors, la fonction ζ (locale) de f est méromorphe sur $\mathrm{Re}(p) \geq 2$ (resp. $\mathrm{Re}(p) \geq 1$) avec un seul pôle simple en $p = 2$ (resp. $p = 1$). De plus, si f est une semimartingale continue et $[f]_t < \infty$, alors dans L^s

$$N^\varepsilon \sim \frac{[f]_t}{2\varepsilon^2} \quad \textit{lorsque } \varepsilon \to 0\,. \tag{1.82}$$

Corollaire 4.11 (Chapitre 4 Corollaire **??**, [79])**.** *Si f est une semi-martingale continue, alors, en espérance, $\mathrm{Pers}_p^p(f)$ admet un pôle d'ordre 1 en $p = 2$ de résidu $[f]_t$.*

La structure d'ordre total sur $\mathbb{R}$ permet une analyse plus fine (voir même des calculs explicites !) des fonctions ζ des processus stochastiques et du comportement asymptotique de N^ε et $N^{x,x+\varepsilon}$ et en particulier des processus de Lévy. En effet, cette structure d'ordre total permet la définition de la suite de temps d'arrêt définie par Neveu [74], qui compte effectivement N^ε (on peut définir une suite de temps d'arrêt analogues qui compte $N^{x,x+\varepsilon}$). Les processus de Lévy sont ici particulièrement bien adaptés, puisque ces temps d'arrêt sont dans ce cas tous i.i.d.. Par conséquent, il suffit de s'intéresser à la variable aléatoire

$$U^\varepsilon := T^\varepsilon + S^\varepsilon \tag{1.83}$$

où T^ε et S^ε sont deux variables aléatoires indépendantes déjà définies dans un énoncé d'un théorème de Picard, à savoir

$$T^\varepsilon := \inf\{t \geq 0 \,|\, f(t) - \inf_{[0,t]} f > \varepsilon\} \quad \text{et} \quad S^\varepsilon := \inf\{t \geq 0 \,|\, \sup_{[0,t]} f - f(t) > \varepsilon\} \tag{1.84}$$

Le comptage de N^ε devient alors à quelques détails près un problème analogue à celui de la théorie du renouvellement (*cf.* [48] pour une introduction approfondie). Par abus de notation nous noterons $U := U^1$. Avec cette notation, il est possible de montrer le résultat suivant.

Théorème 4.4 (Chapitre 4 Théorème 3.11, [79])**.** *Soit f un processus de Lévy tel que f soit presque sûrement nulle part monotone, alors pour tout $n \in \mathbb{N}$*

$$\mathbb{E}[N_t^\varepsilon] = \frac{t}{\mathbb{E}[U^\varepsilon]} + \left(\frac{\mathbb{E}[(U^\varepsilon)^2]}{2\mathbb{E}[U^\varepsilon]^2} - 1\right) + \mathbb{P}(R_t \geq \varepsilon) + o(\rho_\varepsilon^{-n}) \quad \textit{as } \varepsilon \to 0\,, \tag{1.85}$$

où $\rho_\varepsilon \geq -\log(\mathbb{P}(T^\varepsilon > 1) \vee \mathbb{P}(S^\varepsilon > 1))$ est le rayon de convergence de la série de Taylor de $\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]$ autour de $\lambda = 0$ ($\rho_\varepsilon > 1$ pour ε suffisamment petit). Si f est α-stable, la formule ci-dessus devient

$$\mathbb{E}[N_t^\varepsilon] = \frac{t}{\mathbb{E}[U]\,\varepsilon^\alpha} + \frac{\mathbb{E}[U^2]}{2\mathbb{E}[U]^2} + o(\varepsilon^{\alpha n}) \quad \textit{lorsque } \varepsilon \to 0,. \tag{1.86}$$

Ce résultat est différent de celui obtenu par le biais Théorème Central Limite (TCL) des processus de comptage (théorème 3.6), car

- Les processus que l'on considère ici ne sont pas nécessairement auto-similaires, et donc la limite $t \to \infty$ n'est *a priori* pas interchangeable avec la limite $\varepsilon \to 0$ dans l'énoncé du TCL.
- Crucialement, nous avons quantifié l'ordre du reste dans le développement asymptotique ci-dessus.

Définition 4.12. La **fonction ζ sans barre infinie** du processus stochastique f sur $[0,t]$ est définie comme suit

$$\hat{\zeta}_f(p) := \mathbb{E}\left[\mathrm{Pers}_p^p(f) - R_t^p\right]\,. \tag{1.87}$$

Théorème 4.5 (Chapitre 4 Théorème 3.21, [79])**.** *Avec les notations précédentes, soit*

$$B(z) := \frac{\mathbb{E}\left[e^{zU}\right]}{1 - \mathbb{E}[e^{-zU}]}\,, \tag{1.88}$$

alors la fonction ζ sans barre infinie associée à un processus de Lévy α-stable est donnée par

$$\hat{\zeta}_f(p) = \frac{t^{\frac{p}{\alpha}}}{\Gamma(\frac{p}{\alpha})} B^*\left(\frac{p}{\alpha}\right) \tag{1.89}$$

et s'étend à une fonction méromorphe de p sur $\mathbb{C}$ (puisque B^* est lui-même méromorphe), avec un unique pôle simple en $p = \alpha$ de résidu $\mathbb{E}[U]^{-1}\,\alpha t$.

Remarque 4.13. Pourvu qu'une expression explicite pour B^* soit calculée, cette expression permet de calculer les contributions superpolynomiales restantes de la série asymptotique de N^ε. Il sera possible de calculer ces contributions dans le cas du mouvement brownien.

4.3 Conséquences et études de cas particuliers

Nous avons explicitement calculé les fonctions ζ de plusieurs processus. Nous énoncerons ici simplement les résultats que nous avons obtenus pour le mouvement brownien standard sur $[0,t]$, B : les formules explicites pour les fonctions ζ, le développement asymptotique complet de N^ε et de $N^{x,x+\varepsilon}$ pour $x>0$, l'expression explicite de la densité de $\mathbb{E}[\mathrm{Dgm}(B)]$ par rapport à la mesure de Lebesgue, dont l'existence a originellement été démontrée par Chazal et Divol [20], et enfin, l'expression de tous les moments et les lois de la k-ième barre la plus grande de B.

Théorème 4.6 (Théorèmes 4.2 et 4.6)**.** *La fonction ζ du mouvement brownien sur l'intervalle $[0,t]$ admet une extension méromorphe à $\mathbb{C}$ tout entier et est donnée par l'expression suivante*

$$\zeta_B(p) = \frac{4(2^p-3)}{\sqrt{\pi}}\left(\frac{t}{2}\right)^{\frac{p}{2}}\Gamma\left(\frac{p+1}{2}\right)\zeta(p-1) \tag{1.90}$$

pour tout p. En particulier, elle possède un unique pôle simple en $p=2$ de résidu $[B]_t = t$. De plus, la fonction ζ locale du mouvement brownien au niveau $x>0$ admet une extension méromorphe à $\mathbb{C}$ tout entier donnée par

$$\zeta_B^x(p) = 2^{-\frac{3p}{2}}\left(2^p-1\right)t^{\frac{p}{2}}\zeta(p)\Gamma(p+1)\left(\frac{{}_1F_1\left(\frac{-p}{2};\frac{1}{2};\frac{-x^2}{2t}\right)}{\Gamma\left(\frac{p}{2}+1\right)} - \sqrt{\frac{2x^2}{t}}\frac{{}_1F_1\left(\frac{1-p}{2};\frac{3}{2};\frac{-x^2}{2t}\right)}{\Gamma\left(\frac{p+1}{2}\right)}\right). \tag{1.91}$$

où ${}_1F_1$ désigne la fonction hypergéométrique confluente.

En utilisant la correspondance fondamentale, il est possible d'extraire de ces formules des séries asymptotiques pour N^ε autour de 0 et de l'infini.

Corollaire 4.14 (Chapitre 4 Propositions 4.4 et 4.7, [79])**.** *Pour un mouvement brownien standard sur $[0,t]$, $\mathbb{E}[N_t^\varepsilon]$ admet les représentations en série suivantes qui convergent bien pour les grands et petits ε respectivement*

$$\begin{aligned}\mathbb{E}[N_t^\varepsilon] &= 4\sum_{k\geq 1}(2k-1)\,\mathrm{erfc}\left(\frac{(2k-1)\varepsilon}{\sqrt{2t}}\right) - k\ \mathrm{erfc}\left(\frac{2k\varepsilon}{\sqrt{2t}}\right)\\ &= \frac{t}{2\varepsilon^2} + \frac{2}{3} + 2\sum_{k\geq 1}(2(-1)^k-1)\frac{e^{-\pi^2k^2t/2\varepsilon^2}t}{\varepsilon^2}\left[1+\frac{\varepsilon^2}{\pi^2k^2t}\right].\end{aligned}$$

De plus, pour $x>0$,

$$\begin{aligned}\mathbb{E}\left[N^{x,x+\varepsilon}\right] &= \sum_{k=1}^{\infty}\mathrm{erfc}\left(\frac{x+(2k-1)\varepsilon}{\sqrt{2t}}\right)\\ &\sim \frac{1}{2\varepsilon}\int_0^t\varphi(x,s)\,ds + \sum_{k\geq 0}\frac{4(-2)^k\left(2^{2k+1}-1\right)\zeta(2k+2)}{\pi^{2k+2}}\left[\frac{\partial^k}{\partial t^k}\varphi(x,t)\right]\varepsilon^{2k+1}\ \text{as } \varepsilon\to 0\,.\end{aligned}$$

où $\varphi(x,t)$ est la densité d'une variable gaussienne de moyenne 0 et variance t.

Corollaire 4.15 (Chapitre 4 Proposition 4.10, [79])**.** *Pour $x>0$ et $\varepsilon>0$, la densité de $\mathbb{E}[\mathrm{Dgm}(B)]$ par rapport à la mesure de Lebesgue en coordonnées de naissance-persistance (cf. la* *remarque 2.32) est*

$$g(x,\varepsilon)=\sqrt{\frac{2}{\pi t^3}}\sum_{k=1}^{\infty}(2k-1)(x+(2k-1)\varepsilon)\,e^{-\frac{(x+(2k-1)\varepsilon)^2}{2t}}\,. \tag{1.92}$$

Proposition 4.16 (Chapitre 4 Section 4.1, [79])**.** *Pour $k\geq 2$, la distribution de la longueur de la kième plus longue barre de B, $\ell_k(B)$ est caractérisée par sa transformée de Mellin qui est donnée par*

$$\mathbb{E}[\ell_k^p(X)]=\frac{1}{k!}\left.\frac{\partial^k}{\partial z^k}\right|_{z=0}G_2(z;p)\,. \tag{1.93}$$

où

$$G_2(z;p)=8\,\frac{\Gamma(p)}{\Gamma(\frac{p}{2})}\left(\frac{t}{8}\right)^{\frac{p}{2}}z^2\sum_{n=0}^{\infty}\sum_{k\geq 1}\frac{a_k^{(n)}(0)}{n!}\,\frac{z^n}{k^p} \tag{1.94}$$

et

$$a_k^{(n)}=\frac{2^n(-1)^{k+n+1}}{k}\prod_{i=0}^{n}\frac{k^2-i^2}{2i+1} \tag{1.95}$$

Pour les premiers k, nous avons

$$\begin{aligned}
\mathbb{E}[\ell_2^p(B)]&=\frac{2^{3-\frac{5p}{2}}t^{\frac{p}{2}}\Gamma(p)}{\Gamma(\frac{p}{2})}(2^p-2^2)\zeta(p-1)\\
\mathbb{E}[\ell_3^p(B)]&=\frac{2^{4-\frac{5p}{2}}t^{\frac{p}{2}}\Gamma(p)}{3\Gamma(\frac{p}{2})}[(2^p-2^2)\zeta(p-1)-(2^p-2^4)\zeta(p-3)]\,,
\end{aligned}$$

où ζ désigne ici la fonction ζ de Riemann. Nous en déduisons des formules explicites pour $\mathbb{P}(\ell_k\geq\varepsilon)$,

$$\begin{aligned}
\mathbb{P}(\ell_2(B)\geq\varepsilon)&=4\sum_{k\geq 1}k\left[\operatorname{erfc}\left(k\varepsilon\sqrt{\frac{2}{t}}\right)-4\operatorname{erfc}\left(2k\varepsilon\sqrt{\frac{2}{t}}\right)\right]\\
\mathbb{P}(\ell_3(B)\geq\varepsilon)&=\frac{8}{3}\sum_{k\geq 1}k\left[4\left(4k^2-1\right)\operatorname{erfc}\left(2k\varepsilon\sqrt{\frac{2}{t}}\right)-\left(k^2-1\right)\operatorname{erfc}\left(k\varepsilon\sqrt{\frac{2}{t}}\right)\right]\,.
\end{aligned}$$

4.4 Tests statistiques

L'estimation de paramètres est un sujet largement étudié, notamment pour les processus auto-similaires en dimension 1 (une liste non-exhaustive de références sur le sujet est contenue dans la bibliographie de [32]). Une variété de méthodes différentes comme les ondelettes multi-échelles ou bien certaines méthodes analytiques ont été utiles dans le développement de ces estimateurs. Notre approche n'offre donc rien de nouveau à cet égard, ni des meilleurs estimateurs en dimension 1. L'intérêt de notre méthode réside plutôt dans son aptitude à être généralisée à des champs aléatoires en dimension supérieure, pour lesquels l'utilisation des ondelettes n'est plus possible. Un cadre théorique complet nécessiterait une étude quantitative des code-barres des champs aléatoires de dimension supérieure, ce que nous n'avons pas réalisé dans cette thèse. Le développement du test statistique ci-dessous devrait être vu comme une preuve de concept concernant l'utilisation des propriétés de l'homologie persistante pour discriminer des processus aléatoires.

Dans ce qui suit, nous considérerons que f est un processus de Lévy α-stable, dont nous chercherons à estimer le paramètre α. D'après la proposition 2.21, nous savons que presque sûrement

$$N_t^\varepsilon \sim Ct\varepsilon^{-\alpha} \quad \text{lorsque } \varepsilon \to 0\,. \tag{1.96}$$

Remarque 4.17. En fait, le même raisonnement nous permet d'estimer le paramètre de Hurst H d'un mouvement brownien fractionnaire (fBM), qui présente également une auto-similarité. Dans ce cas, l'analogue du résultat asymptotique de la proposition 2.21 est [80, §3]

$$N^\varepsilon \sim Ct\varepsilon^{-\frac{1}{H}} \quad \text{p.s. lorsque } \varepsilon \to 0\,. \tag{1.97}$$

Étant donné un échantillonnage du processus, notre construction d'un estimateur de α se fait comme suit.

1. Nous échantillonnons M trajectoires indépendantes de f à intervalles réguliers de taille $\frac{1}{N}$ pour un certain N et nous interpolons linéairement entre les valeurs échantillonnées ;

2. Nous calculons le code-barres des trajectoires ainsi construites. Pour ce faire, nous voyons cette interpolation linéaire comme un complexe simplicial filtré (qui n'est dans ce cas rien d'autre qu'une chaîne avec $\approx N$ arêtes). Les sommets du complexe sont les points échantillonés directement de f et les interpolations sont vues comme des arêtes joignant les sommets. La filtration sur ce complexe sur chaque sommet a est $f(a)$ et la valeur de la filtration pour une arête reliant le sommet a au sommet b est donnée $f(a) \wedge f(b)$. L'homologie persistante de ce complexe peut être calculée. Cet algorithme peut être implementé à l'aide de `gudhi` [50] sur `python`.

3. Pour *une certaine plage de ε assez petits*, et pour une certaine constante positive $c > 1$, nous calculons la quantité

$$\hat{\alpha}_M := \log_c \left[\frac{\overline{N}_t^{\varepsilon/c} - \overline{N}_t^{2\varepsilon/c}}{\overline{N}_t^{\varepsilon} - \overline{N}_t^{2\varepsilon}} \right]\,. \tag{1.98}$$

   Ici, la notion d'*une certaine plage de ε assez petits* et la constante c dépendent toutes les deux de N, avec la condition limite que lorsque $N \to \infty$, la borne inférieure de l'intervalle des ε admissibles tend vers 0.

La quantité calculée $\hat{\alpha}_M$ est un estimateur du paramètre α (pour fBM, c'est un estimateur de $\frac{1}{H}$). En pratique, il est possible de voir l'intervalle des ε admissibles en traçant la courbe $\overline{N}_t^\varepsilon$ sur une échelle log-log et en considérant les ε correspondant au régime linéaire le plus proche de zéro (*cf.* la figure 1.5).

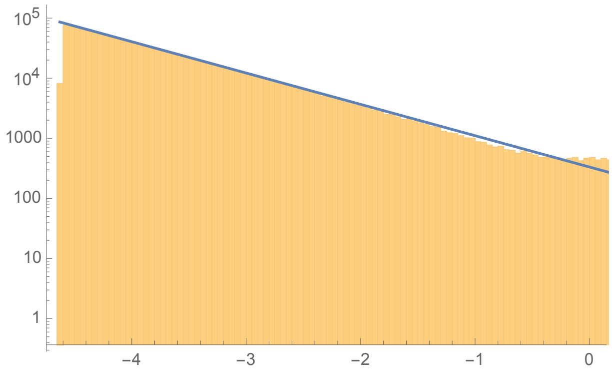


Figure 1.5: En orange, l'histogramme du nombre de barres de longueur $\geq \varepsilon$ en fonction de $\log \varepsilon$ trouvé à partir d'une simulation d'un processus stable de Lévy 1.2 comme une marche aléatoire. En bleu, la fonction $C_{1.2}\varepsilon^{-1.2}$.

Lemme 4.18 (Convergence des moyennes empiriques, Chapitre 4 Lemme 3.27, [79])**.** *La limite*

$$\frac{\overline{N}_t^{\varepsilon/c} - \overline{N}_t^{2\varepsilon/c}}{\overline{N}_t^{\varepsilon} - \overline{N}_t^{2\varepsilon}} \xrightarrow[M\to\infty]{\mathbb{P}} \frac{\mathbb{E}\left[N_t^{\varepsilon/c} - N_t^{2\varepsilon/c}\right]}{\mathbb{E}\left[N_t^{\varepsilon} - N_t^{2\varepsilon}\right]} \tag{1.99}$$

converge à une vitesse $C_s M^{-s}$, pour chaque $1 \leq s \leq 2$ où C_s est une constante dépendant de s et du s-ième moment de $N^{\varepsilon/c}$. En particulier,

$$\hat{\alpha}_M \xrightarrow[M\to\infty]{\mathbb{P}} \alpha + \xi(\varepsilon) \tag{1.100}$$

à la même vitesse, où $\xi(\varepsilon)$ est une fonction superpolynomialement petite de ε.

Remarque 4.19. L'expression alambiquée de l'estimateur $\hat{\alpha}_M$ peut être comprise grâce à la série asymptotique de N^ε pour les processus α-stables. Les soustractions présentes au numérateur et au dénominateur sont effectuées de sorte que les termes constants du développement s'annulent. En ignorant les corrections super-polynomiales (qui sont négligeables), l'argument du log de l'estimateur est

$$c^{\hat{\alpha}_M} \approx \frac{\frac{t}{\mathbb{E}[U](\varepsilon/c)^\alpha} - \frac{t}{\mathbb{E}[U](2\varepsilon/c)^\alpha}}{\frac{t}{\mathbb{E}[U]\varepsilon^\alpha} - \frac{t}{\mathbb{E}[U](2\varepsilon)^\alpha}} \approx c^\alpha\,. \tag{1.101}$$

5 Conclusion et perspectives

Dans cette thèse, nous nous sommes intéressés à la continuité des diagrammes de persistance pour différentes topologies de l'espace de diagrammes par rapport à la norme L^∞ sur l'espace des fonctions. Ce faisant, nous avons étudié les propriétés de la fonctionnelle Pers_p et introduit une notion de dimension qui caractérise bien les propriétés de continuité mentionnées et que nous avons pu lier à des notions de dimension classiques pour les espaces métriques. Nous avons exploité ces résultats afin d'étudier la topologie des ensembles de surniveau des processus stochastiques, et notamment leur stabilité par rapport à des perturbations au niveau de leurs distributions. De plus, dans certains cas particuliers assez simples, nous avons pu décrire explicitement la distribution de ces ensembles de surniveau et nous nous en avons tiré un test de paramètres statistique.

À la lumière de ces résultats, il serait intéressant d'explorer les éléments suivants, qui ne forment en aucun cas une liste exhaustive des ouvertures possibles de cette thèse.

Le théorème de stabilité des diagrammes par rapport à la norme C^0 classique et ses repercussions sur les processus stochastiques sont des outils utiles, mais insuffisants. Pour élargir le champ des possibles, il serait intéressant de prouver des théorèmes de stabilité pour d'autres topologies sur d'autres espaces de fonctions. Dans le cas de la dimension 1, notre intérêt pour les processus de Lévy ainsi que l'intérêt que nous avons porté aux marches aléatoires dans le chapitre 3 suggère qu'une bonne première généralisation du théorème de stabilité en dimension 1 serait l'extension de son régime de validité à toutes les fonctions càdlàg, munies de la topologie de Skorokhod. Ceci permettrait en outre de donner un cadre théorique général aux résultats constatés dans les chapitres 3 et 4. En métrisant cette topologie, il en suivrait des théorèmes analogues au théorème de stabilité stochastique, dont nous avons précédemment discuté. En dimension supérieure, le théorème de continuité de Kolmogorov et les résultats du chapitre 4 suggèrent que les espaces de Sobolev $W^{k,p}$ (et en particulier ceux qui s'injectent dans un certain C^α) sont des bons candidats d'espaces des fonctions, pour lesquels on peut espérer démontrer des résultats de stabilité (stochastique). Cet intérêt fait d'ailleurs écho à un résultat de Polterovitch *et al.* [85, Theorème 6.4.1] qui donne une inégalité entre $\sum_k \mathrm{Pers}_1(H_k(X,f))$ et la norme $W^{2,2}$ pour les fonctions f sur le tore $\mathbb{R}^2/\mathbb{Z}^2$.

Toujours concernant les liens entre la régularité et les indices de persistance, il serait intéressant d'étudier les indices de persistance de fonctions régulières, voir lisses, (génériques) sur lesquelles se concentre la plupart de la littérature existante dans le domaine des champs aléatoires [3, 4, 7].

En vue du théorème 5.9 du chapitre 4 et du développement du test statistique dans le cadre des processus de Lévy en dimension 1, il serait intéressant de concevoir des tests pour des classes de champs aléatoires plus vastes.

Un autre axe de recherche pertinent concerne les moyennes de Fréchet définissables sur les espaces des diagrammes $(\mathcal{D}_p, d_p)$ [96]. Ces moyennes ne concordent en général pas avec la moyenne précédemment définie par dualité et dépendent généralement de p. En revanche, elles présentent l'avantage de ne pas quitter l'espace des diagrammes, contrairement à la moyenne à laquelle nous nous sommes intéressés dans cette thèse, qui vit sur la complétude par convergence vague (ou pour la distance d_p) de cet espace. Un analogue du théorème de stabilité stochastique (théorème 5.9 du chapitre 4) entraînerait les deux conséquences suivantes. D'abord, ça montrerait que les moyennes de Fréchet des diagrammes des processus stochastiques sont des quantités stochastiquement robustes (*i.e.* stables par perturbations aux distributions des processus sous-jacents). Puis, du fait que les inégalités du théorème dépendraient de p et du fait que la convergence en W_p est caractérisée par la convergence vague et la convergence des p-ièmes moments, ces inégalités pourraient orienter un choix pour le paramètre p avec des garanties théoriques pour les applications pratiques.

Un dernier point sur le théorème de stabilité stochastique à élargir concerne une caractérisation de la vitesse de convergence des moyennes empiriques de diagrammes vers le diagramme moyen (défini par dualité). Ceci donnerait des garanties théoriques sur la distance entre les moyennes empiriques, telles que simulées, et les vrais diagrammes moyens de processus stochastiques, qui demeurent à ce jour incalculables en dimension supérieure. Ce faisant, on aurait une théorie effective, complètement applicable dans des cadres pratiques.

Enfin, il est pertinent de s'intéresser aux nuages de points aléatoires (comme l'ont fait Divol et Polonik pour le cube [35]) sur des espaces métriques plus généraux. En effet, si des résultats de généricité au sens de Baire sont vrais dans des espaces plus généraux, on peut espérer que les nuages de points aléatoires fournissent les perturbations qui saturent les inégalités entre les indices de persistance et la dimension upper-box souhaités. Cette démarche pourrait donc facilement généraliser les résultats de généricité à des espaces plus généraux.

Chapter 2

On C^0 persistent homology and trees

Abstract

We revisit and extend some classical results on persistent homology. We start by extending the notion of merge trees to all continuous functions on some general topological spaces. We revisit the concept of homological dimension, previously introduced by other authors and show that the suprema in the definitions of these concepts is attained generically in the sense of Baire. We then generalize the Wasserstein stability theorem to irregular settings, giving explicit bounds on the constants in the theorem and sharp bounds on its regime of validity. Finally, we use this generalized Wasserstein stability theorem to show a stochastic stability theorem for persistence diagrams.

Contents

1 Introduction

1.1 State of the art

The topology of superlevel sets of a function has been a subject of study in different mathematical communities. In the probability theory, the introduction of trees lead to the understanding of connected components of superlevel sets (called *excursions* in the probabilistic setting) of (irregular) random functions on $[0, 1]$ [29, 37, 38, 80].

More recently, so-called *merge trees* have also appeared in topological data analysis (TDA) (*cf.* the books by Chazal *et al.* [24] and Oudot's book [76] for an introduction to TDA). As in the probabilistic case, these trees carry important information about the connected components of superlevel sets and moreover about the 0th degree homology persistence diagram of a Morse function f defined on a compact manifold X [30, 31, 68, 73, 97] (an explicit construction and correspondence between trees and barcodes can be found in [30]).

The construction of these trees is different in both cases: the approach of the probabilists is analytic [29, 37], whereas merge trees are more algebraic in nature [30, 31, 73]. Since these trees capture essentially the same information about the connected components of superlevel sets, one can ask whether both constructions coincide where their regimes of validity intersect.

Parallel to this development, Wasserstein p distances on the space of diagrams (denoted d_p) [39, Chapter VIII.2] have been widely used and studied by the TDA community in different contexts [19, 28, 33, 71, 96]. Recently, Wasserstein distances have been formalized through the use of optimal *partial* transport by Divol and Lacombe [33]. In this formalism, persistence diagrams are regarded as measures and use optimal transport theory to extend the notion of Wasserstein p distances, previously defined on persistence diagrams, to Radon measures on the upper-half plane $\mathcal{X} \subset \mathbb{R}^2$. It is to be noted that regarding persistence diagrams as measures is a point of view which had already been introduced [24, 76] and which has proved fruitful independently from the considerations regarding Wasserstein distances.

The extension to all Radon measures comes with certain advantages, such as having an easily definable and computable notion of "average diagram", defined by duality. This notion was originally introduced by Chazal and Divol in [20] as follows. If f is a random function, seeing $\mathrm{Dgm}(f)$ as a measure, it is possible to define the average diagram of the process by duality in the following way. For every measurable set $B \subset \mathcal{X}$,

$$\mathbb{E}[\mathrm{Dgm}(f)]\,(B) := \mathbb{E}[\mathrm{Dgm}(f)(B)]\ . \tag{2.1}$$

From the definition, $\mathbb{E}[\mathrm{Dgm}(f)]$ encodes every linear functional of the diagram and is easily computed, motivating its introduction. This definition contrasts the Fréchet means approach of other authors (*e.g.* Turner *et al.* [96]), which is non-linear, depends on p and requires a proof of existence and unicity, but does not require the extension of the space of persistence diagrams to the space of arbitrary measures on $\mathcal{X}$.

This dual approach of Chazal and Divol inscribes itself in the more general context of the study of the persistence diagrams of stochastic processes, which have been studied by a wide variety of authors, for instance in [2, 3, 4, 9, 20, 21, 22, 78, 96]. Some of the previously cited results discuss different aspects of random field persistence theory, which include, but are not limited to, computations for canonical processes [9, 78], stability of certain linear functionals with respect to the bottleneck distance [22], the Euler characteristic [4], random complexes [4] and notions of central tendency [20, 96].

Given the widespread use of Wasserstein p distances, it is important to understand whether this notion is stable with respect to perturbations to filtration functions on X. This so-called "Wasserstein stability" of persistence diagrams of functions $f : X \to \mathbb{R}$ has been widely discussed by the TDA community in the context where the space X is triangulable. There are many results in this direction [25, 28, 93], valid with different degrees of generality, covering both X compact [28, 93] and X non-compact [25], but mainly focusing on Lipschitz functions (however, the work of Chen and Edelsbrunner [25] does not require the Lipschitz condition). The first result in this direction was obtained by Cohen-Steiner *et al.* [28] and depends on the following *ad-hoc* condition on X.

Definition 1.1. [28] A (triangulable) metric space X **implies bounded q-total persistence** if, for all $k \in \mathbb{N}$, there exists a constant C_X that depends only on X such that

$$\mathrm{Pers}_q^q(\mathrm{Dgm}_k(f)) < C_X \tag{2.2}$$

for every tame function f with Lipschitz constant $\mathrm{Lip}(f) \leq 1$.

The Pers_p-functional of the definition above is the usual p-persistence used in TDA (a non-exhaustive list of uses of this functional includes [1, 19, 33, 71, 96]) and is defined as the ℓ^p-norm of the length of the bars of the barcode of f. The results obtained thereafter rely heavily on this condition, which is not rendered quantitative (in particular, given X, no upper bound for C_X or lower bound for q had been established in general). Nonetheless, this condition allowed the authors to show a Wasserstein stability result.

Theorem 1.2 (Cohen-Steiner, Edelsbrunner, Harer, [28])**.** *Let X be a triangulable space implying bounded q-total persistence and let f and g be two $\mathbb{R}$-valued Lipschitz functions on X. Then, for all $p > q$, we have*

$$d_p(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq C_X (\mathrm{Lip}(f) \vee \mathrm{Lip}(g))^{\frac{q}{p}} \|f - g\|_\infty^{1-\frac{q}{p}} , \tag{2.3}$$

where $\mathrm{Lip}(f)$ *denotes the Lipschitz constant of f and $a \vee b := \max\{a, b\}$.*

Further results in this direction, such as the ones in [93], also rely on the bounded q-total persistence condition, but give lower bounds on admissible q, finding that $q \geq d$, where d is the maximal dimension of simplices in the triangulation of X. It is also known that, for distance functions to point clouds in $\mathbb{R}^d$, $q = d$.

We will later see that the lower bound for the validity of Wasserstein stability is closely related to a different question regarding the link between the so-called *homological dimensions* of X and the upper-box dimension of X, which we will denote $\overline{\dim}(X)$ (analogously, we will denote $\underline{\dim}(X)$ the lower-box dimension). To the best knowledge of the author, although Yuliy Baryshnikov and Shmuel Weinberger had previously obtained results in this direction (but never published them), this question was first opened and studied by Schweinhart and MacPherson [66] and later studied in more detail by Schweinhart in [91], but has also been addressed by other authors (*cf.* [2] and the references therein).

Definition 1.3 (Schweinhart's definition of $\dim_{\mathrm{PH}}^k$, [91])**.** Let X be a bounded subset of a metric space. The PH_k-dimension of X is

$$\dim_{\mathrm{PH}}^k(X) := \inf_p \{\sup_{\mathbf{x}} \mathrm{Pers}_p(\mathrm{Dgm}_k(d(-,\mathbf{x}))) < \infty\}\,, \tag{2.4}$$

where the supremum is taken over all finite sets of points $\mathbf{x}$ of X.

There are open problems stated in Schweinhart's paper regarding the relation between these notions of dimension and $\overline{\dim}(X)$, some of which we will give a partial answer to in this paper.

1.2 Our contribution

This paper mainly extends previously known results about the persistence theory of filtrating functions f to more irregular settings. More precisely, we focus on relaxing the assumptions on not only the regularity of the underlying functions f, but also on the nature of the compact metric space X over which they are defined. The main contributions of this paper are mostly contained under the scope of the four following theorems, which are not stated in their full generality here for the sake of brevity and clarity, but which contain the general ideas hereby explored. We kindly refer the reader to the corresponding theorems for the full generality of the statements.

Theorem 1.4 (Merge tree extension to C^0-functions (section 2.1 and theorem 2.20))**.** *Let X be a compact, connected, locally path connected topological space and k be a field. There exists a map $T : C^0(X,\mathbb{R}) \to$ **Tree** associating $f \mapsto T_f$ and a functor* Alg : **Tree** $\to$ **PersMod**$_k$ *such that*

$$\mathrm{Alg}(T_f) = H_0(X, f)\,, \tag{2.5}$$

where **Tree** *is the category of rooted $\mathbb{R}$-trees seen as metric spaces, whose morphisms are isometric embeddings preserving the roots and where* **PersMod$_\mathbf{k}$** *is the observable category of q-tame persistence modules over a field k.* ■

This result extends the results of Curry linking merge trees and persistence barcodes [30], by extending the regime of validity of the theory of merge trees from a Morse setting to a C^0 one. The theorem draws its inspiration from the constructions previously made by Le Gall and Curien [29, 37]. We subsequently show that the map assigning a continuous function $f : [0,1] \to \mathbb{R}$ to its constructed tree T_f is a surjection onto the space of trees of finite upper-box dimension and provide an explicit construction of an inverse image (section 2.3). The latter is quite technical, and may be skipped in a first reading.

Following Picard [80] and Schweinhart [91], we introduce the so-called *persistence index of degree k*, $\mathcal{L}_k(f)$, of a function $f : X \to \mathbb{R}$ (definition 3.3) as

$$\mathcal{L}_k(f) := \inf\{p \,|\, \mathrm{Pers}_p(H_k(X, f)) < \infty\}\,, \tag{2.6}$$

and show the following.

Theorem 1.5 (Generic saturation of persistence indices (theorems 3.19 and 3.23))**.** *Let X be a compact Riemannian manifold of dimension d and let $f \in C^\alpha(X, \mathbb{R})$. Then, for any $0 \leq k < d$,*

$$\mathcal{L}_k(f) \leq \frac{d}{\alpha}\,. \tag{2.7}$$

Moreover, this bound is saturated generically in the sense of Baire within $C^\alpha(X, \mathbb{R})$. ■

Modifying Schweinhart's definition for the kth degree homological dimension of X (definition 3.32), we define

$$\dim^k_{\mathrm{PH}}(X) := \sup_{f \in \mathrm{Lip}_1(X)} \mathcal{L}_k(f)\,. \tag{2.8}$$

The generic saturation theorem shows that the supremum over all 1-Lipschitz functions in the definition of $\dim^k_{\mathrm{PH}}$ can be replaced by a supremum over the class of α-Hölder class with bounded Hölder constant for any α, up to a factor of α. In so doing, we answer a question by Schweinhart [91] regarding bounds on homological dimensions and regularity conditions on X for this bound to be sharp (section 3.3).

We give a quantitative version of the Wasserstein stability theorem valid for all degrees of Čech homology on regular enough metric spaces (which in particular include compact smooth manifolds) and for a wider class of regularity than what was previously considered.

Theorem 1.6 (General Wasserstein Stability (theorem 4.13))**.** *Let X be a compact Riemannian manifold of dimension d, then for every $0 \leq k < d$, and all $p > q > \frac{d}{\alpha}$,*

$$d_p^p(H_k(X, f), H_k(X, g)) \leq C_{X,\alpha,k}(\|f\|^q_{C^\alpha} + \|g\|^q_{C^\alpha})\,\|f - g\|^{p-q}_\infty\,, \tag{2.9}$$

with

$$C_{X,\alpha,k} \leq 4^q(M^{k+1} - M^k)\alpha q \int_0^{diam(X)} \varepsilon^{\alpha q - 1}(\mathcal{N}_X(\varepsilon) \vee \mathcal{N}_X(r_C))\, d\varepsilon\,, \tag{2.10}$$

where $\mathcal{N}_X(\varepsilon)$ is the minimal covering number of X by balls of radius ε, r_C denotes the radius of convexity of X and M is the doubling constant of X. ■

Notice that we retrieve the usual stability theorem with respect to the bottleneck distance d_∞ by taking $p \to \infty$ in the above expression. We also show an annex result of stability for the trees constructed from the functions f in terms of the Gromov-Hausdorff distance between the trees (theorem 4.21).

Finally, we discuss some consequences of these results to the stochastic setting (section 5) and prove Chazal *et al.*-like results [21] for the d_p-stability of average diagrams of stochastic processes with an *a priori* hypothesis of regularity, which can in particular be used to infer distances between distributions of diagrams of close stochastic processes.

Theorem 1.7 (Stochastic Wasserstein Stability (theorem 5.9))**.** *Let f and g be two $\mathbb{R}$-valued a.s. C^α stochastic processes on a d-dimensional compact Riemannian manifold X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, for any $0 \leq k < d$, every $\frac{d}{\alpha} < q < p < \infty$ and any $r, s \in]1, \infty[$ satisfying $\frac{1}{r} + \frac{1}{s} = 1$ and $(p-q)s \geq 1$, there exists a constant C_X depending only on X such that*

$$d_p(\mathbb{E}[\mathrm{Dgm}_k(f)], \mathbb{E}[\mathrm{Dgm}_k(g)]) \leq W_{p,d_p}((\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm}_k \circ g)_\sharp \mathbb{P}) \tag{2.11}$$

$$\leq C_X \left[\mathbb{E}[\|f\|_{C^\alpha}^{qr}]^{\frac{1}{r}} + \mathbb{E}[\|g\|_{C^\alpha}^{qr}]^{\frac{1}{r}} \right]^{\frac{1}{p}} W_{(p-q)s,\infty}^{1-\frac{q}{p}}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P}) \tag{2.12}$$

$$\leq C_X \left[\mathbb{E}[\|f\|_{C^\alpha}^{qr}]^{\frac{1}{r}} + \mathbb{E}[\|g\|_{C^\alpha}^{qr}]^{\frac{1}{r}} \right]^{\frac{1}{p}} \|f-g\|_{L^{(p-q)s}(\Omega, L^\infty(X,\mathbb{R}))}^{1-\frac{q}{p}} . \tag{2.13}$$

■

2 Barcodes, diagrams and trees

2.1 Trees stemming from a continuous function

Unless otherwise specified, throughout this section, let X denote a connected, locally path-connected, compact topological space and let $f : X \to \mathbb{R}$ be a continuous function. Let us denote $(X_r)_{r\in\mathbb{R}}$ the filtration of X by the **superlevels** of f, that is

$$X_r := \{x \in X \mid f(x) \geq r\} . \tag{2.14}$$

Notation 2.1. We will denote the open superlevel sets by $X_{>r}$ whenever necessary and denote X_r^z the connected component of X_r containing z.

There exists a pseudo-distance on X, denoted d_f, given by:

Definition 2.2. Let X and f be defined as above. The H_0**-distance**, d_f, is the pseudo-distance

$$d_f(x, y) := f(x) + f(y) - 2 \sup_{\gamma : x \mapsto y} \inf_{t\in[0,1]} f(\gamma(t)) , \tag{2.15}$$

where the supremum runs over every path γ linking x to y.

Remark 2.3. Notice there are different ways of writing this distance. In particular, the sup above is also characterized by

$$\sup_{\gamma : x \mapsto y} \inf_{t\in[0,1]} f(\gamma(t)) = \sup\{r \mid [x]_{H_0(X_r)} = [y]_{H_0(X_r)}\} \tag{2.16}$$

$$= \sup\{r \mid \exists \gamma \in C_1(X_r) \text{ such that } \partial\gamma = x - y\} . \tag{2.17}$$

These equalities hold, since we take the coefficients of homology with respect to $\mathbb{Z}/2\mathbb{Z}$, so we can interpret 1-chains as sums of paths on X.

This pseudo-distance is a generalization of the distance introduced by Curien, Le Gall and Miermont in [29]. d_f has the following properties:

- **(P1) Identification of the connected components of superlevel sets**: $d_f(x, y) = 0$ if and only if there exists $t \in \mathbb{R}$ such that $x, y \in \{f = t\}$ and for every $\varepsilon > 0$, x and y lie in the same connected component of $X_{>t-\varepsilon}$;
- **(P2) Compatibility with the filtration induced by** f: Let $x, y \in X$ and suppose that $f(x) < f(y)$, then if $[x]_{H_0(X_{f(x)})} = [y]_{H_0(X_{f(x)})}$,

$$d_f(x, y) := |f(x) - f(y)| \,. \tag{2.18}$$

(P2) is immediate from the definition of d_f. It remains to show the two following propositions.

Proposition 2.4. *The function $d_f : X^2 \to \mathbb{R}^+$ of definition 2.2 is a pseudo-distance.*

Proof. Checking symmetry and positivity is easy. The only non-obvious point is that the triangle inequality is satisfied by this expression. Let $x, y, z \in X$ and denote

$$[x \mapsto y] := \sup_{\gamma : x \mapsto y} \inf_{t \in [0,1]} f \circ \gamma(t) \,. \tag{2.19}$$

It suffices to show the following inequality

$$[x \mapsto z] + [z \mapsto y] \leq [x \mapsto y] + f(z) \,. \tag{2.20}$$

Let γ be a path from x to z and η be a path from z to y and let $\gamma * \eta$ be the concatenation of these two paths. By definition,

$$\inf_{t \in [0,1]} f \circ (\gamma * \eta)(t) \leq [x \mapsto y] \,, \tag{2.21}$$

from which it follows that

$$[x \mapsto z] \wedge [z \mapsto y] \leq [x \mapsto y] \,. \tag{2.22}$$

Without loss of generality, suppose that $[x \mapsto z]$ achieves the above minimum and note that

$$[z \mapsto y] \leq f(z) \tag{2.23}$$

by definition of $[z \mapsto y]$. Adding the two last inequalities together,

$$[x \mapsto z] + [z \mapsto y] \leq [x \mapsto y] + f(z) \,, \tag{2.24}$$

as desired. ■

Proposition 2.5. *Let f be a continuous function as above, then **(P1)** holds.*

Proof. The ($\Leftarrow$) direction is immediate, so let us show ($\Rightarrow$).

Suppose that $d_f(x, y) = 0$ and that $f(x) \neq f(y)$, then,

$$\sup_{\gamma : x \mapsto y} \inf_{t \in [0,1]} f(\gamma(t)) = \frac{f(x) + f(y)}{2} > f(x) \wedge f(y) \,. \tag{2.25}$$

However,

$$\sup_{\gamma:x\mapsto y}\inf_{t\in[0,1]} f(\gamma(t)) \leq f(x)\wedge f(y)\,, \tag{2.26}$$

which leads to a contradiction, so $f(x) = f(y)$. The condition $d_f(x, y) = 0$ becomes:

$$f(x) = \sup_{\gamma:x\mapsto y}\inf_{t\in[0,1]} f(\gamma(t))\,. \tag{2.27}$$

This is only possible if for every $\varepsilon > 0$ there is a path γ lying entirely in $X^x_{>f(x)-\varepsilon}$, so

$$x, y \in \bigcap_{\varepsilon>0} X^x_{>f(x)-\varepsilon} \tag{2.28}$$

finishing the proof. ■

With these technicalities out of the way, let us consider the metric space

$$(T_f, d_f) := (X/\{d_f = 0\}, d_f)\ , \tag{2.29}$$

where $X/\{d_f = 0\}$ denotes the quotient of X where we identify all points x and y on X satisfying $d_f(x, y) = 0$. Slightly abusing the notation, let d_f denote the distance induced on T_f by the pseudo-distance d_f on X.

The metric structure of T_f turns out to be simple, as T_f is an $\mathbb{R}$-tree. Let us briefly recall the definition of an $\mathbb{R}$-tree.

Definition 2.6 (Chiswell, [26])**.** An **$\mathbb{R}$-tree** (T, d) is a connected metric space such that any of the following equivalent conditions hold:

- T is a geodesic connected metric space and there is no subset of T which is homeomorphic to the circle, $\mathbb{S}_1$;
- T is a geodesic connected metric space and the Gromov 4-point condition holds, *i.e.* :

$$\forall x, y, z, t \in T \quad d(x, y) + d(z, t) \leq \max\left[d(x, z) + d(y, t), d(x, t) + d(y, z)\right]\ ;$$

- T is a geodesic connected 0-hyperbolic space.

A **rooted $\mathbb{R}$-tree** (T, O, d) is an $\mathbb{R}$-tree along with a marked point $O \in T$.

A first important remark is that since X is connected, so is T_f. To show T_f is an $\mathbb{R}$-tree, we will use the first characterization of the definition above and show both conditions, *i.e.* that there are no subspaces of T_f which are homeomorphic to $\mathbb{S}_1$ and that T_f is in fact a geodesic metric space, to be satisfied separately.

Before showing this, it is helpful to introduce some notation.

Notation 2.7. Let $\pi_f : X \to T_f$ denote the canonical projection onto T_f and let O denote the root of T_f (*i.e.* $f(O) = \min f$), let us define the following quantity

$$\ell(\tau) := \inf_X f + d_f(O, \tau)\,, \tag{2.30}$$

where $X^{\tau}_{f(\tau)}$ denotes the connected component of the superlevel set $X_{f(\tau)}$ containing a preimage of τ.

Remark 2.8. These objects are well-defined by definition of d_f and render π_f continuous.

Definition 2.9. The **pseudo-distance topology on** X or the **topology of** d_f is the topology on X generated by the open balls:

$$B(x,r) := \{z \in X \mid d_f(x,z) < r\} \tag{2.31}$$

Despite the fact that the pseudo-distance topology is not in general Hausdorff, it is nonetheless fine enough to be useful, as shown by the two following technical lemmas.

Lemma 2.10. *$X_{>r}$ has the same connected components for the topology of d_f on X and the usual topology of X.*

Proof. By continuity of f, $X_{>r}$ is open in X for both topologies. For the usual topology, it is trivial. For the topology of d_f it is the complement in X of the closed ball $\overline{B(p, r - \inf f)}$, where p is a point on X achieving the infimum of f, which exists by compactness of X (in the usual topology).

Let Y now denote a connected component of $X_{>r}$ for the usual topology. The set Y is connected for the topology of d_f. Otherwise, we could write $Y = U \sqcup V$ for some open sets U and V, but since open sets of the topology of d_f are also open for the usual topology, this leads to a contradiction, as we assumed Y was connected for the usual topology. We will now show that Y is both open and closed in $X_{>r}$ for the topology of d_f. Y is open, since it can be written as the union of open balls

$$Y = \bigcup_{y \in Y} B(y, f(y) - r)\,. \tag{2.32}$$

Additionally, Y is closed since its complement is open, as it can be similarly written as the union of open balls. It follows that Y is also a connected component of $X_{>r}$ for the topology of d_f.

Now, suppose that Y is a connected component of $X_{>r}$ for the topology of d_f. Any ball of the covering above is path connected, but since Y is connected, this implies that Y is path connected (and the paths are completely included within Y), it is thus a path connected component of $X_{>r}$. Since X is connected and locally path connected for the usual topology, Y is a path connected component of the usual topology, rendering it a connected component for the usual topology. ∎

Lemma 2.11. *Denote $T_{>r}$ the open superlevel set on T_f. For the topology of d_f, π_f induces a bijective correspondence between the connected components of $X_{>r}$ and those of $T_{>r}$.*

Proof. Since π_f is surjective and is both open and closed for the topology of d_f (on both X and T_f), lemma 2.10 implies that the map π_f surjectively sends the connected components of $X_{>r}$ onto connected components of $T_{>r}$, since the connected components of $X_{>r}$ for the topology of d_f and the usual topology of X are the same.

It remains to show the injectivity. Note that π_f is open and closed for the topology of d_f on X. The connected components of $X_{>r}$ are either disjoint or equal and, in fact, so are the images by π_f of these connected components. Otherwise, there exists some $\tau \in T_{>r}$ such that there is a preimage of τ lying in two different connected components of $X_{>r}$, which is impossible, as every preimage of τ must lie in the same connected component of $X_{>r}$ in accordance to proposition 2.5. This is equivalent to stating that if Y and Z are two connected components of $X_{>r}$ and $Y \neq Z$, then $\pi_f(Y) \cap \pi_f(Z) = \emptyset$, in particular $\pi_f(Y) \neq \pi_f(Z)$. Symbolically,

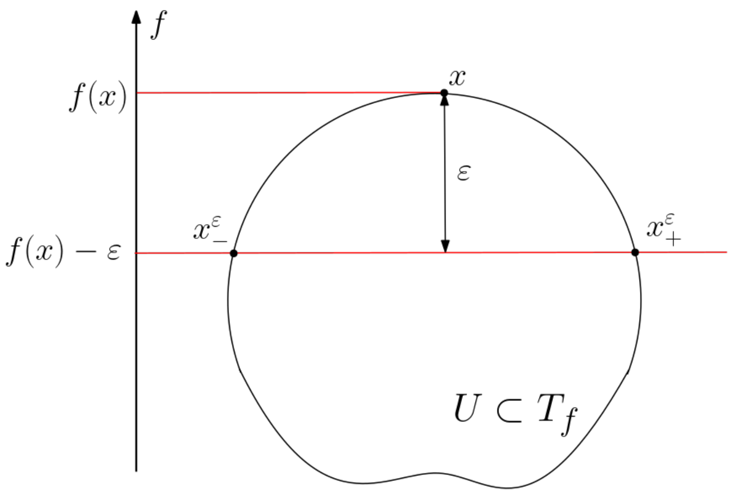
$$Y \neq Z \Rightarrow \pi_f(Y) \neq \pi_f(Z)\,, \tag{2.33}$$

which is the contrapositive of the statement of injectivity. ■

From the above lemmata, we get the following proposition.

Proposition 2.12. *The metric space $T_f := X/\{d_f = 0\}$ equipped with distance d_f possesses no subspace homeomorphic to $\mathbb{S}_1$.*

Proof. We will reason by contradiction. Suppose that T_f contains $U \subset T_f$ such that U is homeomorphic to the circle, $\mathbb{S}_1$. The function f descends to a function on T_f which is not locally constant anywhere by definition of d_f and in particular not locally constant anywhere on U, as the level-sets of f in T are totally discontinuous.



It follows that there exists an element $x \in U$ such that the maximum of f on U is attained at x. For $\varepsilon > 0$ small enough, there are two distinct points x_-^ε and x_+^ε such that $f(x_+^\varepsilon) = f(x) - \varepsilon = f(x_-^\varepsilon)$. Without loss of generality, we pick these points to be the closest ones to x along an arbitrary parametrization of U where this equality occurs. Since U is homeomorphic to $\mathbb{S}_1$, there is a path γ linking x_+ and x_- lying entirely above $f(x) - \varepsilon$ (equal at the endpoints) and passing through x. The image of γ in $T_{>f(x)-\varepsilon}$ is contained within one and only one connected component of $T_{>f(x)-\varepsilon}$, which we will denote S. By lemma 2.11, S corresponds to a unique connected component of $X_{>f(x)-\varepsilon}$ with respect to the topology of d_f, which we will denote X^S. By lemma 2.10, X^S is a connected component of $X_{>f(x)-\varepsilon}$ for the usual topology.

For every $0 < \varepsilon' < \varepsilon$, we can pick points $x_{\pm}^{\varepsilon'}$ on U. The connected component S contains $x_{\pm}^{\varepsilon'}$ for every such ε' and since inverse images of these two points are connected in $X_{>f(x)-\varepsilon}$ by a path $\gamma : x_+^{\varepsilon'} \mapsto x_-^{\varepsilon'}$,

$$d_f(x_+^{\varepsilon'}, x_-^{\varepsilon'}) < 2(f(x) - \varepsilon') - 2 \inf_{t \in [0,1]} f \circ \gamma < 2(\varepsilon - \varepsilon') \tag{2.34}$$

Letting $x_{\pm}^{\varepsilon'} \to x_{\pm}^{\varepsilon}$ in U as $\varepsilon' \to \varepsilon$, we have that $d_f(x_-^{\varepsilon}, x_+^{\varepsilon}) = 0$, leading to a contradiction, since we supposed that x_+^{ε} and x_-^{ε} were disjoint in T_f (and therefore not a distance zero away from one-another). ∎

Proposition 2.13. *The metric space (T_f, d_f) is a rooted $\mathbb{R}$-tree whose root is the unique point point O in the image in T_f of a point $x \in X$ for which the function f is minimal.*

Proof. The only thing left to show is that T_f is a geodesic space. As before, f descends to the quotient and induces a non-locally constant function on T_f. Let $x, y \in X$, if $f(x) = f(y)$ and x and y are in the same path connected component of $X_{>f(x)-\varepsilon}$ for all $\varepsilon > 0$ and there is nothing to show.

Suppose that $f(x) < f(y)$ and that x and y are in the same path connected component of $X_{f(x)}$ and consider a path in $X_{f(x)}$ going from y to x, $\gamma : [0,1] \to X$. The path γ can be modified into a path

$$\tilde{\gamma}(t) := \pi_f\left(\gamma\left(\arg\min_{s \in [0,t]} f \circ \gamma(s)\right)\right) . \tag{2.35}$$

On this modified path f is decreasing implying that $\tilde{\gamma}$ does not self-intersect, although it may be locally constant. The length of $\tilde{\gamma}$ is defined as

$$L(\tilde{\gamma}) = \sup_{(t_i)} \sum_{(t_i)} d_f(\tilde{\gamma}(t_{i+1}), \tilde{\gamma}(t_i)) , \tag{2.36}$$

where the supremum is taken over all finite partitions of $[0,1]$. For any finite partition, this sum is always bounded by $f(y) - f(x)$, since along $\tilde{\gamma}$

$$f(\tilde{\gamma}(t_i)) \geq f(\tilde{\gamma}(t_{i+1})) \implies d_f(\tilde{\gamma}(t_{i+1}), \tilde{\gamma}(t_i)) = f(\tilde{\gamma}(t_i)) - f(\tilde{\gamma}(t_{i+1})) \tag{2.37}$$

by monotonicity of f along $\tilde{\gamma}$. This leads to pairwise cancellation of terms in the sum of equation 2.36. And so,

$$L(\tilde{\gamma}) = d_f(x, y) . \tag{2.38}$$

Now, suppose that x and y are two points on X, such that $f(x) \leq f(y)$ but such that x and y no longer lie in the same path connected component of $X_{f(x)}$ and pick a maximizer γ of the supremum (*cf.* remark 2.14)

$$\sup_{\gamma : x \mapsto y} \inf_{t \in [0,1]} f \circ \gamma(t) . \tag{2.39}$$

Since y is not connected to x in $X_{f(x)}$, by continuity of f, the path γ must eventually go under the level $f(x)$. Let us set

$$t^* := \sup \left\{ \arg\min_{s\in[0,1]} f \circ \gamma(s) \right\} \tag{2.40}$$

and note that $f(\gamma(t^*)) < f(x)$.

On $[0, t^*]$, the path γ lies entirely in $X_{f(\gamma(t^*))}$ and similarly, entirely in $X_{f(\gamma(t^*))}$ on $[t^*, 1]$. On $[0, t^*]$, we can define a modification of γ, $\tilde{\gamma} : [0, t^*] \to T_f$ by

$$\tilde{\gamma}(t) := \pi_f\left(\gamma\left(\arg\min_{s\in[0,t]} f \circ \gamma(s)\right)\right) . \tag{2.41}$$

Analogously, if we define $\eta(t) := \gamma(1-t)$ – the reversed version of γ – it is possible to define a modification of η, $\tilde{\eta} : [0, 1-t^*] \to T_f$, by

$$\tilde{\eta}(s) := \pi_f\left(\eta\left(\arg\min_{r\in[0,s]} f \circ \eta(r)\right)\right) . \tag{2.42}$$

In particular, $\tilde{\eta}(1-t^*) = \tilde{\gamma}(t^*)$. If $\tilde{\eta}_-$ denotes the reversed path along $\tilde{\eta}$, the concatenation (without reparametrization),

$$\zeta := \tilde{\gamma} * \tilde{\eta}_- \tag{2.43}$$

is a path going from $\pi_f(x)$ to $\pi_f(y)$ monotone decreasing on $[0, t^*]$ and monotone increasing on $]t^*, 1]$.

For all $\varepsilon > 0$, $\zeta(t^* + \varepsilon)$ does not lie in the same connected component of $X_{f(\zeta(t^*+\varepsilon))}$ as $\pi_f(x)$, but lies in the same connected component of $X_{f(\zeta(t^*+\varepsilon))}$ as $\pi_f(y)$. We are thus reduced to examine the length of the path along two different sections of ζ, each lying in the same connected component as either $\pi_f(x)$ and $\pi_f(y)$. By the previous argument for points of T_f lying in the same connected component of a superlevel set, the length of ζ is

$$L(\zeta) = f(x) - f(\zeta(t^*)) + f(y) - f(\zeta(t^*)) = d_f(x, y) \tag{2.44}$$

by definition of $d_f(x, y)$. Thus, T_f is indeed geodesic and it is an $\mathbb{R}$-tree, by virtue of proposition 2.12.

Finally, the tree is rooted since for any $r < \inf f$, every single point of $X_r = X$ is identified in the quotient (since X was supposed to be connected), so we can identify the root with the point of T_f achieving this infimum. ■

Remark 2.14. If the supremum in the proof of proposition 2.13 is not achieved, it is still possible to construct a geodesic path in T_f. Let us denote

$$r := \sup_{\gamma: x\mapsto y} \inf_{t\in[0,1]} f \circ \gamma \tag{2.45}$$

and X_r^x and X_r^y the path connected components of of $X_r = \{f \geq r\}$ containing x and y respectively. The sets $X_r^x \cap \{f = r\}$ and $X_r^y \cap \{f = r\}$ are not empty, since we can always find an element in these sets by taking the first and last instances where any path $\gamma : x \mapsto y$ hits the level set $\{f = r\}$. On the T_f we have

$$\pi_f(X_r^x \cap \{f = r\}) = \pi_f(X_r^y \cap \{f = r\})\,. \tag{2.46}$$

To see this, consider $z \in X_r^x \cap \{f = r\}$ and $z' \in X_r^y \cap \{f = r\}$, then there exist two paths $\eta : x \mapsto z$ and $\eta' : y \mapsto z'$ which lie entirely above r. Furthermore, for any $\varepsilon > 0$, there exists a path γ_ε linking x and y whose image lies entirely in $X_{>r-\varepsilon}$. The concatenation $\eta * \gamma_\varepsilon * \overline{\eta}'$ yields a path whose image lies entirely in $X_{>r-\varepsilon}$ which links z to z'. We conclude that $\pi_f(z) = \pi_f(z')$, since these points are at zero d_f-distance away from one another. The geodesic path in T_f can be found by taking a path γ linking x and y in X, stopping γ as soon as it hits an element of $X_r^x \cap \{f = r\}$ and resume following it at the last instance where γ intersects $X_r^y \cap \{f = r\}$. Taking the images of this stopped version of γ as per the construction of the proof above yields a geodesic path in T_f linking $\pi_f(x)$ and $\pi_f(y)$.

Remark 2.15. If $X = [0,1]$, there is only one possible path between any two points $x < y$, so the definition above boils down to

$$d_f(x,y) := f(x) + f(y) - 2 \inf_{t\in[x,y]} f\,, \tag{2.47}$$

which is exactly the distance originally introduced by Le Gall *et al.* [37].

2.2 From trees to barcodes

Given a tree stemming from a continuous function $f : X \to \mathbb{R}$, it is possible to reconstruct the H_0-barcode of f from T_f. If T_f has a finite number of leaves, the relation between the barcode of $H_0(X,f)$ with respect to the superlevel filtration and the tree T_f is given by algorithm 2.

Algorithm 2: A functorial relation between persistence modules and $\mathbb{R}$-trees

```
Result: 𝕍
ℱ ← T ;
𝕍 ← 0 ;
i ← 0 ;
while ℱ ≠ ∅ do
    Find γ the longest path in ℱ starting from a root α and ending in a leaf β ;
    if i = 0 then
        𝕍 ← 𝕍 ⊕ k[ℓ(α), ∞[ ;
    else
        𝕍 ← 𝕍 ⊕ k[ℓ(α), ℓ(β)[ ;
    end
    ℱ ← closure(ℱ \ Im(γ));
    i ← i + 1 ;
end
return 𝕍
```

Definition 2.16. We say T_f is finite if it has a finite number of leaves and say it is infinite otherwise.

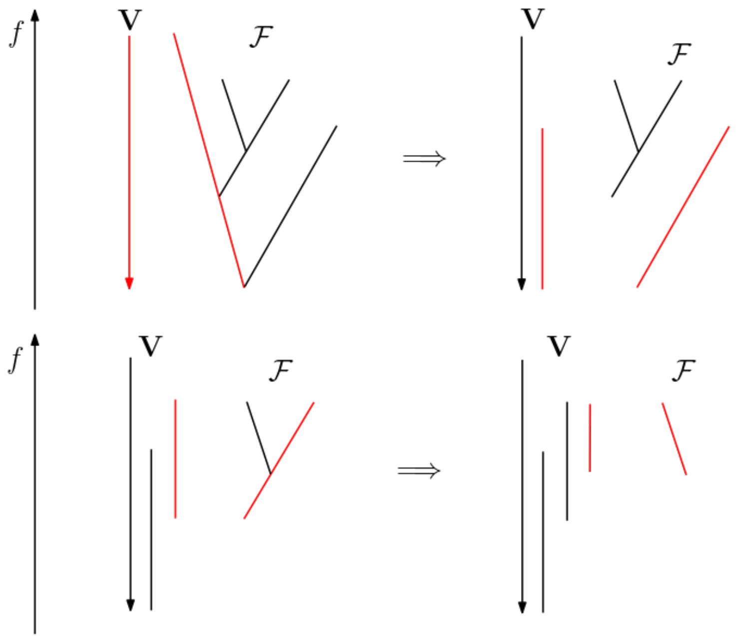


Figure 2.1: The first four iterations of algorithm 2. For every step, in red is the longest branch of the tree, which we use to progressively construct the persistent module $\mathbb{V}$ by associating an interval module whose ends correspond exactly to the values of the endpoints of the branches.

If T_f is infinite, we can still give a correspondence between the barcode and the tree proceeding by approximation. This approximation procedure requires the introduction of so-called ε-trimmings of T_f, of which we briefly recall the definition. Since the results of this section can be easily extended to any compact tree, we formulate the rest of this section in full generality.

For any rooted $\mathbb{R}$-tree (T, d, O), we can define a filtering function $\ell : T \to \mathbb{R}$ by setting

$$\ell(\tau) := d(O, \tau)\,. \tag{2.48}$$

This allows us to define the height above a point τ as follows.

Definition 2.17. The function of **the height above** τ on a rooted $\mathbb{R}$-tree T is a function $h : T \mapsto \mathbb{R}$, defined as

$$h(\tau) := \sup_{\eta \in T^{\tau}_{\ell(\tau)}} d(O, \eta) - \ell(\tau)\,, \tag{2.49}$$

The height above τ allows us to define so-called ε-*trimmings* or ε-*simplifications* of T.

Definition 2.18. The ε**-simplified tree of** T**,** T^{ε} or the ε**-trimmed tree of** T, is the subtree of T defined as

$$T^{\varepsilon} := \{\tau \in T \mid h(\tau) \geq \varepsilon\} \tag{2.50}$$

Provided T is compact, its ε-trimmings are always finite. For a monotone decreasing sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ such that $\varepsilon_n \to 0$, we have the following chain of inclusions

$$T^{\varepsilon_1} \hookrightarrow T^{\varepsilon_2} \hookrightarrow T^{\varepsilon_3} \hookrightarrow \cdots\,. \tag{2.51}$$

Applying algorithm 2, we get a set of maps on the persistence modules induced by these inclusions. More precisely, denoting $\mathrm{Alg}(T^{\varepsilon_n})$ the output of the algorithm

$$\mathrm{Alg}(T^{\varepsilon_1}) \to \mathrm{Alg}(T^{\varepsilon_2}) \to \mathrm{Alg}(T^{\varepsilon_3}) \to \cdots . \tag{2.52}$$

where the morphisms are the maps induced at the level of the interval modules generating $\mathrm{Alg}(T^{\varepsilon_n})$. The interval modules $k[\alpha, \beta_n[$ of $\mathrm{Alg}(T^{\varepsilon_n})$ satisfy that there is exactly one interval module of $\mathrm{Alg}(T^{\varepsilon_m})$ $(m > n)$ such that $[\alpha, \beta_n[\subset [\alpha, \beta_m[$. A natural definition for infinite T is thus

$$\mathrm{Alg}(T) := \varinjlim \mathrm{Alg}(T^{\varepsilon_n}) . \tag{2.53}$$

In categoric terms, the algorithm above in fact is a functor

$$\mathrm{Alg} : \mathbf{Tree} \to \mathbf{PersMod_k} , \tag{2.54}$$

where **Tree** is the category of rooted $\mathbb{R}$-trees seen as metric spaces, whose morphisms are isometric embeddings (which are not required to be surjective) preserving the roots, and where $\mathbf{PersMod_k}$ is the category of q-tame persistence modules over a field k (*cf.* Oudot's book for details on the category of persistence modules [76]). The action of Alg on morphisms between two trees $\zeta : T \to T'$ is defined as follows. Let $\bigoplus_i k[\alpha_i, \beta_i[$ be the interval module decomposition stemming from $\mathrm{Alg}(T)$. By construction, the intervals $[\alpha_i, \beta_i[$ are in bijective correspondence with branches of T. Since both T and T' are finite and since ζ is an isometric embedding and it preserves the root, we can look at the intesection $\mathrm{Im}(\zeta) \cap [\alpha_i, \beta_i[$, where we somewhat abuse the notation, by regarding the intervals as embedded in T'. Somewhat abusing the notation once again to look at these intersections as simple intervals, we define

$$\mathrm{Alg}(\zeta) := \bigoplus_i \mathrm{id}_{k[\alpha_i,\beta_i[\cap \mathrm{Im}(\zeta)} . \tag{2.55}$$

If T is infinite, we extend the above definition by taking successive ε_n-simplifications of T and taking the direct limit of the construction above. This procedure is well-defined since ε_n-simplifications only depend on the function h, which in turn can be taken to only depend on the distance to the root.

Trees stemming from a function

Let us now consider a tree T_f stemming from a function f and show that $\mathrm{Alg}(T_f) = H_0(X, f)$.

Proposition 2.19. *Let τ and η be elements of T_f such that $f(\tau) < f(\eta)$ and let $x \in \pi_f^{-1}(\tau)$ and $y \in \pi_f^{-1}(\eta)$, then*

$$\forall \varepsilon > 0, \exists \text{ path } \gamma : x \mapsto y \text{ s.t. } \forall t,\ f(\gamma(t)) > f(\tau) - \varepsilon \iff h(\tau) \geq f(\eta) - f(\tau) \text{ and } \forall \varepsilon > 0\, x, y \in X^{\tau}_{>f(\tau)-\varepsilon} . \tag{2.56}$$

Proof. Since there exists γ connecting x and y and since γ always stays above $f(\tau) - \varepsilon$ for all $\varepsilon > 0$, we conclude naturally that $\mathrm{Im}(\gamma) \subset X^{\tau}_{>f(\tau)-\varepsilon}$, which implies that for all $\varepsilon > 0$,

$h(\tau) \geq f(\eta) - f(\tau) + \varepsilon$ by definition of $h(\tau)$.

The implication ($\Leftarrow$) is clear since if for all $\varepsilon > 0$, $x, y \in X^{\tau}_{>f(\tau)-\varepsilon}$ and since $X^{\tau}_{>f(\tau)-\varepsilon}$ is connected, by path connectedness of X there exists a path between x and y which stays above $f(\tau) - \varepsilon$ for all $\varepsilon > 0$. ■

This proposition suffices to prove the following theorem on the validity of algorithm 2.

Theorem 2.20. *Let X be a compact, connected, locally path connected topological space and let $f : X \to \mathbb{R}$ be continuous. Then $\mathrm{Alg}(T_f) = H_0(X, f)$ in the observable category of persistence modules (cf. [23]).*

Remark 2.21. This theorem is a slight improvement on the result of Curry in [30, Theorem §2.13]. In the language of [30], this constitutes a proof of the "Elder rule" with less assumptions of regularity. Indeed, in [30], the assumption of a Morse set (or that f is a Morse function) is necessary for the proof, whereas the functions hereby considered are merely required to be continuous.

Remark 2.22. By setting $X = T$ and $\ell = f$, theorem 2.20 states that $\mathrm{Alg}(T) = H_0(T, \ell)$.

Proof. Suppose that T_f is finite, then $\mathrm{Alg}(T_f)$ is a decomposable persistence module $\mathrm{Alg}(T_f) := \mathbb{V}$. The fact that $\mathbb{V}$ is pointwise isomorphic to $H_0(X, f)$ (after Serre localization, *i.e.* up to evanescent modules of the form $k[\alpha, \alpha]$) holds since d_f correctly identifies the connected components of the (open) superlevel sets. This guarantees the existence of a pointwise isomorphism since both spaces have the same (finite) dimension.

Let us now check that $\mathrm{rank}(\mathbb{V}(r \to s)) = \mathrm{rank}(H_0(X_r \to X_s))$. The inclusion $X_r \hookrightarrow X_s$ induces the following long exact sequence in homology

$$
\begin{array}{ccccccc}
\cdots \longrightarrow H_1(X_s) \longrightarrow H_1(X_s, X_r) \longrightarrow & H_0(X_r) & & & & \\
& \downarrow & & & & \\
& H_0(X_s) & \longrightarrow & H_0(X_s, X_r) & \longrightarrow & 0
\end{array}
$$

Since this sequence is exact

$$
\mathrm{rank}(H_0(X_r \to X_s)) = \dim\ker(H_0(X_s) \to H_0(X_s, X_r))\,. \tag{2.57}
$$

For notational simplicity, let us denote $\phi : H_0(X_s) \to H_0(X_s, X_r)$. Note that $\phi[c] = [0]$ if and only if there is a path γ between the representative $c \in X_s$ and an element $b \in X_r$ such that γ stays within X_s. Without loss of generality, let us take c such that $c \in \{f = s\}$. Finding such a path γ is only possible if c and b lie in the same connected component of X_r. By proposition 2.19, this can happen if and only if $h([c]_{T_f}) \geq r - s$. It follows that

$$
\dim\ker\phi = \#\{\tau \in T_f \,|\, h(\tau) \geq r - s \ \text{ and } \ f(\tau) = s\}\,, \tag{2.58}
$$

which concludes the proof for the finite case.

If T_f is infinite, we consider a sequence of ε_n-trimmings of T_f such that $\varepsilon_n \xrightarrow[n\to\infty]{} 0$. For any $r > s$, there exists n such that $r - s > \varepsilon_n$. But $T_f^{\varepsilon_n}$ is finite, so we are reduced to the previous case. ■

2.3 The inverse problem

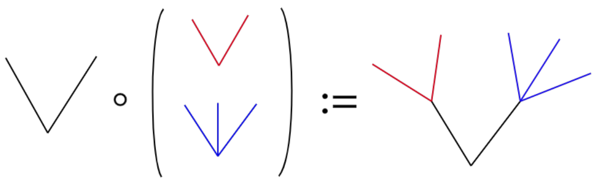
An interesting question is whether every (compact) tree stems from a function $f : X \to \mathbb{R}$. If the tree is a so-called *merge tree* (in particular, we require that it be locally finite and 1-dimensional), a solution has been provided by Curry in [30, §6]. We will now positively answer this question under the assumptions that $\overline{\dim}\, T < \infty$ and that $X = [0, 1]$ by constructing a function $f : [0, 1] \to \mathbb{R}$, which constitute a wider class of trees than merge trees. The rest of this section will focus on proving the following theorem:

Theorem 2.23. *Let T be a compact $\mathbb{R}$-tree such that $\overline{\dim}\, T < \infty$. Then, for any $\delta > 0$ it is possible to construct a continuous function $f : [0, 1] \to \mathbb{R}$ of finite $(\overline{\dim}\, T + \delta)$-variation such that $T = T_f$. In particular, up to a reparametrization, f can be taken to be $\frac{1}{\overline{\dim}\, T+\delta}$-Hölder continuous.*

The idea is to once again use ε-simplifications T^ε for which we can construct a function by taking the contour of the tree. Such a construction is referred to as the Dyck path in the terminology of [94].

Finite trees

We can regard a rooted discrete tree as being an operator with N inputs, where N is the number of leaves of the tree. There is a natural operation on the space of discrete trees which composes these operations by:



These objects are called **operads** and originated in the study of iterated loop spaces [12, 13, 67]. Since then, these objects have been studied in different fields for a variety of purposes [46, 63]. We will not give the explicit definition of an operad here, as a rigorous introduction is unnecessary for our purposes. However, we introduce this notion of composition of trees for notational simplicity.

Given a discrete $\mathbb{R}$-tree T, if we have an embedding of T in $\mathbb{R}^2$, or equivalently, a partial order on its vertices, we can assign to T an interval I of a certain length with N marked points as well as a function $f_T : I \to \mathbb{R}$, where N is the number of leaves of T. Using the terminology of [94], a way to do this is by considering the so-called **Dyck path** or **contour path** where the path around T parametrized by arclength in T. The construction of the Dyck path has been

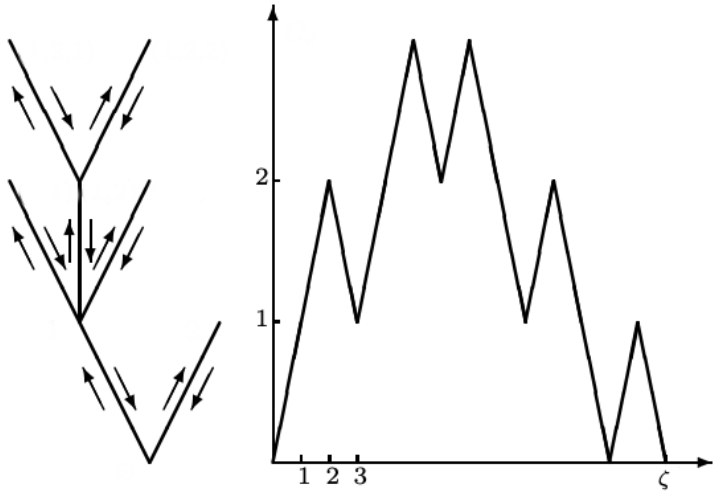


Figure 2.2: The Dyck path is the function f which assigns the height (the distance from the root) of each vertex of the tree as we wrap around the tree following a clockwise contour around it. There is a map $\phi : T \mapsto [0, \zeta]$ where $[0, \zeta]$ is now marked at the points at which f achieves its local maxima. The figure is taken from [37].

carefully detailed in [37, 94], but it is better understood by looking at figure 2.2. By construction the equality: $T_{f_T} = T$ holds for any discrete $\mathbb{R}$-tree T. Here, equality is taken up to isometry.

As per the description of figure 2.2, the construction of the Dyck path yields a map ϕ which to T assigns an interval $\phi(T)$ with N marked points. An example of the action of ϕ is illustrated in figure 2.3.

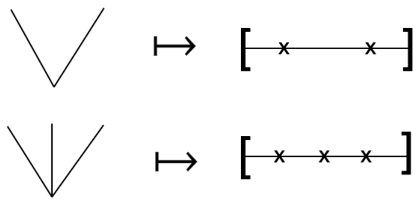


Figure 2.3: The action of ϕ on trees with two and three leaves respectively. The length of the intervals assigned is exactly the length of the contour around the trees and the marked points are the points at which f_T achieves its maxima.

This operation ϕ is in fact a "morphism" with respect to a composition operation on the intervals, defined as follows. If we have an interval I with N marked points and N intervals J_k each with M_j marked points, the result of the operation $I \circ (J_1, \cdots, J_N)$ is the insertion of the marked interval J_k at the kth marked point of I. The length of $I \circ (J_1, \cdots, J_N)$ is

$$|I \circ (J_1, \cdots, J_N)| = |I| + \sum_{k=1}^{n} |J_k| \ , \tag{2.59}$$

where $|\cdot|$ denotes the lengths of the intervals. The fact that ϕ is a "morphism" results from the definitions of compositions for trees and intervals. We can also define a variant of this morphism ϕ, which we will call ϕ_λ, which for any tree T simply scales the (marked) interval $\phi(T)$ by a factor λ.

Given a tree T the Dyck path $f_T : \phi(T) \to \mathbb{R}$ can be transformed into a function $f_T^\lambda : \phi_\lambda(T) \to \mathbb{R}$ by setting

$$f_T^\lambda(x) := f_T(x/\lambda)\,. \tag{2.60}$$

This is a rescaling of the x-axis which means that $T_{f_T^\lambda} = T_{f_T} = T$ still holds. Once again, these equalities are taken up to isometry.

Remark 2.24. The definition of f_T^λ is readily generalizable to forests. If $\mathcal{F}$ denotes a forest, then we define $f_\mathcal{F}^\lambda = \bigsqcup_{T\in\mathcal{F}} f_T^\lambda$.

For discrete trees, there is an upper bound of the number of vertices of the tree given its number of leaves.

Lemma 2.25. *Let T be a rooted discrete tree, $N \geq 2$ be its number of leaves and V be its number of vertices, then*

$$V \leq 2N - 1\,. \tag{2.61}$$

In particular, if the edges of T all have length 1, the contour of the tree can be done over an interval of length at most $4N - 2$

Proof. For binary trees, it is known that [37, 94]

$$V = 2N - 1\,. \tag{2.62}$$

Given a tree with N leaves, we can obtain a binary tree with N leaves by blowing up the vertices which are non-binary. The inequality of the lemma follows. On a binary tree, the Dyck path passes through almost every point in T twice, so the length of the interval is exactly $4N - 2$. Since binary trees are the extremal case, a bound for all trees with N leaves follows. ■

The results above show the result of theorem 2.23 for finite trees, since their upper-box dimension is equal to 1.

Infinite trees

The concatenation of trees can be defined for $\mathbb{R}$-trees too in the obvious way. Given an infinite number of compositions, we can define a limit tree by defining it to be the limit of the partial compositions in the Gromov-Hausdorff sense. Ideally, we would like to have an equality of the following type

$$T = T^a \circ \overline{(T \setminus T^a)}\,, \tag{2.63}$$

where $T \setminus T^a$ now denotes the rooted forest corresponding to the set $T \setminus T^a$. This equality is desirable because by taking infinitely many compositions, we can eventually recover the original tree T, by composing successive ε_n-simplifications with each other. However, this equality does not hold since T^a might not have the right amount of leaves for this operation to be well-defined. Nonetheless, we can decide to count the vertices $T^a \cap \overline{(T \setminus T^a)}$ as leaves with multiplicity, so that the equality above holds.

For an infinite compact tree with $\overline{\dim}\, T < \infty$, the idea is to take some appropriate rapidly decreasing (monotonous) sequence $(\varepsilon_n)_{n\in\mathbb{N}^*}$ such that the interval

$$I = \phi_{\varepsilon_1}(T^{\varepsilon_1}) \circ \phi_{\varepsilon_2}(T^{\varepsilon_2} \setminus T^{\varepsilon_1}) \circ \phi_{\varepsilon_3}(T^{\varepsilon_3} \setminus T^{\varepsilon_2}) \circ \cdots \tag{2.64}$$

has finite length. On each $\phi_{\varepsilon_k}(T^{\varepsilon_k}\setminus T^{\varepsilon_{k-1}})$ we can consider the Dyck path on the forest $T^{\varepsilon_k}\setminus T^{\varepsilon_{k-1}}$. Defining a correct superposition of these Dyck paths, we would be done (*cf.* figure 2.4).

For an infinite tree, it suffices to show that the sequence generated by the procedure of figure 2.4 converges in the Gromov-Hausdorff sense to an interval of finite length I and that $(f_i)_i$ converge in $L^\infty(I)$ to some function f.

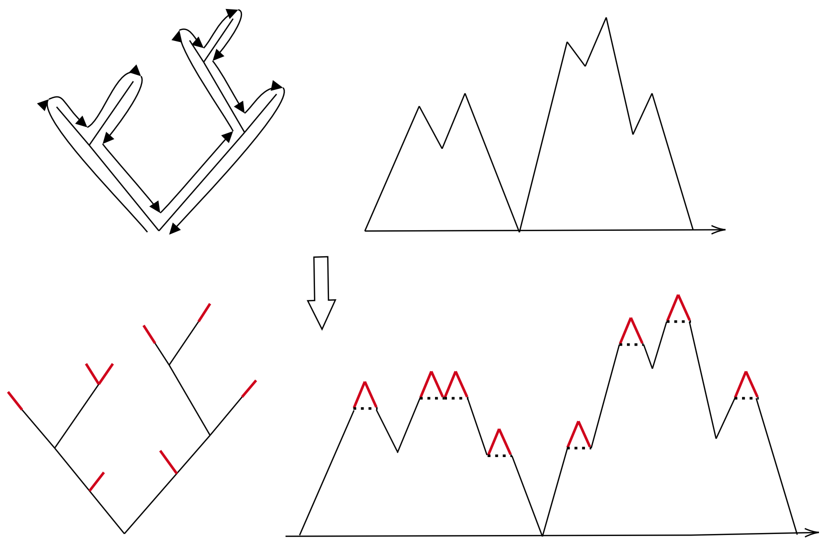


Figure 2.4: Starting from a tree $T^{a/2^k}$ (black) we construct the Dyck path around it in the first step. Then, we look at $T^{a/2^{k+1}}$ which leads to the addition of intervals (dotted), and a correction of the function at the kth step f_k (which is the function depicted in black, extended linearly over the new intervals). We can further define a function by pasting the Dyck paths of the forest over the corresponding leaves, which leads to the function depicted in the second step (red and black).

Detailed construction of the approximants

Definition 2.26. Let $I \subset \mathbb{R}_+$ be a marked interval with n marked points, which we will denote $(i_k)_{\{1\leq k\leq n\}}$. Furthermore, let $(J_k)_{\{1\leq k\leq n\}}$ be a set of n marked intervals of $\mathbb{R}_+$, each with j_k marked points. Define $\sigma_I : I \to I \circ (J_1, \cdots J_n)$ by

$$\sigma_I(x; J_1, \cdots, J_n) := \left[x + \sum_{i=1}^{\arg\max_k \{i_k < x\}} |J_i| \right] \in I \circ (J_1, \cdots, J_n)\,. \tag{2.65}$$

This definition naturally extends to the whole interval I and we can define the image $\sigma_I(I; J_1 \cdots J_n)$, which is diagramatically represented in figure 2.5.

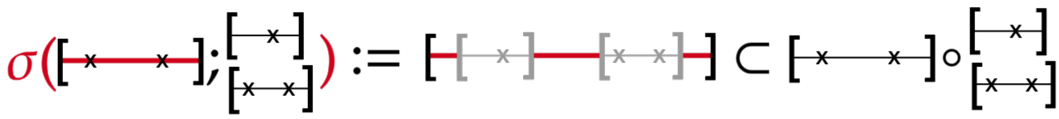


Figure 2.5: $\sigma_I(I; J_1 \cdots J_n)$

Remark 2.27. Fixing $J_1, \cdots, J_n$, σ_I is a bijective map onto its image, meaning every point $y \in \sigma_I(I; J_1, \cdots, J_n)$ admits a preimage in I, which we will denote by $\sigma_I^{-1}(y; J_1, \cdots, J_n)$.

Definition 2.28. Let $f : I \to \mathbb{R}$ be a continuous function from an interval I with n marked points and let $(J_1, \cdots, J_n)$ be intervals with each with j_i marked points as before. Abusing the notation, we define another function $\sigma(-; J_1, \cdots, J_n)$ which assigns a function on I to a function on $\sigma_I(I; J_1, \cdots, J_n)$ via the following formula

$$\sigma_I(f; J_1, \cdots, J_n)(x) := \begin{cases} f(\sigma_I^{-1}(x; J_1, \cdots, J_n)) & x \in \sigma_I(I; J_1, \cdots, J_n) \\ \text{Linearly extend elsewhere} \end{cases} \tag{2.66}$$

Remark 2.29. By continuity of $f : I \to \mathbb{R}$, this linear extension on $I \circ (J_1, \cdots, J_n)$ is in fact constant everywhere outside $\sigma_I(I; J_1, \cdots, J_n)$ (this is the dotted region in figure 2.4). Note also that $\sigma_I(f; J_1, \cdots, J_n)$ is continuous.

Definition 2.30. Given a tree T_f associated to a continuous function $f : I \to \mathbb{R}$, defined on an interval $I \subset \mathbb{R}$. Let $\tau \in T_f$, define the **left preimage of τ,** $\overleftarrow{\tau}$ and the **right preimage of τ by π,** $\overrightarrow{\tau}$ as

$$\overleftarrow{\tau} := \inf \pi^{-1}(\tau) \tag{2.67}$$

$$\overrightarrow{\tau} := \sup \pi^{-1}(\tau)\,. \tag{2.68}$$

Definition 2.31. Let T be a discrete rooted tree and $T' \subset T$ be a subtree sharing roots with T and suppose that we have chosen some embedding of T. Suppose there is a function $f : I \to \mathbb{R}$ on a certain interval I such that $T_f = T'$. Then, the marking of I induced by T is the marking induced by marking the preimage $\pi_f^{-1}(T' \cap \overline{(T \setminus T')})$ chosen in the following way:

- If $\tau \in T' \cap \overline{(T \setminus T')}$ admits a single preimage, choose this preimage;
- Else, if the connected component of τ in $\overline{T \setminus T'}$ is smaller (with respect to the partial order on the tree induced by the embedding of T) than every vertex strictly greater than $\tau \in T'$, choose $\overleftarrow{\tau}$. Otherwise, choose $\overrightarrow{\tau}$. In simpler terms, we choose $\overrightarrow{\tau}$ or $\overleftarrow{\tau}$ depending on whether the subtree of $\overline{T \setminus T'}$ containing τ branches to the right or to the left respectively of T', with the convention that we say that it branches to the left if lies at the top of a leaf of T' (*cf.* figure 2.6).

We will denote this marking operation by $\mu(I; T', T, f)$.

We can also define analogous maps to σ_I, but this time on the intervals J_k as follows.

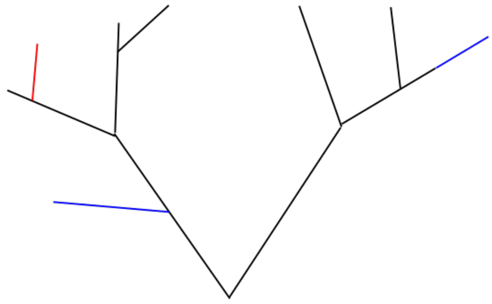


Figure 2.6: A tree T embedded in $\mathbb{R}^2$ with a subtree T' in black, the subtrees highlighted in red branch to the right and those in blue to the left.

Definition 2.32. Let $I \subset \mathbb{R}_+$ be a marked interval with n marked points, which we will denote $(i_k)_{\{1 \le k \le n\}}$. Furthermore, let $(J_k)_{\{1 \le k \le n\}}$ be a set of n marked intervals of $\mathbb{R}_+$, each with j_k marked points. Define $\eta_I^{J_k} : J_k \to I \circ (J_1, \cdots, J_n)$ by

$$\eta_I^{J_k}(x; J_1, \cdots, J_n) := x + i_k + \sum_{j=1}^{k-1} |J_j| \ . \tag{2.69}$$

These maps define a map $\eta_I = \bigsqcup_k \eta_I^{J_k}$ on $\bigsqcup_k J_k$ and η_I also induces a map on the functions $f : \bigsqcup_k J_k \to \mathbb{R}$, defined analogously to σ_I, which we shall also denote η_I.

With this notation, the construction is made in accordance to algorithm 3. A depiction of the mechanism of algorithm 3 can be found in figure 2.4. For an infinite tree, it suffices to show

Algorithm 3: Construction of approximants

Output: A set of unions of intervals $(I_i)_{i \in \{1, \cdots, n\}}$ and a set of functions on I_n, $(f_i : I_n \to \mathbb{R})_{i \in \{1, \cdots, n\}}$
Input: An infinite tree T and $a > 0$.
$I_1 \leftarrow \phi(T^a)$;
$f_1 \leftarrow f_{T^a}$;
$I \leftarrow I_1$;
$i \leftarrow 1$;
while $i \le n$ **do**
  $I_{i+1} := I_i \circ \phi_{\lambda^i}(\overline{T^{a/2^{i+1}} \setminus T^{a/2^i}})$;
  $f \leftarrow \eta_{I_{i+1}}(f^{\lambda^{i+1}}_{T^{a/2^{i+1}} \setminus T^{a/2^i}}; I_1, \cdots I_i)$;
  $I_i \leftarrow \mu(I_i; T^{a/2^{i-1}}, T^{a/2^i}, f_i)$;
  for *j=1; $j \le i$* **do**
    $I_j \leftarrow \sigma(I_j; \phi_{\lambda^i}(\overline{T^{a/2^{i+1}} \setminus T^{a/2^i}}))$;
    $f_j \leftarrow \sigma(f_j; \phi_{\lambda^i}(\overline{T^{a/2^{i+1}} \setminus T^{a/2^i}}))$;
    $j \leftarrow j + 1$;
  end
  $f_{i+1} := f_i + f$;
  $i \leftarrow i + 1$;
end
return $(I_i)_{i \in \{1, \cdots, n\}}$, $(f_i)_{i \in \{1, \cdots, n\}}$.

that the sequence generated by this algorithm converges in the Gromov-Hausdorff sense to an interval of finite length I and that $(f_i)_i$ converge in $L^\infty(I)$ to some function f.

End of the proof

To get the desired convergence we must show the two following lemmata.

Lemma 2.33. *If T is a compact $\mathbb{R}$-tree of finite upper-box dimension, there exist a and λ such that I defined by the construction above has finite length.*

We need to show the convergence of the corresponding functions $(f_n)_n$. This can be done by proving that the sequence is Cauchy.

Lemma 2.34. *Given the definition of functions f_n above, then the sequence $(f_n)_{n\in\mathbb{N}^*}$ is Cauchy in $C^0(I)$, we have*

$$\|f_n - f_m\|_{C^0} \leq a2^{-(n\wedge m)} \tag{2.70}$$

for any n and $m \in \mathbb{N}^*$.

By completeness of C^0, the sequence $(f_n)_{n\in\mathbb{N}^*}$ uniformly converges to a continuous function f. By virtue of stability theorem for trees (theorem 4.21) it follows that T is isometric to T_f. Using Picard's theorem (theorem 3.6)

$$\mathcal{V}(f) = \overline{\dim}\, T_f = \overline{\dim}\, T \tag{2.71}$$

which concludes the proof of theorem 2.23.

Proof of lemma 2.33. Recall that, according to the proof of theorem 3.9, the following equality holds for any tree T

$$\limsup_{\varepsilon\to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 \leq \overline{\dim}\, T := \alpha\,. \tag{2.72}$$

Unpacking the definition of the limit, for any $\delta > 0$ there is a $a > 0$ such that for all $\varepsilon < a$, we have that

$$N^\varepsilon < C\varepsilon^{-\alpha-\delta}\,. \tag{2.73}$$

Let us fix such a δ and pick a small enough so that the condition above holds. For any $n \in \mathbb{N}^*$, the partial composition of intervals has length

$$|I_n| = |\phi(T^a)| + \sum_{k=1}^{n} \left|\phi_{\lambda^k}(T^{a/2^k} \setminus T^{a/2^{k-1}})\right|\,. \tag{2.74}$$

However, we can bound $\left|\phi_{\lambda^k}(T^{a/2^k} \setminus T^{a/2^{k-1}})\right|$ by

$$\begin{aligned}\left|\phi_{\lambda^k}(T^{a/2^k} \setminus T^{a/2^{k-1}})\right| &= \lambda^k \left|\phi(T^{a/2^k} \setminus T^{a/2^{k-1}})\right| \\ &\leq \lambda^k \left(\frac{a}{2^k}\right)(4N^{a/2^k})\,,\end{aligned} \tag{2.75}$$

since on $T^{a/2^k} \setminus T^{a/2^{k-1}}$ the distances between the vertices of each tree are at most $a/2^k$ and there are at most $4N^{a/2^k}$ such edges by virtue of lemma 2.25. Thus,

$$\left|\phi_{\lambda^k}(T^{a/2^k} \setminus T^{a/2^{k-1}})\right| < 4\lambda^k \left(\frac{a}{2^k}\right)^{1-\alpha-\delta} = 4a^{1-\alpha-\delta}\left(2^{\alpha+\delta-1}\lambda\right)^k . \tag{2.76}$$

Setting $\lambda < 2^{1-\alpha-\delta}$ I_n converges to some interval of finite length I, since the partial sums $|I_n|$ converge. ■

Proof of lemma 2.34. Suppose that $n < m$. It is sufficient to show that on I_m the equality holds, since in all further iterations of the algorithm, the functions f_n and f_m are locally constant over the intervals introduced. By definition of f_n, f_n and f_m agree on I_n. Outside of this set, f_n is constant and the difference in the L^∞-norm depends only on what happens above $T^{a/2^n}$, thus we can write

$$\|f_n - f_m\|_{L^\infty} \leq \left\|f_{T^{a/2^m} \setminus T^{a/2^n}}\right\|_{L^\infty} \tag{2.77}$$

by definition of f_n. However, the Dyck path on $T^{a/2^m} \setminus T^{a/2^n}$ can at most reach a height of $a(2^{-n} - 2^{-m}) < a2^{-n}$, which finishes the proof. ■

3 Regularity, persistence index and metric properties of trees

Throughout this section X will be a compact, connected and locally path-connected metric space. On general topological spaces, it is important to specify which homological theory we are using to compute the homology of X. For nice enough spaces, this choice has little to no importance, as most homological theories coincide. However, for abstract metric spaces this is no longer necessarily the case. For our purposes, we will always consider the homology of the space X to be its Čech homology. *A priori*, this might pose some problems, as Čech homology does not always satisfy the axioms of a proper homological theory in the sense of Eilenberg-Steenrod. For this to be the case, a sufficient condition is to consider X to be compact and the homology to be taken over a field. These are not the only conditions for which Čech homology gives rise to a proper homological theory, as in general the exactness axiom might fail, but suffices for our purposes. For more on these technical details, we encourage the reader to consult Eilenberg's book [40, Chapter 7].

Remark 3.1. If we wish to consider more general topological spaces where the exactness axiom does indeed fail for the Čech homology, there are multiple options. We could either consider more elaborate homology theories such as singular homology or strong homology (which fixes the issue with the exactness axiom of Čech homology), or we could rewrite this paper in cohomological terms and use Čech cohomology, for which this problem doesn't present itself.

With this technicality out of the way, let us now define the main objects which will concern us for the rest of this paper.

Definition 3.2. Let X be a compact, connected, locally path connected topological space and consider $f : X \to \mathbb{R}$ be a continuous function. The kth Pers_p**-functional of** f is

$$\mathrm{Pers}_p(H_k(X,f)) := \left(\sum_{b \in H_k(X,f)} \ell(b \cap [\inf(f), \sup(f)])^p \right)^{1/p}, \tag{2.78}$$

where $\ell(b)$ denotes the length of the bar b and $H_k(X,f)$ denotes the H_k-barcode (or diagram) stemming from the superlevel filtration. Abusing the notation, we will denote $\mathrm{Pers}_p(f) := \mathrm{Pers}_p(H_0(X,f))$. If we further assume that there exists n such that for all $m > n$, $H_m(X) = 0$, we define the **total** Pers_p **functional of** f as

$$\mathrm{TPers}_p(f) := \sum_{k=0}^{n} \mathrm{Pers}_p(H_k(X,f))\,. \tag{2.79}$$

Definition 3.3. Let $f : X \to \mathbb{R}$ be a continuous function. The k**th-persistence index of** f is defined as

$$\mathcal{L}_k(f) := \inf\{p \geq 1 \mid \mathrm{Pers}_p(H_k(X,f)) < \infty\}\,. \tag{2.80}$$

We will sometimes write $\mathcal{L}(f) := \mathcal{L}_0(f)$. Provided that higher degrees of homology identically vanish, we may also talk about the **total persistence index of** f, defined as

$$\mathcal{L}_{Tot}(f) := \inf\{p \geq 1 \mid \sum_k \mathrm{Pers}_p(H_k(X,f)) < \infty\}\,. \tag{2.81}$$

3.1 1D case: a connection with the p-variation

Definition 3.4. Let $f : [0,1] \to \mathbb{R}$ be a continuous function. The **true** p**-variation of** f is defined as

$$\|f\|_{p-\mathrm{var}} := \left[\sup_D \sum_{t_k \in D} |f(t_k) - f(t_{k-1})|^p \right]^{1/p}, \tag{2.82}$$

where the supremum is taken over all finite partitions D of the interval $[0,1]$.

Remark 3.5. We talk about *true* p-variation to make the distinction with the notion of variation typically considered in probabilistic contexts (more precisely, stochastic calculus), where instead of the supremum over all partitions, we have a probable limit as the mesh of the partition considered tends to zero.

Proposition 3.6 (Picard, §3 [80]). *Let $f : [0,1] \to \mathbb{R}$ be a continuous function, then $\|f\|_{p-var}$ is finite as soon as $\mathrm{Pers}_p(f)$ is finite. In fact, for any p*

$$\|f\|_{p-var}^p \leq 2\,\mathrm{Pers}_p^p(f)^p\,. \tag{2.83}$$

Furthermore, if $\|f\|_{(p-\delta)\text{-}var}$ is finite for some $\delta > 0$, $\mathrm{Pers}_p(f)$ is also finite.

In fact, Picard showed that on the interval $[0,1]$, the persistence index of f is linked to the regularity of f.

Theorem 3.7 (Picard, §3 [80])**.** *Let $f : [0,1] \to \mathbb{R}$ be a continuous function and denote*

$$\mathcal{V}(f) := \inf\{p \,|\, \|f\|_{p\text{-}var} < \infty\}\,. \tag{2.84}$$

Then,

$$\mathcal{V}(f) = \mathcal{L}(f) = \limsup_{\varepsilon\to 0} \frac{\log(\lambda(T_f^\varepsilon)/\varepsilon)}{\log(1/\varepsilon)} + 1 = \limsup_{\varepsilon\to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 = \overline{\dim}\, T_f \tag{2.85}$$

where $a \vee b := \max\{a, b\}$, N^ε is the number of leaves of the ε-trimmed tree T_f^ε, $\lambda(T_f^\varepsilon)$ denotes the length of T_f^ε and $\overline{\dim}$ denotes the upper-box dimension.

Remark 3.8. More generally, we can define λ as the unique atomless Borel measure on T_f characterized by the fact that the measure of a geodesic is given by the length of the geodesic [80].

3.2 More general spaces

Connected, locally path-connected, compact topological spaces

Theorem 3.9. *Let X be a connected, locally path-connected, compact topological space and let $f : X \to \mathbb{R}$ be a continuous function. With the same notation as above and supposing that $\overline{\dim}\, T_f$ is finite, the following chain of equalities holds*

$$\mathcal{L}(f) = \limsup_{\varepsilon\to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 = \limsup_{\varepsilon\to 0} \frac{\log(\lambda(T_f^\varepsilon)/\varepsilon)}{\log(1/\varepsilon)} + 1 = \overline{\dim}\, T_f\,. \tag{2.86}$$

Furthermore,

$$\liminf_{\varepsilon\to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 \leq \underline{\dim}\, T_f \leq \liminf_{\varepsilon\to 0} \frac{\log(\lambda(T_f^\varepsilon)/\varepsilon)}{\log(1/\varepsilon)} + 1\,, \tag{2.87}$$

where $\underline{\dim}$ is the lower-box dimension. For $\underline{\dim}\, T_f > 1$, these inequalities turn into equalities if either:

$$\limsup_{\varepsilon\to 0} \frac{N^{2\varepsilon}}{N^\varepsilon} < 1 \quad \textit{or} \quad \limsup_{\varepsilon\to 0} \frac{\lambda(T_f^{2\varepsilon})}{\lambda(T_f^\varepsilon)} < 1\,. \tag{2.88}$$

Remark 3.10. The study of N^ε is in fact completely equivalent to the study of $\mathrm{Pers}_p^p(f)$. Indeed,

$$\mathrm{Pers}_p^p(f) = p\int_0^\infty \varepsilon^{p-1} N^\varepsilon \, d\varepsilon\,, \tag{2.89}$$

which is finite as soon as $p > \mathcal{L}(f)$. This is nothing other than the Mellin transform of N^ε. By the Mellin inversion theorem, for any $c > \mathcal{L}(f)$, we have

$$N^\varepsilon = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \mathrm{Pers}_p^p(f)\, \varepsilon^{-p}\, \frac{dp}{p}\,. \tag{2.90}$$

Proof of theorem 3.9. By the procedure detailed in section 2.3, since $\overline{\dim}\, T_f$ is finite we can construct a function $\hat{f} : [0,1] \to \mathbb{R}$ such that T_f and $T_{\hat{f}}$ are isometric. Applying Picard's theorem to $T_{\hat{f}}$ and noting that $\mathcal{L}(f)$ depends only on the T_f, we have that

$$\mathcal{L}(f) = \limsup_{\varepsilon \to 0} \frac{\log(\lambda(T_f^\varepsilon)/\varepsilon)}{\log(1/\varepsilon)} + 1 = \limsup_{\varepsilon \to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 = \overline{\dim}\, T_f \,. \tag{2.91}$$

Let us now show the inequalities for the lim inf. Since

$$\lambda(T_f^\varepsilon) = \int_\varepsilon^\infty N^a \, da \,, \tag{2.92}$$

the following inequality holds

$$\liminf_{\varepsilon \to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 \le \liminf_{\varepsilon \to 0} \frac{\log(\lambda(T_f^\varepsilon)/\varepsilon)}{\log(1/\varepsilon)} + 1 \,. \tag{2.93}$$

Additionally,

$$N^\varepsilon \le \mathcal{N}(\varepsilon/2) \tag{2.94}$$

where $\mathcal{N}(\varepsilon)$ denotes the minimal number of balls of radius ε necessary to cover T_f. This inequality holds as above each leaf of T_f^ε, at least one ball of radius $\frac{\varepsilon}{2}$ is necessary to cover this section of the tree. It follows that

$$\liminf_{\varepsilon \to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1 \le \underline{\dim}\, T_f \,. \tag{2.95}$$

We can bound this minimal number of balls $\mathcal{N}(\varepsilon)$ by the following

$$\mathcal{N}(\varepsilon) \le N^{\varepsilon/2} + \frac{\lambda(T_f^{\varepsilon/2})}{\varepsilon/2} \le 2\, N^{\varepsilon/2} \vee \left\lceil \frac{\lambda(T_f^{\varepsilon/2})}{\varepsilon/2} \right\rceil \,, \tag{2.96}$$

which holds since, at most N^ε balls are needed to cover $T_f \setminus T_f^\varepsilon$. To cover T_f^ε, at most: $\left\lceil \lambda(T_f^{\varepsilon/2})/(\varepsilon/2) \right\rceil$ balls are needed, so the inequality above follows by further majorizing the terms. This implies that

$$\underline{\dim}\, T_f \le \left[\liminf_{\varepsilon \to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} \vee 1\right] \vee \left[\liminf_{\varepsilon \to 0} \frac{\log(\lambda(T_f^\varepsilon)/\varepsilon)}{\log(1/\varepsilon)} + 1\right] \,, \tag{2.97}$$

but by inequality 2.95 this means that

$$\underline{\dim}\, T_f \le \liminf_{\varepsilon \to 0} \frac{\log(\lambda(T_f^\varepsilon)/\varepsilon)}{\log(1/\varepsilon)} + 1 \,. \tag{2.98}$$

Finally,

$$\begin{aligned} \frac{\lambda(T_f^\varepsilon) - \lambda(T_f^{2\varepsilon})}{\varepsilon} &= \frac{1}{\varepsilon}\left[\int_\varepsilon^\infty N^a \, da - \int_{2\varepsilon}^\infty N^a \, da\right] \\ &= \frac{1}{\varepsilon} \int_\varepsilon^{2\varepsilon} N^a \, da \le N^\varepsilon \,, \end{aligned} \tag{2.99}$$

since N^ε is monotone decreasing. This reasoning also gives a lower bound

$$N^{2\varepsilon} \leq \frac{\lambda(T_f^\varepsilon) - \lambda(T_f^{2\varepsilon})}{\varepsilon} \leq N^\varepsilon\,, \tag{2.100}$$

which entails that

$$\liminf_{\varepsilon\to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} = \liminf_{\varepsilon\to 0} \frac{\log\left[\frac{\lambda(T_f^\varepsilon)-\lambda(T_f^{2\varepsilon})}{\varepsilon}\right]}{\log(1/\varepsilon)}\,. \tag{2.101}$$

Suppose that this limit is larger than 1. Rearranging, we get

$$\frac{\varepsilon N^{2\varepsilon}}{\lambda(T_f^\varepsilon)} \leq 1 - \frac{\lambda(T_f^{2\varepsilon})}{\lambda(T_f^\varepsilon)} \leq \frac{\varepsilon N^\varepsilon}{\lambda(T_f^\varepsilon)}\,, \tag{2.102}$$

from which it follows that if any of these quantities admits a lim inf which is stricly greater than zero, we have

$$\liminf_{\varepsilon\to 0} \frac{\log N^\varepsilon}{\log(1/\varepsilon)} = \liminf_{\varepsilon\to 0} \frac{\log \lambda(T_f^\varepsilon)}{\log(1/\varepsilon)} + 1\,. \tag{2.103}$$

Noticing another equivalent condition for the validity of this equality is whether

$$\limsup_{\varepsilon\to 0} \frac{N^{2\varepsilon}}{N^\varepsilon} < 1\,, \tag{2.104}$$

finishes the proof. ■

Remark 3.11. If $\overline{\dim} = \underline{\dim}$, all the limits of the above theorem are well-defined, yielding exact asymptotics for $\lambda(T_f^\varepsilon)$ and N^ε. This is in particular the case if $\overline{\dim} = \dim_H$, where $\dim_H$ denotes the Hausdorff dimension.

The functional $\lambda(T_f^\varepsilon)$ is what some authors [84, 85] refer to as the Banach indicatrix and its asymptotics have a topological interpretation as described in the statement of the theorem. It is interesting to note that the study of the upper-box dimension is natural in the tree approach. Additionally, $\overline{\dim}$ has also been used in the context of persistent homology by Schweinhart [91], Schweinhart and MacPherson [66] and by Adams *et al.* [2] in a probabilistic setting.

LLC metric spaces

It is possible to further extend Picard's theorem by some rudimentary considerations and by imposing the so-called locally linearly connected condition on X.

Definition 3.12. A **locally linearly connected (LLC) metric space** (X, d), is a connected metric space such that for all $r > 0$ and for all $z \in X$, for all $x, y \in B(z, r)$, there exists an arc connecting x and y such that the diameter of this arc is linear in $d(x, y)$.

With this extra assumption, we can prove the following lemma.

Lemma 3.13 (Regularity-dimension)**.** *Let X be a compact LLC metric space. Keeping the same notations as in theorem 3.9, the following inequality holds*

$$\mathcal{L}(f) = \overline{\dim}\, T_f \leq \mathcal{H}(f)\, \overline{\dim}\, X\,, \tag{2.105}$$

where:

$$\mathcal{H}(f) := \inf\left\{ \frac{1}{\alpha} \,\middle|\, \exists \lambda \in \mathrm{Homeo}(X)\,,\ \|f \circ \lambda\|_{C^\alpha} < \infty \right\} \tag{2.106}$$

The proof of this lemma relies itself on two lemmata, which are interesting in and of themselves.

Lemma 3.14. *Let X and Y be two metric spaces such that there is a surjective map $\pi : X \to Y$ such that $\pi \in C^\alpha(X, Y)$, then*

$$\overline{\dim}\, Y \leq \frac{1}{\alpha}\, \overline{\dim}\, X \quad \textit{and} \quad \underline{\dim}\, Y \leq \frac{1}{\alpha}\, \underline{\dim}\, X\,. \tag{2.107}$$

and if we denote K the Hölder constant of π, the following inequality holds

$$\mathcal{N}_Y(\varepsilon) \leq \mathcal{N}_X\left(\left(\frac{\varepsilon}{K}\right)^{1/\alpha}\right)\,. \tag{2.108}$$

Lemma 3.15. *Let X be a compact locally linearly connected (LLC) metric space (cf. definition 3.12) and let $f : X \to \mathbb{R}$ be a continuous function, then*

$$f \in C^\alpha(X, \mathbb{R}) \Longrightarrow \pi_f \in C^\alpha(X, T_f)\,. \tag{2.109}$$

Let us show that lemmata 3.14 and 3.15 imply lemma 3.13.

Proof of lemma 3.13. If, up to precomposition, $f \notin C^\alpha(X, \mathbb{R})$ for any α, there is nothing to show, since the statement is vacuous. Otherwise, since T_f is preserved by precomposition by a homeomorphism, we may suppose without loss of generality that $f \in C^\alpha(X, \mathbb{R})$. The projection onto the tree of f, $\pi_f : X \to T_f$ is in $C^\alpha(X, T_f)$ according to lemma 3.15. It follows from lemma 3.14 that

$$\overline{\dim}\, T_f \leq \frac{1}{\alpha}\, \overline{\dim}\, X\,. \tag{2.110}$$

The statement of the theorem follows by taking the infimum over $\frac{1}{\alpha}$. ■

All that remains to show is the two remaining lemmata.

Proof of lemma 3.14. Since $\pi : X \to Y$ is surjective and $C^\alpha(X, Y)$, for any $x \in X$

$$\pi\left(B_X\left(x, \left(\frac{\varepsilon}{K}\right)^{1/\alpha}\right)\right) \subset B_Y(\pi(x), \varepsilon) \tag{2.111}$$

for some constant K. It follows that the minimal number of balls needed to cover X, $\mathcal{N}_X$

dominates the minimal number of balls needed to cover Y, $\mathcal{N}_Y$. More precisely

$$\mathcal{N}_Y(\varepsilon) \leq \mathcal{N}_X\left(\left(\frac{\varepsilon}{K}\right)^{1/\alpha}\right) \iff \alpha \frac{\mathcal{N}_Y(\varepsilon)}{\log(1/\varepsilon)+\log(K)} \leq \frac{\mathcal{N}_X\left(\left(\frac{\varepsilon}{K}\right)^{1/\alpha}\right)}{\log\left(\left(\frac{K}{\varepsilon}\right)^{1/\alpha}\right)} .$$

The statement of the lemma follows. ■

Proof of lemma 3.15. Suppose that $f : X \to \mathbb{R}$ is in $C^\alpha(X, \mathbb{R})$ with Hölder constant Λ and let $x, y \in X$. Without loss of generality, suppose that $f(x) < f(y)$. Since T_f is a geodesic space, the distance $d_f(\pi_f(x), \pi_f(y))$ is the length of the geodesic arc in T_f linking $\pi_f(x)$ and $\pi_f(y)$. By compactness of this geodesic path, there is a point $\tau \in T_f$ where f achieves its minimum, thus

$$d_f(\pi_f(x), \pi_f(y)) = f(x) - f(\tau) + f(y) - f(\tau) . \tag{2.112}$$

This minimum $f(\tau)$ has the particularity that

$$f(\tau) = \sup_{\gamma:x\mapsto y} \inf_{t\in[0,1]} f \circ \gamma , \tag{2.113}$$

where the supremum is taken over all paths on X linking x and y. From the LLC condition, we know that there is a path $\eta : x \mapsto y$ whose diameter is controlled by $d_X(x, y)$ and $z \in X$ achieving the minimum of f over η. In particular,

$$f(\tau) \geq \inf_{t\in[0,1]} f \circ \eta =: f(z) . \tag{2.114}$$

Since f is α-Hölder on X,

$$f(x) - f(\tau) \leq f(x) - f(z) \leq \Lambda\, d(x, z)^\alpha \leq \Lambda \operatorname{diam}(\eta)^\alpha \leq C\Lambda\, d(x, y)^\alpha \tag{2.115}$$

for some constant C determined by the LLC condition and we have an analogous inequality for $f(y) - f(\tau)$. Putting everything together we have that:

$$d_f(\pi_f(x), \pi_f(y)) \leq 2C\Lambda\, d_X(x, y)^\alpha , \tag{2.116}$$

which finishes the proof. ■

Lemma 3.13 is sharp, since Brownian sample paths almost surely saturate this inequality. However, there is no hope to prove equality for every f. Indeed, for any $f \in C^1(\mathbb{T}^2, \mathbb{R})$ having a finite amount of bars, T_f is a finite tree and has upper-box dimension 1, but

$$\overline{\dim}\, T_f = 1 < 2 = \mathcal{H}(f)\, \overline{\dim}\, \mathbb{T}^2 . \tag{2.117}$$

Nonetheless, it is possible to show that lemma 3.13 holds generically. This is a consequence of a generalization of work never published by Weinberger and Baryshnikov. We extend their result to homogenous enough spaces in the following sense.

Definition 3.16. A metric space (X, d) is said to **admit a homogeneous set** (for a certain property) if there exists an open set $U \subset X$ where for every ball $B(x, r) \subset U$, the property of the ball is the same as the property of the space X.

Remark 3.17. In the previous definition, one can for instance take any notion of dimension, entropy, *etc.*

The following proposition will be useful in simplifying the assumptions of the theorem.

Proposition 3.18. *Let (X, d) be a compact metric space and $N_P(\varepsilon)$ denote the cardinality of the maximal packing of X by balls of radius ε. Then,*

$$\mathcal{N}_X(2\varepsilon) \leq N_P(\varepsilon) \leq \mathcal{N}_X(\varepsilon) \tag{2.118}$$

and in particular,

$$\underline{\dim}(X) = \liminf_{\varepsilon \to 0} \frac{\log(N_P(\varepsilon))}{\log(1/\varepsilon)} \quad \textit{and} \quad \overline{\dim}(X) = \limsup_{\varepsilon \to 0} \frac{\log(N_P(\varepsilon))}{\log(1/\varepsilon)} \,. \tag{2.119}$$

Proof. Let M_ε be a maximal packing of X by balls of radius ε. For every $x \in X \setminus (\cup_{V \in M_\varepsilon} V)$ there exists $U \in M_\varepsilon$ such that $d(x, U) \leq \varepsilon$, otherwise, $B(x, \varepsilon) \cup M_\varepsilon$ would also be a packing of X with cardinality strictly greater than $|M_\varepsilon|$. It follows that the balls of radius 2ε of centers that of the maximal packing of radius ε is a covering of X, proving the first inequality.

For the second inequality, we reason by contradiction. Suppose there is a maximal packing P_ε and a minimal covering C_ε such that $|P_\varepsilon| \geq |C_\varepsilon| + 1$. Then, since C_ε covers X, by the pigeonhole principle there are at least two centers of balls of P_ε inside a ball of C_ε. But the triangle inequality implies that the balls around these two centers of radius ε have non-empty intersection (as the center of the ball of C_ε in which they are contained is in the intersection), thereby contradicting that P_ε is a packing, showing the result. ∎

Theorem 3.19. *Let X be a compact LLC metric space admitting a set of homogeneous upper-box dimension, then for any $0 < \alpha \leq 1$*

$$\sup_{f \in C^\alpha(X, \mathbb{R})} \alpha \mathcal{L}(f) = \overline{\dim}(X) \,. \tag{2.120}$$

Moreover, the supremum is attained generically in the sense of Baire, i.e. the set over which $\alpha\mathcal{L}(f) < \overline{\dim}(X)$ is meagre in $C^\alpha(X, \mathbb{R})$.

Once again, we split the proof along key lemmata.

Lemma 3.20. *Let X be a compact LLC space, then the functional $\mathrm{Pers}^p_{p,\varepsilon} : C^\alpha_\Lambda(X, \mathbb{R}) \to \mathbb{R}_+$ defined by*

$$f \mapsto \sum_{\substack{b \in \mathcal{B}(f) \\ \ell(b) \geq \varepsilon}} \ell(b)^p \tag{2.121}$$

is continuous.

Proof. We start by noting that the total number of bars of length $\geq \varepsilon$ that a function $f \in C^\alpha_\Lambda(X,\mathbb{R})$ can have is uniformly bounded above by virtue of the proof of lemma 3.14 by a constant $C_{X,\alpha,\varepsilon}$. By lemma 3.15, we know $\pi_f : X \to T_f$ is α-Hölder, with Hölder constant K depending only on Λ and X. This fact, combined with the inequality $N^\varepsilon_f \leq \mathcal{N}_{T_f}(\varepsilon/2)$ entails that for any f,

$$N^\varepsilon_f \leq \mathcal{N}_{T_f}(\varepsilon/2) \leq \mathcal{N}_X\left(\left(\frac{\varepsilon}{2K}\right)^{1/\alpha}\right) =: C_{X,\alpha,\varepsilon}\,. \tag{2.122}$$

It follows that for any $f, g \in C^\alpha_\Lambda(X,\mathbb{R})$, by choosing to sum along the d_∞-matching, we have

$$\begin{aligned}\left|\mathrm{Pers}^p_{p,\varepsilon}(f) - \mathrm{Pers}^p_{p,\varepsilon}(g)\right| &\leq \sum_{\substack{b_f\in\mathcal{B}(f)\,,\, b_g\in\mathcal{B}(g)\\ \ell(b_f),\ell(b_g)\geq\varepsilon}} |\ell(b_f)^p - \ell(b_g)^p| \\ &\leq \sum p \underbrace{|\ell(b_f) - \ell(b_g)|}_{\leq\|f-g\|_\infty \text{ by stability}} \max\{\ell(b_f)^{p-1}, \ell(b_g)^{p-1}\} \\ &\leq p\,\|f-g\|_\infty \underbrace{\sum \max\{\ell(b_f)^{p-1}, \ell(b_g)^{p-1}\}}_{\leq C_{X,\alpha,\varepsilon}\Lambda^{p-1}\,\mathrm{diam}(X)^{\alpha(p-1)} \text{ by global } \alpha\text{-Hölderness}} \\ &\leq C_{X,\alpha,\varepsilon}\,\Lambda^{p-1}\mathrm{diam}(X)^{\alpha(p-1)}\,p\,\|f-g\|_\infty\,.\end{aligned}$$

■

Lemma 3.21. *Let X be a compact, LLC, admitting a set of homogeneous lower-box dimension. Then, for all $p < \overline{\dim}(X)$ and $M \geq 0$, the set of functions*

$$\{f \in C^\alpha(X,\mathbb{R}) \mid \mathrm{Pers}^p_p(f) > M\} \tag{2.123}$$

is dense in $C^\alpha(X,\mathbb{R})$.

Proof. Without loss of generality, suppose that the uniform set is a ball of radius 1 inside X, denoted $B \subset X$ and construct a function h of persistence $> M$ on this ball. Noting $d = \overline{\dim}(X)$, by proposition 3.18 and the definition of the upper-box dimension, for some subsequence of $(\varepsilon_n)_n$ decreasing to 0, we have

$$\tilde{C}\varepsilon_n^{-(d-\delta)} \leq N_P(\varepsilon) \leq C\varepsilon_n^{-(d+\delta)} \tag{2.124}$$

for some constants C and $\tilde{C}$. Note E_ε the centers of the balls of a maximal packing of radius ε and define $h_n : B \to \mathbb{R}$ as

$$h_n(x) := d^\alpha(x, E_{\varepsilon_n}) \tag{2.125}$$

The Pers_p-functional of these functions can be bounded below by

$$\mathrm{Pers}^p_p(h_n) \geq N_P(\varepsilon_n)\varepsilon_n^p \geq \tilde{C}\varepsilon_n^{p\alpha-d+\delta} \tag{2.126}$$

for all $\delta > 0$. Since $\alpha p < d$, this quantity can be made as large as we want and in particular $> M$ by picking a large enough n. By the assumptions of the theorem, it is possible to choose the original ball of the construction to have as small a radius as we wish. Note we may perturb any function $f \in C^\alpha(X, \mathbb{R})$ by a function close to it which is locally constant on a small enough ball and on this ball, add h_n for n large enough. Since the ball of the construction can be chosen as small as we want, any neighborhood of f contains a function satisfying the condition of the lemma. ∎

Proof of theorem 3.19. We are interested in showing that for $p < \overline{\dim}(X)$, the set

$$\mathcal{S}(p) := \{f \in C^\alpha(X, \mathbb{R}) \,|\, \mathrm{Pers}_p^p(f) < \infty\} \tag{2.127}$$

is meager in $C^\alpha(X, \mathbb{R})$. Let us start by noticing that

$$\mathcal{S}(p) = \bigcup_{\Lambda \geq 0} \bigcup_{M \geq 0} \mathcal{S}(p, \Lambda, M)\,, \tag{2.128}$$

where the union is taken over an increasing diverging sequences of Λ and M and

$$\mathcal{S}(p, \Lambda, M) := \{f \in C^\alpha_\Lambda(X, \mathbb{R}) \,|\, \mathrm{Pers}_p^p(f) \leq M\}\,. \tag{2.129}$$

Furthermore,

$$\mathcal{S}(p, \Lambda, M) = \bigcap_{k \geq 1} \{f \in C^\alpha_\Lambda(X, \mathbb{R}) \,|\, \mathrm{Pers}^p_{p, \frac{1}{k}}(f) \leq M\}\,. \tag{2.130}$$

By lemma 3.20, $\mathrm{Pers}_{p, \frac{1}{k}}$ is continuous, thereby guaranteeing that these sets are closed in $C^\alpha(X, \mathbb{R})$, and therefore so is their intersection. It remains to show that the $\mathcal{S}(p, \Lambda, M)$ are nowhere dense, but this amounts to finding a dense set of functions for which

$$\mathrm{Pers}^p_{p, \frac{1}{k}}(f) \leq M \tag{2.131}$$

is violated for infinitely many k. It suffices to find a dense set of functions for which the total $\mathrm{Pers}_p^p(f) > M$ (for $p < \overline{\dim}(X)$), but the existence of such a dense family is given by lemma 3.21, showing the result. ∎

Remark 3.22. The space defined by

$$E_p = \{f \in C^0(X, \mathbb{R}) \,|\, \mathcal{L}(f) \leq p\} \tag{2.132}$$

is **not** a linear space.

Doubling spaces with small convex balls

One could ask whether the results of genericity of theorem 3.19 hold in every degree of homology for f within some class of regularity. This question has been considered in [28] and more recently in [93] with different degrees of generality. The following theorem is a slight generalization of the two cited results.

Theorem 3.23. *Let X be a compact connected geodesic doubling space whose small enough balls are geodesically convex. Denote $d = \overline{\dim}(X)$, $k \in \mathbb{N}$ and let $f \in C^\alpha(X, \mathbb{R})$, then $\mathcal{L}_k(f) \leq \frac{d}{\alpha}$.*

Remark 3.24. The doubling assumption is satisfied for Riemannian manifolds whose Ricci curvature is bounded below, by the Bishop-Gromov inequality. By considering Gromov-Hausdorff limits of Riemannian manifolds with Ricci curvature bounded below, we obtain spaces satisfying the doubling property. Spaces included in this class include, but are not limited to, Riemannian manifolds with conic singularities. In general, it is also possible to obtain less well-behaved spaces. For more on poorly behaved examples, we refer the reader to the works of Xavier Menguy [69, 70] and to even more recent and poorly behaved examples, such as those described in [53].

The proof relies on the two following well-known lemmata.

Lemma 3.25 (Nerve lemma, Lemma 4.11 [76]). *Let X be a paracompact space, and let $\mathcal{U}$ be an open cover of X such that the $(k+1)$-fold intersections of elements of $\mathcal{U}$ are either empty or contractible for all $k \in \mathbb{N}$. Then, there is a homotopy equivalence between the nerve of $\mathcal{U}$ and X.*

Lemma 3.26. *Let (X, d) be a geodesic metric space whose balls of radius $\leq \varepsilon$ are geodesically convex. Then, minimal coverings of X by balls of radius $\leq r_C$ are such that the $(k+1)$-fold intersections of elements of $\mathcal{U}$ are either empty or contractible for all $k \in \mathbb{N}$. We call the maximal radius for which balls are geodesically convex the* ***convexity radius*** *of X and we denote it r_C.*

Proof of theorem 3.23. The proof is an immediate consequence of the proof of theorem 3.19, where we only need to modify the proof of lemma 3.20. For this, it is sufficient to bound the number of bars in the persistence diagram of the kth degree in homology of length $\geq \varepsilon$, N_k^ε. On a given a minimal covering $\mathcal{U}$ of X by balls of radius $\left(\frac{\varepsilon}{4\|f\|_{C^\alpha}}\right)^{1/\alpha}$, f varies by at most $\frac{\varepsilon}{2}$ inside each ball. Given any $r \in \mathbb{R}$, construct the set $\mathcal{U}_r$ consisting in the union of all balls of $\mathcal{U}$ which intersect X_r. From this, we get a chain of inclusions

$$X_r \hookrightarrow \mathcal{U}_r \hookrightarrow X_{r-\varepsilon}\,, \tag{2.133}$$

which induces a chain of maps at the homology level. More precisely, these inclusions entail that the map $H_*(X_r \hookrightarrow X_{r-\varepsilon})$ factorizes through $H_*(\mathcal{U}_r)$, *i.e.*

$$H_*(X_r \hookrightarrow X_{r-\varepsilon}) = H_*(X_r \hookrightarrow \mathcal{U}_r) \circ H_*(\mathcal{U}_r \hookrightarrow X_{r-\varepsilon})\,. \tag{2.134}$$

In particular, for all $0 \neq [\alpha] \in \mathrm{Im}(H_*(X_r \to X_{r-\varepsilon}))$, there exists a non-trivial cycle in $H_*(\mathcal{U}_r)$ homologous to $[\alpha]$, which is representable as a cycle of the nerve of the minimal covering by virtue of the isomorphism provided by the nerve lemma (lemma 3.25), here applicable by virtue of lemma 3.26 for $\varepsilon \leq 4\,\|f\|_{C^\alpha}\, r_C^\alpha$. In what follows, we will always identify

$$H_*(\mathcal{U}_r) \cong H_*(\mathcal{N}(\mathcal{U}_r))\,, \tag{2.135}$$

where $\mathcal{N}(\mathcal{U}_r)$ denotes the nerve of the covering $\mathcal{U}_r$ via this isomorphism.

Denote $([\alpha_i])_{1\leq i\leq N_k^\varepsilon}$ the homology classes represented by the N_k^ε bars (noting each of these is itself represented by a persistent cycle α_i). We may in particular order these classes or cycles by their births $b(\alpha_i)$ as follows

$$b(\alpha_{N_k^\varepsilon}) \leq \cdots \leq b(\alpha_1)\,. \tag{2.136}$$

We note that if $b(\alpha_i) = b(\alpha_{i+1}) = \cdots = b(\alpha_k)$, then the $([\alpha_j])_{i\leq j\leq k}$ are independent homology classes and thus the α_j are homologically independent cycles.

We now show that the representations of these cycles as cycles in the k-skeleton of the nerve of the covering are themselves independent, *i.e.* that the family $\{H_k(X_{b(\alpha_i)} \to \mathcal{U}_{b(\alpha_i)})(\alpha_i)\}_{1\leq i\leq m}$ is independent.

By induction, suppose that for a certain i, $b(\alpha_i) < b(\alpha_{i-1})$ and that the family

$$\{H_k(X_{b(\alpha_j)} \to \mathcal{U}_{b(\alpha_j)})(\alpha_j)\}_{1\leq j\leq i} \tag{2.137}$$

is dependent, then

$$H_k(X_{b(\alpha_i)} \to \mathcal{U}_{b(\alpha_i)})(\alpha_i) = \sum_{j=1}^{i-1} c_j\, H_k(X_{b(\alpha_i)} \to \mathcal{U}_{b(\alpha_i)})(\alpha_j)\,. \tag{2.138}$$

Composing both sides by $H_k(\mathcal{U}_{b(\alpha_i)} \to X_{b(\alpha_i)-\varepsilon})$, we have

$$H_k(X_{b(\alpha_i)} \to X_{b(\alpha_i)-\varepsilon})(\alpha_i) = \sum_{j=1}^{i-1} c_j\, H_k(X_{b(\alpha_i)} \to X_{b(\alpha_i)-\varepsilon})(\alpha_j), \tag{2.139}$$

which contradicts that the cycle α_i has persistence $\geq \varepsilon$. Suppose now that $b(\alpha_i) = b(\alpha_{i-1})$ and consider the minimal index $k < i$ for which $b(\alpha_k) = b(\alpha_i)$. We suppose dependence once again, so that after composing with $H_k(\mathcal{U}_{b(\alpha_i)} \to X_{b(\alpha_i)-\varepsilon})$ we get

$$\sum_{j=k}^{i} a_j\, H_k(X_{b(\alpha_i)} \to X_{b(\alpha_i)-\varepsilon})(\alpha_j) = \sum_{j=1}^{k-1} c_j\, H_k(X_{b(\alpha_i)} \to X_{b(\alpha_i)-\varepsilon})(\alpha_j)\,, \tag{2.140}$$

which once again contradicts that the cycles have persistence $\geq \varepsilon$. Since the cycles $(\alpha_j)_{k\leq j\leq i}$ are independent, we conclude that the family up to i is independent. This finishes showing the independence of the representations of the cycles in the k-skeleton of the nerve of the cover.

Since the family of N_k^ε k-cycles is independent in the nerve, we conclude that N_k^ε is bounded above by the cardinatily of the k-skeleton of the nerve of the minimal covering. The M-doubling property of the space yields the upper bound

$$N_k^\varepsilon \leq (M^{k+1} - M^k)\, \mathcal{N}_X\left(\left(\frac{\varepsilon}{4\,\|f\|_{C^\alpha}}\right)^{1/\alpha}\right)\,, \tag{2.141}$$

for $\varepsilon \leq 4\,\|f\|_{C^\alpha}\, r_C^\alpha$. The rest of the proof follows from previous arguments without extra difficulty. ∎

From the proof of theorem 3.23, we extract the following useful lemma.

Lemma 3.27. *Let X be a compact connected geodesic doubling space whose small enough balls* *are geodesically convex. Denote $d = \overline{\dim}(X)$, $k \in \mathbb{N}$ and let $f \in C^\alpha(X, \mathbb{R})$ and r_C denotes the convexity radius of X. Then,*

$$N^{\varepsilon}_{H_k(X,f)} \leq (M^{k+1} - M^k) \left[\mathcal{N}_X \left(\left(\frac{\varepsilon}{4 \|f\|_{C^\alpha}} \right)^{1/\alpha} \right) \vee \mathcal{N}_X(r_C) \right]$$

$$\mathrm{Pers}_p^p(H_k(X,f)) \leq 4^p \|f\|_{C^\alpha}^p \alpha p \int_0^{diam(X)} z^{\alpha p - 1} \left[\mathcal{N}_X(z) \vee \mathcal{N}_X(r_C) \right] \, dz \, .$$

Remark 3.28. If the space is not supposed to be doubling, the only bound we have on N_k^ε is given by $\mathcal{N}_X^{k+1}$, which yields an analogous statement for $\mathcal{L}_k(f) \leq \frac{d(k+1)}{\alpha}$.

Under a supplementary assumption, we can show that the inequality obtained in theorem 3.23 is in fact generically an equality. As before, the genericity result relies on the existence of functions whose Pers_p functional for $p < \frac{d}{\alpha}$ is arbitrarily large. For this we rely on the following theorem of Divol and Polonik.

Theorem 3.29 (Divol and Polonik, [35])**.** *Let μ be a bounded probability measure on $[0,1]^d$ and let $\mathbf{X}_n := (X_1, \cdots, X_n)$ be a vector of i.i.d. samples of μ, then for $0 < p < d$ and $0 \leq k < d$ then almost surely,*

$$\lim_{n\to\infty} n^{-1+\frac{p}{d}} \, \mathrm{Pers}_p^p(H_k([0,1]^d, d(-, \mathbf{X}_n)) \to \mathrm{Pers}_p^p(\nu_p^\mu) \, , \tag{2.142}$$

for some non-degenerate Radon measure depending on p and the probability measure μ, ν_p^μ on $\mathcal{X}$.

With this result we are now ready to prove the following theorem.

Theorem 3.30. *Let X be a compact Riemannian manifold of dimension d. Then, generically in the sense of Baire in $C^\alpha(X, \mathbb{R})$, for any $0 \leq k < d$, $\mathcal{L}_k(f) = \frac{d}{\alpha}$.*

Proof. The proof of genericity is essentially the same as that of theorem 3.19, with the exception that we now need to modify lemma 3.21. The existence of a function h with arbitrarily large Pers_p-functional for $p < \frac{d}{\alpha}$ on any small ball is given by Divol and Polonik's construction by tweaking the filtration in their proofs from being the distance d to d^α. As before, this entails the genericity result for the set of functions of C^α satisfying $\mathcal{L}_k(f) \geq \frac{d}{\alpha}$. Compact Riemannian manifolds have strictly positive convexity radii and Ricci curvature bounded below, and so satisfy the hypotheses of theorem 3.23, applying the theorem yields the desired equality. ■

3.3 A partial answer to a question by Schweinhart

In [91], Schweinhart introduces a notion of persistent homology dimension of a metric space X, defined as follows.

Definition 3.31 (Schweinhart's definition of $\dim_{\mathrm{PH}}^k$, [91])**.** Let X be a bounded subset of a metric space. The kth homological dimension of X is

$$\dim_{\mathrm{PH}}^k(X) := \inf_p \{ \sup_{\mathbf{x}} \mathrm{Pers}_p(H_k(X, d(-, \mathbf{x}))) < \infty \} \, , \tag{2.143}$$

where the supremum is taken over all finite sets of points $\mathbf{x}$ of X.

Given our previous results, we suggest the following modification to this definition, for reasons which will become apparent later.

Definition 3.32 (kth homological dimension of X)**.** Let X be a bounded subset of a metric space. The kth homological dimension of X is defined as

$$\dim_{\mathrm{PH}}^{k}(X) := \sup_{f \in \mathrm{Lip}_1(X)} \mathcal{L}_k(f)\,, \tag{2.144}$$

where $\mathrm{Lip}_1(X)$ denotes the set of Lipschitz functions with Lipschitz constant ≤ 1.

Theorem 3.23 already allows us to partially answer Schweinhart's Question 5 [91]. However, this is not a complete answer, because one should make sure that there are Lipschitz functions on X on the class of metric spaces as those of those of theorem 3.23 such that the inequality $\mathcal{L}_k(f) \leq d$ is saturated, or saturated to within δ for all $\delta > 0$. Without the assumption that X is doubling, an interesting question is whether the bound found is optimal: the proof of the theorem suggests that if such metric spaces exist, they cannot be of "bounded geometry" and are relatively pathological.

As we saw in theorem 3.30, this bound is saturated for any integer $0 \leq k < \overline{\dim}(X)$ under the assumption that X is a compact manifold. Thereby entailing

$$\dim_{\mathrm{PH}}^{k}(X) = \overline{\dim}(X) \tag{2.145}$$

for such X. Here, the notions of homological dimension of Schweinhart and our own coincide exactly, as the genericity result is proven via distance functions to point clouds. This thus establishes sufficient conditions for this equality to hold, albeit not necessary ones.

4 Distance notions and stability properties of trees and diagrams

4.1 Some elements of optimal transport

Defining optimal partial transport

Let us follow the exposition by Divol and Lacombe [33], and quickly introduce optimal *partial* transport, which extends optimal transport to measures of *a priori* different masses (which may be potentially infinite), for a detailed account of the theory, we refer the reader to the cited article, but also to the works of different authors [27, 43, 58]. Divol and Lacombe build on the work of Figalli [44] and extend Wasserstein distances to Radon measures supported on open proper subsets $\mathcal{X}$ of $\mathbb{R}^n$, whose boundary is denoted by $\partial\mathcal{X}$ (and $\overline{\mathcal{X}} := \mathcal{X} \sqcup \partial\mathcal{X}$). The general idea is that we should look at $\partial\mathcal{X}$ as a reservoir of infinite mass, capable of accomodating for any disparity in the mass of the measures considered. In this way, if two Radon measures μ and ν have different mass, we can form still define a transport map from one measure to the other by sending the mass surplus to the boundary $\partial\mathcal{X}$. Symbolically,

Definition 4.1. [44, Problem 1.1] Let $p \in [1, +\infty)$. Let μ, ν be two Radon measures supported on $\mathcal{X}$ satisfying

$$\int_{\mathcal{X}} d(x, \partial\mathcal{X})^p \, d\mu(x) < +\infty, \quad \int_{\mathcal{X}} d(x, \partial\mathcal{X})^p \, d\nu(x) < +\infty.$$

The set of **admissible transport plans** $\Gamma(\mu, \nu)$ is defined as the set of Radon measures π on $\overline{\mathcal{X}} \times \overline{\mathcal{X}}$ satisfying

$$\pi(A \times \overline{\mathcal{X}}) = \mu(A) \quad \text{and} \quad \pi(\overline{\mathcal{X}} \times B) = \nu(B).$$

for all Borel sets $A, B \subset \mathcal{X}$. Furthermore, the cost of $\pi \in \Gamma(\mu, \nu)$ is defined as

$$C_p(\pi) := \int_{\overline{\mathcal{X}} \times \overline{\mathcal{X}}} d(x, y)^p \, d\pi(x, y). \tag{2.146}$$

The **optimal transport distance** $d_p(\mu, \nu)$ is defined as

$$d_p(\mu, \nu) := \left(\inf_{\pi \in \Gamma(\mu, \nu)} C_p(\pi) \right)^{1/p}. \tag{2.147}$$

Plans $\pi \in \Gamma(\mu, \nu)$ realizing the infimum in equation 2.147 are called **optimal**.

Definition 4.2. The space of Radon measures on $\mathcal{X}$ will be denoted $\mathcal{D}(\mathcal{X})$ (or simply $\mathcal{D}$ if $\mathcal{X}$ is clear from context). We also introduce the following spaces

$$\mathcal{D}_p := \left\{ \mu \in \mathcal{D} \,\middle|\, \int_{\mathcal{X}} d^p(x, \partial\mathcal{X}) \, d\mu(x) < \infty \right\}. \tag{2.148}$$

We further define $\mathcal{D}_\infty$ as the space of Radon measures with compact support.

Remark 4.3. A proof by Théo Lacombe shows that for optimal partial transport distances d_p also satisfy $d_p \xrightarrow{p \to \infty} d_\infty$. Indeed, for any $\pi \in \Gamma(\mu, \nu)$

$$C_p(\pi) \xrightarrow{p \to \infty} C_\infty(\pi) \tag{2.149}$$

The space $\Gamma(\mu, \nu)$ is sequentially compact [33, Proposition 3.2], so up to extraction of a subsequence, $(\pi_p)_p$ admits a limit π_∞. Finally, if π^* is an optimal transport for the cost function C_∞, then

$$C_\infty(\pi^*) = \lim_{p \to \infty} C_p(\pi^*) \geq \lim_{p \to \infty} C_p(\pi_p) = C_\infty(\pi_\infty), \tag{2.150}$$

so π_∞ also achieves $\inf_\pi C_\infty(\pi)$, showing the desired result.

When considering optimal partial transport, there may be complications with respect to the conventional theory of optimal transport, because the measures may have infinite mass. This poses some problems, among others because of the unavalaibility of Jensen's inequality, which may render certain results of the classical theory false, or require alternative proofs. Luckily, most classical results we will need can be adapted to this more general setting.

Some results on optimal transport distances

To distinguish the theory of optimal transport from that of optimal *partial* transport, let us introduce the following notation.

Notation 4.4. Let (X, δ) be a Polish metric space. Denote $\mathcal{P}(X)$ (or simply $\mathcal{P}$ is X is clear from context) the set of **probability measures on** X and define

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P} \,\middle|\, \int_X \delta^p(x, x_0)\, d\mu(x) < \infty \right\} \tag{2.151}$$

for some $x_0 \in X$ (note that this definition does not depend on x_0). Once again, we may omit X if it is clear from context. For any two measures $\mu, \nu \in \mathcal{P}(X)$, slightly abusing then notation, we may define the **space of transport maps** $\Gamma(\mu, \nu)$ to be the space of probability measures on X^2 having marginals μ and ν. We equip the space $\mathcal{P}_p(X)$ with a Wasserstein distance, defined as

$$W_{p,\delta}(\mu, \nu) := \inf_{\pi \in \Gamma(\mu,\nu)} \|\delta\|_{L^p(\pi)} \ . \tag{2.152}$$

For the rest of this paper, the distance indicated by W will always reserved to classical Wasserstein distances between *probability* measures, whereas the distance denoted d_p will always refer to the notion of Wasserstein distances between general Radon measures, previously described in the context of optimal *partial* transport.

Many statements are valid whether we are in the optimal transport or the optimal *partial* transport setting. For this reason, we introduce the following generic notation along with the following dictionary to transpose statements to one setting or another.

Generic notation	Optimal transport	Optimal partial transport
(Y, d)	(X, δ)	$(\mathcal{X}, d)$
∂Y	$x_0 \in X$	$\partial \mathcal{X}$
OT_p	$W_{p,\delta}$	d_p
$\mathcal{M}(Y)$	$\mathcal{P}(X)$	$\mathcal{D}(\mathcal{X})$
$\mathcal{M}_p(Y)$	$\mathcal{P}_p(X)$	$\mathcal{D}_p(\mathcal{X})$

Table 2.1: Dictionary between optimal and optimal partial transport.

Proposition 4.5. *For any $1 \le p < \infty$, OT_p^p is convex, in the sense that for every $\mu_1, \mu_2, \nu \in \mathcal{M}_p$ and $t \in [0, 1]$,*

$$\mathrm{OT}_p^p(t\mu_1 + (1-t)\mu_2, \nu) \le t\, \mathrm{OT}_p^p(\mu_1, \nu) + (1-t)\, \mathrm{OT}_p^p(\mu_2, \nu)\,. \tag{2.153}$$

Moreover, if $\nu_1, \nu_2 \in \mathcal{M}_p$,

$$\mathrm{OT}_p^p(t\mu_1 + (1-t)\mu_2, t\nu_1 + (1-t)\nu_2) \le t\, \mathrm{OT}_p^p(\mu_1, \nu_1) + (1-t)\, \mathrm{OT}_p^p(\mu_2, \nu_2)\,. \tag{2.154}$$

Proof. For every $\pi_i \in \Gamma(\mu_i, \nu)$, $t\pi_1 + (1-t)\pi_2 \in \Gamma(t\mu_1 + (1-t)\mu_2, \nu)$, so

$$\mathrm{OT}_p^p(t\mu_1 + (1-t)\mu_2, \nu) \leq t \int_{Y^2} d(x,y)^p \, d\pi_1(x,y) + (1-t) \int_{Y^2} d(x,y)^p \, d\pi_2(x,y) \,, \tag{2.155}$$

which yields the result by taking the infimum over π_1 and π_2 on the right-hand side. The second convexity result is obtained by an analogous proof. ■

Remark 4.6. Convexity does not hold for $p = \infty$. By taking the $\frac{1}{p}$-th power of both sides and letting $p \to \infty$ in the inequality above, all that we may conclude is that

$$\mathrm{OT}_\infty(t\mu_1 + (1-t)\mu_2, \nu) \leq \max\{\mathrm{OT}_\infty(\mu_1, \nu), \mathrm{OT}_\infty(\mu_2, \nu)\} \,. \tag{2.156}$$

Theorem 4.7 (OT_p for $p = \infty$, [47])**.** *The distance obtained on $\mathcal{M}_\infty(Y)$ from OT_p by taking $p \to \infty$ is well-defined and coincides with the distance defined by*

$$\mathrm{OT}_\infty(\mu, \nu) = \inf_{\pi \in \Gamma(\mu,\nu)} \|d\|_{L^\infty(\pi)} \,. \tag{2.157}$$

Furthermore, we have the following characterization of OT_∞

$$\mathrm{OT}_\infty(\mu, \nu) = \inf\{r > 0 \mid \forall U \subset Y \text{ open, } \mu(U) \leq \nu(U^r) \text{ and } \nu(U) \leq \mu(U^r)\} \,, \tag{2.158}$$

where U^r denotes an open tubular neighborhood of radius r around U.

Remark 4.8. The topology of OT_∞ is finer than that of weak convergence.

Proposition 4.9. *Let $f : (Y, \delta) \to (Y', \delta')$ be an α-Hölder map with Hölder constant Λ and let $\mu, \nu \in \mathcal{P}(Y)$ then*

$$W_{p,\delta'}(f_\sharp\mu, f_\sharp\nu) \leq \Lambda \, W^\alpha_{p\alpha,\delta}(\mu, \nu) \,. \tag{2.159}$$

Proof. The inequality is an immediate consequence of the Hölder continuity of f. ■

Persistence measures

Coming back to persistence theory, recall that it is possible to see persistence diagrams as measures on

$$\mathcal{X} := \{(x,y) \in \mathbb{R}^2 \mid y > x\} \,. \tag{2.160}$$

Henceforth, $\mathcal{X}$ will always refer to this half space. Seen as measures, persistence diagrams are nothing other than a sum of Dirac measures. Closing this space with respect to the topology of vague convergence, we retrieve the set of Radon measures on $\mathcal{X}$.

Definition 4.10. The set of **persistence measures** $\mathcal{D}$ is the set of Radon measures (of potentially infinite mass) on $\mathcal{X} := \{(x,y) \in \mathbb{R}^2 \mid y > x\}$.

Equipping $\mathcal{X}$ with the ℓ^∞-distance on $\mathbb{R}^2$ defined by

$$d((p,q),(r,s)) = \max\{|p-r|\,, |q-s|\} \,, \tag{2.161}$$

optimal partial transport distances d_p between persistence measures become definable. The repercussions of this have been explored by Divol and Lacombe in [33].

The extension from the space of persistence diagrams to the space of persistence measures has three main advantages. First, that, as shown in [33], it is possible to use the machinery of optimal transport to address problems in persistence theory. Second, that $\mathcal{D}$ is a linear space, which renders taking means and combinations of diagrams possible and easy. Finally, that it is well-adapted to the stochastic setting, because of the linearity property and Tonelli's theorem: two key properties which we will exploit repeatedly.

Remark 4.11. The notion of average as defined in the linear space of persistence measures in general exits the space of persistence diagrams. This can for instance be seen by considering a sequence of measures which vaguely tend to a measure which is absolutely continuous with respect to the Lebesgue measure on $\mathcal{X}$. In this case, it is impossible to reconstruct a function whose diagram agrees with the desired measure. This is obvious in the 1D case where it is impossible to construct any tree from such a persistence measure, and so by extension, to construct any function. Nonetheless, this notion of average has the advantage of encoding the averages of all linear functionals of the diagrams (one can in fact see this as a definition of this notion of average by adopting a dual point of view). Some authors have considered alternative notions of central tendencies adapted to metric spaces (and in particular the space of diagrams), such as Fréchet means defined on the spaces of diagrams (*cf.* for instance the work of Turner *et al.* [96]). While this notion stays in the space where persistence diagrams are defined, it depends on the distance chosen on $\mathcal{D}$ and moreover also on the exponent chosen for the cost function in the definition of Fréchet means.

4.2 Stability of Wasserstein p-distances on diagrams

With respect to optimal transport distances, we have some "stability theorems" the most classical of which is

Theorem 4.12 (Bottleneck stability with respect to L^∞, Corollary 3.6 [76])**.** *Let $f, g : X \to \mathbb{R}$ be two continuous functions, then*

$$d_\infty(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq \|f - g\|_\infty \tag{2.162}$$

where $\mathrm{Dgm}(f)$ *and* $\mathrm{Dgm}(g)$ *denote the diagrams of f and g respectively.*

Theorem 4.13 (Wasserstein p stability)**.** *Let X be a compact LLC metric space of $\overline{\dim}(X) = d$ and consider $f, g \in C^\alpha(X, \mathbb{R})$. Then, for all $p > q > \frac{d}{\alpha}$,*

$$d_p^p(H_0(X, f), H_0(X, g)) \leq C_{X,\alpha} \left(\|f\|_{C^\alpha} \vee \|g\|_{C^\alpha}\right)^q \|f - g\|_\infty^{p-q} . \tag{2.163}$$

If X is assumed to be geodesic and is such that small enough balls of X are geodesically convex, then for every $k \in \mathbb{N}^*$, and all $p > q > \frac{d(k+1)}{\alpha}$

$$d_p^p(H_k(X, f), H_k(X, g)) \leq C_{X,\alpha,k} \left(\|f\|_{C^\alpha} \vee \|g\|_{C^\alpha}\right)^q \|f - g\|_\infty^{p-q} . \tag{2.164}$$

Finally, if X is further supposed to be doubling, then the inequality above holds for all $p > q > \frac{d}{\alpha}$.

Proof. Let $\Lambda = (\|f\|_{C^\alpha} \vee \|g\|_{C^\alpha})$. The first part of the proof is essentially as in [28]. Start by picking the bottleneck matching between the diagrams of f and g and denote it by $\gamma : \mathrm{Dgm}(f) \to \mathrm{Dgm}(g)$. Then for any $p > q > \frac{d}{\alpha}$,

$$\begin{aligned}
d_p^p(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) &\leq \sum_{b \in \mathrm{Dgm}(f)} d_{\mathcal{X},\infty}(b, \gamma(b))^p \\
&\leq \|f-g\|_\infty^{p-q} \sum_{b \in \mathrm{Dgm}(f)} d_{\mathcal{X},\infty}(b, \gamma(b))^q \\
&\leq 2^q \|f-g\|_\infty^{p-q} \sum_{b \in \mathrm{Dgm}(f)} d_{\mathcal{X},\infty}(b, \Delta)^q + d_{\mathcal{X},\infty}(\gamma(b), \Delta)^q \\
&= 2^q \|f-g\|_\infty^{p-q} (\mathrm{Pers}_q^q(f) + \mathrm{Pers}_q^q(g))
\end{aligned}$$

But both $\mathrm{Pers}_q^q(f)$ and $\mathrm{Pers}_q^q(g)$ are bounded above by a global constant for the class $C^\alpha_\Lambda(X, \mathbb{R})$, since by the proof of lemma 3.20

$$N_f^\varepsilon \leq \mathcal{N}_X\left(\left(\frac{\varepsilon}{2C\Lambda}\right)^{1/\alpha}\right), \tag{2.165}$$

where C is a constant stemming from the quantitative LLC condition on X. This inequality entails that

$$\begin{aligned}
\mathrm{Pers}_q^q(f) = q\int_0^\infty \varepsilon^{q-1} N_f^\varepsilon \, d\varepsilon &\leq q\int_0^{\Lambda\,\mathrm{diam}(X)^\alpha} \varepsilon^{q-1}\mathcal{N}_X\left(\left(\frac{\varepsilon}{2C\Lambda}\right)^{1/\alpha}\right) d\varepsilon \\
&= (2C\Lambda)^q \alpha q \int_0^{\frac{\mathrm{diam}(X)}{(2C)^{1/\alpha}}} \varepsilon^{q\alpha-1}\mathcal{N}_X(\varepsilon)\, d\varepsilon\,,
\end{aligned}$$

which is finite as soon as $q > \frac{d}{\alpha}$ since $\mathcal{N}_X(\varepsilon) = O(\varepsilon^{-d-\delta})$ as $\varepsilon \to 0$ for all $\delta > 0$, by definition of the upper-box dimension. The constant in the statement of the theorem is bounded above by the above estimate. The statements for with the supplementary assumptions of the theorem, the proof follows from the same reasononing by using the proof of theorem 3.23 and remark 3.28 ■

Remark 4.14. More generally, the proof of the theorem adapts with ease to accomodate any compact set of $C^0(X, \mathbb{R})$ admitting a global modulus of continuity dominated by a Hölder modulus of continuity. It is worth mentioning that such a theorem is impossible to prove for any regularity strictly worse than Hölder, as in such a class of regularity, there are functions f of infinite persistence index, so the theorem is vacuous.

Wasserstein stability results are common in the literature and are typically stated by making the following assumption on the underlying metric space X.

Definition 4.15. [28] A metric space X **implies bounded q-total persistence** if, for all $k \in \mathbb{N}$, there exists a constant C_X that depends only on X such that

$$\mathrm{Pers}_q^q(H_k(X, f)) < C_X \tag{2.166}$$

for every tame function f with Lipschitz constant $\mathrm{Lip}(f) \leq 1$.

The regime of validity of Wasserstein stability thus depends solely on this condition on X. We can thus see theorem 4.13 as a theorem giving explicit bounds on the q such that X implies bounded q-total persistence (in fact, it does so for every degree in homology independently). Following [28], it follows clearly from the proof of Wasserstein stability that this definition implies bounded persistence stability for Lipschitz functions.

Corollary 4.16. *Let X be a compact LLC metric space of $\overline{\dim}(X) = d$ of LLC constant C. Then, for all $f \in \mathrm{Lip}_1(X)$ and $p > q > d$,*

$$\mathrm{Pers}_q^q(H_0(X, f)) \leq (2C)^q q \int_0^{\frac{diam(X)}{2C}} \varepsilon^{q-1} \mathcal{N}_X(\varepsilon)\, d\varepsilon\,. \tag{2.167}$$

If X is assumed to be geodesic and is such that small enough balls of X are geodesically convex, then for every $k \in \mathbb{N}^*$, and all $p > q > d(k+1)$

$$\mathrm{Pers}_q^q(H_k(X, f)) \leq 4^q q \int_0^{diam(X)} \varepsilon^{q-1} (\mathcal{N}_X(\varepsilon) \vee \mathcal{N}_X(r_C))^k\, d\varepsilon\,, \tag{2.168}$$

where r_C denotes the convexity radius of X. Finally, if X is further supposed to be M-doubling, then for all $p > q > d$,

$$\mathrm{Pers}_q^q(H_k(X, f)) \leq 4^q q (M^{k+1} - M^k) \int_0^{diam(X)} \varepsilon^{q-1} (\mathcal{N}_X(\varepsilon) \vee \mathcal{N}_X(r_C))\, d\varepsilon\,. \tag{2.169}$$

Some other Wasserstein p stability results have been reported in the literature: Chen and Edelsbrunner [25] studied functions on non-compact domains of $\mathbb{R}^d$, obtaining a stability result which holds for $p > d$. The condition $p > d$ also appears in stability results for Čech filtrations for point clouds in $\mathbb{R}^d$ and the case of Vietoris-Rips filtrations was recently addressed in [93] by Skraba and Turner.

4.3 Distance notion and stability for trees

Definition 4.17. Let X and Y be two compact metric spaces, the **Gromov-Hausdorff distance,** $d_{GH}(X, Y)$ between X and Y, is defined as

$$d_{GH}(X, Y) := \inf_{\substack{f: X \to Z \\ g: Y \to Z}} \max \left\{ \sup_{x \in X} \inf_{y \in Y} d_Z(f(x), g(y)), \sup_{y \in Y} \inf_{x \in X} d_Z(f(x), g(y)) \right\}. \tag{2.170}$$

where the infimum is taken over all metric spaces Z and all isometric embeddings $f : X \to Z$ and $g : Y \to Z$.

The Gromov-Hausdorff distance quantifies how far away two metric spaces X and Y are from being isometric to each other. However, it is practically impossible to compute this distance with the above definition. To somewhat alleviate this, we will use the following characterization of the Gromov-Hausdorff distance:

Proposition 4.18 (Burago *et al.*, §7 [18]). *The Gromov-Hausdorff distance is characterized by*

$$d_{GH}(X,Y) = \frac{1}{2} \inf_{\mathfrak{R}} \sup_{\substack{(x,y)\in\mathfrak{R} \\ (x',y')\in\mathfrak{R}}} \left| d_X(x,x') - d_Y(y,y') \right| \,, \tag{2.171}$$

where the infimum is taken over all correspondences, i.e. subsets $\mathfrak{R} \subset X \times Y$ *such that for every* $x \in X$ *there is at least one* $y \in Y$ *such that* $(x,y) \in \mathfrak{R}$ *and a symmetric condition for every* $y \in Y$.

Remark 4.19. Given two surjective maps $\pi_X : Z \to X$ and $\pi_Y : Z \to Y$, it is possible to build a correspondence between X and Y by considering the set $\{(\pi_X(z), \pi_Y(z)) \in X \times Y \mid z \in Z\}$.

A natural question is to ask whether we have an equivalent statement about the stability of d_{GH} with respect to $\|\cdot\|_{L^\infty}$ and whether the two notions of distances are in some sense "compatible". We will positively answer this first question. In general d_∞ and d_{GH} are not compatible, in the sense that no inequality between the two holds in all generality (*cf.* remark 4.22). Le Gall and Duquesne [37] gave a first stability result of d_{GH} with respect to the L^∞-norm on continuous functions on $[0,1]$:

Theorem 4.20 (L^∞-stability of trees, [37]). *Let* $f, g : [0,1] \to \mathbb{R}$ *be two continuous functions. Then*

$$d_{GH}(T_f, T_g) \leq 2 \|f - g\|_{L^\infty} \,. \tag{2.172}$$

This result for functions on $[0,1]$ generalizes to more general topological spaces.

Theorem 4.21 (Stability theorem for trees). *Let* X *be a compact, connected and locally path connected topological space and let* f *and* $g : X \to \mathbb{R}$ *be two continuous functions, then*

$$d_{GH}(T_f, T_g) \leq 2 \|f - g\|_{L^\infty} \,. \tag{2.173}$$

Proof. We will use the distortion characterization of the Gromov-Hausdorff distance, which yields the following inequality

$$d_{GH}(T_f, T_g) \leq \frac{1}{2} \sup_{x,y\in X} |d_f(x,y) - d_g(x,y)| \,. \tag{2.174}$$

Following the logic of the proof of lemma 3.15, the distance between $\pi_f(x)$ and $\pi_f(y)$ is of the form

$$d_f(\pi_f(x), \tau) + d_f(\tau, \pi_f(y)) = f(x) - f(\tau) + f(y) - f(\tau) \tag{2.175}$$

where τ is the lowest point of the geodesic path in T_f between $\pi_f(x)$ and $\pi_f(y)$. This geodesic path on T_f admits preimages by π_f which are paths connecting x to y. These paths achieve the following supremum

$$\sup_{\gamma : x \mapsto y} \inf_{t\in[0,1]} f \circ \gamma = f(\tau) \leq f(x) \wedge f(y) \tag{2.176}$$

where $a \wedge b := \min\{a, b\}$ since by construction γ must always stay above $f(\tau)$ and since for $r > f(\tau)$, x and y lie in different connected components of X_r. If ν is the analogous vertex to τ

on T_g,

$$\begin{aligned} d_{GH}(T_f, T_g) &\leq \frac{1}{2} \sup_{x,y \in X} |d_f(x,y) - d_g(x,y)| \\ &= \frac{1}{2} \sup_{x,y \in X} |f(x) - g(x) + f(y) - g(y) - 2f(\tau) + 2g(\nu)| \\ &\leq \|f-g\|_{L^\infty} + \sup_{x,y\in X} \left| \sup_{\gamma : x \mapsto y} \inf_{t\in[0,1]} f \circ \gamma - \sup_{\eta : x \mapsto y} \inf_{t \in [0,1]} g \circ \eta \right| \\ &\leq 2 \|f-g\|_{L^\infty} , \end{aligned} \tag{2.177}$$

as desired. ■

Remark 4.22. One can be tempted to establish a general inequality between d_{GH} and d_∞ since both of these distances are bounded by the L^∞-norm. However, this is not possible.

Indeed, there is a simple counter-example to $d_{GH} \geq d_\infty$. To illustrate this consider two barcodes over a field k, $k[s, -\infty[$ and $k[s+\varepsilon, -\infty[$. The bottleneck distance between these two is clearly $\geq \varepsilon$. But supposing that the functions f and g generating these barcodes are such that $f = g + \varepsilon$ the trees T_g and T_f are isometric, so $d_{GH}(T_f, T_g) = 0 < \varepsilon \leq d_\infty(\mathcal{B}(f), \mathcal{B}(g))$.

Conversely, there are also counter-examples to $d_\infty \geq d_{GH}$, as this inequality would imply that two trees which have the same barcode are isometric. This is clearly false, as one can "glue" the bars of a given barcode is many different ways to give a tree, which generically will not be isometric.

5 Stochastic processes

As we have previously seen, the study of diagrams of continuous functions involves understanding their regularity. Many stochastic processes are *almost* Hölder continuous in the following sense.

Definition 5.1. The class of **almost α-Hölder continuous functions from X to $\mathbb{R}$**, denoted $E^\alpha(X, \mathbb{R})$ is the class of functions defined by

$$E^\alpha(X, \mathbb{R}) := \bigcap_{0 \leq \beta < \alpha} C^\beta(X, \mathbb{R}) \tag{2.178}$$

For example, Brownian motion and fractional Brownian motion are in a certain E^α for some value of α and moreover, as shown by Kahane [54, Chapter 7], random subgaussian Fourier series on torii of any dimension also tend to have E^α regularities. The ubiquity of E^α-regularities in the context of stochastic processes partially motivate this definition.

Notation 5.2. In what will follow, we will denote $f_\sharp \mathbb{P}$ the pushforward measure of $\mathbb{P}$ by f.

5.1 A change in perspective

Remark 5.3. Slightly abusing the notation, throughout this section, when we talk about a (continuous) stochastic process, we will talk about a measurable function $f : \Omega \to C^0(X, \mathbb{R})$ (where $(\Omega, \mathcal{F}, \mathbb{P})$ is some probability space).

Random diagrams, or more precisely, probability measures on the space of diagrams (or on the space of persistence measures) have been studied under many different contexts in the persistence theory literature [20, 21, 34, 42, 96]. Since ultimately we are interested in studying random processes on some base space X, the space of probability measures on the space of diagrams is far too large, as not all diagrams stem from (continuous) functions. In all practical applications, we are never given an abstract persistence diagram. Rather, we compute the persistence diagram from a certain continuous function (on which we may postulate further regularity assumptions, typically that the function is inside some $E^\alpha(X, \mathbb{R})$). This motivates studying subspaces of the full space of persistence diagrams of the form $\cup_k \operatorname{Dgm}_k(E^\alpha(X, \mathbb{R})) \subset \mathcal{D}$. This perspective turns out to have notable advantages. For instance, it is known that $(\mathcal{D}, d_\infty)$ is not a separable space [17, Theorem 5], but adopting this point of view we can show the opposite.

Proposition 5.4. *Let $K \subset (C^0(X, \mathbb{R}), \|\cdot\|_{L^\infty})$, be a closed subset, then $(\overline{\operatorname{Dgm}(K)}, d_\infty)$ is a Polish metric space.*

Proof. We start by noticing that the map Dgm is continuous and that the continuous image of a separable metric space is separable [98, Theorem 16.4a]. Moreover, $\overline{\operatorname{Dgm}(K)}$ remains separable, since the countable dense subset of $\operatorname{Dgm}(K)$ remains dense in the completion. ∎

Remark 5.5. If the subset K is compact, then $\operatorname{Dgm}(K) = \overline{\operatorname{Dgm}(K)}$. Notice also that the compact subsets of $C^0(X, \mathbb{R})$ are sets having a uniform modulus of continuity, by virtue of Ascoli's theorem. In particular, spaces such as $C^\alpha_\Lambda(X, \mathbb{R})$ are compact.

Consider now continuous $\mathbb{R}$-valued stochastic processes on X, f, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the space of probability measures on diagrams is also too large, as the probability measures we are concerned with must be of the form $(\operatorname{Dgm}_k \circ f)_\sharp \mathbb{P}$. For convenience, we could take the closure of this space induced by measures of this form with respect to the topology of vague convergence, or with respect to some Wasserstein distance $W_{p,\delta}$ (on the space of probability measures on diagrams). This is a technical point, but allows us to avoid making hypotheses on the probability measures on the space of diagrams, which are in practice almost never verifiable, and instead give hypotheses on the stochastic processes from which the diagrams stem from.

This point of view is particularly well-suited to look at stochastic processes supported on compact subsets of $C^0(X, \mathbb{R})$ (in fact, $E^\alpha(X, \mathbb{R})$, for reasons which will become apparent later). An easy first result in this direction is that

Proposition 5.6. *Let K be a compact subset of $C^0(X, \mathbb{R})$, then $\operatorname{Dgm}_k(K) \subset \mathcal{D}_\infty$.*

This restriction to compact sets can be seen as a considerable limitation. For example, Brownian motion on the interval $[0, 1]$ does not satisfy this hypothesis of compactness. However, by virtue of the tightness of probability measures on $C^0(X, \mathbb{R})$, we may restrict ourselves to a compact K_ε of $C^0(X, \mathbb{R})$ in which the process lies with probability $1 - \varepsilon$ and make probable statements there, or, alternatively, make conditional statements.

Furthermore,

Proposition 5.7. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and f be a $\mathbb{R}$-valued, a.s. E^α stochastic process on a d-dimensional compact manifold X. Then, for all $\varepsilon > 0$, $(\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P} \in \mathcal{P}(\mathcal{D}_{\frac{d}{\alpha}+\varepsilon} \cap \mathcal{D}_\infty)$ and a fortiori in $\mathcal{P}(\mathcal{D}_r)$ for every $\frac{d}{\alpha} < r < \infty$. Furthermore, if $\frac{d}{\alpha} < q < \infty$ and for all $\beta < \alpha$, $\mathbb{E}\left[\|f\|^q_{C^\beta(X,\mathbb{R})}\right] < \infty$, then, $\mathbb{E}[\mathrm{Dgm}_k(f)] \in \bigcap_{\frac{d}{\alpha} < p \leq q} \mathcal{D}_p$.*

Proof. Since $f \in E^\alpha(X, \mathbb{R})$ a.s., it is a.s. $C^\beta(X, \mathbb{R})$ for every $\beta < \alpha$, and so a.s. bounded by compactness of X. By theorem 3.23 and the previous remark, it follows that for every $k \in \mathbb{N}$, $\mathrm{Dgm}_k(f) \in \mathcal{D}_{\frac{d}{\beta}} \cap \mathcal{D}_\infty$, proving the first result.

Next, we remark that if $\mathbb{E}\left[\|f\|^q_{C^\beta(X,\mathbb{R})}\right]$ is finite so is the pth moment of the norm for every $1 \leq p \leq q$ by a simple application of Jensen's inequality. To show the result, it suffices to show that for such p,

$$\mathbb{E}\left[\mathrm{Pers}^p_p(f)\right] < \infty\,. \tag{2.179}$$

But applying Tonelli's theorem and using lemma 3.27,

$$\begin{aligned}\mathrm{Pers}^p_p(f) &= p\int_0^\infty \varepsilon^{p-1}\mathbb{E}\left[N^\varepsilon_f\right]\, d\varepsilon \\ &\leq 4^p \beta p \,\|f\|^p_{C^\beta} \int_0^{\mathrm{diam}(X)} \varepsilon^{p\beta-1}[\mathcal{N}_X(\varepsilon) \vee \mathcal{N}_X(r_C)]\, d\varepsilon\,.\end{aligned}$$

The integral on $[0, 1]$ is finite as soon as $p > \frac{d}{\alpha}$ since the dimension of X is d. Taking the expectation of both sides,

$$\mathbb{E}\left[\mathrm{Pers}^p_p(f)\right] \leq \tilde{C}_{X,p,\beta}\mathbb{E}\left[\|f\|^p_{C^\beta}\right]\,, \tag{2.180}$$

which is finite as soon as the moments of the C^β-norm of f are finite, exactly as supposed in the proposition. Finally, the *a fortiori* inclusion in $\mathcal{D}_r$ is a consequence of the Wasserstein interpolation theorem [79]. ∎

Corollary 5.8. *With the same hypotheses for f and p as in the previous proposition, for every $r \geq 1$ and every $\beta < \alpha$,*

$$\mathbb{E}\left[\mathrm{Pers}^r_p(f)\right] \leq C_{X,p,\beta}\mathbb{E}[\|f\|^r_{C^\beta}] \tag{2.181}$$

5.2 Consequences of stability

Equipped with some of the elementary facts from optimal transport theory, we may come back to persistence measures and diagrams. The main goal of this section will be to prove the following theorem.

Theorem 5.9 (Stability of random fields under Wasserstein perturbations)**.** *Let f and g be two $\mathbb{R}$-valued a.s. E^α stochastic processes on a d dimensional compact Riemannian manifold X on* *a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *Then, for any* $k \in \mathbb{N}$ *and any* $1 \leq p \leq \infty$,

$$W_{p,d_\infty}((\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm}_k \circ g)_\sharp \mathbb{P}) \leq W_{p,L^\infty}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P}) . \tag{2.182}$$

Moreover, for every $\frac{d}{\alpha} < q < p < \infty$ *and any* $r, s \in]1, \infty[$ *satisfying* $\frac{1}{r} + \frac{1}{s} = 1$ *and* $(p-q)s \geq 1$, *there exists a constant* C_X *depending only on* X *such that*

$$d_p(\mathbb{E}[\mathrm{Dgm}_k(f)], \mathbb{E}[\mathrm{Dgm}_k(g)]) \leq W_{p,d_p}((\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm}_k \circ g)_\sharp \mathbb{P}) \tag{2.183}$$

$$\leq C_X \left[\mathbb{E}\left[\|f\|_{C^\beta}^{qr} \right]^{\frac{1}{r}} + \mathbb{E}\left[\|g\|_{C^\beta}^{qr} \right]^{\frac{1}{r}} \right]^{\frac{1}{p}} W_{(p-q)s,\infty}^{1-\frac{q}{p}}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P}) \tag{2.184}$$

$$\leq C_X \left[\mathbb{E}\left[\|f\|_{C^\beta}^{qr} \right]^{\frac{1}{r}} + \mathbb{E}\left[\|g\|_{C^\beta}^{qr} \right]^{\frac{1}{r}} \right]^{\frac{1}{p}} \|f - g\|_{L^{(p-q)s}(\Omega, L^\infty(X,\mathbb{R}))}^{1-\frac{q}{p}} . \tag{2.185}$$

Finally, if the supports of $f_\sharp \mathbb{P}$ *and* $g_\sharp \mathbb{P}$ *are compact in* $E^\alpha(X, \mathbb{R})$, *then*

$$d_\infty(\mathbb{E}[\mathrm{Dgm}_k(f)], \mathbb{E}[\mathrm{Dgm}_k(g)]) \leq W_{\infty,d_\infty}((\mathrm{Dgm}_k \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm}_k \circ g)_\sharp \mathbb{P}) . \tag{2.186}$$

Remark 5.10. The proof of this theorem uses some of the techniques from [21, Lemma 15]. It differs from this result, as it concerns the d_p-stability as opposed to simply d_∞-stability, but also because the statement of theorem 5.9 gives a bound on the distance between expected diagrams, as opposed to a linear functional of the latter. However, necessary and sufficient conditions for the continuity of linear functionals of $\mathbb{E}[\mathrm{Dgm}(f)] \in (\mathcal{D}_p, d_p)$ has been studied by Divol and Lacombe in [33].

Proof of theorem 5.9. The first inequality is a simple consequence of a change of variables and an application of the bottleneck stability theorem. Next, notice that if $f_\sharp \mathbb{P}$ and $g_\sharp \mathbb{P}$ have compact support in E^α, then f and g are almost surely uniformly bounded functions, so $\mathbb{E}[\mathrm{Dgm}(f)]$ and $\mathbb{E}[\mathrm{Dgm}(g)]$ are both in $\mathcal{D}_\infty$.

Notice that,

$$\mathbb{E}[\mathrm{Dgm}(f)] = \int_{E^\alpha} \mathrm{Dgm}(h) \, df_\sharp \mathbb{P}(h) = \int_{(E^\alpha)^2} \mathrm{Dgm}(h) \, d\pi(h, \tilde{h}) , \tag{2.187}$$

for any $\pi \in \Gamma(f_\sharp \mathbb{P}, g_\sharp \mathbb{P})$ and an analogous equality holds for $\mathbb{E}[\mathrm{Dgm}(g)]$. Since d_p^p is convex, applying Jensen's inequality

$$\begin{aligned} d_p^p(\mathbb{E}[\mathrm{Dgm}(f)], \mathbb{E}[\mathrm{Dgm}(g)]) &= d_p^p \left(\int_{(E^\alpha)^2} \mathrm{Dgm}(h) \, d\pi(h, \tilde{h}), \int_{(E^\alpha)^2} \mathrm{Dgm}(\tilde{h}) \, d\pi(h, \tilde{h}) \right) \\ &\leq \int_{(E^\alpha)^2} d_p^p(\mathrm{Dgm}(h), \mathrm{Dgm}(\tilde{h})) \, d\pi(h, \tilde{h}) \\ &= \int_{(\mathrm{Dgm}(E^\alpha))^2} d_p^p(x, y) \, d\mathrm{Dgm}_\sharp^{\otimes 2} \pi(x, y) . \end{aligned}$$

Taking the infimum over every π of this inequality and taking the pth root,

$$d_p(\mathbb{E}[\mathrm{Dgm}(f)]\,, \mathbb{E}[\mathrm{Dgm}(g)]) \leq W_{p,d_p}((\mathrm{Dgm} \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm} \circ g)_\sharp \mathbb{P})\,.$$

Under the hypothesis of compactness, the result for $p = \infty$ is obtained by taking the limit $p \to \infty$, justified by remark 4.3 and the fact that the stochastic processes and their distributions in E^α are uniformly bounded.

Going back to the non-compact setting, keeping the same notation, let $\pi \in \Gamma((\mathrm{Dgm} \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm} \circ g)_\sharp \mathbb{P})$ be an optimal transport. For any $\beta < \alpha$, applying the Wasserstein p stability theorem, for all $p > q > \frac{d}{\beta}$,

$$\int_{(E^\alpha)^2} d_p^p(\mathrm{Dgm}(h), \mathrm{Dgm}(k))\, d\pi(h,k) \leq C_X \int_{(E^\alpha)^2} (\|h\|_{C^\beta}^q + \|k\|_{C^\beta}^q)\, \|h-k\|_\infty^{p-q}\, d\pi(h,k)\,. \tag{2.188}$$

By virtue of Hölder's inequality, for any $r, s \in]1,\infty[$ satisfying $\frac{1}{r} + \frac{1}{s} = 1$ and $(p-q)s \geq 1$,

$$\begin{aligned}
\int_{(E^\alpha)^2} \|h\|_{C^\beta}^q\, \|h-k\|_\infty^{p-q}\, d\pi(h,k) &\leq \left[\int_{(E^\alpha)^2} \|h\|_{C^\beta}^{qr}\, d\pi(h,k)\right]^{\frac{1}{r}} \left[\int_{(E^\alpha)^2} \|h-k\|_\infty^{(p-q)s}\, d\pi(h,k)\right]^{\frac{1}{s}} \\
&= \left[\int_{E^\alpha} \|f\|_{C^\beta}^{qr}\, d\mathbb{P}(\omega)\right]^{\frac{1}{r}} \left[\int_{(E^\alpha)^2} \|h-k\|_\infty^{(p-q)s}\, d\pi(h,k)\right]^{\frac{1}{s}} \\
&= \mathbb{E}\Big[\|f\|_{C^\beta}^{qr}\Big]^{\frac{1}{r}}\, W_{(p-q)s,\infty}^{p-q}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P})\,,
\end{aligned}$$

where the equality on the second line is valid since we know the marginals of π. Putting everything together we retrieve the statement of the theorem, namely that for a universal constant C_X depending only on X,

$$W_{p,d_p}((\mathrm{Dgm} \circ f)_\sharp \mathbb{P}, (\mathrm{Dgm} \circ g)_\sharp \mathbb{P}) \leq C_X \left[\mathbb{E}\Big[\|f\|_{C^\beta}^{qr}\Big]^{\frac{1}{r}} + \mathbb{E}\Big[\|g\|_{C^\beta}^{qr}\Big]^{\frac{1}{r}}\right]^{\frac{1}{p}} W_{(p-q)s,\infty}^{1-\frac{q}{p}}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P})\,.$$

The last inequality in the theorem is obtained by virtue of proposition 5.12. ■

This shows the following proposition.

Proposition 5.11. *Let $\mathcal{B}$ be a Banach space and $\Psi : \mathcal{D}_p \to \mathcal{B}$ be an α-Hölder continuous functional. Let $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_{\alpha q}(\mathcal{D}_p)$, then*

$$\|\mathbb{E}_\mathbb{P}[\Psi] - \mathbb{E}_\mathbb{Q}[\Psi]\|_\mathcal{B} \leq W_{q,\|\cdot\|_\mathcal{B}}(\Psi_\sharp \mathbb{P}, \Psi_\sharp \mathbb{Q}) \leq \|\Psi\|_{C^\alpha(\mathcal{D}_p,\mathcal{B})}\, W_{q\alpha,d_p}^\alpha(\mathbb{P},\mathbb{Q})\,. \tag{2.189}$$

Proposition 5.12 (Control of W_{p,L^∞})**.** *Let f and g be two $\mathbb{R}$-valued a.s. E^α stochastic processes on a d dimensional compact Riemannian manifold X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the following inequality holds*

$$W_{p,L^\infty}(f_\sharp \mathbb{P}, g_\sharp \mathbb{P}) \leq \|f-g\|_{L^p(\Omega, L^\infty(X,\mathbb{R}))} \tag{2.190}$$

Proof. The map $F : \Omega \to E^\alpha(X,\mathbb{R})^2$ which sends $\omega \mapsto (f(\omega), g(\omega))$ induces a transport map

$F_\sharp\mathbb{P} \in \Gamma(f_\sharp\mathbb{P}, g_\sharp\mathbb{P})$ and

$$W^p_{p,L^\infty}(f_\sharp\mathbb{P}, g_\sharp\mathbb{P}) \leq \int_{E^\alpha(X,\mathbb{R})^2} \|h-k\|^p \, dF_\sharp\mathbb{P}(h,k) = \int_\Omega \|f(\omega)-g(\omega)\|^p_\infty \, d\mathbb{P}(\omega)$$
$$= \|f-g\|^p_{L^p(\Omega, L^\infty(X,\mathbb{R}))} ,$$

which finishes the proof. ■

Remark 5.13. Proposition 5.12 yields an easy way to estimate the value of Wasserstein distances between stochastic processes. Using the results of [78] and other results on rates of convergence of random processes (which could be obtained by using results such as those of Kahane [54]), this instantly gives estimates for Wasserstein distances between distributions for a panoply of processes.

Corollary 5.14 (A remark on discretization)**.** *Keeping the same notation, fix a triangulation P of X whose* 0*-skeleton has n points and such that the* 0*-skeleton of P is an ε-net of X (this constrains $n \geq \mathcal{N}_X(\varepsilon)$) and define a new process $\hat{f}$ which is equal to f on the* 0*-skeleton of P and linearly interpolate in between. Then,*

$$W_{p,L^\infty}(f_\sharp\mathbb{P}, \hat{f}_\sharp\mathbb{P}) \leq \mathbb{E}\Big[\|f\|^p_{C^\beta}\Big]\, \varepsilon^{\beta p} . \tag{2.191}$$

If $p=\infty$ and that $\|f\|_{C^\beta}$ is uniformly bounded by L, then

$$W_{\infty,L^\infty}(f_\sharp\mathbb{P}, \hat{f}_\sharp\mathbb{P}) \leq L\varepsilon^\alpha \tag{2.192}$$

and theorem 5.9 applies.

Proof. Clearly, $\hat{f} : \Omega \to \mathrm{Lip}_{\Lambda_\varepsilon}(X,\mathbb{R})$ of law $\hat{f}_\sharp\mathbb{P}$. By proposition 5.12, for any $\beta < \alpha$,

$$W^p_{p,L^\infty}(f_\sharp\mathbb{P}, \hat{f}_\sharp\mathbb{P}) \leq \mathbb{E}\Big[\Big\|f-\hat{f}\Big\|^p_\infty\Big] \leq \mathbb{E}\Big[\|f\|^p_{C^\beta}\Big]\, \varepsilon^{\beta p} .$$

Taking $p \to \infty$, provided that the distribution of $\|f\|_{C^\beta}$ has bounded support, we can bound the support of this distribution by L, we get $W_{\infty,L^\infty}(f_\sharp\mathbb{P}, \hat{f}_\sharp\mathbb{P}) \leq L\varepsilon^\alpha$. In particular, the expected diagrams differ from less than $L\varepsilon^\alpha$ in d_∞. ■

Remark 5.15. The topology on the measures on $C^0(X,\mathbb{R})$ defined by Wasserstein distances may be too weak. Indeed, note that W_{p,L^∞}-balls around any measure μ supported on some $E^\alpha(X,\mathbb{R})$ include probability measures whose support intersects sets of $C^0(X,\mathbb{R})$ whose number of small bars grows faster than any polynomial (or indeed any computable function!). To see why, it suffices to exhibit an example of such a function (let us denote it h), and notice that if a stochastic process f has law μ, if ξ denotes a standard gaussian random variable, then $f + \varepsilon\xi h$ is (up to rendering f locally constant on some small ball) an arbitrarily small L^∞-perturbation of f whose number of small bars grows arbitrarily fast. In particular, this perturbation is not in any $\mathcal{D}_p$ for any p, but the law of this perturbed process is included within a W_{p,L^∞}-ball of arbitrarily small radius.

However, by changing topology to that of a Sobolev space which injects itself onto some $C^\alpha(X,\mathbb{R})$, we can avoid this problem. With this change in topology, it might be superfluous to require that the processes lie in $E^\alpha(X,\mathbb{R})$, as it might follow from an argument ressembling that of the proof of the Kolmogorov-Chentsov theorem (theorem 5.16).

5.3 Establishing classes of regularity

A sufficient and easily verifiable condition for a stochastic process to be almost surely E^α is given by the Kolmogorov-Chentsov theorem.

Theorem 5.16 (Kolmogorov-Chentsov Theorem for compact manifolds, [5, 6])**.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, $\mathcal{B}$ be a Banach space, X be a d-dimensional compact Riemannian manifold (without boundary) with distance d_X and $f:\Omega\times X\to\mathcal{B}$ be a $\mathcal{B}$-valued separable stochastic process. Suppose there exists constants $C>0$, $\varepsilon>0$ and $\delta>1$ such that for all $x,y\in X$,*

$$\mathbb{E}\Big[\|f(x)-f(y)\|_{\mathcal{B}}^{\delta}\Big] \leq C d_X(x,y)^{d+\varepsilon}\,, \tag{2.193}$$

then there exists a modification of f such that for all $\alpha\in[0,\frac{\varepsilon}{\delta}[$, f is almost surely α-Hölder continuous.

The proof uses the same idea of [5] to use the Sobolev embedding theorem. For compact Riemannian manifolds, the required Sobolev embedding theorem is given by [6, Theorem 2.20] (in fact, within [6], one can actually find Sobolev embedding theorems valid for wider classes of manifolds). Let us give a sketch of the proof.

Sketch of proof of theorem 5.16. First, by virtue of Markov's inequality, the estimation on the moments above entails that the process is continuous in probability. We may therefore assume that, up to taking a modification of f, the process f is measurable on $\Omega\times X$. Fix γ a real number, then Tonelli's theorem and the estimation of the moments above implies that

$$\begin{aligned}\mathbb{E}\left[\int_X\int_X \frac{\|f(x)-f(y)\|_{\mathcal{B}}^{\delta}}{d_X(x,y)^{d+\gamma\delta}}\,dx\,dy\right] &= \int_X\int_X \frac{\mathbb{E}\Big[\|f(x)-f(y)\|_{\mathcal{B}}^{\delta}\Big]}{d_X(x,y)^{d+\gamma\delta}}\,dx\,dy\\ &\leq C\int_X\int_X d_X(x,y)^{\varepsilon-\gamma\delta}\,dx\,dy\end{aligned}$$

which is finite as soon as $\gamma<\frac{d+\varepsilon}{\delta}$. Notice that the bounded quantity is nothing other than the norm of f in $L^\delta(\Omega,W^{\gamma,\delta}(X,\mathcal{B}))$, so that almost surely, $f_\omega\in W^{\gamma,\delta}(X,\mathcal{B})$. There is a Sobolev injection of $W^{\gamma,\delta}(X,\mathcal{B})\hookrightarrow C^\alpha(X,\mathcal{B})$ for all $\alpha<\gamma-\frac{d}{\delta}$, so for every $\alpha<\frac{\varepsilon}{\delta}$, there is a measurable set $\Omega_0\subset\Omega$ of probability measure 1 on which for every $\omega\in\Omega_0$, f_ω is α-Hölder almost everywhere on X. The corresponding modification can be obtained by making the trajectories continuous everywhere. Since the process f is measurable on $\Omega\times X$, we can set

$$g_\omega(h,x) := \frac{1}{\mathrm{Vol}(B(x,h))}\int_{B(x,h)} f_\omega(y)\,dy\,, \tag{2.194}$$

and consider the set

$$B = \{(\omega, x) \in \Omega \times X \mid (g_\omega(h,x))_h \text{ converges as } h \to 0\} \tag{2.195}$$

and set the continuous modification of f to be

$$g_\omega(x) := \begin{cases} \lim_{h\to 0} g_\omega(h,x) & (\omega, x) \in B \\ 0 & \text{else} \end{cases} . \tag{2.196}$$

Finally, it is easy to check this function is indeed α-Hölder everywhere on Ω_0 and to check that $\mathbb{P}(g(x) = f(x)) = 1$ almost everywhere on X. ■

Remark 5.17. If $\mathcal{B} = \mathbb{R}$, the same idea works (as shown in [5]) to prove results on the existence of modifications of processes such that the modification is almost surely of class C^k.

Provided that we have control over all moments of $\|f(x) - f(y)\|$, the Kolmogorov-Chentsov theorem constrains the regularity of the process to live within some family

$$\bigcap_{0 \le \alpha < \alpha^*} C^\alpha(X, \mathbb{R}) \tag{2.197}$$

for some α^*. As an immediate corollary,

Corollary 5.18. *With the same hypotheses and notation of theorem 5.16 where now $\mathcal{B} = \mathbb{R}$, denoting $\alpha^* := \sup_{\varepsilon,\delta} \frac{\varepsilon}{\delta}$, almost surely,*

$$\mathcal{L}_{Tot}(f) \le \frac{d}{\alpha^*} . \tag{2.198}$$

Chapter 3

On the persistent homology of almost surely C^0 stochastic processes

Abstract

We investigate the propreties of the persistence diagrams stemming from almost surely continuous random processes on $[0,t]$. We focus our study on two variables which together characterize the barcode : the number of points of the persistence diagram inside a rectangle $]-\infty, x] \times [x+\varepsilon, \infty[$, $N^{x,x+\varepsilon}$ and the number of bars of length $\geq \varepsilon$, N^ε. For processes with the strong Markov property, we show both of these variables admit a moment generating function and in particular moments of every order. Switching our attention to semimartingales, we show the asymptotic behaviour of N^ε and $N^{x,x+\varepsilon}$ as $\varepsilon \to 0$ and of N^ε as $\varepsilon \to \infty$. Finally, we study the repercussions of the classical stability theorem of barcodes and illustrate our results with some examples, most notably Brownian motion and empirical functions converging to the Brownian bridge.

Contents

1 Introduction

1.1 State of the art

One of the questions of interest in the theory of persistent homology is the following: given a random function on some topological space X, what can we say about the barcode $\mathcal{B}(X)$ of this process? The study of the topology of (super)level-sets of random functions has been a subject of interest in probability theory for a long time [3, 7, 54, 61, 72, 87]. Most prominently for this paper, by Le Gall and Duquesne, who gave a construction of a tree from any continuous function $f : [0,1] \to \mathbb{R}$ [37], and who interpreted different properties of these trees to give fine results about Lévy processes [38]. Picard later linked the upper-box dimension of these trees to the regularity of the function f [80]. In essence, these trees have proved to be a fruitful and natural setting from which many results regarding the topology of the superlevel sets of the function f stem [29, 30, 41, 74, 77].

A natural question is whether, or indeed how, these results are applicable to the persistent homology of stochastic processes. The answer turns out to be total: the study of barcodes and trees are completely equivalent in degree 0 of homology. This has been established in [77], in which a dictionary between H_0-barcodes and the trees of Le Gall and Duquesne was constructed.

In this paper, we focus mainly on two barcode-related quantities, namely the number of bars in the barcode of f of length $\geq \varepsilon$, which we will denote N_f^ε and the number of bars in the barcode of f whose intersection with the interval $[x, x+\varepsilon]$ is non-empty, which we will denote $N_f^{x,x+\varepsilon}$. Whenever the function is implicit from the context, we may omit the subscript f.

With the dictionary established in [77], it is possible to interpret the results previously obtained by Picard in the context of trees. Most notably, Picard showed that the regularity of functions is closely related to their small-bar asymptotics.

Theorem 1.1 (Picard, §3[80]). *Given a continuous function $f : [0,1] \to \mathbb{R}$,*

$$\mathcal{V}(f) = \overline{\dim}\, T_f = \limsup_{\varepsilon \to 0} \frac{\log N_f^\varepsilon}{\log(1/\varepsilon)} \vee 1 \tag{3.1}$$

where $\overline{\dim}$ denotes the upper-box dimension, $a \vee b = \max\{a,b\}$,

$$\mathcal{V}(f) := \inf\{p \geq 1 \mid \|f\|_{p-var} < \infty\}\,. \tag{3.2}$$

Moreover, for processes which are self-similar in distribution, Picard [80] shows the almost sure small-bar asymptotics is closely related to this self-similarity (the results are shown for trees, but are immediately interpretable in terms of barcodes by the results of [77]).

Theorem 1.2 (Picard, §3[80]). *Let $X : [0,1] \to \mathbb{R}$ be Brownian motion, a Lévy α-stable process or fractional Brownian motion, then, almost surely, there exists a constant C_H such that*

$$N_X^\varepsilon \sim \frac{C_H}{\varepsilon^{1/H}} \quad \text{as } \varepsilon \to 0\,, \tag{3.3}$$

where H is the self-similary index, i.e. the H for which X satisfies $X_{\lambda t} = \lambda^H X_t$ for every $\lambda > 0$

and $t \geq 0$ in distribution.

Parallel to these developments, some results regarding the persistent homology of Brownian motion have also been provided by the topological data analysis (TDA) community. In particular, for Brownian motion B Chazal and Divol gave a formula for the distribution of the number of points $N_B^{x,y}$ lying inside a given rectangle $]-\infty, x] \times [y, \infty[$ in the persistence diagram of Brownian motion, $\mathrm{Dgm}(B)$ [20].

Proposition 1.3 (Chazal, Divol, [20])**.** *For $0 < x < y$, the distribution of $N_B^{x,y}$ is*

$$\mathbb{P}(N_B^{x,y} \geq k) = \int_{\Sigma_{2k-1}} \psi(x, t_1)\psi(y-x, s_1)\psi(y-x, t_2)\cdots\psi(y-x, t_k) \prod_{i=1}^{k-1} dt_i \prod_{j=1}^{k-2} ds_j\,, \tag{3.4}$$

where Σ_{2k-1} denotes the corner of the domain bounded by the $(2k-1)$-simplex and we note a vector in $\mathbb{R}^{2k-1}$ by $(t_1, s_1, \cdots, s_{k-2}, t_{k-1})$ and

$$\psi(x,t) := \frac{x}{\sqrt{2\pi t^3}} e^{-\frac{x^2}{2t}}\,. \tag{3.5}$$

Using this result Chazal and Divol established that $\mathbb{E}[N^{x,y}]$ was C^1 in x and y. In the context of our own results, we will show this relation to be analytic and give an explicit expression for $N^{x,x+\varepsilon}$ for $x > 0$ and $\varepsilon > 0$.

Similar results were obtained by Baryshnikov [9], who computed exactly $\mathbb{E}[N_{B^\mu}^{x,y}]$ for the Brownian motion with a strictly positive linear drift $B_t^\mu := \mu t + B_t$ for $\mu > 0$ over the whole ray $[0, \infty[$.

Proposition 1.4 (Baryshnikov, [9])**.** *The expected value of $N_{B^\mu}^{x,x+\varepsilon}$ for $x > 0$ and $\mu > 0$ over the whole ray $[0, \infty[$ is given by*

$$\mathbb{E}\Big[N_{B^\mu}^{x,x+\varepsilon}\Big] = \frac{1}{e^{2\mu\varepsilon} - 1}\,. \tag{3.6}$$

In particular, as $\varepsilon \to 0$, the following asymptotic relation holds

$$\mathbb{E}\Big[N_{B^\mu}^{x,x+\varepsilon}\Big] \sim \frac{1}{2\mu\varepsilon} - \frac{1}{2} + \frac{1}{6}\mu\varepsilon + O(\mu^3\varepsilon^3) \quad \textit{as } \varepsilon \to 0\,. \tag{3.7}$$

In this paper, we will be concerned with almost surely C^0 processes, but it is noteworthy that the study of smooth random fields (and their topology) is currently an open field of research in probability theory. A good introduction to the smooth setting is provided by the celebrated books of Adler and Taylor [3], and that of Azaïs and Wschebor [7].

Whenever possible and necessary, we give the definitions and most important results necessary to make the proofs in this paper self-contained. However, for the sake of brevity, we do not introduce *all* the probabilistic concepts necessary for this paper (most notably, we do not define Brownian motion or stochastic processes, filtrated probability spaces, nor stopping times and the strong Markov property). We kindly refer the reader unfamiliar with these concepts to Le Gall's [61] and Revuz and Yor's [86] books on stochastic calculus for a comprehensive introduction to the subject.

1.2 Our contribution

Our contribution can be summarized along the following points.

Theorem 1.5 (Theorem 2.13)**.** *let X be an almost surely continuous stochastic process on $[0,t]$ with the strong Markov property. Define its range to be the random variable*

$$R_t = \sup_{[0,t]} X - \inf_{[0,t]} X \,. \tag{3.8}$$

Assume there exist ε^* such that $P(R_t > \varepsilon^*) < 1$. Then for all $\varepsilon > \varepsilon^*$, N_X^ε and $N_X^{x,x+\varepsilon}$ have moments of all order and admit a moment generating function.

This theorem answers the question of the existence of moments (and their quantification) for the quantities N^ε and $N^{x,x+\varepsilon}$ in the context of almost surely continuous Markov processes. It also allows us to show the behaviour of large bars of the barcode of such a process in expectation.

Corollary 1.6 (Corollary 2.17)**.** *Under the same assumptions,*

$$\mathbb{E}[N_X^\varepsilon] \sim \mathbb{P}(R_t \geq \varepsilon) \quad \text{as } \varepsilon \to \infty \,. \tag{3.9}$$

Having answered the question of large bars (at least partially) in this very general context, we switch our attention to the behaviour of small bars. Indeed, as the following result shows, the latter is very regular and closely related to the nature of the noise at hand.

Theorem 1.7 (Theorem 3.9)**.** *Let $X = M + A$ be a continuous semimartingale on $[0,t]$ and suppose that for $s \geq 1$*

$$\mathbb{E}\left[[M]_t^{s/2} + \left(\int_0^t |dA|_s\right)^s\right] < \infty \,. \tag{3.10}$$

Then in $L^s(\Omega)$,

$$\begin{aligned} N_X^{x,x+\varepsilon} &\sim \frac{L_X^x(t)}{2\varepsilon} + O(1) \quad \text{as } \varepsilon \to 0 \\ N_X^\varepsilon &\sim \frac{[X]_t}{2\varepsilon^2} + O(\varepsilon^{-1}) \quad \text{as } \varepsilon \to 0 \,, \end{aligned}$$

where $L_X^x(t)$ denotes the local time of X on $[0,t]$ at x.

For particular instances of local martingales, we have a full description of the barcode of these processes, which depends on their quadratic variation.

Theorem 1.8 (Theorem 5.3)**.** *For any continuous local martingale M on $[0,t]$ having deterministic and strictly increasing quadratic variation $[M]_t$ such that $[M]_\infty = \infty$,*

$$\begin{aligned} \mathbb{E}[N_M^\varepsilon] &= 4\sum_{k\geq 1}(2k-1)\,\mathrm{erfc}\left(\frac{(2k-1)\varepsilon}{\sqrt{2[M]_t}}\right) - k\ \mathrm{erfc}\left(\frac{2k\varepsilon}{\sqrt{2[M]_t}}\right) \\ &= \frac{[M]_t}{2\varepsilon^2} + \frac{2}{3} + 2\sum_{k\geq 1}(2(-1)^k - 1)\frac{e^{-\pi^2 k^2 [M]_t/2\varepsilon^2}[M]_t}{\varepsilon^2}\left[1 + \frac{\varepsilon^2}{\pi^2 k^2 [M]_t}\right] . \end{aligned}$$

Moreover on $[0,t]$ *and for* $x>0$,

$$\mathbb{E}\left[N_M^{x,x+\varepsilon}\right] = \sum_{k=1}^{\infty} \operatorname{erfc}\left(\frac{x+(2k-1)\varepsilon}{\sqrt{2[M]_t}}\right)$$
$$\sim \frac{1}{2\varepsilon}\int_0^{[M]_t} \varphi(x,s)\,ds + \sum_{k\geq 0} \frac{4(-2)^k\left(2^{2k+1}-1\right)\zeta(2k+2)}{\pi^{2k+2}} \left[\frac{\partial^k}{\partial s^k}\Big|_{s=[M]_t} \varphi(x,s)\right] \varepsilon^{2k+1} \text{ as } \varepsilon \to 0\,,$$

where $\varphi(x,t)$ *is the density of a centered Gaussian random variable of variance* t *and* ζ *denotes the Riemann zeta function.*

This theorem allows us to explain exactly the experimental observations we made through simulation regarding the barcode of Brownian motion (*cf.* figures 3.1 and 3.4).

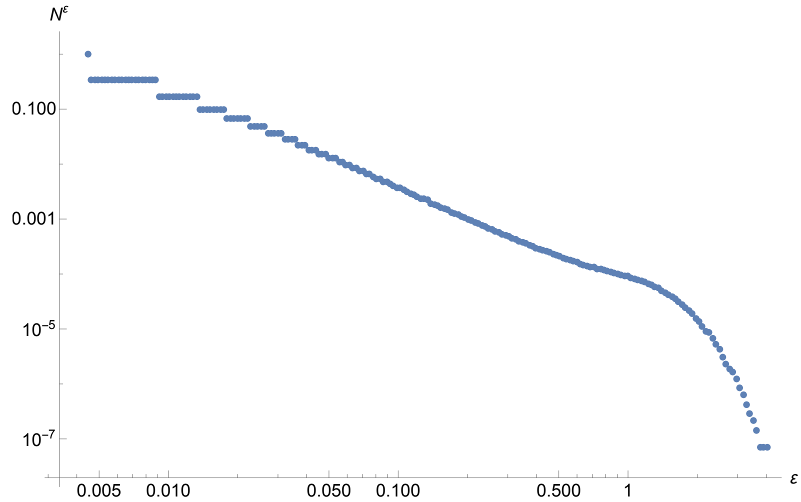


Figure 3.1: $\mathbb{E}[N_B^\varepsilon]$ as approximated with simulations of random Rademacher walks.

We use the classical stability theorem of barcodes [24] and complete our description of random processes by examining sequences converging to the processes above, which constitute universal limits for many random processes (most notably, random walks and empirical processes).

Theorem 1.9 (Propositions 4.1 and 4.3)**.** *Let* (M,d) *be a compact Polish metric space and let* X *be an almost surely continuous stochastic process on* M*, defined on a probability space* $(\Omega,\mathcal{F},\mathbb{P})$*. Let* $(X_n)_{n\in\mathbb{N}}$ *be any sequence of continuous stochastic processes defined on the same probability space and suppose there exists a* $p\geq 1$ *such that*

$$\delta_n := \|X-X_n\|_{L^p(\Omega,L^\infty(M,\mathbb{R}))} \xrightarrow[n\to\infty]{} 0\,. \tag{3.11}$$

Then, with probability $\geq 1-\frac{1}{a^p}$*, for every* $\varepsilon \geq 2a\delta_n$*,*

$$\left|N_{X_n}^\varepsilon - N_X^\varepsilon\right| \leq N_X^{\varepsilon-2a\delta_n} - N_X^{\varepsilon+2a\delta_n}\,. \tag{3.12}$$

Suppose further that $\mathbb{E}[N_X^\varepsilon]$ *is continuous in* ε. *Then,*

$$\mathbb{E}\Big[|N_{X_n}^\varepsilon - N_X^\varepsilon| \;\Big|\; \|X - X_n\|_\infty \le a\delta_n\Big] \le \omega_\varepsilon(2a\delta_n)\,. \tag{3.13}$$

where $\omega_\varepsilon(\delta) := \mathbb{E}\Big[N_X^{\varepsilon-\delta} - N_X^{\varepsilon+\delta}\Big]$. *Moreover,*

$$\mathbb{P}(|N_{X_n}^\varepsilon - N_X^\varepsilon| \ge k) \le \frac{\omega_\varepsilon(2a\delta_n)}{k} + \frac{1}{a^p} \quad \textit{and} \quad \mathbb{P}(|N_{X_n}^\varepsilon - N_X^\varepsilon| \ge k\;,\; \|X - X_n\|_\infty \le a\delta_n) \le \frac{\omega_\varepsilon(2a\delta_n)}{k}\,. \tag{3.14}$$

If $p = \infty$ *and still assuming* $\mathbb{E}[N_X^\varepsilon] < \infty$ *and is continuous, for any* $\varepsilon \ge 2\delta_n$,

$$N_{X_n}^\varepsilon \xrightarrow[n\to\infty]{L^1} N_X^\varepsilon \quad \textit{and} \quad \mathbb{E}[|N_X^\varepsilon - N_{X_n}^\varepsilon|] \le \omega_\varepsilon(2\delta_n)\,.$$

With analogous hypotheses, the same statement holds for $N_X^{x,x+\varepsilon}$.

Finally, we use the previous results to give statements about well-known processes, such as Brownian motion, Itô processes and some limiting processes (most notably the Fourier decomposition of Brownian motion and the convergence of empirical processes to the Brownian bridge), *cf.* section 5.

2 Some generalities about H_0 homology and trees

Let us briefly recall the construction of a tree from a continuous function $f : M \to \mathbb{R}$ detailed in [77].

Definition/Proposition 2.1. Let M denote a connected, locally path-connected, compact topological space, $f : M \to \mathbb{R}$ be a continuous function and let $x, y \in X$, then the function

$$d_f(x,y) := f(x) + f(y) - 2 \sup_{\gamma : x \mapsto y} \min_{t\in[0,t]} f(\gamma(t))\,, \tag{3.15}$$

where the supremum runs over all paths $\gamma : [0,1] \to M$ is a pseudo-distance on M and the quotient metric space

$$T_f := M/\{x \sim y \iff d_f(x,y) = 0\} \tag{3.16}$$

equipped with the distance d_f is a rooted $\mathbb{R}$-tree, whose root coincides with the image in T_f of the point in $[0,1]$ at which f achieves its infimum.

Definition 2.2. Let M denote a connected, locally path-connected, compact topological space and $f : M \to \mathbb{R}$ be a continuous function. We denote $\pi_f : M \to T_f$ the canonical projection.

The tree T_f has the particularity that its branches correspond to connected components of the superlevel sets of f, as illustrated by figure 3.2. To define N^ε on this tree, it is first necessary to introduce the so-called ε-simplified or ε-trimmed tree of T_f^ε. This object is obtained by "giving a haircut" of length ε to T_f. More precisely, if we define a function $h : T_f \to \mathbb{R}$ which to a point $\tau \in T_f$ associates the distance from τ to the highest leaf above τ with respect to the filtration on T_f induced by f, then

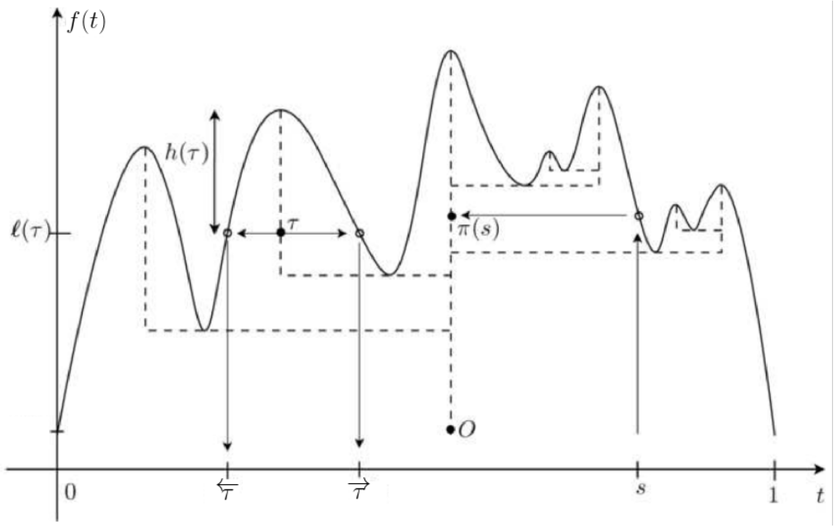


Figure 3.2: A function $f : [0, 1] \to \mathbb{R}$ and its associated tree T_f in dashed lines.

Definition 2.3. Let $\varepsilon \geq 0$. An ε**-trimming** or ε**-simplification** of T_f is the metric subspace of T_f defined by

$$T_f^\varepsilon := \{\tau \in T_f \mid h(\tau) \geq \varepsilon\} \tag{3.17}$$

With this definition, we can interpret N^ε geometrically as being equal to the number of leaves of T_f^ε. The reason for this is explicited in [77]. The idea is that, starting from T_f, we can look at the longest branch (starting from the root) of T_f. This branch corresponds to the longest bar of $\mathcal{B}(f)$, since branches of T_f correspond to connected components of the superlevel sets of f. Next, we erase this longest branch and, on the remaning (rooted) forest, look for the next longest branch. This will be the second longest bar of the barcode. Proceeding iteratively in this way, we retrieve $\mathcal{B}(f)$. An illustration of this can be found in figure 3.3.

Definition/Proposition 2.4. The number of bars alive at x persisting through $x + \varepsilon$,

$$N_f^{x,x+\varepsilon} := \operatorname{rank}(H_0(\{f \geq x + \varepsilon\} \to \{f \geq x\})) = \#\{\tau \in T_f \mid f(\tau) = x \text{ and } h(\tau) \geq \varepsilon\}\,. \tag{3.18}$$

Proof. By the results of [77], there is a bijective correspondence between points in T_f and points in the barcode (as seen as points in the collection of intervals of the barcode). It follows that every element of the set

$$\{\tau \in T_f \mid f(\tau) = x \text{ and } h(\tau) \geq \varepsilon\}\,, \tag{3.19}$$

has a corresponding image in one and only one bar of the barcode, whose intersection with $[x, x + \varepsilon]$ is not empty. Conversely, for every bar in the barcode with non-empty intersection with the interval $[x, x + \varepsilon]$ there is a unique image on the tree at height x, τ, and, since the bar persists a length $\geq \varepsilon$ after x, $h(\tau) \geq \varepsilon$. ■

Remark 2.5 (Link with traditional persistence diagrams and sublevel set filtrations). In the tree formalism, it is typical to consider **superlevel** filtrations as opposed to sublevel ones, as is typically done in traditional persistence theory. Considering one or the other poses no problem

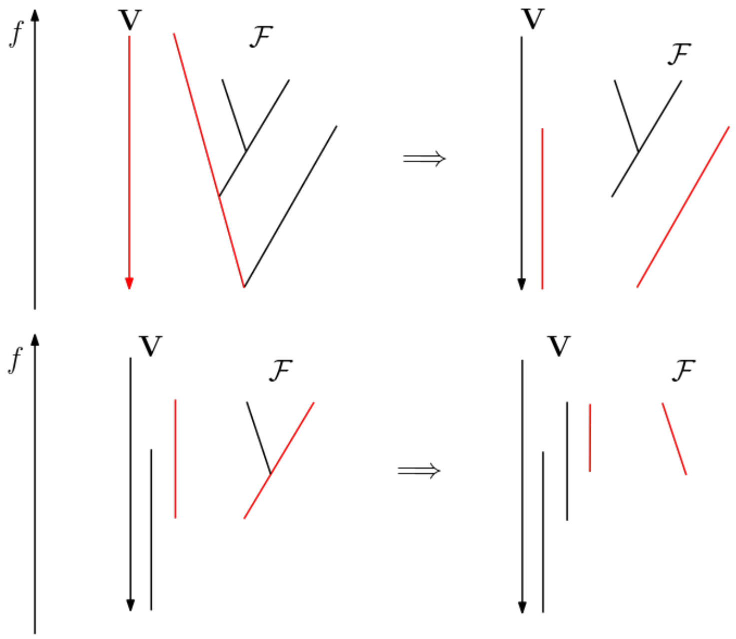


Figure 3.3: A depiction of the first steps of the algorithm which assigns a barcode $\mathcal{B}(f)$ to a tree T_f.

for us, as one can pass from one filtration to the other by switching f into $-f$. Diagrams (in the sense of collections of points (b,d) of moments of birth b and death d of bars of the barcode) of a filtration by superlevel sets lie *below* the diagonal, as the moment of birth occurs *higher* than the moment of death. Given the diagram of $-f$ as computed with the *superlevel* filtration (using for instance the tree), we can retrieve the diagram associated to *sublevels* of f by sending each point in the diagram (of the superlevel filtration of $-f$)

$$(b,d) \mapsto (-d,-b)\,. \tag{3.20}$$

For most results, this subtlety makes little to no difference, as most of the examples considered (typically Brownian motion B) are processes which are symmetric, *i.e.* $B_t = -B_t$ in distribution for all t. This is notably the case for the results obtained by Chazal and Divol [20].
Finally, with respect to the superlevel (resp. sublevel) set filtration, we will henceforth **always** consider that the infinite bars of the barcode are capped at $\inf(f)$ (resp. $\sup(f)$). Equivalently, in terms of barcodes we will always consider for any bar

$$b = b \cap [\inf f, \sup f]\,. \tag{3.21}$$

On a tree T_f, we can define a notion of integration by defining the unique atomless Borel measure λ which is characterized by the property that every geodesic segment on T_f has measure equal to its length. Formally, we can express λ in two ways [80]

$$\lambda = \int_{\mathbb{R}} dx \sum_{\substack{\tau\in T_f \\ f(\tau)=x}} \delta_\tau \quad \text{and} \quad \lambda = \int_0^\infty d\varepsilon \sum_{\substack{\tau\in T_f \\ h(\tau)=\varepsilon}} \delta_\tau\,. \tag{3.22}$$

Proposition 2.6.

$$\lambda(T_f^\varepsilon) = \int_\varepsilon^\infty N_f^a \, da = \int_{\mathbb{R}} N_f^{x,x+\varepsilon} \, dx \tag{3.23}$$

Proof. By using the second identity for λ,

$$\lambda(T_f^\varepsilon) = \int_\varepsilon^\infty N_f^a \, da \,, \tag{3.24}$$

since every sum in the second expression is finite for all $\varepsilon > 0$ and has N^ε terms. Writing it using the first identity, we must restrict the sum in the identity to

$$\sum_{\substack{\tau \in T_f \\ f(\tau)=x \\ h(\tau)\geq\varepsilon}} \delta_\tau \,, \tag{3.25}$$

which is finite for all $\varepsilon > 0$. There are exactly $N_f^{x,x+\varepsilon}$ terms in this sum, therefore

$$\lambda(T_f^\varepsilon) = \int_{\mathbb{R}} N_f^{x,x+\varepsilon} \, dx \,. \tag{3.26}$$

■

2.1 Persistent homology of processes on $[0, t]$

Consider now $f : [0, t] \to \mathbb{R}$. Given the total order on $\mathbb{R}$, the preimages of T_f by $\pi_f : [0, t] \to T_f$ inherit a natural order structure. This allows us to define

Definition 2.7. The **right** (resp. **left**) **preimage of** $\tau \in T_f$ **by** π_f is

$$\overrightarrow{\tau} := \sup \pi_f^{-1}(\tau) \quad (\text{resp. } \overleftarrow{\tau} := \inf \pi_f^{-1}(\tau)) \,. \tag{3.27}$$

Remark 2.8. Notice we can have $\overleftarrow{\tau} = \overrightarrow{\tau}$. See figure 3.2 for a depiction of these preimages.

Whenever $f : [0, t] \to \mathbb{R}$, we can compute N_f^ε is by counting the number of times we go up by at least ε from a local minimum and down by at least ε from a local maximum. This idea can be formalized by the following sequence, originally introduced by Neveu *et al.* [74].

Definition 2.9. Setting $S_0^\varepsilon = T_0^\varepsilon = 0$, we define a sequence of times by induction

$$T_{i+1}^\varepsilon := \inf \left\{ s \geq S_i \;\middle|\; \sup_{[S_i^\varepsilon, s]} f - f(s) > \varepsilon \right\}$$

$$S_{i+1}^\varepsilon := \inf \left\{ s \geq T_{i+1} \;\middle|\; f(s) - \inf_{[T_{i+1}^\varepsilon, s]} f > \varepsilon \right\}$$

Lemma 2.10. *If $k \geq 2$, and $f : [0, t] \to \mathbb{R}$ is continuous,*

$$N_f^\varepsilon \geq k \iff S_{k-1}^\varepsilon \leq t \,. \tag{3.28}$$

Proof. ($N^\varepsilon \geq k \implies S^\varepsilon_{k-1} \leq t$): We start by noticing we can order the bars by the value of their preimages by virtue of the total order on $\mathbb{R}$. Since $N^\varepsilon \geq k$, there are at least $(k-1)$ right preimages and left preimages by π_f of leaves of T_f^ε stemming from the first $(k-1)$ bars, which we denote $\{(\overleftarrow{\tau}_i, \overrightarrow{\tau}_i)\}_{1\leq i\leq k-1}$ satisfying

$$\overleftarrow{\tau}_1 \leq \overrightarrow{\tau}_1 = T_1^\varepsilon < S_1^\varepsilon \leq \overleftarrow{\tau}_2 \leq \overrightarrow{\tau}_2 = T_2^\varepsilon < \cdots \leq \overleftarrow{\tau}_{k-1} \leq \overrightarrow{\tau}_{k-1} = T^\varepsilon_{k-1}\,. \tag{3.29}$$

Note that $\overrightarrow{\tau}_1 > 0$ as soon as $k \geq 2$ and $\varepsilon > 0$. But $N^\varepsilon \geq k$, therefore there must exist a preimage $\overleftarrow{\tau}_k \leq t$ corresponding to the kth bar. But $S^\varepsilon_{k-1} \leq \overleftarrow{\tau}_k \leq t$ by definition of S^ε_{k-1}.

($N^\varepsilon \geq k \impliedby S^\varepsilon_{k-1} \leq t$): Since $S^\varepsilon_{k-1} > T^\varepsilon_{k-1} = \overrightarrow{\tau}_{k-1}$, there are at least $(k-1)$ distinct bars. We easily check that, by definition, $d_f(S^\varepsilon_{k-1}, T^\varepsilon_{k-1}) > 0$, implying that $\pi_f(S^\varepsilon_{k-1})$ and $\pi_f(T^\varepsilon_{k-1})$ lie on different branches of T_f, and therefore S^ε_{k-1} is a preimage of a distinct bar of length $\geq \varepsilon$, implying $N^\varepsilon \geq k$. ■

Definition 2.11. Let $f : [0,t] \to \mathbb{R}$ be a continuous function. Setting $US_0^{x,\varepsilon} = UT_0^{x,\varepsilon} = 0$, we define a sequence of times recursively

$$\begin{aligned} UT^{x,\varepsilon}_{i+1} &:= \inf\left\{ s \geq US_i^{x,\varepsilon} \;\middle|\; f(s) \leq x \wedge (x+\varepsilon)\right\} \\ US^{x,\varepsilon}_{i+1} &:= \inf\left\{ s \geq UT_{i+1}^{x,\varepsilon} \;\middle|\; f(s) \geq x \vee (x+\varepsilon)\right\}. \end{aligned}$$

The maximum i for which $US_i^{x,\varepsilon} \leq t$ is called the **number of upcrossings by** f **from** x **to** $x+\varepsilon$ and we denote it $U_f^{x,x+\varepsilon}$. Similarly, setting $DS_0^{x,\varepsilon} = DT_0^{x,\varepsilon} = 0$ and defining

$$\begin{aligned} DT^{x,\varepsilon}_{i+1} &:= \inf\left\{ s \geq DS_{i+1}^{x,\varepsilon} \;\middle|\; f(s) \leq x \wedge (x+\varepsilon)\right\} \\ DS^{x,\varepsilon}_{i+1} &:= \inf\left\{ s \geq DT_{i}^{x,\varepsilon} \;\middle|\; f(s) \geq x \vee (x+\varepsilon)\right\}, \end{aligned}$$

we can define the **number of downcrossings by** f **from** x **to** $x+\varepsilon$ and denote it $D_f^{x,x+\varepsilon}$ as the maximum i for which $DT_i^{x,\varepsilon} \leq t$.

Proposition 2.12. *Let $f : [0,t] \to \mathbb{R}$ be a continuous function. Then,*

$$N_f^{x,x+\varepsilon} = U_f^{x,x+\varepsilon} \vee D_f^{x,x+\varepsilon} \leq D_f^{x,x+\varepsilon} + 1\,. \tag{3.30}$$

If $x \geq 0$, and $f(0) = 0$, $N_f^{x,x+\varepsilon} = U_f^{x,x+\varepsilon}$.

Proof. Each bar alive at $x+\varepsilon$ persisting until x (seen as a branch of T_f) admits a left and right preimage (which can be sometimes equal), this shows $N^{x,x+\varepsilon} \geq U_f^{x,x+\varepsilon} \vee D_f^{x,x+\varepsilon}$. Conversely, every downcrossing (resp. upcrossing) generates a distinct homology class alive at $x+\varepsilon$ persisting until x, which shows $U_f^{x,x+\varepsilon} \vee D_f^{x,x+\varepsilon} \geq N^{x,x+\varepsilon}$. The last inequality is a consequence of the fact that, by definition $\left|U_f^{x,x+\varepsilon} - D_f^{x,x+\varepsilon}\right| \leq 1$. Finally, if $x \geq 0$, and $f(0) = 0$, by continuity of f, there is a bijective correspondence between upcrossings and the bars alive at $x+\varepsilon$ persisting until x. ■

Theorem 2.13. *Let X be a non-constant stochastic process on $[0,t]$ defined on the filtered probability space $(\Omega,\mathcal{F},\mathbb{P})$ allowing an almost surely continuous modification and satisfying the strong Markov property. Denote*

$$R_t := \sup_{[0,t]} X - \inf_{[0,t]} X\,, \tag{3.31}$$

Then,

$$\mathbb{P}(N^\varepsilon \geq k) \leq \mathbb{P}(R_t \geq \varepsilon)^{2(k-1)}\,. \tag{3.32}$$

Suppose further that X that there exists some ε^* such that for all $\varepsilon > \varepsilon^*$, $\mathbb{P}(R_t \geq \varepsilon^*) < 1$. Then for every $\varepsilon > \varepsilon^*$ and every $x \in \mathbb{R}$, all the moments of the random variables N^ε and $N^{x,x+\varepsilon}$ are finite and their moment generating functions $M(\lambda)$ converge uniformly and absolutely for every $\lambda \in \mathbb{C}$ such that $\mathrm{Re}(\lambda) > 2\log(\mathbb{P}(R_t \geq \varepsilon))$.

Proof. For simplicity, let us set $t=1$. The probability $\mathbb{P}(N^\varepsilon \geq k)$ can be written in terms of the stopping times T_i^ε and S_i^ε and their increments by the (strong) Markov property of X. By lemma 2.10, $\mathbb{P}(N^\varepsilon \geq k) = \mathbb{P}(S_{k-1}^\varepsilon \leq 1)$ for $k \geq 2$, so

$$\mathbb{P}(N^\varepsilon \geq k) = \int_{\Sigma_{2k-2}} \mathbb{P}(T_1^\varepsilon = t_1)\mathbb{P}(S_1 = s_1\,|\,T_1 = t_1)\cdots\mathbb{P}(S_{k-1} = s_{k-1}\,|\,T_{k-1} = t_{k-1})\prod_{i=1}^{k-1} ds_i dt_i\,. \tag{3.33}$$

where Σ_{2k-2} denotes the simplex

$$\Sigma_{2k-2} := \left\{(t_1, s_1, \cdots, s_{k-1}) \in \mathbb{R}^{2k-1} \mid 0 \leq t_1 \leq s_2 \leq \cdots \leq s_{k-1} \leq 1\right\}\,. \tag{3.34}$$

By the definition of these stopping times we know that

$$\begin{aligned}
\mathbb{P}(s \leq T_i^\varepsilon \leq t | S_{i-1} = s) &= \mathbb{P}\left(\sup_{\tau\in[s,t]}\left[\sup_{[s,\tau]} X - X_\tau\right] \geq \varepsilon\right)\\
\mathbb{P}(t \leq S_i^\varepsilon \leq s | T_i = t) &= \mathbb{P}\left(\sup_{\tau\in[t,s]}\left[X_\tau - \inf_{[t,\tau]} X\right] \geq \varepsilon\right)
\end{aligned}$$

Both of these expressions are dominated by $\mathbb{P}(R_1 \geq \varepsilon)$. Indeed,

$$\mathbb{P}\left(\sup_{\tau\in[s,s']}\left[\sup_{[s,\tau]} X - X_\tau\right] \geq \varepsilon\right) \leq \mathbb{P}\left(\sup_{s\in[0,1]}\left[\sup_{[0,s]} X - X_s\right] \geq \varepsilon\right) \tag{3.35}$$

and the supremum on the right hand side is dominated by R_1. Thus,

$$\mathbb{P}\left(\sup_{\tau\in[s,s']}\left[\sup_{[s,\tau]} X - X_\tau\right] \geq \varepsilon\right) \leq \mathbb{P}(R_1 \geq \varepsilon)\,. \tag{3.36}$$

Integrating the expression as a nested integral of $\mathbb{P}(N^\varepsilon \geq k)$, the variable s_{k-1} between t_{k-1} and

1

$$\mathbb{P}(N^\varepsilon \geq k) = \int_{\Sigma_{2k-3}} \mathbb{P}(T_1^\varepsilon = t_1)\mathbb{P}(S_1 = s_1 \,|\, T_1 = t_1) \cdots \mathbb{P}(t_{k-1} \leq S_{k-1} \leq 1 \,|\, T_{k-1} = t_{k-1})\, dt_{k-1} \prod_{i=1}^{k-2} ds_i dt_i$$
$$\leq \mathbb{P}(R \geq \varepsilon) \int_{\Sigma_{2k-3}} \mathbb{P}(T_1^\varepsilon = t_1)\mathbb{P}(S_1 = s_1 \,|\, T_1 = t_1) \cdots \mathbb{P}(T_{k-1} = t_{k-1} \,|\, S_{k-2} = s_{k-2}) dt_{k-1} \prod_{i=1}^{k-2} ds_i dt_i$$

Carrying out the subsequent $2k-3$ integrations and by repeated use of the inequality given in equation 3.36, we obtain the result

$$\mathbb{P}(N^\varepsilon \geq k) \leq \mathbb{P}(R \geq \varepsilon)^{2k-2} \,. \tag{3.37}$$

By the hypothesis of the theorem, for all $\varepsilon \geq \varepsilon^*$, $\mathbb{P}(R \geq \varepsilon) < 1$ so the above condition guarantees the summability (and absolute and uniform convergence) of the series $\mathbb{E}\left[e^{\lambda N^\varepsilon}\right]$ on the half plane $\mathrm{Re}(\lambda) > 2\log(\mathbb{P}(R_1 \geq \varepsilon))$. Moreover the same summability condition holds for $N^{x,x+\varepsilon}$, since the latter is dominated by N^ε. ■

2.2 *A priori* estimates on the asymptotic behaviour of small and large bars

If f is a continuous function, the asymptotic behaviour of N^ε is closely related to the regularity of f. For functions $f : [0,t] \to \mathbb{R}$, the correct notion of regularity to look at is the p-variation, of which we recall the definition.

Definition 2.14. Let $f : [0,t] \to \mathbb{R}$ be a function. The **(true)** p**-variation of** f is defined as

$$\|f\|_{p\text{-var}} := \sup_{\mathcal{P}} \left[\sum_{t_k \in \mathcal{P}} |f(t_k) - f(t_{k-1})|^p \right]^{1/p} \tag{3.38}$$

where the supremum ranges over all finite partitions of $[0,t]$.

The p-variation can be used to infer something about the asymptotic behaviour of N_f^ε, as shown by theorem 1.1.

Remark 2.15. A general version of theorem 1.1, applicable on more general metric spaces exists [77].

Corollary 2.16. *For any deterministic function* $f : [0,1] \to \mathbb{R}$ *and for every* $\delta > 0$,

$$N_f^\varepsilon = O(\varepsilon^{-\mathcal{V}(f)-\delta}) \quad \textit{as } \varepsilon \to 0 \,. \tag{3.39}$$

Having characterized the behavious of N^ε for small ε, we can characterize the behaviour of large bars (in expectation) using theorem 2.13.

Corollary 2.17. *Let* X *be a non-constant stochastic process on* $[0,t]$ *defined on the filtered probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *allowing an almost surely continuous modification and satisfying the* *strong Markov property. Suppose also that there exists some ε^* such that for all $\varepsilon > \varepsilon^*$, $\mathbb{P}(R_t \geq \varepsilon^*) < 1$, then,*

$$\mathbb{E}[N^\varepsilon] \sim \mathbb{P}(R_t \geq \varepsilon) \quad \textit{as } \varepsilon \to \infty\,. \tag{3.40}$$

Proof. Note $p_\varepsilon = \mathbb{P}(R_t \geq \varepsilon)$. From the theorem, we deduce,

$$p_\varepsilon \leq \mathbb{E}[N^\varepsilon] \leq p_\varepsilon + \frac{p_\varepsilon^2}{1-p_\varepsilon^2} \quad \implies \quad \mathbb{E}[N^\varepsilon] \sim p_\varepsilon \text{ as } \varepsilon \to \infty\,. \tag{3.41}$$

■

3 Continuous semimartingales, local times and asymptotic behaviour of barcodes

In the previous section, we quantified the asymptotics of $N^{x,x+\varepsilon}$ an N^ε solely based on the regularity of the functions considered. For stochastic processes on $[0,t]$, we can refine this analysis by focusing on (continuous) semimartingales. Semimartingales constitute the largest class of processes with respect to which the Itô and Stratonovich integrals can be defined. In other words, they are a class of processes rich enough to be worthy of particular attention. For a comprehensive introduction to these objects and the probabilistic concepts included in this section, we kindly refer the reader to classical references on stochastic calculus [61, 86].

Definition 3.1 (Local martingales and semimartingales)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a filtered probability space and let $\mathcal{F}_* := (\mathcal{F}_t)_{t\geq 0}$ be the filtration of $\mathcal{F}$. An $\mathcal{F}_*$-adapted process $M : [0,\infty] \times \Omega \to \mathbb{R}$ is an **$\mathcal{F}_*$-local martingale** if there exists a sequence of stopping times τ_k such that

1. The τ_k are a.s. increasing, *i.e.* $\mathbb{P}(\tau_k < \tau_{k+1}) = 1$;

2. The τ_k are a.s. divergent, *i.e.* $\mathbb{P}(\lim_{k\to\infty} \tau_k = \infty) = 1$;

3. The process $M_{t\wedge\tau_k}$ is an $\mathcal{F}_*$-martingale, *i.e.* for all $s < t$,

$$\mathbb{E}[M_{t\wedge\tau_k} \,|\, \mathcal{F}_s] = M_{s\wedge\tau_k}\,. \tag{3.42}$$

A process $(X_t)_{t\geq 0}$ is a **continuous semimartingale** if it can be written in the form

$$X_t = M_t + A_t \tag{3.43}$$

where A_t is a process of finite variation and M_t is a continuous local martingale.

Throughout this section, we will use two key concepts stemming from the theory of stochastic integration: the local time of a (continuous) semimartingale and the quadratic variation. The latter is the easiest to define, given our introduction of the **true** quadratic variation in the previous section.

3.1 Quadratic variation and *a priori* bounds

Definition 3.2 (Quadratic variation, Theorem 4.9 [61])**.** Suppose that $(X_t)_t$ is a $\mathbb{R}$-valued stochastic process indexed by $\mathbb{R}_+$. The **quadratic variation of** X is the process defined as

$$[X]_t := \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2 \,, \tag{3.44}$$

where $\mathcal{P}$ ranges over the set of finite partitions of the interval $[0, t]$ and $\|\mathcal{P}\|$ the mesh of the partition $\mathcal{P}$. If it exists, the limit is taken in the sense of convergence in probability. Whenever X is a continuous semimartingale, the quadratic variation always exists.

Remark 3.3. This definition is strictly weaker than the **true** 2-variation, which is defined as a supremum over the set of all finite partitions. In fact, for Brownian motion B, it is possible to show that $[B]_t = t$, but on the interval $[0, t]$, $\|B\|_{2-\text{var}} = \infty$ almost surely.

Understanding the quadratic variation is in some sense enough to understand local martingales, as the following theorem shows.

Theorem 3.4 (Dambis-Dubins-Schwarz)**.** *Let M be a continuous local martingale vanishing at* 0 *such that $[M]_\infty = \infty$, then there exists a Brownian motion B such that, a.s. for all $t \geq 0$,*

$$M_t = B_{[M]_t} \,. \tag{3.45}$$

This theorem allows us to give *a priori* bounds on the small bar asymptotics of the barcode of a semimartingale. Indeed, since Brownian motion is a.s. $(\frac{1}{2} - \delta)$-Hölder continuous for every $\delta > 0$, the p-variation of any semimartingale is almost surely finite as soon as $p > 2$ (since p-variation does not depend on parametrization). In particular we expect

$$N_X^\varepsilon = O(\varepsilon^{-2-\delta}) \quad \text{as } \varepsilon \to 0 \,. \tag{3.46}$$

for every $\delta > 0$, by virtue of Picard's theorem. We may further refine this result by introducing the local time.

3.2 The local time and sharp asymptotics

The local time of a continuous semimartingale on an interval $[0, t]$ can be informally understood as the "time spent" by the process X around a level $x \in \mathbb{R}$, in an equation

$$L_X^x(t) = \int_0^t \delta(x - X_s)\, d[X]_s \,. \tag{3.47}$$

The above equation is informal and ill-defined, but gives an intuitive insight on what the local time represents and how it behaves. Formally, we define the local time as follows.

Definition/Proposition 3.5 (Proposition 9.2 [61])**.** Let X be a continuous semimartingale and $x \in \mathbb{R}$. There exists an increasing process $(L_X^x(t))_{t\geq 0}$ such that the three following identities hold

$$\begin{aligned}|X_t - x| &= |X_0 - x| + \int_0^t \operatorname{sgn}(X_s - x)\, dX_s + L_X^x(t)\\ (X_t - x)_+ &= (X_0 - x)_+ + \int_0^t \mathbf{1}_{X_s > x}\, dX_s + \frac{1}{2} L_X^x(t)\\ (X_t - x)_- &= (X_0 - x)_- - \int_0^t \mathbf{1}_{X_s \le x}\, dX_s + \frac{1}{2} L_X^x(t)\end{aligned}$$

The increasing process $(L_X^x(t))_{t\geq 0}$ is called the **local time of X at level x**. Furthermore, for every stopping time T, the local time at x of the stopped process X at T, $L_{X^T}^x(t) = L_X^x(t \wedge T)$.

The local time of a process is useful, as it allows us to exchange time and space in integrations. A first useful result in this direction, which we will later use, is the following proposition.

Proposition 3.6 (Density occupation formula, Corollary 9.7 [61])**.** *Almost surely, for every non-negative, measurable function ϕ on $\mathbb{R}$,*

$$\int_0^t \phi(X_s)\, d[X]_s = \int_{\mathbb{R}} \phi(a) L_X^a(t)\, da\,. \tag{3.48}$$

This result can be informally derived using the informal definition of the local time we gave earlier. For a rigorous proof, we refer the reader to the cited reference. From the informal description we gave about the local time as the time spent by the process around a level x, we could expect the local time to be related to $N_X^{x,x+\varepsilon}$ for small ε. This turns out to be a well-known fact.

Proposition 3.7 (Approximation of the local time by downcrossings, §VI, Theorem 1.10 [86])**.** *For every $t \geq 0$ and $s \geq 1$, if $X = M + A$ is a continuous semimartingale on $[0, t]$ and suppose that for $s \geq 1$*

$$\mathbb{E}\Big[[M]_t^{s/2} + \Big(\int_0^t |dA|_s\Big)^s\Big] < \infty\,. \tag{3.49}$$

Then,

$$\varepsilon D_X^{x,x+\varepsilon} \xrightarrow[\varepsilon\to 0]{L^s} \frac{1}{2} L_X^x(t)\,. \tag{3.50}$$

Moreover, in $L^s(\Omega)$,

$$D_X^{x,x+\varepsilon} \sim \frac{L_X^x(t))}{2\varepsilon} + O(1) \quad \textit{as } \varepsilon \to 0\,. \tag{3.51}$$

Remark 3.8. An analogous statement can be proven for upcrossings.

Theorem 3.9. *Let $X = M + A$ be a continuous semimartingale on $[0, t]$ and suppose that for $s \geq 1$*

$$\mathbb{E}\Big[[M]_t^{s/2} + \Big(\int_0^t |dA|_s\Big)^s\Big] < \infty\,. \tag{3.52}$$

Then in $L^s(\Omega)$,

$$\begin{aligned}N_X^{x,x+\varepsilon} &\sim \frac{L_X^x(t)}{2\varepsilon} + O(1) \quad \textit{as } \varepsilon \to 0\\ N_X^{\varepsilon} &\sim \frac{[X]_t}{2\varepsilon^2} + O(\varepsilon^{-1}) \quad \textit{as } \varepsilon \to 0\,.\end{aligned}$$

Proof. Since $0 \leq N_X^{x,x+\varepsilon} - D_X^{x,x+\varepsilon} \leq 1$, the statement of proposition 3.7 is applicable to $N^{x,x+\varepsilon}$ yielding the first result. It follows that in $L^s(\Omega)$,

$$\left|2\varepsilon\lambda(T_X^\varepsilon) - \int_{\mathbb{R}} L_X^x(t)\, dx\right| = O(\varepsilon)\,. \tag{3.53}$$

Note that while the integral carries over $\mathbb{R}$, its support is contained within the compact $[\inf X, \sup X]$. We now use the density occupation formula for the local time where $\phi = 1$, so

$$2\varepsilon\lambda(T_X^\varepsilon) = [X]_t + O(\varepsilon) \quad \text{ as } \varepsilon \to 0 \text{ in } L^s(\Omega) \tag{3.54}$$

We now use the fact that

$$\lambda(T_X^\varepsilon) = \int_\varepsilon^\infty N^a\, da\,. \tag{3.55}$$

By monotonicity of N^ε, for every $\delta > 0$ small enough, almost surely,

$$N^{\varepsilon(1+\delta)} \leq \frac{\lambda(T_X^\varepsilon) - \lambda(T_X^{\varepsilon(1+\delta)})}{\delta\varepsilon} \leq \frac{1}{\delta\varepsilon}\int_\varepsilon^{\varepsilon(1+\delta)} N^a\, da \leq N^\varepsilon\,. \tag{3.56}$$

It follows that in $L^s(\Omega)$, for every $\delta > 0$ small enough,

$$\frac{[X]_t}{1+\delta} + O(\varepsilon) \leq 2\varepsilon^2 N^\varepsilon \leq \frac{[X]_t}{1-\delta} + O(\varepsilon) \quad \text{ as } \varepsilon \to 0\,. \tag{3.57}$$

Taking $\delta \to 0$, we obtain the desired result. ■

4 Limiting processes

An interesting point of view relative to these stability results and the results of this paper is to consider these irregular processes as almost sure C^0-limits of smooth processes, which have traditionally been more difficult to study. In this way, we can make affirmations about the barcodes of smooth processes up to some (small) error. This way of thinking is inspired by the study of trees, ultralimits and asymptotic cones in geometric group theory [89].

Proposition 4.1. *Let (M, d) be a compact Polish metric space and let X be an almost surely continuous stochastic process on M, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $(X_n)_{n\in\mathbb{N}}$ be any sequence of continuous stochastic processes defined on the same probability space and suppose*

$$\delta_n := \|X - X_n\|_{L^\infty(\Omega, L^\infty(M,\mathbb{R}))} \xrightarrow[n\to\infty]{} 0\,. \tag{3.58}$$

If for all $\varepsilon > 0$, $\mathbb{E}[N_X^\varepsilon] < \infty$ and is continuous in ε, then for any $\varepsilon \geq 2\delta_n$,

$$N_{X_n}^\varepsilon \xrightarrow[n\to\infty]{L^1} N_X^\varepsilon\,,$$

Morever,

$$\mathbb{E}[|N_X^\varepsilon - N_{X_n}^\varepsilon|] \leq \omega_\varepsilon(\delta_n)\,, \tag{3.59}$$

where $\omega_\varepsilon(\delta) := \mathbb{E}\Big[N_X^{\varepsilon-\delta} - N_X^{\varepsilon+\delta}\Big]$. *With analogous hypotheses, the same statement holds for* $N^{x,x+\varepsilon}$.

Proof. The L^∞-stability of barcodes with respect to the L^∞-distance tells us that δ_n controls the bottleneck distance between the two barcodes [24, 76]. This implies that there exists a δ_n-matching between the barcodes of X_n and that of X, therefore:

- If $\alpha \in \mathcal{B}(X)$ has length $|\alpha| \geq 2\delta_n$, then $\exists!\, \beta \in \mathcal{B}(X_n)$ such that (α, β) are matched and the difference $||\alpha| - |\beta|| \leq 2\delta_n$;
- If $\beta \in \mathcal{B}(X_n)$ has length $|\beta| \geq 2\delta_n$, then $\exists!\, \alpha \in \mathcal{B}(X)$ such that (α, β) are matched and the difference $||\alpha| - |\beta|| \leq 2\delta_n$;
- If $\alpha \in \mathcal{B}(X)$ or $\beta \in \mathcal{B}(X_n)$ in unmatched, then they have length $\leq 2\delta_n$.

It follows that for $\varepsilon \geq 2\delta_n$ we have inequalities

$$\begin{aligned} N_{X_n}^{\varepsilon+2\delta_n} &\leq N_X^\varepsilon \\ N_X^{\varepsilon+2\delta_n} &\leq N_{X_n}^\varepsilon \,, \end{aligned}$$

from which we obtain bounds on $N_{X_n}^\varepsilon$

$$N_X^{\varepsilon+2\delta_n} \leq N_{X_n}^\varepsilon \leq N_X^{\varepsilon-2\delta_n} \,. \tag{3.60}$$

These inequalities imply

$$|N_{X_n}^\varepsilon - N_X^\varepsilon| \leq \Big|N_X^{\varepsilon+2\delta_n} - N_X^\varepsilon\Big| \vee \Big|N_X^{\varepsilon-2\delta_n} - N_X^\varepsilon\Big|$$

Taking expectations of both sides, we have

$$\begin{aligned} \mathbb{E}[|N_{X_n}^\varepsilon - N_X^\varepsilon|] &\leq \mathbb{E}\Big[\Big|N_X^{\varepsilon+2\delta_n} - N_X^\varepsilon\Big| \vee \Big|N_X^{\varepsilon-2\delta_n} - N_X^\varepsilon\Big|\Big] \\ &\leq \mathbb{E}\Big[\Big|N_X^{\varepsilon+2\delta_n} - N_X^\varepsilon\Big|\Big] + \mathbb{E}\Big[\Big|N_X^{\varepsilon-2\delta_n} - N_X^\varepsilon\Big|\Big] \\ &= \mathbb{E}\Big[N_X^\varepsilon - N_X^{\varepsilon+2\delta_n}\Big] + \mathbb{E}\Big[N_X^{\varepsilon-2\delta_n} - N_X^\varepsilon\Big] \\ &\leq \omega_\varepsilon(2\delta_n) \end{aligned}$$

by monotonicity of N_X^ε. The right hand side of the inequality tends to 0 as $n \to \infty$ by continuity of $\mathbb{E}[N_X^\varepsilon]$, so $N_{X_n}^\varepsilon \xrightarrow[n\to\infty]{L^1} N_X^\varepsilon$. ∎

Remark 4.2. The condition of uniform convergence over Ω can be quite restrictive, but covers some cases of distributions with compact supports in $C^0(X, \mathbb{R})$. Moreover, it covers the case of processes derived from random point clouds P (the filtration function is given by $d(-, P)$) stemming from a distribution with compact support on any Polish metric space M.

We can adapt the proof of the above proposition to a setting relaxing the $L^\infty(\Omega, L^\infty(M, \mathbb{R}))$ convergence condition.

Proposition 4.3. *Keeping the same notation as before, suppose there exists a $p \geq 1$ such that*

$$\delta_n := \|X - X_n\|_{L^p(\Omega, L^\infty(M,\mathbb{R}))} \xrightarrow[n\to\infty]{} 0\,. \tag{3.61}$$

Then, with probability $\geq 1 - \frac{1}{a^p}$, for every $\varepsilon \geq 2a\delta_n$,

$$\left|N^\varepsilon_{X_n} - N^\varepsilon_X\right| \leq N_X^{\varepsilon - 2a\delta_n} - N_X^{\varepsilon+2a\delta_n}\,. \tag{3.62}$$

If $\mathbb{E}[N^\varepsilon_X]$ is continuous in ε,

$$\mathbb{E}\Big[\left|N^\varepsilon_{X_n} - N^\varepsilon_X\right| \;\Big|\; \|X - X_n\|_\infty \leq a\delta_n\Big] \leq \omega_\varepsilon(2a\delta_n)\,. \tag{3.63}$$

Moreover,

$$\mathbb{P}(\left|N^\varepsilon_{X_n} - N^\varepsilon_X\right| \geq k) \leq \frac{\omega_\varepsilon(2a\delta_n)}{k} + \frac{1}{a^p} \quad \textit{and} \quad \mathbb{P}(\left|N^\varepsilon_{X_n} - N^\varepsilon_X\right| \geq k\;,\; \|X - X_n\|_\infty \leq a\delta_n) \leq \frac{\omega_\varepsilon(2a\delta_n)}{k}\,. \tag{3.64}$$

Proof. The proof goes as the previous one for a few minor exceptions. To obtain the probable bound on $\|X - X_n\|_{L^\infty(M,\mathbb{R})}$, we apply the Markov inequality. The rest of the proof follows by noticing that the conditional expectation is a contractive projection on L^1 and that

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) \leq \mathbb{P}(A \cap B) + \mathbb{P}(B^c)\,. \tag{3.65}$$

■

Remark 4.4. By an application of the Borel-Cantelli lemma, if X_n tends to X in $L^p(\Omega, L^\infty(M,\mathbb{R}))$ for $p > 1$ at a rate $O(r(n))$ (where r is a function tending to 0 at infinity), then almost surely $X_n \xrightarrow[n\to\infty]{L^\infty} X$ at a rate $O(r(n))$ as well.

5 Applications

5.1 Brownian motion and local martingales with deterministic strictly increasing quadratic variation

For Brownian motion, it is possible to compute our quantities of interest exactly.

Theorem 5.1 (Perez, Proposition 4.4 [79])**.** *For Brownian motion B on $[0,t]$, $\mathbb{E}[N^\varepsilon_B]$ admits the following series representations which converge well for large and small ε respectively*

$$\begin{aligned}
\mathbb{E}[N^\varepsilon_B] &= 4\sum_{k\geq 1}(2k-1)\,\mathrm{erfc}\Big(\frac{(2k-1)\varepsilon}{\sqrt{2t}}\Big) - k\;\mathrm{erfc}\Big(\frac{2k\varepsilon}{\sqrt{2t}}\Big) \\
&= \frac{t}{2\varepsilon^2} + \frac{2}{3} + 2\sum_{k\geq 1}(2(-1)^k - 1)\frac{e^{-\pi^2k^2t/2\varepsilon^2}t}{\varepsilon^2}\left[1 + \frac{\varepsilon^2}{\pi^2k^2t}\right]\,.
\end{aligned}$$

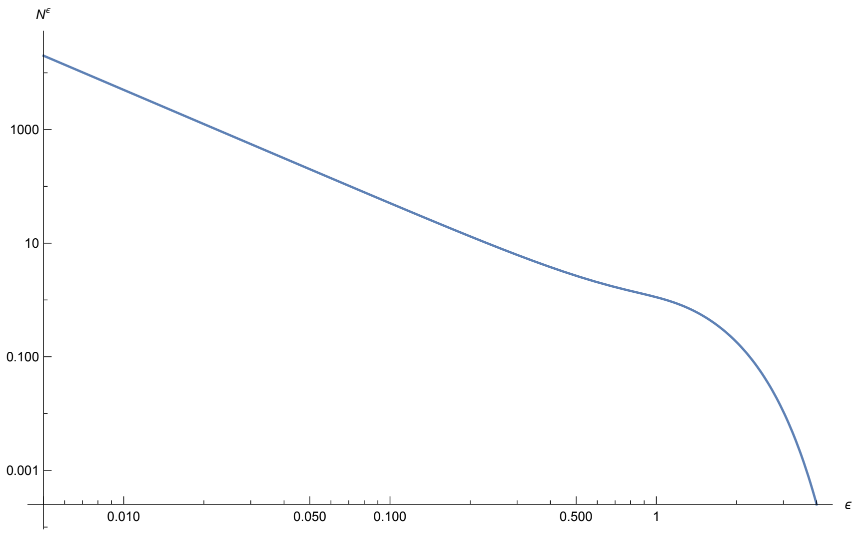


Figure 3.4: The expected number of bars of length $\geq \varepsilon$, $\mathbb{E}[N_B^\varepsilon]$, as a function of ε.

Remark 5.2. In particular, $\mathbb{E}[N_B^\varepsilon]$ is analytic in ε on an open wedge around the positive real axis in the complex plane, thereby confirming and extending Divol and Chazal's results.

Proof. We start by writing

$$\mathbb{E}[N_B^\varepsilon] = \sum_{k=1}^{\infty} \mathbb{P}(N^\varepsilon \geq k) = \mathbb{P}(R_t \geq \varepsilon) + \sum_{k=2}^{\infty} \mathbb{P}(S_{k-1}^\varepsilon \leq t)$$

Using the scale invariance of Brownian motion,

$$\mathbb{E}[N_B^\varepsilon] = \mathbb{P}(R_1 \geq t^{-\frac{1}{2}}\varepsilon) + \sum_{k=2}^{\infty} \mathbb{P}(S_{k-1}^1 \leq t/\varepsilon^2) \tag{3.66}$$

We now notice that the stopping times

$$S_{k-1}^\varepsilon = \sum_{i=1}^{k-1} (S_i^\varepsilon - T_i^\varepsilon) + (T_i^\varepsilon - S_{i-1}^\varepsilon) \tag{3.67}$$

and that the increments $(S_i^\varepsilon - T_i^\varepsilon)$ and $(T_i^\varepsilon - S_{i-1}^\varepsilon)$ are independent and identically distributed. Moreover, in distribution,

$$\sup_{[0,t]} B - B_t = |B_t| \ , \tag{3.68}$$

so that the $(S_i^\varepsilon - T_i^\varepsilon)$ and $(T_i^\varepsilon - S_{i-1}^\varepsilon)$ are distributed like the first hitting time of ε by $|B_t|$. It is a classical result [15, p.641] that this hitting time H^ε satisfies

$$\mathbb{E}\left[e^{\lambda H^\varepsilon}\right] = \operatorname{sech}(\varepsilon\sqrt{2\lambda})\,. \tag{3.69}$$

Similarly, it is also well-known that the range of Brownian motion satisfies

$$\mathcal{L}_t(\mathbb{P}(R_t \geq \varepsilon))(\lambda) = \frac{\operatorname{sech}^2(\varepsilon\sqrt{\frac{\lambda}{2}})}{\lambda} \tag{3.70}$$

We now take the Laplace transform with respect to the time variable of $\mathbb{E}[N_B^\varepsilon]$

$$\begin{aligned}\mathcal{L}_t(\mathbb{E}[N_B^\varepsilon])(\lambda) &= \frac{\operatorname{sech}^2(\varepsilon\sqrt{\frac{\lambda}{2}})}{\lambda} + \frac{1}{\lambda}\sum_{k=2}^{\infty}\mathbb{E}\Big[e^{-\lambda\varepsilon^2 S^1_{k-1}}\Big] \\ &= \frac{\operatorname{sech}^2(\varepsilon\sqrt{\frac{\lambda}{2}})}{\lambda} + \frac{1}{\lambda}\sum_{k=2}^{\infty}\mathbb{E}\Big[e^{-\lambda\varepsilon^2 H^1}\Big]^{2(k-1)},\end{aligned}$$

where the last equality holds by virtue of i.i.d. character of the increments $(S_i^\varepsilon - T_i^\varepsilon)$ and $(T_i^\varepsilon - S_{i-1}^\varepsilon)$. Replacing the value of $\mathbb{E}\Big[e^{\lambda H^\varepsilon}\Big]$,

$$\mathcal{L}_t(\mathbb{E}[N_B^\varepsilon])(\lambda) = \frac{\operatorname{sech}^2(\varepsilon\sqrt{\frac{\lambda}{2}}) + \operatorname{csch}^2(\varepsilon\sqrt{2\lambda})}{\lambda}$$

The result is obtained by taking the inverse Laplace transform. ■

Persistent homology is invariant under reparametrization. In particular if $\lambda : [0,t] \to [0,\lambda(t)]$ is an increasing bijection, $N_{f\circ\lambda}^\varepsilon$ as calculated on $[0,\lambda(t)]$ is exactly equal to N_f^ε as calculated on $[0,t]$. Invoking the Dambis-Dubins-Schwarz theorem (theorem 3.4), for every continuous local martingale M (such that $M_0 = 0$) with deterministic and strictly increasing quadratic variation $[M]_t$, almost surely,

$$N_M^\varepsilon[0,t] = N_B^\varepsilon[0,[M]_t]\,. \tag{3.71}$$

Theorem 5.3. *For any continuous local martingale M on $[0,t]$ having deterministic and strictly increasing quadratic variation $[M]_t$ such that $[M]_\infty = \infty$, the same formula holds by replacing t by $[M]_t$ in the result of theorem 5.1.*

We can perform a similar calculation for $N^{x,x+\varepsilon}$ when $x > 0$,

Proposition 5.4 (Proposition 4.7)**.** *For Brownian motion on $[0,t]$ and $x > 0$,*

$$\begin{aligned}\mathbb{E}\Big[N_B^{x,x+\varepsilon}\Big] &= \sum_{k=1}^{\infty} erfc\left(\frac{x+(2k-1)\varepsilon}{\sqrt{2t}}\right) \\ &\sim \frac{1}{2\varepsilon}\int_0^t \varphi(x,s)\,ds + \sum_{k\geq 0}\frac{4(-2)^k\Big(2^{2k+1}-1\Big)\zeta(2k+2)}{\pi^{2k+2}}\left[\frac{\partial^k}{\partial t^k}\varphi(x,t)\right]\varepsilon^{2k+1} \;\; as\ \varepsilon \to 0\,,\end{aligned}$$

where $\varphi(x,t)$ is the density of a centered Gaussian random variable of variance t and ζ denotes the Riemann zeta function.

Remark 5.5. In the appropriate limits, we retrieve the results established for semimartingales. By virtue of the Dambis-Dubins-Schwarz theorem and by the invariance under reparametrization of persistent homology, we may also immediately extend the results and formulæ of [79] to local martingales with deterministic and strictly increasing quadratic variation by replacing $t \mapsto [M]_t$.

5.2 Itô processes

Definition 5.6. An **Itô process** is the solution to a stochastic differential equation of the form

$$dX_t = \mu_t \, dt + \sigma_t \, dB_t \tag{3.72}$$

where μ_t and σ_t are adapted processes.

Every Itô process of this form has the strong Markov property and moreover, it is a semi-martingale. We deduce that

$$N_X^\varepsilon \sim \frac{[X]_t}{2\varepsilon^2} + O(\varepsilon^{-1}) \text{ as } \varepsilon \to 0\,, \tag{3.73}$$

and an analogous expression for $N^{x,x+\varepsilon}$ in terms of the local time of X. In expectation, we may always say something about $N_X^{x,x+\varepsilon}$ by virtue of the following proposition.

Proposition 5.7. *Let X be an Itô process with deterministic quadratic variation, then*

$$\mathbb{E}[L_X^x(t)] = \int_0^t \varphi_X(x,s)\; d[X]_s\,, \tag{3.74}$$

where $\varphi_X(-,s)$ is the density function of the random variable X_s (which itself is a solution of the Fokker-Planck PDE associated to the SDE).

Proof. We take the expectation of both sides of the occupation density formula with $\phi(a) = e^{-i\lambda a}$.

$$\int_0^t \mathbb{E}\Big[e^{-i\lambda X_s}\Big]\, d[X]_s = \int_{\mathbb{R}} e^{-i\lambda a}\, \mathbb{E}[L_X^a(t)] \;\, da \tag{3.75}$$

The result follows from taking the inverse Fourier transform of both sides. ∎

Proposition 5.8 (Asymptotics of barcodes of Itô processes)**.** *Let X be an Itô process with deterministic quadratic variation on $[0,t]$, then,*

$$\mathbb{E}\Big[N^{x,x+\varepsilon}\Big] \sim \frac{1}{2\varepsilon}\int_0^t \varphi_X(x,s)\; d[X]_s + O(1) \quad \text{as } \varepsilon \to 0\,. \tag{3.76}$$

where $\varphi_X(-,s)$ is the density function of the random variable X_s.

∎

As a Markov processes, Itô processes also satisfy

$$\mathbb{E}[N_X^\varepsilon] \sim \mathbb{P}(R_t \geq \varepsilon) \quad \text{as } \varepsilon \to \infty\,. \tag{3.77}$$

Finally, when $\mu_t = 0$ and $\sigma_t > 0$ for all $t > 0$ and is deterministic, an Itô process is a local martingale with quadratic variation

$$[X]_t = \int_0^t \sigma_s^2 \, ds \,, \tag{3.78}$$

and, in particular, its barcode is then known completely.

5.3 Limiting processes

The Karhunen-Loève theorem [55, 64] shows that many well-known C^0-processes can be seen as limits of smooth processes and that the logic of studying smooth objects with their more irregular C^0 limits is a logic which can also be applicable in higher dimensions. Brownian motion itself can be seen as such a limit [61, 72]. Indeed, Paul Lévy showed that if $(\xi_k)_{k\in\mathbb{N}}$ is a sequence of i.i.d. standard normal variables, the series

$$\xi_0 t + \frac{\sqrt{2}}{\pi} \sum_{k=1}^{\infty} \xi_k \frac{\sin(\pi k t)}{k} \tag{3.79}$$

almost surely uniformly converges to the standard Brownian motion on $[0,1]$. Noting

$$S_n B := \xi_0 t + \frac{\sqrt{2}}{\pi} \sum_{k=1}^{n-1} \xi_k \frac{\sin(\pi k t)}{k} \,, \tag{3.80}$$

it can be shown (for instance using [54, Chapter 15, Theorem 4]) that for $p \geq 1$ there exists a finite constant C_p such that

$$\|B - S_n B\|_{L^p(\Omega, L^\infty([0,1],\mathbb{R}))} \leq C_p n^{-\frac{1}{2}} \log^{\frac{1}{2}}(n) \,. \tag{3.81}$$

Applying the results of proposition 4.3 and optimizing in a, we have

$$\mathbb{P}(|N^\varepsilon_{S_n B} - N^\varepsilon_B| \geq k) = O\left(\left[\frac{C_p \, n^{-\frac{1}{2}} \log^{\frac{1}{2}}(n)}{p \, \varepsilon^3 k} \right]^{\frac{p}{p+1}} \right) \quad \text{as } \varepsilon \to 0. \tag{3.82}$$

In particular, we know that through this approximation yields a curve on the log-chart which doesn't stray far away from the Brownian motion's for $\varepsilon \gtrsim (n^{-1}\log(n))^{\frac{1}{2}}$. Brownian motion can also be approximated by random walks. The Komlós–Major–Tusnády (KMT) theorem provides a sharp estimate of the rate of convergence of these empirical processes to the Brownian bridge (which we will denote W_t).

Definition 5.9. Let $(X_n)_{n\in\mathbb{N}^*}$ be a sequence of (reduced, centered) i.i.d random variables. The **empirical process** defined by X is the process

$$\alpha_n^X(t) := \left[\frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{]-\infty,t]}(X_k) \right] - t \,. \tag{3.83}$$

Theorem 5.10 (KMT Theorem, [57])**.** *Let $(U_n)_{n\in\mathbb{N}^*}$ be an i.i.d. sequence of uniform random variables on $[0,1]$. Then, there exists a Brownian bridge $(W_t)_{1\geq t\geq 0}$ such that for all $n \in \mathbb{N}^*$ and* *all $x > 0$*

$$\mathbb{P}\Big(\Big\|\alpha_n^U(t) - W_t\Big\|_{L^\infty} > n^{-\frac{1}{2}}(C\log(n) + x)\Big) \leq Le^{-\lambda x} \tag{3.84}$$

for some universal positive constants C, L and λ which are explicitly known [16].

Since W is a semimartingale with quadratic variation $[W]_t = t$ on $[0,1]$, as it can be obtained as the solution to the SDE

$$dW_t = \frac{-W_t}{1-t}\, dt + dB_t\,. \tag{3.85}$$

We deduce that $N_W^\varepsilon \sim \frac{1}{2\varepsilon^2}$ as $\varepsilon \to 0$. The KMT theorem and the same reasoning behind proposition 4.3 imply

$$\mathbb{P}\Big(\Big|N^\varepsilon_{\alpha_n^U} - N^\varepsilon_W\Big| \geq k\Big) = O\left(\frac{n^{-\frac{1}{2}}\log(n)}{\varepsilon^3 k}\right) \quad \text{as } \varepsilon \to 0\,. \tag{3.86}$$

In particular, this approximation yields the curve $N^\varepsilon_{\alpha_n^U}$ on the log-chart which doesn't stray far away from N^ε_W's for $\varepsilon \gtrsim n^{-\frac{1}{2}}\log(n)$.

Chapter 4

ζ-functions and the topology of superlevel sets of stochastic processes

Abstract

We introduce the so-called stochastic ζ-functions, defined in terms of Pers_p-functional of the average diagram of a stochastic process X. For Lévy α-stable processes, these ζ-functions always admit a meromorphic extension to the entire complex plane with a single pole at α of known residue. The analytic properties of these ζ-functions are related to the asymptotic expansion of a dual variable counting the number of bars of the barcode of X of size $\geq \varepsilon$, from which we devise a new statistical parameter test.

Contents

1 Introduction

The problem of the characterization of the topology of superlevel sets of random functions has been a long studied topic in the theory of random fields. While a complete description has been thus far unknown, partial descriptors of the topology of superlevel sets, such as their Euler characteristic, have been described for certain classes of random processes [3, 7, 9, 20, 74, 80]. Thus far, the study of the homology of superlevel sets of random functions in dimension one has focused on either smooth random (Gaussian) fields [3, 7], or irregular processes which are in some sense canonical, such as Brownian motion [9, 20, 80]. In this paper, following the universality reasoning detailed in [78, §3], we will adopt the second point of view while enlarging the category of processes considered to objects acting as universal limits of random processes in 1D.

This work is another stage in a program started in [77] and later continued in [78], which aimed to characterize the barcodes of random functions as completely as possible (in dimension one). To do this, we adopted the tree formalism originally developped by Le Gall [37, 38], which brings benefits in the probabilistic setting. This formalism allowed us to partially study the case of Markov or self-similar processes, and to processes admitting the two latter as limits [78]. In this paper, we further develop the theory to describe almost completely the case of (α-stable) Lévy processes.

The understanding of so-called *topological noise* is an active area of research in Topological Data Analysis (TDA) (*cf.* section 2.3 for a quick introduction, or [24, 76] for a more comprehensive one). This topological noise is characterized by the behaviour of the small bars of the barcode of a function and its role is particularly difficult to grasp. Nonetheless, topological noise has proved useful in a variety of different applications, which seem to exploit the information contained therein, sometimes directly, or indirectly through the use of Wasserstein p metrics on the space or functionals such as Pers_p [19, 28, 33, 71, 96], despite the absence of general stability results. The program described above is based around the intuition that while it might be hopeless to find a general stability theorem, there should be a form of *statistical stability* of barcodes.

By statistical stability we mean that any two samples drawn from the same probability distribution should have "close" or "similar" behaviour of its topological noise. For example, this paper shows that, at least in the simple setting of 1D and for a fairly wide range of distributions, this topological noise does indeed exhibit the robustness sought. More precisely, we show that the number of bars of length $\geq \varepsilon$, N^ε, is a dual to Pers_p^p, and that it exhibits the statistical robustness required, such as an almost sure asymptotic behaviour as $\varepsilon \to 0$.

If this notion of statistical stability is a good one, a natural subsequent question is whether it is possible to differentiate stochastic processes given their topological signature. In this paper, we partially answer this question positively through the development of a statistical test constructed with the functional N^ε, which can differentiate α-stable processes for different values of α. In dimension one, while interesting, this development is unlikely to do better than wavelet analysis or other known techniques (*cf.* [32] and the references therein). However, further developments in this direction could eventually lead to "*topological statistics*", *i.e.* robust statistical tests for random fields, for which all known techniques do not generalize, but for which barcodes are easily computable.

This discussion and the results of this paper hint at the fact that statistical stability is a correct notion of stability to consider, and there are many open questions to be tackled, some of these questions are:

- **Dimensionality**: are the statistical stability results of this paper particular to dimension one, or do they generalize in some way to higher dimensional random fields?
- **Signal vs. noise problem**: Given the regular structure of topological noise shown in this paper, does this allow us to detect the presence of an underlying topological signal?
- **Statistical robustness of topological noise**: What can we say, quantitatively, about the variation of topological noise induced by perturbations of the distribution of the noise (for instance in some Wasserstein metric)?
- **Best proxys for topological signatures**: much of this paper was inspired by what is used in practice, namely Wasserstein p metrics and the Pers_p functional and its dual N^ε. However, there is no guarantee this yields the best possible proxys to answer the two previous problems. In this regard, is it possible to prove or disprove which proxys do best in what context?

From a more probabilistic point of view, this paper introduces so-called ζ-functions associated to a stochastic process, constructed using the Pers_p^p functional we previously discussed. The main, and perhaps most important, departure from the conventional TDA theory is that we will consider this quantity for complex p for reasons which will become evident throughout this paper, but which are analogous to the ones behind the complexification of the Riemann ζ-function in analytic number theory. There are some similarities to the results of Pitman, Yor and Biane in [11, 81, 82] regarding the probabilistic interpretation of the ζ-function and more generally L-functions based on connections with some families of infinitely divisible distributions connected to Brownian motion. However, the ζ-functions herein are of different nature to those considered by Pitman, Yor and Biane, as they stem from a different construction. For Brownian motion, we fall back on one of the infinitely divisible distributions considered by these authors. Renewal theory [48] intervenes at many different steps in this paper and combined with the results of Pitman, Yor and Biane, it is *a posteriori* perhaps not surprising that the ζ-functions hereby introduced share some of the analycity properties of the Riemann ζ-function.

1.1 Our contribution

More precisely, our contribution can be split along the following lines:

1. We establish a duality relation with respect to the Mellin transform between the study of Pers_p^p and the number of leaves of a ε-trimmed tree $\geq \varepsilon$, N^ε (*cf.* section 2.5). With the help of a correct notion of integration on trees developped in [80], it is possible to prove an interpolation theorem for Pers_p^p (proposition 2.18);

2. We introduce ζ-functions for stochastic processes and for persistent measures (*cf.* section 2.6 and section 2.7 respectively). We show an interpolation theorem for Wasserstein p-distances between diagrams (proposition 2.40) and a characterization of convergence between diagrams in these Wasserstein distances in terms of the ζ-functions introduced (theorem 2.44).

3. We show that in the context of α-stable Lévy processes, the associated (tail) ζ-functions always admit a meromorphic extension to the entire complex plane, with a unique pole at $p = \alpha$ with known residue (theorem 3.21). By duality, this meromorphic extension implies the existence of an asymptotic series for N^ε as $\varepsilon \to 0$, which we explicitly calculate up to superpolynomial (*i.e.* smaller than any polynomial) corrections (theorem 3.11). An explicit form of the meromorphic continuation of $\hat{\zeta}$ is shown to be related to the superpolynomial corrections to the asymptotic expansion of theorem 3.11 (*cf.* section 2.1). We also define a generating function for the length of the kth longest bar (*cf.* section 3.2);

4. We apply the theory above to different stochastic processes, such as Brownian motion, reflected Brownian motion. We derive explicit formulæ for the respective ζ-functions of these processes and infer the associated asymptotic expansions of N^ε (theorems 4.2, 4.11 propositions 4.4 and 4.12) and in the case of Brownian motion, the explicit distribution of the length of the kth longest bar (*cf.* section 4.1).

5. We design a statistical test for the parameter α of α-stable Lévy processes by using the theory previously described (*cf.* section 3.2);

6. We study local trees and introduce local ζ-functions (*cf.* section 2.6) and deduce formulæ for the number of points contained in the rectangle $]-\infty, x] \times [x+\varepsilon, \infty[$ of the persistence diagram, $N^{x,x+\varepsilon}$, by introducing the notion of propagators, which, for Markov processes, reduces the problem of the study of $N^{x,x+\varepsilon}$ to the study of hitting times of the process (*cf.* section 3.3), in particular, we link the regularity of these propagators with meromorphic extensions of the local ζ-functions of the process (proposition 3.35);

7. Finally, we apply the theory above to different stochastic processes, such as Brownian motion, reflected Brownian motion, Brownian motion with drift and the Ornstein-Uhlenbeck process for which explicit computations are possible. We derive explicit formulæ for the respective (local, global) ζ-functions of these processes and infer the associated asymptotic expansions of $N^{x,x+\varepsilon}$ (theorems 4.6, 4.11 propositions 4.13 and 4.12 and section 4.3). We also infer formulæ regarding the Ornstein-Uhlenbeck process, in particular concerning its local time (*cf.* section 4.4).

2 Generalities

2.1 The Mellin transform

Definition 2.1. Let f be a locally integrable function over the ray $]0,\infty[$. The **Mellin transform** of f is

$$\mathcal{M}[f(x)](s) := \int_0^\infty x^{s-1} f(x)\, dx\,. \tag{4.1}$$

Note that $d\log(x) = \frac{dx}{x}$ is the Haar measure of $(\mathbb{R}_+, \times)$. The Mellin transform reflects the Pontryagin duality with respect to this locally compact abelian group. Its theory is analogous to that of the bilateral Laplace transform, as the map $\log : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$ induces an isomorphism of abelian groups.

Notation 2.2. For convenience, we will also employ the shorthand notation $\mathcal{M}[f](s) = f^*(s)$.

Definition 2.3. The **fundamental strip** of f, $\langle \alpha, \beta \rangle$ is the maximal set

$$\langle \alpha, \beta \rangle := \{ z \in \mathbb{C} \,|\, \alpha < \mathrm{Re}(z) < \beta \} \tag{4.2}$$

where $f^*(s)$ is well defined.

The Mellin transform can be inverted by virtue of the following theorem, which follows from the Laplace inversion theorem.

Theorem 2.4 (Mellin inversion,[36, 75]). *Let f have fundamental strip $\langle \alpha, \beta \rangle$ and let $c \in]\alpha, \beta[$. Then*

1. *If f is integrable and $f^*(c+it)$ is integrable, then for almost every x*

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f^*(s) x^{-s}\, ds \tag{4.3}$$

   If f is continuous, the equality holds everywhere.

2. *If f is locally integrable and of bounded variation in a neighbourhood of x, then*

$$\frac{f(x^+) + f(x^-)}{2} = \lim_{T\to\infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} f^*(s) x^{-s}\, ds \tag{4.4}$$

A sufficient condition for the Mellin transform to be well-defined on $\langle \alpha, \beta \rangle$ is that the function is such that

$$f(x) = O(x^{-\alpha}) \text{ as } x \to 0 \quad \text{and} \quad f(x) = O(x^{-\beta}) \text{ as } x \to \infty\,. \tag{4.5}$$

In fact, Mellin transforms are a good tool to study asymptotic expansions as suggested by the following theorem.

Theorem 2.5 (Fundamental correspondence, [45]). *Let $f :]0, \infty[\to \mathbb{C}$ be a continuous function with non-empty fundamental strip $\langle \alpha, \beta \rangle$. Then,*

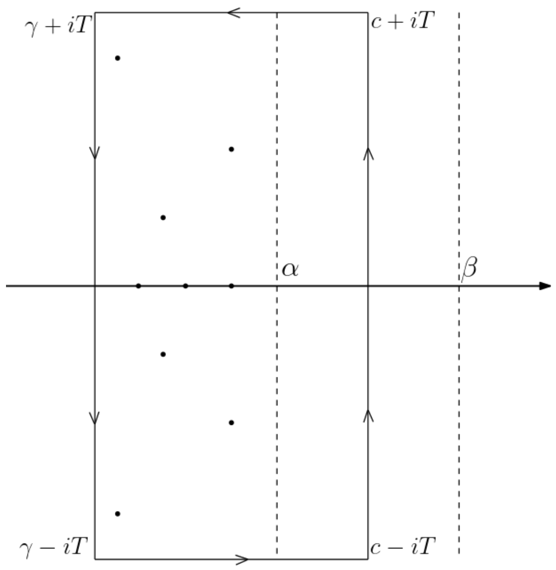


Figure 4.1: Contour for the evaluation of the Bromwich integral of the inverse Mellin transform.

- *Assume that $f^*(s)$ admits a meromorphic continuation to the strip $\langle \gamma, \beta \rangle$ for $\gamma < \alpha$, that it has only a finite amount of poles there and that it is analytic on* $\mathrm{Re}(s) = \gamma$*. Assume also that there exists $\eta \in]\alpha, \beta[$ such that along a denumerable set of horizontal segments with $|\mathrm{Im}(s)| = T_i$ where $T_i \to \infty$,*

$$f^*(s) = O(|s|^{-r}) \text{ with } r > 1 \text{ as } |s| \to \infty \text{ and } s \in \langle \gamma, \eta \rangle . \tag{4.6}$$

  Indexing the poles on $\langle \gamma, \beta \rangle$ by their location ξ and by their order k and denoting $c_{\xi,k}$ the kth coefficient in the Laurent expansion around ξ of $f^*(s)$, we have an asymptotic expansion of f around 0

$$f(x) \sim \sum_{(\xi,k)} c_{\xi,k} \frac{(-1)^{k-1}}{(k-1)!} x^{-\xi} \log^k(x) + O(x^{-\gamma}) \quad \text{as } x \to 0 . \tag{4.7}$$

- *Conversely, if the function f has such an asymptotic expansion around* 0*, then $f^*(s)$ has a meromorphic continuation to the strip $\langle \gamma, \beta \rangle$.*

Furthermore, an analogous statement holds true for asymptotic expansions around ∞ and meromorphic continuations beyond β.

Sketch of proof. It suffices to perform contour integration using the contour of figure 4.1. The estimates of the theorem allow us to discard the top and bottom integrals and to state that the integral of the path along $\mathrm{Re}(p) = \gamma$ is $O(x^{-\gamma})$. Conversely, consider

$$f(x) \sim \sum_{(\xi,k)} c_{\xi,k} x^{\xi} \log^k(x) + O(x^{-\gamma}) \quad \text{as } x \to 0 \tag{4.8}$$

for some $\gamma < \alpha$. It follows that

$$f^*(s) = \sum_{(\xi,k)} c_{\xi,k} \frac{(-1)^k k!}{(s+\xi)^{k+1}} + \int_1^\infty x^{s-1} f(x)\, dx \\ + \int_0^1 x^{s-1} \underbrace{\left[f(x) - \sum_{(\xi,k)} c_{\xi,k}\; x^\xi \log^k(x) \right]}_{=O(x^{-\gamma})} dx\,,$$

which is well-defined on the strip $\langle \gamma, \beta \rangle$. ■

$f(x)$	$f^*(s)$	$\langle \alpha, \beta \rangle$
$x^\nu f(x)$	$f^*(s+\nu)$	$\langle \alpha-\nu, \beta-\nu \rangle$
$f(x^\nu)$	$\frac{1}{\nu} f^*(\frac{s}{\nu})$	$\langle \nu\alpha, \nu\beta \rangle$
$f(x^{-1})$	$f^*(-s)$	$\langle -\beta, -\alpha \rangle$
$f(\lambda x)$	$\lambda^{-s} f^*(s)$	$\langle \alpha, \beta \rangle$
$\frac{\partial}{\partial x} f(x)$	$-(s-1) f^*(s-1)$	
$\int_0^x f(t)\, dt$	$-\frac{1}{s} f^*(s+1)$	

Table 4.1: Functional properties of the Mellin transform

$f(x)$	$f^*(s)$	$\langle \alpha, \beta \rangle$
e^{-x}	$\Gamma(s)$	$\langle 0, \infty \rangle$
e^{-x^2}	$\frac{1}{2}\Gamma(\frac{s}{2})$	$\langle 0, \infty \rangle$
$\mathrm{erfc}(x)$	$2^{-s} \frac{\Gamma(s)}{\Gamma(1+\frac{s}{2})}$	$\langle 0, \infty \rangle$
$\mathrm{csch}(x)$	$2^{1-s}\,(2^s-1)\,\Gamma(s)\zeta(s)$	$\langle 1, \infty \rangle$
$\mathrm{csch}^2(x)$	$2^{2-s}\Gamma(s)\zeta(s-1)$	$\langle 2, \infty \rangle$
$\frac{1}{e^x-1}$	$\Gamma(s)\zeta(s)$	$\langle 1, \infty \rangle$

Table 4.2: A short dictionary of Mellin transforms.

Analytic continuation

As stated by the fundamental correspondence (theorem 2.5), the existence of an asymptotic expansion around 0 of $f(x)$ entails a meromorphic continuation of $f^*(s)$ to a larger strip. If $f(x)$ admits a converging Laurent series (with finite singular part) on some open disk around the origin, then this extension is in fact valid over all of $\mathbb{C}$, and the residues of the poles of $f^*(s)$ will be related to the Laurent coefficients of $f(x)$. It turns out that in this context, one can even write an explicit integral representation for the extension of $f^*(s)$.

Lemma 2.6 (Integral representation of f^*)**.** *Let f be a meromorphic function admitting a Laurent series at 0, with singular part of degree n, holomorphic on a neighbourhood of $\mathbb{R}_+^*$ and integrable over the Hankel contour (cf. figure 4.2). Suppose further that its fundamental strip $\langle n, \beta\rangle$ is non-empty. Then, the function f^* admits a meromorphic continuation on $\langle -\infty, \beta\rangle$ given by*

$$\begin{aligned} f^*(s) &= \frac{e^{-i\pi s}}{2i\sin(\pi s)} \oint_H z^{s-1} f(z)\, dz \\ &= -\frac{\Gamma(s)\Gamma(1-s)}{2\pi i} \oint_H (-z)^{s-1} f(z)\, dz\,, \end{aligned}$$

where H denotes the Hankel contour.

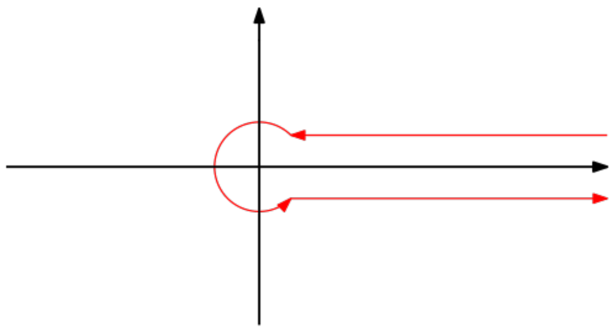


Figure 4.2: The Hankel contour H.

Proof. We start by splitting the Hankel contour into three pieces.

1. A segment from $\infty + i\varepsilon$ to $\nu + i\varepsilon$;
2. A circle C_ν around the origin of radius ν;
3. A segment from $\nu - i\varepsilon$ to $\infty - i\varepsilon$.

For $s \in \langle n, \beta\rangle$, f is holomorphic everywhere on this contour, so that we may take $\varepsilon = 0$ according to Cauchy's theorem. Notice also that

$$\int_{C_\nu} z^{s-1} B(z)\, dz = O(\nu^{\mathrm{Re}(s)-n}) \to 0 \quad \text{as } \nu \to 0\,. \tag{4.9}$$

It follows that for $\mathrm{Re}(s) > n$

$$\begin{aligned} \oint_H z^{s-1} f(z)\, dz &= \lim_{\nu\to 0} \left\{ \int_\infty^\nu + \int_{C_\nu} + \int_{\nu e^{2\pi i}}^{\infty e^{2\pi i}} \right\} z^{s-1} f(z)\, dz \\ &= (e^{2\pi i(s-1)} - 1) \int_0^\infty z^{s-1} f(z)\, dz \\ &= 2ie^{i\pi s} \sin(\pi s) f^*(s)\,, \end{aligned}$$

as desired. The integral over the complex contour H converges for all $s \in \langle -\infty, \beta \rangle \setminus \{n\}$. Finally, the second expression for f^* is obtained through Euler's reflection formula, namely

$$\Gamma(p)\Gamma(1-p) = \frac{\pi}{\sin(\pi p)}\,, \tag{4.10}$$

which after some simplification yields the desired expression. ■

Remark 2.7. If f has fundamental strip $\langle n, \infty \rangle$, then the extension given by this procedure holds over $\mathbb{C}$.

Furthermore, if f posseses a meromorphic continuation to $\mathbb{C} \setminus \mathbb{R}_+$ (*i.e.* we admit the possibility of a branch cut on the positive real axis), then we can find a more explicit formulation for the Hankel representation of f^*.

Lemma 2.8 (Functional equation of f^*). *Suppose f posseses a meromorphic continuation to $\mathbb{C} \setminus \mathbb{R}_+$ and denote $\mathcal{P}$ the set of poles of f not including 0. Suppose further that f has the following decay condition : for all $s \in \langle n, \beta \rangle$ and for some monotone increasing sequence of radii $r_n \to \infty$ as $n \to \infty$,*

$$\int_{C_{r_n,\varepsilon}} \left| z^{s-1} f(z) \right| \, dz \xrightarrow[n\to\infty]{} 0\,. \tag{4.11}$$

where $C_{r_n,\varepsilon}$ is the circle of radius r_n minus a small (symmetric) arc of length ε around the positive real axis (cf. figure 4.3). Then,

$$f^*(s) = \Gamma(s)\Gamma(1-s) \sum_{z_0 \in \mathcal{P}} \mathrm{Res}((-z)^{s-1} f(z); z_0)\,. \tag{4.12}$$

Proof. The proof relies on the use of the residue theorem by completing the Hankel contour into a Pac-Man (*cf.* figure 4.3), whose circular contribution is going to zero, due to the assumption of the lemma. By the residue theorem, we then have

$$f^*(s) = \Gamma(s)\Gamma(1-s) \sum_{z_0 \in \mathcal{P}} \mathrm{Res}((-z)^{s-1} f(z); z_0)\,,$$

as desired. ■

2.2 Connected components of superlevel sets of stochastic processes

Let us briefly recall the construction of a tree from a continuous function $f : [0,1] \to \mathbb{R}$. For a more complete description of this, the reader is welcome to consult [37, 77].

Definition/Proposition 2.9 ([37]). Let $x < y \in [0,1]$, the function

$$d_f(x,y) := f(x) + f(y) - 2 \min_{t \in [x,y]} f(t) \tag{4.13}$$

is a pseudo-distance on $[0,1]$ and the quotient metric space

$$T_f := [0,1]/\{d_f = 0\} \tag{4.14}$$

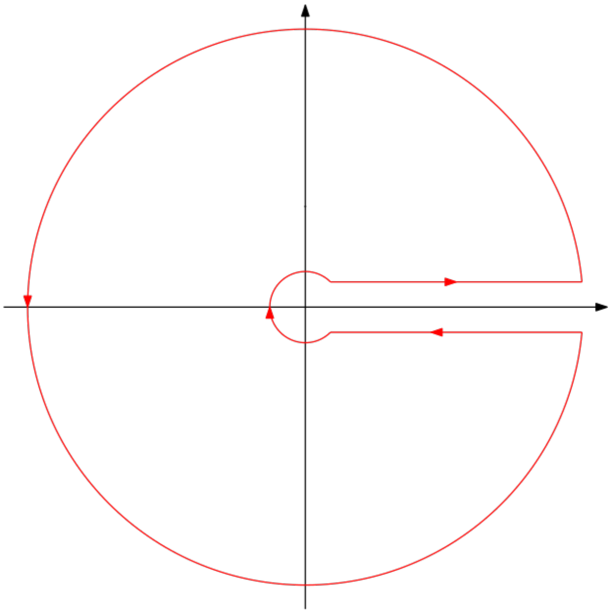


Figure 4.3: The Pac-Man contour.

with distance d_f is a rooted $\mathbb{R}$-tree, whose root coincides with the image in T_f of the point in $[0,1]$ at which f achieves its infimum.

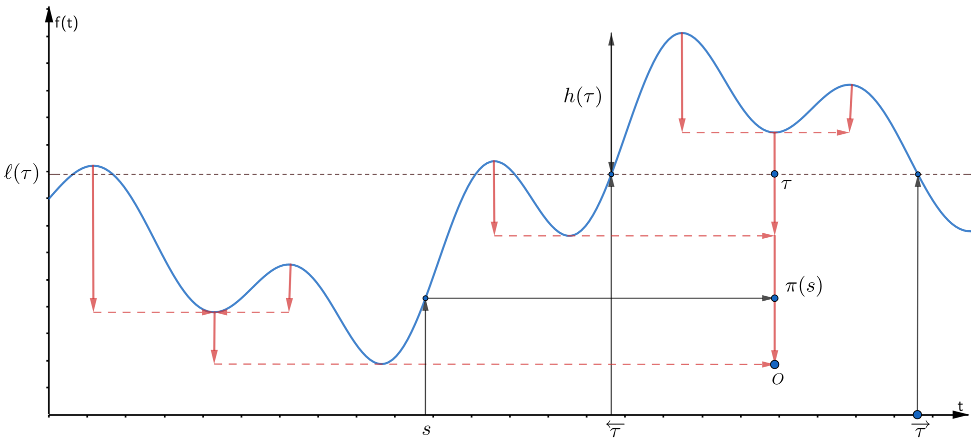


Figure 4.4: A function f and its associated tree T_f in red.

The tree T_f has the particularity that its branches correspond to connected components of the superlevel sets of f, as illustrated by figure 4.4. Let us now introduce the so-called ε-simplified or ε-trimmed tree of T_f^ε. This object is obtained by "giving a haircut" of length ε to T_f. More precisely, if we define a function $h : T_f \to \mathbb{R}$ which to a point $\tau \in T_f$ associates the distance from τ to the highest leaf above τ with respect to the filtration on T_f induced by f, then

Definition 2.10. Let $\varepsilon \geq 0$. An **ε-trimming** or **ε-simplification** of T_f is the metric subspace of T_f defined by

$$T_f^\varepsilon := \{\tau \in T_f \,|\, h(\tau) \geq \varepsilon\} \tag{4.15}$$

Notation 2.11. Let us denote N^ε the number of leaves of T_f^ε.

2.3 A crashcourse in persistent homology

Throughout this section, we will detail and give the ideas behind persistent homology. A proper introduction to this is out of the scope of this paper, so we encourage the reader to consult the following classical references about this topic [24, 49, 65, 76]. This section aims nonetheless to give a brief introduction compiling the main results and intuition behind this field. To do so, it is convenient to break down the topic along its title. First, we will briefly recall what homology is and how it can be defined, and then we will explain the *persistent* aspect of persistent homology. It goes without saying that a reader familiar with these concepts may skip this section entirely.

Homology

In general, the motivation behind the introduction of objects in algebraic topology such as homology is to study topological spaces through algebra. That is, to attach an algebraic object (such as a module, a group, *etc.*) to a topological space, in such a way that, loosely speaking, this algebraic object remains invariant for any two homeomorphic topological spaces. Furthermore, we would like this invariant to behave well with respect to continuous maps. Namely, if we have a continuous map between two topological spaces $f : X \to Y$, we would like to have an induced morphism at the level of the two invariants we attached to M and Y. The most famous such invariant for topological spaces is the fundamental group $\pi_1(X)$ first introduced by Poincaré. Some useful references for further reading are [49, 65].

In the terms of category theory, the above discussion is equivalent to saying that we use *functors* between the category of topological spaces, **Top**, and a category of algebraic objects, such as the category of groups, **Grp**, or that of modules over some ring R, $\mathbf{Mod}_R$. In this sense, homology is nothing other than a functor $H_* : \mathbf{Top} \to \mathbf{Mod}_R$. For our purposes, it is sufficient to consider the ring R to be a field k, so that we are really working over the category of vector spaces over this field $\mathbf{Vect}_k$. We will not detail the precise definitions of these objects, as we will not really need them, but a good reference as an introduction to category theory is given by Mac Lane in [60].

Recall that finding such a functor entails attaching a vector space to a topological space, in such a way that continuous maps between topological spaces induce linear maps at the level of vector spaces. To render this practical, let us first focus on *triangulable* spaces. Namely, spaces which are homeomorphic to a simplicial complex (a set of *oriented* simplices glued to one-another along edges, n-faces or points). Given a simplicial complex M, we can define a so-called *chain complex*, which is nothing other than a sequence of vector spaces, denoted $C_*(M,k)$, called the space of *chains* (the star denotes an index, which we call the *degree*), where C_n is the free k-vector space generated by the set of n-faces in the simplicial complex. For the sake of notational simplicity, whenever the simplicial complex we are talking about and the field over which we are working on is clear, we may drop M and k and denote $C_*(M,k) = C_*$.

We can define a linear map $\partial : C_* \to C_*$ called the *boundary map*, which is defined degree-by-degree as follows. The boundary map ∂ sends the generator of an n-face (an element of C_n) to the (signed) sum of the generators of its boundary (which are elements of C_{n-1}), where the sign in front of each generator is determined by the compatibility of its orientation with the orientation of the n-face. With this definition $\partial^2 = 0$, which reflects the fact that the boundary of a boundary is always empty. So, we can see a chain complex as some graded vector space C_* along with the map ∂. Looking at the restriction of ∂ to each degree of C_*, we can write ∂ as a chain of morphisms

$$\cdots \to C_2 \xrightarrow{\partial} C_1 \xrightarrow{\partial} C_0 \xrightarrow{\partial} 0\,, \tag{4.16}$$

with the property that $\partial^2 = 0$. This property implies in particular that $\mathrm{Im}(\partial) \subset \ker(\partial)$. We call $\ker(\partial)$ the space of *cycles*, reflecting the fact that elements of $\ker(\partial)$ tend to be "loops" or "cycles" of n-chains. Homology quantifies exactly which cycles are not boundaries. More precisely, fixing a degree n, we can define the nth homology group over the field k as

$$H_n(M,k) = \ker(\partial|_{C_n})/\,\mathrm{Im}(\partial|_{C_{n+1}})\,, \tag{4.17}$$

where $\partial|_{C_n}$ denotes the restriction of ∂ to C_n. In some sense, this gives a definition of what we mean by an n-dimensional hole. All of this discussion is best illustrated by an example. In order to avoid sign problems, it is often practical to work over $k = \mathbb{Z}_2$, so as to simplify calculations. In general, this is restrictive, but it is enough for our purposes and let us fix the simplicial complex M of figure 4.5. In this case, we see that $C_0 = \langle A, B, C\rangle_{\mathbb{Z}_2}$ and $C_1 = \langle a, b, c\rangle_{\mathbb{Z}_2}$ with all of the

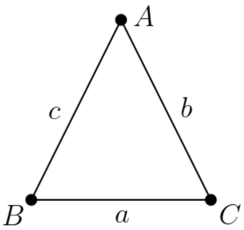


Figure 4.5: An example of a simplicial complex M.

higher chains being 0. Furthermore, the boundary operator sends

$$\partial : a \mapsto B + C\ ,\ b \mapsto A + C\ ,\ c \mapsto A + B\,, \tag{4.18}$$

with all other generators being sent to zero. From this, it is clear that $\ker(\partial|_{C_1}) = (a+b+c)\mathbb{Z}_2$, which represents the cycle which loops around the simplicial complex and $\ker(\partial|_{C_0}) = C_0$. It follows that $H_1(X) = (a+b+c)\mathbb{Z}_2$, so that we indeed detect a hole inside the complex. If we now fill that hole with a cell which fills the triangle (let us note it Δ), we now have $C_2 = \langle \Delta \rangle_{\mathbb{Z}_2}$, and its boundary $\partial(\Delta) = a + b + c$, so that by filling this "hole", we have effectively killed the H_1. As for H_0, it always quantifies the number of connected components, as (finite) simplicial complexes are connected if and only if they are path connected, so these paths constitute cycles which map any two elements of C_0 to each other. In particular, if we suppose that X is path connected, then $\mathrm{Im}(\partial|_{C_1})$ is always generated by all the sums of pairs of vertices in C_0, so we get one generator of H_0 for every connected component.

At this point, given some triangulable space Y, the reader might be worried whether this definition depends on the triangulation we chose for Y, but it is a theorem that this is a well-defined invariant of topological spaces, *cf.* [49] for details. In fact, we have adopted a rather restrictive point of view throughout this discussion, as in reality we can make sense on how homology can be defined in more general settings [65].

Persistent homology

Now that we have roughly sketched out what homology is, let us introduce the idea behind persistent homology. Once again, we refer the reader to consult the following references if he or she desires a more detailed description of the theory [24, 76]. In this case, instead of dealing with a simple chain complex $C_*(M)$, we induce a filtration on this complex, which is typically done by giving a function on the underlying topological space M. For example, if M is a differentiable manifold and $f : M \to \mathbb{R}$ is a smooth function, then we can filter the complex $C_*(M)$ by considering the subsets $M_r = \{f > r\}$ and considering $C_*(M_r)$. We call this filtration of the complex the *superlevel filtration* (an analogous definition can be given for *sublevel filtrations*). Of course, we may then compute the homology of M_r for every r, which gives us a family of vector spaces indexed by r. However, by the functorial nature of H_*, the fact that we have a (continuous) inclusion $i_{r,s} : M_r \hookrightarrow M_s$ for every $r > s$ yields a linear map between the vector spaces $H_*(i_{r,s}) : H_*(M_r) \to H_*(M_s)$.

These morphisms are of interest to us, as they tell us about how the homology changes as we vary the level r. This motivates the study of the so-called *persistent homology*, which is nothing other than the family of vector spaces $(H_*(M_r))_r$ and the family of morphisms $(H_*(i_{r,s}))_{r>s}$. This is more comfortably expressed in the language of category theory : persistent homology is a *functor* $H_* : \mathbb{R} \to \mathbf{Vect}_k$, where $\mathbb{R}$ is seen as a small category (the objects are elements of $\mathbb{R}$ and there is a morphism between $r \to s$ if and only if $r > s$) defined by $H_*(r) = H_*(M_r)$ and $H_*(r \to s) = H_*(i_{r,s})$.

If the function f is nice enough, for instance C^1 and M is compact, the persistent homology induced by the superlevel filtrations of f can be decomposed into so-called *interval modules*. The latter are themselves functors defined as follows. Fixing a field k and if A is an interval of $\mathbb{R}$, then

$$k_A(r) := \begin{cases} k & \text{if } r \in A \\ 0 & \text{else} \end{cases} \qquad k_A(r \to s) = \begin{cases} \mathrm{id} & \text{if } r, s \in A \\ 0 & \text{else} \end{cases} \tag{4.19}$$

The decomposition theorem states the following (*cf.* Oudot's book [76] for a more complete description).

Theorem 2.12 (Decomposition theorem, Auslander, Ringel, Tachikawa, Gabriel, Azumaya)**.** *Under some conditions for M and f, if $H_*(M, f)$ denotes the persistent homology with values in $\mathbf{Vect}_k$, then $H_*(M, f)$ it is isomorphic to a (possibly infinite) direct sum of interval modules. Moreover, this decomposition is unique up to isomorphism and permutation of the terms.*

This theorem entails that if the filtration function f and the space M are nice enough, the persistent homology functor H_* in fact decomposes as a direct sum of interval modules, more precisely, fixing a degree in homology we have

$$H_n = \bigoplus_i k_{A_i}\,, \tag{4.20}$$

where the A_i are intervals of $\mathbb{R}$. Notice then that this means that the information contained in the persistence module can be encoded by these intervals A_i. This collection of intervals is what we call the *barcode* associated to a function $f : M \to \mathbb{R}$, typically denoted $\mathcal{B}(f)$. Another way of representing this information is by keeping track of the endpoints of the interval. In this way, we may represent the intervals as a collection of points in the half-plane

$$\mathcal{X} := \{(x, y) \in (\mathbb{R} \cup \{\pm\infty\})^2 \,|\, x < y\}\,. \tag{4.21}$$

This collection of points is called the *persistence diagram* associated to f, and is typically denoted $\mathrm{Dgm}(f)$.

This is not exactly the full story, as there are some technical caveats to this. Indeed, the theorem requires "nice enough" M and f. Throughout this paper we will be dealing with C^0 funtions (or in C^0, up to a finite number of discontinuities), for which these spaces could be of infinite dimension. Crucially, however, if M is compact and f is C^0, the rank of the maps $H_*(i_{r,s})$ is always finite. Under this condition, the decomposition theorem above applies so we may consider our modules to be decomposable (*cf.* [23, 76] for details).

Specializing all of this to H_0 amounts to talking about connected components of superlevel sets. In this sense, $\dim H_0(M_r)$ is exactly equal to the number of connected components of M_r. The rank of $H_0(i_{r,s})$ corresponds to the number of connected components of M_s which contain all of M_r. The decomposition theorem can also be easily understood in this setting : bars in the barcode (or equivalently points in the persistence diagram) indicate when a connected component was "born" and when it gets absorbed by another one, with the rule that the "eldest" connected component is the one which always "survives".

Persistent homology as a measure

There is one last point to be discussed, which will come in useful in section 2.7. That is, it can be useful to consider persistence diagrams as *measures* over the upper half-plane of $\mathbb{R}^2$. For q-tame modules, we can define such a measure as follows : if $a < b \leq c < d \in \mathbb{R}$, then

$$\mathrm{Dgm}(f)([a, b] \times [c, d]) := \mathrm{rank}(H_*(i_{b,c})) - \mathrm{rank}(H_*(i_{a,c})) + \mathrm{rank}(H_*(i_{b,d})) - \mathrm{rank}(H_*(i_{a,d}))\,. \tag{4.22}$$

This turns out to be a measure, once we have ironed out some details, as done in [24]. As an example, if the module is decomposable, then the *persistence measure* is nothing other than the sum of Dirac masses at every point of the persistence diagram. As we will see, considering persistent diagrams as measures has considerable advantages. The reason for this is because the space of measures is a linear space, so that we can operate on the space of diagrams with greater ease than in the algebraic context. This injection onto a linear space yields some desirable properties for persistence diagrams : for example, the ability to define Fréchet means [96], or a notion of mean and expectation. This suggests that this injection is crucial in simplifying the study of persistence diagrams in a probabilistic setting.

Furthermore, the fact that it is a space of measures allows us to use tools, such as optimal transport. In particular, there are natural notions of distances – so-called Wasserstein distances – which stem from this point of view. This approach has been explored in [33], to which we refer the reader for further details. The idea is to use the diagonal as a source of infinite mass, and study the optimal transport distances between diagrams (now seen as measures), where the distance on the base space (the upper-half plane) is the ℓ^∞-distance on the plane, namely

$$d_{\mathbb{R}^2,\infty}((p,q),(r,s)) = \max\{|p-r|\,, |q-s|\}\,. \tag{4.23}$$

Definition 2.13 (Wasserstein p distances for diagrams)**.** If μ, ν are two persistent measures (to which we have added the diagonal Δ as a source of infinite mass), then the Wasserstein p distance between two diagrams is given by

$$d_p(\mu,\nu) := \left[\inf_{\pi\in\Gamma(\mu,\nu)} \int_{\overline{\mathcal{X}}^2} d^p_{\mathbb{R}^2,\infty}(x,y)\; d\pi(x,y)\right]^{1/p} \tag{4.24}$$

where $\Gamma(\mu,\nu)$ denotes the set of all measures on $\overline{\mathcal{X}}^2$ whose marginals are μ and ν.

If $p=\infty$ we take the sup-norm, which is exactly the so-called bottleneck distance referred to in [24, 76], which we may define as follows, for any two persistent measures μ, ν,

$$d_\infty(\mu,\nu) := \inf_{\pi\in\Gamma(\mu,\nu)} \sup_{(x,y)\in\mathrm{supp}(\pi)} d_{\mathbb{R}^2,\infty}(x,y)\,. \tag{4.25}$$

An important result testifying of why persistent homology is interesting is that it is a stable construct in the following sense.

Theorem 2.14 (Stability theorem, [24])**.** *Let (X,d) be a compact, triangulable metric space and let $f,g\in C^0(X,\mathbb{R})$, then if we see* $\mathrm{Dgm}(f)$ *and* $\mathrm{Dgm}(g)$ *as persistence measures, then*

$$d_\infty(\mathrm{Dgm}(f),\mathrm{Dgm}(g)) \le \|f-g\|_\infty\,. \tag{4.26}$$

Fixing a smaller functional space, it is possible to prove a stability theorem for d_p as well.

Theorem 2.15 (Wasserstein p stability, [93])**.** *Let (X,d) be a compact, triangulable metric space of dimension D and let* $\mathrm{Lip}(\Lambda,X)$ *denote the space of real valued Λ-Lipschitz functions on X. Then if $f,g\in\mathrm{Lip}(\Lambda,X)$, for every $n\ge D$*

$$d_p(\mathrm{Dgm}_k(f),\mathrm{Dgm}_k(g)) \le C_{X,\Lambda,p}\,\|f-g\|_\infty^{1-\frac{n}{p}}\,, \tag{4.27}$$

where $\mathrm{Dgm}_k(f)$ *denotes the persistence measure associated to the persistent homology in degree*

k of f, $H_k(X, f)$.

2.4 Trees and barcodes

There is a correspondence between trees and barcodes described in full detail in [77]. Starting from T_f, we can look at the longest branch (starting from the root) of T_f. This branch corresponds to the longest bar of $\mathcal{B}(f)$ since branches of T_f correspond to connected components of the superlevel sets of f. Next, we erase this longest branch and, on the remaining (rooted) forest, look for the next longest branch. This will be the second longest bar of the barcode. Proceeding iteratively in this way, we retrieve $\mathcal{B}(f)$. An illustration of this algorithm can be found in figure 3.3.

We can interpret N^ε geometrically as being equal to the number of leaves of T_f^ε. In terms of the barcode, the same N^ε counts the number of bars of length $\geq \varepsilon$ with the caveat that we count the infinite bar as having length equal to the range of f. As we will see, reasoning in terms of trees has some major advantages, so in what will follow we will adopt the following convention

Convention 2.16. The length of the infinite bar of $\mathcal{B}(f)$ will be set to $\sup f - \inf f$.

2.5 Integration on trees and the duality between N^ε and Pers_p^p

Let us recall the following simple remark made in [77]. On a tree T_f, we can define a notion of integration by defining the unique atomless Borel measure λ which is characterized by the property that every geodesic segment on T_f has measure equal to its length. Formally, we can express λ in two ways [80]

$$\lambda = \int_{\mathbb{R}} dx \sum_{\substack{\tau \in T_f \\ f(\tau)=x}} \delta_\tau \quad \text{and} \quad \lambda = \int_0^\infty d\varepsilon \sum_{\substack{\tau \in T_f \\ h(\tau)=\varepsilon}} \delta_\tau \tag{4.28}$$

By using the second way of writing λ, the identity

$$\lambda(T_f^\varepsilon) = \int_\varepsilon^\infty N^a \, da \tag{4.29}$$

is clear, as every sum in the second expression is finite for all $\varepsilon > 0$ and has N^ε terms. Of course, we could very well have written it using the first sum, but this poses the difficulty that if T_f is infinite, so is the sum considered in this formal expression for at least some value of x. However, the restricted sum

$$\sum_{\substack{\tau \in T_f \\ f(\tau)=x \\ h(\tau)\geq\varepsilon}} \delta_\tau \tag{4.30}$$

is finite for all $\varepsilon > 0$ and there are exactly $N^{x,x+\varepsilon}$ terms in this sum. We thus obtain an alternative expression for $\lambda(T_f^\varepsilon)$

$$\lambda(T_f^\varepsilon) = \int_{\mathbb{R}} N^{x,x+\varepsilon} \, dx \,. \tag{4.31}$$

We deduce that more information is contained in $N^{x,x+\varepsilon}$ than in $\lambda(T_f^\varepsilon)$ (and by extension than in N^ε). The calculation above provides the connection between Chazal and Divol and Baryshnikov's functional $N^{x,x+\varepsilon}$ and the functional detailed in this paper N^ε, since N^ε is nothing other than the derivative of $\lambda(T_f^\varepsilon)$.

The study of N^ε is in fact completely equivalent to the study of $\mathrm{Pers}_p^p(f)$. Indeed,

$$\mathrm{Pers}_p^p(f) = p\int_{T_f} h(\tau)^{p-1}\,\lambda(d\tau) = p\int_0^\infty \varepsilon^{p-1} N^\varepsilon\,d\varepsilon\,, \tag{4.32}$$

where $h : T_f \to \mathbb{R}$ associating to $\tau \in T_f$ the distance between τ and the highest leaf (with respect to the filtration of f) above τ in T_f. We immediately recognize the above integral as being the Mellin transform of N^ε. Allowing for complex p, this integral relation can be inverted by virtue of the Mellin inversion theorem, provided that the fundamental strip of N^ε is not empty. For compact intervals and continuous functions f, this fundamental strip is never empty (provided $\mathcal{L}(f) < \infty$) and in fact is exactly equal to $\langle \mathcal{L}(f), \infty\rangle$. Thus, for any real number $c > \mathcal{L}(f)$,

$$N^\varepsilon = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \mathrm{Pers}_p^p(f)\,\varepsilon^{-p}\,\frac{dp}{p}\,, \tag{4.33}$$

which estabilished the duality relation desired. Notice also that Pers_p^p is a norm in the sense that

$$\mathrm{Pers}_p^p(f) = p\,\|h\|_{L^{p-1}(\lambda)}^{p-1}\,, \tag{4.34}$$

For any (deterministic) continuous function f, $\mathrm{Pers}_p^p(f)$ is nothing other than a sum of the bars of the barcode to the power p. An in depth explanation of this is provided in [77, §2.2], but let us briefly give some intuition for this. By the algorithm depicted in figure 3.3, if we denote b any of the bars of the barcode, seen as embedded in the tree T_f the length of the branch, $\ell(b)$, can be written as

$$p\int_b h(\tau)^{p-1}\lambda(d\tau) = \ell(b)^p\,. \tag{4.35}$$

The bars of the barcode partition the tree T_f, so that the integration present in the definition of Pers_p^p is nothing other than the sum of the $\ell(b)^p$'s.

Remark 2.17. This definition of Pers_p^p coincides perfectly with a definition of Pers_p^p typically used in persistent homology [19, 28, 33, 71, 96], as long as we consider that the infinite bar has the length of the range (*i.e.* the $\sup - \inf$) of the function f. Of course, within this framework an equally valid definition for Pers_p^p would have been to exclude the infinite bar from being counted all-together, and to consider only the bars of finite length. This approach turns out to give the correct definition for the Pers_p^p-functional in the definition of *tail* ζ-functions (*cf.* definition 3.20), which is necessary to study Lévy α-stable processes for $\alpha < 2$.

Additionally, by the usual inequalities of L^p-spaces,

Proposition 2.18. Pers_p^p *is almost* log*-convex, i.e. let* $p_0 < p_1$ *and* $\theta \in [0,1]$ *and set* $p = (1-\theta)p_0 + \theta p_1$*, then,*

$$\mathrm{Pers}_p^p \leq \frac{p}{p_0^{1-\theta}p_1^\theta}\ \mathrm{Pers}_{p_0}^{p_0(1-\theta)}\ \mathrm{Pers}_{p_1}^{p_1\theta}\,. \tag{4.36}$$

Proof. The statement follows directly from an application of Lyapunov's inequality for L^p-spaces. ■

More generally, it is always true that one can express the $L^p(\mu)$-norm of a function f as the Mellin transform of the repartition function of $|f|$, $\mu(|f| > x)$.

Calculation of N^ε in dimension one

In dimension one, it is possible to use the total order of $\mathbb{R}$ and count N^ε by counting the number of times we go up by at least ε from a local minimum and down by at least ε from a local maximum. This idea can be formalized by the following sequence, originally introduced by Neveu *et al.* [74].

Definition 2.19. Setting $S_0^\varepsilon = T_0^\varepsilon = 0$, we define a sequence of times recursively

$$T_{i+1}^\varepsilon := \inf\left\{ t \geq S_i^\varepsilon \;\middle|\; \sup_{[S_i^\varepsilon, t]} f - f(t) > \varepsilon \right\}$$

$$S_{i+1}^\varepsilon := \inf\left\{ t \geq T_{i+1}^\varepsilon \;\middle|\; f(t) - \inf_{[T_{i+1}^\varepsilon, t]} f > \varepsilon \right\}$$

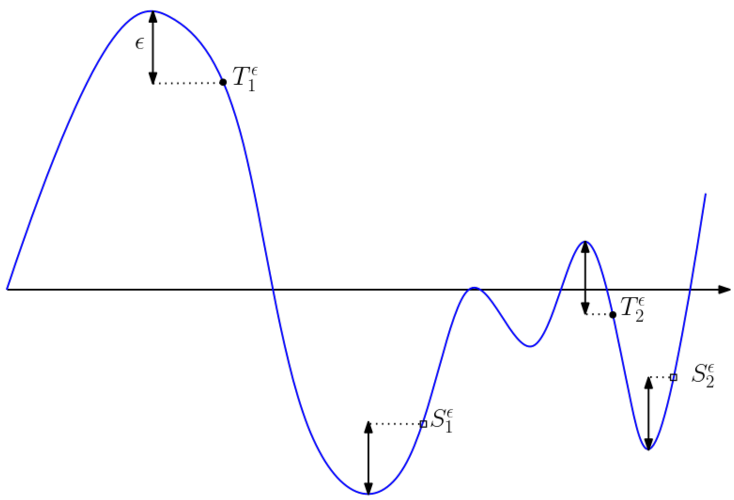


Figure 4.6: A function f along with the times T_i^ε and S_i^ε indicated. Because of the boundary this function has exactly 3 bars of length $\geq \varepsilon$ and not just 2.

Counting the number of bars of length ε is thus exactly to count the number of up and downs we make. More precisely

$$N^\varepsilon = \inf\{i \,|\, T_i^\varepsilon \text{ or } S_i^\varepsilon = \inf \emptyset\} \tag{4.37}$$

by which we mean that it is the smallest i such that the set over which T_i^ε or S_i^ε are defined as infima is empty.

Notation 2.20. We denote the range of X R. Symbolically,

$$R_t := \sup_{[0,t]} X - \inf_{[0,t]} X \,. \tag{4.38}$$

Moreoever, denote N_t^ε the number N^ε of the process X restricted to the interval $[0,t]$.

Intuitively, this calculation process hints at the fact that if ε is small, the number of bars N^ε should strongly depend on the regularity of the process, as ultimately N^ε counts the number of "oscillations" of size ε. In a very precise sense, regularity almost fully determines the asymptotics of N^ε in the $\varepsilon \to 0$ regime. This intuition is corroborated by theorem 1.1. For the rest of this paper it is exactly the functional Pers_p^p which shall occupy us.

Remark 2.21. The Pers_∞ functional is stable under L^∞ perturbations of f. However, it is unknown whether a similar stability result exists for $p < \infty$.

2.6 ζ-functions associated to stochastic processes

Definition 2.22. Let f be a stochastic process on some compact topological space X. Its **ζ-function** ζ_f is defined by:

$$\zeta_f(p) := \mathbb{E}\Big[\mathrm{Pers}_p^p(f)\Big] = p \int_0^\infty \varepsilon^{p-1} \mathbb{E}[N^\varepsilon] \; d\varepsilon \,. \tag{4.39}$$

for $p \in \langle \mathcal{L}(f), \infty \rangle$.

This is reminiscent of the structure of the ζ-function, but *a priori* not enough to draw any parallels. However, it turns out that this nomenclature turns out to have a meaning for stochastic processes. Similarly, we could consider the Pers_p^p functional of $T_f^{>x}$, which denotes the forest

$$T_f^{>x} = \{\tau \in T_f \mid f(\tau) > x\} \tag{4.40}$$

Remark 2.23. In the tree setting the number $N^{x,x+\varepsilon}$ is also the number of branches of length $\geq \varepsilon$ in the forest $T_f^{>x}$.

Following this analogy, it is natural to define

Definition 2.24. The **local ζ-function** associated to f at x, ζ_f^x is defined as

$$\zeta_f^x(p) := \mathbb{E}\Big[\mathrm{Pers}_p^p(T_f^{>x})\Big] = p \int_0^\infty \varepsilon^{p-1} \mathbb{E}\Big[N^{x,x+\varepsilon}\Big] \; d\varepsilon \,. \tag{4.41}$$

Proposition 2.25. *The two ζ-functions we have so far defined are related via the following formula*

$$\zeta_f(p) = p \int_{\mathbb{R}} \zeta_f^x(p-1) \; dx \,. \tag{4.42}$$

Proof. Let us start by noticing that N^ε is nothing other than

$$N^\varepsilon = -\frac{\partial}{\partial \varepsilon} \lambda(T_f^\varepsilon) = -\frac{\partial}{\partial \varepsilon} \int_{\mathbb{R}} N^{x,x+\varepsilon} \; dx \,, \tag{4.43}$$

where this derivative is defined and locally constant almost everywhere. From this and the fact that the derivative $\frac{\partial N^{x,x+\varepsilon}}{\partial \varepsilon}$ is also defined and locally constant almost everywhere and

$$N^{\varepsilon} = -\int_{\mathbb{R}} \frac{\partial N^{x,x+\varepsilon}}{\partial \varepsilon} \, dx \, . \tag{4.44}$$

Applying the Mellin transform to both sides and applying Tonelli's theorem,

$$\int_0^{\infty} \varepsilon^{p-1} N^{\varepsilon} \, d\varepsilon = -\int_{\mathbb{R}} dx \int_0^{\infty} \varepsilon^{p-1} N^{x,x+\varepsilon} \, d\varepsilon \, . \tag{4.45}$$

From the derivation functional property of the Mellin transform (*cf.* table 4.1) we get

$$\int_0^{\infty} \varepsilon^{p-1} N^{\varepsilon} \, d\varepsilon = \int_{\mathbb{R}} dx \; (p-1) \int_0^{\infty} \varepsilon^{p-2} N^{x,x+\varepsilon} \, d\varepsilon \, . \tag{4.46}$$

Applying the expectation to both sides, multiplying times p and applying Tonelli's theorem once again, we have the desired result, namely,

$$\zeta_f(p) = p \int_{\mathbb{R}} \zeta_f^x(p-1) \, dx \, . \tag{4.47}$$

∎

Remark 2.26. If the process starts at 0, it is in general easy to compute ζ_f^x for $x > 0$, but more challenging to do so for $x < 0$. In what will follow, we will always focus on $x > 0$.

Notation 2.27. For the rest of this paper we will take the following conventions. First, we will sometimes omit the subscript t of N_t^{ε} whenever convenient. The Laplace transform $\mathcal{L}$ is always taken with respect to the variable t and its conjugate variable will always be λ. Similarly, Mellin transforms will always be taken with respect to the variable ε and its conjugate variable will be p.

Lemma 2.28. *Let $f(x,t) : [0,\infty[^2 \to \mathbb{R}_+$ such that the functions $f(x,-)$ and $f(-,t)$ are monotone in their arguments. Then, denoting $\mathcal{L}_t$ the Laplace transform with respect to t,*

$$\mathcal{M}_x \mathcal{L}_t[f] = \mathcal{L}_t \mathcal{M}_x[f] \tag{4.48}$$

Proof. The monotonicity of f in its arguments ensures that f is a measurable, positive function. The statement holds by virtue of Tonelli's theorem. ∎

Remark 2.29. Notice this last lemma is applicable to N_t^{ε}, $\mathbb{E}[N_t^{\varepsilon}]$, $\mathbb{P}(N_t^{\varepsilon} \geq k)$ and other such quantities.

2.7 Average diagrams and characterizations of Wasserstein convergence of diagrams

Notation 2.30. For the rest of this section, denote

$$\mathcal{X} := \{(x,y) \in \mathbb{R}^2 \,|\, y > x\} \text{ and } \Delta := \{(x,x) \in \mathbb{R}^2\} \, , \tag{4.49}$$

which we equip with the metric ℓ^∞-metric on $\mathbb{R}^2$ recalled in equation 4.23. Let $\varepsilon > 0$ and denote $\Delta_\varepsilon \subset \overline{\mathcal{X}}$ an open tubular neighbourhood of radius ε around Δ inside $\overline{\mathcal{X}}$ and denote its complement in $\mathcal{X}$ by Δ_ε^c. Finally, denote

$$R_{x,\varepsilon} :=]-\infty, x] \times [x+\varepsilon, \infty[\,. \tag{4.50}$$

By looking at persistence diagrams of functions as measures (*cf.* section 2.3 for details), it is possible to define a notion of an average diagram of a stochastic process.

Definition 2.31. Let M be a compact, triangulable metric space, $\mathcal{E}(M)$ be a Polish metric space of (q-tame) functions on M and let $\mathcal{D}$ denote the space of measures on the upper half-plane $\mathcal{X}$. Let $\phi : \mathcal{E}(M) \to \mathcal{D}$ be the map taking $f \mapsto \mathrm{Dgm}(f)$, where $\mathrm{Dgm}(f)$ is seen as a measure, and let $X : \Omega \to \mathcal{E}(M)$ be a stochastic process of law μ on $\mathcal{E}(M)$. Then, the average diagram of X is given by

$$\mathbb{E}[\mathrm{Dgm}(X)] := \mathbb{E}_\mu[\phi] = \int_{\mathcal{E}(M)} \phi(f)\, d\mu(f)\,. \tag{4.51}$$

Remark 2.32. Note that $\mathbb{E}[\mathrm{Dgm}(X)]$ is itself a measure. Whenever this measure is absolutely continuous with respect to the Lebesgue measure on $\mathcal{X}$, there exists $g : \mathcal{X} \to \mathbb{R}$ such that for any test function $f : \mathcal{X} \to \mathbb{R}$

$$\mathbb{E}[\mathrm{Dgm}(X)]\,(f) = \int_{\mathcal{X}} g(x,y) f(x,y)\, dx\, dy\,. \tag{4.52}$$

A sufficient condition to ensure the existence of g is that $\partial_x \partial_y \mathbb{E}[N^{x,y}]$ exists (this is equivalent to requiring the existence of $\partial_x \partial_\varepsilon \mathbb{E}\left[N^{x,x+\varepsilon}\right]$). Instead of using birth-death coordinates, we may also express this density in terms of birth-persistence coordinates. In these coordinates, we will denote this density function by $g(x,\varepsilon)$, somewhat abusing the notation.

Definition 2.33. The space of measures on $\overline{\mathcal{X}}$ with finite Pers_p will be denoted $\mathcal{D}_p$. Symbolically,

$$\mathcal{D}_p := \left\{\mu \in \mathcal{D} \,\middle|\, \mathrm{Pers}_p(\mu) := d_p(\mu, \Delta) < \infty\right\}\,, \tag{4.53}$$

where d_p is the Wasserstein p-distance on the space of diagrams defined in [33] and explicited in definition 2.13.

This allows us to define a ζ-function for $\mu \in \mathcal{D}_p$ as follows.

Definition 2.34. Let $\mu \in \mathcal{D}_p$ and suppose that for some $q > p$

$$\mu(\Delta_\varepsilon^c) = O(\varepsilon^{-q}) \quad \text{as } \varepsilon \to \infty\,, \tag{4.54}$$

then define the **ζ-function associated to μ** by

$$\zeta_\mu(p) := \mathrm{Pers}_p^p(\mu) = p\,\mathcal{M}[\mu(\Delta_\varepsilon^c)]\,(p)\,. \tag{4.55}$$

Remark 2.35. *A priori, $\zeta_\mu(p)$ is defined on the strip $\langle p, q\rangle \subset \mathbb{C}$.* This could have also been guaranteed by replacing the condition of decay of $\mu(\Delta_\varepsilon^c)$ by requiring that $\mu \in \mathcal{D}_p \cap \mathcal{D}_q$.

As before, we may also define a local ζ-function.

Definition 2.36. The **local ζ-function associated to** μ at x is defined as

$$\zeta_\mu^x(p) := p\,\mathcal{M}[\mu(R_{x,\varepsilon})](p)\,. \tag{4.56}$$

Remark 2.37. These definitions are compatible with the notions of ζ-functions defined for a stochastic process. Seeing $\mathbb{E}[\mathrm{Dgm}(X)]$ as a measure on $\overline{\mathcal{X}}$, $\zeta_X = \zeta_{\mathbb{E}[\mathrm{Dgm}(X)]}$ (and the same holds for local ζ-functions).

A characterization of the topology metrized by the distance d_p is useful and has been investigated by Divol and Lacombe in [33]. The reader familiar with optimal transport will recognize this as an adaptation of the known characterization for probability measures of Wasserstein topology by vague convergence and convergence of pth-moments. That this equivalence holds for measures of *a priori* infinite mass is, however, a non-trivial extension.

Lemma 2.38 (Characterization of the topology metrized by d_p, [33])**.** *Let $(\mu_n)_n \subset \mathcal{D}_p$ and $\mu \in \mathcal{D}_p$. Then, the following equivalence holds*

$$\left\{d_p(\mu_n,\mu) \xrightarrow{n\to\infty} 0\right\} \iff \left\{\mu_n \xrightarrow[n\to\infty]{v} \mu \text{ and } \mathrm{Pers}_p(\mu_n) \xrightarrow[n\to\infty]{} \mathrm{Pers}_p(\mu)\right\}\,. \tag{4.57}$$

Remark 2.39. Of course, given our choice of notation, if $p < \infty$, we can rewrite $\mathrm{Pers}_p(\mu_n) \to \mathrm{Pers}_p(\mu)$ as $\zeta_{\mu_n}(p) \to \zeta_\mu(p)$.

Furthermore, it is possible to show that

Proposition 2.40 (Interpolation for optimal transport)**.** *Let $1 \le p < q \le \infty$ and $\theta \in]0,1[$. Define p_θ by*

$$\frac{1}{p_\theta} = \frac{\theta}{p} + \frac{1-\theta}{q}\,. \tag{4.58}$$

Then, for $\mu,\nu \in \mathcal{D}_p \cap \mathcal{D}_q$

$$d_{p_\theta}(\mu,\nu) \le 2^{1-\theta}\, d_p^\theta(\mu,\nu)\,(\mathrm{Pers}_q(\mu) + \mathrm{Pers}_q(\nu))^{1-\theta}\,. \tag{4.59}$$

Consequently, if $p \le r \le q$, then $\mathcal{D}_p \cap \mathcal{D}_q \subset \mathcal{D}_r$.

Remark 2.41. For probability measures, Wasserstein interpolation follows trivially from an application of Jensen's inequality. However, since diagrams are *a priori* of infinite mass, this interpolation result needs to be shown.

Proof. Let π be an optimal transport for d_p. Applying Littlewood's inequality,

$$\begin{aligned} d_{p_\theta}(\mu,\nu) &\leq \left\| d_{\mathbb{R}^2,\infty} \right\|_{L^{p_\theta}(\pi)} \leq \left\| d_{\mathbb{R}^2,\infty} \right\|_{L^p(\pi)}^{\theta} \left\| d_{\mathbb{R}^2,\infty} \right\|_{L^q(\pi)}^{1-\theta} \\ &= d_p(\mu,\nu)^\theta \left[\int_{\overline{\mathcal{X}}^2} d^q_{\mathbb{R}^2,\infty}(z,z')\, d\pi(z,z') \right]^{\frac{1-\theta}{q}} \\ &\leq d_p(\mu,\nu)^\theta \left[2^q \int_{\overline{\mathcal{X}}^2} d^q_{\mathbb{R}^2,\infty}(z,\Delta) + d^q_{\mathbb{R}^2,\infty}(\Delta,z')\, d\pi(z,z') \right]^{\frac{1-\theta}{q}} , \end{aligned}$$

where this equality holds everywhere on the support of π. This can be shown by defining

$$S = \{(z,z') \in \overline{\mathcal{X}}^2 \cap \operatorname{supp}(\pi) \,|\, d_{\mathbb{R}^2,\infty}(z,z') > d_{\mathbb{R}^2,\infty}(z,\Delta) + d_{\mathbb{R}^2,\infty}(z',\Delta)\} . \tag{4.60}$$

This set S either has null or positive measure. If it has positive measure, then we can modify the transport plan π by sending the projections of S to the diagonal, thereby producing a transport plan of strictly inferior cost to that of π, which is a contradiction. Hence, S is of null measure, so the equality holds over the support of the measure. Finally, this entails

$$d_{p_\theta}(\mu,\nu) = 2^{1-\theta}\ d_p(\mu,\nu)^\theta\,(\operatorname{Pers}_q(\mu) + \operatorname{Pers}_q(\nu))^{1-\theta} .$$

If $q=\infty$, since π is an optimal transport between μ and ν and $\mu,\nu \in \mathcal{D}_\infty$, π itself must have compact support and the diameter of the support is bounded above by $\operatorname{Pers}_\infty(\mu) \vee \operatorname{Pers}_\infty(\nu)$, so the inequality of the proposition follows. ■

Lemma 2.42 (Sequential continuity of the inverse Mellin transform)**.** *Let $f_k :]0,\infty[\to \mathbb{C}$ be a sequence of functions uniformly bounded by a function $g :]0,\infty[\to \mathbb{R}_+$, whose Mellin transform g^* is defined over some non-empty strip $\langle \alpha,\beta \rangle \subset \mathbb{C}$. Suppose further that there is a function $f :]0,\infty[\to \mathbb{C}$ such that $f^* : \langle\alpha,\beta\rangle \to \mathbb{R}_+$ and $f_k^* \to f^*$ uniformly on every compact set of $\langle\alpha,\beta\rangle$. Then, $f_k \to f$ almost everywhere.*

Proof. For every k, $|f_k| \leq g \vee |f|$ and so the sequence $(f_k)_k$ is bounded in a weighted L^1 space. Up to extraction of a subsequence, f_k converges a.e. to some function h, also bounded above by $g \vee |f|$. By dominated convergence, along this subsequence, $f_k^* \to h^*$, which entails that $h = f$ a.e. since the Mellin transform is injective and $(h-f)^* = 0$ identically on $\langle\alpha,\beta\rangle$, since along *any* subsequence $f_k^* \to f^*$ uniformly on every compact set of $\langle\alpha,\beta\rangle$. It follows that the sequence $(f_k)_k$ has f as its only accumulation point, finishing the proof. ■

Remark 2.43. It is sufficient to consider that for all $s \in]\alpha,\beta[$, the sequence $(x^{s-1}f_k(x))_k$ is absolutely uniformly integrable.

Both of these lemmas allow for a comprehensive characterization of the objects we have thus far been concerned with throughout this paper.

Theorem 2.44 (ζ characterization of Wasserstein p-convergence)**.** *Let $(\mu_n)_n \subset \mathcal{D}_p \cap \mathcal{D}_q$ be a sequence of q-tame measures and $\mu \in \mathcal{D}_p \cap \mathcal{D}_q$ and suppose that the sequence $(\mu_n(\Delta_\varepsilon^c))_n$ can be uniformly bounded above by a function $g :]0,\infty[\to \mathbb{R}_+$ such that for $\varepsilon \in]0,1]$ $g(\varepsilon) = O(\varepsilon^{-p})$ as $\varepsilon \to 0$ and on $[1,\infty[$, $g(\varepsilon) = O(\varepsilon^{-q})$ as $\varepsilon \to \infty$. Then, the following are equivalent:*

1. *There exists* $p < r < q$ *such that* $d_r(\mu_n, \mu) \xrightarrow{n\to\infty} 0$.

2. *There exists* $p < r < q$ *such that* $\mu_n \xrightarrow[n\to\infty]{v} \mu$ *and* $\zeta_{\mu_n}(r) \to \zeta_\mu(r)$.

3. *For almost every* $x \in \mathbb{R}$ *and* $\varepsilon > 0$, $\mu_n(\Delta_\varepsilon^c) \to \mu(\Delta_\varepsilon^c)$ *and* $\mu_n(R_{x,\varepsilon}) \to \mu(R_{x,\varepsilon})$.

4. *For almost every* $x \in \mathbb{R}$, $\zeta_{\mu_n}^x \to \zeta_\mu^x$ *and* $\zeta_{\mu_n} \to \zeta_\mu$ *uniformly on every compact of* $\langle p, q\rangle$.

5. *For all* $p < r < q$, $d_r(\mu_n, \mu) \xrightarrow{n\to\infty} 0$.

Proof. Let us break down the proof in different steps.

- (1) $\iff$ (2) by lemma 2.38.

- (2) $\implies$ (3). By the Portmanteau theorem, the vague convergence of μ_n to μ entails that for any continuity set A of μ, $\mu_n(A) \to \mu(A)$. In particular, since the μ_n and μ are q-tame, $R_{x,\varepsilon}$ is a continuity set of μ for almost every x and ε, so $\mu_n(R_{x,\varepsilon}) \to \mu(R_{x,\varepsilon})$ and $\mu_n(\Delta_\varepsilon^c) \to \mu(\Delta_\varepsilon^c)$ a.e..

- (3) $\implies$ (4). Since $\mu_n(\Delta_\varepsilon^c) \to \mu(\Delta_\varepsilon^c)$ a.e., if we let $s \in \langle \alpha, \beta\rangle$, then

$$\begin{aligned}|\zeta_{\mu_n}(s) - \zeta_\mu(s)| &\le |s| \int_0^\infty \varepsilon^{\mathrm{Re}(s)-1} \left|\mu_n(\Delta_\varepsilon^c) - \mu(\Delta_\varepsilon^c)\right| \, d\varepsilon \\ &= |s| \left\{ \int_0^1 + \int_1^\infty \right\} \varepsilon^{\mathrm{Re}(s)-1} \left|\mu_n(\Delta_\varepsilon^c) - \mu(\Delta_\varepsilon^c)\right| \, d\varepsilon\end{aligned}$$

  The domination conditions of the theorem on $\mu_n(\Delta_\varepsilon^c)$ and $\mu(\Delta_\varepsilon^c)$ guarantee that this quantity is integrable as soon as $p < \mathrm{Re}(s) < q$. By dominated convergence, $\zeta_{\mu_n} \to \zeta_\mu$ on every compact of $\langle p, q\rangle$. We apply the same reasoning to the $R_{x,\varepsilon}$ by noting that the measure of $R_{x,\varepsilon}$ is always dominated by that of Δ_ε^c.

- (4) $\implies$ (5). Fix $r \in \,]p, q[$ and take a compact set K around r contained in $\langle p, q\rangle$. It suffices thus to show that the convergence of these ζ-functions entails vague convergence of the μ_n, which will show the result by lemma 2.38. Denoting $A_\varepsilon \in \{R_{x,\varepsilon}, \Delta_\varepsilon^c\}$, the boundedness condition of the theorem on $\mu_n(A_\varepsilon)$ entails that, by lemma 2.42, $\mu_n(A_\varepsilon) \to \mu(A_\varepsilon)$ almost everywhere. Since the collection of $R_{x,\varepsilon}$ and Δ_ε^c together forms a π-system, so for every continuity set of μ, we have convergence of their measures μ_n to μ, and so by applying the Portmanteau theorem once again, $\mu_n \to \mu$ vaguely, which shows the result.

- (5) $\implies$ (1) trivially.

■

Remark 2.45. Of course, by interpolation, it is sufficient that for all μ and $\mathrm{Re}(s) > p$, $\zeta_{\mu_n}(s)$ be bounded to guarantee that for any $s \in \,]p, \infty[$, $d_s(\mu_n, \mu) \to 0$, if $d_p(\mu_n, \mu) \to 0$.

Remark 2.46. Theorem 2.44 applies to average diagrams from any stochastic process satisfying the hypotheses of the theorem. This entails that if the ζ-functions of a process converge on some open set $\langle \alpha, \beta\rangle \subset \mathbb{C}$, the underlying expected diagrams themselves converge in some Wasserstein distance.

3 ζ-functions of Lévy processes

In [78], we have already studied the functional N^ε for Markov processes. Let us briefly recall some useful known facts about N^ε.

Proposition 3.1 (P, [78]). *Using summation by parts, it is possible to write*

$$\mathbb{E}[(N_t^\varepsilon)^s] = \sum_{k\geq 1}(k^s - (k-1)^s)\mathbb{P}(N_t^\varepsilon \geq k) \tag{4.61}$$

For processes on the interval which are not periodic (in the sense of [78]), if $k \geq 2$

$$\mathbb{P}(N_t^\varepsilon \geq k) = \mathbb{P}(S_{k-1}^\varepsilon \leq t)\,, \tag{4.62}$$

and $\mathbb{P}(N_t \geq 1) = \mathbb{P}(R_t \geq \varepsilon)$. Furthermore if X has the strong Markov property,

$$\mathbb{E}[N_t^\varepsilon] \sim \mathbb{P}(R_t \geq \varepsilon) \quad \text{as } \varepsilon \to \infty\,. \tag{4.63}$$

Finally, for $k \geq 2$ the Laplace transform (with respect to time, as per our convention) of equation 4.62 is

$$\mathcal{L}(\mathbb{P}(N_t^\varepsilon \geq k))(\lambda) = \frac{\mathbb{E}\left[e^{-\lambda S_{k-1}^\varepsilon}\right]}{\lambda}\,. \tag{4.64}$$

Proof of proposition 3.1. The only thing to prove is the result of equation 4.64, as the previous statements are all proved in [78]. Since we are dealing with processes which are not periodic in the sense of [78], then

$$\mathbb{P}(N_t^\varepsilon \geq k) = \mathbb{P}(S_{k-1}^\varepsilon \leq t)\,, \tag{4.65}$$

since as soon as the hitting time S_{k-1}^ε is attained we have at least k bars (due to the boundary of the interval). Using standard functional properties of the Laplace transform it is easy to see that

$$\mathcal{L}[\mathbb{P}(S_{k-1}^\varepsilon \leq t)](\lambda) = \frac{\mathcal{L}[\mathbb{P}(S_{k-1}^\varepsilon = t)](\lambda)}{\lambda}\,, \tag{4.66}$$

where $\mathbb{P}(S_{k-1}^\varepsilon = t)$ denotes the probability density function of S_{k-1}^ε. However, the Laplace transform of this density function is nothing other than the moment generating function $\mathbb{E}\left[e^{-\lambda S_{k-1}^\varepsilon}\right]$, since S_{k-1}^ε is a positive random variable. ■

Remark 3.2. If the process has the strong Markov property, we can write $\mathbb{E}\left[e^{-\lambda S_{k-1}^\varepsilon}\right]$ as the product of the Laplace transform of the distribution of its increments

$$S_{k-1}^\varepsilon = \sum_{i=0}^{k}(S_i^\varepsilon - T_i^\varepsilon) + (T_i^\varepsilon - S_{i-1}^\varepsilon)\,. \tag{4.67}$$

The expression of $\mathbb{E}\left[e^{-\lambda S_{k-1}^\varepsilon}\right]$ is particularly simple as soon as these increments are independent and identically distributed.

Remark 3.3. Ordering the bars of the barcode of a function f by their length, and denoting the length of the kth longest branch by ℓ_k, the following equivalence holds

$$N_t^\varepsilon \geq k \iff \ell_k \geq \varepsilon\,. \tag{4.68}$$

The probability distribution of both of the events above are thus the same. Consequently, there is a one-to-one correspondence between the elements of the sums

$$\mathbb{E}\Big[\mathrm{Pers}_p^p\Big] = \sum_{k\geq 1}\mathbb{E}[\ell_k^p] = p\sum_{k\geq 1}\mathcal{M}[\mathbb{P}(N_t^\varepsilon \geq k)](p) \tag{4.69}$$

whenever these quantities are defined. In particular, the distribution of each bar is in principle readily available, since $\mathbb{E}\Big[\ell_k^{p-1}\Big]$ is the Mellin transform of the distribution of ℓ_k. We will later see that in particular cases, we can gain access to the explicit distribution of bars in this way (*cf.* section 3.2).

3.1 Renewal theory

Throughout this section, we shall consider that the stopping times S_{k-1}^ε are such that the sequence $(U_k^\varepsilon)_{k\geq 0}$ is a sequence of i.i.d. atomless random variables, where

$$U_k^\varepsilon := S_k^\varepsilon - S_{k-1}^\varepsilon\,. \tag{4.70}$$

We shall adapt and enounce some theorems from renewal theory (*cf.* Allan Gut's book [48] for a detailed account of this) Let us define $\hat{N}_t^\varepsilon$ by

$$\hat{N}_t^\varepsilon := \max\{k\,|\,S_k \leq t\}\,. \tag{4.71}$$

Tautologically, for every k, $\hat{N}_t^\varepsilon \geq k \iff S_k^\varepsilon \leq t$. In fact, adopting the convention that the infinite bar has length the range of the process, we can write, always under convention 2.16 that

$$N_t^\varepsilon = 1_{\{R_t\geq\varepsilon\}} + \hat{N}_t^\varepsilon \tag{4.72}$$

Remark 3.4. Notice that convention 2.16 for the infinite bar might not be the most natural or more convenient from this point of view. Indeed, if we had adopted the convention that the infinite bar *always* has length at least ε, then

$$N_t^\varepsilon = 1 + \hat{N}_t^\varepsilon \tag{4.73}$$

and N_t^ε is a first passage time process defined by

$$N_t^\varepsilon := \min\{k\,|\,S_k > t\}\,. \tag{4.74}$$

Nonetheless, as previously stated, we will keep adopting convention 2.16 for the rest of this paper.

According to renewal theory, $\hat{N}_t^\varepsilon$ posses various desirable limit theorems which we shall briefly recall and refer the reader to [48] for complete proofs of the latter.

Theorem 3.5 (Strong Law of Counting Processes, Theorem 5.1 [48]). *Let* $0 < \mathbb{E}[U_1^\varepsilon] < \infty$, *then*

$$\frac{\hat{N}_t^\varepsilon}{t} \xrightarrow[a.s]{t\to\infty} \frac{1}{\mathbb{E}[U_1^\varepsilon]} \quad and \quad \frac{\mathbb{E}\Big[(\hat{N}_t^\varepsilon)^s\Big]}{t^s} \xrightarrow{t\to\infty} \frac{1}{(\mathbb{E}[U_1^\varepsilon])^s} \tag{4.75}$$

for all $s > 0$. *If* $\mathbb{E}[U_1^\varepsilon] = \infty$, *then the limits are* 0.

Theorem 3.6 (CLT for Counting Processes, Theorem 5.2 [48]). *Let* $0 < \mathbb{E}[U_1^\varepsilon] = \mu < \infty$ *and* $\sigma^2 = Var(U_1^\varepsilon) < \infty$, *then*

$$\frac{\hat{N}_t^\varepsilon - t/\mu}{\sqrt{\frac{\sigma^2 t}{\mu^3}}} \xrightarrow[\mathbb{P}]{t\to\infty} \mathcal{N}(0,1) \tag{4.76}$$

Furthermore,

$$\mathbb{E}\Big[\hat{N}_t^\varepsilon\Big] = \frac{t}{\mu} + \frac{\sigma^2 - \mu^2}{2\mu^2} + o(1) \quad as\ t \to \infty$$
$$Var(\hat{N}_t^\varepsilon) = \frac{\sigma^2 t}{\mu^3} + o(t) \quad as\ t \to \infty\,.$$

These results are in particular applicable for Lévy processes, where it is possible to show that the requirement that the sequence $(U_k)_k$ is i.i.d. is satisfied. In particular, the theorems above give the asymptotic long-time behaviour of the number of bars.

3.2 Lévy processes

For Lévy processes, the small scale asymptotics of N^ε can also be studied up to the following caveat : a wide range of Lévy processes have almost surely discontinuous paths (but nonetheless càdlàg), but our construction of trees (as done in [77]) is based on continuous functions. For this reason, it is necessary to define what tree we associate to a process X when X_t has almost surely discontinuous paths. Luckily, this caveat has been treated for càdlàg processes in [37, 80]. We will adopt the approach taken by Picard in [80], where the reader can find the details of the construction. Loosely speaking, Picard's approach consists in "completing" the function at the discontinuity points by joining an imaginary line linking the points of discontinuity (*cf.* figure 4.7).

In any case, it has been shown that on some fixed interval $[0,t]$, it is possible to obtain the behaviour of the number of bars of length $\geq \varepsilon$ as $\varepsilon \to 0$.

Proposition 3.7 (Picard, §3 [80]). *Let X be a Lévy process and suppose that, almost surely X has no interval on which it is monotone. Define*

$$\xi(\varepsilon) := \mathbb{E}[S^\varepsilon + T^\varepsilon] \tag{4.77}$$

for

$$S^\varepsilon := \inf\{t \mid X_t - \inf_{[0,t]} X > \varepsilon\} \quad and\ T^\varepsilon := \inf\{t \mid \sup_{[0,t]} X - X_t > \varepsilon\}\,, \tag{4.78}$$

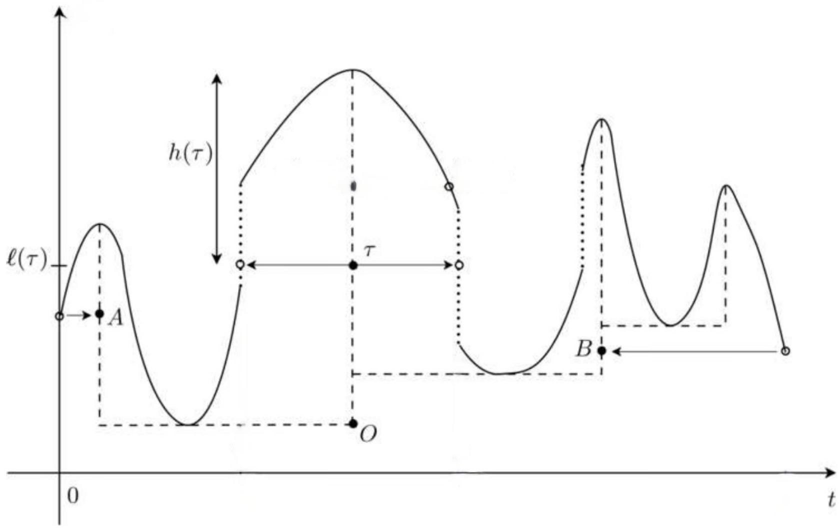


Figure 4.7: A depiction of the construction of a tree associated to a càdlàg function. The figure is taken from [80]

then $\xi(\varepsilon)N^\varepsilon \to 1$ *as* $\varepsilon \to 0$ *in probability. If* $\xi(\varepsilon) = O(\varepsilon^\alpha)$ *for some* α*, then the convergence is almost sure.*

Remark 3.8. The hypothesis on X is satisfied if X or $-X$ is not the sum of a subordinator and a compound Poisson process, in which case T_X is finite, so N^ε is bounded. Furthermore, the convergence is always almost sure for α-stable processes for which $|X|$ is not a subordinator by the scaling property. In fact, in that case there exists a constant C_α such that almost surely,

$$N^\varepsilon \sim \frac{C_\alpha}{\varepsilon^\alpha} \quad \text{as } \varepsilon \to 0\,. \tag{4.79}$$

If we can quantify correction terms to this asymptotic relation in L^1, this gives rise to a statistical test for α by using the stability results discussed in [78], we will explore this in more detail in section 3.2. By the self-similarity of α-stable process following the arguments of [80, §3], we can already at least conclude that

$$\left|\mathbb{E}[N_\alpha^\varepsilon] - C_\alpha \varepsilon^{-\alpha}\right| \leq 1\,. \tag{4.80}$$

Notation 3.9. In what will follow, we will denote by S^ε and T^ε two independent random variables distributed as the analogously denoted ones in proposition 3.7. Furthermore, define $U^\varepsilon = T^\varepsilon + S^\varepsilon$. In particular, if $\varepsilon = 1$, abusing the notation we will denote $U^1 = U$.

Remark 3.10. Henceforth, unless otherwise specified, we will always assume that X almost surely has no interval on which it is monotone.

This result by Picard is exactly the Strong Law of Counting Processes (theorem 3.5) applied to the stopping times we previously considered. Self-similarity allows us to trade the $t \to \infty$ limit by an $\varepsilon \to 0$ limit. However, Picard's result is more general that what could've been deduced with self-similarity, as it applies to Lévy processes which are not necessarily α-stable. In this setting, one could ask whether it is possible to prove an analogous theorem to the CLT of Counting Processes (theorem 3.6).

Theorem 3.11. *Let X be a Lévy process such that almost surely X has no interval on which it is monotone, using the notation defined in 3.9,*

$$\mathbb{E}[N_t^\varepsilon] = \frac{t}{\mathbb{E}[U^\varepsilon]} + \left(\frac{\mathbb{E}[(U^\varepsilon)^2]}{2\mathbb{E}[U^\varepsilon]^2} - 1 \right) + \mathbb{P}(R_t \geq \varepsilon) + o(\rho_\varepsilon^{-n}) \quad \text{as } \varepsilon \to 0 \,. \tag{4.81}$$

for any $n \in \mathbb{N}$, where ρ_ε denotes the radius of convergence of the Taylor series of $\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]$ around $\lambda = 0$, which can be bounded below by $-\log(\mathbb{P}(T^\varepsilon > 1) \vee \mathbb{P}(S^\varepsilon > 1))$, which is larger than 1 for ε small enough. Furthermore, if X is α-stable, the formula above becomes

$$\mathbb{E}[N_t^\varepsilon] = \frac{t}{\mathbb{E}[U]\,\varepsilon^\alpha} + \frac{\mathbb{E}[U^2]}{2\mathbb{E}[U]^2} + o(\varepsilon^{\alpha n}) \quad \text{as } \varepsilon \to 0 \,. \tag{4.82}$$

Let us stress that the main addition over the CLT of Counting Processes (theorem 3.6) is two-fold :

- We have found a similar limit for a processes which is not self-similar, and so for which the $t \to \infty$ limit was *a priori* not interchangeable with the $\varepsilon \to 0$ limit.
- We have further specified the rate of decay of the remainder in terms of ε, and this will turn out to be of importance;

To show the theorem, it is convenient to show first some technical lemmata, one of which is a slight refinement to a technical lemma proved in [80].

Lemma 3.12 (Picard, Proposition 3.14 [80])**.** *The variables S^ε and T^ε admit finite moments of order k for all k and the moment generating function $\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]$ is well defined on a neighborhood of zero. Furthermore, the radius of convergence of $\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]$ around 0, ρ_ε, then*

$$\rho_\varepsilon \geq -\log(\mathbb{P}(S^\varepsilon > 1) \vee \mathbb{P}(T^\varepsilon > 1)) \,, \tag{4.83}$$

if X is α-stable, then $\rho_\varepsilon = \rho\varepsilon^{-\alpha}$, for some constant $\rho > 0$ (which might be infinite). Finally, there exists $1 \leq C_k \leq 2^k \operatorname{Li}_{-k}(\frac{1}{2})$ such that

$$\mathbb{E}[U^\varepsilon]^k \leq \mathbb{E}\left[(U^\varepsilon)^k\right] \leq C_k\, \mathbb{E}[U^\varepsilon]^k \ . \tag{4.84}$$

Remark 3.13. The bound on the constant in this lemma is not optimal. Notice also that this result entails $\rho_\varepsilon \to \infty$ as $\varepsilon \to 0$.

Lemma 3.14. *Keeping the same notation (cf. notation 3.9) as before,*

$$U^\varepsilon \xrightarrow[\varepsilon \to 0]{L^r} 0 \quad \text{and} \quad U^\varepsilon \xrightarrow[\varepsilon \to 0]{a.s.} 0 \,. \tag{4.85}$$

for every $r \geq 1$.

Lemma 3.15. *For every k, there exists a constant D_k such that*

$$1 - D_k e^{-\lambda} \mathbb{E}[U^\varepsilon]^k \lesssim \mathbb{E}\left[e^{-\lambda U^\varepsilon}\right] \leq 1 \quad \text{as } \varepsilon \to 0 \,. \tag{4.86}$$

Furthermore, for real γ and σ,

$$1 - \mathbb{E}\Big[e^{-\gamma U^\varepsilon}\Big] \leq \Big|1 - \mathbb{E}\Big[e^{-(\gamma+i\sigma)U^\varepsilon}\Big]\Big| \,, \tag{4.87}$$

In particular,

$$\left|\frac{\mathbb{E}\Big[e^{-(\gamma+i\sigma)U^\varepsilon}\Big]}{1-\mathbb{E}\big[e^{-(\gamma+i\sigma)U^\varepsilon}\big]}\right| \leq \frac{\mathbb{E}\Big[e^{-\gamma U^\varepsilon}\Big]}{1-\mathbb{E}\big[e^{-\gamma U^\varepsilon}\big]} \,. \tag{4.88}$$

Lemma 3.16. *For any $\delta > 0$, weakly in $L^2([\delta, \infty[)$, for any $k \geq 0$*

$$\frac{1}{2\pi i}\int_{\gamma-iT}^{\gamma+iT} e^{\lambda t}\lambda^k \, d\lambda \xrightarrow[T\to\infty]{\mathcal{D}} 0 \tag{4.89}$$

at a rate $O(T^{-n})$ for any $n \in \mathbb{N}$. Furthermore, for any $k \geq 0$

$$\frac{1}{2\pi i}\int_{\gamma-iT}^{\gamma+iT} \frac{e^{\lambda t}}{\lambda^k} \, d\lambda \xrightarrow[T\to\infty]{\mathcal{D}} \frac{t^{k-1}}{(k-1)!} \tag{4.90}$$

weakly at a rate $O(T^{-(n+k-1)})$ and the convergence is strong as soon as $k \geq 2$.

Proof of lemma 3.12. It is sufficient to show it for S^ε knowing that an analogous treatment is possible for T^ε. The points in $[0, S^\varepsilon]$ are characterized by the fact that

$$X_t - \inf_{[0,t]} X < \varepsilon \,. \tag{4.91}$$

In particular, the supremum over all t ranging within $[0, S^\varepsilon]$ of this quantity is also less than ε.

$$\mathbb{P}(S^\varepsilon > a) = \mathbb{P}\left[\sup_{0\leq t\leq a}\left(X_t - \inf_{[0,t]} X\right) < \varepsilon\right] \,. \tag{4.92}$$

Consider now an interval $[0, k\mu]$, where $\mu > 0$ and k is an integer and let us slice this interval into k segments of length μ. It is clear that the following inequality holds

$$\sup_{0\leq t\leq k\mu}\left(X_t - \inf_{[0,t]} X\right) \geq \sup_{1\leq j\leq k}\left(\sup_{(j-1)\mu\leq t\leq j\mu}\left(X_t - \inf_{[(j-1)\mu,t]} X\right)\right) \,, \tag{4.93}$$

since over each smaller interval we can have only smaller spread than over the entire interval. However, the right hand side is a supremum over i.i.d. random variables, so that

$$\begin{aligned}\mathbb{P}(S^\varepsilon > k\mu) &\leq \mathbb{P}\left[\sup_{1\leq j\leq k}\left(\sup_{(j-1)\mu\leq t\leq j\mu}\left(X_t - \inf_{[(j-1)\mu,t]} X\right)\right) < \varepsilon\right] \\ &= \mathbb{P}\left[\sup_{0\leq t\leq \mu}\left(X_t - \inf_{[0,t]} X\right) < \varepsilon\right]^k = \mathbb{P}(S^\varepsilon > \mu)^k \,.\end{aligned}$$

By the non-monotonicity of X, $\mathbb{P}(S^\varepsilon > \mu) < 1$. In particular, if we let $\mu = 1$, and denote

$c = \mathbb{P}(S^\varepsilon > 1) < 1$, then:

$$\lim_{k\to\infty} e^{\lambda k}\,\mathbb{P}(S^\varepsilon > k) \le \lim_{k\to\infty} e^{\lambda k}\,\mathbb{P}(S^\varepsilon > 1)^k = \lim_{k\to\infty} (e^\lambda c)^k = 0 \tag{4.94}$$

as soon as $\lambda < \log(1/c)$. It follows that $\mathbb{E}\left[e^{-\lambda S^\varepsilon}\right]$ is well-defined for λ in some neighbourhood of zero and in particular all moments of S^ε are well-defined and finite. Finally, combining the above remark with an application of Markov's inequality we get

$$\mathbb{P}(S^\varepsilon > 2k\mathbb{E}[S^\varepsilon]) \le 2^{-k}\,. \tag{4.95}$$

Almost surely, $\frac{S^\varepsilon}{2\mathbb{E}[S^\varepsilon]} \le G$, where G is a geometric random variable. The moments of G can easily be calculated, yielding the estimation in the lemma for the moments..

This shows that the radius of convergence of the Taylor series $\mathbb{E}\left[e^{-\lambda S^\varepsilon}\right]$ is bounded below by $-\log(\mathbb{P}(S^\varepsilon > 1))$, since Taylor series converge up to their nearest singularity. If X is α-stable, we can use the relation which tells us that, in distribution $U^\varepsilon = \varepsilon^\alpha U$, so that by the ratio test,

$$\limsup_{k\to\infty} \frac{|\lambda|}{k+1}\,\frac{\mathbb{E}\left[(U^\varepsilon)^{k+1}\right]}{\mathbb{E}[(U^\varepsilon)^k]} = \varepsilon^\alpha\,|\lambda| \limsup_{k\to\infty} \frac{1}{k+1}\,\frac{\mathbb{E}\left[U^{k+1}\right]}{\mathbb{E}[U^k]}\,. \tag{4.96}$$

This limit is equal to $0 \le \rho^{-1} < \infty$, since $\mathbb{E}\left[e^{-\lambda U}\right]$ is analytic on a non-trivial disk around 0, by lemma 3.12 and usual properties of the Laplace transform.

$$\limsup_{k\to\infty} \frac{|\lambda|}{k+1}\,\frac{\mathbb{E}\left[(U^\varepsilon)^{k+1}\right]}{\mathbb{E}[(U^\varepsilon)^k]} < 1\,, \tag{4.97}$$

whenever $|\lambda| < \rho\varepsilon^{-\alpha}$, which shows the desired result.

■

Proof of lemma 3.14. The statement in L^r follows from the following observation.

$$0 \le \mathbb{E}[(U^\varepsilon)^r] \le \sum_{k=1}^{\infty} k^r\mathbb{P}(U^\varepsilon \ge k) \le \sum_{k=1}^{\infty} k^r\mathbb{P}(U^\varepsilon \ge 1)^k \tag{4.98}$$

by the arguments of lemma 3.12. This sum converges, since $\mathbb{P}(U^\varepsilon \ge 1) < 1$. As $\varepsilon \to 0$, $\mathbb{P}(U^\varepsilon > 1) \to 0$, since X is almost surely nowhere monotone, so that the entire sum tends to 0. The almost sure statement follows from the fact that U^ε is monotone, since both T^ε and S^ε are monotone functions of ε. Since L^r convergence implies almost sure convergence along a subsequence ε_n, for $\varepsilon_{n+1} < \varepsilon < \varepsilon_{n+1}$ by monotonicity of U^ε

$$U^{\varepsilon_{n+1}} < U^\varepsilon < U^{\varepsilon_n}\,, \tag{4.99}$$

so the convergence is almost sure. ■

Proof of lemma 3.15. The first inequality of the lemma relies on the fact that

$$\begin{aligned}\mathbb{E}\Big[e^{-\lambda T^\varepsilon}\Big] &\le \sum_{k\ge 0} e^{-\lambda k}\mathbb{P}(T^\varepsilon > k) \le \sum_{k\ge 0}\Big[e^{-\lambda}\mathbb{P}(T^\varepsilon>1)\Big]^k \\ &= \frac{1}{1-e^{-\lambda}\mathbb{P}(T^\varepsilon>1)} \sim 1-e^{-\lambda}\mathbb{P}(T^\varepsilon>1) \quad \text{as } \varepsilon\to 0\,,\end{aligned}$$

since $\mathbb{P}(T^\varepsilon > 1) \xrightarrow{\varepsilon\to 0} 0$. Notice an analogous inequality holds for S^ε. By Markov's inequality, we know that

$$\mathbb{P}(T^\varepsilon > 1) \le \mathbb{E}\Big[(T^\varepsilon)^k\Big] \le C_k\mathbb{E}[T^\varepsilon]^k \tag{4.100}$$

from which the first inequality follows by lemma 3.14. The second and third inequalities follow from noticing that for any x and y

$$||x|-|y|| \le |x-y| \tag{4.101}$$

and applying Jensen's inequality. ■

Proof of lemma 3.16. Consider a test function $\varphi \in C_c^\infty([\delta,\infty[)$, then integrating by parts

$$\begin{aligned}\frac{1}{2\pi i}\int_{\mathbb{R}}\left[\int_{\gamma-iT}^{\gamma+iT} e^{\lambda t}\lambda^k\, d\lambda\right]\varphi(t)\; dt &= \frac{(-1)^k}{2\pi i}\int_{\mathbb{R}} dt\; \varphi^{(k)}(t)\int_{\gamma-iT}^{\gamma+iT} e^{\lambda t}\, d\lambda \\ &= (-1)^k\int_{\mathbb{R}}\frac{e^{\gamma t}\varphi^{(k)}(t)}{\pi t}\sin(Tt)\; dt\end{aligned} \tag{4.102}$$

By performing the change of variables $y = Tt$, we see that the integral is weakly approaching 0, as φ is not supported at 0. Additionally, away from 0, the function

$$\frac{e^{\gamma t}\varphi^{(k)}(t)}{\pi t} \tag{4.103}$$

is a compactly supported C^∞-function, integrating by parts n subsequent times equation 4.102 yields bounds of this integral by $C_\varphi T^{-n}$, where C_φ is a constant which depends on the test function and its support.

Let us now show that

$$\frac{1}{2\pi i}\int_{\gamma-iT}^{\gamma+iT}\frac{e^{\lambda t}}{\lambda^k}\, d\lambda \xrightarrow[T\to\infty]{\mathcal{D}} \frac{t^{k-1}}{(k-1)!}\,. \tag{4.104}$$

Once again integrating by parts,

$$\int_{\mathbb{R}} dt\;\varphi(t)\left[\frac{1}{2\pi i}\int_{\gamma-iT}^{\gamma+iT}\frac{e^{\lambda t}}{\lambda^k}\, d\lambda - \frac{t^{k-1}}{(k-1)!}\right] = (-1)^n\int_{\mathbb{R}} dt\;\varphi^{(n)}(t)\left[\frac{1}{2\pi i}\int_{\gamma-iT}^{\gamma+iT}\frac{e^{\lambda t}}{\lambda^{k+n}}\, d\lambda - \frac{t^{n+k-1}}{(n+k-1)!}\right].$$

Applying the residue theorem to evaluate the complex integral we get, for $T > \gamma$,

$$\frac{1}{2\pi i}\int_{\gamma-iT}^{\gamma+iT}\frac{e^{\lambda t}}{\lambda^{k+n}}\, d\lambda = \frac{t^{n+k-1}}{(n+k-1)!} + \frac{1}{2\pi i}\int_{C_T}\frac{e^{\lambda t}}{\lambda^{k+n}}\, d\lambda\,, \tag{4.105}$$

where C_T is the circle of center $\lambda = \gamma$ and radius T. By the estimation lemma, the contribution of this integral is bounded by $e^{\gamma t}T^{-(n+k-1)}$. It follows that

$$\left|\int_{\mathbb{R}} dt\, \varphi(t) \left[\frac{1}{2\pi i}\int_{\gamma-iT}^{\gamma+iT} \frac{e^{\lambda t}}{\lambda^k}\, d\lambda - \frac{t^{k-1}}{(k-1)!}\right]\right| \leq T^{-(n+k-1)} \left\|e^{\gamma t}\varphi^{(n)}(t)\right\|_{L^1}, \tag{4.106}$$

thereby giving the speed of convergence desired. ■

Proof of theorem 3.11. Throughout this proof, we shall denote

$$F(\lambda,\varepsilon) := \frac{1}{\lambda}\frac{\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]}{1-\mathbb{E}[e^{-\lambda U^\varepsilon}]}. \tag{4.107}$$

The assumption of non-monotonicity of the Lévy process ensures that, almost surely, S^ε and T^ε both tend to 0 as $\varepsilon \to 0$. Consider now the times T_i^ε and S_i^ε given in definition 2.19. Since X is Lévy, $T_{i+1}^\varepsilon - S_i^\varepsilon$ and $S_i^\varepsilon - T_i^\varepsilon$ are independent from one another, and are both equal in distribution to T^ε and S^ε respectively.

By lemma 3.12, S^ε and T^ε admit finite moments for all k and the function $\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]$ is well defined, so, by equation 4.64

$$\begin{aligned}\mathcal{L}(\mathbb{E}[N_t^\varepsilon])(\lambda) &= \mathcal{L}(\mathbb{P}(R_t \geq \varepsilon))(\lambda) + \frac{1}{\lambda}\sum_{k\geq 1}\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]^k \\ &= \mathcal{L}(\mathbb{P}(R_t \geq \varepsilon))(\lambda) + \frac{1}{\lambda}\frac{1}{\mathbb{E}[e^{-\lambda U^\varepsilon}]^{-1}-1}\end{aligned} \tag{4.108}$$

If we denote ρ_ε the radius of convergence of the Taylor series at zero associated to $\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]$, for $|\lambda| < \rho_\varepsilon$,

$$\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right] = \sum_{k=0}^{\infty}\frac{(-\lambda)^k\mathbb{E}\left[(U^\varepsilon)^k\right]}{k!}. \tag{4.109}$$

This radius of convergence ρ_ε can be bounded below with the results of lemma 3.12 by

$$-\log(\mathbb{P}(S^\varepsilon > 1) \vee \mathbb{P}(T^\varepsilon > 1)) < \rho_\varepsilon, \tag{4.110}$$

which entails that $\rho_\varepsilon \to \infty$ as $\varepsilon \to 0$. We deduce from this series the Laurent series associated to $\lambda^{-1}(\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]^{-1} - 1)^{-1}$, namely

$$F(\lambda,\varepsilon) = \frac{1}{\lambda^2\mathbb{E}[U^\varepsilon]} + \frac{1}{\lambda}\left[\frac{\mathbb{E}[(U^\varepsilon)^2]}{2\mathbb{E}[U^\varepsilon]^2} - 1\right] + \frac{3\mathbb{E}[(U^\varepsilon)^2]^2 - 2\mathbb{E}[U^\varepsilon]\,\mathbb{E}[(U^\varepsilon)^3]}{12\mathbb{E}[U^\varepsilon]^3} + O(\lambda).$$

where the remainder in λ is an analytic function of λ for $|\lambda| < \rho_\varepsilon$. By the inequalities of lemma 3.15, the function doesn't admit any poles on the half plane $\mathrm{Re}(\lambda) > 0$, so that the Taylor series above converges over the same disk as that of $\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]$.

Observe now that for some small $\gamma > 0$, the inverse Laplace transform of $F(\lambda,\varepsilon)$ can be written as

$$\mathcal{L}^{-1}[F](t,\varepsilon) = \frac{1}{2\pi i}\left\{\int_{\gamma-i\rho_\varepsilon}^{\gamma+i\rho_\varepsilon} + \int_{\gamma+i\rho_\varepsilon}^{\gamma+i\infty} + \int_{\gamma-i\infty}^{\gamma-i\rho_\varepsilon}\right\} e^{\lambda t}F(\lambda,\varepsilon)\, d\lambda. \tag{4.111}$$

Weakly, the integrals going off to infinity are of order $o(\rho_\varepsilon^{-n})$ for any $n \in \mathbb{N}$, since for any test function $\varphi \in C_c^\infty([0,\infty[)$, integrating by parts

$$\int_{\mathbb{R}} dt\, \varphi(t) \left[\frac{1}{2\pi i} \int_{\gamma+i\rho_\varepsilon}^{\gamma+i\infty} e^{\lambda t} F(\lambda,\varepsilon)\, d\lambda\right] = \frac{(-1)^n}{2\pi i} \int_{\mathbb{R}} dt\, \varphi^{(n)}(t) \int_{\gamma+i\rho_\varepsilon}^{\gamma+i\infty} d\lambda\, \frac{e^{\lambda t}}{\lambda^n} F(\lambda,\varepsilon)\, d\lambda \quad (4.112)$$

But using lemma 3.15,

$$\left| \int_{\gamma+i\rho_\varepsilon}^{\gamma+i\infty} d\lambda\, \frac{e^{\lambda t}}{\lambda^n} F(\lambda,\varepsilon)\, d\lambda \right| \leq e^{\gamma t} \int_{\gamma+i\rho_\varepsilon}^{\gamma+i\infty} \left| \frac{F(\lambda,\varepsilon)}{\lambda^n} \right| d\lambda = O(\rho_\varepsilon^{-n-2})\,, \quad (4.113)$$

which entails that the integrals going to infinity in equation 4.111 converge weakly to 0 at a rate $o(\rho_\varepsilon^{-n})$ for any $n \in \mathbb{N}$. Thus, asymptotically as $\varepsilon \to 0$, for $t > 0$,

$$\mathbb{E}[N_t^\varepsilon] = \frac{t}{\mathbb{E}[U^\varepsilon]} + \left(\frac{\mathbb{E}\left[(U^\varepsilon)^2\right]}{2\mathbb{E}[U^\varepsilon]^2} - 1\right) + \mathbb{P}(R_t \geq \varepsilon) + o(\rho_\varepsilon^{-n})\,, \quad (4.114)$$

for any $n \in \mathbb{N}$. If the process is α-stable, then $(X_{c^\alpha t})_{t\geq 0} = (cX_t)_{t\geq 0}$ in distribution for all c, so that $U^\varepsilon = \varepsilon^\alpha U$ in distribution and

$$\mathbb{E}[N_t^\varepsilon] = \frac{t}{\mathbb{E}[U]\,\varepsilon^\alpha} + \frac{\mathbb{E}[U^2]}{2\mathbb{E}[U]^2} + o(\varepsilon^{\alpha n}) \quad \text{as } \varepsilon \to 0\,, \quad (4.115)$$

for all $n \in \mathbb{N}$. As $\varepsilon \to 0$, $1 - \mathbb{P}(R_t \geq \varepsilon) = o(\varepsilon^n)$ for any n, since

$$\mathbb{P}(R_t \leq \varepsilon) \leq \mathbb{P}(T^\varepsilon > t) \leq \frac{\varepsilon^{\alpha k}\mathbb{E}\left[(T^1)^k\right]}{t^k} \quad (4.116)$$

for any k by Markov's inequality. ■

Remark 3.17. A similar theorem can be proven in $L^s(\Omega)$ for α-stable processes. For instance, if X is α-stable and $s = 2$ one obtains that for every $n \in \mathbb{N}$,

$$\begin{aligned}\mathrm{Var}(N_t^\varepsilon) \sim &\left[\frac{\mathrm{Var}(U) - 2\mathbb{E}[U]^2}{\mathbb{E}[U]^3}\right] \frac{t}{\varepsilon^\alpha} \\ &+ \frac{5\mathrm{Var}(U)^2}{4\mathbb{E}[U]^4} + \frac{\mathrm{Var}(U)}{\mathbb{E}[U]^2} - \frac{2\mathbb{E}[U^3]}{3\mathbb{E}[U]^3} + \frac{7}{4} + o(\varepsilon^{\alpha n}) \quad \text{as } \varepsilon \to 0\,.\end{aligned}$$

Interestingly, there is a constant term appearing in this expansion, which can be understood as induced by the boundary. This interpretation comes from Picard's analysis of the problem [80], in which the first term of this asymptotic series was also derived (*cf.* proposition 3.7).

If X has almost surely discontinuous paths, X_t exhibits macroscopic jumps. These will turn out to bring significative contributions, so much so that

Corollary 3.18. *If $\alpha \neq 2$, the ζ-function of any α-stable Lévy process is ill-defined for any $p \in \mathbb{C}$.*

Proof. The ζ-function of a stochastic process X can be written as

$$\zeta_X(p) = \mathbb{E}[R_t^p] + \mathbb{E}\left[\sum_{k\geq 2} \ell_k^p(X_t)\right] , \tag{4.117}$$

if X is α-stable, the first term can be written as

$$\mathbb{E}[R_t^p] = t^{\frac{p}{\alpha}} \mathbb{E}[R_1^p] \geq t^{\frac{p}{\alpha}} \mathbb{E}[|X_1|^p] , \tag{4.118}$$

where we have momentarily taken $p \in \mathbb{R}$. Since X_1 has a Lévy α-stable distribution, taking p now complex, $\mathbb{E}[R_t^p]$ is infinite as soon as $\mathrm{Re}(p) \geq \alpha$, since X_1 does not admit any moments of order (of real part) larger than α. Applying theorem 3.11, we know that the second term in the above decomposition of ζ_X is only defined for $\mathrm{Re}(p) > \alpha$, so the fundamental strip of $\mathcal{M}\mathbb{E}[N_t^\varepsilon]$ is empty. ■

In fact, it is possible to show that $\mathbb{P}(X_1 > \varepsilon) \sim \mathbb{P}(R_1 > \varepsilon)$ as $\varepsilon \to \infty$. It turns out that the distribution of R_1 is dominated by the probability of having one large jump, which confirms our previous statement on the effect of the discontinuity of Lévy processes on the distribution of R. This is the so-called single big jump principle.

Proposition 3.19 (Single big jump principle, Bertoin, [10])**.** *If X is an α-stable process ($\alpha < 2$), there exists a constant k such that*

$$\mathbb{P}(R_1 \geq \varepsilon) \sim \frac{k}{\varepsilon^\alpha} \quad \textit{as } \varepsilon \to \infty . \tag{4.119}$$

Loosely speaking, it is intuitive to think that a corrective asymptotic power series for $\mathbb{P}(R_t \geq \varepsilon)$ of the form

$$\mathbb{P}(R_1 \geq \varepsilon) \sim \sum_{k\geq 1} a_k \varepsilon^{-k\alpha} \qquad \text{as } \varepsilon \to \infty \tag{4.120}$$

should exist for the following reason. By the single big jump principle, the probability that the range exceeds ε for large ε is dominated by the probability of a single big jump. However, it is also possible to have n large jumps of size $J_k\varepsilon$ where $\sum_k^n J_k \geq 1$. The probability of each of these jumps happening is of order $O(\varepsilon^{-\alpha})$ and by independence, the probability that k jumps of size $O(\varepsilon)$ happen is $O(\varepsilon^{-\alpha k})$. In general, we cannot expect these events to be disjoint from one another, so the coefficients a_k of this sum may be negative. Finally, by the scaling invariance it is sufficient to show that this is so for R_1. Corrective terms to the above asymptotic relation should thus in principle exist, but the explicit calculation of these terms is out of the scope of this paper.

By contrast, we will now show that $\mathbb{E}[N_t^\varepsilon] - \mathbb{P}(R_t \geq \varepsilon)$ is well-behaved. This motivates the following definition

Definition 3.20. The **tail ζ-function** of the stochastic process X on $[0, t]$ is defined as

$$\hat{\zeta}_X(p) := \mathbb{E}\Big[\mathrm{Pers}_p^p(X) - R_t^p\Big] . \tag{4.121}$$

Theorem 3.21. *The tail ζ-function associated to an α-stable Lévy process is given by*

$$\hat{\zeta}_X(p) = \frac{t^{\frac{p}{\alpha}}}{\Gamma(\frac{p}{\alpha})} B^*\left(\frac{p}{\alpha}\right) , \tag{4.122}$$

which extends to a meromorphic function of p to the entire complex plane (since B^* is itself meromorphic), with a unique simple pole at $p = \alpha$ of residue $\mathbb{E}[U]^{-1} \alpha t$.

Proof of theorem 3.21. To show that this quantity is well-defined, let us start by noticing that

$$\mathcal{L}(\mathbb{E}[N_t^\varepsilon] - \mathbb{P}(R_t \geq \varepsilon))(\lambda) = \frac{\mathbb{E}\left[e^{-\lambda\varepsilon^\alpha U}\right]}{\lambda(1 - \mathbb{E}[e^{-\lambda\varepsilon^\alpha U}])} \tag{4.123}$$

which for $\mathrm{Re}(\lambda) > 0$ goes to zero (uniformly in λ) exponentially fast as $\varepsilon \to \infty$, showing that $\mathbb{E}[N_t^\varepsilon] - \mathbb{P}(R_t \geq \varepsilon)$ does as well for $\varepsilon \to \infty$ by an application of Markov's inequality. We can also compute the contribution of the second term of equation 4.117. First, notice that

$$\mathcal{L}\left(\mathbb{E}\left[\sum_{k\geq 2} \ell_k^p(X_t)\right]\right)(\lambda) = \frac{p}{\lambda}\mathcal{M}\left[\frac{\mathbb{E}\left[e^{-\lambda\varepsilon^\alpha U}\right]}{1 - \mathbb{E}[e^{-\lambda\varepsilon^\alpha U}]}\right](p) \tag{4.124}$$

Using the scaling property of the Mellin transform $\mathcal{M}_z[f(\lambda z)](p) = \lambda^{-p} f^*(p)$ and inverting the Laplace transform

$$\mathbb{E}\left[\mathrm{Pers}_p^p(X) - R_t^p\right] = \frac{pt^{\frac{p}{\alpha}}}{\Gamma(1+\frac{p}{\alpha})}\mathcal{M}\left[\frac{\mathbb{E}\left[e^{-\varepsilon^\alpha U}\right]}{1 - \mathbb{E}[e^{-\varepsilon^\alpha U}]}\right](p) \, . \tag{4.125}$$

Finally, setting

$$B(z) := \frac{\mathbb{E}\left[e^{-zU}\right]}{1 - \mathbb{E}[e^{-zU}]} \quad \text{and} \quad B^*(p) := \mathcal{M}_z[B(z)](p) \, , \tag{4.126}$$

the polynomial scaling property of the Mellin transform, $\mathcal{M}_z[f(z^\alpha)](p) = \frac{1}{\alpha} f^*(\frac{p}{\alpha})$ yields the final result. ■

Remark 3.22. Theorem 3.21 can be used to give an alternative proof for the series expansion of theorem 3.11.

Alternate proof of theorem 3.11. By lemma 3.12 and the analyticity of the expresion of B with respect to $\mathbb{E}\left[e^{-zU}\right]$, B admits a Laurent series on some non-trivial annulus around zero with a single simple pole at $z = 0$. By the fundamental correspondence (theorem 2.5), the existence of this Laurent expansion guarantees that $B^*(\frac{p}{\alpha})$ admits a meromorphic continuation to the whole complex plane with only simple poles at every $p = -n\alpha$ for every $n \in \mathbb{N}$ and at $p = \alpha$. The poles at the negative integer multiples of α are compensated exactly by those of the Γ-function in the denominator of the expression of $\hat{\zeta}$, leaving only a pole at α. Now, recalling that

$$\hat{\zeta}(p) = p\mathcal{M}[\mathbb{E}[N_t^\varepsilon] - \mathbb{P}(R_t \geq \varepsilon)](p) \, , \tag{4.127}$$

$\mathcal{M}[\mathbb{E}[N_t^\varepsilon] - \mathbb{P}(R_t \geq \varepsilon)]$ has a supplementary pole at $p = 0$. Admitting that $\hat{\zeta}(p)/p$ has the decay condition to apply the fundamental correspondence by inverting the Mellin transform we get the asymptotic relation desired. ■

Exponential corrections

The fundamental correspondence is limited in that it allows us only to describe $\mathbb{E}[N^\varepsilon]$ asymptotically up to terms smaller than any polynomial. However, in accordance to the discussion of section 2.1, a finer study of the analytic properties of $\hat{\zeta}$ can yield the superpolynomial corrections to our estimate, assuming that $B(z)$ admits a meromorphic extension to the whole complex plane. Using lemmata 2.6 and 2.8,

$$\hat{\zeta}_X(p) = t^{\frac{p}{\alpha}}\Gamma\left(1 - \frac{p}{\alpha}\right)\sum_{z_0\in\mathcal{P}} \mathrm{Res}((-z)^{\frac{p}{\alpha}-1}B(z); z_0) \tag{4.128}$$

$$\mathcal{M}(\mathbb{E}[N_t^\varepsilon] - \mathbb{P}(R_t \geq \varepsilon))(p) = -\frac{t^{\frac{p}{\alpha}}\Gamma(-\frac{p}{\alpha})}{\alpha}\sum_{z_0\in\mathcal{P}} \mathrm{Res}((-z)^{\frac{p}{\alpha}-1}B(z); z_0)\,. \tag{4.129}$$

Recognizing that

$$\mathcal{M}_z\left[\frac{e^{z_0/z}}{z_0}\right](p) = -\Gamma(-p)\left(\frac{-1}{z_0}\right)^{1-p}\,, \tag{4.130}$$

we may formally invert the Mellin transform if all the z_0's are simple poles to obtain the exponentially small corrections

$$\mathbb{E}[N_t^\varepsilon] - \mathbb{P}(R_t \geq \varepsilon) - \frac{t}{\mathbb{E}[U]\,\varepsilon^\alpha} - \left[\frac{\mathbb{E}[U^2]}{2\mathbb{E}[U]^2} - 1\right] \sim \sum_{z_0\in\mathcal{P}} \frac{e^{tz_0/\varepsilon^\alpha}}{\alpha z_0}\mathrm{Res}(B(z); z_0) \quad \text{as } \varepsilon \to 0\,. \tag{4.131}$$

Generally, the poles are not simple so the corrective terms to this series stem from residues of higher order poles (the corrections remain nonetheless superpolynomially small as $\varepsilon \to 0$).

Distribution of the length of branches

The distribution of the length of the kth branch (in the sense of figure 3.3) can be calculated. Recall that

$$\mathbb{E}[\ell_k^p(X)] = p\mathcal{M}[\mathbb{P}(N_t^\varepsilon \geq k)](p)\,. \tag{4.132}$$

For $k \geq 2$,

$$\mathcal{L}[\mathbb{E}[\ell_k^p(X)]](\lambda) = \frac{p}{\lambda}\mathcal{M}\left[\mathbb{E}\left[e^{-\lambda U^\varepsilon}\right]^{k-1}\right](p)\,. \tag{4.133}$$

If we suppose once again that X is a Lévy α-stable process, this can be simplified to yield

$$\mathbb{E}[\ell_k^p(X)] = \frac{t^{\frac{p}{\alpha}}}{\Gamma(\frac{p}{\alpha})}\mathcal{M}\left[\mathbb{E}\left[e^{-\varepsilon U}\right]^{k-1}\right]\left(\frac{p}{\alpha}\right)\,, \tag{4.134}$$

Taking the Mellin transform of a power is in general difficult. Inversion is also in general complicated due to the presence of the Γ-function in the denominator of the above expression.

To remediate the first problem, we can form the generating function yielding the distribution for the kth bar,

$$G_\alpha(z;p) := \sum_{k\geq 2} \mathbb{E}[\ell_k^p(X)]\, z^k = \frac{t^{\frac{p}{\alpha}}}{\Gamma(\frac{p}{\alpha})}\, \mathcal{M}\left[\frac{z}{(z\mathbb{E}[e^{-\varepsilon U}])^{-1} - 1}\right]\left(\frac{p}{\alpha}\right), \tag{4.135}$$

which allows us to express

Proposition 3.23. *For $k \geq 2$, the distribution of the length of the kth longest branch is characterized by its Mellin transform which is given by*

$$\mathbb{E}[\ell_k^p(X)] = \frac{1}{k!}\, \frac{\partial^k}{\partial z^k}\bigg|_{z=0} G_\alpha(z;p)\,. \tag{4.136}$$

Whenever convenient, the expression above can also be evaluated by considering a circular contour of small enough radius r around the origin C_r and evaluating

$$\mathbb{E}[\ell_k^p(X)] = \frac{1}{2\pi i} \oint_{C_r} \frac{G_\alpha(z;p)}{z^{k+1}}\, dz\,. \tag{4.137}$$

Statistical parameter testing for α-stable processes and perspectives

What we will aim to do in this section is to illustrate by example why barcodes can be a robust statistical tools for parameter testing. Parameter testing is a widely studied subject, notably for self-similar processes, where the problem has been treated in dimension 1 (a non-comprehensive list of references is [32] and the references therein). A variety of different methods, such as multi-scale wavelet analysis, have been used to produce these results (although other methods such as the ones of [32] have also been used), so our approach does not offer anything new in this respect. The interest of our method lies in possible applications to higher dimensional random fields, for which wavelet analysis is not an effective tool. A complete theoretical framework for this would require the study of the trees of higher dimensional random fields, which are out of the scope of this paper : instead, this section acts as a proof of concept for the use of topological estimators and their utility, by studying what happens in dimension 1.

In what follows, we will consider X to be an α-stable Lévy process, of which we will aim to estimate the parameter α. From proposition 3.7 we know that almost surely

$$N_t^\varepsilon \sim Ct\varepsilon^{-\alpha} \quad \text{as } \varepsilon \to 0\,. \tag{4.138}$$

In particular, given some sample we may compute the sampled value of N_t^ε, which we will denote $\hat{N}_t^\varepsilon$ explicitly. A close inspection of the behaviour of the sample mean $\overline{N}_t^\varepsilon$ should thus yield an estimation for the parameter α of the process X.

Remark 3.24. In fact, the same reasoning allows us to estimate the Hurst parameter H of a fractional Brownian motion (fBM), which also exhibits self-similarity. In this case, the analogue of the asymptotic result of proposition 3.7 is [80, §3]

$$\text{a.s.} \quad N^\varepsilon \sim Ct\varepsilon^{-\frac{1}{H}} \quad \text{as } \varepsilon \to 0\,. \tag{4.139}$$

More precisely, given a sample, our test consists in performing the following steps.

1. Sample M paths of the stochastic process X (for example at regular intervals of size $\frac{1}{N}$ for some N) ;

2. Compute the barcode of the sampled paths. To do this, first construct a filtered simplicial complex (which is in this case nothing other than a chain with $\sim N$ links) by taking each point to be a vertex of a complex and joining adjacent sampling points with an edge. The filtration on this complex is the value of the process at the edge (for an edge connecting vertex a to vertex b, the value of the filtration is $X_a \wedge X_b$). Finally, the persistent homology of this complex can be computed with the `gudhi` package [50], which incidentally also offers a convenient implementation of filtered simplicial complexes due to Boissonnat and Maria [14].

3. For *some range of small enough* ε, and for some positive constant $c > 1$ compute the quantity
$$\hat{\alpha}_M := \log_c \left[\frac{\overline{N}_t^{\varepsilon/c} - \overline{N}_t^{2\varepsilon/c}}{\overline{N}_t^{\varepsilon} - \overline{N}_t^{2\varepsilon}} \right]\,. \tag{4.140}$$
Here, the notion of *some range of small enough* ε and the constant c both depend on N, with the limiting condition that as $N \to \infty$, the lower bound on the range of valid ε goes to zero.

Our claim is that the computed quantity $\hat{\alpha}$ is a valid estimation of the parameter α (for fBM, the quantity obtained in this way is an estimate of $\frac{1}{H}$).

Lemma 3.25 (Convergence of the sample means)**.** *The quotient*

$$\frac{\overline{N}_t^{\varepsilon/c} - \overline{N}_t^{2\varepsilon/c}}{\overline{N}_t^{\varepsilon} - \overline{N}_t^{2\varepsilon}} \xrightarrow[M\to\infty]{\mathbb{P}} \frac{\mathbb{E}\left[N_t^{\varepsilon/c} - N_t^{2\varepsilon/c}\right]}{\mathbb{E}\left[N_t^{\varepsilon} - N_t^{2\varepsilon}\right]} \tag{4.141}$$

at a rate $C_s M^{-s}$, for every $1 \leq s \leq 2$ where C_s is a constant depending on s and the sth moment of $N^{\varepsilon/c}$. In particular,

$$\hat{\alpha}_M \xrightarrow[M\to\infty]{\mathbb{P}} \alpha + \xi(\varepsilon) \tag{4.142}$$

at the same rate, where $\xi(\varepsilon)$ is a superpolynomially small function of ε.

Remark 3.26. The at first seemingly convoluted expression for the estimator $\hat{\alpha}_M$ can be explained due to the results of theorem 3.11. The substraction present in the numerator and denominator is performed so that the constant terms of equation 4.82 vanish. Ignoring the superpolynomial contributions to this expression which remain small, we then have that the argument inside the log of the estimator is roughly

$$c^{\hat{\alpha}_M} \approx \frac{\frac{t}{\mathbb{E}[U](\varepsilon/c)^\alpha} - \frac{t}{\mathbb{E}[U](2\varepsilon/c)^\alpha}}{\frac{t}{\mathbb{E}[U]\varepsilon^\alpha} - \frac{t}{\mathbb{E}[U](2\varepsilon)^\alpha}} \approx c^\alpha \, . \tag{4.143}$$

With this in mind, let us now formally prove the statement of lemma 3.25.

Proof. That the numerator and the denominator tend to the respective expected values holds by a simple application of the weak law of large numbers, since N_t^ε is a random variable in L^s for $s \geq 1$. The rate of convergence of this limit can be obtained via a simple application of Markov's inequality, by noting first that the summands in the denominator tend to their limits faster than those of the numerator, as the latter's sth moments are always larger than the former's. From theorem 3.11, we see that the limit can be expressed as

$$\frac{\mathbb{E}\left[N_t^{\varepsilon/c} - N_t^{2\varepsilon/c}\right]}{\mathbb{E}\left[N_t^\varepsilon - N_t^{2\varepsilon}\right]} = c^\alpha \, \frac{1+g(c\varepsilon)}{1+g(\varepsilon)} \, , \tag{4.144}$$

where g is a function tending to 0 superpolynomially fast as $\varepsilon \to 0$, determined by the superpolynomial corrections to the results of theorem 3.11. The statement of the lemma ensues. ∎

Lemma 3.27 (Probable L^∞-distance of sampling). *Denote $\hat{X}$ the trajectory samples of the α-stable process X at every interval of length $\frac{1}{N}$. More precisely, somewhat abusing the notation we can write,*

$$\hat{X}_t = \sum_{n=0}^{N-1} 1_{[\frac{n}{N}, \frac{n+1}{N}[}(t) \, X_{\frac{n}{N}} \, . \tag{4.145}$$

There exists a constant k such that

$$\mathbb{P}\left(\sup_{t\in\mathbb{R}} \left|X_t - \hat{X}_t\right| \leq \varepsilon\right) \gtrsim 1 - \left(\frac{k}{N\varepsilon^\alpha}\right)^N \qquad \textit{as } \varepsilon N^{1/\alpha} \to \infty \, . \tag{4.146}$$

Remark 3.28. The asymptotic dependence above fixes admissible values of ε as a function of N as holding whenever the asymptotic dependence above is valid (it must be valid between ε/c and 2ε). Furthermore, the parameter c we chose above is also further constrained by the requirement that the asymptotic relation of theorem 3.11 holds between ε/c and 2ε. More precisely, we fix c and ε such that the superpolynomial contributions in the expansion of theorem 3.11 are negligible with respect to the term in $\varepsilon^{-\alpha}$ *and* by imposing that $\varepsilon N^{1/\alpha}$ is large enough so that the asymptotic relation of lemma 3.27 simultaneously holds within the range $[\frac{\varepsilon}{c}, 2\varepsilon]$. In practice, we may fix c and ε by looking at a log-log chart of N^ε, the regime of validity of ε and the value of c become clear, as shown in figure 4.8.

Remark 3.29. With respect to the barcode, linear interpolation between values of X or the consideration of the process $\hat{X}$ is equivalent.

Remark 3.30. It is not a priori obvious that the event above is measurable. However, continuity in probability and the a.s. existence of a càdlàg modification of the process allows us to interpret this event to be a supremum over every $t \in \mathbb{Q}$, rendering the event measurable.

Proof. It suffices to show the result over the interval $[0,1]$. Noticing that the sampling coincides with the value of the path at every $t = \frac{1}{N}$, it suffices to evaluate the probability that over N intervals of length $\frac{1}{N}$ the real sampled path $X_t(\omega)$ (notice the absence of a hat) strays away from the sampled path $\hat{X}_t(\omega)$. Focusing on a single interval $[0, \frac{1}{N}]$, the single big jump principle 3.19 states that there exists a constant k such that this probability of straying away in this interval is

$$\mathbb{P}(R_{\frac{1}{N}} \geq \varepsilon) \sim \frac{k\varepsilon^{-\alpha}}{N} \quad \text{as } \varepsilon N^{1/\alpha} \to \infty \,. \tag{4.147}$$

By independence, over N such intervals

$$\mathbb{P}\left(\sup_{t\in\mathbb{R}} \left|X_t - \hat{X}_t\right| \leq \varepsilon\right) \gtrsim 1 - \left(\frac{k}{N\varepsilon^\alpha}\right)^N \sim 1 \quad \text{as } \varepsilon N^{1/\alpha} \to \infty \,, \tag{4.148}$$

as desired. ■

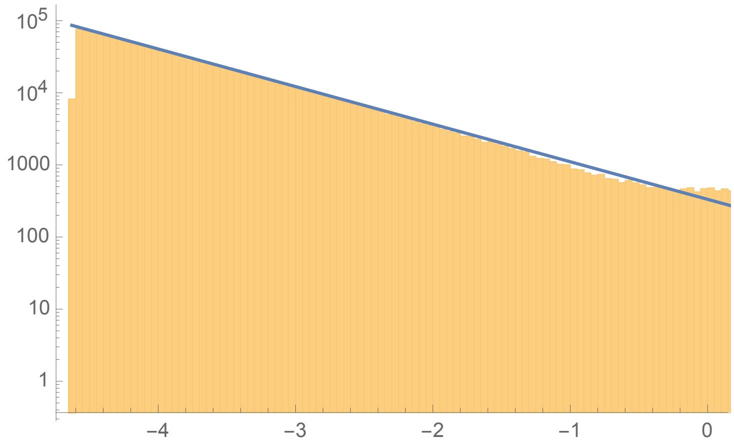


Figure 4.8: In orange, a histogram of the number of bars of length $\geq \varepsilon$, N^ε found as a function of $\log \varepsilon$ from a simulation of a Lévy 1.2-stable process as a random walk. In blue, the function $C_{1.2}\varepsilon^{-1.2}$.

Now, we use proposition 4.3, which can be specialized given our two lemmas above. The theorem provides bounds on $N^\varepsilon_{\hat{X}}$ provided that we know the value of δ_N, since if ε is small enough, N^ε_X has some almost sure asymptotic behaviour. On the other hand, by virtue of lemma 3.27 we have a *probable* estimate of δ_N, *i.e.* with probability q, we may give a bound of δ_N, rendering the statement quantitative. The second part of the statement of theorem 4.1 provides bounds on the L^1 distance between $N^\varepsilon_{\hat{X}}$ and N^ε_X, provided that we know that $\mathbb{E}[N^\varepsilon_X]$ is continuous. This happens to be the case for Brownian motion, as shown in [78]. Showing it in full generality for Lévy processes requires a closer study of the range of Lévy processes and the continuity of the inverse Mellin transform of $\hat{\zeta}_X(p)/p$. However, for the purposes of the construction of our statistical test, lemma 3.25 suffices, as it provides us with a quantitative guarantee that the parameter α is well estimated by our estimator $\hat{\alpha}_M$.

3.3 Propagators and local ζ-functions

In dimension one, it is possible to use the total order of $\mathbb{R}$ and count $N^{x,x+\varepsilon}$ by counting the number of times we go up from x to $x+\varepsilon$. This idea can be formalized by the following sequence of stopping times already introduced in the literature of classical probability theory (*cf.* for instance [86]).

Definition 3.31. Setting $S_0^{x,\varepsilon} = T_0^{x,\varepsilon} = 0$, we define a sequence of times recursively

$$T_{i+1}^{x,\varepsilon} := \inf\left\{ t \geq S_i^{x,\varepsilon} \;\middle|\; f(t) \leq x \wedge (x+\varepsilon) \right\}$$
$$S_{i+1}^{x,\varepsilon} := \inf\left\{ t \geq T_{i+1}^{x,\varepsilon} \;\middle|\; f(t) \geq x \vee (x+\varepsilon) \right\} . \tag{4.149}$$

Counting the number of bars of length ε is thus exactly to count the number of up and downs we make. More precisely,

$$N^{x,x+\varepsilon} = \inf\{ i \,|\, T_i^{x,\varepsilon} \text{ or } S_i^{x,\varepsilon} = \inf \emptyset \} , \tag{4.150}$$

by which we mean that it is the smallest i such that the set over which T_i^ε or S_i^ε are defined as infima is empty.

With the aforementioned, it is possible to define a Feynman-like formalism to perform the computation of $\mathbb{E}\left[N^{x,x+\varepsilon}\right]$.

Definition 3.32. Let X be a (strong Markov) stochastic process and let

$$T^a := \inf\{ t \geq 0 \;|\; X_t > a \} \tag{4.151}$$

be the hitting time of a by X. We define the **propagator from x to y** by

$$\langle x | y \rangle := \mathbb{E}_x\left[e^{-\lambda T^y} \right] \tag{4.152}$$

whenever this exists.

Remark 3.33. Whenever convenient, we may take $\lambda = i\omega$, and modify the subsequent expressions appropriately.

If the process X has the strong Markov property and the increments between the stopping times $T_i^{x,\varepsilon}$ and $S_i^{x,\varepsilon}$ are identically distributed, we can once again apply renewal theory. In particular, the Laplace transform of the occupation numbers $\mathcal{L}N_t^{x,x+\varepsilon}$ can be understood in terms of the propagators above. For $x > 0$,

$$\begin{aligned}
\mathcal{L}(\mathbb{E}\left[N_t^{x,x+\varepsilon}\right])(\lambda) &= \frac{\langle 0 | x+\varepsilon \rangle}{\lambda} \sum_{k \geq 0} (\langle x+\varepsilon | x \rangle \langle x | x+\varepsilon \rangle)^k \\
&= \frac{\langle 0 | x+\varepsilon \rangle}{\lambda (1 - \langle x+\varepsilon | x \rangle \langle x | x+\varepsilon \rangle)}
\end{aligned}$$

Similarly, all moments of the distribution can in principle be calculated

$$\mathcal{L}(\mathbb{E}\Big[(N_t^{x,x+\varepsilon})^{s-1}\Big])(\lambda) = \frac{\langle 0|x+\varepsilon\rangle(1-\langle x+\varepsilon|x\rangle\langle x|x+\varepsilon\rangle)}{\lambda}\,\mathrm{Li}_{-s+1}(\langle x+\varepsilon|x\rangle\langle x|x+\varepsilon\rangle)$$

where Li denotes the polylogarithm.

Remark 3.34. If $X_t - X_s = X_{t-s}$ in distribution, we can rewrite the above as,

$$\mathcal{L}(\mathbb{E}\Big[(N^{x,x+\varepsilon})^{s-1}\Big])(\lambda) = \frac{\langle 0|x+\varepsilon\rangle(1-\langle\varepsilon|0\rangle\langle 0|\varepsilon\rangle)}{\lambda}\,\mathrm{Li}_{-s+1}(\langle\varepsilon|0\rangle\langle 0|\varepsilon\rangle) \tag{4.153}$$

If $\langle x+\varepsilon|x\rangle\langle x|x+\varepsilon\rangle$ and $\langle 0|x+\varepsilon\rangle$ admit an asymptotic expansion for small ε (alternatively, we can suppose that this function is a smooth enough function of ε), then the Mellin transform of the expression above admits a meromorphic continuation. Furthermore,

$$\langle x+\varepsilon|x\rangle\langle x|x+\varepsilon\rangle = 1 + o(1) \quad \text{as } \varepsilon \to 0 \tag{4.154}$$

so that the order of the divergence of $\mathcal{L}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big])$ is dictated exclusively by the order of the first correction in ε to the product above. In reality,

Corollary 3.35. *Suppose that the increments in the sequence of stopping times defined by $(T_i^{x,\varepsilon}, S_i^{x,\varepsilon})$ are i.i.d. and that $\langle 0|x+\varepsilon\rangle$ and $\langle x+\varepsilon|x\rangle\langle x|x+\varepsilon\rangle$ have asymptotic expansions of the form of that of the fundamental correspondence (cf. theorem 2.5), then, the function $\mathcal{ML}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big])(p,\lambda)$ has a meromorphic extension on the half plane* $\mathrm{Re}(p) > -k$ *for any $k \in \mathbb{N}$.*

Proof. It's a simple application of the fundamental correspondence. ∎

Itô diffusions

Computing the local ζ-functions is sometimes easier than computing the ζ-function associated with the process X. This is because the computation of the distribution of U^ε might not always be straightforward. An example of this is the case of Itô diffusions, *i.e.* solutions to stochastic differential equations of the form

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dB_t\,, \tag{4.155}$$

for some smooth functions μ and σ. These processes have the strong Markov property and the sequence of stopping times $(T_i^\varepsilon, S_i^\varepsilon)$ defined above are identically distributed as μ and σ do not depend explicitly on time. Furthermore, these processes have infinitesimal generator

$$\mathcal{G} = \mu(x)\frac{\partial}{\partial x} + \frac{\sigma^2(x)}{2}\frac{\partial^2}{\partial x^2}\,. \tag{4.156}$$

We can use the theory of diffusion developped by Itô and McKean [51] to find explicit expressions for the propagators $\langle x|y\rangle$. The propagator is exactly the fundamental solution associated with the equation

$$(\mathcal{G}-\lambda)\rho(x,\lambda)=0 \text{ subject to } \begin{cases} \rho(y,\lambda)=1 & \rho(x,\lambda) \xrightarrow{x\to-\infty} 0 \quad \text{for } x<y \\ \rho(y,\lambda)=1 & \rho(x,\lambda) \xrightarrow{x\to+\infty} 0 \quad \text{for } x>y \end{cases}$$

The solution to the above boundary value problem has been shown to be of the form [51, p.130]

$$\rho(x,\lambda)=\frac{\Psi_\lambda(x)}{\Psi_\lambda(y)}=\langle x|y\rangle\,, \tag{4.157}$$

where if $x<y$ (resp. $x>y$) $\Psi_\lambda(z)$ is (up to some constant) the unique increasing (resp. decreasing) positive solution of the equation

$$\mathcal{G}\Psi_\lambda=\lambda\Psi_\lambda\,. \tag{4.158}$$

Noting Ψ_λ the solution for $x<y$ and Φ_λ the solution for $x>y$, for $x>0$,

$$\mathcal{L}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big])(\lambda)=\frac{\Psi_\lambda(0)}{\lambda\Psi_\lambda(x+\varepsilon)}\frac{1}{1-\frac{\Psi_\lambda(x)}{\Phi_\lambda(x)}\frac{\Phi_\lambda(x+\varepsilon)}{\Psi_\lambda(x+\varepsilon)}}\,. \tag{4.159}$$

These solutions $\Psi_\lambda(z)$ and $\Phi_\lambda(z)$ are smooth in z, so that the (Laplace transform of the) local ζ-function ζ_X^x admits a meromorphic continuation to $\mathbb{C}$.

Obtention of the local time for continuous semimartingales

The obtention of an expression for the propagators of a semimartingale is in principle sufficient to obtain an expression for the Laplace transform (in time) of the local time. This is possible due to the following theorem.

Theorem 3.36 (Revuz,Yor, Ch. VI Theorem 1.10 [86])**.** *Let X be a continuous semimartingale and let $L_X^x(t)$ be its local time on the interval $[0,t]$ at level x. Writing $X=M+A$ the Doob decomposition of X, suppose that for $s\geq 1$*

$$\mathbb{E}\Big[[M]_t^{s/2}+\Big(\int_0^t |dA|_s\Big)^s\Big]<\infty\,, \tag{4.160}$$

then in L^s,

$$2\varepsilon N_t^{x,x+\varepsilon}\xrightarrow[\varepsilon\to0]{L^s} L_X^x(t)\,, \tag{4.161}$$

in particular, this convergence holds in distribution as well.

Notation 3.37. Whenever the underlying process is implicitly clear, we will denote the local time by L_t^x.

This theorem entails in particular that for $s\geq 2$,

$$(2\varepsilon)^{s-1}\mathbb{E}\Big[(N_t^{x,x+\varepsilon})^{s-1}\Big]\xrightarrow[\varepsilon\to0]{}\mathbb{E}\Big[(L_X^x(t))^{s-1}\Big]\,. \tag{4.162}$$

Under the technical hypothesis that if for some $a > 0$, $\mathrm{Re}(\lambda) > a$ for every $\varepsilon > 0$ the function $e^{-\lambda t}\mathbb{E}\Big[(2\varepsilon N_t^{x,x+\varepsilon})^{s-1}\Big]$ is bounded above by some integrable function of t, the dominated convergence theorem entails that the Laplace transform of the limit and the limit of the Laplace transforms also coincide. Alternatively, we may also check whether $\mathbb{E}\Big[(2\varepsilon N_t^{x,x+\varepsilon})^{s-1}\Big]$ is a monotone function of ε over some neighbourhood for ε small enough and apply the monotone convergence theorem. This allows us to conclude that

$$\mathcal{L}[\mathbb{E}\Big[(L_X^x(t))^{s-1}\Big]](\lambda) = \lim_{\varepsilon\to 0}(2\varepsilon)^{s-1}\mathcal{L}[\mathbb{E}\Big[(N_t^{x,x+\varepsilon})^{s-1}\Big]](\lambda)\,. \tag{4.163}$$

Finally, this has consequences for the distribution of $L_X^x(t)$ since $\mathbb{E}\big[(L_X^x(t))^{s-1}\big]$ is exactly the Mellin transform of the distribution of the local time.

4 Examples of applications

4.1 Brownian motion

For the rest of this section, B will denote a standard Brownian motion started at 0.

Associated ζ-function and asymptotic expansions for N^ε

Let us start by remarking that, in distribution

$$\sup_{[0,t]} B - B_t = |B_t|\ . \tag{4.164}$$

The stopping times T^ε and S^ε of theorem 3.11 are identically distributed and are distributed as the hitting times of ε by a reflected Brownian motion. An application of Doob's stopping theorem (*cf.* [15, p.641]) shows that

$$\mathbb{E}\Big[e^{-\lambda U^\varepsilon}\Big] = \mathrm{sech}^2(\varepsilon\sqrt{2\lambda})\,. \tag{4.165}$$

The term $\mathbb{P}(N_t^\varepsilon \geq 1) = \mathbb{P}(R_t \geq \varepsilon)$ can also be computed by considering the fundamental solution of the corresponding heat equation with Dirichlet boundary conditions. We obtain [78]

$$\mathbb{P}(R_t \geq \varepsilon) = 4\sum_{k=1}^{\infty}(-1)^{k-1}k\,\mathrm{erfc}\left[\frac{k\varepsilon}{\sqrt{2t}}\right]\,. \tag{4.166}$$

Respectively, since Brownian motion is a 2-stable Lévy process, using equation 4.122 (here, $B(z) = \mathrm{csch}^2(\sqrt{2z})$) we can write

$$\hat{\zeta}_B(p) = 2^{3-\frac{3p}{2}}t^{\frac{p}{2}}\frac{\Gamma(p)}{\Gamma(\frac{p}{2})}\zeta(p-1)\,. \tag{4.167}$$

Remark 4.1. This can be obtained by using the functional properties of the Mellin transform (scaling, power of the argument) shown in table 4.1 and by the results of table 4.2.

Putting everything together, we get

Theorem 4.2. *The ζ-function of Brownian motion on the interval $[0, t]$ admits an meromorphic extension to the whole complex plane. Furthermore, it is exactly equal to*

$$\zeta_B(p) = \frac{4(2^p - 3)}{\sqrt{\pi}} \left(\frac{t}{2}\right)^{\frac{p}{2}} \Gamma\left(\frac{p+1}{2}\right) \zeta(p-1) \tag{4.168}$$

for all p and has a unique simple pole at $p = 2$ of residue t.

Remark 4.3. That the Riemann ζ-function appears in this expression is *a posteriori* not surprising. Loosely speaking, this connection between the Riemann ζ-function and Brownian motion appears through the relation of Brownian motion with Jacobi's ϑ-function, which is a fundamental solution of the heat equation [11]. Work connecting stable distributions whose Laplace transform is of the form of equation 4.165 have been widely studied by Pitman, Yor and Biane [11, 81, 82]. In [11], one can also find connections between stable distributions of this form and number theoretical L-functions. It is also interesting to recall that there exists a probabilistic interpretation of the Riemann hypothesis through Li's criterion [62] as detailed in [11, §2.3].

Proof of theorem 4.2. Taking the Mellin transform of $\mathbb{P}(R_t \geq \varepsilon)$

$$\mathcal{M}(\mathbb{P}(R_t \geq \varepsilon))(p) = 2^{2-\frac{3p}{2}}(2^p - 4)t^{\frac{p}{2}} \frac{\Gamma(p)}{\Gamma(\frac{p}{2}+1)} \zeta(p-1)\,.$$

Multiplying the above expression by p and adding both terms and using the fact that $\Gamma(z+1) = z\Gamma(z)$ and the Legendre duplication formula, we obtain the result. ∎

The meromorphic extension of ζ allows us to directly compute correction terms for the asymptotic series given in [78]. $\mathcal{M}(\mathbb{E}[N_t^\varepsilon])(p)$ has only two poles, one at $p = 0$ and one at $p = 2$. Furthermore, along a vertical strip, $\mathcal{M}(\mathbb{E}[N_t^\varepsilon])(p)$ decays rapidly enough to use the fundamental correspondence (theorem 2.5). Using contour integration and the Mellin inversion theorem, we can conclude that

$$\mathbb{E}[N_t^\varepsilon] = \frac{t}{2\varepsilon^2} + \frac{2}{3} + O(\varepsilon^n) \quad \text{as } \varepsilon \to 0\,, \tag{4.169}$$

for any $n \in \mathbb{N}$, as prescribed by theorem 3.11. The expectations in the expression of the theorem can be read on the expansion

$$\operatorname{sech}^2(\sqrt{2\varepsilon}) = 1 - 2\varepsilon + \frac{1}{2!}\frac{16}{3}\varepsilon^2 + O(\varepsilon^3)\,. \tag{4.170}$$

As previously shown, the analyticity of ζ_B beyond $\operatorname{Re}(p) = 2$ guarantees that there are no polynomial corrections in ε to $\mathbb{E}[N_t^\varepsilon]$ as $\varepsilon \to 0$. The analyticity of ζ_B on the half plane $\operatorname{Re}(p) > 2$ suggests that $\mathbb{E}[N_t^\varepsilon]$ is rapidly decreasing as $\varepsilon \to \infty$. This is corroborated by the more general approximation of proposition 3.1 for Markov processes found in [78], namely $\mathbb{E}[N_t^\varepsilon] \sim \mathbb{P}(R_t \geq \varepsilon)$ as $\varepsilon \to \infty$.

Applying the observations made in section 2.1, the superpolynomial corrections to the asymptotic series can be found by looking carefully at the meromorphic extension of ζ_B.

Proposition 4.4. *For Brownian motion* $\mathbb{E}[N_t^\varepsilon]$ *admits the following series representations which converge well for large and small* ε *respectively*

$$\begin{aligned}\mathbb{E}[N_t^\varepsilon] &= 4\sum_{k\geq 1}(2k-1)\,\mathrm{erfc}\left(\frac{(2k-1)\varepsilon}{\sqrt{2t}}\right) - k\ \mathrm{erfc}\left(\frac{2k\varepsilon}{\sqrt{2t}}\right)\\ &= \frac{t}{2\varepsilon^2} + \frac{2}{3} + 2\sum_{k\geq 1}(2(-1)^k - 1)\frac{e^{-\pi^2k^2t/2\varepsilon^2}t}{\varepsilon^2}\left[1 + \frac{\varepsilon^2}{\pi^2k^2t}\right].\end{aligned}$$

Proof. This asymptotic formula for $\mathbb{E}[N_t^\varepsilon]$ can be obtained by using the arguments of section 2.1. Indeed, $B(z) = \mathrm{csch}^2(\sqrt{2z})$ admits a meromorphic continuation to the entire complex plane and has poles of order two at $z = -\frac{\pi^2n^2}{2}$ for every $n \in \mathbb{Z}\setminus\{0\}$. It follows that

$$\mathrm{Res}\left((-z)^{\frac{p}{2}-1}B(z), -\frac{\pi^2n^2}{2}\right) = 2^{1-\frac{p}{2}}(2\pi)^{p-2}(p-1)\,. \tag{4.171}$$

Taking the inverse Mellin transform of equation 4.129, we obtain the desired result. The second expression converging fast for large ε is obtained by using the functional equation of the ζ-function and taking the inverse Mellin transform of the expression obtained. ∎

Alternative proof of proposition 4.4. Note that

$$\mathcal{L}(\mathbb{E}[N_t^\varepsilon])(\lambda) = \frac{4}{\lambda}\sum_{k\geq 1}(2k-1)e^{-(2k-1)\varepsilon\sqrt{2\lambda}} - ke^{-2k\varepsilon\sqrt{2\lambda}} \tag{4.172}$$

$$= \left[\frac{2\cosh(\varepsilon\sqrt{2\lambda}) - 1}{\lambda}\right]\mathrm{csch}^2(\varepsilon\sqrt{2\lambda})\,. \tag{4.173}$$

By inverting the Laplace transform in equation 4.172 (this can be done by first decomposing the hyperbolic expressions into a series of exponential terms, of which the inverse Laplace transform can be found by virtue of a table or using some computational software such as `Mathematica`. The normal convergence of the resulting series guarantees that the inverse transform of the expression is exactly the series of the inverse transforms of the resulting exponentials), we obtain

$$\mathbb{E}[N_t^\varepsilon] = 4\sum_{k\geq 1}(2k-1)\,\mathrm{erfc}\left(\frac{(2k-1)\varepsilon}{\sqrt{2t}}\right) - k\ \mathrm{erfc}\left(\frac{2k\varepsilon}{\sqrt{2t}}\right), \tag{4.174}$$

which converges quickly for large ε. For $\varepsilon \to 0$, we can get a quickly converging expression by recalling the Mittag-Leffler expansion of the hyperbolic cosecant,

$$\begin{aligned}\frac{\mathrm{csch}^2(\varepsilon\sqrt{2\lambda})}{\lambda} &= \frac{1}{\lambda}\sum_{k\in\mathbb{Z}}\frac{1}{(\varepsilon\sqrt{2\lambda} - i\pi k)^2}\\ &= \frac{1}{2\varepsilon^2\lambda^2} + \frac{1}{\lambda}\sum_{k\geq 1}\frac{4\varepsilon^2\lambda - 2\pi^2k^2}{(2\varepsilon^2\lambda + \pi^2k^2)^2}\,.\end{aligned}$$

We can take the inverse Laplace transform termwise by using the residue theorem, due to the absolute and uniform convergence of the expression. After some algebra, this operation results in

$$\mathbb{E}[N_t^\varepsilon] = \frac{t}{2\varepsilon^2} + \frac{2}{3} + 2\sum_{k\geq 1}(2(-1)^k - 1)\frac{e^{-\pi^2k^2t/2\varepsilon^2}t}{\varepsilon^2}\left[1 + \frac{\varepsilon^2}{\pi^2k^2t}\right], \tag{4.175}$$

which confirms that $\mathbb{E}[N_t^\varepsilon]$ is extremely well approximated by $\frac{t}{2\varepsilon^2} + \frac{2}{3}$ when ε is small. ■

From the alternative proof of proposition 4.4 and the formulas above give the functional equation of ζ_B we can naturally retrieve the functional equation of the Riemann ζ-function. This is part of the usual folklore of ζ and similar functions, where functional equations are essentially related by Poisson summation.

Proposition 4.5. *Defining*

$$\eta_B(p) := (3\cdot 2^p - 8)(p-2)(2\pi t)^{-\frac{p}{2}}\zeta_B(p), \tag{4.176}$$

The functional equation of ζ_B is given by

$$\eta_B(p) = \eta_B(3-p)\,. \tag{4.177}$$

In particular, as expected from the symmetry of ζ, the axis of symmetry of η_B is $\mathrm{Re}(p) = \frac{3}{2}$.

Finally, we can calculate the moments of N_t^ε. After some calculations, we obtain

$$\mathcal{L}(\mathbb{E}[(N^\varepsilon)^s])(\lambda) = \frac{1}{\lambda}\left[\sinh^2(\varepsilon\sqrt{2\lambda})\,\mathrm{Li}_{-s}(\mathrm{sech}^2(\varepsilon\sqrt{2\lambda})) - \tanh^2(\varepsilon\sqrt{\lambda/2})\right]. \tag{4.178}$$

Distribution of the length of the kth longest branch

Following the discussion of section 3.2, we know that for $k \geq 2$, we can calculate the moment generating function $G_\alpha(z;p)$, noticing that Brownian motion is a 2-stable process ($\alpha = 2$). Then,

$$\frac{z}{(z\mathbb{E}[e^{-\varepsilon U}])^{-1} - 1} = \frac{2z^2}{\cosh\left(2\sqrt{2\varepsilon}\right) - 2z + 1}\,. \tag{4.179}$$

However, taking the Mellin transform of this expression is not easily feasible. To do so, we will write the above expression as a geometric series of decaying exponentials. Denoting $y := e^{-2\sqrt{2\varepsilon}}$, we can write the expression above as

$$\frac{4z^2y}{y^2 - 2(2z-1)y + 1} = \frac{4z^2y}{(y-y_+)(y-y_-)}, \tag{4.180}$$

where $y_\pm$ are the roots of the polynomial in the denominator of the expression, namely

$$y_\pm = 2z - 1 \pm 2i\sqrt{z(1-z)}\,. \tag{4.181}$$

Partial fraction decomposition entails

$$\frac{4z^2y}{y^2 - 2(2z-1)y + 1} = \frac{A(z)y}{y-y_+} - \frac{A(z)y}{y-y_-}, \tag{4.182}$$

where

$$A(z)=\frac{4z^2}{y_+-y_-}=\frac{z^2}{\sqrt{z(z-1)}}\,. \tag{4.183}$$

Finally, we may express each of the terms above as a geometric series. Summing both terms,

$$-A(z)\sum_{k\geq 1}\left(\frac{y_-^k-y_+^k}{y_+^k y_-^k}\right)y^k=4z^2\sum_{k\geq 1}\frac{y_+^k-y_-^k}{y_+-y_-}y^k \tag{4.184}$$

Recalling that $y=e^{-2\sqrt{2\varepsilon}}$ and taking the Mellin transform with respect to ε

$$\mathcal{M}\left[\frac{z}{(z\mathbb{E}[e^{-\varepsilon U}])^{-1}-1}\right](p)=2^{3-3p}\Gamma(2p)z^2\frac{\mathrm{Li}_{2p}(y_+(z))-\mathrm{Li}_{2p}(y_-(z))}{y_+(z)-y_-(z)}\,. \tag{4.185}$$

Finally, the generating function can be written as

$$G_2(z;p)=8\,\frac{\Gamma(p)}{\Gamma(\frac{p}{2})}\left(\frac{t}{8}\right)^{\frac{p}{2}}z^2\,\frac{\mathrm{Li}_p(y_+(z))-\mathrm{Li}_p(y_-(z))}{y_+(z)-y_-(z)}\,. \tag{4.186}$$

When z is in the vicinity of 0, y_+ and y_- are both complex, it is thus *a priori* not obvious that the quantity defined above should remain real for real p. However, this must be so, since

$$\frac{y_+^k-y_-^k}{y_+-y_-}=a_k(z)\,, \tag{4.187}$$

where $a_k(z)$ is the solution to the following difference equation

$$a_k(z)=2(2z-1)\,a_{k-1}-a_{k-2}\,, \tag{4.188}$$

with seed $a_0=0$ and $a_1=1$. In fact, it is possible to express $a_k(z)$ as defined in equation 4.188 in terms of the Chebyshev polynomials of the second kind U_k

$$a_k(z)=U_{k-1}(2z-1)\quad\text{and}\quad a_k(0)=(-1)^{k-1}k\,, \tag{4.189}$$

and incidentally $\frac{a_k}{a_{k-1}}$ corresponds to the kth convergent of the continuous fraction

$$2(2z-1)-\cfrac{1}{2(2z-1)-\cfrac{1}{2(2z-1)-\cfrac{1}{\ddots}}}\quad. \tag{4.190}$$

Using these relations, it is possible to rewrite G_2 as

$$G_2(z;p)=8\,\frac{\Gamma(p)}{\Gamma(\frac{p}{2})}\left(\frac{t}{8}\right)^{\frac{p}{2}}z^2\sum_{k\geq 1}\frac{a_k(z)}{k^p}\,. \tag{4.191}$$

Since ultimately what interests us are the derivatives of this function at 0, we can rewrite $G_2(z;p)$ formally as

$$G_2(z;p) = 8\,\frac{\Gamma(p)}{\Gamma(\frac{p}{2})}\left(\frac{t}{8}\right)^{\frac{p}{2}} z^2 \sum_{n=0}^{\infty}\sum_{k\geq 1} \frac{a_k^{(n)}(0)}{n!}\,\frac{z^n}{k^p}\,. \tag{4.192}$$

The problem thus boils down to effectively computing the coefficients $a_k^{(n)}(0)$. To do so, we can consider augmenting the recurrence problem to phase space $(a_k, a_k^{(1)}, a_k^{(2)}, \cdots, a_k^{(n)})$ noticing the following relation

$$a_k^{(n)}(z) = 2(2z-1)a_{k-1}^{(n)}(z) + 4na_{k-1}^{(n-1)}(z) - a_{k-2}^{(n)}(z) \tag{4.193}$$

Setting $z = 0$ and $a_k^{(n)}(0) = a_k^{(n)}$ in this relation yields

$$a_k^{(n)} = -2a_{k-1}^{(n)} + 4na_{k-1}^{(n-1)} - a_{k-2}^{(n)}\,. \tag{4.194}$$

For example, it is easy to verify that

$$a_k^{(1)} = \frac{2}{3}(-1)^k k(k^2-1) \quad \text{and} \quad a_k^{(2)} = \frac{2^2}{15}(-1)^{k-1}k(k^2-1)(k^2-2^2)\,. \tag{4.195}$$

In general, the following formula holds

$$a_k^{(n)} = \frac{2^n(-1)^{k+n+1}}{k}\prod_{i=0}^{n}\frac{k^2-i^2}{2i+1}\,. \tag{4.196}$$

Thus, the Mellin transforms of the distributions of the longest branches may be expressed as linear combinations of shifted and twisted ζ-functions, since the $a_k^{(n)}$'s are polynomials of degree $2n+1$ in k. Computing the first few terms yields

$$\begin{aligned}
\mathbb{E}[\ell_2^p(B)] &= \frac{2^{3-\frac{5p}{2}}t^{\frac{p}{2}}\Gamma(p)}{\Gamma(\frac{p}{2})}(2^p-2^2)\zeta(p-1)\\
\mathbb{E}[\ell_3^p(B)] &= \frac{2^{4-\frac{5p}{2}}t^{\frac{p}{2}}\Gamma(p)}{3\Gamma(\frac{p}{2})}[(2^p-2^2)\zeta(p-1)-(2^p-2^4)\zeta(p-3)]\\
\mathbb{E}[\ell_4^p(B)] &= \frac{2^{4-\frac{5p}{2}}t^{\frac{p}{2}}\Gamma(p)}{15\Gamma(\frac{p}{2})}[4(2^p-2^2)\zeta(p-1)-5(2^p-2^4)\zeta(p-3)+(2^p-2^6)\zeta(p-5)]\,,
\end{aligned}$$

and so on. These Mellin transforms can be inverted to yield explicit expressions of $\mathbb{P}(\ell_k \geq \varepsilon)$,

$$\begin{aligned}
\mathbb{P}(\ell_2(B)\geq\varepsilon) &= 4\sum_{k\geq 1}k\left[\operatorname{erfc}\left(k\varepsilon\sqrt{\frac{2}{t}}\right)-4\operatorname{erfc}\left(2k\varepsilon\sqrt{\frac{2}{t}}\right)\right]\\
\mathbb{P}(\ell_3(B)\geq\varepsilon) &= \frac{8}{3}\sum_{k\geq 1}k\left[4\left(4k^2-1\right)\operatorname{erfc}\left(2k\varepsilon\sqrt{\frac{2}{t}}\right)-\left(k^2-1\right)\operatorname{erfc}\left(k\varepsilon\sqrt{\frac{2}{t}}\right)\right]\\
\mathbb{P}(\ell_4(B)\geq\varepsilon) &= \frac{8}{15}\sum_{k\geq 1}k\left[\left(k^4-5k^2+4\right)\operatorname{erfc}\left(k\varepsilon\sqrt{\frac{2}{t}}\right)-16\left(4k^4-5k^2+1\right)\operatorname{erfc}\left(2k\varepsilon\sqrt{\frac{2}{t}}\right)\right]\,.
\end{aligned}$$

These calculations can be performed for any ℓ_k without any additional difficulty.

Local ζ-functions and corrections to the local time

In what will follow, we will denote

$$\varphi(x,t) := \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} . \tag{4.197}$$

Theorem 4.6. *The local ζ-function of Brownian motion at the level $x > 0$ admits a meromorphic continuation to the entire complex plane given by*

$$\zeta_B^x(p) = 2^{-\frac{3p}{2}} \left(2^p - 1\right) t^{\frac{p}{2}} \zeta(p)\Gamma(p+1) \left(\frac{{}_1F_1\left(\frac{-p}{2}; \frac{1}{2}; \frac{-x^2}{2t}\right)}{\Gamma\left(\frac{p}{2}+1\right)} - \sqrt{\frac{2x^2}{t}} \frac{{}_1F_1\left(\frac{1-p}{2}; \frac{3}{2}; \frac{-x^2}{2t}\right)}{\Gamma\left(\frac{p+1}{2}\right)} \right) . \tag{4.198}$$

Proof. It suffices to calculate the propagatators, which in this case can be done explicitly. Using the reflecting property, for any $a > 0$,

$$\langle a|0\rangle = \langle 0|a\rangle = e^{-a\sqrt{2\lambda}} . \tag{4.199}$$

So by our previous work

$$\mathcal{L}(\mathbb{E}\left[N_t^{x,x+\varepsilon}\right])(\lambda) = \frac{e^{-(x+\varepsilon)\sqrt{2\lambda}}}{\lambda(1 - e^{-2\varepsilon\sqrt{2\lambda}})} . \tag{4.200}$$

The Mellin transform of this expression can be calculated easily to be

$$\mathcal{ML}(\mathbb{E}\left[N_t^{x,x+\varepsilon}\right])(p,\lambda) = 2^{-\frac{3p}{2}} \left(2^p - 1\right) \lambda^{-1-\frac{p}{2}} e^{-x\sqrt{2\lambda}}\, \Gamma(p)\zeta(p) . \tag{4.201}$$

Using lemma 2.28 and inverting the Laplace transform we get the desired expression. ∎

We can also extract asymptotic relations for $\mathbb{E}\left[N_t^{x,x+\varepsilon}\right]$.

Proposition 4.7. *For Brownian motion and $x > 0$,*

$$\begin{aligned}\mathbb{E}\left[N^{x,x+\varepsilon}\right] &= \sum_{k=1}^{\infty} \mathit{erfc}\left(\frac{x + (2k-1)\varepsilon}{\sqrt{2t}}\right) \\ &\sim \frac{1}{2\varepsilon}\int_0^t \varphi(x,s)\, ds + \sum_{k\geq 0} \frac{4(-2)^k \left(2^{2k+1} - 1\right)\zeta(2k+2)}{\pi^{2k+2}} \left[\frac{\partial^k}{\partial t^k}\varphi(x,t)\right] \varepsilon^{2k+1} \quad \text{as } \varepsilon \to 0 .\end{aligned}$$

Proof. To deduce an asymptotic expression of $\mathbb{E}\left[N^{x,x+\varepsilon}\right]$, it is much easier to consider the problem in terms of the dual variables p and λ. Notice that $\mathcal{ML}(\mathbb{E}\left[N^{x,x+\varepsilon}\right])$ only has poles (in p) at every odd negative integer (stemming from those of Γ and the non-presence of trivial zeros of ζ) and at 1 (stemming from the pole of the ζ-function). The residues of these simple poles can be calculated

$$\mathrm{Res}(\mathcal{ML}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big]),1) = \frac{e^{-x\sqrt{2\lambda}}}{2\sqrt{2}\lambda^{3/2}}$$
$$\mathrm{Res}(\mathcal{ML}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big]),-(2k+1)) = \frac{2\left(2^{2k+1}-1\right)\zeta(-2k-1)}{(2k+1)!}e^{-x\sqrt{2\lambda}}(2\lambda)^{k-\frac{1}{2}}\,.$$

The rapid enough decay of $\mathcal{ML}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big])$ on the upper and lower ends of a vertical strip in the complex plane allows us to apply the inverse Mellin theorem and contour integration to obtain an asymptotic series of $\mathcal{L}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big])(\lambda)$

$$\mathcal{L}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big])(\lambda) \sim e^{-x\sqrt{2\lambda}}\left[\frac{1}{2\varepsilon\lambda^{3/2}\sqrt{2}} + \sum_{k=0}^{n-1}\frac{4(-1)^k\left(2^{2k+1}-1\right)\zeta(2k+2)}{\pi^{2k+2}}(2\lambda)^{k-\frac{1}{2}}\varepsilon^{2k+1} + O(\varepsilon^{2n+1})\right]\,.$$

Noticing that

$$\mathcal{L}\left[2^k\frac{\partial^k}{\partial t^k}\varphi(x,t)\right](\lambda) = e^{-x\sqrt{2\lambda}}(2\lambda)^{k-\frac{1}{2}}\ , \tag{4.202}$$

one can write that as $\varepsilon \to 0$

$$\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big] \sim \frac{1}{2\varepsilon}\int_0^t \varphi(x,s)\,ds + \sum_{k\geq 0}\frac{4(-2)^k\left(2^{2k+1}-1\right)\zeta(2k+2)}{\pi^{2k+2}}\left[\frac{\partial^k}{\partial t^k}\varphi(x,t)\right]\varepsilon^{2k+1}$$

This series formally diverges, so we consider it as an asymptotic series only. From this formula we retrieve some classical results from probability theory [51], namely that the local time is a good first order approximation of $N_t^{x,x+\varepsilon}$ (in fact almost surely) as well as an explicit expression for the expected value of the local time of Brownian motion. One can find a converging representation of $\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big]$ by expanding out its Laplace transform as a sum of exponentials as done for N_t^{ε}. In this case, one obtains a series which converges absolutely and uniformly on every compact set of $x \geq 0$ and $\varepsilon > 0$

$$\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big] = \sum_{k=1}^{\infty}\mathrm{erfc}\left(\frac{x+(2k-1)\varepsilon}{\sqrt{2t}}\right)\,. \tag{4.203}$$

■

The distribution of $N_t^{x,x+\varepsilon}$ is at this point in principle accessible given out previous work. After some algebra, we find that

$$\mathcal{L}(\mathbb{E}\Big[(N_t^{x,x+\varepsilon})^{s-1}\Big])(\lambda) = \frac{2e^{-x\sqrt{2\lambda}}}{\lambda}\mathrm{Li}_{-s+1}\left(e^{-2\varepsilon\sqrt{2\lambda}}\right)\sinh(\varepsilon\sqrt{2\lambda}) \tag{4.204}$$

where Li denotes the polylogarithm. Now, recalling the discussion of section 3.3,

$$2\varepsilon N_t^{x,x+\varepsilon} \xrightarrow[\varepsilon\to 0]{L^s} L_t^x(B)\,, \tag{4.205}$$

where $L_t^x(B)$ denotes the local time of the Brownian motion.

Remark 4.8. In fact, for Brownian it is true that

$$2\varepsilon N_t^{x,x+\varepsilon} \xrightarrow[\varepsilon\to 0]{\text{a.s.}} L_t^x(B)\,. \tag{4.206}$$

We have

$$\mathcal{L}(\mathbb{E}\Big[(2\varepsilon N_t^{x,x+\varepsilon})^{s-1}\Big])(\lambda) = 2^{\frac{1-s}{2}}\lambda^{-\frac{s+1}{2}}e^{-x\sqrt{2\lambda}}\,\Gamma(s) + o(1) \quad \text{as } \varepsilon \to 0\,, \tag{4.207}$$

so in particular, following the discussion of section 3.3 this entails that

$$\mathcal{ML}(\mathbb{P}(L_t^x = w))(s,\lambda) = 2^{\frac{1-s}{2}}\lambda^{-\frac{s+1}{2}}e^{-x\sqrt{2\lambda}}\,\Gamma(s)\,. \tag{4.208}$$

Inverting the Mellin transform can be easily done in this case. Doing so for $w > 0$, we obtain

$$\mathcal{L}(\mathbb{P}(L_t^x = w))(\lambda) = \frac{e^{-(x+w)\sqrt{2\lambda}}}{\sqrt{\lambda}}\,. \tag{4.209}$$

Finally, the inverse Laplace transform of this now familiar expression is

$$\mathbb{P}(L_t^x = w) = 2\varphi(x+w,t) = \frac{2}{\sqrt{2\pi t}}e^{-(x+w)^2/2t}\,. \tag{4.210}$$

Of course, this is a classical result that can be obtained in variety of different ways. Nevertheless, for more complicated processes, it might be possible to retrieve the distribution of their local times in this manner. Notice that the distribution of L_t^x has an atom at $w = 0$, corresponding to the probability of not reaching level x, which is why the integral of the probability density above is not 1.

Remark 4.9. Inverting the Laplace transform first in equation 4.208, we immediately retrieve a formula for all the moments of this distribution.

Calculation of the average persistence diagram

In accordance to the discussion of section 2.7, using the results of proposition 4.7 the density of the persistence diagram of Brownian motion can be computed to be

Proposition 4.10. *For $x > 0$ and $\varepsilon > 0$, the density of the average persistence diagram of Brownian motion in birth-persistence coordinates (cf. remark 2.32) is*

$$g(x,\varepsilon) = \sqrt{\frac{2}{\pi t^3}}\sum_{k=1}^{\infty}(2k-1)(x+(2k-1)\varepsilon)\,e^{-\frac{(x+(2k-1)\varepsilon)^2}{2t}} \tag{4.211}$$

4.2 Reflected Brownian motion

We can carry out a similar procedure for the reflected Brownian motion.

Associated ζ-function and N^ε

Theorem 4.11. *The ζ-function of the process $|B|$ is*

$$\zeta_{|B|}(p) = \frac{2^{1-\frac{p}{2}}(2^p-2)t^{\frac{p}{2}}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right)\zeta(p-1) \tag{4.212}$$

which has a unique pole at $p=2$ of residue t.

Proof. The theorem immediately follows from applying proposition 2.25. ∎

We immediately deduce by inverting the Mellin transform that

Proposition 4.12. *The function $\mathbb{E}[N_t^\varepsilon]$ admits the following representations for reflected Brownian motion*

$$\begin{aligned}\mathbb{E}[N_t^\varepsilon] &= \sum_{k\geq 1} 2k\left[\operatorname{erfc}\left(\frac{k\varepsilon}{\sqrt{2t}}\right) - 2\operatorname{erfc}\left(\frac{2k\varepsilon}{\sqrt{2t}}\right)\right] \\ &= \frac{1}{2\varepsilon^2}+\frac{1}{6}+\sum_{k\geq 1}\frac{4\pi^2k^2e^{-\frac{2\pi^2k^2}{\varepsilon^2}}+\varepsilon^2e^{-\frac{2\pi^2k^2}{\varepsilon^2}}-2e^{-\frac{\pi^2k^2}{2\varepsilon^2}}\left(\pi^2k^2+\varepsilon^2\right)}{\pi^2k^2\varepsilon^2}\,.\end{aligned}$$

Calculating the propagators stem from classical results [15],

$$\langle x+\varepsilon|x\rangle\langle x|x+\varepsilon\rangle = \frac{e^{-\varepsilon\sqrt{2\lambda}}\cosh(x\sqrt{2\lambda})}{\cosh((x+\varepsilon)\sqrt{2\lambda})}\,, \tag{4.213}$$

so

$$\mathcal{L}(\mathbb{E}\left[N_t^{x,x+\varepsilon}\right])(\lambda) = \frac{e^{-x\sqrt{2\lambda}}}{\lambda}\operatorname{csch}(\varepsilon\sqrt{2\lambda}) \tag{4.214}$$

Inverting the Laplace transform, we get

Proposition 4.13. *For $x>0$ and $\varepsilon>0$, we can express $\mathbb{E}[N^{x,x+\varepsilon}]$ in the following form*

$$\mathbb{E}\left[N^{x,x+\varepsilon}\right] = 2\sum_{k\geq 0}\operatorname{erfc}\left[\frac{x+(2k+1)\varepsilon}{\sqrt{2t}}\right]\,. \tag{4.215}$$

4.3 Brownian motion with drift

Throughout this section, we will denote $B_t^{\mu,\sigma}$ the Lévy process defined by

$$B_t^{\mu,\sigma} = \mu t + \sigma B_t\,. \tag{4.216}$$

We will assume that $\sigma > 0$ and without loss of further generality that $\mu \geq 0$. By the scale invariance and almost sure $(\frac{1}{2}-\delta)$-Hölder continuity of Brownian motion, it follows that the process $B^{\mu,\sigma}$ almost surely leaves and never returns to any compact set $[0,x]$ when studied over the ray $[0,\infty[$. Using this and the Markov property of $B^{\mu,\sigma}$, the following was shown by Baryshnikov.

Proposition 4.14 (Baryshnikov, [9]). *For $\sigma = 1$ and $x > 0$, $B^{\mu,1}$ satisfies*

$$\mathbb{E}\left[N_{B^{\mu,1}}^{x,x+\varepsilon}\right] = \frac{1}{e^{2\mu\varepsilon} - 1} \tag{4.217}$$

on the infinite ray $[0, \infty[$.

Following this result, we may deduce the local ζ-function associated to $B^{\mu,1}$ is given by the following expression

$$\zeta_{B^{\mu,1}}^{x}(p) = (2\mu)^{-p}\Gamma(p+1)\zeta(p)\,. \tag{4.218}$$

Of course, we would like to have a similar result over a compact set $[0, t]$. However, the computations in this restricted space turn out to be somewhat more challenging. Applying our theory of propagators detailed in section 3.3, and noticing that $B^{\mu,\sigma}$ is Lévy, it suffices to compute $\langle 0|a\rangle$ to gain access to the Laplace transform of $\mathbb{E}\left[N_t^{x,x+\varepsilon}\right]$.

Proposition 4.15. *The propagators of the process $B^{\mu,\sigma}$ for $a > 0$ are given by*

$$\begin{aligned}\langle 0|a\rangle &= e^{\frac{a\mu - a\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}}\\ \langle a|0\rangle &= e^{\frac{-a\mu - a\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}}\,.\end{aligned}$$

Proof. Since $B^{\mu,\sigma}$ is a Lévy process, we start by noticing that $\langle a|0\rangle = \langle 0|-a\rangle$, so that it suffices to calculate $\langle 0|a\rangle$ and $\langle 0|-a\rangle$. Denote M_t^θ the martingale defined by

$$M_t^\theta = \frac{e^{\theta B_t^{\mu,\sigma}}}{\mathbb{E}\left[e^{\theta B_t^{\mu,\sigma}}\right]}\,. \tag{4.219}$$

We may compute the expectation in the denominator to be exactly

$$\mathbb{E}\left[e^{\theta B_t^{\mu,\sigma}}\right] = e^{\theta\mu t + \frac{1}{2}\theta^2\sigma^2 t}\,. \tag{4.220}$$

Further denoting

$$T^a = \inf\{t \geq 0 \,|\, B_t^{\mu,\sigma} = a\}\,, \tag{4.221}$$

we can calculate the propagators by considering the two following cases

1. $\langle 0|a\rangle$. In this case, M_t^θ is bounded for $\theta > 0$ and by an application of Doob's stopping theorem,
$$\langle 0|a\rangle = \mathbb{E}\left[e^{-\lambda T^a}\right] = e^{\frac{a\mu - a\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}}\,; \tag{4.222}$$
2. $\langle 0|-a\rangle$. In this case, $M_{t\wedge T^a}^\theta$ is bounded for $\theta < 0$, so that once again by the stopping theorem
$$\langle 0|-a\rangle = \mathbb{E}\left[e^{-\lambda T^{-a}}\right] = e^{\frac{-a\mu - a\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}}\,, \tag{4.223}$$

which finishes the proof. ∎

Following the discussion of section 3.3,

$$\mathcal{L}(\mathbb{E}\Big[N_t^{x,x+\varepsilon}\Big])(\lambda) = \frac{e^{x\frac{\mu-\sqrt{\mu^2+2\lambda\sigma^2}}{\sigma^2}}}{2\lambda} e^{\frac{\varepsilon\mu}{\sigma^2}} \operatorname{csch}\left(\frac{\varepsilon\sqrt{\mu^2+2\lambda\sigma^2}}{\sigma^2}\right). \tag{4.224}$$

Taking the Mellin transform we get

$$\mathcal{M}\mathcal{L}(\mathbb{E}\Big[N_{B^{\mu,\sigma}}^{x,x+\varepsilon}\Big])(p,\lambda) = \frac{e^{\frac{x\left(\mu-\sqrt{\mu^2+2\lambda\sigma^2}\right)}{\sigma^2}}}{\lambda}\left(\frac{\sigma^2}{2\sqrt{\mu^2+2\lambda\sigma^2}}\right)^p \Gamma(p)\zeta\left(p, \frac{1}{2} - \frac{\mu}{2\sqrt{\mu^2+2\lambda\sigma^2}}\right),$$

where ζ now denotes the Hurwitz ζ-function, namely

$$\zeta(p,z) = \sum_{k\geq 1} \frac{1}{(k+z)^p}, \tag{4.225}$$

which also admits a meromorphic continuation to the entire complex plane, so we can conclude that $\zeta^x_{B^{\mu,\sigma}}$ does as well. Again, we notice the presence of poles at every negative integer and at $p = 1$. However, inverting the Laplace transform of this expression is difficult as soon as $\mu > 0$, however, as expected, setting $\mu = 0$ and $\sigma = 1$, we retrieve the local ζ-function of Brownian motion.

Distribution of $N_{B^{\mu,\sigma}}^{x,x+\varepsilon}$ and the local time $L_t^x(B^{\mu,\sigma})$

As before, it is possible to calculate the moments of the occupation numbers $N_{B^{\mu,\sigma}}^{x,x+\varepsilon}$, the result of this calculation is

$$\mathcal{L}(\mathbb{E}\Big[(N_{B^{\mu,\sigma}}^{x,x+\varepsilon})^{s-1}\Big])(\lambda) = \frac{2e^{\frac{x}{\sigma^2}\left(\mu-\sqrt{\mu^2+2\lambda\sigma^2}\right)}e^{\frac{\mu\varepsilon}{\sigma^2}}}{\lambda} \operatorname{Li}_{-s+1}\left(e^{-\frac{2\varepsilon}{\sigma^2}\sqrt{\mu^2+2\lambda\sigma^2}}\right)\sinh\left(\frac{\varepsilon\sqrt{\mu^2+2\lambda\sigma^2}}{\sigma^2}\right), \tag{4.226}$$

which implies that for $x > 0$,

$$\mathcal{M}_w\mathcal{L}(\mathbb{P}(L_t^x = w))(\lambda, s) = e^{\frac{x\left(\mu-\sqrt{2\lambda\sigma^2+\mu^2}\right)}{\sigma^2}}\left(\frac{\sigma^2}{\sqrt{\mu^2+2\lambda\sigma^2}}\right)^{s-1}\Gamma(s) \tag{4.227}$$

So for $w > 0$,

$$\mathcal{L}(\mathbb{P}(L_t^x = w))(\lambda) = \frac{e^{\frac{x\mu}{\sigma^2}}\sqrt{\mu^2+2\lambda\sigma^2}}{\lambda\sigma^2} e^{-\frac{(x+w)}{\sigma^2}\sqrt{\mu^2+2\lambda\sigma^2}} \tag{4.228}$$

Leading to

$$\mathbb{P}(L_t^x = w) = e^{\frac{x\mu}{\sigma^2}}\int_0^t \frac{(x+w)^2 - s\sigma^2}{\sqrt{2\pi s^5\sigma^6}} e^{-\frac{(x+w)^2+\mu^2 s^2}{2s\sigma^2}}\, ds\,.$$

4.4 Ornstein-Uhlenbeck process

Let us now consider the case of the Ornstein-Uhlenbeck process, *i.e.* the Itô diffusion satisfying the following SDE (for $\theta > 0$ and $\sigma > 0$)

$$dX_t = -\theta X_t dt + \sigma dB_t \,. \tag{4.229}$$

From our discussion in section 3.3, it suffices to find the solutions to the differential equation

$$\frac{\sigma^2}{2}\frac{\partial^2\Psi}{\partial x^2} - \theta x \frac{\partial \Psi}{\partial x} = \lambda \Psi \,. \tag{4.230}$$

The solutions to the above equations are known, as this is nothing other than Hermite's differential equation. We can identify the two linearly independent solutions satisfying the respective boundary conditions to be

$$\Psi_\lambda(x) = {}_1F_1\left(\frac{\lambda}{2\theta}; \frac{1}{2}; \frac{x^2\theta}{\sigma^2}\right) \quad \text{and } \Phi_\lambda(x) = \mathrm{He}\left(-\frac{\lambda}{\theta}; \frac{x\sqrt{\theta}}{\sigma}\right) . \tag{4.231}$$

where ${}_1F_1$ denotes the confluent hypergeometric function and $\mathrm{He}(\mu; z)$ denotes the μth Hermite polynomial. The propagators of the Ornstein-Uhlenbeck process are readily given products of these functions, as mentionned in section 3.3, and expressions for $\mathcal{L}\mathbb{E}\left[N_t^{x,x+\varepsilon}\right]$ can be deduced from equation 4.159. For $x = 0$, the expression obtained simplifies considerably to give

$$\mathcal{L}[\mathbb{E}\left[N_t^{0,\varepsilon}\right]](\lambda) = \frac{\sigma}{\lambda\varepsilon\sqrt{\theta}} \frac{\Gamma\left(\frac{\lambda}{2\theta}\right)}{\Gamma\left(\frac{\theta+\lambda}{2\theta}\right) {}_1F_1\left(\frac{\theta+\lambda}{2\theta}; \frac{3}{2}; \frac{\varepsilon^2\theta}{\sigma^2}\right)} \,. \tag{4.232}$$

This Laplace transform can in principle be formally inverted by virtue of the residue theorem. The expression above admits poles at every $\lambda = -2\theta k$, but there are also poles stemming from the zeros of the confluent hypergeometric function in the denominator, which are more difficult to locate. Nonetheless denoting $\mathcal{P}$ the set of poles stemming from ${}_1F_1$, we get the following formal expression

$$\begin{aligned}\mathbb{E}\left[N_t^{0,\varepsilon}\right] = &\frac{\sqrt{\pi}\sigma(2\theta t + \log(4))\,\mathrm{erfi}\left(\frac{\varepsilon\sqrt{\theta}}{\sigma}\right) - 2\varepsilon\sqrt{\theta}{}_1F_1^{(1,0,0)}\left(\frac{1}{2}; \frac{3}{2}; \frac{\theta\varepsilon^2}{\sigma^2}\right)}{\pi^{3/2}\sigma \mathrm{erfi}\left(\frac{\varepsilon\sqrt{\theta}}{\sigma}\right)^2} \\ &- \frac{\sigma}{2\varepsilon\sqrt{\theta}} \sum_{k\geq 1} \frac{(-1)^k e^{-2\theta k t}}{k^2\,\Gamma\left(\frac{1-2k}{2}\right)\Gamma(k)\,{}_1F_1\left(\frac{1-2k}{2}; \frac{3}{2}; \frac{\varepsilon^2\theta}{\sigma^2}\right)} + \sum_{z\in\mathcal{P}} \mathrm{Res}(e^{\lambda t}\mathcal{L}\mathbb{E}\left[N^{0,\varepsilon}\right], z) \,.\end{aligned}$$

More interestingly, equation 4.159 can be used to yield the distribution of the local time of the Ornstein-Uhlenbeck process. The expresions for $x > 0$ are rather involved, so we will study the case $x = 0$ explicitly only. The simplification that occurs for the 0 level is due to the supplementary scaling symmetry of the latter. Taking the limit as $\varepsilon \to 0$ of $\mathcal{L}\mathbb{E}\left[(N_t^{0,\varepsilon})^{s-1}\right]$ we obtain

$$\mathcal{M}_w\mathcal{L}(\mathbb{P}(L_t^0 = w))(\lambda) = \frac{\Gamma(s)}{\lambda}\left(\frac{\sigma\Gamma\left(\frac{\lambda}{2\theta}\right)}{\sqrt{\theta}\Gamma\left(\frac{\theta+\lambda}{2\theta}\right)}\right)^{s-1} . \tag{4.233}$$

This expression yields the Laplace transform of all the moments of L_t^0. For integer $s \geq 2$, these Laplace transforms can be formally inverted by using the residue theorem and computing the residues of the expression above. For example,

$$\begin{aligned}
\mathbb{E}\Big[L_t^0\Big] &= \frac{\sigma}{\sqrt{\theta}}\frac{(2\theta t + \log(4))}{\sqrt{\pi}} + \frac{\sigma}{\sqrt{\theta}}\sum_{k\geq 1}\frac{(-1)^{k+1}e^{-2\theta k t}}{k^2\Gamma(\frac{1-2k}{2})\Gamma(k)} \\
\mathbb{E}\Big[(L_t^0)^2\Big] &= \frac{2\sigma^2}{\theta}\frac{\left(6(\theta t + \log(4))^2 - \pi^2\right)}{3\pi} - \frac{2\sigma^2}{\theta}\sum_{k\geq 1}\frac{e^{-2\theta k t}\left(-2kH_{-k-\frac{1}{2}} + 2kH_k + 2\theta k t + 1\right)}{k^4\Gamma(\frac{1-2k}{2})^2\Gamma(k)^2} ,
\end{aligned}$$

and so on. The H_k above denote the kth Harmonic number, *i.e.*

$$H_k = \sum_{n=1}^{k}\frac{1}{n} = \int_0^1 \frac{1-z^k}{1-z}\, dz , \tag{4.234}$$

where the second representation of H_k holds for non-integer k. Finally, taking the inverse Mellin transform we get

$$\mathcal{L}[\mathbb{P}(L_t^0 = w)](\lambda) = \frac{\sqrt{\theta}\Gamma\left(\frac{\theta+\lambda}{2\theta}\right)}{\lambda\sigma\Gamma\left(\frac{\lambda}{2\theta}\right)}\exp\left[-\frac{w\sqrt{\theta}\Gamma\left(\frac{\theta+\lambda}{2\theta}\right)}{\sigma\Gamma\left(\frac{\lambda}{2\theta}\right)}\right] . \tag{4.235}$$

This expression exhibits essential singularities at every $\lambda = -(2k-1)\theta$ for every $k \in \mathbb{N}$. For $x > 0$, it is possible to treat the problem analogously, but the expressions involved – while expressible in terms of Hermite polynomials and confluent hypergeometric functions – remain quite involved. Nonetheless, the elementary methods introduced in this paper allow us to calculate such quantities, which have not been quantitatively studied until recently in [56] via perturbative methods.

For the sake of completeness we give the expression of $\mathcal{M}_w\mathcal{L}\mathbb{P}(L_t^x = w)(\lambda, s)$ for $x > 0$,

$$\mathcal{M}_w\mathcal{L}\mathbb{P}(L_t^x = w)(\lambda, s) = \frac{\Gamma(s)}{{}_1F_1(\frac{\lambda}{2\theta};\frac{1}{2};\frac{x^2\theta}{\sigma^2})\lambda}\left[\frac{\lambda}{\sigma\sqrt{\theta}}\frac{\mathrm{He}(-\frac{\lambda+\theta}{\lambda},\frac{x\sqrt{\theta}}{\sigma})}{\mathrm{He}(-\frac{\lambda}{\theta},\frac{x\sqrt{\theta}}{\sigma})} + \frac{x\lambda}{\sigma^2}\frac{{}_1F_1\left(1+\frac{\lambda}{2\theta};\frac{3}{2};\frac{x^2\theta}{\sigma^2}\right)}{{}_1F_1\left(\frac{\lambda}{2\theta};\frac{1}{2};\frac{x^2\theta}{\sigma^2}\right)}\right]^{1-s} . \tag{4.236}$$

Studying the poles in λ of this expression, it is in principle possible to give expressions for the moments of the local times at x, as we did in the case of $x = 0$.

Bibliography

[1] H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research*, 18(8):1–35, 2017. URL http://jmlr.org/papers/v18/16-337.html.

[2] H. Adams, M. Aminian, E. Farnell, M. Kirby, J. Mirth, R. Neville, C. Peterson, and C. Shonkwiler. A fractal dimension for measures via persistent homology. *Abel Symposia*, pages 1–31, 2020. ISSN 2197-8549. doi: 10.1007/978-3-030-43408-3_1. URL http://dx.doi.org/10.1007/978-3-030-43408-3_1.

[3] R. J. Adler and J. E. Taylor. *Random Fields and Geometry*. Springer New York, 2007. ISBN 9780387481166. doi: 10.1007/978-0-387-48116-6. URL http://dx.doi.org/10.1007/978-0-387-48116-6.

[4] R. J. Adler, O. Bobrowski, M. S. Borman, E. Subag, and S. Weinberger. Persistent homology for random fields and complexes. In *Institute of Mathematical Statistics Collections*, pages 124–143. Institute of Mathematical Statistics, 2010. doi: 10.1214/10-imscoll609. URL https://doi.org/10.1214%2F10-imscoll609.

[5] R. Andreev and A. Lang. Kolmogorov-Chentsov theorem and differentiability of random fields on manifolds. *Potential Analysis*, 41(3):761–769, feb 2014. doi: 10.1007/s11118-014-9392-y. URL https://doi.org/10.1007%2Fs11118-014-9392-y.

[6] T. Aubin. *Some Nonlinear Problems in Riemannian Geometry*. Springer Berlin Heidelberg, 1998. doi: 10.1007/978-3-662-13006-3. URL https://doi.org/10.1007%2F978-3-662-13006-3.

[7] J.-M. Azaïs and M. Wschebor. *Level Sets and Extrema of Random Processes and Fields*. John Wiley & Sons, Inc., Jul 2008. ISBN 9780470409336. doi: 10.1002/9780470434642. URL http://dx.doi.org/10.1002/9780470434642.

[8] S. A. Barannikov. The framed morse complex and its invariants, dec 1994. URL https://doi.org/10.1090%2Fadvsov%2F021%2F03.

[9] Y. Baryshnikov. Time series, persistent homology and chirality. *arXiv:1909.09846*, 2019.

[10] J. Bertoin. *Lévy Processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 1996. ISBN 9780521562430. URL https://books.google.fr/books?id=A2CMMQEACAAJ.

[11] P. Biane, J. Pitman, and M. Yor. Probability laws related to the jacobi theta and riemann zeta functions, and brownian excursions. *Bulletin of the American Mathematical Society*,

38(4):435–465, jun 2001. doi: 10.1090/s0273-0979-01-00912-0. URL https://doi.org/10.1090%2Fs0273-0979-01-00912-0.

[12] J. M. Boardman and R. M. Vogt. Homotopy-everything h-spaces. *Bulletin of the American Mathematical Society*, 74(6):1117–1123, nov 1968. doi: 10.1090/s0002-9904-1968-12070-1. URL https://doi.org/10.1090%2Fs0002-9904-1968-12070-1.

[13] J. M. Boardman and R. M. Vogt. *Homotopy Invariant Algebraic Structures on Topological Spaces*. Springer Berlin Heidelberg, 1973. doi: 10.1007/bfb0068547. URL https://doi.org/10.1007%2Fbfb0068547.

[14] J.-D. Boissonnat and C. Maria. The simplex tree: An efficient data structure for general simplicial complexes. *Algorithmica*, 70(3):406–427, may 2014. doi: 10.1007/s00453-014-9887-3. URL https://doi.org/10.1007%2Fs00453-014-9887-3.

[15] A. N. Borodin and P. Salminen. *Handbook of Brownian Motion - Facts and Formulae*. Birkhäuser Basel, 2002. ISBN 9783034881630. doi: 10.1007/978-3-0348-8163-0. URL http://dx.doi.org/10.1007/978-3-0348-8163-0.

[16] J. Bretagnolle and P. Massart. Hungarian constructions from the nonasymptotic viewpoint. *The Annals of Probability*, 17(1):239–256, jan 1989. doi: 10.1214/aop/1176991506. URL https://doi.org/10.1214%2Faop%2F1176991506.

[17] P. Bubenik and T. Vergili. Topological spaces of persistence modules and their properties. *Journal of Applied and Computational Topology*, 2(3-4):233–269, dec 2018. doi: 10.1007/s41468-018-0022-4. URL https://doi.org/10.1007%2Fs41468-018-0022-4.

[18] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. ISBN 0-8218-2129-6. doi: 10.1090/gsm/033. URL https://mathscinet.ams.org/mathscinet-getitem?mr=1835418.

[19] M. Carriere, S. Oudot, and M. Ovsjanikov. Sliced Wasserstein Kernel for Persistence Diagrams. In *ICML 2017 - Thirty-fourth International Conference on Machine Learning*, pages 1–10, Sydney, Australia, Aug. 2017. URL https://hal.inria.fr/hal-01203716.

[20] F. Chazal and V. Divol. The density of expected persistence diagrams and its kernel based estimation. In B. Speckmann and C. D. Tóth, editors, *34th International Symposium on Computational Geometry (SoCG 2018)*, volume 99 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:15, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 978-3-95977-066-8. doi: 10.4230/LIPIcs.SoCG.2018.26. URL http://drops.dagstuhl.de/opus/volltexte/2018/8739.

[21] F. Chazal, M. Glisse, C. Labruère, and B. Michel. Convergence rates for persistence diagram estimation in topological data analysis. In E. P. Xing and T. Jebara, editors, *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 163–171, Bejing, China, 22–24 Jun 2014. PMLR. URL https://proceedings.mlr.press/v32/chazal14.html.

[22] F. Chazal, B. Fasy, F. Lecci, B. Michel, A. Rinaldo, and L. Wasserman. Subsampling methods for persistent homology. In F. Bach and D. Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 2143–2151, Lille, France, 07–09 Jul 2015. PMLR. URL https://proceedings.mlr.press/v37/chazal15.html.

[23] F. Chazal, W. Crawley-Boevey, and V. de Silva. The observable structure of persistence modules. *Homology, Homotopy and Applications*, 18(2):247–265, 2016. doi: 10.4310/hha.2016.v18.n2.a14. URL https://doi.org/10.4310%2Fhha.2016.v18.n2.a14.

[24] F. Chazal, V. de Silva, M. Glisse, and S. Oudot. *The Structure and Stability of Persistence Modules*. Springer International Publishing, 2016. ISBN 9783319425450. doi: 10.1007/978-3-319-42545-0. URL http://dx.doi.org/10.1007/978-3-319-42545-0.

[25] C. Chen and H. Edelsbrunner. Diffusion runs low on persistence fast. In *2011 International Conference on Computer Vision*, pages 423–430, 2011. doi: 10.1109/ICCV.2011.6126271. URL https://doi.org/10.1109/ICCV.2011.6126271.

[26] I. Chiswell. *Introduction to* Λ*-Trees*. World Scientific, feb 2001. doi: 10.1142/4495. URL https://doi.org/10.1142%2F4495.

[27] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard. Unbalanced optimal transport: Dynamic and Kantorovich formulations. *Journal of Functional Analysis*, 274(11):3090–3123, 2018. ISSN 0022-1236. doi: https://doi.org/10.1016/j.jfa.2018.03.008. URL https://www.sciencedirect.com/science/article/pii/S0022123618301058.

[28] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and Y. Mileyko. Lipschitz functions have L^p-stable persistence. *Foundations of Computational Mathematics*, 10(2):127–139, Jan 2010. ISSN 1615-3383. doi: 10.1007/s10208-010-9060-6. URL http://dx.doi.org/10.1007/s10208-010-9060-6.

[29] N. Curien, J.-F. Le Gall, and G. Miermont. The Brownian cactus I. Scaling limits of discrete cactuses. *Ann. Inst. H. Poincaré Probab. Statist.*, 49(2):340–373, 05 2013. doi: 10.1214/11-AIHP460. URL https://doi.org/10.1214/11-AIHP460.

[30] J. Curry. The fiber of the persistence map for functions on the interval. *Journal of Applied and Computational Topology*, 2(3-4):301–321, dec 2018. doi: 10.1007/s41468-019-00024-z. URL https://doi.org/10.1007%2Fs41468-019-00024-z.

[31] J. Curry, H. Hang, W. Mio, T. Needham, and O. B. Okutan. Decorated merge trees for persistent topology. *Journal of Applied and Computational Topology*, 6(3):371–428, mar 2022. doi: 10.1007/s41468-022-00089-3. URL https://doi.org/10.1007%2Fs41468-022-00089-3.

[32] T. T. N. Dang and J. Istas. Estimation of the hurst and the stability indices of a \$h\$-self-similar stable process. *Electronic Journal of Statistics*, 11(2), jan 2017. doi: 10.1214/17-ejs1357. URL https://doi.org/10.1214%2F17-ejs1357.

[33] V. Divol and T. Lacombe. Understanding the topology and the geometry of the space of persistence diagrams via optimal partial transport. *Journal of Applied and Computational Topology*, 5(1):1–53, oct 2020. doi: 10.1007/s41468-020-00061-z. URL https://doi.org/10.1007%2Fs41468-020-00061-z.

[34] V. Divol and T. Lacombe. Estimation and quantization of expected persistence diagrams. In M. Meila and T. Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 2760–2770. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/divol21a.html.

[35] V. Divol and W. Polonik. On the choice of weight functions for linear representations of persistence diagrams. *Journal of Applied and Computational Topology*, 3(3): 249–283, aug 2019. doi: 10.1007/s41468-019-00032-z. URL https://doi.org/10.1007%2Fs41468-019-00032-z.

[36] G. Doetsch. *Handbuch der Laplace-Transformation*. Birkhäuser Basel, 1950. doi: 10.1007/978-3-0348-6984-3. URL https://doi.org/10.1007%2F978-3-0348-6984-3.

[37] T. Duquesne and J.-F. Le Gall. *Random trees, Lévy processes and spatial branching processes*. Number 281 in Astérisque. Société mathématique de France, 2002. URL http://www.numdam.org/item/AST_2002__281__R1_0.

[38] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probability Theory and Related Fields*, 131(4):553–603, Nov 2004. ISSN 1432-2064. doi: 10.1007/s00440-004-0385-4. URL http://dx.doi.org/10.1007/s00440-004-0385-4.

[39] H. Edelsbrunner and J. Harer. *Computational Topology: An Introduction*. Springer-Verlag Berlin Heidelberg, 01 2010. ISBN 978-0-8218-4925-5. doi: 10.1007/978-3-540-33259-6_7. URL https://doi.org/10.1007/978-3-540-33259-6_7.

[40] S. Eilenberg and N. Steenrod. *Foundations of Algebraic Topology*. Princeton University Press, dec 1952. doi: 10.1515/9781400877492. URL https://doi.org/10.1515%2F9781400877492.

[41] S. N. Evans. *Probability and Real Trees*. Springer Berlin Heidelberg, 2008. ISBN 9783540747987. doi: 10.1007/978-3-540-74798-7. URL http://dx.doi.org/10.1007/978-3-540-74798-7.

[42] B. T. Fasy, F. Lecci, A. Rinaldo, L. Wasserman, S. Balakrishnan, and A. Singh. Confidence sets for persistence diagrams. *The Annals of Statistics*, 42(6), dec 2014. doi: 10.1214/14-aos1252. URL https://doi.org/10.1214%2F14-aos1252.

[43] A. Figalli. The optimal partial transport problem. *Archive for Rational Mechanics and Analysis*, 195(2):533–560, jan 2009. doi: 10.1007/s00205-008-0212-7. URL https://doi.org/10.1007%2Fs00205-008-0212-7.

[44] A. Figalli and N. Gigli. A new transportation distance between non-negative measures, with applications to gradients flows with dirichlet boundary conditions. *Journal de Mathématiques Pures et Appliquées*, 94(2):107–130, 2010. ISSN 0021-7824. doi: https://doi.org/10.1016/j.matpur.2009.11.005. URL https://www.sciencedirect.com/science/article/pii/S0021782409001627.

[45] P. Flajolet, X. Gourdon, and P. Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144(1):3–58, 1995. ISSN 0304-3975. doi: https://doi.org/10.1016/0304-3975(95)00002-E. URL https://www.sciencedirect.com/science/article/pii/030439759500002E.

[46] V. Ginzburg and M. Kapranov. Koszul duality for operads. *Duke Mathematical Journal*, 76 (1):203–272, oct 1994. doi: 10.1215/s0012-7094-94-07608-4. URL https://doi.org/10.1215%2Fs0012-7094-94-07608-4.

[47] C. R. Givens and R. M. Shortt. A class of Wasserstein metrics for probability distributions. *Michigan Mathematical Journal*, 31(2), jan 1984. doi: 10.1307/mmj/1029003026. URL https://doi.org/10.1307%2Fmmj%2F1029003026.

[48] A. Gut. *Stopped Random Walks*. Springer New York, 2009. doi: 10.1007/978-0-387-87835-5. URL https://doi.org/10.1007%2F978-0-387-87835-5.

[49] A. Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002. ISBN 0-521-79160-X; 0-521-79540-0.

[50] INRIA. Gudhi library. https://gudhi.inria.fr/.

[51] K. Itô and H. P. McKean. Diffusion processes and their sample paths. *Classics in Mathematics*, 1996. ISSN 1431-0821. doi: 10.1007/978-3-642-62025-6. URL http://dx.doi.org/10.1007/978-3-642-62025-6.

[52] J. Jaquette and B. Schweinhart. Fractal dimension estimation with persistent homology: A comparative study. *Communications in Nonlinear Science and Numerical Simulation*, 84:105163, 2020. ISSN 1007-5704. doi: https://doi.org/10.1016/j.cnsns.2019.105163. URL https://www.sciencedirect.com/science/article/pii/S1007570419304824.

[53] H. Jiang and Y.-H. Yang. Manifolds of positive ricci curvature with quadratically asymptotically nonnegative curvature and infinite topological type. *Communications in Analysis and Geometry*, 29(5):1233–1253, 2021. doi: 10.4310/cag.2021.v29.n5.a7. URL https://doi.org/10.4310%2Fcag.2021.v29.n5.a7.

[54] J.-P. Kahane. *Some random series of functions*. Cambridge University Press, Feb 1986. doi: 10.1017/s0013091500017521. URL http://dx.doi.org/10.1017/S0013091500017521.

[55] K. Karhunen. Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.*, (37), 1947.

[56] G. Kishore and A. Kundu. Local time of an Ornstein-Uhlenbeck particle. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(3):033218, mar 2021. doi: 10.1088/1742-5468/abe93d. URL https://doi.org/10.1088%2F1742-5468%2Fabe93d.

[57] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV's, and the sample DF. I. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 32(1-2):111–131, 1975. doi: 10.1007/bf00533093. URL https://doi.org/10.1007%2Fbf00533093.

[58] S. Kondratyev, L. Monsaingeon, and D. Vorotnikov. A new optimal transport distance on the space of finite Radon measures. *Adv. Differ. Equ.*, 21(11-12):1117–1164, 2016. ISSN 1079-9389.

[59] G. Kozma, Z. Lotker, and G. Stupp. The minimal spanning tree and the upper box dimension. *Proceedings of the American Mathematical Society*, 134(4):1183–1187, 2006. ISSN 00029939, 10886826. URL http://www.jstor.org/stable/4098085.

[60] S. M. Lane. *Categories for the Working Mathematician*. Springer New York, 1978. doi: 10.1007/978-1-4757-4721-8. URL https://doi.org/10.1007%2F978-1-4757-4721-8.

[61] J.-F. Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer International Publishing, 2016. ISBN 9783319310893. doi: 10.1007/978-3-319-31089-3. URL http://dx.doi.org/10.1007/978-3-319-31089-3.

[62] X.-J. Li. The Positivity of a Sequence of Numbers and the Riemann Hypothesis. *Journal of Number Theory*, 65(2):325–333, 1997. ISSN 0022-314X. doi: https://doi.org/10.1006/jnth.1997.2137. URL https://www.sciencedirect.com/science/article/pii/S0022314X97921375.

[63] J.-L. Loday. La renaissance des opérades. In *Séminaire Bourbaki. Volume 1994/95. Exposés 790-804*, pages 47–74, ex. Paris: Société Mathématique de France, 1996.

[64] M. Loève. *Probability Theory II*, volume 46 of *Graduate Studies in Mathematics*. Springer-Verlag New York, 4th edition, 1978. ISBN 978-0-387-90262-3.

[65] S. MacLane. *Homology*. Springer Berlin Heidelberg, 1963. ISBN 9783642620294. doi: 10.1007/978-3-642-62029-4. URL http://dx.doi.org/10.1007/978-3-642-62029-4.

[66] R. MacPherson and B. Schweinhart. Measuring shape with topology. *Journal of Mathematical Physics*, 53(7):073516, Jul 2012. ISSN 1089-7658. doi: 10.1063/1.4737391. URL http://dx.doi.org/10.1063/1.4737391.

[67] J. P. May. *The Geometry of Iterated Loop Spaces*. Springer Berlin Heidelberg, 1972. doi: 10.1007/bfb0067491. URL https://doi.org/10.1007%2Fbfb0067491.

[68] F. Mémoli and O. B. Okutan. Reeb posets and tree approximations. *Discrete Mathematics*, 343(2):111658, 2020. ISSN 0012-365X. doi: https://doi.org/10.1016/j.disc.2019.111658. URL https://www.sciencedirect.com/science/article/pii/S0012365X1930336X.

[69] X. Menguy. Examples of strictly weakly regular points. *Geometric and Functional Analysis*, 11(1):124–131, apr 2001. doi: 10.1007/pl00001667. URL https://doi.org/10.1007%2Fpl00001667.

[70] X. C. Menguy. *Examples of manifolds and spaces with positive Ricci curvature*. PhD thesis, New York University, 2000. URL https://www.proquest.com/dissertations-theses/examples-manifolds-spaces-with-positive-ricci/docview/304606534/se-2?accountid=15867.

[71] Y. Mileyko, S. Mukherjee, and J. Harer. Probability measures on the space of persistence diagrams. *Inverse Problems*, 27(12):124007, Nov 2011. ISSN 1361-6420. doi: 10.1088/0266-5611/27/12/124007. URL http://dx.doi.org/10.1088/0266-5611/27/12/124007.

[72] P. Mörters and Y. Peres. *Brownian Motion*. Cambridge University Press, Jan 2001. ISBN 9780511750489. doi: 10.1017/cbo9780511750489. URL http://dx.doi.org/10.1017/CBO9780511750489.

[73] E. Munch and A. Stefanou. The ℓ^∞-Cophenetic Metric for Phylogenetic Trees As an Interleaving Distance. In *Association for Women in Mathematics Series*, pages 109–127. Springer International Publishing, 2019. doi: 10.1007/978-3-030-11566-1_5. URL https://doi.org/10.1007%2F978-3-030-11566-1_5.

[74] J. Neveu and J. Pitman. Renewal property of the extrema and tree property of the excursion of a one-dimensional Brownian motion. *Séminaire de probabilités de Strasbourg*, 23:239–247, 1989. URL http://www.numdam.org/item/SPS_1989__23__239_0.

[75] A. C. Offord. Introduction to the theory of Fourier integrals. *Nature*, 141(3561):183, jan 1938. doi: 10.1038/141183a0. URL https://doi.org/10.1038%2F141183a0.

[76] S. Y. Oudot. *Persistence Theory - From Quiver Representations to Data Analysis*, volume 209 of *Mathematical surveys and monographs*. American Mathematical Society, 2015. ISBN 978-1-4704-2545-6. URL http://bookstore.ams.org/surv-209/.

[77] D. Perez. On C^0-persistent homology and trees. *arXiv e-prints*, art. arXiv:2012.02634, Dec. 2020.

[78] D. Perez. On the persistent homology of almost surely C^0 stochastic processes. *arXiv e-prints*, art. arXiv:2012.09459, Dec. 2020.

[79] D. Perez. ζ-functions and the topology of superlevel sets of stochastic processes. *arXiv e-prints*, art. arXiv:2110.10982, Oct. 2021.

[80] J. Picard. A tree approach to p-variation and to integration. *The Annals of Probability*, 36 (6):2235–2279, Nov 2008. ISSN 0091-1798. doi: 10.1214/07-aop388. URL http://dx.doi.org/10.1214/07-AOP388.

[81] J. Pitman and M. Yor. The Law of the Maximum of a Bessel Bridge. *Electronic Journal of Probability*, 4(none), jan 1999. doi: 10.1214/ejp.v4-52. URL https://doi.org/10.1214%2Fejp.v4-52.

[82] J. Pitman and M. Yor. Infinitely divisible laws associated with hyperbolic functions. *Canadian Journal of Mathematics*, 55(2):292–330, 2003. doi: 10.4153/CJM-2003-014-x.

[83] H. Poincaré. Analysis situs. *Journal de l'École Polytechnique*, 2(1):1–123, 1895.

[84] I. Polterovich, L. Polterovich, and V. Stojisavljević. Persistence barcodes and Laplace eigenfunctions on surfaces. *Geometriae Dedicata*, 201(1):111–138, Aug 2018. ISSN 1572-9168. doi: 10.1007/s10711-018-0383-9. URL http://dx.doi.org/10.1007/s10711-018-0383-9.

[85] L. Polterovich, D. Rosen, K. Samvelyan, and J. Zhang. Topological Persistence in Geometry and Analysis. *arXiv e-prints*, art. arXiv:1904.04044, Apr 2019.

[86] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer Berlin Heidelberg, 1999. doi: 10.1007/978-3-662-06400-9. URL https://doi.org/10.1007%2F978-3-662-06400-9.

[87] S. O. Rice. Mathematical analysis of random noise. *The Bell System Technical Journal*, 23 (3):282–332, 1944. doi: 10.1002/j.1538-7305.1944.tb00874.x. URL https://doi.org/10.1002/j.1538-7305.1944.tb00874.x.

[88] V. Robins. *Computational topology at multiple resolutions: foundations and applications to fractals and dynamics*. PhD thesis, University of Colorado at Boulder, 2000.

[89] J. Roe. *Lectures on Coarse Geometry*, volume 31 of *University Lecture Series*. American Mathematical Society (AMS), 2003. ISBN 978-0-8218-3332-2. doi: http://dx.doi.org/10.1090/ulect/031. URL https://doi.org/10.1090/ulect/031MathSciNet.

[90] A. Romero, J. Heras, J. Rubio, and F. Sergeraert. Defining and computing persistent $\mathbb{Z}$-homology in the general case. *ArXiv*, abs/1403.7086, 2014.

[91] B. Schweinhart. Persistent homology and the upper box dimension. *Discrete & Computational Geometry*, Nov 2019. ISSN 1432-0444. doi: 10.1007/s00454-019-00145-3. URL http://dx.doi.org/10.1007/s00454-019-00145-3.

[92] B. Schweinhart. Fractal dimension and the persistent homology of random geometric complexes. *Advances in Mathematics*, 372:107291, 2020. ISSN 0001-8708. doi: https://doi.org/10.1016/j.aim.2020.107291. URL https://www.sciencedirect.com/science/article/pii/S0001870820303170.

[93] P. Skraba and K. Turner. Wasserstein stability for persistence diagrams. *arXiv e-prints*, 2020. doi: 10.48550/ARXIV.2006.16824. URL https://arxiv.org/abs/2006.16824.

[94] R. P. Stanley. *Enumerative Combinatorics*. Cambridge University Press, 2009. doi: 10.1017/cbo9781139058520. URL https://doi.org/10.1017%2Fcbo9781139058520.

[95] J. M. Steele. Growth rates of euclidean minimal spanning trees with power weighted edges. *The Annals of Probability*, 16(4):1767–1787, 1988. ISSN 00911798. URL http://www.jstor.org/stable/2243991.

[96] K. Turner, Y. Mileyko, S. Mukherjee, and J. Harer. Fréchet means for distributions of persistence diagrams. *Discrete & Computational Geometry*, 52(1):44–70, Jul 2014. ISSN 1432-0444. doi: 10.1007/s00454-014-9604-7. URL http://dx.doi.org/10.1007/s00454-014-9604-7.

[97] S. Wang, Y. Wang, and R. Wenger. The JS-graphs of join and split trees. In *Proceedings of the thirtieth annual symposium on Computational geometry*. ACM, jun 2014. doi: 10.1145/2582112.2582162. URL https://doi.org/10.1145%2F2582112.2582162.

[98] S. Willard. *General topology*. Addison-Wesley Publishing Co., 1970. ISBN 978-0-201-08707-9.

RÉSUMÉ

Cette thèse étudie l'homologie persistante des fonctions continues à valeurs réelles f sur des espaces topologiques compacts X. L'introduction des *indices homologiques* et des *dimensions homologiques* permet de lier la théorie de la persistance à des quantités métriques de l'espace compact X, telles que sa dimension. L'étude de ces quantités permet de d'étendre les résultats de stabilité des distances Wasserstein p sur l'espace des diagrammes de persistance aux fonctions höldériennes sur des espaces métriques plus généraux que ceux précédemment établis par la littérature (qui incluent en particulier toutes les variétés riemanniennes compactes) avec des constantes explicites. En degré d'homologie zéro, une étude plus approfondie peut être réalisée à l'aide d'arbres associés à f, qui généralisent les *merge trees* définissables lorsque f est de Morse. Il est possible de lier la dimension de ces arbres à l'indice de persistance de f et à son code-barres. Nous appliquons ces résultats déterministes au cadre stochastique pour en tirer des conséquences sur les code-barres de fonctions aléatoires de régularité prescrite. Ces conséquences permettent en outre d'élaborer des tests de discrimination de distribution des processus, dont nous présentons un exemple particulier. Enfin, nous définissons les fonctions ζ associés à un processus stochastique et nous calculons ces fonctions d'autres quantités annexes pour plusieurs processus en dimension 1, ainsi que le mouvement Brownien et les processus de Lévy α-stables.

MOTS CLÉS

homologie, persistance, code-barres, topologie, probabilité

ABSTRACT

This thesis studies the persistent homology of $\mathbb{R}$-valued continuous functions f on compact topological spaces X. The introduction of *homological indices* and *homological dimensions* allows us to link persistence theory to metric quantities of the compact space X, such as its upper-box dimension. These quantities give a precise framework to the Wasserstein p-stability results known in the literature, but also extend them to Hölder functions on more general spaces (including all compact Riemannian manifolds) with explicit constants and whose regime for p is optimal. In degree zero of homology, a more in-depth study can be made using trees associated to f, which generalize the *merge trees* definable when f is Morse. It is possible to link the dimension of these trees to the persistence index of f and to its barcode. We apply these deterministic results to the stochastic setting to draw consequences about the barcodes of random functions of prescribed regularity. These consequences also allow us to develop distributional discrimination tests for the processes, of which we present a particular example. Finally, we define the ζ-functions associated with a stochastic process and compute these functions and other related quantities for several processes in dimension one, including the Brownian motion and the α-stable Lévy processes.

KEYWORDS

homology, persistence, barcodes, topology, probability