Title
Spectra of Random Trees, Coalescing Non-Brownian Particles and Geometric Influences of Boolean Functions

Author
Sen, Arnab

Publication Date
2010

Peer reviewed|Thesis/dissertation

Spectra of Random Trees, Coalescing Non-Brownian Particles and Geometric Influences of Boolean Functions

by

Arnab Sen

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Statistics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Steven N. Evans, Co-chair
Professor Elchanan Mossel, Co-chair
Professor David J. Aldous
Professor Luca Trevisan

Spring 2010

Spectra of Random Trees, Coalescing Non-Brownian Particles and Geometric Influences of Boolean Functions

Abstract

Spectra of Random Trees, Coalescing Non-Brownian Particles and Geometric Influences of Boolean Functions

by

Arnab Sen
Doctor of Philosophy in Statistics

University of California, Berkeley

Professor Steven N. Evans, Co-chair
Professor Elchanan Mossel, Co-chair

In the first part of this dissertation, we analyze the eigenvalues of the adjacency matrices of a wide variety of random trees as the size of the trees gets larger. We show that the empirical spectral distributions for many random tree models converge to a deterministic (model dependent) limit as the number of vertices goes to infinity. Our proof uses a suitable notion of local weak convergence for an ensemble of random trees which is known as probability fringe convergence. We conclude for ensembles such as the linear preferential attachment models, random recursive trees, and the uniform random trees that the limiting spectral distribution has a set of atoms that is dense in the real line. We employ a simplified version of Karp-Sipser algorithm to obtain lower bounds on the mass assigned to zero by the empirical spectral measures. For the the linear preferential attachment model with parameter $a > -1$, we show that for any fixed k, the k largest eigenvalues jointly converge in distribution to a non-trivial limit when suitably rescaled.

A well-known result of Arratia shows that one can make rigorous the notion of starting an independent Brownian motion at every point of an arbitrary closed subset of the real line and then building a set-valued process by requiring particles to coalesce when they collide. Arratia noted that the value of this process will be almost surely a locally finite set at all positive times, and a finite set almost surely if the initial value is compact. In the second part of this dissertation, we study the set-valued coalescing processes when the underlying process is not Brownian motion on the real line but is one of the following two examples of self-similar processes: Brownian motions on the Sierpinski gasket and stable processes on the real line with stable index greater than one. We show that Arratia's conclusion is still valid for these two examples.

Finally in the third and last part of this dissertation we present a new definition of influences of boolean functions in product spaces of continuous distributions. Our definition is geometric, and for monotone sets it is equal to the measure of the boundary with respect to uniform enlargement. We prove analogues of the Kahn-Kalai-Linial

(KKL) bound and Russo-type formula for the new definition. As a consequence, we establish a sharp threshold phenomenon for monotone increasing events in the product Gaussian space with respect to the mean parameter and give a statistical application of it. We also obtain isoperimetric inequality for the Gaussian measure on $\mathbb{R}^n$ and the class of sets invariant under transitive permutation group of the coordinates.

For my parents

Contents

List of Figures

Acknowledgments

Foremost, deep appreciation and gratitude are due to my two advisors - Steve Evans and Elchanan Mossel. Excellent teachers, amazing mentors and great collaborators, Elchanan and Steve have always been enthusiastic, bestowed endless amounts of advice, encouraged and supported me throughout my entire PhD life. My research has been greatly enriched by the many exciting projects they have provided me and by their deep insight into the subject. They have shown by example that academia can be incisive, invigorating and fun all at once. Above all, they are among the nicest persons I have been fortunate enough to meet in my life.

This dissertation is based on the joint work with my coauthors - Shankar Bhamidi, Nathan Keller, Ben Morris along with my advisors. Without them I would not have completed even a fraction of what I accomplished. I am also grateful to the other members of my dissertation committee, Professors David Aldous and Luca Trevisan for their prompt feedback and valuable suggestions.

I am specially indebted to my Master thesis advisor Professor Arup Bose who is responsible for my 'real' introduction to the beautiful subject of Probability and who inspired me to think about choosing a career in Probability.

I had the opportunity to take numerous outstanding courses in Berkeley which have immensely enhanced my knowledge in Mathematics and Statistics. Professors Craig Evans and Alistair Sinclair taught two of the best courses I have ever taken in my life. I am also very grateful to Professor Jim Pitman for his continuous encouragement and valuable advice. As a GSI for Professor Roger Purves, I gained invaluable teaching experience for which I am earnestly thankful to him. I am grateful to the entire Statistics department of Berkeley for providing a simulating yet friendly environment for research. I will also never forget the kindness and help I have received from the department staff members.

I am deeply grateful to Ani Adhikari for her support and her words of wisdom. Talking to ever smiling Shanky has always been a pure joy and I thank him for all his help and words of encouragement throughout the initial stages of my PhD. Sourav Chatterjee has always been a guiding force throughout my PhD career. I have learnt so many things from Sourav - ranging from concentration inequality to Chinese history - during many wonderful interactions with him at Berkeley. Also many thanks to him for giving me countably infinitely many treats!

There were several devoted teachers in Indian Statistical Institute who taught me to love mathematics as a young man, and whose encouragement, in spite of the difficulties I present as a student, gave me the motivation and confidence I needed to achieve my goals. These include Professors S.C. Bagchi, Alok Goswami, A.R. Rao, B.V. Rao, Debashis Sengupta and I remain forever grateful to them.

The scores of stimulating discussions that I had with Nick Crawford and Nayantara Bhatnagar are too good to forget. I also thank my colleagues and friends in the Evans Hall including Josh Abramson, Partha Dey, Jian Ding, Subhroshekhar Ghosh, Chris

Haulk, Karl Rohe, Steven Troxler for all the laughs and experiences. I consider myself extremely lucky to have Partha by my side as a true friend, nice flat-mate and helpful comrade throughout the entire journey of my PhD.

I can never overemphasize the contributions of my wonderful friends without whose support and friendship it would have been impossible for me to travel across the treacherous road of PhD. Roshni, Samriddhi, Shirshendu Chatterjee, Sirshendu Maiti, Subarata have stood by me at best and worst of times - I am specially thankful to Roshni and Samridhhi who patiently listened to my long complaints and stressed-out monologues and helped me come out of hard times. Ashis, Debashis, Prantik, Soumik have always been there for me in their usual hilarious self to cheer me up in my loneliness even when we were oceans apart. I am grateful to Koushik, Sayak, Sharmodeep who have really made my years in Berkeley memorable.

Last but not the least, I would like to thank my parents. Though they do not understand even a single line of the math I have written, their love, support and sacrifice is hidden in every page of this dissertation. In the insecure world of research, they have instilled in me a sense of security for I know that they are always there for me to fall back upon. Thank you for believing in me even when I did not believe in myself.

Chapter 1

Introduction

This dissertation essentially consists of three self-contained chapters on diverse topics in Probability theory.

In Chapter 2, we look at the eigenvalues of the adjacency matrices of large random trees. The study of large random matrices and their eigenvalues is one of the primary themes of current research in probability. It finds applications in such diverse fields as number theory, random partitions, free probability and operator algebras, high-dimensional statistical analysis, nuclear physics, signal processing, wireless communication, quantum percolation, and the operation of search engines. Some recent book length expositions are [56, 26, 106, 138, 80, 73].

The objects of interest in this field are usually random real symmetric or complex Hermitian matrices. For example, one of the most popular models is the *Gaussian unitary ensemble (GUE)*, where the matrices are Hermitian, the entries above the diagonal are independent, identically distributed, standard complex Gaussian random variables, and the entries on the diagonal are independent, identically distributed, standard real Gaussian random variables. Much is now known about the asymptotic behavior of objects such as the empirical distribution of the eigenvalues and the behavior of the maximal eigenvalue.

Here we investigate random matrices with substantially greater structure and complexity than the GUE, namely the adjacency matrices of random graphs, although our methods are also applicable to the closely related Laplacian matrices. The recent availability of large amounts of data has led to an explosion in the number of models used to model real-world networks, and dynamically grown models such as various preferential attachment schemes have attracted significant interest from the computer science and mathematical physics community. It is known (see, for example, the monographs [50, 78, 24, 52, 54, 53]) that a surprising diversity of features of a graph are determined, at least in part, by the behavior of the eigenvalues of its adjacency and Laplacian matrices.

We concentrate on the adjacency matrices of various ensembles of random trees. Random trees arise in numerous contexts, ranging from the analysis of database and

search algorithms in computer science to models of phylogenies (that is, evolutionary family trees) in biology. Moreover, many of the preferential attachment schemes for networks are also largely random models of growing trees (see, for example, [23] for a survey of some of the more popular schemes). We note that, although trees are rather simple graphs, the analysis of their eigenvalues is still rather challenging, and even in the case of highly symmetric deterministic trees explicit formulae for spectra have only been found recently [118, 119, 117, 120, 115, 116, 121].

We introduce a general technique based on the concept of *probability fringe convergence* for showing that the *spectral distributions* (that is, empirical distributions of the eigenvalues) of the adjacency matrices of an ensemble of random trees converge in the topology of weak convergence of probability measures on the line to a deterministic limit as the number of vertices goes to infinity, and we show how this technique applies in several natural examples.

The notion of probability fringe convergence is a type of *local weak convergence* for random graphs that involves the convergence in distribution of suitably defined neighborhoods of a vertex picked uniformly from a random graph as the size of the graph goes to infinity. Surveys of this general methodology are [2, 5]. Such convergence results for random trees where the limit is described in terms of a continuous-time branching processes go back to [89, 112]. The first (to our knowledge) use of such techniques in various general models of preferential attachment is [122]. Such notions are further explored in [23].

The key algebraic element of our proof of convergence of the spectral distributions is the set of interlacing inequalities between the eigenvalues of a Hermitian matrix and the eigenvalues of one of its principal sub-matrices. The interlacing inequalities allow us to break a large tree up into a forest of smaller trees by deleting a small proportion of edges and conclude that the spectral distribution of the tree is close to that of the forest which, in turn, is a convex combination of the spectral distributions of its component sub-trees. If the decomposition into a forest is done appropriately, then the resulting sub-trees are "locally defined" in a sense that allows probability fringe convergence to be brought to bear to show that the spectral distribution of the forest converges.

We note that interlacing has found other applications in algebraic graph theory [33, 69, 81].

Another interesting technical aspect of our work is that the *method of moments*, one of the most commonly used tools in random matrix theory, fails for some natural ensembles because, as we observe in Subsection 2.3.1, expected values of moments of the spectral distribution go to infinity.

While our method for showing that the spectral distribution converges is quite general, it does not provide any sort of characterization of the limiting distribution. In Section 2.7 we look at an extremely simple random tree that is obtained by taking the tree consisting of a path of n points and independently connecting an edge to each point with equal probability, so that the resulting tree resembles a comb with some

of its teeth missing. Our probability fringe convergence methodology does not apply immediately to this ensemble of random trees, but a straightforward modification of it does. We compute the asymptotic moments of the spectral distribution for this ensemble and show that even in this simple case the answer is rather intricate, indicating that we should perhaps not expect simple characterizations of the limit for more complex models.

Quite closely related to our results for the spectral distribution is the recent work [37], where similar local weak convergence techniques are combined with Stieltjes transform methods to prove various limiting results for families of random graphs.

We extend our results on the convergence of the spectral distribution in two different directions.

First, we show for any $\gamma \in \mathbb{R}$ that the proportion of eigenvalues that have the value γ converges to a constant under the assumption of probability fringe convergence. Moreover, we give a simple sufficient condition for the limit to be positive and apply this condition to show for several models that the limiting spectral distribution has a set of atoms that is dense in $\mathbb{R}$. We pay particular attention to the proportion of zero eigenvalues, a quantity of importance in areas such as quantum percolation [17, 18]. It is possible to obtain much more exact information on the limiting proportion because of the connection between the number of zero eigenvalues of the adjacency matrix of a tree and the cardinality of a maximal matching. In particular, we use a simplified version of the Karp-Sipser algorithm [94] to construct maximal matchings. Incidentally, the Karp-Sipser algorithm has been also used in a recent work [38] to study the limiting proportion of zero eigenvalues of random sparse graphs. We also use our methods to obtain the asymptotic behavior of the total weight of a maximal weighted matching when the edge weights are given by independent, identically distributed, non-negative random variables.

Second, we obtain results on the joint convergence in distribution of the suitably normalized k largest eigenvalues for the preferential attachment tree. These results extend and sharpen those in [49, 48, 71, 70], where it was shown that the k largest eigenvalues are asymptotically equivalent to the square roots of the k largest out-degrees. The weak convergence of the suitably rescaled maximum out-degree was obtained in [109] using martingale methods. However it is not clear how to extend this technique to describe the asymptotics for the k largest out-degrees for $k \geq 2$. We prove our more general results using an approach that is essentially completely different.

Chapter 3 concerns with coalescing particle systems. A construction due to Richard Arratia [10, 11] shows that it is possible to make rigorous sense of the informal notion of starting an independent Brownian motion at each point of the real line and letting particles coalesce when they collide.

Arratia proved that the set of particles remaining at any positive time is locally finite almost surely. Arratia's argument is based on the simple observation that at the time two particles collide, one or the other must have already collided with each

particle that was initially between them. The same argument shows that if we start an independent circular Brownian motion at each point of the circle and let particles coalesce when they collide, then, almost surely, there are only finitely many particles remaining at any positive time.

Arratia established something even stronger: it is possible to construct a flow of random maps $(\Phi_{s,t})_{s<t}$ from the real line to itself in such a way that for each fixed s, the process $(\Phi_{s,s+u})_{u\geq 0}$ is given by the above particle system. Arratia's flow has since been studied by several authors such as [136, 127, 128, 101, 137, 72, 86, 134] for purposes as diverse as giving a rigorous definition of a one-dimensional self-repelling Brownian motion or a more general system of coalescing Brownian particles called "Brownian Web" where particles start from each point the real line as well as from each time point to providing examples of noises that are, in some sense, completely "orthogonal" to those produced by Poisson processes or Brownian motions.

Coalescing systems of more general Markov processes have been investigated because of their appearance as the duals of models in genetics of the stepping stone type, see, for example, [99, 66, 65, 59, 142, 140, 84, 111, 143].

Arratia's "topological" argument for instantaneous coalescence to a locally finite set fails when one considers Markov processes on the line or circle with discontinuous sample paths or Markov processes with state spaces that are not locally like the real line. We show, however, that analogous conclusions holds for coalescing Brownian motions on the "finite" and "infinite" (that is, compact and non-compact) Sierpinski gaskets and stable processes on the circle and line – provided, of course, that the stable index is greater than one, so that an independent pair of such motions collides with positive probability. Similar methods will apply to Markov processes on more general state spaces when the process and the state space have suitable local self-similarity properties, but we have not pursued a result with more encompassing hypotheses. As well as providing an interesting test case of a process with continuous sample paths on a state space that is not locally one-dimensional but is such that two independent copies of the process will collide with positive probability, the Brownian motion on the Sierpinski gasket was introduced as a model for diffusion in disordered media and it has since attracted a considerable amount of attention. The reader can get a feeling for this literature by consulting some of the earlier works such as [15, 102, 14] and more recent papers such as [82, 100] and the references therein.

In Chapter 4, we turn to an entirely new topic - influences of boolean functions. The notion of influences of variables on boolean functions defined on the n-dimensional discrete cube $\{0,1\}^n$ equipped with product Bernoulli measure is one of the central concepts in the theory of discrete harmonic analysis. The influence of the ith variable on a boolean function f of n boolean variables is the probability that after assigning the $n-1$ variables values at random the value of f changes if we flip the ith variable. This definition was introduced by Ben-Or and Linial [19] in the context of "collective coin-flipping". In the last two decades, the study of influences found several applications in diverse fields, including Combinatorics, Theoretical Computer Science,

Statistical Physics, Social Choice Theory, etc. (see, for example, the survey article [93]). The influences have numerous properties that allow to use them in applications. The following three properties are amongst the most fundamental ones:

1. **Geometric Meaning.** The influences of a boolean function f on the discrete cube $\{0,1\}^n$ with uniform measure have a clear geometric meaning. The sum of influences of f is the size of *edge boundary* of the set $A := \{x \in \{0,1\}^n : f(x) = 1\}$ divided by a normalizing factor 2^{n-1}. The edge boundary of A is the number of edges between A and A^c. The influence of the ith variable on f is the number of the edges emitting from A parallel to the i-th direction, divided by 2^{n-1}.

2. **The KKL Theorem.** In the remarkable paper [91], Kahn, Kalai, and Linial proved that for any boolean function $f : \{0,1\}^n \to \{0,1\}$, there exists a variable i whose influence is at least $ct(1-t)\log n/n$, where $t = \mathbb{E}f$ and c is a universal constant. Many applications of influences make use of the KKL theorem or of related results such as [135, 76] in one way or another.

3. **The Russo Lemma.** Let μ_p denote the Bernoulli measure where 0 is given weight $1-p$ and 1 is given weight p. Clearly if $A \subseteq \{0,1\}^n$ is increasing then $\mu_p^{\otimes n}(A)$ is monotone increasing as function of p. The question of understanding how $\mu_p^{\otimes n}(A)$ varies with p has important applications in the theory of random graphs and in percolation theory. Russo's Lemma [103, 124] asserts that the derivative of $\mu_p^{\otimes n}(A)$ with respect to p is the sum of influences of $f = 1_A$.

The basic results on influences were obtained for functions on the discrete cube, but some applications required generalization of the results to more general product spaces. Unlike the discrete case, where there exists a single natural definition of influence, for general product spaces several definitions were presented in different papers, see for example [43, 83, 95, 110]. While each of these definitions has its advantages, in general all of them lack geometric interpretation for continuous probability spaces.

Here we present a new definition of the influences in product spaces of absolutely continuous probability measures on the real line, that has a clear geometric meaning. We call them *geometric influences*. We show that for important classes of product measures, including the Gaussian measure, our definition allows us to obtain analogues of the KKL theorem and Russo-type formulas. These, in turn, can be combined to establish sharp threshold of monotone increasing sets with respect to the location parameter underlying measure. We provide an application of the above in testing statistical hypothesis regarding the mean of a Gaussian distribution. Moreover, for Gaussian measure on $\mathbb{R}^n$ we obtain a dimension-dependent isoperimetric inequality when we restrict our attention to a class of sets with many symmetries - the sets invariant under transitive permutation group of the coordinates.

Chapter 2

Spectra of Random Trees

2.1 Some representative random tree models

An enormous number of random tree models have been developed by computer scientists working on the analysis of algorithms and the mathematical modeling of real world networks: see [2, 23] for a description of some of the more popular models. Although our methods apply quite generally, it will be useful to have the following models in mind when it comes to checking how the hypotheses of our results may be verified in particular instances.

Random recursive tree: This is the simplest model of constructing a rooted tree sequentially via the addition of a new vertex at each stage. Start with a single vertex (the root) at time 1. Label the vertex added at stage n by n, so the tree $\mathcal{T}_n$ that has been constructed by stage n has vertex set $[n] := \{1, 2, \ldots, n\}$. Construct the tree at stage $n+1$ by adding an edge from vertex $n+1$ to a vertex chosen uniformly among the vertices $1, 2, \ldots, n$. We refer the reader to [126] for a survey of some of the properties of the random recursive tree.

Linear preferential attachment tree: This is another sequential construction. As before, start with a single vertex (the root) at time 1. Suppose the tree on n vertices labeled by $[n]$ has been constructed. Think of the edges as directed away from the root and let $D(v, n)$ be the out-degree of vertex $v \in [n]$ at time n (that is, $D(v, n)$ is the number of children of vertex v at time n). Construct a tree on $n+1$ vertices via the addition of an edge between the new vertex $n+1$ and the vertex v in $[n]$ with probability proportional to $D(v, n) + 1 + a$, where $a > -1$ is a parameter of the process. There is an enormous amount of recent literature on this model. We refer the reader to [34, 62, 23] for relevant references.

Uniform random rooted unordered labeled tree: By Cayley's theorem, there are n^{n-1} rooted trees on n labeled vertices (we think of trees as abstract graphs and

so we don't consider a particular embedding of a tree in the plane when it comes to deciding whether two trees are "the same" or "different" – this is the import of the adjective "unordered"). Choose one of these trees uniformly at random. Since we are interested only in the structure of the tree, the labeling will be irrelevant.

Random binary tree: There are various models of random rooted binary trees. The one we shall consider is the following sequential construction. Start at time 1 with the three vertex tree consisting of a root and two leaves. At each stage, choose a leaf uniformly and attach two new leaves to it by two new edges.

2.2 Probability fringe convergence of random trees

The key to understanding the asymptotic properties of the spectra of random trees such as those introduced in Section 2.1 is that they converge "locally" to appropriate locally finite infinite trees. We define the relevant notion of local convergence in this section, and then show how it applies to the models of Section 2.1.

We first need to be precise about what we mean by the term *finite rooted tree*. So far, we have talked about trees as particular types of graphs. That is, we have thought of a tree as being described by a finite set of vertices and a finite set of edges that are unordered pairs of vertices. A rooted tree has then been defined as such an object with a particular distinguished vertex that we call the root. This point of view is useful for describing constructions of random trees. However, we will often wish to consider two trees as being the same if they are *isomorphic* in the usual graph-theoretic sense: that is, if they have the same *shape* and only differ by a labeling of the vertices. A tree in this latter sense is thus an isomorphism class of trees thought of as graphs. When we wish to distinguish these two notions we will use standard terminology and speak of *labeled* and *unlabeled* trees, respectively. Continuing in this vein, we take two rooted trees (thought of as graphs) to be the same if there is a graph-theoretic isomorphism from one to the other that preserves the root, and we call the corresponding equivalence classes *unlabeled rooted trees*. Even more generally, we may consider unlabeled trees with several distinguished vertices.

Let $\mathbb{T}$ be the countable space of all finite unlabeled rooted trees. Set $\mathbb{T}_* = \mathbb{T} \sqcup \{*\}$, where $*$ is an adjoined point. Equip $\mathbb{T}$ and $\mathbb{T}_*$ with the respective discrete topologies, and equip the Cartesian products $\mathbb{T}^\infty$ and $\mathbb{T}_*^\infty$ with the usual product topologies.

Consider a finite unlabeled rooted tree $\mathbf{t} \in \mathbb{T}$ with root ρ and another distinguished vertex v that is at distance h from the root (v may coincide with ρ, in which case $h = 0$). Let $(v = v_0, v_1, \ldots, v_h = \rho)$ denote the unique path from the vertex v to the root. Write t_0 for the subtree rooted at $v_0 = v$ that consists of all vertices for which the path to the root passes through v_0, and for $1 \leq k \leq h$, write t_k for the subtree rooted at v_k that consists of all vertices for which the path from the root passes through v_k but not through v_{k-1}. Write $\Phi(\mathbf{t}, \cdot)$ for the probability distribution

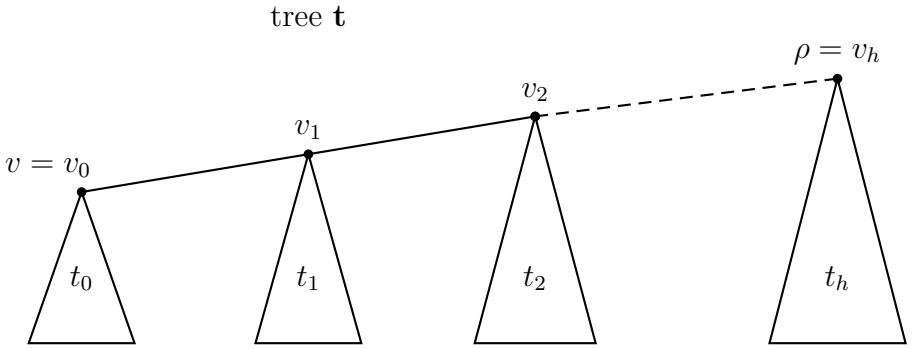

Figure 2.1: Fringe decomposition of a finite rooted tree

on $\mathbb{T}_*^\infty$ that places mass $(\#\mathbf{t})^{-1}$ at each of the sequences $(t_0, t_1, \ldots, t_h, *, *, \ldots) \in \mathbb{T}_*^\infty$ as v ranges over the $\#\mathbf{t}$ vertices of $\mathbf{t}$. It is clear that Φ is a probability kernel from $\mathbb{T}$ to $\mathbb{T}_*^\infty$.

Definition 2.2.1. *Let $(\mathcal{T}_n)_{n=1}^\infty$ be a sequence of random finite unlabeled rooted trees, and suppose that $\mathcal{T}$ is a $\mathbb{T}^\infty$-valued random variable. The sequence $(\mathcal{T}_n)_{n=1}^\infty$ converges in the* probability fringe sense *to $\mathcal{T}$ if the sequence $\Phi(\mathcal{T}_n, \cdot)$ of random probability measures on $\mathbb{T}_*^\infty$ converges weakly to the distribution of $\mathcal{T}$ in the topology of weak convergence of probability measures on $\mathbb{T}_*^\infty$.*

Remark 1. *The definition requires that the empirical distribution of the sub-trees below the various vertices of $\mathcal{T}_n$ converges. However, it demands much more than this: for each $k \geq 1$, the joint empirical distribution of the sub-tree below a vertex and the sub-trees below each of its k most recent ancestors must also converge.*

Remark 2. *Note that any sequence $(t_0, t_1, \ldots) \in \mathbb{T}^\infty$ may be thought of as a locally finite unlabeled rooted tree with one end (that is, with a single semi-infinite path) via the identification of the roots of t_k, $k \in \mathbb{Z}^+$, as the successive vertices on the unique semi-infinite path from the root. We call such trees* `sin`-trees *(for single infinite path trees).*

Remark 3. *The terminology "probability fringe convergence" is not standard. In the literature, the convergence of the local structure around a* **uniformly chosen** *vertex of $\mathcal{T}_n$ to the structure around the root for some limiting random* `sin`*-tree is an instance of what has been termed "local weak convergence" by Aldous, see [5]. Our definition is somewhat stronger.*

A powerful technique for establishing probability fringe convergence of an ensemble of random trees is to first show that each member of the ensemble can be

constructed as the family tree of a suitable stopped continuous-time branching process. (For us, a continuous-time branching process is the sort of object considered in [88]: individuals give birth to a possibly random number of offspring at the arrival times of a point process up to a possibly infinite death time, and those offspring go on to behave as independent copies of their parent.) The next result describes such embeddings for the ensembles of Section 2.1.

Proposition 2.2.1. *(a)* [Random recursive tree] *Consider a continuous time branching process that starts with a single progenitor, individuals live forever, and individuals produce a single offspring at each arrival time of a unit rate Poisson process (this process is sometimes called the* Yule process*, but the usage of that terminology is not completely consistent in the literature). Write $\mathcal{F}(t) \in \mathbb{T}$ for the corresponding family tree at time $t \geq 0$. Set $T_n := \inf\{t > 0 : \#\mathcal{F}(t) = n\}$. Then $\mathcal{F}(T_n)$ has the same distribution as $\mathcal{T}_n$, where $\mathcal{T}_n$ is the random recursive tree on n vertices.*

(b) [Linear preferential attachment tree] *Consider a continuous time branching process that starts with a single progenitor, individuals live forever, and the point process representing the offspring distribution of any individual is a pure birth point process started at 0 that can be described as follows: Whenever any individual has already given birth to k direct offspring, the individual produces a new offspring at rate $k + 1 + a$. In particular, at the time an individual is born, the individual generates new offspring at rate $1 + a$. Thus, the times that elapse between the birth of an individual and the successive births of the individual's offspring, say $(\beta_1, \beta_2, \ldots)$, may be written as $\beta_i = \sum_{j=0}^{i-1} \eta_j$, where the successive η_j are independent exponential random variables and η_j has rate $j+1+a$. Each individual in the population has its own independent and identically distributed copy of the above offspring point process. Write $\mathcal{F}(t) \in \mathbb{T}$ for the corresponding family tree at time $t \geq 0$. Set $T_n := \inf\{t > 0 : \#\mathcal{F}(t) = n\}$. Then, $\mathcal{F}(T_n)$ has the same distribution as $\mathcal{T}_n$, where $\mathcal{T}_n$ is the linear preferential attachment tree on n vertices with parameter $a > -1$.*

(c) [Uniform random rooted unordered labeled tree] *Let Z_∞ be the complete family tree for a (discrete-time) Galton-Watson branching process with mean 1 Poisson offspring distribution. Note that Z_∞ is finite almost surely. The distribution of Z_∞ conditioned on $\#Z_\infty = n$ is the same as that of $\mathcal{T}_n$, where $\mathcal{T}_n$ is the objected obtained by taking the uniform random rooted unordered tree on n labeled vertices and removing the labeling.*

(d) [Random binary tree] *Consider a continuous-time branching process that starts with a single progenitor, individuals live until a rate 1 exponential time, at which time they produce two offspring (we will refer to this process as the* random binary splitting process*). Write $\mathcal{F}(t) \in \mathbb{T}$ for the corresponding family tree at time $t \geq 0$. Set $T_n := \inf\{t > 0 : \#\mathcal{F}(t) = n\}$. Then, $\mathcal{F}(T_n)$ has the same distribution as $\mathcal{T}_n$, where $\mathcal{T}_n$ is the random binary tree on n vertices.*

Proof. Parts (a), (b) and (d) follow from the comparison of the rates of the production of the offspring and the corresponding growth dynamics of the associated tree $\mathcal{T}_n$. Part

(c) is well-known and follows from randomly ordering the offspring of each individual to obtain an ordered (that is, planar) tree, computing the conditional probability distribution of the resulting rooted ordered tree, randomly labeling the vertices of the rooted ordered tree, and verifying that the randomly labeled tree is uniformly distributed using Cayley's theorem for the number of rooted labeled trees on n vertices (see, for example, [3]). □

We now describe briefly the limiting `sin`-trees for the models considered above. Recall that a `sin`-tree can be thought of as an element of $\mathbb{T}^\infty$. The following proposition follows from well-known results, and we give the appropriate references for each specific construction.

Proposition 2.2.2. *Each of the four ensembles of Section 2.1. converges in the probability fringe sense, (as defined in Definition 2.2.1). The limiting random* `sin`*-tree for each model is described explicitly in Construction 2.2.3.*

Construction 2.2.3. (a) [*Random recursive tree: [89, 112, 2]*] Let $\mathcal{F}_i(\cdot)$ be independent rate one Yule processes. Let $X_0, X_1, \ldots$ be independent rate 1 exponential random variables and put $S_i = \sum_{j=0}^{i} X_j$. Then, the limiting `sin`-tree has the distribution of $(\mathcal{F}_i(S_i))_{i=0}^{\infty}$.

(b) [*Linear referential attachment: [112, 89, 23]*] Let $(X_i)_{i=0}^{\infty}$ be independent exponential random variables, where X_0 has rate $2+a$ and each X_i, $i>0$, has rate $1+a$. Let $(\mathcal{F}_i)_{i=0}^{\infty}$ be continuous time branching process that are conditionally independent given $(X_i)_{i=0}^{\infty}$, with the conditional distribution of $\mathcal{F}_i$ being that in part (b) of Proposition 2.2.1 subject to the minor modifications that the point process describing the times at which the root individual gives birth is conditioned to have a birth at time X_i and the offspring born at this time and all its descendants are removed from the population. All other vertices give birth to according to the original offspring point process. Then, the limiting `sin`-tree has the distribution of $(\mathcal{F}_i(\sum_{j=0}^{i} X_j))_{i=0}^{\infty}$.

(c) [*Uniform random tree: [79]*] The limiting `sin`-tree has the distribution of an infinite sequence of independent copies of the critical Poisson Galton-Watson tree Z_∞ of part (c) of Proposition 2.2.1.

(d) [*Random binary tree: [2]*] Let $(\mathcal{F}_i)_{i=0}^{\infty}$ be independent random binary splitting processes as in part (d) of Proposition 2.2.1. Let $(X_i)_{i=0}^{\infty}$ be independent rate 1 exponential random variables and set $S_i = \sum_{j=0}^{i} X_j$. Define $\mathbb{T}$-valued random variables $(\mathcal{U}_i)_{i=0}^{\infty}$ as follows. Put $\mathcal{U}_0 = \mathcal{F}_0(S_0)$. For $i \geq 1$, $\mathcal{U}_i$ is constructed by attaching a new vertex ρ_i to the root of $\mathcal{F}_i(S_{i-1})$ and re-rooting the resulting tree at ρ_i. Then, the limiting `sin`-tree has the distribution of $(\mathcal{U}_i)_{i=0}^{\infty}$.

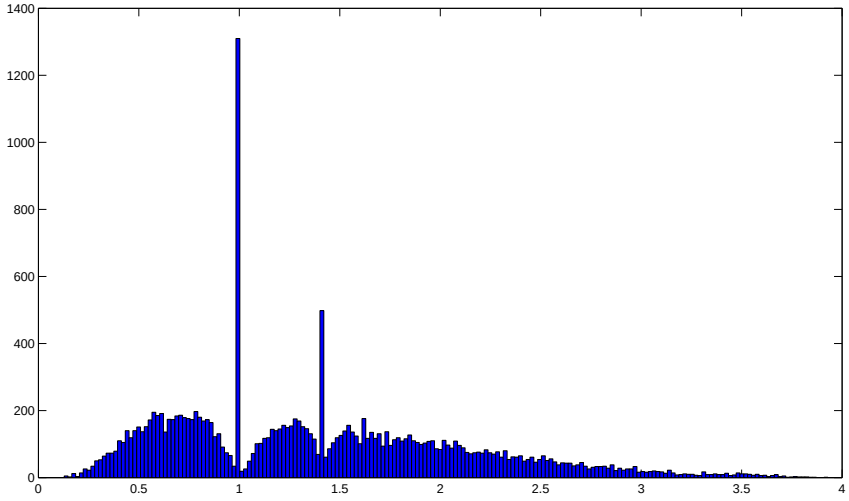

Figure 2.2: Empirical distribution for the positive eigenvalues of the random recursive tree with 200 vertices, averaged over 200 realizations.

2.3 Statement of results

2.3.1 Convergence of the spectral distribution

Theorem 2.3.1. *Suppose that $(\mathcal{T}_n)_{n=1}^{\infty}$ is a sequence of random finite unlabeled rooted trees that converges in the probability fringe sense to a* `sin`*-tree $\mathcal{T}$. Let F_n denote the spectral distribution of the adjacency matrix of $\mathcal{T}_n$. Then there exists a (model dependent) deterministic probability distribution F such that F_n converges in distribution to F in the topology of weak convergence of probability measures on $\mathbb{R}$.*

Simulations of the expected value the spectral distribution for various finite random trees are shown in Figure 2.2, Figure 2.3, Figure 2.4 and Figure 2.5.

Remark 4. *Recall that the* graph Laplacian *of a tree* $\mathbf{t}$ *with adjacency matrix A is the matrix $A - D$, where D is the diagonal matrix recording the degrees of the vertices of* $\mathbf{t}$ *(we caution the reader that some authors refer to the negative of this matrix as the Laplacian). The methods we use to establish Theorem 2.3.1 can also be used to show that if the sequence $(\mathcal{T}_n)_{n=1}^{\infty}$ converges in the probability fringe sense, then the spectral distribution of the Laplacian matrix of $\mathcal{T}_n$ converges in distribution to a deterministic probability distribution on $\mathbb{R}$.*

Remark 5. *One of the interesting aspects of our proof of Theorem 2.3.1 is that it does not employ the usual work-horse of random matrix theory, namely the method of*

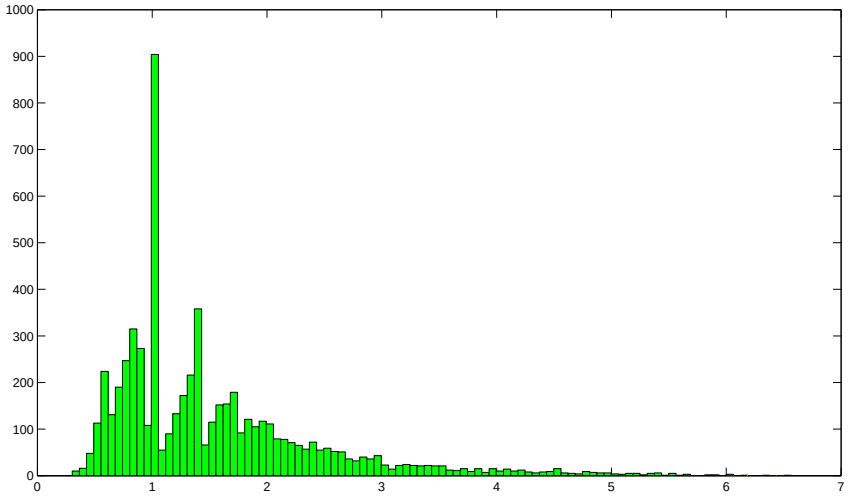

Figure 2.3: Empirical distribution for the positive eigenvalues of the preferential attachment tree with 100 vertices, averaged over 200 realizations.

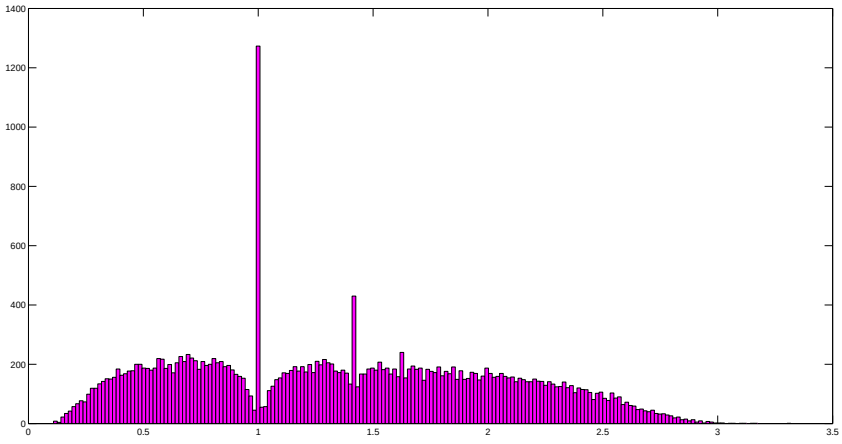

Figure 2.4: Empirical distribution for the positive eigenvalues of the uniform random tree with 200 vertices, averaged over 300 realizations.

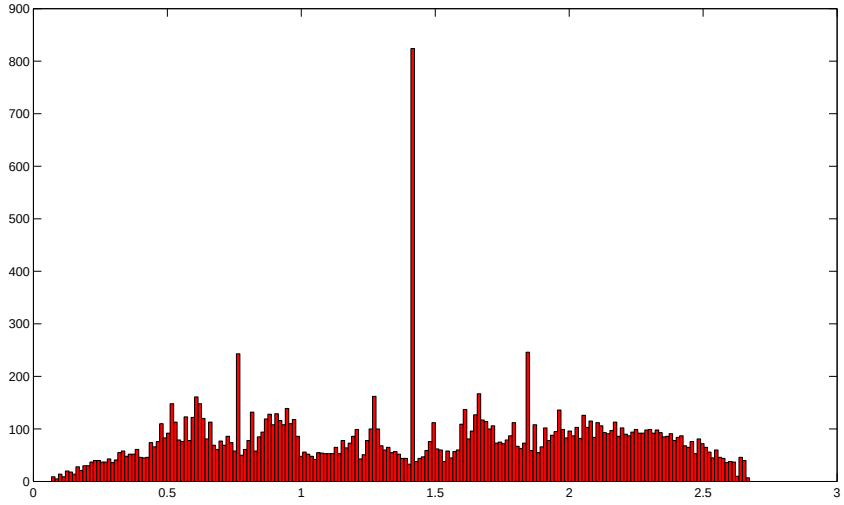

Figure 2.5: Empirical distribution for the positive eigenvalues of the random binary tree with 401 vertices, averaged over 100 realizations.

moments. A standard method for establishing convergence of the spectral distribution F_n of an $n \times n$ random matrix A_n to a deterministic probability distribution F as $n \to \infty$ is to show for all $k \geq 1$ that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\mathrm{tr}(A^k)\right] \lim_{n\to\infty} = \mathbb{E}\left[\int_{\mathbb{R}} x^k dF_n(x)\right] = a_k,$$

where the constants $(a_k)_{k=1}^{\infty}$ are the moments of a unique distribution F, and then show that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n^2}\mathrm{tr}(A^k)^2\right] = a_k^2.$$

The following result shows that this strategy cannot be used for linear preferential attachment model with parameter $a = 0$ and indicates why it is necessary for us to develop other techniques.

Lemma 2.3.2. *Let A_n be the adjacency matrix of the linear preferential attachment tree $\mathcal{T}_n$ with $a = 0$. Then*

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\mathrm{tr}(A_n^4)\right] = \infty.$$

Proof. Recall that $D(k, n)$ is the out-degree of the vertex k in $\mathcal{T}_n$ and that the vertex $n + 1$ is connected to the vertex k with probability proportional to $D(k, n) + 1$.

Let $\mathbb{I}$ be the (random) vertex amongst the vertices $\{1, 2, \ldots, n\}$ to which the vertex

$(n+1)$ is attached. Because $n+1$ is a leaf in $\mathcal{T}_{n+1}$, every path in $\mathcal{T}_{n+1}$ from $n+1$ to itself of length 4 is of the form $n+1 \to \mathbb{I} \to i \to \mathbb{I} \to n+1$ for some $i \in \{1, 2, \ldots, n+1\} \setminus \{\mathbb{I}\}$, so that every path of length 4 in $\mathcal{T}_{n+1}$ from $n+1$ to itself corresponds to a unique path of length 2 in $\mathcal{T}_{n+1}$ from $\mathbb{I}$ to itself. Thus, $A^4_{n+1}(n+1, n+1) = A^2_{n+1}(\mathbb{I}, \mathbb{I})$. Also, $A^2_{n+1}(\mathbb{I}, \mathbb{I}) = A^2_n(\mathbb{I}, \mathbb{I}) + 1$ because the paths of length 2 in $\mathcal{T}_{n+1}$ from $\mathbb{I}$ to itself consist of the paths of length 2 in $\mathcal{T}_n$ from $\mathbb{I}$ to itself plus the additional path $\mathbb{I} \to n+1 \to \mathbb{I}$. Hence, $A^4_{n+1}(n+1, n+1) \geq A^2_n(\mathbb{I}, \mathbb{I})$.

Therefore,

$$\operatorname{tr}(A^4_{n+1}) - \operatorname{tr}(A^4_n) \geq A^4_{n+1}(n+1, n+1) \geq A^2_n(\mathbb{I}, \mathbb{I}) = D(\mathbb{I}, n) + 1_{\{\mathbb{I} \neq 1\}},$$

and so, setting $C_n = \sum_{k=1}^n D(k, n)^2$ and using the fact that $\sum_{k=1}^n D(k, n) = n - 1$,

$$\begin{aligned}\mathbb{E}[\operatorname{tr}(A^4_{n+1})] - \mathbb{E}[\operatorname{tr}(A^4_n)] \geq \mathbb{E}[D(\mathbb{I}, n)] &= \sum_{k=1}^n D(k, n) \frac{D(k, n) + 1}{2n - 1} \\ &= \frac{1}{2n-1} \mathbb{E}[C_n] + \frac{n-1}{2n-1} \geq \frac{1}{2n} \mathbb{E}[C_n].\end{aligned}$$

Now,

$$\frac{1}{n+1} \mathbb{E}[\operatorname{tr}(A^4_{n+1})] = \frac{1}{n+1} \sum_{m=1}^n \left(\mathbb{E}[\operatorname{tr}(A^4_{m+1})] - \mathbb{E}[\operatorname{tr}(A^4_m)] \right) \geq \frac{1}{n+1} \sum_{m=1}^n \frac{1}{2m} \mathbb{E}[C_m],$$

and so it will suffice to show that

$$\lim_{n \to \infty} \frac{1}{n} \mathbb{E}[C_n] = \infty. \tag{2.1}$$

Note that

$$\begin{aligned}\mathbb{E}[C_{n+1}] - \mathbb{E}[C_n] &= \mathbb{E}[D(n+1, n+1)^2] + \mathbb{E}[(1 + D(\mathbb{I}, n))^2 - D(\mathbb{I}, n)^2] \\ &= 1 + 2\,\mathbb{E}[D(\mathbb{I}, n)] \geq 1 + \frac{1}{n} \mathbb{E}[C_n],\end{aligned}$$

and hence

$$\mathbb{E}[C_{n+1}] \geq n + \sum_{m=1}^n \frac{1}{m} \mathbb{E}[C_m].$$

It follows that

$$\liminf_{n \to \infty} \frac{\mathbb{E}[C_n]}{n} \geq 1 + \liminf_{n \to \infty} \frac{\mathbb{E}[C_n]}{n},$$

and so (2.1) holds, as required. □

2.3.2 Atoms in the spectral distribution

Theorem 2.3.3. *Let $(\mathcal{T}_n)_{n=1}^{\infty}$ a sequence of random finite unlabeled rooted trees that converges in the probability fringe sense to a* `sin`*-tree. For any $\gamma \in \mathbb{R}$, $F_n(\{\gamma\})$ converges in distribution to a (model dependent) constant c_γ as $n \to \infty$. If F is the limiting spectral distribution guaranteed by Theorem 2.3.1, then $F(\{\gamma\}) \geq c_\gamma$.*

Theorem 2.3.4. *Suppose that $(\mathcal{T}_n)_{n=1}^{\infty}$ is a sequence of random finite unlabeled rooted trees that converges in the probability fringe sense to a* `sin`*-tree $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \ldots)$. Consider a forest $\mathbf{u}$ composed of finitely many finite unlabeled rooted trees, and assume that some eigenvalue γ of the adjacency matrix of $\mathbf{u}$ has multiplicity $L > 1$. Write $\mathcal{U}$ for the random forest obtained by deleting the root of $\mathcal{T}^0$ from $\mathcal{T}^0$, and suppose that $\mathbf{P}\{\mathcal{U} = \mathbf{u}\} > 0$. Then, the constant c_γ of Theorem 2.3.3 is strictly positive and hence γ is an atom of the limiting spectral distribution F of Theorem 2.3.1.*

Remark 6. *The condition that the adjacency matrix of the forest $\mathbf{u}$ has an eigenvalue γ with multiplicity greater than 1 certainly holds if γ is an eigenvalue of more than one of the trees that make up $\mathbf{u}$. In particular, it holds if two or more of the trees that make up $\mathbf{u}$ are equal to some common tree $\mathbf{t}$, and γ is an eigenvalue of $\mathbf{t}$. It is clear for the random recursive tree, the linear preferential attachment tree, and the uniform random tree, that if $\mathbf{u}$ is any forest of finite unlabeled rooted trees, then $\mathbf{P}\{\mathcal{U} = \mathbf{u}\} > 0$, and so any number γ that is the eigenvalue of the adjacency matrix of a finite tree will be an atom of the limiting spectral distribution F for these models. From Theorem 7 of [119], the eigenvalues of the adjacency matrix of the rooted tree in which every non-leaf vertex has out-degree d and each leaf is distance $k-1$ from the root are*

$$2\sqrt{d}\cos\left(\frac{\pi\ell}{j+1}\right), \quad j = 1, \ldots, k,\ \ell = 1, \ldots, j,$$

with given multiplicities. It follows that the atoms of the limiting spectral distribution F are dense in $\mathbb{R}$ for the random recursive tree, the linear preferential attachment tree, and the uniform random tree. A similar argument shows that the limiting spectral distribution for the random binary tree has a set of atoms that is dense in the interval $[-2\sqrt{2}, 2\sqrt{2}]$. Because we can embed any binary tree into a complete binary tree of suitable height, we see that the limiting spectral measure has this interval as its support.

Remark 7. *In light of the previous remark, it is natural to inquire whether the limiting spectral distribution F is purely discrete or whether it also has a continuous component. Our methods do not suffice to resolve this question.*

2.3.3 The proportion of zero eigenvalues and maximal matchings

Theorem 2.3.3 and Theorem 2.3.4 show that the limiting spectral distribution F will typically have many atoms. However, Theorem 2.3.4 provides a rather crude lower bounds on the mass of each atom. We obtain better lower bounds on the limiting proportion of zero eigenvalues in Subsection 2.5.3. The key tool we use is the intimate connection we recall in Subsection 2.5.1 between the number of zero eigenvalues of the adjacency matrix of a tree and *maximal matchings* on the tree – a notion that we now review briefly.

Suppose that G is a graph with vertex set V and edge set E and for each edge $e \in E$ there is a corresponding *weight* $w(e)$. Recall that a *matching* of G is a subset of $S \subseteq E$ such that no two edges in S share a common vertex. A matching S^* is *maximal* for the system of weights $\{w(e) : e \in E\}$ if $\sum_{e \in S^*} w(e) \geq \sum_{e \in S} w(e)$ for any other matching S. There may be several maximal matchings but the total weight $\sum_{e \in S^*} w(e)$ is, of course, the same for all of them. When no weights are mentioned explicitly, they are assumed to be all 1, and the total weight of a maximal matching in this case is just the maximal possible cardinality of a matching.

Although we only need the case when all the weights are 1 to investigate the proportion of zero eigenvalues, our methods establish the following more general result without much further effort.

Theorem 2.3.5. *Consider a sequence $(\mathcal{T}_n)_{n=1}^{\infty}$ of random trees that converge in the probability fringe sense to a random* `sin`*-tree $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \ldots)$. Write M_n for the number of vertices of $\mathcal{T}_n$ and $M(\mathcal{T}_n)$ for the total weight of a maximal matching on $\mathcal{T}_n$ when the associated system of edge weights is a collection of independent and identically distributed $\mathbb{R}_+$-valued random variables $X_n(e)$ with a common distribution ν that has finite expected value. Then, $M_n^{-1} M(\mathcal{T}_n)$ converges in distribution to a (model dependent) constant $c_{\mathcal{T},\nu}$ as $n \to \infty$.*

Using their *objective method*, Aldous and Steele [5] show that $M_n^{-1}\mathbb{E}[M(\mathcal{T}_n)]$ converges in the case of the ensemble of uniform random trees. Moreover, they characterize the limit in terms of the fixed points of certain distributional identities.

2.3.4 Isospectrality

A result of Schwenk [125] states that the probability the adjacency matrix of a realization of the uniform random tree has the same spectrum as some other (non-isomorphic) tree converges to one as the number of vertices goes to infinity. Schwenk's method was developed further in [41]. The key idea is to first establish that a certain pair of non-isomorphic finite rooted trees $\mathbf{t}_1$ and $\mathbf{t}_2$ with the same number of vertices have the following *exchange property*: If $\mathbf{t}'$ is any finite rooted tree with $\mathbf{t}_1$ as a subtree, then replacing $\mathbf{t}_1$ by $\mathbf{t}_2$ produces a tree $\mathbf{t}''$ with

the same adjacency matrix spectrum as that of $\mathbf{t}'$. If one can then show that a given sequence $(\mathcal{T}_n)_{n=1}^{\infty}$ is such that $\mathbf{P}\{\mathbf{t}_1 \text{ is a subtree of } \mathcal{T}_n\} \to 1$ as $n \to \infty$, then $\mathbf{P}\{\mathcal{T}_n \text{ shares its spectrum with another tree}\} \to 1$ as $n \to \infty$. Pairs of trees with exchange property are exhibited in [125, 41]. Pairs of binary trees (that is, every non-leaf vertex has out-degree 2) with the exchange property are found in [105]. The following result is sufficiently obvious that we will not provide a proof. It applies to all four of the models in Section 2.1, with the pair $\mathbf{t}_1, \mathbf{t}_2$ being, for example, the binary trees in [105].

Proposition 2.3.6. *Consider a sequence $(\mathcal{T}_n)_{n=1}^{\infty}$ of random finite unlabeled rooted trees that converges in the probability fringe sense to a* `sin`*-tree $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \ldots)$. Suppose for some pair $\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{T}$ with the exchange property that $\mathbf{P}\{\mathcal{T}^0 = \mathbf{t}_1\} > 0$. Then,*

$$\lim_{n\to\infty} \mathbf{P}\{\mathcal{T}_n \text{ shares its spectrum with another tree}\} = 1.$$

2.3.5 Largest eigenvalues and largest degrees

The following result is proved in [71, 70, 48] in the case $a = 0$. The proof extends readily to general $a > -1$.

Theorem 2.3.7. *Let $(\mathcal{T}_n)_{n=1}^{\infty}$ be the ensemble of linear preferential attachment trees. Fix any $k \geq 1$. Write $\lambda_{n,1} \geq \lambda_{n,2} \geq \ldots \geq \lambda_{n,k}$ for the k largest eigenvalues of the adjacency matrix of $\mathcal{T}_n$ and denote by $\Delta_{n,1} \geq \Delta_{n,2} \geq \ldots \geq \Delta_{n,k}$ the k largest out-degrees of $\mathcal{T}_n$. Then, $\lambda_{n,i}/\sqrt{\Delta_{n,i}}$ converges in distribution to 1 as $n \to \infty$ for $1 \leq i \leq k$.*

We complement this result by establishing the following theorem. Recall that the linear preferential attachment model depends on a parameter $a > -1$. Define the corresponding *Malthusian parameter* by

$$\gamma_a := a + 2. \tag{2.2}$$

Theorem 2.3.8. *There exist random variables $X_1 \geq X_2 \geq \cdots X_k > 0$ that such that*

$$\left(\frac{\Delta_{n,1}}{n^{1/\gamma_a}}, \frac{\Delta_{n,2}}{n^{1/\gamma_a}}, \ldots, \frac{\Delta_{n,k}}{n^{1/\gamma_a}}\right)$$

converges in distribution to $(X_1, X_2, \ldots, X_k)$ as $n \to \infty$. Hence,

$$\left(\frac{\lambda_{n,1}}{n^{1/2\gamma_a}}, \frac{\lambda_{n,2}}{n^{1/2\gamma_a}}, \ldots, \frac{\lambda_{n,k}}{n^{1/2\gamma_a}}\right)$$

converges in distribution to $(\sqrt{X_1}, \sqrt{X_2}, \ldots, \sqrt{X_k})$ as $n \to \infty$.

2.4 Convergence of spectral distributions

2.4.1 Interlacing inequalities and some of their consequences

Suppose that A is an $m \times m$ Hermitian matrix and B is an $n \times n$ principal submatrix of A for $1 \leq n \leq m$ (that is, B is formed by deleting $m-n$ rows and columns of A with the same indices).

Write $\mu_1 \leq \ldots \leq \mu_m$ for the eigenvalues of A and $\nu_1 \leq \ldots \leq \nu_n$ for the eigenvalues of B. The interlacing theorem (see, for example, [85]) gives that $\mu_k \leq \nu_k \leq \mu_{k+m-n}$ for $1 \leq k \leq n$.

Write $P := \frac{1}{m}\sum_{i=1}^m \delta_{\mu_i}$ for the spectral distribution of A and $Q := \frac{1}{n}\sum_{i=1}^n \delta_{\nu_i}$ for the spectral distribution of B.

We wish to compare P and Q. To this end, we recall that the Lévy distance between two probability measures σ and τ on $\mathbb{R}$ is given by

$$d(\sigma,\tau) := \inf\{\varepsilon > 0 : S(x-\varepsilon) - \varepsilon < T(x) < S(x+\varepsilon) + \varepsilon,\ \forall x \in \mathbb{R}\},$$

where S and T are the cumulative distribution functions of σ and τ, respectively – see, for example, [144]. The Lévy distance is a metric that metrizes weak convergence of probability measures on $\mathbb{R}$, and the space of probability measures on $\mathbb{R}$ is complete with respect to this metric. Note also that if $\sigma \preceq \tau$, where $\preceq$ denotes stochastic domination (that is, $S(x) \geq T(x)$ for all $x \in \mathbb{R}$), and σ' and τ' are two other probability measures with $\sigma \preceq \sigma' \preceq \tau$ and $\sigma \preceq \tau' \preceq \tau$, then $d(\sigma',\tau') \leq d(\sigma,\tau)$. Finally, observe that if X and Y are two random variables on the same probability space that have marginal distributions σ and τ, respectively, then $d(\sigma,\tau) \leq \mathbf{P}\{X \neq Y\}$.

We return to the comparison of P and Q. Define two new probability measures by $P_* = \frac{1}{n}\sum_{i=1}^n \delta_{\mu_i}$ and $P^* = \frac{1}{n}\sum_{i=1}^n \delta_{\mu_{i+m-n}}$. It is clear that $P_* \preceq P \preceq P^*$ and, by the interlacing theorem, that $P_* \preceq Q \preceq P^*$. Thus,

$$d(P,Q) \leq d(P_*, P^*).$$

Suppose that on some probability space we have a random variable I that is uniformly distributed on $\{1,\ldots,n\}$. Define a pair of random variables (Z_*, Z^*) by

$$(Z_*, Z^*) = \begin{cases} (\mu_I, \mu_I), & \text{on the event } \{I \geq m-n+1\}, \\ (\mu_I, \mu_{n+I}), & \text{on the event } \{I \leq m-n\}. \end{cases}$$

Note that Z_* has marginal distribution P_* and Z^* has marginal distribution P^*. By construction, $\mathbf{P}\{Z_* \neq Z^*\} \leq [(m-n) \wedge n]/n$. We have therefore established the following result.

Proposition 2.4.1. *In the above notation,* $d(P,Q) \leq (\frac{m}{n} - 1) \wedge 1$.

Corollary 2.4.2. *Consider a sequence $(A_k)_{k=1}^{\infty}$ of Hermitian matrices, with $(A_k)_{k=1}^{\infty}$ being $m_k \times m_k$ and having spectral distribution P_k. For each $\varepsilon > 0$, let $(B_k^{\varepsilon})_{k=1}^{\infty}$ be such that B_k^{ε} is an $n_k^{\varepsilon} \times n_k^{\varepsilon}$ principal sub-matrix of A_k with spectral distribution Q_k^{ε}. Suppose for every $\varepsilon > 0$ that $Q_{\infty}^{\varepsilon} = \lim_{k\to\infty} Q_k^{\varepsilon}$ exists and $\limsup_{k\to\infty} m_k/n_k^{\varepsilon} \leq 1+\varepsilon$. Then, $P_{\infty} = \lim_{k\to\infty} P_k$ exists and is given by $P_{\infty} = \lim_{\varepsilon\downarrow 0} Q_{\infty}^{\varepsilon}$.*

Proof. From Proposition 2.4.1,

$$\begin{aligned}\limsup_{k,\ell\to\infty} d(P_k, P_\ell) &\leq \limsup_{k\to\infty} d(P_k, Q_k^{\varepsilon}) \\ &\quad + \limsup_{k,\ell\to\infty} d(Q_k^{\varepsilon}, Q_\ell^{\varepsilon}) + \limsup_{\ell\to\infty} d(Q_\ell^{\varepsilon}, P_\ell) \\ &\leq 2\varepsilon\end{aligned}$$

for each $\varepsilon > 0$. The sequence $(P_k)_{k=1}^{\infty}$ is thus Cauchy in the Lévy metric, and hence it converges weakly to a limit P_{∞}.

Moreover,

$$d(P_{\infty}, Q_{\infty}^{\varepsilon}) = \lim_{k\to\infty} d(P_k, Q_k^{\varepsilon}) \leq \varepsilon,$$

and so $P_{\infty} = \lim_{\varepsilon\downarrow 0} Q_{\infty}^{\varepsilon}$. □

Proposition 2.4.3. *In the notation at the beginning of this subsection,*

$$\begin{aligned}&|\#\{1 \leq k \leq m : \mu_k = \gamma\} - \#\{1 \leq k \leq n : \nu_k = \gamma\}| \\ &\quad = |mP(\{\gamma\}) - nQ(\{\gamma\})| \\ &\quad \leq (m-n)\end{aligned}$$

for all $\gamma \in \mathbb{R}$.

Proof. Suppose that $p = \#\{1 \leq k \leq m : \mu_k = \gamma\}$, with $\mu_{a+1} = \ldots \mu_{a+p} = \gamma$, and $q = \#\{1 \leq k \leq n : \nu_k = \gamma\}$, with $\nu_{b+1} = \ldots \nu_{b+q} = \gamma$. It follows from the interlacing inequalities that $\nu_{a+1} \leq \mu_{a+1}$, provided $a+1 \leq n$, and $\nu_{a+p-(m-n)} \leq \mu_{a+p}$ provided $a+p-(m-n) \geq 1$. Hence, $q \geq p-(m-n)$. Similarly, $\nu_{b+1} \leq \mu_{b+1+(m-n)}$ and $\mu_{b+q} \leq \nu_{b+q}$, so that $p \geq q-(m-n)$. Thus, $|p-q| \leq (m-n)$, as required. □

Corollary 2.4.4. *Consider a sequence $(A_k)_{k=1}^{\infty}$ of Hermitian matrices, with $(A_k)_{k=1}^{\infty}$ being $m_k \times m_k$ and having spectral distribution P_k. For each $\varepsilon > 0$, let $(B_k^{\varepsilon})_{k=1}^{\infty}$ be such that B_k^{ε} is an $n_k^{\varepsilon} \times n_k^{\varepsilon}$ principal sub-matrix of A_k with spectral distribution Q_k^{ε}. Suppose for some fixed $\gamma \in \mathbb{R}$ that for every $\varepsilon > 0$ the limit $\lim_{k\to\infty} Q_k^{\varepsilon}(\{\gamma\})$ exists and $\limsup_{k\to\infty} m_k/n_k^{\varepsilon} \leq 1+\varepsilon$. Then, $\lim_{k\to\infty} P_k(\{\gamma\})$ exists and is given by $\lim_{\varepsilon\downarrow 0} \lim_{k\to\infty} Q_k^{\varepsilon}(\{\gamma\})$.*

Proof. From Proposition 2.4.3,

$$|m_k P_k(\{\gamma\}) - n_k^{\varepsilon} Q_k^{\varepsilon}(\{\gamma\})| \leq (m_k - n_k^{\varepsilon}),$$

and so

$$|P_k(\{\gamma\}) - Q_k^\varepsilon(\{\gamma\})| \le \left(1 - \frac{n_k^\varepsilon}{m_k}\right) + \left(\frac{m_k}{n_k^\varepsilon} - 1\right).$$

An argument using completeness similar to that in the proof of Corollary 2.4.2 finishes the proof. □

Corollary 2.4.5. *Consider a forest* $\mathbf{u}$ *made up of finitely many finite unlabeled rooted trees, and assume that some eigenvalue* γ *of the adjacency matrix of* $\mathbf{u}$ *has multiplicity* L*. Suppose that* A *is the adjacency matrix of a finite unlabeled rooted tree* $\mathbf{t}$ *with* m *vertices, and suppose that there are* K *vertices* v *of* $\mathbf{t}$ *such that the forest formed by deleting* v *from the subtree below* v *produces the forest* $\mathbf{u}$*. Then,* γ *is an eigenvalue of the matrix* A *with multiplicity at least* $KL - m + (m - K) = K(L - 1)$*.*

Proof. The proof follows immediately by applying Proposition 2.4.3 to the matrix B that is the adjacency matrix of the graph obtained by deleting the K designated vertices from $\mathbf{t}$. The matrix B is block diagonal, and some of its blocks can be collected into K identical larger blocks that each form a copy of the adjacency matrix of the forest $\mathbf{u}$. It remains to observe that the set of eigenvalues of a block diagonal matrix is the union (including multiplicities) of the sets of eigenvalues of the respective blocks. □

2.4.2 Proof of Theorem 2.3.1

Suppose that the random tree $\mathcal{T}_n$ has M_n vertices and adjacency matrix A_n.

Fix a positive integer K. The construction of several objects in the proof will depend on K, but our notation will not record this.

Denote by W_n the set of vertices v of $\mathcal{T}_n$ such that the sub-tree below v (including v) contains at most K vertices. Put $N_n := \#W_n$. In the notation of Section 2.2, $N_n/M_n = \Phi(\mathcal{T}_n, \{(t_0, t_1, \ldots) : \#t_0 \le K\})$.

In order to avoid conflicting notation, write the limit `sin`-tree $\mathcal{T}$ as $(\mathcal{T}^0, \mathcal{T}^1, \ldots)$. By the assumption of probability fringe convergence, N_n/M_n converges in distribution to the constant $\mathbf{P}\{\#\mathcal{T}^0 \le K\}$. The latter constant can be made arbitrarily close to 1 by choosing K sufficiently large.

Denote by $\mathcal{U}_n$ the subgraph of $\mathcal{T}_n$ induced by the set of vertices W_n. That is, the graph $\mathcal{U}_n$ has vertex set W_n and two vertices in $\mathcal{U}_n$ are connected by an edge if they are connected by an edge in $\mathcal{T}_n$. The graph $\mathcal{U}_n$ is a forest.

Write X_{nk}, $1 \le k \le K$, for the set of vertices v of $\mathcal{T}_n$ with the following two properties:

- the subtree below v contains k vertices,
- if w is first vertex (other than v) on the path to the root from v, then w is on the path to the root for more than K vertices (that is, the subtree below w contains more than K vertices).

The set of roots of the trees in the the forest $\mathcal{U}_n$ is the disjoint union $\bigcup_{k=1}^{K} X_{nk}$. Put $R_{nk} := \#X_{nk}$, so that $N_n = \sum_{k=1}^{K} kR_{nk}$. It follows from the assumption of probability fringe convergence that $R_{nk}/M_n = \Phi(\mathcal{T}_n, \{(t_0, t_1, \ldots) : \#t_0 = k,\, \#t_0 + \#t_1 > K\})$ converges in distribution to the constant $p_k := \mathbf{P}\{\#\mathcal{T}^0 = k,\, \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$. Of course, the value of p_k depends on K and may be 0. However,

$$\begin{aligned}\sum_{k=1}^{K} kp_k &= \lim_{n\to\infty} \sum_{k=1}^{K} k\frac{R_{nk}}{M_n} \\ &= \lim_{n\to\infty} \frac{N_n}{M_n} \\ &= \mathbf{P}\{\#\mathcal{T}^0 \le K\}.\end{aligned}$$

Moreover, if we write

$$\Xi_{nk} := \frac{M_n}{R_{nk}} \Phi(\mathcal{T}_n, \cdot \cap \{(t_0, t_1, \ldots) : \#t_0 = k,\, \#t_0 + \#t_1 > K\})$$

for the empirical distribution of the subtrees rooted at the vertices in X_{nk} (with some suitable convention when $R_{nk} = 0$), then Ξ_{nk} is concentrated on the finite set of trees with k vertices and $\Xi_{nk}(\{\mathbf{t}\})$ converges in distribution when $p_k > 0$ to the constant

$$\Xi_k(\{\mathbf{t}\}) := \mathbf{P}\{\mathcal{T}^0 = \mathbf{t} \mid \#\mathcal{T}^0 = k,\, \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$$

for each such tree.

Denote by λ_k the distribution of an eigenvalue picked independently and uniformly at random from the k eigenvalues (counting possible multiplicities) of the $k \times k$ adjacency matrix of a k-vertex random tree with distribution Ξ_k. The probability measure λ_k is concentrated on the finite set of real numbers that are the possible eigenvalues of some tree with k vertices.

Write B_n for the adjacency matrix of the forest $\mathcal{U}_n$. This is a block diagonal matrix with R_{nk} many $k \times k$ blocks for $1 \le k \le K$. Recall that the set of eigenvalues of a block diagonal Hermitian matrix is the union of the eigenvalues of the blocks (including multiplicities). Thus, the spectral distribution of B_n converges in distribution to the deterministic probability measure

$$\frac{\sum_{k=1}^{K} kp_k\lambda_k}{\sum_{k=1}^{K} kp_k}$$

as $n \to \infty$.

An application of Corollary 2.4.2 completes the proof.

Remark 8. *It is instructive to consider what the various objects that appeared in the proof look like in a simple example. Suppose that $\mathcal{T}_n$ is the deterministic tree*

with $2^{n+1}-1$ vertices in which every non-leaf vertex has out-degree 2 and each leaf is distance n from the root. We say that $\mathcal{T}_n$ is a complete binary tree of height n. It is clear that $\mathcal{T}_n$ converges in the probability fringe sense to a random `sin`*-tree $(\mathcal{T}^0, \mathcal{T}^1, \ldots)$, where $\mathcal{T}^0$ is a complete binary tree of height H with $\mathbf{P}\{H = h\} = 2^{-h}$, $h = 0, 1, \ldots$, and $\mathcal{T}^i$ consists of a root connected by an edge to the root of a complete binary tree of height $H + i - 1$ for $i \geq 1$.*

If $2^{n+1} - 1 \geq K$ and ℓ is the unique integer such that $2^{\ell+1} - 1 \leq K < 2^{\ell+2} - 1$, then W_n is the set of vertices of $\mathcal{T}_n$ that are within distance at most ℓ of the leaves. Thus, $N_n = 2^{n-\ell}(2^{\ell+1} - 1)$. Moreover, the set X_{nk} is empty unless $k = 2^{\ell+1}$, in which case X_{nk} is the set of vertices of $\mathcal{T}_n$ that are at distance exactly ℓ from the leaves and $R_{nk} = 2^{n-\ell}$.

The sub-probability distribution $(p_k)_{k=1}^K$ assigns mass $2^{-\ell}$ to $2^{\ell+1} - 1$ and 0 elsewhere, while the probability measure Ξ_k is the point mass at the complete binary tree of height h when k is of the form $2^{h+1} - 1$. The spectral distribution of B_n converges to the spectral distribution of the complete binary tree of height ℓ.

2.4.3 Proof of Theorem 2.3.3

The proof is almost identical to that of Theorem 2.3.1 in Subsection 2.4.2. Recall from that proof the constant K, the probabilities $p_1, \ldots, p_K$, the probability distributions λ_k, $1 \leq k \leq K$, on $\mathbb{R}$, and the random adjacency matrix B_n with distribution depending on K and n. Recall also that the probability measure λ_k is concentrated on the finite set of real numbers that are the possible eigenvalues of some tree with k vertices.

It follows from the argument in Subsection 2.4.2 that the mass assigned by the spectral distribution of B_n to $\gamma \in \mathbb{R}$ converges in distribution to the deterministic probability measure

$$\frac{\sum_{k=1}^K k p_k \lambda_k(\{\gamma\})}{\sum_{k=1}^K k p_k}$$

as $n \to \infty$.

An application of Corollary 2.4.4 completes the proof.

2.4.4 Proof of Theorem 2.3.4

It follows from Corollary 2.4.5 that multiplicity of γ as an eigenvalue of the adjacency matrix of $\mathcal{T}_n$ is at least $(L-1)$ times the number of vertices v of $\mathcal{T}_n$ such that the forest formed by deleting v from the subtree below v produces the forest $\mathbf{u}$. By the assumption of probability fringe convergence, the proportion of eigenvalues of the adjacency matrix of $\mathcal{T}_n$ that have the value γ (that is, $F_n(\{\gamma\})$) satisfies

$$\mathbf{P}\{F_n(\{\gamma\}) > (L-1)\mathbf{P}\{\mathcal{U} = \mathbf{u}\} - \varepsilon\} \to 1$$

as $n \to \infty$ for any $\varepsilon > 0$. Moreover, because F_n converges weakly to F in distribution by Theorem 2.3.1,

$$\mathbf{P}\{F(\{\gamma\}) > F_n(\{\gamma\}) - \varepsilon\} \to 1$$

as $n \to \infty$ for any $\varepsilon > 0$. Combining these observations establishes that

$$F(\{\gamma\}) \geq (L-1)\mathbf{P}\{\mathcal{U} = \mathbf{u}\} > 0,$$

as required.

2.5 Maximal matchings and the number of zero eigenvalues

2.5.1 Combinatorial preliminaries

The following lemma is standard, but we include the proof for completeness.

Lemma 2.5.1. *Consider a tree* $\mathbf{t}$ *with* n *vertices and adjacency matrix* A*. Let* $\delta(\mathbf{t})$ *denote the number of zero eigenvalues* A*. Then*

$$\delta(\mathbf{t}) = n - 2M(\mathbf{t}),$$

where $M(\mathbf{t})$ *is the cardinality of a maximal matching of* $\mathbf{t}$*.*

Proof. It follows from the usual expansion of the determinant that the characteristic polynomial of the adjacency matrix of $\mathbf{t}$ is given by

$$\det(zI - A) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k N_k(\mathbf{t}) z^{n-2k},$$

where $N_k(\mathbf{t})$ is the number of matchings of $\mathbf{t}$ that contain k edges (see, for example, Addition Result 7b of [24]), and the result follows immediately. □

Our analysis of the cardinality of a maximal matching for a tree relies on the following "greedy" algorithm for producing a maximal matching of a forest. It is a simplification of one due to Karp and Sipser [94] that is intended to find approximate maximal matchings of more general sparse graphs. The algorithm takes an initial forest and iteratively produces forests with the same set of vertices but smaller sets of edges while at the same time adding edges to a matching of the initial forest. We stress that a leaf of a forest is a vertex with degree one.

- Input a forest $\mathbf{f}$ with vertices $V(\mathbf{f})$ and edges $E(\mathbf{f})$.

- Initialize $S \leftarrow \emptyset$.
- While $E(\mathbf{f}) \neq \emptyset$ do
  - Choose a leaf, say x, and let $\{x, y\}$ be the unique edge in $\mathbf{f}$ incident to x.
  - Set $E(\mathbf{f}) \leftarrow \{e \in E(\mathbf{f}) : e \cap \{x, y\} = \emptyset\}$, and $S \leftarrow S \cup \{\{x, y\}\}$.
- Output the matching S.

Lemma 2.5.2. *The algorithm produces a maximal matching as its output.*

Proof. Let x be any leaf of the forest, and write $\{x, y\}$ for the unique incident edge. Note that every maximal matching either contains the edge $\{x, y\}$ or an edge of the form $\{y, z\}$ for some vertex $z \neq x$, because otherwise $\{x, y\}$ could be added to a putative maximal matching that contains no edge of the form $\{y, w\}$ to produce a matching with a larger cardinality. Also, note that replacing any edge of the form $\{y, z\}$ with $z \neq x$ that appears in some matching by the edge $\{x, y\}$ results in a collection of edges that is also a matching and has the same cardinality. It follows that the edge $\{x, y\}$ must belong to at least one maximal matching.

The result now follows by induction on the number of edges in the forest. □

Note that we are free to take any current leaf at each iteration of the "while" step of the algorithm. We start with some initial set of leaves and each iteration of the while step removes some leaves (by turning them into isolated vertices) as well as sometimes producing new leaves. We can therefore think of the leaves present after the completion of each while step as being labeled with the number of the step at which that vertex became a leaf, where the leaves in the initial forest are labeled with 0. We adopt the convention that in any iteration of the while step we take one of the current leaves with the lowest label.

Put $i_0 = 0$ and define $i_1, i_2, \ldots$ inductively by setting i_{k+1} to be the number of iterations of the while step required until all of the leaves with labels at most i_k are turned into isolated vertices, where $i_{k+1} = i_k$ if the forest after i_k iterations already consists of only isolated vertices. The numbers i_k are eventually constant and this final value is the cardinality of a maximal matching.

Each of the iterations $i_k + 1, \ldots, i_{k+1}$ of the while step is one or the other of the following two types.

Type I: The iteration removes all of the edges of the form $\{y, z\}$, where the vertex y is not a leaf with label at most i_k and there is a leaf x with label at most i_k such that $\{y, x\}$ is an edge (so that y is at graph distance 1 from the leaves of the forest present after i_k iterations).

Type II: The iteration removes an edge of the form $\{y, z\}$ such that y and z are both leaves with label at most i_k (we say that $\{y, z\}$ is an *isolated edge* in the forest present after i_k iterations).

Therefore, the cardinality of a maximal matching is the number of vertices that will be at graph distance 1 from the current leaves after i_k iterations of the while step for some k plus the number of edges in the initial forest that will eventually become isolated edges after i_k iterations of the while step for some k. We next introduce some notation to describe the sets of vertices and edges we have just characterized.

Write $\mathbf{f}_k$, $E_k(\mathbf{f})$, $L_k(\mathbf{f})$, and $I_k(\mathbf{f})$, respectively, for the forest, the set of edges, the set of leaves, and the set of isolated vertices after i_k iterations of the while step starting from the initial forest $\mathbf{f}$. Note that $E_k(\mathbf{f})$ is obtained by removing all edges $\{y, z\} \in E_{k-1}(\mathbf{f})$ such that there exists $x \in L_{k-1}(\mathbf{f})$ with $\{x, y\} \in E_{k-1}(\mathbf{f})$. Equivalently, $E_k(\mathbf{f})$ consists of exactly those edges $\{u, v\} \in E_{k-1}(\mathbf{f})$ such that both vertices u and v are at graph distance at least 2 from $L_{k-1}(\mathbf{f})$ in $\mathbf{f}_{k-1}$. This means that vertices that are distance 0 or 1 from $L_{k-1}(\mathbf{f})$ in $\mathbf{f}_{k-1}$ are isolated in $\mathbf{f}_k$, and vertices that are at graph distance 2 or greater from $L_{k-1}(\mathbf{f})$ have degree in $\mathbf{f}_k$ equal to the number of their neighbors in $\mathbf{f}_{k-1}$ that are at graph distance 2 or greater from $L_{k-1}(\mathbf{f})$.

We further introduce new sets $G_k(\mathbf{f}), H_k(\mathbf{f})$ and $J_k(\mathbf{f})$ as follows:

$$\begin{aligned} G_k(\mathbf{f}) &:= \{u \in L_k(\mathbf{f}) : \exists v \in L_k(\mathbf{f}) \text{ so that } \{u, v\} \in E_k(\mathbf{f})\}, \\ H_k(\mathbf{f}) &:= \{u \in V(\mathbf{f}) \setminus L_k(\mathbf{f}) : \exists v \in L_k(\mathbf{f}) \text{ so that } \{u, v\} \in E_k(\mathbf{f})\}, \\ J_k(\mathbf{f}) &:= (I_{k+1}(\mathbf{f}) \setminus I_k(\mathbf{f})) \setminus (G_k(\mathbf{f}) \cup H_k(\mathbf{f})). \end{aligned}$$

In words, $G_k(\mathbf{f})$ is the set of leaves that are one of the two leaves of an isolated edge present after i_k iterations of the while step – these are the vertices that become isolated during iterations $i_k + 1, \ldots, i_{k+1}$ due to Type II steps, $H_k(\mathbf{f})$ is the set of vertices that are graph distance 1 from the leaves after i_k iterations of the while step – these are the non-leaf vertices that become isolated during iterations $i_k + 1, \ldots, i_{k+1}$ due to Type I steps, and $J_k(\mathbf{f})$ is the remaining set of vertices that become isolated during iterations $i_k + 1, \ldots, i_{k+1}$ (all due to Type I steps). Note that $V(\mathbf{f})$ is the disjoint union of $I_0(\mathbf{f})$ and $G_k(\mathbf{f}), H_k(\mathbf{f}), J_k(\mathbf{f})$, $k \geq 0$, and so

$$\#V(\mathbf{f}) = \#I_0(\mathbf{f}) + \sum_{k=0}^{\infty} \left(\#G_k(\mathbf{f}) + \#H_k(\mathbf{f}) + \#J_k(\mathbf{f})\right).$$

Clearly, all the above objects can also be defined for an infinite forest $\mathbf{f}$ such that every vertex is at a finite graph distance from a leaf, (that is, a vertex of degree one).

The discussion above leads immediately to the following result.

Lemma 2.5.3. *The cardinality of a maximal matching of a finite forest* $\mathbf{f}$ *is*

$$M(\mathbf{f}) = \sum_{k=0}^{\infty} \#H_k(\mathbf{f}) + \frac{1}{2} \sum_{k=0}^{\infty} \#G_k(\mathbf{f}).$$

Consequently, the number of zero eigenvalues of the adjacency matrix of a finite tree

t *is*

$$\delta(\mathbf{t}) = \#V(\mathbf{t}) - 2M(\mathbf{t}) = \sum_{k=0}^{\infty} \#J_k(\mathbf{t}) - \sum_{k=0}^{\infty} \#H_k(\mathbf{t}).$$

Example 2.5.4. Consider the tree $\mathbf{t}$ with vertices $\{1, \ldots, m\}$ and edges connecting successive integers. The cardinality of a maximal matching is obviously $(m-1)/2$ when m is odd and $m/2$ when is even (so that $\delta(\mathbf{t})$ is 1 when m is odd and 0 when is even). There are four cases to consider in checking that this agrees with the formula of Lemma 2.5.3.

Case I: m is odd and $(m-1)/2$ is odd ($\Leftrightarrow m \equiv 3 \mod 4$).

Then, $H_0(\mathbf{t}) = \{2, m-1\}$, $H_1(\mathbf{t}) = \{4, m-3\}$, ..., $H_{(m-3)/4}(\mathbf{t}) = \{(m+1)/2\}$, all other $H_k(\mathbf{t})$ are empty, and all $G_k(\mathbf{t})$ are empty. The formula of Lemma 2.5.3 gives $2 \times (m-3)/4 + 1 = (m-1)/2$.

Case II: m is odd and $(m-1)/2$ is even ($\Leftrightarrow m \equiv 1 \mod 4$).

Then, $H_0(\mathbf{t}) = \{2, m-1\}$, $H_1(\mathbf{t}) = \{4, m-3\}$, ..., $H_{(m-5)/4}(\mathbf{t}) = \{(m-1)/2, (m+3)/2\}$, all other $H_k(\mathbf{t})$ are empty, and all $G_k(\mathbf{t})$ are empty. The formula of Lemma 2.5.3 gives $2 \times ((m-5)/4 + 1) = (m-1)/2$.

Case III: m is even and $(m-2)/2$ is odd ($\Leftrightarrow m \equiv 0 \mod 4$).

Then, $H_0(\mathbf{t}) = \{2, m-1\}$, $H_1(\mathbf{t}) = \{4, m-3\}$, ..., $H_{(m-4)/4}(\mathbf{t}) = \{m/2, (m+2)/2\}$, all other $H_k(\mathbf{t})$ are empty, and all $G_k(\mathbf{t})$ are empty. The formula of Lemma 2.5.3 gives $2 \times ((m-4)/4 + 1) = m/2$.

Case IV: m is even and $(m-2)/2$ is even ($\Leftrightarrow m \equiv 2 \mod 4$)

Then, $H_0(\mathbf{t}) = \{2, m-1\}$, $H_1(\mathbf{t}) = \{4, m-3\}$, ..., $H_{(m-6)/4}(\mathbf{t}) = \{(m-2)/2, (m+4)/2\}$, all other $H_k(\mathbf{t})$ are empty, $G_{(m-2)/4}(\mathbf{t}) = \{(m/2, (m+2)/2\}$, and all other $G_k(\mathbf{t})$ are empty. The formula of Lemma 2.5.3 gives $2 \times ((m-6)/4 + 1) + 1 = m/2$.

2.5.2 Maximal weighted matchings: Proof of Theorem 2.3.5

We will use the same construction as we used in the proof of Theorem 2.3.1 in Subsection 2.4.2.

Recall that for a fixed positive integer K this construction produced for each n a set of vertices W_n of $\mathcal{T}_n$ with cardinality N_n such that N_n/M_n, where M_n is the number of vertices of $\mathcal{T}_n$, converged in distribution to $\mathbf{P}\{\#\mathcal{T}^0 \leq K\}$ – a constant that can be made arbitrarily close to 1 by choosing K sufficiently large.

The subgraph of $\mathcal{T}_n$ induced by W_n was the forest $\mathcal{U}_n$ rooted at the points $\bigcup_{k=1}^{K} X_{nk}$ and $\#X_{nk}/M_n$ converged in distribution to the constant $p_k := \mathbf{P}\{\#\mathcal{T}^0 = k,\ \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$.

Moreover, the random probability measure Ξ_{nk} given by the empirical distribution of the subtrees rooted at the vertices in X_{nk} was concentrated on the finite set of trees with k vertices and $\Xi_{nk}(\{\mathbf{t}\})$ converged in distribution when $p_k > 0$ to the constant

$$\Xi_k(\{\mathbf{t}\}) := \mathbf{P}\{\mathcal{T}^0 = \mathbf{t} \mid \#\mathcal{T}^0 = k,\ \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$$

for each such tree $\mathbf{t}$.

Write $M(\mathcal{T}_n)$ (respectively, $M(\mathcal{U}_n)$ for the total weight of a maximal matching on $\mathcal{T}_n$ (respectively, $\mathcal{U}_n$) for the independent, identically distributed edge weights $X_n(e)$, where e ranges over the edges of $\mathcal{T}_n$.

Note that a maximal matching on $\mathcal{U}_n$ is obtained by separately constructing maximal matchings on each component subtree of $\mathcal{U}_n$. It follows from Lemma 2.5.5 below that $M_n^{-1}M(\mathcal{U}_n)$ converges in distribution to

$$\sum_{k=1}^{K} p_k \sum_{\mathbf{t}:\#\mathbf{t}=k} \Xi_k(\{\mathbf{t}\})\mu(\mathbf{t}),$$

where $\mu(\mathbf{t})$ is the expected value of the total weight of a maximal matching on $\mathbf{t}$ when the weights of the edges are independent and identically distributed with common distribution ν.

Observe that any matching on $\mathcal{U}_n$ is also a matching on $\mathcal{T}_n$ and that the restriction of any matching on $\mathcal{T}_n$ to $\mathcal{U}_n$ is a matching on $\mathcal{U}_n$. Thus,

$$M(\mathcal{U}_n) \le M(\mathcal{T}_n) \le M(\mathcal{U}_n) + \sum_{e\in E(\mathcal{T}_n)\setminus E(\mathcal{U}_n)} X_n(e),$$

where $E(\mathcal{T}_n)$ (respectively, $E(\mathcal{U}_n)$) is the set of edges of $\mathcal{T}_n$ (respectively, $\mathcal{U}_n$).

There is an element of $E(\mathcal{T}_n)\setminus E(\mathcal{U}_n)$ for each vertex of $\mathcal{T}_n$ other than the root that is not a vertex of $\mathcal{U}_n$ and one for each root of a subtree in the forest $\mathcal{U}_n$. Thus, writing μ for the common expected value of the edge weights,

$$\mathbb{E}\left[M_n^{-1} \sum_{e\in E(\mathcal{T}_n)\setminus E(\mathcal{U}_n)} X_n(e) \;\middle|\; \mathcal{T}_n \right] = M_n^{-1}\left[(M_n - N_n - 1)_+ + \sum_{k=1}^{K} \#X_{nk} \right]\mu.$$

Note from above that $1-M_n^{-1}N_n$ converges in distribution to the constant $\mathbf{P}\{\#\mathcal{T}^0 > K\}$ and $M_n^{-1}\sum_{k=1}^{K}\#X_{nk}$ converges in distribution to the constant $\sum_{k=1}^{K} p_k = \mathbf{P}\{\#\mathcal{T}^0 \le K, \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$ as $n\to\infty$. Both of these constants converge to 0 as $K\to\infty$. It follows that

$$\lim_{K\to\infty}\lim_{n\to\infty} \mathbf{P}\left\{ M_n^{-1} \sum_{e\in E(\mathcal{T}_n)\setminus E(\mathcal{U}_n)} X_n(e) > \varepsilon \right\} = 0$$

for all $\varepsilon > 0$.

Therefore, $M_n^{-1}M(\mathcal{T}_n)$ converges in distribution as $n \to \infty$ to the constant

$$\lim_{K\to\infty} \sum_{k=1}^{K} p_k \sum_{\mathbf{t}:\#\mathbf{t}=k} \Xi_k(\{\mathbf{t}\})\mu(\mathbf{t}),$$

where we stress that p_k and Ξ_k depend on K, even though this is not indicated by our notation.

The following lemma, which we used above, is a straightforward consequence of the strong law of large numbers.

Lemma 2.5.5. *For $i = 1, 2, \ldots$ let L^i be a positive integer-valued random variable and $\theta^i_1, \ldots, \theta^i_{L^i}$ be random variables taking values in a finite set Θ. Suppose that as $i \to \infty$ the random variable L^i converges in distribution to ∞ and for each $\theta \in \Theta$ the random variable*

$$\frac{\#\{1 \le j \le L^i : \theta^i_j = \theta\}}{L^i}$$

converges in distribution to a constant $\pi(\theta)$. Let $\xi^i_1, \ldots, \xi^i_{L^i}$ be $\mathbb{R}_+$-valued random variables that are conditionally independent given $\theta^i_1, \ldots, \theta^i_{L^i}$, and such that

$$\mathbf{P}\{\xi^i_j \in A \mid \theta^i_1, \ldots, \theta^i_{L^i}\} = \Pi(\theta^i_j; A)$$

for some collection of Borel probability measures $(\Pi(\theta; \cdot))_{\theta\in\Theta}$. Suppose that

$$\upsilon(\theta) := \int_{\mathbb{R}_+} x\Pi(\theta; dx) < \infty$$

for all $\theta \in \Theta$. Then,

$$\frac{\sum_{j=1}^{L^i} \xi^i_j}{L^i}$$

converges in distribution to

$$\sum_{\theta\in\Theta} \pi(\theta)\upsilon(\theta)$$

as $i \to \infty$.

Remark 9. *Let a sequence $(\mathcal{T}_n)_{n=1}^{\infty}$ of random unlabeled rooted trees converge in the probability fringe sense to a random* `sin`*-tree $\mathcal{T}$. Let M_n be the number of vertices in $\mathcal{T}_n$. Consider the case when each edge-weight is identically one. Write $I(\mathcal{T}_n)$ for the cardinality of a maximal independent set for $\mathcal{T}_n$. By König's theorem [36], for a general bipartite graph the cardinality of a maximal matching is equal to the cardinality of a minimal vertex cover. On the other hand, complementation of a minimal vertex cover in any graph always yields a maximal independent set. Thus, $I(\mathcal{T}_n) = M_n - M(\mathcal{T}_n)$ in our case. Consequently, $M_n^{-1}I(\mathcal{T}_n)$ also converges in distribution to a (model-dependent) constant $\kappa_{\mathcal{T}} \ge 1/2$ as $n \to \infty$.*

2.5.3 Asymptotics of the number of zero eigenvalues

If we combine Theorem 2.3.5 on the rescaled convergence of the total weight of a maximal weighted matching with Lemma 2.5.1 on the connection between the cardinality of a maximal matching and the number of zero eigenvalues of the adjacency matrix, then we get another proof of Theorem 2.3.3 on the convergence of $F_n(\{\gamma\})$ in the special case when $\gamma = 0$. We now improve this result by using Lemma 2.5.3 to give a formula for the limit in terms of features of the limit `sin`-tree. We then show that how this formula may be used to get explicit lower bounds on the limit.

Proposition 2.5.6. *Consider a sequence $(\mathcal{T}_n)_{n=1}^{\infty}$ of random unlabeled rooted trees, where $\mathcal{T}_n$ has M_n vertices. Suppose that $(\mathcal{T}_n)_{n=1}^{\infty}$ converges in the probability fringe sense to a random* `sin`*-tree $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \ldots)$ and write R for the root of $\mathcal{T}^0$. Then $F_n(\{0\})$ converges in distribution as $n \to \infty$ to*

$$\sum_{k=0}^{\infty} \Big(\mathbf{P}\{R \in J_k(\mathcal{T})\} - \mathbf{P}\{R \in H_k(\mathcal{T})\} \Big).$$

Proof. In view of Theorem 2.3.3, its enough to prove the convergence of $F_n(\{0\})$ to the desired quantity in expectation. If V is a vertex chosen uniformly at random from $\mathcal{T}_n$, then, by Lemma 2.5.3 we can write $\mathbb{E}[F_n(\{0\})] = M_n^{-1}\,\mathbb{E}[\delta(\mathcal{T}_n)]$ as

$$\begin{aligned} &\sum_{k=0}^{\infty} \left(M_n^{-1}\,\mathbb{E}[\#J_k(\mathcal{T}_n)] - M_n^{-1}\,\mathbb{E}[\#H_k(\mathcal{T}_n)] \right) \\ &\quad = \sum_{k=0}^{\infty} \left(\mathbf{P}\{V \in J_k(\mathcal{T}_n)\} - \mathbf{P}\{V \in H_k(\mathcal{T}_n)\} \right). \end{aligned} \tag{2.3}$$

Given a tree $\mathbf{t} \in \mathbb{T}$ with root ρ and a vertex $v \in \mathbf{t}$, write $\mathcal{N}_k(v, \mathbf{t})$ for the subtree of $\mathbf{t}$ induced by vertices that are at graph distance at most k from v. Note that whether or not a vertex v of $\mathbf{t}$ belongs to the sets $H_k(\mathbf{t})$ or $J_k(\mathbf{t})$ can be determined by examining the neighborhood $\mathcal{N}_{2k+4}(v, \mathbf{t})$. Observe also that $(t_0, t_1, \ldots, t_h, *, *, \ldots) \in \mathbb{T}_*^{\infty}$ is the decomposition of $\mathbf{t}$ relative to ρ and v, then $\mathcal{N}_k(v, \mathbf{t})$ can be reconstructed from $(t_0, t_1, \ldots, t_{k \wedge h})$.

Recall that $J_k(\mathcal{T}_n)$ and $H_k(\mathcal{T}_n)$ are both subsets of $I_{k+1}(\mathcal{T}_n) \setminus I_k(\mathcal{T}_n)$, and so

$$|\mathbf{P}\{V \in J_k(\mathcal{T}_n)\} - \mathbf{P}\{V \in H_k(\mathcal{T}_n)\}| \leq \mathbf{P}\{V \in I_{k+1}(\mathcal{T}_n) \setminus I_k(\mathcal{T}_n)\}.$$

Moreover, for any nonnegative integer m,

$$\begin{aligned}
&\sum_{k=m}^{\infty} \mathbf{P}\{V \in I_{k+1}(\mathcal{T}_n) \setminus I_k(\mathcal{T}_n)\} \\
&\quad = \mathbf{P}\{V \in V(\mathcal{T}_n) \setminus I_m(\mathcal{T}_n)\} \\
&\quad \le \mathbf{P}\{\text{the subtree of } \mathcal{T}_n \text{ below } V \text{ contains at least } 2m-1 \text{ vertices }\} \\
&\quad \to \mathbf{P}\{\#\mathcal{T}^0 \ge 2m-1\}
\end{aligned}$$

as $n \to \infty$, by the assumption of probability fringe convergence. The last term clearly converges to 0 as $m \to \infty$. A similar argument shows that the analogous series involving the limiting `sin`-tree is also absolutely convergent.

Finally, it follows from the assumption of probability fringe convergence and our observations above about membership of $H_k(\mathcal{T})$ and $J_k(\mathcal{T})$ being locally determined that for each $k \ge 1$, the first k terms of the series (2.3) converge to the corresponding terms of the desired infinite series involving the limiting `sin`-tree. □

By construction, for any tree $\mathbf{t}$, $G_k(\mathbf{t}) \subseteq L_k(\mathbf{t})$ and $L_k(\mathbf{t}) \setminus G_k(\mathbf{t}) \subseteq J_k(\mathbf{t})$. Set $K_k(\mathbf{t}) := J_k(\mathbf{t}) \setminus (L_k(\mathbf{t}) \setminus G_k(\mathbf{t}))$. That is, $K_k(\mathbf{t})$ consists of vertices that become isolated due to Type I steps during iterations $i_k + 1, \ldots, i_{k+1}$ of the Karp-Sipser algorithm but are not leaves in the forest present after iteration i_k; for example, if $\mathbf{t}$ has vertices $\{1,2,3,4,5\}$ and adjacent integers are joined by edges, then $J_0(\mathbf{t}) = \{1,3,5\}$, $L_0(\mathbf{t}) \setminus G_0(\mathbf{t}) = L_0(\mathbf{t}) = \{1,5\}$, and $K_0(\mathbf{t}) = \{3\}$. Note that for each $v \in H_k(\mathbf{t})$ there exists $u \in L_k(\mathbf{t}) \setminus G_k(\mathbf{t})$ such that $\{v,u\} \in E_k(\mathbf{t})$. Also, if v_1, v_2 are distinct elements of $H_k(\mathbf{t})$ and $u_1, u_2 \in L_k(\mathbf{t}) \setminus G_k(\mathbf{t})$ are such that $\{v_1,u_1\}, \{v_2,u_2\} \in E_k(\mathbf{t})$, then u_1 and u_2 are also distinct. Consequently, $\#L_k(\mathbf{t}) - \#G_k(\mathbf{t}) - \#H_k(\mathbf{t}) \ge 0$. Applying this observation to $\mathcal{T}_n$, dividing by M_n, and taking the limit as $n \to \infty$, we deduce that the formula in Proposition 2.5.6 for the limit of $F_n(\{0\})$ may be written as a sum over k of the sum of the two nonnegative terms $\mathbf{P}\{R \in L_k(\mathcal{T})\} - \mathbf{P}\{R \in G_k(\mathcal{T})\} - \mathbf{P}\{R \in H_k(\mathcal{T})\}$ and $\mathbf{P}\{R \in K_k(\mathcal{T})\}$. We may give good lower bounds for the first few of these summands with relative ease.

We first find lower bound on $\mathbf{P}\{R \in L_0(\mathcal{T})\} - \mathbf{P}\{R \in G_0(\mathcal{T})\} - \mathbf{P}\{R \in H_0(\mathcal{T})\}$. Note for any tree $\mathbf{t}$ with 3 or more vertices that $G_0(\mathbf{t}) = \emptyset$. Observe also that

$$\begin{aligned}
&\#L_0(\mathcal{T}_n) - \#H_0(\mathcal{T}_n) \\
&= \sum_{m=2}^{\infty} (m-1) \times \#\{u \in H_0(\mathcal{T}_n) : u \text{ is connected to exactly } m \text{ vertices in } L_0(\mathcal{T}_n)\} \\
&\ge \sum_{m=2}^{\infty} (m-1) \times \#\{u \in V(\mathcal{T}_n) : \text{the subtree below } u \text{ is an } m\text{-star}\},
\end{aligned}$$

where by a m-star we mean a unlabeled rooted tree with $(m+1)$ vertices in which

the root is connected to each of the other m vertices via an edge. Therefore,

$$\begin{aligned}\mathbf{P}\{R \in L_0(\mathcal{T})\} - \mathbf{P}\{R \in H_0(\mathcal{T})\} &= \lim_{n\to\infty} M_n^{-1}(\mathbb{E}[\#L_0(\mathcal{T}_n)] - \mathbb{E}[\#H_0(\mathcal{T}_n)]) \\ &\geq \sum_{m=2}^{\infty}(m-1) \times \mathbf{P}\{\mathcal{T}^0 \text{ is an } m\text{-star}\}.\end{aligned}$$

On the other hand, it is easy to check that $\mathbf{P}\{R \in K_0(\mathcal{T})\} \geq \mathbf{P}\{\mathcal{T}^0 \in \mathbb{T}'', \mathcal{T}^1 \in \mathbb{T}'\}$, where $\mathbb{T}' \subseteq \mathbb{T}$ is the set of finite unlabeled rooted trees for which the root has at least one child that has no children, and $\mathbb{T}'' \subseteq \mathbb{T}$ is the set of finite unlabeled rooted trees for which the root has single child and that child in turn has at least one child that has no children.

As one might expect, finding good lower bounds on the terms $\mathbf{P}\{R \in L_k(\mathcal{T})\} - \mathbf{P}\{R \in G_k(\mathcal{T})\} - \mathbf{P}\{R \in H_k(\mathcal{T})\}$ and $\mathbf{P}\{R \in K_k(\mathcal{T})\}$ becomes increasingly difficult as k gets larger. However, we can still get crude lower bounds by computing the probability of appearance of special kinds of trees in the first few fringes in the limiting `sin`-tree. For example,

$$\begin{aligned}&\mathbf{P}\{R \in L_k(\mathcal{T})\} - \mathbf{P}\{R \in G_k(\mathcal{T})\} - \mathbf{P}\{R \in H_k(\mathcal{T})\} \\ &\quad \geq \mathbf{P}\{\mathcal{T}^0 \text{ is a complete binary tree of depth } (2k+1) \\ &\qquad \text{and } \mathcal{T}^i = \bullet \text{ for } 1 \leq i \leq 2k-2\}.\end{aligned}$$

where $\bullet$ denotes the rooted tree with a single vertex. The proof follows along the same lines as the $k = 0$ case above. Furthermore,

$$\begin{aligned}&\mathbf{P}\{R \in K_k(\mathcal{T})\} \\ &\quad \geq \mathbf{P}\{\mathcal{T}^0 \text{ is a path of length } (2k+2)\ , \\ &\qquad \mathcal{T}^i = \bullet \text{ for } 1 \leq i \leq 2k, \\ &\qquad \text{and } \mathcal{T}^{2k+1} \text{ is a 1-star}\}.\end{aligned}$$

For the ensemble of linear preferential attachment trees with parameter $a = 0$, it is well known (see, for example, [62]) that the proportion of vertices with degree d converges in distribution to $p_d = 4/d(d+1)(d+2)$. Specializing to $d = 1$, we see that $n^{-1}\#L_0(\mathcal{T}_n)$ converges in distribution to $2/3$, and so $\mathbf{P}\{R \in L_0(\mathcal{T})\} = 2/3$. Hence,

$$\lim_{n\to\infty} F_n(\{0\}) \geq \mathbf{P}\{R \in L_0(\mathcal{T})\} - \mathbf{P}\{R \in H_0(\mathcal{T})\} \geq 2\mathbf{P}\{R \in L_0(\mathcal{T})\} - 1 = 1/3.$$

Now consider the ensemble of random recursive trees. Recall Construction 2.2.3(a). Let $\xi_i, \xi_i', i \geq 1$ and X be i.i.d. exponential random variables with rate 1. To get a lower bound on $\lim_{n\to\infty} F_n(\{0\})$, we may use the inequality $\lim_{n\to\infty} F_n(\{0\}) \geq$

$\sum_{m=2}^{\infty}(m-1)\times\mathbf{P}\{\mathcal{T}^0 \text{ is an } m\text{-star}\}$ where

$$\begin{aligned}\mathbf{P}\{\mathcal{T}^0 \text{ is an } m\text{-star}\} &\geq \mathbb{E}\left[\mathbf{P}\left\{\sum_{i=1}^{m}\xi_i \leq X, \sum_{i=1}^{m+1}\xi_i > X \,\middle|\, X\right\}\prod_{i=1}^{m}\mathbf{P}\{\xi_i' > X \mid X\}\right]\\ &= \mathbb{E}\left[\mathbf{P}\left\{\sum_{i=1}^{m}\xi_i \leq X, \sum_{i=1}^{m+1}\xi_i > X \,\middle|\, X\right\}e^{-mX}\right]\\ &= \mathbb{E}\left[e^{-X}\frac{X^m}{m!}e^{-mX}\right]\\ &= \frac{1}{m!}\int_0^{\infty} x^m e^{-(m+2)x}\,dx\\ &= (m+2)^{-(m+1)}.\end{aligned}$$

For the uniform random trees, we can easily obtain lower bounds for various terms using the description of the fringes of the limiting `sin`-tree in terms of critical Poisson Galton-Watson trees. For example,

$$\mathbf{P}\{\mathcal{T}^0 \text{ is an } m\text{-star}\} = \frac{e^{-1}}{m!}\times(e^{-1})^m = \frac{e^{-(m+1)}}{m!},$$

$$\mathbf{P}\{\mathcal{T}^1 \in \mathbb{T}'\} = 1 - \sum_{i=0}^{\infty}\frac{e^{-1}}{i!}\times(1-e^{-1})^i = 1 - e^{-1}e^{1-e^{-1}},$$

and

$$\mathbf{P}\{\mathcal{T}^0 \in \mathbb{T}''\} = e^{-1}\times(1-e^{-1}e^{1-e^{-1}}).$$

Therefore,

$$\begin{aligned}\lim_n F_n(\{0\}) &\geq \sum_{m=2}^{\infty}(m-1)\times\mathbf{P}\{\mathcal{T}^0 \text{ is a } m\text{-star}\} + \mathbf{P}\{\mathcal{T}^0 \in \mathbb{T}'', \mathcal{T}^1 \in \mathbb{T}'\}\\ &= e^{-1}(1-(1-e^{-1})e^{e^{-1}}) + e^{-1}(1-e^{-1}e^{1-e^{-1}})^2.\end{aligned}$$

2.6 Largest eigenvalues: Proof of Theorem 2.3.8

We first recall from Proposition 2.2.1(b) how $\mathcal{T}_n$, the linear preferential attachment tree on n vertices with parameter $a > -1$, can be constructed from a particular continuous-time branching process.

Denote by $N_a = (N_a(t))_{t\geq 0}$ a pure birth process that starts with a single progenitor and when there have been k births a new birth occurs at rate $k+1+a$. Recall that $\mathcal{F}(t) \in \mathbb{T}$ is the family tree at time $t \geq 0$ of the continuous-time branching process

in which the birth process of each individual is a copy of N_a. Then, $\mathcal{T}_n$ has the same distribution as $\mathcal{F}(T_n)$, where $T_n := \inf\{t > 0 : \#\mathcal{F}(t) = n\}$.

We now record some useful facts about the birth process N_a. Recall the Malthusian rate of growth parameter $\gamma_a := a + 2$.

Lemma 2.6.1. *(a) For any fixed time $t \geq 0$, the random variable $\mathbf{P}\{N_0(t) = k\} = (1-e^{-t})^k e^{-t}$, $k = 0, 1, \ldots$. That is, $N_0(t)$ is distributed as the number of failures until the first success in a sequence of independent Bernoulli trials with common success probability e^{-t}.*
(b) For $a > -1$, set $A := \lceil a + 1 \rceil$. Then,

$$\mathbf{P}\{N_a(t) > Ke^t\} \leq Ae^{-\frac{K}{A}}$$

for all $K > 0$ and $t \geq 0$.

Proof. For part(a), note that $N_0 + 1$ is a Yule process – the birth rate in state ℓ is ℓ – and the claimed distribution is well-known.

To prove part (b), suppose that $M = (M(t))_{t\geq 0}$ is a Yule process started in state A (that is, M is pure birth process and the birth rate in state ℓ is ℓ). Then, $(M(t)-A)_{t\geq 0}$ is a pure birth process that starts in state 0 and has birth rate $\ell + A \geq \ell + 1 + a$ in state ℓ. It is therefore possible to couple M and N_a in such a way that $N_a(t) \leq M(t) - A$ for all $t \geq 0$. Observe that M has the same distribution as $\sum_{i=1}^{A}(N_0^i + 1)$, and so $M - A$ has the same distribution as $\sum_{i=1}^{A} N_0^i$. We could prove (b) using the fact that $M(t)$ is distributed as the number of trials before the A^{th} success in in a sequence of independent Bernoulli trials with common success probability e^{-t}, but it is more straightforward to use a simple union bound.

Observe that from part(a) and the inequality $1-x \leq \exp(-x)$ that, for any $C \geq 0$,

$$\begin{aligned}
\mathbf{P}\{N_0(t) > Ce^t\} &= (1 - e^{-t})^{\lfloor Ce^t \rfloor + 1} \\
&\leq (\exp(-e^{-t}))^{\lfloor Ce^t \rfloor + 1} \\
&= \exp(-e^{-t}(\lfloor Ce^t \rfloor + 1)) \\
&\leq e^{-C},
\end{aligned}$$

and hence,

$$\begin{aligned}
\mathbf{P}\{N_a(t) > Ke^t\} &\leq \sum_{i=1}^{A} \mathbf{P}\{N_0^i(t) > \frac{K}{A}e^t\} \\
&\leq Ae^{-\frac{K}{A}}.
\end{aligned}$$

□

Theorem 2.6.2. *(a) There exists a random variable $Z_a > 0$ such that*

$$\lim_{t\to\infty} \frac{\#\mathcal{F}(t)}{e^{\gamma_a t}} = Z_a \quad \textit{almost surely.}$$

(b) There exists a constant C such that $\mathbb{E}[\#\mathcal{F}(t)] \le Ce^{\gamma_a t}$.
(c) For the random variable Z_a of part (a),

$$\lim_{n\to\infty} T_n - \frac{1}{\gamma_a}\log n = -\log Z_a \quad \textit{almost surely.}$$

(d) There exists a random variable $W_a > 0$ such that

$$\lim_{t\to\infty} \frac{N_a(t)}{e^t} = W_a \quad \textit{almost surely.}$$

Proof. Parts (a) and (b) (in a more general context) can be found in [23], so we shall not give the proof here. They essentially follow from the general theory of continuous time branching processes developed by Jagers and Nerman.

Part (c) follows immediately from part (a) and the relationship $\#\mathcal{F}(T_n) = n$.

Turning to part (d), note that

$$N_a(t) - \int_0^t (N_a(s) + 1 + a)\, dt, \quad t \ge 0,$$

is a local martingale with bounded variation. Stochastic calculus shows that the process $(e^{-t} \cdot (N_a(t) + 1 + a))_{t\ge 0}$ is also a local martingale with bounded variation. The fact that the latter process is bounded in $\mathbb{L}^2$ and hence, in particular, a true martingale follows from Lemma 2.6.1(b).

It follows from the martingale convergence theorem that $e^{-t}N_a(t)$ converges almost surely and in $\mathbb{L}^2$ to a random variable W_a.

It remains to show that W_a is strictly positive almost surely. Consider first the case $a \ge 0$. From a comparison of branching rates similar to that in the proof of Lemma 2.6.1, it is possible to couple N_a and N_0 so that $N_0(t) \le N_a(t)$ for all $t \ge 0$. Note that W_0 has an exponential distribution with mean 1, and so W_a is certainly almost surely positive.

Consider now the case $-1 < a < 0$. Let $\tilde{N}_a$ be N_a started in the initial state 1 rather than 0, and put $\hat{N}_a = \tilde{N}_a - 1$. Another comparison of branching rates shows that it is possible to couple $\hat{N}_a$ and N_0 so that $N_0(t) \le \hat{N}_a(t)$ for all $t \ge 0$. Thus, $\lim_{t\to\infty} e^{-t}N_a(t)$ is stochastically greater than the strictly positive random variable $e^{-\tau}W_0$, where the random variable τ is independent of W_0 and has the same distribution as the time taken for N_a to go from 0 to 1 (that is, τ has an exponential distribution with rate $1 + a$). □

Fix $k \geq 1$. Recall that $\Delta_{n,1} \geq \Delta_{n,2} \geq \cdots \geq \Delta_{n,k}$ are the k largest out-degrees in $\mathcal{T}_n$ (the out-degree of a vertex is its number of children). We will show that the vertices with these out-degrees occur in a *finite* neighborhood about the root and that the out-degrees of vertices in a finite neighborhood about the root converge in distribution when properly normalized.

Lemma 2.6.3. *In the branching process construction of the linear preferential attachment tree $\mathcal{T}_n$, let $\Delta_{n,1}^S$ denote the maximum out-degree in $\mathcal{T}_n$ among all vertices born before time S. Given any $\varepsilon > 0$, there exists a finite constant S_ε such that*

$$\liminf_{n\to\infty} \mathbf{P}\{\Delta_{n,1}^{S_\varepsilon} = \Delta_{n,1}\} \geq 1 - \varepsilon.$$

Proof. Fix $\varepsilon > 0$. Write $A_{n,S}$ for the event

> *There exists a vertex in $\mathcal{T}_n$ which was born after time S and has out-degree greater than the root.*

The claim of the lemma may be rephrased as a statement that there is a finite constant S_ε such that

$$\limsup_{n\to\infty} \mathbf{P}\{A_{n,S_\varepsilon}\} \leq \varepsilon.$$

Note from Theorem 2.6.2(c) that there is a constant B_ε such that

$$\limsup_{n\to\infty} \mathbf{P}\left\{\left|T_n - \frac{1}{\gamma_a}\log n\right| > B_\varepsilon\right\} \leq \varepsilon/2.$$

Set $t_-^n := \frac{1}{\gamma_a}\log n - B_\varepsilon$ and $t_+^n := \frac{1}{\gamma_a}\log n + B_\varepsilon$. It is enough to prove that there exists a finite constant S_ε such that

$$\limsup_{n\to\infty} \mathbf{P}\{A'_{n,S_\varepsilon}\} \leq \varepsilon/2,$$

where $A'_{n,S}$ is the event

> *There exists a vertex born after time S that has out-degree greater than the root for some time t in the interval $[t_-^n, t_+^n]$.*

Furthermore, since the out-degrees of vertices increase with time, it is enough to show that

$$\limsup_{n\to\infty} \mathbf{P}\{A''_{n,S_\varepsilon}\} \leq \varepsilon/2,$$

where $A''_{n,S}$ is the event

> *There exists a vertex born after time after time S such that the out-degree of the vertex at time t_+^n is greater than the out-degree of the root at time t_-^n.*

For $t \geq 0$ and a time interval $I \subseteq [0,t]$ denote by $Z(I,t)$ the maximum out-degree at time t of all vertices born in the time interval I. Let $\zeta_\rho(t)$ denote the out-degree of the root at time t. Note that

$$A''_{n,S} = \{Z([S,t^n_+],t^n_+) > \zeta_\rho(t^n_-)\}.$$

Observe also, that for any constant K,

$$\begin{aligned}\mathbf{P}\{A''_{n,S}\} &\leq \mathbf{P}\left\{\zeta_\rho(t^n_-) \leq K \text{ or } Z([S,t^n_+],t^n_+) > K\right\}\\ &\leq \mathbf{P}\{\zeta_\rho(t^n_-) \leq K\} + \mathbf{P}\{Z([S_\varepsilon,t^n_+],t^n_+) > K\}.\end{aligned}$$

It thus suffices to show that there is a sequence K_n and a constant S_ε such that

$$\limsup_{n\to\infty} \mathbf{P}\{\zeta_\rho(t^n_-) \leq K_n\} \leq \varepsilon/4 \tag{2.4}$$

and

$$\limsup_{n\to\infty} \mathbf{P}\{Z([S,t^n_+],t^n_+) > K_n\} \leq \varepsilon/4. \tag{2.5}$$

It follows from Theorem 2.6.2(d) that the inequality (2.4) holds with $K_n = K_\varepsilon n^{1/\gamma_a}$ for a suitable constant $K_\varepsilon > 0$.

Turning to the inequality (2.5), assume without loss of generality that S and t^n_+ are integers. In that case,

$$Z(S,t^n_+) = \max_{S \leq m \leq t^n_+ - 1} Z([m,m+1],t^n_+).$$

Note that, by the union bound,

$$\mathbf{P}\{Z([m,m+1],t^n_+) > K_n\} \leq \mathbb{E}[\#\mathcal{F}(m+1)]\ \mathbf{P}\{N_a(t^n_+ - m) > K_n\}.$$

Applying Theorem 2.6.2(b) and Lemma 2.6.1(b) gives

$$\mathbf{P}\{Z(S,t^n_+) > K_n\} \leq \sum_{m=S}^{t^n_+-1} Ce^{\gamma_a(m+1)} Ae^{-C'e^m},$$

where $C' = K_\varepsilon/(Ae^{B_\varepsilon})$. The inequality (2.5) follows upon choosing $S = S_\varepsilon$ large enough. □

A slightly more detailed analysis shows that Lemma 2.6.3 can be generalized to the k maximal out-degrees for any fixed k. Let $\Delta^S_{n,1} \geq \Delta^S_{n,2} \geq \cdots \geq \Delta^S_{n,k}$ be the k largest out-degrees in $\mathcal{T}_n$ from among the vertices that are born before time S, with the convention that $\Delta^{S,n}_i = 0$ for $i \geq \#\mathcal{F}(S)$ when $\#\mathcal{F}(S) < k$. We leave the proof of the following result to the reader.

Lemma 2.6.4. *For any $\varepsilon > 0$ there exists a finite constant S_ε such that*

$$\liminf_{n\to\infty} \mathbf{P}\{(\Delta_{n,1}^{S_\varepsilon}, \ldots, \Delta_{n,k}^{S_\varepsilon}) = (\Delta_{n,1}, \ldots, \Delta_{n,k})\} \geq 1 - \varepsilon.$$

Proposition 2.6.5. *Fix $S > 0$ and consider the marked tree $\mathcal{T}_n^{\#}$ constructed by marking each vertex v of the tree $\mathcal{F}(S)$ with $n^{-1/\gamma_a} D(v,n) \in \mathbb{R}_+$, where $D(v,n)$ is the out-degree of v in $\mathcal{T}_n$. Then, $\mathcal{T}_n^{\#}$ converges almost surely as $n \to \infty$ to the tree $\mathcal{F}(S)$ equipped with a set of marks that are strictly positive almost surely.*

Proof. For any vertex $v \in \mathcal{F}(S)$, write $\zeta_v(t)$ for the out-degree (that is, number of offspring) of v at time t, so that $\zeta_v(T_n) = D(v,n)$. Note that $\zeta_v(S)$ can be computed by only looking at $\mathcal{F}(S)$ (recall that our trees are rooted). Conditional on $\mathcal{F}(S)$, the processes $\hat{\zeta}_v := (\zeta_v(S+t) - \zeta_v(S))_{t\geq 0}$, $v \in \mathcal{F}(S)$, are conditionally independent. Note that the conditional distribution of $\hat{\zeta}_v$ is that of a pure birth process that starts in state 0 and has birth rate $\zeta_v(S) + \ell + 1 + a$ in state ℓ. It follows from Theorem 2.6.2(d) that $e^{-t}\hat{\zeta}_v(t)$ converges almost surely as $t \to \infty$ to a random variable that has a conditional distribution which is that of the strictly positive random variable $W_{a+\zeta_v(S)}$. Hence, by Theorem 2.6.2(c),

$$\lim_{n\to\infty} e^{-(T_n - S)} \hat{\zeta}_v(T_n - S) = \lim_{n\to\infty} e^S Z_a n^{-1/\gamma_a} \hat{\zeta}_v(T_n - S)$$

exists almost surely and the limit is strictly positive almost surely.

The result follows because $n^{-1/\gamma_a} D(v,n) = n^{-1/\gamma_a}(\zeta_v(S) + \hat{\zeta}_v(T_n - S))$.

□

Corollary 2.6.6. *The random vector $n^{-1/\gamma_a}(\Delta_{n,1}^S, \ldots, \Delta_{n,k}^S)$ converges almost surely to a random vector $(Y_1^S, Y_2^S, \ldots, Y_k^S)$ as $n \to \infty$, where $Y_1^S \geq Y_2^S \geq \ldots \geq Y_k^S > 0$ almost surely.*

Completion of the proof of Theorem 2.3.8. Given Corollary 2.6.6 and Lemma 2.6.4, the proof is completed by applying the following elementary result with $X_{n,i} = n^{-1/\gamma_a}\Delta_{n,i}$ and $Y_{n,i}^\varepsilon = n^{-1/\gamma_a}\Delta_{n,i}^{S_\varepsilon}$.

Lemma 2.6.7. *Let $(\mathbf{X}_n)_{n=1}^\infty = ((X_{n,1}, \ldots, X_{n,k}))_{n=1}^\infty$ be a sequence of $\mathbb{R}^k$-valued random variables. Suppose for each fixed $\varepsilon > 0$ that there exists a sequence of $\mathbb{R}^k$-valued random variables $(\mathbf{Y}_n^{(\varepsilon)})_{n=1}^\infty = ((Y_{n,1}^\varepsilon, \ldots, Y_{n,k}^\varepsilon))_{n=1}^\infty$ on the same probability space such that*

$$\liminf_{n\to\infty} \mathbf{P}\{\mathbf{X}_n = \mathbf{Y}_n^{(\varepsilon)}\} \geq 1 - \varepsilon.$$

Suppose further that for each $\varepsilon > 0$ there exists a random vector $\mathbf{Y}_\infty^{(\varepsilon)}$ such that $\mathbf{Y}_n^{(\varepsilon)}$ converges in probability to $\mathbf{Y}_\infty^{(\varepsilon)}$ as $n \to \infty$. Then, there exists an $\mathbb{R}^k$-valued random variable $\mathbf{X}_\infty$ such that $\mathbf{X}_n$ converges in probability to $\mathbf{X}_\infty$ as $n \to \infty$.

Proof. Convergence in probability for the space of $\mathbb{R}^k$-valued random variables is metrized by the metric r, where $r(\mathbf{X}, \mathbf{Y}) := \mathbb{E}[|\mathbf{X} - \mathbf{Y}| \wedge 1]$ and where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^k$. Moreover, this metric space is complete. By assumption,

$$\begin{aligned}\limsup_{m,n\to\infty} r(\mathbf{X}_m, \mathbf{X}_n) &\le \limsup_{m\to\infty} r(\mathbf{X}_m, \mathbf{Y}_m^{(\varepsilon)}) \\ &\quad + \limsup_{m,n\to\infty} r(\mathbf{Y}_m^{(\varepsilon)}, \mathbf{Y}_n^{(\varepsilon)}) + \limsup_{n\to\infty} r(\mathbf{Y}_n^{(\varepsilon)}, \mathbf{X}_n) \\ &\le 2 \limsup_{n\to\infty} \mathbf{P}\{\mathbf{Y}_n^{\varepsilon} \neq \mathbf{X}_n\} \\ &\le 2\varepsilon.\end{aligned}$$

The sequence $(\mathbf{X}_n)_{n=1}^{\infty}$ is thus Cauchy in the metric r, and hence it converges in probability to a limit $\mathbf{X}_\infty$. □

2.7 An example

Consider the following construction of a random finite graph $\mathcal{G}_n$ for a given positive integer n. The vertices of $\mathcal{G}_n$ are $\{1, \ldots, 2n\}$. Let $\varepsilon_{n1}, \ldots, \varepsilon_{nn}$ be independent, identically distributed random variables, with $\mathbf{P}\{\varepsilon_{nk} = 0\} = q$ and $\mathbf{P}\{\varepsilon_{nk} = 1\} = 1 - q$, where $0 < q < 1$. There is always an edge between the vertices $2k - 1$ and $2k + 1$ for $1 \le k \le n - 1$, and there is an edge between the vertices $2k - 1$ and $2k$ for $1 \le k \le n$ if and only if $\varepsilon_{nk} = 1$. There are no other edges.

The graph $\mathcal{G}_n$ consists of a large connected component $\mathcal{T}_n$ with vertices

$$\{1, 3, \ldots, 2n - 1\} \cup \{2k : 1 \le k \le n,\ \varepsilon_{nk} = 1\}$$

plus the (possibly empty) set of isolated points $\{2k : 1 \le k \le n,\ \varepsilon_{nk} = 0\}$. The graph $\mathcal{T}_n$ is a tree. It resembles a comb with some of the teeth missing, see Figure 2.6.

No matter how we chose a root, the sequence of random finite trees $(\mathcal{T}_n)_{n=1}^{\infty}$ would not converge in the probability fringe sense because the cardinality of the subtree below a uniformly chosen point would not converge in distribution to a finite random variable. However, if we look at the empirical distribution of the tree within graph distance k of a vertex v as v ranges over the $\mathcal{T}_n$, then it is not hard to see that this random measure converges to a deterministic limit for every k. This observation and the fact that the vertices of $\mathcal{T}_n$ have degree at most 3 shows that the moments of the spectral distribution of $\mathcal{T}_n$ converge as $n \to \infty$ to finite constants, and these constants are the moments of a distribution that is determined by its moments. Therefore, the spectral distribution of $\mathcal{T}_n$ converges in distribution to a deterministic limit as $n \to \infty$, and in order to compute the moments of that limit, it suffices to compute the limits of the expectations of the moments of the spectral distribution of $\mathcal{T}_n$

Write $Z_n = \#\{1 \le k \le n : \varepsilon_{nk} = 0\}$. By permuting indices, it is possible to

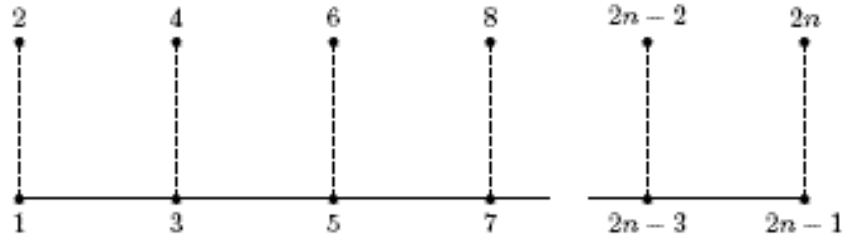

Figure 2.6: Comb tree: In the above graph $\mathcal{G}_n$, every horizontal edge always occurs whereas every vertical (dotted) edge occurs independently with probability $1-q$. The comb tree $\mathcal{T}_n$ is the unique connected component of $\mathcal{G}_n$ that is not an isolated point.

re-write the adjacency matrix of $\mathcal{G}_n$ in block form, where the upper-left block has dimensions $(2n - Z_n) \times (2n - Z_n)$ and is the adjacency matrix of $\mathcal{T}_n$, while the lower-right block is the zero matrix of dimensions $Z_n \times Z_n$. Therefore, if we write F_n (respectively, H_n) for the empirical distribution of the eigenvalues of the adjacency matrix of $\mathcal{T}_n$ (respectively, $\mathcal{G}_n$), then

$$H_n = \frac{Z_n}{2n}\delta_0 + \left(1 - \frac{Z_n}{2n}\right) F_n,$$

where δ_0 is the unit point mass at 0. Since Z_n/n converges in probability to q as $n \to \infty$, the limiting behavior of F_n is determined by that of H_n and *vice-versa*: H_n converges in probability to a non-random probability measure H and

$$H = \frac{q}{2}\delta_0 + \left(1 - \frac{q}{2}\right) F,$$

where the probability measure F is the limit of the sequence F_n.

Define a random infinite graph $\mathcal{G}$ with vertex set $\mathbb{Z}$ as follows. Let ε_k, $k \in \mathbb{Z}$, be independent, identically distributed random variables, with $\mathbf{P}\{\varepsilon_k = 0\} = q$ and $\mathbf{P}\{\varepsilon_k = 1\} = 1 - q$. There is an edge between the vertices $2k - 1$ and $2k + 1$ for all $k \in \mathbb{Z}$, and there is an edge between the vertices $2k - 1$ and $2k$ for $k \in \mathbb{Z}$ if and only if $\varepsilon_k = 1$. There are no other edges.

Let B_n (resp. B) denote the adjacency matrix of $\mathcal{G}_n$ (resp. $\mathcal{G}$). For each non-

negative integer m,

$$\begin{aligned}
&\int x^m \, H(dx) \\
&\quad = \lim_{n\to\infty} \int x^m \, H_n(dx) \\
&\quad = \lim_{n\to\infty} \mathbb{E}\left[\int x^m \, H_n(dx)\right] \\
&\quad = \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{2n}\operatorname{tr} B_n^m\right] \\
&\quad = \mathbb{E}\left[\frac{1}{2}(B^m)_{11} + \frac{1}{2}(B^m)_{22}\right]
\end{aligned}$$

(because $B_{ij} = 0$ for $|i-j| > 2$, there is no problem defining B^m).

Divide the matrix B into 2×2 blocks with the rows (resp. columns) of the $(i,j)^{\text{th}}$ block indexed by $\{2i+1, 2i+2\}$ (resp. $\{2j+1, 2j+2\}$). This block form matrix is block tridiagonal. The entries in the diagonals above and below the main diagonal are always the matrix

$$\Pi_x := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

We use this notation because we can think of Π_x as the matrix for the orthogonal projection onto the x-axis in a two-dimensional (x,y) coordinate system. The entry in the (k,k) diagonal block is the matrix

$$\begin{pmatrix} 0 & \varepsilon_{k+1} \\ \varepsilon_{k+1} & 0 \end{pmatrix}.$$

If $\varepsilon_{k+1} = 1$, then this is the matrix

$$\Sigma := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

that permutes the two coordinates. Otherwise, it is is the 2×2 zero matrix. By analogy with the definition of Π_x, set

$$\Pi_y := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

We have the following relations between these matrices

$$\begin{aligned}
\Pi_x^2 &= \Pi_x \\
\Pi_y^2 &= \Pi_y \\
\Pi_x \Pi_y &= \Pi_y \Pi_x = 0 \\
\Sigma^2 &= I \\
\Sigma \Pi_x &= \Pi_y \Sigma \\
\Sigma \Pi_y &= \Pi_x \Sigma \\
\operatorname{tr} \Pi_x &= \operatorname{tr} \Pi_y = 1 \\
\operatorname{tr} \Sigma \Pi_x &= \operatorname{tr} \Pi_x \Sigma = 0 \\
\operatorname{tr} \Sigma \Pi_y &= \operatorname{tr} \Pi_y \Sigma = 0.
\end{aligned}$$

A consequence of these relations is the following.

Lemma 2.7.1. *For $a \geq 0$,*

$$\operatorname{tr} \Sigma^a = \begin{cases} 2, & \text{if } a \text{ is even,} \\ 0, & \text{if } a \text{ is odd.} \end{cases}$$

For any $r \geq 1$,

$$\operatorname{tr} \left(\Sigma^{a_1} \Pi_x \Sigma^{a_2} \Pi_x \cdots \Sigma^{a_r} \Pi_x \Sigma^{a_{r+1}}\right) = \begin{cases} 1, & \text{if } a_1 + a_{r+1}, a_2, \ldots, a_r \text{ are all even,} \\ 0, & \text{otherwise.} \end{cases}$$

Proof. The first claim is obvious, because Σ^a either the 2×2 identity matrix or Σ depending on whether a is even or odd.

For the second claim, first observe that the product in question may be rewritten as

$$\Sigma^{b_1} \Pi_x \Sigma^{b_2} \Pi_x \cdots \Sigma^{b_r} \Pi_x \Sigma^{b_{r+1}},$$

where b_ℓ is 0 or 1 depending on whether a_ℓ is even or odd. This in turn may be rewritten as

$$\Pi_{z_1} \Pi_{z_2} \cdots \Pi_{z_r} \Sigma^c,$$

where z_ℓ, $1 \leq \ell \leq r$, is x or y depending on whether $b_1 + \cdots + b_\ell$ is even or odd, and c is 0 or 1 depending on whether $b_1 + \cdots + b_{r+1}$ is even or odd.

The product is non-zero if and only if $z_1 = z_2 = \ldots = z_r$. This is equivalent to either $b_1 = 0$ and $b_2 = \ldots = b_r = 0$, in which case the product is $(\Pi_x)^r \Sigma^c = \Pi_x \Sigma^c$, or $b_1 = 1$ and $b_2 = \ldots = b_r = 0$, in which case the product is $(\Pi_y)^r \Sigma^c = \Pi_y \Sigma^c$.

Furthermore, even if the product is non-zero, and hence of the form $\Pi_x \Sigma^c$ or $\Pi_y \Sigma^c$, the trace is zero if $c = 1$. Otherwise, the trace is 1.

Thus, the trace is zero unless $b_2 = \ldots = b_r = 0$ and $b_1 + \cdots + b_{r+1}$ is even, in

which case the trace is 1. This condition is equivalent to $a_2, \ldots, a_r$ and $a_1 + \cdots + a_{r+1}$ all being even, and the result follows. □

In order to carry out the combinatorial calculations involved in computing the expected value of $(B^m)_{11} + (B^m)_{22}$, it will be helpful to recast them as computations involving a certain killed random walk on a random subinterval of the integers.

Let $U_1, U_2, \ldots$ be independent, identically distributed random variables that are independent of the random infinite matrix B and have common distribution

$$\mathbf{P}\{U_k = -1\} = \mathbf{P}\{U_k = 0\} = \mathbf{P}\{U_k = +1\} = \frac{1}{3}.$$

Set $W_m = U_1 + \cdots + U_m$ (with $W_0 = 0$). Note that

$$\begin{aligned}\mathbb{E}&\left[\begin{pmatrix}(B^m)_{11} & (B^m)_{12}\\ (B^m)_{21} & (B^m)_{22}\end{pmatrix}\right]\\ &= 3^m \mathbb{E}\left[\Xi_1\Xi_2\cdots\Xi_m \mathbf{1}\{W_m = 0 \ \&\ \alpha < W_\ell < \beta,\, 1 \le \ell \le m\}\right],\end{aligned}$$

where

$$\begin{aligned}\alpha &:= \max\{k \le 0 : \varepsilon_{k+1} = 0\},\\ \beta &:= \min\{k \ge 0 : \varepsilon_{k+1} = 0\}\end{aligned}$$

and

$$\Xi_\ell := \begin{cases}\Pi_x, & \text{if } U_\ell \in \{-1, +1\},\\ \Sigma, & \text{if } U_\ell = 0.\end{cases}$$

Lemma 2.7.2. *The random variables α and β satisfy*

$$\mathbf{P}\{\beta - \alpha = N\} = \begin{cases}q, & \text{if } N = 0,\\ 0, & \text{if } N = 1,\\ (N-1)(1-q)^{N-1}q^2, & \text{if } N \ge 2,\end{cases}$$

and

$$\mathbf{P}\{0 - \alpha = j \,|\, \beta - \alpha = N\} = \frac{1}{N-1}, \quad N \ge 2,\ 1 \le j \le N-1.$$

Proof. It is clear that

$$\mathbf{P}\{\beta - \alpha = 0\} = \mathbf{P}\{\varepsilon_1 = 0\} = q$$

and $\mathbf{P}\{\beta - \alpha = 1\} = 0$.

For $N \geq 2$ and $1 \leq j \leq N-1$,

$$\begin{aligned}
&\mathbf{P}\{\beta - \alpha = N,\ 0 - \alpha = j\} \\
&\quad = \mathbf{P}\{\varepsilon_{-j+1} = 0,\ \varepsilon_{-j+2} = \ldots = \varepsilon_{-j+N} = 1,\ \varepsilon_{-j+N+1} = 0\} \\
&\quad = (1-q)^{N-1}q^2 \\
&\quad = (N-1)(1-q)^{N-1}q^2 \times \frac{1}{N-1},
\end{aligned}$$

and the remainder of the result follows. □

Consider the moment generating function $\varphi(t) := \sum_{m=0}^{\infty} \frac{t^m}{m!} \int x^m \, H(dx)$. From the above, $\varphi(t) = \varphi(-t)$ for all $t \in \mathbb{R}$ and

$$\exp(-3t)\varphi(t) = \frac{1}{2}\mathbb{E}\left[\mathrm{tr}\,(\Xi_1\Xi_2\cdots\Xi_M)\mathbf{1}\{W_M = 0 \ \&\ \alpha < W_\ell < \beta,\ 1 \leq \ell \leq M\}\right]$$

for $t \geq 0$, where M is a Poisson distributed random variable with expected value $3t$ that is independent of the random infinite matrix B and the increments $U_1, U_2, \ldots$.

Put $\mathcal{I}' := \{1 \leq \ell \leq M : U_\ell = \pm 1\}$ and $\mathcal{I}'' := \{1 \leq \ell \leq M : U_\ell = 0\}$. Then $M' := \#\mathcal{I}'$ and $M'' := \#\mathcal{I}''$ are independent Poisson distributed random variables with expected values $2t$ and t, respectively. Write $\mathcal{I}' = \{\ell_1, \ell_2, \ldots, \ell_{M'}\}$, with $\ell_1 < \ell_2 < \ldots < \ell_{M'}$. The set $\mathcal{I}''$ is the disjoint union of sets $\mathcal{I}''_1, \mathcal{I}''_2, \ldots, \mathcal{I}''_{M'+1}$, where

$$\mathcal{I}''_1 := \{1 \leq \ell \leq M : \ell < \ell_1\},$$

$$\mathcal{I}''_j := \{1 \leq \ell \leq M : \ell_{j-1} < \ell < \ell_j\}, \quad 1 \leq j \leq M',$$

and

$$\mathcal{I}''_{M'+1} := \{1 \leq \ell \leq M : \ell_{M'} < \ell\}.$$

Conditional on $M' = m'$, the random variables $\#\mathcal{I}''_1, \#\mathcal{I}''_2, \ldots, \#\mathcal{I}''_{m'+1}$ are independent and identically distributed with a common distribution that is Poisson with expected value $t/(m'+1)$.

In view of Lemma 2.7.1, the following well-known result will be useful.

Lemma 2.7.3. *The probability that a Poisson random variable with expected value λ takes an even value is $\frac{1}{2}[1 + \exp(-2\lambda)]$.*

Proof. The probability in question is

$$\sum_{k=0}^{\infty} \exp(-\lambda)\frac{\lambda^{2k}}{(2k)!} = \exp(-\lambda)\frac{1}{2}[\exp(\lambda) + \exp(-\lambda)] = \frac{1}{2}[1 + \exp(-2\lambda)].$$

□

The conditional distribution of the sequence $0 = W_0, W_{\ell_1}, \ldots, W_{\ell_{m'}}$ given $M' = m'$ is that of a symmetric simple random walk with m' steps, and we need the following result.

Lemma 2.7.4. *Fix an integer $N \geq 2$. Suppose that $S_0, S_1, S_2, \ldots$ is a symmetric simple random walk. Then*

$$\sum_{j=1}^{N-1} \mathbf{P}\{S_n = j \ \& \ 0 < S_\ell < N,\, 1 \leq \ell \leq n \mid S_0 = j\} = \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N}\right)^n$$

for $n = 0, 1, 2, \ldots$. This quantity is zero when n is odd.

Proof. Consider the transition matrix P of a symmetric simple random walk killed on exiting $\{1, 2, \ldots, N-1\}$. That is, P is the $(N-1) \times (N-1)$ symmetric matrix with rows and columns both indexed by $\{1, 2, \ldots, N-1\}$ and entries

$$P_{ij} = \begin{cases} \frac{1}{2}, & \text{if } |i-j| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

It is a straightforward (and well-known) calculation that the eigenvalues of P are $\cos \frac{k\pi}{N}$, $1 \leq k \leq N-1$, and the corresponding normalized eigenvectors are $\sqrt{\frac{2}{N}}(\sin \frac{k\pi}{N} j)_{j=1}^{N-1}$. Hence,

$$(P^n)_{ij} = \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N}\right)^n \frac{2}{N} \left(\sin \frac{k\pi}{N} i\right) \left(\sin \frac{k\pi}{N} j\right).$$

Thus,

$$\begin{aligned} \sum_{j=1}^{N-1} (P^n)_{jj} &= \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N}\right)^n \frac{2}{N} \sum_{j=1}^{N-1} \left(\sin \frac{k\pi}{N} j\right)^2 \\ &= \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N}\right)^n, \end{aligned}$$

as claimed. □

Observe from Lemma 2.7.1 that for $t \geq 0$ the moment generating function φ of H

satisfies

$$\begin{aligned}\exp(-3t)\varphi(t) &= \frac{1}{2}\mathbb{E}\left[\mathrm{tr}\left(\Xi_1\Xi_2\cdots\Xi_M\right)\mathbf{1}\{W_M = 0 \ \&\ \alpha < W_\ell < \beta,\, 1 \le \ell \le M\}\right] \\ &= \frac{1}{2}\mathbb{E}\Big[\mathrm{tr}\,\Big(\Sigma^{\#\mathcal{I}_1''}\Pi_x\Sigma^{\#\mathcal{I}_2''}\cdots\Sigma^{\#\mathcal{I}_{M'}''}\Pi_x\Sigma^{\#\mathcal{I}_{M'+1}''}\Big) \\ &\quad \times \mathbf{1}\{W_M = 0 \ \&\ \alpha < W_\ell < \beta,\, 1 \le \ell \le M\}\Big] \\ &= \frac{1}{2}\Big(2 \times \mathbf{P}\left\{M' = 0,\, M'' \text{ is even}, \ \&\ \alpha < 0 < \beta\right\} \\ &\quad + 1 \times \mathbf{P}\big\{M' > 0,\, \#\mathcal{I}_1'' + \#\mathcal{I}_{M'+1}'', \#\mathcal{I}_2'', \ldots, \#\mathcal{I}_{M'}'' \text{ are all even}, \\ &\qquad S_{M'} = 0, \ \&\ \alpha < S_k < \beta,\, 1 \le k \le M'\big\}\Big)\end{aligned}$$

where $S_k = W_{\ell_k}$ for $1 \le k \le M'$.

Conditioning on M' taking the even value $2h$ and $\beta - \alpha$ taking the value N, we hence, from Lemmas 2.7.3 and 2.7.4,

$$\begin{aligned}\exp(-3t)\varphi(t) &= \exp(-2t)\frac{1}{2}\left[1 + \exp(-2t)\right](1-q) \\ &\quad + \frac{1}{2}\sum_{h=1}^{\infty}\exp(-2t)\frac{(2t)^{2h}}{2h!} \\ &\quad \times \left\{\frac{1}{2}\left[1 + \exp\left(-\frac{4t}{2h+1}\right)\right]\right\}\left\{\frac{1}{2}\left[1 + \exp\left(-\frac{2t}{2h+1}\right)\right]\right\}^{2h-1} \\ &\quad \times \sum_{N=2}^{\infty}(N-1)(1-q)^{N-1}q^2\frac{1}{N-1}\sum_{k=1}^{N-1}\left(\cos\frac{k\pi}{N}\right)^{2h}\end{aligned}$$

for $t \ge 0$. Thus,

$$\begin{aligned}\varphi(t) &= (1-q)\cosh t \\ &\quad + \frac{1}{2}\sum_{h=1}^{\infty}\frac{(2t)^{2h}}{2h!} \\ &\quad \times \cosh\left(-\frac{2t}{2h+1}\right)\left\{\cosh\left(-\frac{t}{2h+1}\right)\right\}^{2h-1} \\ &\quad \times \sum_{N=2}^{\infty}(1-q)^{N-1}q^2\sum_{k=1}^{N-1}\left(\cos\frac{k\pi}{N}\right)^{2h}.\end{aligned}$$

for $t \ge 0$. The right-hand side thought of as a function of $t \in \mathbb{R}$ is an even function, and it follows that the right-hand side is the correct expression for $\varphi(t)$ for all $t \in \mathbb{R}$.

Finally, if we write ψ for the moment generating function of F, then

$$\psi(t) = \frac{\varphi(t) - \frac{q}{2}}{1 - \frac{q}{2}}.$$

Chapter 3

Coalescing Systems of Non-Brownian Particles

3.1 Sketch of the argument

We give a sketch of the argument why there are only finitely many particles left at any positive time, with probability one, after starting from countably many coalescing particles for the simplest case of stable processes on the circle in order to motivate some of the estimates that we must develop before we can present proofs in each of the various settings.

Observe from the scaling property of stable processes on the line that if X' and X'' are independent copies of a stable processes of index $\alpha > 1$ on the unit circle $\mathbb{T}$ starting from two distinct points that are $2\pi\varepsilon$ apart, then there exist positive constants β and p such that the probability the processes will collide by time $\beta\varepsilon^{\alpha}$ is bounded below by p.

Suppose we start with $n+1$ stable particles in some configuration on $\mathbb{T}$. By the pigeonhole principle, there will be (at least) one pair of particles that are distance at most $2\pi/n$ apart. Therefore, with probability at least p, these two particles in isolation would collide with each other by time $\beta n^{-\alpha}$. Hence, in the coalescing system the probability that there is at least one collision between some pair of particles within the time interval $[0, \beta n^{-\alpha}]$ is certainly at least p (either the two distinguished particles collide with each other and no others or some other particle(s) collides with one or both of the distinguished particles). Moreover, if there is no collision between any pair of particles after time $\beta n^{-\alpha}$, then we can again find at time $\beta n^{-\alpha}$ a possibly different pair of particles that are within distance $2\pi/n$ from each other, and the probability that this pair of particles will collide within the time interval $[\beta n^{-\alpha}, 2\beta n^{-\alpha}]$ is again at least p. By repeating this argument and using the Markov property, we see that if we let τ_n^{n+1} be the first time there are n surviving particles starting from $(n+1)$

particles, then, regardless of the particular initial configuration of the $n+1$ particles,

$$\mathbf{P}\left\{\tau_n^{n+1} \geq k\beta n^{-\alpha}\right\} \leq (1-p)^k.$$

In particular, the expected time needed to reduce the number of particles from $n+1$ to n is bounded above by $Cn^{-\alpha}$ for a suitable constant C.

Thus, if we start with N particles somewhere on $\mathbb{T}$, then the probability that after some positive time t the number of particles remaining is greater than m is, by Markov's inequality, bounded above by

$$\frac{1}{t}\sum_{n=m}^{N-1}\mathbb{E}\left[\tau_n^{n+1}\right] \leq \frac{C}{t}\sum_{n=m}^{N-1} n^{-\alpha} \leq \frac{C'}{t}m^{1-\alpha}$$

for some constant C'. Letting $N \to \infty$ and then letting $m \to \infty$, we conclude that by time t there are only finitely many particles almost surely.

The above reasoning uses the compactness of $\mathbb{T}$ in a crucial way and it cannot be applied as it stands to deal with, say, coalescing stable processes on the real line. The primary difficulty is that the argument bounds the time to coalesce from some number of particles to a smaller number by considering a particular sequence of coalescent events, and while waiting for such an event to occur the particles might spread out to such an extent that the pigeon hole argument can no longer be applied. We overcome this problem by using a somewhat more sophisticated pigeonhole argument to assign the bulk of the particles to a collection of suitable disjoint pairs (rather than just selecting a single suitable pair) and then employing a simple large deviation bound to ensure that with high probability at least a certain fixed proportion of the pairs will have collided over an appropriate time interval.

3.2 Countable systems of coalescing Feller processes

In this section we develop some general properties of coalescing systems of Markov processes that we will apply later to Brownian motions on the Sierpinski gasket and stable processes on the line or circle.

3.2.1 Vector-valued coalescing process

Fix $N \in \mathbb{N} \cup \{\infty\}$, where, as usual, $\mathbb{N}$ is the set of positive integers. Write $[N]$ for the set $\{1, 2, \ldots, N\}$ when N is finite and for the set $\mathbb{N}$ when $N = \infty$.

Fix a locally compact, second-countable, Hausdorff space E. Note that E is metrizable. Let d be a metric giving the topology on E. Denote by $D := D(\mathbb{R}_+, E)$ the usual Skorokhod space of E-valued càdlàg paths. Fix a bijection $\sigma : N \to N$. We will call σ a *ranking* of $[N]$. Define a mapping $\Lambda_\sigma : D^N \to D^N$ by setting $\Lambda_\sigma \boldsymbol{\xi} = \boldsymbol{\zeta}$

for $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots) \in D^N$, where $\boldsymbol{\zeta}$ is defined inductively as follows. Set $\zeta_{\sigma(1)} \equiv \xi_{\sigma(1)}$. For $i > 1$, set

$$\tau_i := \inf\Big\{t \geq 0 : \xi_{\sigma(i)}(t) \in \{\zeta_{\sigma(1)}(t), \zeta_{\sigma(2)}(t), \ldots, \zeta_{\sigma(i-1)}(t)\}\Big\},$$

with the usual convention that $\inf \emptyset = \infty$. Put

$$J_i := \min\Big\{j \in \{1, 2, \ldots, i-1\} : \xi_{\sigma(i)}(\tau_i) = \zeta_{\sigma(j)}(\tau_i)\Big\} \quad \text{if } \tau_i < \infty.$$

For $t \geq 0$, define

$$\zeta_{\sigma(i)}(t) := \begin{cases} \xi_{\sigma(i)}(t) & \text{if } t < \tau_i \\ \zeta_{\sigma(J_i)}(t) & \text{if } t \geq \tau_i. \end{cases}$$

We call the map Λ_σ a *collision rule*. It produces a vector of "coalescing" paths from of a vector of "free" paths: after the free paths labeled i and j collide, the corresponding coalescing paths both subsequently follow either the path labeled i or the path labeled j, according to whether $\sigma(i) < \sigma(j)$ or $\sigma(i) > \sigma(j)$. Note for each $n < N$ that the value of $(\zeta_{\sigma(i)})_{1 \leq i \leq n}$ is unaffected by the value of $(\xi_{\sigma(j)})_{j>n}$.

Suppose from now on that the paths $\xi_1, \xi_2, \ldots$ are realizations of independent copies of a Feller Markov process X with state space E.

A priori, the distribution of the finite or countable coalescing system $\boldsymbol{\zeta} = \Lambda_\sigma \boldsymbol{\xi}$ depends on the ranking σ. However, we have the following result, which is a consequence of the strong Markov property of $\boldsymbol{\xi}$ and the observation that if we are given a bijection $\pi : [N] \to [N]$ and define a map $\Sigma_\pi : D^N \to D^N$ by $(\Sigma_\pi \boldsymbol{\xi})_i = \xi_{\pi(i)}$, $i \in [N]$, then $\Sigma_\pi \Lambda_\sigma = \Lambda_{\sigma\pi^{-1}} \Sigma_\pi$.

Lemma 3.2.1 ([10, 11])**.** *The distribution of $\boldsymbol{\zeta} = \Lambda_\sigma \boldsymbol{\xi}$ is the same for all bijections $\sigma : [N] \to [N]$.*

From now on, we will, unless we explicitly state otherwise, take $\sigma = \mathrm{id}$, where $\mathrm{id} : [N] \to [N]$ is the identity bijection. To simplify notation, we will write Λ for the collision rule Λ_{id}.

It is intuitively clear that the coalescing system $\boldsymbol{\zeta}$ is Markov. For the sake of completeness, we establish this formally in the next lemma, the proof of which is essentially an argument from [10, 11].

Define the right-continuous filtration $(\mathcal{F}_t)_{t \geq 0}$ by

$$\mathcal{F}_t := \bigcap_{\varepsilon > 0} \sigma\{\xi_i(s) : s \leq t + \varepsilon, i \geq 1\}.$$

Lemma 3.2.2. *The stochastic process $\boldsymbol{\zeta} = \Lambda \boldsymbol{\xi}$ is strong Markov with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$.*

Proof. Define maps $m : \{1, 2, \ldots, N\} \times E^N \to \{1, 2, \ldots, N\}$ and $\Pi : E^N \times E^N \to E^N$ by setting $m(i, \mathbf{x}) := \min\{j : x_j = x_i\}$ and $\Pi(\mathbf{x}, \mathbf{y})_i := y_{m(i,\mathbf{x})}$. Note that

$$\Pi(\Lambda\boldsymbol{\eta}(t), \boldsymbol{\eta}(t)) := \Lambda\boldsymbol{\eta}(t), \quad \boldsymbol{\eta} \in D^N,\ t \geq 0.$$

Define a map $\widetilde{\Pi} : E^N \times D^N \to D^N$ by

$$\widetilde{\Pi}(\mathbf{x}, \boldsymbol{\eta})(t) = \Pi(\mathbf{x}, \boldsymbol{\eta}(t)), \quad \mathbf{x} \in E^N,\ \boldsymbol{\eta} \in D^N,\ t \geq 0.$$

Writing $\{\theta_s\}_{t\geq 0}$ for the usual family of shift operators on D^N, that is, $(\theta_s\boldsymbol{\eta})(t) = \boldsymbol{\eta}(s + t)$, we have

$$\theta_s\Lambda\boldsymbol{\eta} = \Lambda\widetilde{\Pi}(\Lambda\boldsymbol{\eta}(s), \theta_s\boldsymbol{\eta}), \quad \boldsymbol{\eta} \in D^N,\ s \geq 0.$$

Fix a bounded measurable function on $f : D^N \to \mathbb{R}$ and set

$$g(\mathbf{x}, \mathbf{y}) = \mathbb{E}^{\mathbf{y}}\left[f\left(\Lambda\widetilde{\Pi}(\mathbf{x}, \boldsymbol{\xi})\right)\right].$$

Note that since the components of $\boldsymbol{\xi}$ are independent, if $\Pi(\mathbf{x}, \mathbf{y}) = \mathbf{x}$, then $g(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{x})$. Thus, for a finite $(\mathcal{F}_t)_{t\geq 0}$ stopping time S we have from the strong Markov property of $\boldsymbol{\xi}$ that

$$\begin{aligned}
\mathbb{E}^{\mathbf{x}}\left[f(\theta_S\Lambda\boldsymbol{\xi}) \,\middle|\, \mathcal{F}_S\right] &= \mathbb{E}^{\mathbf{x}}\left[f\left(\Lambda\widetilde{\Pi}(\Lambda\boldsymbol{\xi}(S), (\theta_S\boldsymbol{\xi}))\right) \,\middle|\, \mathcal{F}_S\right] \\
&= g(\Lambda\boldsymbol{\xi}(S), \boldsymbol{\xi}(S)) \\
&= g(\Lambda\boldsymbol{\xi}(S), \Lambda\boldsymbol{\xi}(S)) \\
&= \mathbb{E}^{\Lambda\boldsymbol{\xi}(S)}[f(\Lambda\boldsymbol{\xi})],
\end{aligned}$$

as required. □

3.2.2 Set-valued coalescing process

Write $\mathcal{K} = \mathcal{K}(E)$ for the set of nonempty compact subsets of E equipped with the usual Hausdorff metric d_H defined by

$$d_H(K_1, K_2) := \inf\{\varepsilon > 0 : K_1^\varepsilon \supseteq K_2 \text{ and } K_2^\varepsilon \supseteq K_1\},$$

where $K^\varepsilon := \{y \in E : \exists x \in K,\ d(y, x) < \varepsilon\}$. The metric space $(\mathcal{K}, d_H)$ is complete. It is compact if E is.

If the locally compact space E is not compact, write $\mathcal{C} = \mathcal{C}(E)$ for the set of nonempty closed subsets of E. Identify the elements of $\mathcal{C}$ with their closures in the one-point compactification $\bar{E}$ of E. Write d_C for the metric on $\mathcal{C}$ that arises from the Hausdorff metric on the compact subsets of $\bar{E}$ corresponding to some metric on $\bar{E}$

that induces the topology of $\bar{E}$.

Let $\Xi_t \subseteq E$ denote the closure of the set $\{\zeta_i(t) : i = 1, 2, \ldots\}$ in E, where $\boldsymbol{\zeta} = \Lambda\boldsymbol{\xi}$.

The following result is an almost immediate consequence of Lemma 3.2.1.

Lemma 3.2.3. *If $\mathbf{x}', \mathbf{x}'' \in E^N$ are such that the sets $\{x'_i : i \in [N]\}$ and $\{x''_i : i \in [N]\}$ are equal, then the distributions of the process Ξ under $\mathbf{P}^{\mathbf{x}'}$ and $\mathbf{P}^{\mathbf{x}''}$ are also equal.*

For the remainder of this section, we will make the following assumption.

Assumption 3.2.4. *The Feller process X is such that if X' and X'' are two independent copies of X, then, for all $t_0 > 0$ and $x' \in E$,*

$$\lim_{x'' \to x'} \mathbf{P}^{x',x''}\big\{X'_t = X''_t \text{ for some } t \in [0, t_0]\big\} = 1.$$

Proposition 3.2.5. *Let $\mathbf{x}', \mathbf{x}'' \in E^N$ be such that the sets $\{x'_i : i \in [N]\}$ and $\{x''_i : i \in [N]\}$ have the same closure. Then, the process Ξ has the same distribution under $\mathbf{P}^{\mathbf{x}'}$ and $\mathbf{P}^{\mathbf{x}''}$.*

Proof. We will consider the case where E is compact. The non-compact case is essentially the same, and we leave the details to the reader.

We need to show for any finite set of times $0 < t_1 < \ldots < t_k$ that the distribution of $(\Xi_{t_1}, \ldots, \Xi_{t_k})$ is the same under $\mathbf{P}^{\mathbf{x}'}$ and $\mathbf{P}^{x''}$.

We may suppose without loss of generality that $x'_1, x'_2, \ldots$ (resp. $x''_1, x''_2, \ldots$) are distinct.

Fix $n \in [N]$ and $\delta > 0$. Given $\varepsilon > 0$ that will be specified later, choose $y''_1, y''_2, \ldots, y''_n \in \{x''_i : i \in [N]\}$ such that $d(x'_i, y''_i) \leq \varepsilon$ for $1 \leq i \leq n$. Let $\boldsymbol{\eta}'$ (resp. $\boldsymbol{\eta}''$) be an E^n-valued process with coordinates that are independent copies of X started at $(x'_1, \ldots, x'_n)$ (resp. $(y''_1, y''_2, \ldots, y''_n)$).

By the Feller property, there is a time $0 < t_0 \leq t_1$ that depends on $x'_1, \ldots, x'_n$ such that for all ε sufficiently small

$$\mathbf{P}\{\eta''_i(t) = \eta''_j(t) \text{ for some } 1 \leq i \neq j \leq n \text{ and } 0 < t \leq t_0\} \leq \frac{\delta}{2}.$$

By our standing Assumption 3.2.4, if we take ε sufficiently small, then

$$\mathbf{P}\{\eta'_i(t) \neq \eta''_i(t) \text{ for all } 0 < t \leq t_0\} \leq \frac{\delta}{2n}, \quad 1 \leq i \leq n.$$

Write Ξ' (resp. Ξ'', $\hat{\Xi}$, $\check{\Xi}$) for the set-valued processes constructed from $\boldsymbol{\eta}'$ (resp. $\boldsymbol{\eta}''$, $(\boldsymbol{\eta}', \boldsymbol{\eta}'')$, $(\boldsymbol{\eta}'', \boldsymbol{\eta}')$) in the same manner that Ξ is constructed from $\boldsymbol{\xi}$. We have

$$\mathbf{P}\{\check{\Xi}_t = \Xi''_t \text{ for all } t \geq t_0\} \geq 1 - \delta,$$

$$\Xi'_t \subseteq \hat{\Xi}_t, \quad \text{for all } t \geq 0,$$

and, by Lemma 3.2.3,

$$\hat{\Xi} \stackrel{d}{=} \check{\Xi}.$$

For each $z \in E$, define a continuous function $\phi_z : \mathcal{K} \to \mathbb{R}_+$ by

$$\phi_z(K) := \inf\{d(z,w) : w \in K\}.$$

Note that $K' \subseteq K''$ implies that $\phi_z(K') \geq \phi_z(K'')$ for any $z \in E$. It follows that for points $z_{\ell p} \in E$, $1 \leq p \leq q_\ell$, $1 \leq \ell \leq k$,

$$\begin{aligned}
\mathbb{E}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\Xi'_{t_\ell})\right] &\geq \mathbb{E}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\hat{\Xi}_{t_\ell})\right] \\
&= \mathbb{E}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\check{\Xi}_{t_\ell})\right] \\
&\geq \mathbb{E}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\Xi''_{t_\ell})\right] - \delta\left(\sup\{d(z,w) : z,w \in E\}\right)^{\sum_\ell q_\ell}
\end{aligned}$$

Observe that

$$\mathbb{E}^{\mathbf{x}'}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\Xi_{t_\ell})\right] = \lim_{n\to\infty}\mathbb{E}^{\mathbf{x}'}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\Xi'_{t_\ell})\right]$$

and

$$\mathbb{E}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\Xi''_{t_\ell})\right] \geq \mathbb{E}^{\mathbf{x}''}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\Xi_{t_\ell})\right].$$

Since δ is arbitrary,

$$\mathbb{E}^{\mathbf{x}'}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\Xi_{t_\ell})\right] \geq \mathbb{E}^{\mathbf{x}''}\left[\prod_{\ell=1}^{k}\prod_{p=1}^{q_\ell}\phi_{z_{\ell p}}(\Xi_{t_\ell})\right].$$

Moreover, we see from interchanging the roles of $\mathbf{x}'$ and $\mathbf{x}''$ that the last inequality is actually an equality.

It remains to observe from the Stone-Weierstrass theorem that the algebra of continuous functions generated by the constants and the set $\{\phi_z : z \in E\}$ is uniformly dense in the space of continuous functions on E. □

With Proposition 3.2.5 in hand, it makes sense to talk about the distribution of the process Ξ for a given initial state Ξ_0. The following result follows immediately from Dynkin's criterion for a function of Markov process to be also Markov.

Corollary 3.2.6. *The process $(\Xi_t)_{t\geq 0}$ is strong Markov with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$.*

3.2.3 Coalescing particles in the marked space and a coupling

Let $(\xi_i)_{i=1}^n$ be n independent Feller Markov processes on E. We assume the usual conditions on E. We are interested in the situation where each particle carries a *mark* or *label* which is an element from some set M. Let $m_i \in M$ be the mark carried by the particle i and it remains fixed during the evolution of the particle. Let $\hat{\xi}_i = (\xi, m_i)$ be the marked process corresponding to ξ_i which takes values in marked space $\hat{E} := E \times M$. Set $\boldsymbol{\xi} = (\xi_i)_{i=1}^n$ and $\hat{\boldsymbol{\xi}} = (\hat{\xi}_i)_{i=1}^n$.

Fix a ranking σ among the n particles. Let Λ_σ and $\hat{\Lambda}_\sigma$ be the collision rule for the vector-valued processes $\boldsymbol{\xi}$ and $\hat{\boldsymbol{\xi}}$ respectively. Let $\boldsymbol{\zeta} := \Lambda_\sigma \boldsymbol{\xi}$ and $\hat{\boldsymbol{\zeta}} = \hat{\Lambda}_\sigma \hat{\boldsymbol{\xi}}$. We will call $\hat{\boldsymbol{\zeta}}$ the vector-valued coalescing process in the marked space.

Claim. We can couple $\boldsymbol{\xi}$ and $\hat{\boldsymbol{\xi}}$ such that almost surely,

$$\{\zeta_i(t) : 1 \leq i \leq n\} \subseteq \{\zeta_i'(t) : \hat{\zeta}(t) = (\zeta_i'(t), m_i),\ 1 \leq i \leq n\}$$

for all $t \geq 0$.

In Subsection 3.2.1, we defined the collision rule Λ_σ for any ranking σ which is a permutation on $[n]$. But for the present purposes it will be more convenient if we allow the ranking σ to be any injective function from $[n]$ to $\mathbb{N}$ in which case the definition of the collision rule remains exactly the same. The coupling will be done via building an appropriate coalescing process in $\hat{E}^n$ but instead of using the fixed ranking σ we suitably update the ranking based on the information from the past during the evolution of the coalescing process.

A verbal description of the coupling. We classify the particles into three groups - *active, partially dead and dead* and the groups are updated during the evolution of the particles. All particles are initially active. An active particle can become either a partially dead particle or a dead particle. A partially dead particle can only become a dead particle. A dead particle can not change its status.

(i) If an active particle collides with another active particle, the path of particle with smaller rank (at the time of the collision) remains unchanged and it is still active. If the two colliding particles have same mark, then the particle with the higher rank becomes dead and follow the path of the other particle thereafter. There is no change in the ranking. But if the colliding particles have different marks, the particle with the higher rank becomes partially dead, it follows its own path and its ranking along with all the particles that have already coalesced with it is increased by n while the rankings of the all other particle remains unchanged.

(ii) If a partially dead particle collides with an active particle with same mark, it becomes dead and follow the path of the active particle thereafter. If a partially dead particle collides with an active particle with same mark, it becomes dead and follow the path of the active particle thereafter. If two partially dead particles sharing the same mark collide, then the particle with the higher rank becomes dead and follow the path of the particle with lower rank thereafter.

(iii) If there is a collision between any pair of particles not described above, both of the colliding particles follow their own paths and there is no change in the ranking.

We now give a more formal description of the above process. Let $\Gamma_0 = 0 < \Gamma_1 < \Gamma_2 < \cdots$ be the successive collision times for the process $\boldsymbol{\zeta} = \Lambda_\sigma \boldsymbol{\xi}$, that is,

$$\Gamma_{i+1} := \inf\{t > \Gamma_i : \zeta_j(t) = \zeta_k(t), \zeta_j(\Gamma_i) \neq \zeta_k(\Gamma_i)\}. \tag{3.1}$$

To build our coalescing system, we first define ranking-valued process $(\sigma_t)_{t\geq 0}$ which starts from σ_0 at time $t = 0$ and is $\mathcal{F}_t$-adapted. $(\sigma_t)_{t\geq 0}$ is constant on each interval $[\Gamma_i, \Gamma_{i+1})$. For $i \geq 1$ and $k \in [n]$,

$$\sigma_{\Gamma_i}(k) := \begin{cases} \sigma_{\Gamma_i -}(k) + n & \text{if } \exists j \in [n], \text{ such that } \zeta_j(t) = \zeta_k(t), \\ & \quad \zeta_j(\Gamma_{i-1}) \neq \zeta_k(\Gamma_{i-1}), m_j \neq m_k, \sigma(j) < \sigma(k), \\ \sigma_{\Gamma_i -}(k) & \text{otherwise.} \end{cases}$$

We can now define the new coalescing process $\tilde{\boldsymbol{\zeta}}$ on $\hat{E}^n$ as follows. If $\Gamma_i \leq t \leq \hat{\Gamma}_{i+1}$, then

$$\tilde{\boldsymbol{\zeta}}(t) := \hat{\Lambda}_{\sigma_{\Gamma_i}}\Big(\theta_{\Gamma_i - \Gamma_{i-1}} \circ \hat{\Lambda}_{\sigma_{\Gamma_{i-1}}} \ldots \big(\theta_{\Gamma_2 - \Gamma_1} \circ \hat{\Lambda}_{\sigma_{\Gamma_1}}(\theta_{\Gamma_1 - \Gamma_0} \circ \hat{\Lambda}_{\sigma_{\Gamma_0}} \hat{\boldsymbol{\xi}})\big)\Big)(t - \Gamma_i),$$

where θ_t's are the shift operators on the path space.

For the process $\tilde{\boldsymbol{\zeta}}$, it can be easily seen that the particle i is active at time t if and only if $\sigma_t(i) = \sigma(i)$. The following observations prove the claim.

1. The rank of a partially dead particle is always higher than that of an active particle.

2. During the evolution of the process we never change the relative ranking of the active particles. Thus the set of the locations of the active particles present at time t evolves as the set-valued coalescing process corresponding to $\boldsymbol{\zeta} = \Lambda_\sigma \boldsymbol{\xi}$.

3. The vector-valued process $\tilde{\boldsymbol{\zeta}}$ is distributed same as the coalescing process $\hat{\boldsymbol{\zeta}}$. This fact essentially follows from strong Markov property of $\hat{\boldsymbol{\xi}}$ with respect to filtration $\mathcal{F}_t$. Indeed if $\hat{\Gamma}_i, i \geq 0$ are the successive collision times for the process $\hat{\boldsymbol{\zeta}}$ which can be similarly defined as (3.1), then one can prove via induction on

$i \geq 0$ that the distribution of the process $(\hat{\Gamma}_i \wedge t, \hat{\boldsymbol{\zeta}}_{\hat{\Gamma}_i \wedge t})_{t \geq 0}$ does not depend on the ranking process $(\sigma_{\Gamma_i \wedge t})_{t \geq 0}$ as long as $(\sigma_t)_{t \geq 0}$ is $(\mathcal{F}_t)_{t \geq 0}$-adapted.

3.3 Brownian motion on the Sierpinski gasket

3.3.1 Definition and properties of the gasket

Let

$$G_0 := \{(0,0), (1,0), (1/2, \sqrt{3}/2)\}$$

be the vertices of the unit triangle in $\mathbb{R}^2$ and denote by H_0 the closed convex hull of G_0. The *Sierpinski gasket*, which we also call the *finite gasket*, is a fractal subset of the plane that can be constructed via the following Cantor-like cut-out procedure. Let $\{b_0, b_1, b_2\}$ be the midpoints of three sides of H_0 and let A be the interior of the triangle with vertices $\{b_0, b_1, b_2\}$. Define $H_1 := H_0 \setminus A$ so that H_1 is the union of 3 closed upward facing triangles of side length 2^{-1}. Now repeat this operation on each of the smaller triangles to obtain a set H_2, consisting of 9 upward facing closed triangles, each of side 2^{-2}. Continuing this fashion, we have a decreasing sequence of closed non-empty sets $\{H_n\}_{n=0}^{\infty}$ and we define the Sierpinski gasket as

$$G := \bigcap_{n=0}^{\infty} H_n.$$

We call each of the 3^n triangles of side 2^{-n} that make up H_n an *n-triangle* of G. Denote by $\mathcal{T}_n$ the collection of all n-triangles of G. Let $\mathcal{V}_n$ be the set of vertices of the n-triangles.

We call the unbounded set

$$\widetilde{G} := \bigcup_{n=0}^{\infty} 2^n G$$

the *infinite gasket* (where, as usual, we write $cB := \{cx : x \in B\}$ for $c \in \mathbb{R}$ and $B \subseteq \mathbb{R}^2$) . The concept of n-triangle, where n may now be a negative integer, extends in the obvious way to the infinite gasket. Denote the set of all n-triangles of $\widetilde{G}$ by $\widetilde{\mathcal{T}}_n$. Let $\widetilde{\mathcal{V}}_n$ be the vertices of $\widetilde{\mathcal{T}}_n$.

Given a pathwise connected subset $A \in \mathbb{R}^2$, let ρ_A be the *shortest-path metric* on A given by

$$\rho_A(x,y) := \inf\{|\gamma| : \gamma \text{ is a path between } x \text{ and } y \text{ and } \gamma \subseteq A\},$$

where $|\gamma|$ denote the length (that is, the 1-dimensional Hausdorff measure) of γ. For the finite gasket G, ρ_G is comparable to the usual Euclidean metric $|\cdot|$ (see, for

example, [14, Lemma 2.12]) with the relation,

$$|x-y| \le \rho_G(x,y) \le c|x-y|, \quad \forall x, y \in G,$$

for a suitable constant $1 < c < \infty$. It is obvious that the same is also true for the metric $\rho_{\widetilde{G}}$ on the infinite gasket.

Let μ denote the d_f-dimensional Hausdorff measure on $\widetilde{G}$ where $d_f := \log 3/\log 2$ is the *fractal* or *mass dimension* of the gasket. For the finite gasket G we have $0 < \mu(G) < \infty$ and, with a slight abuse of notation, we will also use the notation μ to denote the restriction of this measure to G. Moreover, we have the following estimate on the volume growth of μ

$$C' r^{d_f} \le \mu(B(x,r)) \le C r^{d_f} \quad \text{for } x \in \widetilde{G},\ 0 < r < 1, \tag{3.2}$$

where $B(x,r) \subseteq \widetilde{G}$ is the open ball with center x and radius r in the Euclidean metric and C, C' are suitable constants (see [15]).

3.3.2 Brownian motions

We construct a graph G_n (respectively, $\widetilde{G}_n$) embedded in the plane with vertices $\mathcal{V}_n$ (resp. $\widetilde{\mathcal{V}}_n$) by adding edges between pairs of vertices that are distance 2^{-n} apart from each other. Let X^n (resp. $\widetilde{X}^n$) be the natural random walk on G_n (resp. $\widetilde{G}_n$); that is, the discrete time Markov chain that at each step chooses uniformly at random from one of the neighbors of the current state. It is known (see [15, 14]) that the sequence $(X^n_{\lfloor 5^n t\rfloor})_{t\ge 0}$ (resp. $(\widetilde{X}^n_{\lfloor 5^n t\rfloor})_{t\ge 0}$) converges in distribution as $n \to \infty$ to a limiting process $(X_t)_{t\ge 0}$ (resp. $(\widetilde{X}_t)_{t\ge 0}$) that is a G-valued (resp. $\widetilde{G}$-valued) strong Markov process (indeed, a Feller process) with continuous sample paths. The processes X and $\widetilde{X}$ are called, for obvious reasons, the Brownian motion on the finite and infinite gaskets, respectively. The Brownian motion on the *infinite* gasket has the following scaling property:

$$(2\widetilde{X}_t)_{t\ge 0} \text{ under } \mathbf{P}^x \text{ has same law as } (\widetilde{X}_{5t})_{t\ge 0} \text{ under } \mathbf{P}^{2x}. \tag{3.3}$$

The process $\widetilde{X}$ has a family $\widetilde{p}(t,x,y)$, $x, y \in \widetilde{G}$, $t > 0$, of transition densities with respect to the measure μ that is jointly continuous on $(0,\infty) \times \widetilde{G} \times \widetilde{G}$. Moreover, $\widetilde{p}(t,x,y) = \widetilde{p}(t,y,x)$ for all $x, y \in \widetilde{G}$ and $t > 0$, so that the process $\widetilde{X}$ is symmetric with respect to μ.

Let $d_w := \log 5/\log 2$ denote the *walk dimension* of the gasket. The following

crucial "heat kernel bound" is established in [15]

$$c_1' t^{-d_f/d_w} \exp\left(-c_2'\left(\frac{|x-y|^{d_w}}{t}\right)^{1/(d_w-1)}\right) \le \widetilde{p}(t,x,y) \tag{3.4}$$

$$\le c_1 t^{-d_f/d_w} \exp\left(-c_2\left(\frac{|x-y|^{d_w}}{t}\right)^{1/(d_w-1)}\right), \quad \forall x,y \in \widetilde{G}, t>0. \tag{3.5}$$

Because the infinite gasket $\widetilde{G}$ and the associated Brownian motion $\widetilde{X}$ both have re-scaling invariances that G and X do not, it will be convenient to work with $\widetilde{X}$ and then use the following observation to transfer our results to X.

Lemma 3.3.1 (Folding lemma)**.** *There exists a continuous mapping $\psi : \widetilde{G} \to G$ such that ψ restricted to G is the identity, ψ restricted to any 0-triangle is an isometry, and $|\psi(x)-\psi(y)| \le |x-y|$ for arbitrary $x, y \in \widetilde{G}$. Moreover, if the $\widetilde{G}$-valued process $\widetilde{X}$ is started at an arbitrary $x \in \widetilde{G}$, then the G-valued process $\psi \circ \widetilde{X}$ has the same distribution the process X started at $\psi(x)$.*

Proof. Let L be the subset of the plane formed by the set of points of the form $n_1(1,0)+n_2(1/2,\sqrt{3}/2)$, where n_1, n_2 are non-negative integers, and the line segments that join such points that are distance 1 apart. It is easy to see that there is a unique labeling of the vertices of L by $\{1,\omega,\omega^2\}$ that has the following properties.

- Label $(0,0)$ with 1.
- If vertex v is labeled $\mathfrak{a} \in \{1,\omega,\omega^2\}$, then the vertex $v+(1,0)$ are labeled with $\mathfrak{a}\mathbb{P}$.
- If we think of the labels as referring to elements of the cyclic group of order 3, then if vertex v is labeled $\mathfrak{a} \in \{1,\omega,\omega^2\}$, then vertex $v+(1/2,\sqrt{3}/2)$ is labeled with $\mathfrak{a}\mathbb{P}^2$.

Indeed, the label of the vertex $n_1(1,0)+n_2(1/2,\sqrt{3}/2)$ is $\omega^{n_1+2n_2}$.

Given a vertex $v \in L$, let $\iota(v)$ be the unique vertex in $\{(0,0),(1,0),(1/2,\sqrt{3}/2)\}$ that has the same label as v. If the vertices $v_1, v_2, v_3 \in L$ are the vertices of a triangle with side length 1, then $\iota(v_1), \iota(v_2), \iota(v_3)$ are all distinct.

With the above preparation, let us now define the map ψ. Given $x \in \widetilde{G}$, let $\Delta \in \widetilde{\mathcal{T}}_0$ be a triangle with vertices v_1, v_2, v_3 that contains x (if x belongs to $\widetilde{\mathcal{V}}_n$, then there may be more than one such triangle, but the choice will not matter). We may write x as a unique convex combination of the vertices v_1, v_2, v_3,

$$x = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3, \quad \sum_{i=1}^{3} \lambda_i = 1, \lambda_i \ge 0.$$

The triple $(\lambda_1, \lambda_2, \lambda_3)$ is the vector of *barycentric* coordinates of x. We define $\psi(x)$ by

$$\psi(x) := \lambda_1 \iota(v_1) + \lambda_2 \iota(v_2) + \lambda_3 \iota(v_3).$$

It is clear that $\psi : \widetilde{G} \to G$ is well-defined and has the stated properties.

Recall that $\widetilde{X}^{(n)}$ be the natural random walk on $\widetilde{G}_n$. It can be verified easily that the projected process $\psi \circ \widetilde{X}^{(n)}$ is the natural random walk on G_n. The result follows by taking the limit as $n \to \infty$ and using the continuoity of ψ. □

Lemma 3.3.2 (Maximal inequality)**.** *(a) Let $\widetilde{X}^i$, $1 \le i \le n$, be n independent Brownian motions on the infinite gasket $\widetilde{G}$ starting from the initial states x^i, $1 \le i \le n$. For any $t > 0$,*

$$\mathbf{P}\left\{ \sup_{0 \le s \le t} |\widetilde{X}^i_s - x^i| > r, \text{ for some } 1 \le i \le n \right\} \le 2nc_1 \exp\Big(- c_2 (r^{d_w}/t)^{1/(d_w - 1)} \Big),$$

where $c_1, c_2 > 0$ are constants and $d_w = \log 5 / \log 2$ is the walk dimension of the gasket.
(b) The same estimate holds for the case of n independent Brownian motions X^i, $1 \le i \le n$, on the finite gasket G starting from the initial states x^i, $1 \le i \le n$.

Proof. (a) Let $\widetilde{X} = (\widetilde{X}_t)_{t \ge 0}$ be a Brownian motion on $\widetilde{G}$. Then for $x \in \widetilde{G}$, $t > 0$, and $r > 0$,

$$\begin{aligned} \mathbf{P}^x \left\{ \sup_{0 \le s \le t} |\widetilde{X}_s - x| > r \right\} &\le \mathbf{P}^x \{ |\widetilde{X}_t - x| > r/2 \} \\ &\quad + \mathbf{P}^x \big\{ |\widetilde{X}_t - x| \le r/2, \sup_{0 \le s \le t} |\widetilde{X}_s - x| > r \big\}. \end{aligned}$$

Writing $S := \inf\{ s > 0 : |\widetilde{X}_s - x| > r \}$, the second term above equals

$$\mathbb{E}^x \Big[1_{\{S < t\}} \mathbf{P}^{\widetilde{X}_S} \{ |\widetilde{X}_{t-S} - x| \le r/2 \} \Big] \le \sup_{y \in \partial B(x,r)} \sup_{s \le t} \mathbf{P}^y \{ |\widetilde{X}_{t-s} - y| > r/2 \},$$

so that

$$\begin{aligned} \mathbf{P}^x \left\{ \sup_{0 \le s \le t} |\widetilde{X}_s - x| > r \right\} &\le 2 \sup_{y \in \widetilde{G}} \sup_{s \le t} \mathbf{P}^y \{ |\widetilde{X}_s - y| > r/2 \} \\ &\le 2c_1 \exp\Big(- c_2 (r^{d_w}/t)^{1/(d_w - 1)} \Big), \end{aligned}$$

where the last estimate is taken from [14, Theorem 2.23(e)]. The lemma now follows by a union bound.

(b) This is immediate from part (a) and Lemma 3.3.1. □

3.3.3 Collision time estimates

We first show that two independent copies of $\widetilde{X}$ collide with positive probability.

Proposition 3.3.3. *Let $\widetilde{X}'$ and $\widetilde{X}''$ be two independent copies of $\widetilde{X}$. Then,*

$$\mathbf{P}^{(x',x'')}\{\exists t > 0 : \widetilde{X}'_t = \widetilde{X}''_t\} > 0$$

for all $(x', x'') \in \widetilde{G} \times \widetilde{G}$.

Proof. Note that $\widetilde{\mathbf{X}} = (\widetilde{X}', \widetilde{X}'')$ is a Feller process on the locally compact separable metric space $\widetilde{G} \times \widetilde{G}$ that is symmetric with respect to the Radon measure $\mu \otimes \mu$ and has transition densities $\widetilde{p}(t, x', y') \times \widetilde{p}(t, x'', y'')$. The corresponding α-potential density is

$$u_\alpha(\mathbf{x}, \mathbf{y}) := \int_0^\infty e^{-\alpha t} \widetilde{p}(t, x_1, y_1) \times \widetilde{p}(t, x_2, y_2)\, dt \quad \text{for } \alpha > 0,$$

where $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$. A standard potential theoretic result says that a compact set $B \subseteq \widetilde{G} \times \widetilde{G}$ is non-polar if there exists a non-zero finite measure ν that is supported on B and has finite energy, that is,

$$\int \int u^\alpha(\mathbf{x}, \mathbf{y})\, \nu(d\mathbf{x})\, \nu(d\mathbf{y}) < \infty.$$

Take $B = \{(x', x'') \in G \times G : x' = x''\}$ and ν to be the 'lifting' of the Hausdorff measure μ on the finite gasket onto B. We want to show that

$$\int_G \int_G \int_0^\infty e^{-\alpha t} \widetilde{p}^2(t, x, y)\, dt\, \mu(dx)\mu(dy) < \infty.$$

It will be enough to show that

$$\int_G \int_G \int_0^\infty \widetilde{p}^2(t, x, y)\, dt\, \mu(dx)\, \mu(dy) < \infty.$$

It follows from the transition density estimate (3.4) and Lemma 3.3.4 below that

$$\int_0^\infty \widetilde{p}^2(t, x, y)\, dt \le C|x - y|^{-\gamma}$$

for some constant C, where $\gamma := 2d_f - d_w$. Thus,

$$\begin{aligned}
&\int_G \int_G \int_0^\infty \widetilde{p}^2(t,x,y)\,dt\,\mu(dx)\,\mu(dy)\\
&\quad\le C\int_G\int_G |x-y|^{-\gamma}\,\mu(dx)\,\mu(dy)\\
&\quad\le C\int_G\int_0^\infty \mu\{x\in G : |x-y|^{-\gamma} > s\}\,ds\,\mu(dy)\\
&\quad\le C\int_G\int_0^\infty \mu\{x\in G : |x-y| < s^{-1/\gamma}\}\,ds\,\mu(dy)\\
&\quad\le C + C\int_G\int_1^\infty \mu\{x\in G : |x-y| < s^{-1/\gamma}\}\,ds\,\mu(dy)\\
&\quad\le C + C_1\int_G\int_1^\infty s^{-d_f/\gamma}\,ds\,\mu(dy) \quad [\text{ By (3.2)}]\\
&\quad\le C + C_2\int_1^\infty s^{-d_f/\gamma}\,ds.
\end{aligned}$$

It remains to note that $\gamma - d_f = (2\log 3/\log 2 - \log 5/\log 2) - (\log 3/\log 2) = (\log 3 - \log 5)/\log 2 < 0$, and so $d_f/\gamma < 1$.

This shows that $\mathbf{P}^{(x',x'')}\{\widetilde{\mathbf{X}} \text{ hits the diagonal}\} > 0$ for some $(x', x'') \in \widetilde{G}\times\widetilde{G}$. Because $\widetilde{p}^2(t,x,y) > 0$ for all $x, y \in \widetilde{G}$ and $t > 0$, we even have $\mathbf{P}^{(x',x'')}\{\widetilde{\mathbf{X}} \text{ hits the diagonal}\} > 0$ for all $(x', x'') \in \widetilde{G}\times\widetilde{G}$. □

We needed the following elementary result in the proof of Proposition 3.3.3.

Lemma 3.3.4. *For $\alpha > 1, \beta > 0$ and $A > 0$,*

$$\int_0^\infty t^{-\alpha}\exp(-A/t^\beta)dt = \Gamma\left(\frac{\alpha-1}{\beta}\right)\beta^{-1}A^{-\frac{\alpha-1}{\beta}} < \infty.$$

Proof. Make the change of variables $u = At^{-\beta}$ in the integration. □

We next establish a uniform lower bound on the collision probability of a pair of independent Brownian motions on the infinite gasket as long as the distance between their starting points remains bounded.

Theorem 3.3.5. *There exist constants $\beta > 0$ and $\underline{p} > 0$ such that if $\widetilde{X}'$ and $\widetilde{X}''$ are two independent Brownian motions on $\widetilde{G}$ starting from any two points x, y belonging to the same n-triangle of $\widetilde{G}$, then*

$$\mathbf{P}^{(x,y)}\{\widetilde{X}'_t = \widetilde{X}''_t \text{ for some } t \in (0, \beta 5^{-n})\} \ge \underline{p}.$$

This result will require a certain amount of work, so we first note that it leads easily to an analogous result for the finite gasket.

Corollary 3.3.6. *If X' and X'' are two independent Brownian motions on G starting from any two points x, y belonging to the same n-triangle of G, then*

$$\mathbf{P}^{(x,y)}\{X'_t = X''_t \text{ for some } t \in (0, \beta 5^{-n})\} \geq \underline{p},$$

where $\beta > 0$ and $\underline{p} > 0$ are the constants given in Theorem 3.3.5.

Proof. The proof follows immediately from Lemma 3.3.1, because if $\widetilde{X}'_t = \widetilde{X}''_t$ for some t, then it is certainly the case that $\psi \circ \widetilde{X}'_t = \psi \circ \widetilde{X}''_t$. □

Definition 3.3.7 (Extended triangles for the infinite gasket)**.** *Recall that $\widetilde{\mathcal{T}}_n$ is the set of all n-triangles of $\widetilde{G}$. Given $\Delta \in \widetilde{\mathcal{T}}_0$ such that Δ does not have the origin as one its vertices, we define the corresponding* extended triangle *$\Delta^e \subset \widetilde{G}$ as the interior of the union of the original 0-triangle Δ with the three neighboring 1-triangles in $\widetilde{G}$ which share one vertex with Δ and are not contained in Δ. Note that for the (unique) triangle Δ in $\widetilde{\mathcal{T}}_n$ having the origin as one of its vertices, there are two neighboring 1-triangles in $\widetilde{G}$ that share one vertex with it which are not contained in Δ. In this case, by Δ^e, we mean the interior of the union of Δ and these two triangles.*

Fix some $\Delta \in \widetilde{\mathcal{T}}_0$. Let $\widetilde{Z}$ be the Brownian motion on Δ^e killed when it exits Δ^e. It follows from arguments similar to those on [60, page 590], that $\widetilde{Z}$ has transition densities $\widetilde{p}_K(t, x, y)$, $t > 0$, $x, y \in \Delta^e$, with respect to the restriction of μ to Δ^e, and these densities have the following properties:

- $\widetilde{p}_K(t, x, y) = \widetilde{p}_K(t, y, x)$ for all $t > 0$, $x, y \in \Delta^e$.
- $\widetilde{p}_K(t, x, y) \leq \widetilde{p}(t, x, y)$, for all $t > 0$, $x, y \in \Delta^e$.
- $y \mapsto \widetilde{p}_K(t, x, y)$ is continuous for all $t > 0$, $x \in \Delta^e$, and $x \mapsto \widetilde{p}_K(t, x, y)$ is continuous for all $t > 0$, $y \in \Delta^e$.

It follows that the process $\widetilde{Z}$ is Feller and symmetric with respect to the measure μ.

Lemma 3.3.8. *Let $\widetilde{Z}', \widetilde{Z}''$ be two independent copies of the killed Brownian motion $\widetilde{Z}$. Given any $\epsilon > 0$, there exists $0 < \delta < \epsilon$ such that the set of $(x, y) \in \Delta^e \times \Delta^e$ for which*

$$\mathbf{P}^{(x,y)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t \in (\delta, \epsilon)\} > 0$$

has positive $\mu \otimes \mu$ mass.

Proof. An argument similar to that in the proof of Proposition 3.3.3 shows that

$$\mathbf{P}^{(x_0,y_0)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t > 0\} > 0$$

for some $(x_0, y_0) \in \Delta^e \times \Delta^e$.

Thus, for any $\epsilon > 0$, we can partition the interval $(0, \infty)$ into the subintervals $(0, \epsilon)$, $[i\epsilon, (i+1)\epsilon)$, $i \geq 0$ and use the Markov property to deduce that there exists a point $(x_1, y_1) \in \Delta^e \times \Delta^e$ such that

$$\mathbf{P}^{(x_1,y_1)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t \in (0, \epsilon)\} > 0. \tag{3.6}$$

By continuity of probability, we can find $0 < \eta < \epsilon < \infty$ such that

$$\mathbf{P}^{(x_1,y_1)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t \in (\eta, \epsilon)\} > 0.$$

By the Markov property,

$$\begin{aligned} 0 &< \mathbf{P}^{(x_1,y_1)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t \in (\eta, \epsilon)\} \\ &= \int_{\Delta^e}\int_{\Delta^e} \widetilde{p}_K(\eta/2, x_1, x)\widetilde{p}_K(\eta/2, y_1, y) \\ &\quad \times \mathbf{P}^{(x,y)}\{\widetilde{Z}'_t = \widetilde{Z}''_t, \text{ for some } t \in (\eta/2, \epsilon - \eta/2)\}\, \mu(dx)\, \mu(dy). \end{aligned}$$

Therefore, the initial points $(x, y) \in \Delta^e \times \Delta^e$ for which the probability

$$\mathbf{P}^{(x,y)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t \in (\eta/2, \epsilon - \eta/2)\}$$

is positive form a set with positive $\mu \otimes \mu$ measure. The proof now follows by taking $\delta = \eta/2$. □

We record the following result for the reader's ease of reference.

Lemma 3.3.9 (Lemma 3.35 of [14])**.** *There exists a constant $c_1 > 1$ such that if $x, y \in \Delta^e$, $r = |x - y|$, then*

$$\mathbf{P}^x\{\widetilde{X}_t = y \ \text{ for some } \ t \in (0, r^{d_w}) \ \text{ and } \ |\widetilde{X}_t - x| \leq c_1 r \ \text{ for all } \ t \leq r^{d_w}\} > 0.$$

Lemma 3.3.10. *There exists a constant $c > 0$ such that for each point $x \in \Delta$, each open subset $U \subset \Delta^e$, and each time $0 < t \leq c$*

$$\mathbf{P}^x\{\widetilde{Z}_t \in U\} > 0.$$

In particular, $\widetilde{p}_K(t, x, y) > 0$ for all $x, y \in \Delta^e$ and $0 < t \leq c$.

Proof. The following three steps combined with the strong Markov property establish the lemma.

Step 1. There exists a constant $c > 0$ such that starting from $x \in \Delta^e$, the unkilled Brownian motion on the infinite gasket $\widetilde{X}$ will stay within Δ^e up to time c with positive probability.

Step 2. Fix $y \in U$. For all sufficiently small $\eta > 0$,

$$\mathbf{P}^y\{\widetilde{X} \text{ does not exit } U \text{ before time } \eta\} > 0.$$

Step 3. For any $\delta > 0$, $z, y \in \Delta^e$

$$\mathbf{P}^z\{\widetilde{Z} \text{ hits } y \text{ before } \delta\} > 0.$$

Consider Step 1. Note that if $x \in \widetilde{G}$, then (see [14, Equation 3.11]) there exists a constant $c > 0$ such that for the unkilled process $\widetilde{X}$, we have,

$$\mathbf{P}^x\{|\widetilde{X}_t - x| \leq 1/4 \ \text{ for } t \in [0, c]\} > 0.$$

But if $x \in \Delta^e$, then

$$\mathbf{P}^x\{\widetilde{X}_t \in \Delta^e \ \text{ for } t \in [0, c]\} \geq \mathbf{P}^x\{|\widetilde{X}_t - x| \leq 1/4 \ \text{ for } t \in [0, c]\},$$

and the claim follows.

Step 2 is obvious from the right continuity of the paths of the killed Brownian motion $\widetilde{Z}$ at time 0.

Consider Step 3. Fix $z, y \in \Delta^e$ and $0 < \delta \leq |z - y|$. Let $\mathcal{S}_n$ be the n-th approximating graph of $\widetilde{G}$ with the set of vertices $\mathcal{V}_n$. Choose n large enough so that we can find points z_0 and y_0 in $\mathcal{V}_n$ close to z and y respectively so that

$$|z - z_0| \leq \frac{\delta}{3}, \quad |y - y_0| \leq \frac{\delta}{3}$$

and

$$B(z, c_1|z - z_0|) \subseteq \Delta^e, \quad B(y_0, c_1|y - y_0|) \subseteq \Delta^e$$

where c_1 is as in Lemma 3.3.9 and the notation $B(u, r)$ denotes the intersection with the infinite gasket $\widetilde{G}$ of the closed ball in the plane of radius r around the point u.

The length of a shortest path Γ lying $\mathcal{S}_n$ between z_0 and y_0 is the same as their distance in the original metric $\rho_{\widetilde{G}}(z_0, y_0)$. Moreover, for any two points p and p' on Γ, the length of the segment of Γ between p and p' is the same as their distance in the original metric $\rho_{\widetilde{G}}(p, p')$.

Thus, we can choose $m + 1$ equally spaced points $z_0, z_1, \ldots, z_m = y_0$ on Γ such that

$$\rho_{\widetilde{G}}(z_{i+1}, z_i) = \frac{1}{m}\rho_{\widetilde{G}}(z_0, y_0) \quad \text{ for each } i.$$

Since Γ is compact, $\text{dist}(\Gamma, \partial\Delta^e) > 0$. Thus we can choose m large so that

$$B(z_i, c_1|z_{i+1} - z_i|) \subseteq \Delta^e \quad \text{ for each } i.$$

By repeated application of Lemma 3.3.9 and the strong Markov property, we conclude that the probability that $\widetilde{Z}$ hits y starting from z before the time

$$T_m := |z - z_0|^{d_w} + |y - y_0|^{d_w} + \sum_{i=0}^{m-1} |z_{i+1} - z_i|^{d_w}$$

is strictly positive. Step 3 follows immediately since

$$T_m \le \left(\frac{\delta}{3}\right)^{d_w} + \left(\frac{\delta}{3}\right)^{d_w} + \text{constant} \times m \times \frac{1}{m^{d_w}} |z_0 - y_0|^{d_w} \le \delta$$

for m sufficiently large, because $d_w > 1$. □

Lemma 3.3.11. *Let $\widetilde{Z}'$ and $\widetilde{Z}''$ be two independent copies of the killed Brownian motion $\widetilde{Z}$. For any $0 < \delta < \beta$, the map*

$$(x, y) \mapsto \mathbf{P}^{(x,y)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \quad \textit{for some } t \in (\delta, \beta)\}$$

is continuous on $\Delta^e \times \Delta^e$.

Proof. We have

$$\begin{aligned}
&\mathbf{P}^{(x,y)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \quad \text{for some } t \in (\delta, \beta)\} \\
&\quad = \int_{\Delta^e} \int_{\Delta^e} \widetilde{p}_K(\delta, x, x')\, \widetilde{p}_K(\delta, y, y') \\
&\qquad \times \mathbf{P}^{(x',y')}\{\widetilde{Z}'_t = \widetilde{Z}''_t \quad \text{for some } t \in (0, \beta - \delta)\}\, \mu(dx')\mu(dy'),
\end{aligned}$$

and the result follows from the continuity of $z \mapsto \widetilde{p}_K(\delta, z, z')$ for each $z' \in \Delta^e$. □

Proof of Theorem 3.3.5. For any $x, y \in \Delta$,

$$\begin{aligned}
&\mathbf{P}^{(x,y)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t \in (\delta, \beta)\} \\
&\quad = \int_{\Delta^e} \int_{\Delta^e} \widetilde{p}_K(\delta/2, x, x') \widetilde{p}_K(\delta/2, y, y') \\
&\qquad \times \mathbf{P}^{(x',y')}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t \in (\delta/2, \beta - \delta/2)\}\, \mu(dx')\, \mu(dy') > 0,
\end{aligned} \tag{3.7}$$

by Lemmas 3.3.8, 3.3.10 and 3.3.11.

Applying Lemma 3.3.11 and equation (3.7) and the fact that a continuous function achieves its minimum on a compact set, we have for any $\Delta \in \widetilde{\mathcal{T}}_0$ that

$$\underline{q}(\Delta) := \inf_{x,y \in \Delta} \mathbf{P}^{(x,y)}\{\widetilde{Z}'_t = \widetilde{Z}''_t \text{ for some } t \in (0, \beta)\} > 0.$$

Note that for any two $\Delta_1, \Delta_2 \in \widetilde{\mathcal{T}}_0$ which do not contain the origin, there exists

a *local isometry* between the corresponding extended triangles Δ_1^e, Δ_2^e. Since the unkilled Brownian motion $\widetilde{X}$ in $\widetilde{G}$ is invariant with respect to local isometries,

$$\underline{q}(\Delta_1) = \underline{q}(\Delta_2).$$

Given two independent copies $\widetilde{X}'$ and $\widetilde{X}''$ of $\widetilde{X}$, set

$$\underline{p} := \inf_{\Delta \in \widetilde{\mathcal{T}}_0} \inf_{x,y \in \Delta} \mathbf{P}^{(x,y)}\{\widetilde{X}'_t = \widetilde{X}''_t \text{ for some } t \in (0,\beta)\}.$$

The above observations enable us to conclude that $\underline{p} > 0$.

For the infinite gasket, if $\Delta \in \widetilde{\mathcal{T}}_n$, then $2^n\Delta \in \widetilde{\mathcal{T}}_0$ and the scaling property of Brownian motion on the infinite gasket gives us that for any $\Delta \in \widetilde{\mathcal{T}}_n$

$$\begin{aligned}&\inf_{x,y\in\Delta} \mathbf{P}^{(x,y)}\{\widetilde{X}'_t = \widetilde{X}''_t \text{ for some } t \in (0,5^{-n}\beta)\}\\ &\quad = \inf_{x,y\in 2^n\Delta} \mathbf{P}^{(x,y)}\{\widetilde{X}'_t = \widetilde{X}''_t \text{ for some } t \in (0,\beta)\}.\end{aligned}$$

Therefore, for any $\Delta \in \widetilde{\mathcal{T}}_n$ and any $x, y \in \Delta$,

$$\mathbf{P}^{(x,y)}\{\widetilde{X}'_t = \widetilde{X}''_t \text{ for some } t \in (0,5^{-n}\beta)\} \geq \underline{p}. \tag{3.8}$$

□

Corollary 3.3.12. *The Brownian motions $\widetilde{X}$ and X on the infinite and finite gaskets both satisfy Assumption 3.2.4.*

Proof. By Theorem 3.3.5 and the Blumenthal zero-one law, we have for two independent Brownian motions $\widetilde{X}'$ and $\widetilde{X}''$ on $\widetilde{G}$ and any point $(x,x) \in \widetilde{G} \times \widetilde{G}$ that

$$\mathbf{P}^{(x,x)}\{\text{for all } \epsilon > 0, \exists\ 0 < t < \epsilon \text{ such that } \widetilde{X}'_t = \widetilde{X}''_t\ \} = 1.$$

Lemma 3.3.11 then gives the claim for $\widetilde{X}$. The proof for X is similar. □

3.4 Instantaneous coalescence on the gasket

We will establish the following three results in this section after obtaining some preliminary estimates.

Theorem 3.4.1 (Instantaneous Coalescence)**.** *(a) Let Ξ be the set-valued coalescing Brownian motion process on $\widetilde{G}$ with Ξ_0 compact. Almost surely, Ξ_t is a finite set for all $t > 0$.*
(b) The conclusion of part (a) also holds for the set-valued coalescing Brownian motion process on G

Theorem 3.4.2 (Continuity at time zero)**.** *(a) Let Ξ be the set-valued coalescing Brownian motion process on $\widetilde{G}$ with Ξ_0 compact. Almost surely, Ξ_t converges to Ξ_0 as $t \downarrow 0$.*
(b) The conclusion of part (a) also holds for the set-valued coalescing Brownian motion process on G.

Theorem 3.4.3 (Instantaneous local finiteness)**.** *Let Ξ be the set-valued coalescing Brownian motion process on $\widetilde{G}$ with Ξ_0 a possibly unbounded closed set. Almost surely, Ξ_t is a locally finite set for all $t > 0$.*

Lemma 3.4.4 (Pigeon hole principle)**.** *Place M balls in m boxes and allow any two balls to be paired off together if they belong to the same box. Then, the maximum number of disjoint pairs of balls possible is at least $(M - m)/2$.*

Proof. Note that in an optimal pairing there can be at most one unpaired ball per box. It follows that the number of paired balls is at least $M - m$ and hence the number of pairs is at least $(M - m)/2$. □

Define the ε-fattening of a set $A \subseteq \widetilde{G}$ to be the set $A^\varepsilon := \{y \in \widetilde{G} : \exists x \in A, |y-x| < \varepsilon\}$. Define the ε-fattening of a set $A \subseteq G$ in G similarly. Recall the constants $\underline{p}$ and β from Theorem 3.3.5. Set $\Gamma := 1/(1 - \underline{p}/5) > 1$. Given a finite subset A of $\widetilde{G}$ or G and a time-interval $I \subseteq \mathbb{R}_+$, define the random variable $\mathcal{R}(A; I)$ to be the *range* of the set-valued coalescing process Ξ in the finite or the infinite gasket during time I with initial state A; that is,

$$\mathcal{R}(A; I) := \bigcup_{s \in I} \Xi_s.$$

Define a stopping time for the same process Ξ by $\tau_m^A := \inf\{t : \#\Xi_t \leq m\}$.

Lemma 3.4.5. *(a) Let Ξ be the set-valued coalescing Brownian motion process in the infinite gasket with $\Xi_0 = A$, where $A \subset \widetilde{G}$ of cardinality n such that A^ε for some $\varepsilon > 0$ is contained in an extended triangle Δ^e of $\widetilde{G}$. Then, there exist constants C_1 and C_2 which may depend on ε but are independent of A such that*

$$\begin{aligned} &\mathbf{P}\left\{\tau^A_{\lceil n\Gamma^{-1}\rceil} > 25\beta n^{-\log_3 5} \text{or } \mathcal{R}(A, [0, \tau_{\lceil n\Gamma^{-1}\rceil}]) \not\subseteq A^{\varepsilon n^{-(1/6)\log_3 5}}\right\} \\ &\quad \leq C_1 \exp(-C_2 n^{1/3}). \end{aligned} \tag{3.9}$$

(b) The same inequality holds for the set-valued coalescing coalescing Brownian motion process in the finite gasket.

Proof. (a) For any integer $b \geq 1$, the set A can be covered by at most 2×3^b b-triangles. Put

$$b_n := \max\{b : 2 \times 3^b \leq n/2\},$$

or, equivalently,

$$b_n = \lfloor \log_3(n/4) \rfloor.$$

By Lemma 3.4.4, at time $t = 0$ it is possible to form at least $n/2 - n/4 = n/4$ disjoint pairs of particles, where two particles are deemed eligible to form a pair if they belong to the same b_n-triangle. Fix such an (incomplete) pairing of particles. Define a new "partial" coalescing system involving n particles, where a particle is only allowed to coalesce with the one it has been paired up with and after such a coalescence occurs the two partners in the pair both follow the path of the particle having the lower rank among the two. Evidently this new system is same as the coalescing system in the marked space where two particles have the same mark if and only if they have been paired up. From the discussion in subsection 3.2.3 the number of surviving distinct particles in this partial coalescing system stochastically dominates the number of surviving particles in the original coalescing system.

By Theorem 3.3.5, the probability that a pair in the partial coalescing system coalesces before time $t_n := \beta 5^{-b_n}$ is at least $\underline{p}$, independently of the other pairs. Thus, the number of coalescence by time t_n in the partial coalescing system stochastically dominates a random variable that is distributed as the number of successes in $n/4$ independent Bernoulli trials with common success probability $\underline{p}$. By Hoeffding's inequality, the probability that a random variable with the latter distribution takes a value $n\underline{p}/5$ or greater is at least $1 - e^{-C_1' n}$ for some constant $C_1' > 0$. Thus, the probability that the number of surviving particles in the original coalescing system drops below $\lceil (1 - \underline{p}/5)n \rceil = \lceil n\Gamma^{-1} \rceil$ by time $t_n \le 25\beta n^{-\log_3 5}$ is at least $1 - e^{-C_1' n}$.

From Lemma 3.3.2(a) and the fact that during a fixed time interval the maximum displacement of particles in the coalescing system is always bounded by the maximum displacement of independent particles starting from the same initial configuration, the probability that over a time interval of length $25\beta n^{-\log_3 5}$ one of the coalescing particles has moved more than a distance $\varepsilon n^{-(1/6)\log_3 5}$ from its original position is bounded by

$$\begin{aligned}
&2nc_1 \exp\Big(- c_2((\varepsilon n^{-(1/6)\log_3 5})^{d_w} (25\beta n^{-\log_3 5})^{-1})^{1/(d_w - 1)} \Big) \\
&\quad \le 2\exp\Big(\log n - C_2'(n^{(1/2)\log_3 5})^{1/(d_w-1)} \Big) \\
&\quad \le C_1 \exp(-C_2 n^{(1/4)\log_3 5}) \\
&\quad \le C_1 \exp(-C_2 n^{1/3}).
\end{aligned}$$

(b) The proof is identical to part (a). It uses Corollary 3.3.6 in place of Theorem 3.3.5 and Lemma 3.3.2(b) in place of Lemma 3.3.2(a). □

Lemma 3.4.6. *(a) Let Ξ be the set-valued coalescing Brownian motion process in the infinite gasket with $\Xi_0 = A$. Fix $\varepsilon > 0$. Set $\nu_i := \varepsilon\gamma^{-(1/6)\log_3 5 \times i}$ and $\eta_i = 25\beta\Gamma^{-i\log_3 5}$*

for $i \geq 1$. There are positive constants $C_1 = C_1(\varepsilon)$ and $C_2 = C_2(\varepsilon)$ such that

$$\mathbf{P}\left\{\tau^A_{\lceil\Gamma^k\rceil} > \sum_{i=k+1}^{m} \eta_i \text{ or } \mathcal{R}(A;[0,\tau^A_{\lceil\Gamma^k\rceil}]) \not\subseteq (A)^{\sum_{i=k+1}^m \nu_i}\right\} \leq \sum_{i=k+1}^{m} C_1 \exp(-C_2\Gamma^{i/3}),$$

uniformly for all sets A of cardinality $\lceil\Gamma^m\rceil$ such that the fattening $A^{\sum_{i=k+1}^m \nu_i}$ is contained in some extended triangle Δ^e of $\widetilde{G}$.
(b) The analogous inequality holds for the set-valued coalescing Brownian motion process in the finite gasket.

Proof. Fix an extended triangle Δ^e of the infinite gasket and a set A such that $\#A = \lceil\Gamma^m\rceil$ and $A^{\sum_{i=k+1}^m \nu_i} \subseteq \Delta^e$. We will prove the bound by induction on m. By the strong Markov property and Lemma 3.4.5, we have, using the notation $A_{\tau,m-1} := \Xi_{\tau^A_{\lceil\Gamma^{m-1}\rceil}}$,

$$\begin{aligned}
&\mathbf{P}\Big\{\tau^A_{\lceil\Gamma^k\rceil} > \sum_{i=k+1}^{m} \eta_i \text{ or } \mathcal{R}(A;[0,\tau^A_{\lceil\Gamma^k\rceil}]) \not\subseteq (A)^{\sum_{i=k+1}^m \nu_i}\Big\} \\
&\quad\leq \mathbf{P}\Big\{\tau^A_{\lceil\Gamma^{m-1}\rceil} > \eta_m \text{ or } \mathcal{R}(A;[0,\tau^A_{\lceil\Gamma^{m-1}\rceil}]) \not\subseteq A^{\nu_m}\Big\} \\
&\qquad + \mathbb{E}\Big[1\Big\{A_{\tau,m-1} \subseteq A^{\nu_m}\Big\} \\
&\qquad\quad \times \mathbf{P}\Big\{\tau^{A_{\tau,m-1}}_{\lceil\Gamma^k\rceil} > \sum_{i=k+1}^{m-1} \eta_i \text{ or } \mathcal{R}(A_{\tau,m-1};[0,\tau^{A_{\tau,m-1}}_{\lceil\Gamma^k\rceil}]) \not\subseteq A^{\sum_{i=k+1}^{m-1}\nu_i}\Big\}\Big] \\
&\quad\leq C_1\exp(-C_2\Gamma^{m/3}) \\
&\qquad + \sup_{A_1:|A_1|=\lfloor\Gamma^{m-1}\rfloor, A_1\subseteq A^{\nu_m}} \mathbf{P}\left\{\tau^{A_1}_{\lceil\Gamma^k\rceil} > \sum_{i=k+1}^{m-1} \eta_i \text{ or } \mathcal{R}(A_1;[0,\tau^{A_1}_{\lceil\Gamma^k\rceil}]) \not\subseteq A_1^{\sum_{i=k+1}^{m-1}\nu_i}\right\}.
\end{aligned}$$

Since $(A^{\nu_m})^{\nu_{m-1}} \subseteq A^{\nu_m+\nu_{m-1}} \subseteq \Delta^e$, the second term on the last expression can be bounded similarly as

$$\begin{aligned}
&\sup_{A_1:|A_1|=\lfloor\Gamma^{m-1}\rfloor, A_1\subseteq A^{\nu_m}} \mathbf{P}\left\{\tau^{A_1}_{\lceil\Gamma^k\rceil} > \sum_{i=k+1}^{m-1} \eta_i \text{ or } \mathcal{R}(A_1;[0,\tau^{A_1}_{\lceil\Gamma^k\rceil}]) \not\subseteq A_1^{\sum_{i=k+1}^{m-1}\nu_i}\right\} \\
&\leq C_1\exp(-C_2\Gamma^{(m-1)/3}) \\
&+ \sup_{A_2:|A_2|=\lfloor\Gamma^{m-2}\rfloor, A_2\subseteq A^{\nu_m+\nu_{m-1}}} \mathbf{P}\left\{\tau^{A_2}_{\lceil\Gamma^k\rceil} > \sum_{i=k+1}^{m-2} \eta_i \text{ or } \mathcal{R}(A_2;[0,\tau^{A_2}_{\lceil\Gamma^k\rceil}]) \not\subseteq A_2^{\sum_{i=k+1}^{m-2}\nu_i}\right\}.
\end{aligned}$$

Iterating the above argument, the assertion follows.
(b) Same as part (a). □

Proof of Theorem 3.4.1. (a) We may assume that $Q := \Xi_0$ is infinite, since otherwise

there is nothing to prove. By scaling, it is enough to prove the theorem when Q is contained in G. Let $Q_1 \subseteq Q_2 \subseteq \ldots \subseteq Q$ be a sequence of finite sets such that $\#Q_m = \lceil \Gamma^m \rceil$ and Q is the closure $\bigcup_{m=1}^{\infty} Q_m$. By assigning suitable rankings to a system of independent particles starting from each point in $\bigcup_{m=1}^{\infty} Q_m$, we can obtain coupled set-valued coalescing processes $\Xi^1, \Xi^2, \ldots$ and Ξ with the property that $\Xi_0^m = Q_m$, $\Xi_0 = Q$, and for each $t > 0$,

$$\Xi_t^1 \subseteq \Xi_t^2 \subseteq \ldots \subseteq \Xi_t$$

and Ξ_t is the closure of $\bigcup_{m=1}^{\infty} \Xi_t^m$.

Fix $\varepsilon > 0$ so that $Q^{\varepsilon \sum_{i=0}^{\infty} \gamma^{-(1/6)\log_3 5 \times i}}$ is contained in the extended triangle corresponding to G. Set $\nu_i := \varepsilon\gamma^{-(1/6)\log_3 5 \times i}$ and $\eta_i := 25\beta\Gamma^{-i\log_3 5}$. Fix $t > 0$. Choose k_0 so that $\sum_{i=k_0+1}^{\infty} \eta_i \leq t$. By Lemma 3.4.6 and the fact that $s \mapsto \#\Xi_s^m$ is non-increasing, we have, for each $k \geq k_0$,

$$\mathbf{P}\left\{\#\Xi_t^m \leq \lceil \Gamma^k \rceil\right\} \geq 1 - \sum_{i=k+1}^{m} C_1 \exp(-C_2\Gamma^{i/3}).$$

By the coupling, the sequence of events $\{\#\Xi_t^m \leq \lceil \Gamma^k \rceil\}$ decreases to the event $\{\#\Xi_t \leq \lceil \Gamma^k \rceil\}$. Consequently, letting $m \to \infty$, we have, for each $k \geq k_0$,

$$\mathbf{P}\left\{\#\Xi_t \leq \lceil \Gamma^k \rceil\right\} \geq 1 - \sum_{i=k+1}^{\infty} C_1 \exp(-C_2\Gamma^{i/3}).$$

Finally letting $k \to \infty$, we conclude that

$$\mathbf{P}\left\{\#\Xi_t < \infty\right\} = 1.$$

(b) Same as part (a). □

Proof of Theorem 3.4.2. (a) Assume without loss of generality that $Q := \Xi_0$ is infinite and contained in the 1-triangle that contains the origin. By Theorem 3.4.1, Ξ_t is almost surely finite and hence it can be considered as a random element in $(\mathcal{K}, d_H)$. It is enough to prove that $\lim_{t\downarrow 0} d_H(\Xi_t, \Xi_0) = 0$ almost surely.

Let $Q_1 \subseteq Q_2 \subseteq \cdots$ be a nested sequence of finite approximating sets of Q chosen as in the proof of Theorem 3.4.1, and let Ξ^m be the corresponding coupled sequence of set-valued processes.

Fix $\delta > 0$. Choose m sufficiently large that $Q \subseteq Q_m^{\delta/2}$. By the right-continuity of the finite coalescing process, we have

$$\lim_{t\downarrow 0} d_H(\Xi_t^m, Q_m) \to 0 \quad a.s.$$

Thus, with probability one, $(\Xi_t^m)^{\delta/2} \supseteq Q_m$ when t is sufficiently close to 0. But, by

the choice of Q_m, with probability one,

$$(\Xi_t^m)^\delta \supseteq (Q_m)^{\delta/2} \supseteq Q \tag{3.10}$$

for t sufficiently close to 0.

Conversely, choose $\varepsilon > 0$ sufficiently small so that $\sum_{i=}^{\infty} \nu_i < \delta/2$ where ν_i is defined as in Lemma 3.4.6. Set $s_k := \sum_{i=k+1}^{\infty} \eta_i \sim C\Gamma^{-k\log_3 5}$. From Lemma 3.4.6, we have

$$\begin{aligned}
&\mathbf{P}\Big\{\mathcal{R}(Q_m;[0,s_k]) \not\subseteq (Q)^\delta\Big\} \\
&\le \mathbf{P}\Big\{\tau^{Q_m}_{\lceil\Gamma^k\rceil} > \sum_{i=k+1}^{m} \eta_i \ \text{ or } \mathcal{R}(Q_m;[0,\tau^{Q_m}_{\lceil\Gamma^k\rceil}]) \not\subseteq (Q)^{\delta/2}\Big\} \\
&+ \mathbf{P}\Big\{\text{max displacement of } \lceil\Gamma^k\rceil \text{ independent particles in } [0,s_{k-1}] > \delta/2\Big\} \\
&\le \sum_{i=k+1}^{m} C_1\exp(-C_2\Gamma^{i/3}) + C_1'\lceil\Gamma^k\rceil\exp(-C_2'\Gamma^k) \\
&\le C_3\exp(-C_2\Gamma^{k/3}).
\end{aligned} \tag{3.11}$$

By Theorem 3.4.1 $\#\Xi_s < \infty$ almost surely, and hence $\Xi_s^m = \Xi_s$ for all m sufficiently large almost surely. Therefore, by letting $m \to \infty$ in (3.11), we obtain

$$\mathbf{P}\Big\{\mathcal{R}(Q;[0,s_k]) \not\subseteq (Q)^\delta\Big\} \le C_3\exp(-C_2\Gamma^{k/3}).$$

Letting $k \to \infty$, we deduce that, with probability one,

$$\Xi_t \subseteq Q^\delta$$

for t sufficiently close to 0. Combined with (3.10), this gives the desired claim. □

Proof of Theorem 3.4.3. By scaling, it suffices to show that almost surely, the set $\Xi_t \cap G$ is finite for all $t > 0$. Fix any $0 < t_1 < t_2$. We will show that almost surely, the set $\Xi_t \cap G$ is finite for all $t \in [t_1, t_2]$.

Set $J_{0,1} := G$. Now for $r \ge 1$, the set $2^rG \setminus 2^{r-1}G$ can be covered by exactly $2 \times 3^{r-1}$ many 0-triangles that we will denote by $J_{r,\ell}$ for $1 \le \ell \le 2 \times 3^{r-1}$. The collection $\{J_{r,\ell}\}$ forms a covering of the infinite gasket.

Put $Q := \Xi_0$ and let D be a countable dense subset of Q. Associate each point of D with one of the (at most two) 0-triangles to which it belongs. Denote by $D_{r,\ell}$ the subset of D consisting of particles associated with $J_{r,\ell}$. Construct a partial coalescing system starting from D such that two particles coalesce if and only if they collide and both of their initial positions belonged to the same set $D_{r,\ell}$. Let $(\Xi_t^{r,\ell})_{t\ge0}$ denote the set-valued coalescing process consisting of the (possibly empty) subset of the particles associated with $J_{r,\ell}$.

Note that $(\bigcup_{(r,\ell)} \Xi_t^{r,\ell})_{t\geq 0}$ is the set-valued coalescing process in the marked space where two particles have same mark if and only if both of them originate from the same $D_{r,\ell}$. Approximate the set D by a sequence of increasing finite sets. By appealing to the same kind of reasoning as in Theorem 3.4.1, we can find an increasing sequence of set-valued coalescing processes in the original (resp. marked) space starting from this sequence of increasing finite sets which 'approximates' the process $(\Xi_t)_{t\geq 0}$ (resp. $(\bigcup_{(r,\ell)} \Xi_t^{r,\ell})_{t\geq 0}$) in the limit. Now using the coupling involving finitely many particles given in Subsection 3.2.3 and then passing to the limit, it follows that

$$\mathbf{P}\{\#\Xi_t \cap G < \infty \ \forall t \in [t_1, t_2]\} \geq \mathbf{P}\{\#\bigcup_{(r,\ell)} \Xi_t^{r,\ell} \cap G < \infty \ \forall t \in [t_1, t_2]\}.$$

It thus suffices to prove that almost surely, the set $G \cap \bigcup_{(r,\ell)} \Xi_t^{r,\ell}$ is finite for all $t \in [t_1, t_2]$.

Fix $\Delta = J_{r,\ell} \in \widetilde{\mathcal{T}}_0$. Recall the notation of Lemma 3.4.6. Find $\varepsilon > 0$ such that $\Delta^{\sum_{i=0}^{\infty} \nu_i} \subset \Delta^e$. Let $A_1 \subseteq A_2 \subseteq \ldots$ be an increasing sequence of sets such that $\bigcup_m A_m = D_{r,\ell}$. Construct coupled set-valued coalescing processes $\widetilde{\Xi}^1 \subseteq \widetilde{\Xi}^2 \subseteq \ldots \subseteq \Xi^{r,\ell}$ such that $\widetilde{\Xi}_0^m = A_m$. Note that by Lemma 3.4.6

$$\begin{aligned}
&\mathbf{P}\Big\{\Xi_t^{r,\ell} \cap G \neq \emptyset \text{ for some } t \in [t_1, t_2]\Big\} \\
&\quad = \lim_{m\to\infty} \mathbf{P}\Big\{\widetilde{\Xi}_t^m \cap G \neq \emptyset \text{ for some } t \in [t_1, t_2]\Big\} \\
&\quad \leq \limsup_{m\to\infty} \mathbf{P}\Big\{\tau^{A_m}_{\lceil \Gamma^r \rceil} > \sum_{i=r+1}^{\infty} \eta_i \text{ or } \Xi^m_{\tau^{A_m}_{\lceil \Gamma^r \rceil}} \not\subseteq \Delta^e \text{ or max displacement} \\
&\qquad \text{of the remaining } \lceil \Gamma^r \rceil \text{ coalescing particles in } [\tau^{A_m}_{\lceil \Gamma^r \rceil}, t_2] > (r - 3/2)\Big\} \\
&\quad \leq \limsup_{m\to\infty} \mathbf{P}\Big\{\tau^{A_m}_{\lceil \Gamma^r \rceil} > \sum_{i=r+1}^{\infty} \eta_i \text{ or } \Xi^m_{\tau^{A_m}_{\lceil \Gamma^r \rceil}} \not\subseteq \Delta^e\Big\} \\
&\qquad + \mathbf{P}\Big\{\text{max displacement of } \lceil \Gamma^r \rceil \text{ independent particles in } [0, t_2] > (r - 3/2)\Big\} \\
&\quad \leq C_1' \exp(-C_2 \Gamma^{r/3}) + 2c_1 \lceil \Gamma^r \rceil \exp\Big(- c_2((r - 3/2)^{d_w}/t_2)^{1/(d_w - 1)}\Big) \\
&\quad \leq C_3 \exp(-C_4 \Gamma^{r/3})
\end{aligned}$$

for some constants $C_3, C_4 > 0$ that may depend on t_2 but are independent of r and ℓ. The first of the above inequalities follows from the fact that

$$\inf_{x\in J_{r,\ell},\ y\in G} |x - y| \geq 2^{r-1} - 1 \geq r - 1,$$

which implies that Δ^e is at least at a distance $(r - 3/2)$ away from G.

Now by a union bound,

$$\mathbf{P}\Big\{\Xi_t^{r,\ell} \cap G \neq \emptyset \text{ for some } t \in [t_1, t_2] \text{ and for some } \ell\Big\} \leq 2 \times 3^{r-1} C_3 \exp(-C_4 \Gamma^{r/3}).$$

By the Borel-Cantelli lemma, the events $\Xi_t^{r,\ell} \cap G \neq \emptyset$ for some $t \in [t_1, t_2]$ happen for only finitely many (r, ℓ) almost surely. This combined with the fact that $\#\Xi_t^{r,\ell} < \infty$ for all $t > 0$ almost surely gives that

$$\#\bigcup_{(r,\ell)} (G \cap \Xi_t^{r,\ell}) < \infty \text{ for all } t \in [t_1, t_2]$$

almost surely. □

3.5 Instantaneous coalescence of stable particles

3.5.1 Stable processes on the real line and unit circle

Let $X = (X_t)_{t \geq 0}$ be a (strictly) stable process with index $\alpha > 1$ on $\mathbb{R}$. The characteristic function of X_t can be expressed as $\exp(-\Psi(\lambda)t)$ where $\Psi(\cdot)$ is called the characteristic exponent and has the form

$$\Psi(\lambda) = c|\lambda|^\alpha \big(1 - i\upsilon \mathrm{sgn}(\lambda) \tan(\pi\alpha/2)\big), \quad \lambda \in (-\infty, \infty), i = \sqrt{-1}.$$

where $c > 0$ and $\upsilon \in [-1, 1]$. The Lévy measure of Π is absolutely continuous with respect to Lebesgue measure, with density

$$\Pi(dx) = \begin{cases} c^+ x^{-\alpha-1} dx & \text{if } x > 0, \\ c^- |x|^{-\alpha-1} dx & \text{if } x < 0, \end{cases}$$

where c^+, c^- are two nonnegative real numbers such that $\upsilon = (c^+ - c^-)/(c^+ + c^-)$. The process is symmetric if $c^+ = c^-$ or equivalently $\upsilon = 0$. The stable process has the scaling property

$$X \stackrel{d}{=} (c^{-1/\alpha} X_{ct})_{t \geq 0}$$

for any $c > 0$. If we put $Y_t := e^{2\pi i X_t}$, then the process $(Y_t)_{t \geq 0}$ is the stable process with index $\alpha > 1$ on the unit circle $\mathbb{T}$.

We define the distance between two points on $\mathbb{T}$ as the length of the shortest path between them and continue to use the same notation $|\cdot|$ as for the Euclidean metric on the real line.

Theorem 3.5.1 (Instantaneous Coalescence)**.** *(a) Let Ξ be the set-valued coalescing stable process on $\mathbb{R}$ with Ξ_0 compact. Almost surely, Ξ_t is a finite set for all $t > 0$. (b) The conclusion of part (a) holds for the set-valued coalescing stable process on $\mathbb{T}$.*

Theorem 3.5.2 (Continuity at time zero)**.** *(a) Let Ξ be the set-valued coalescing stable process on $\mathbb{R}$ with Ξ_0 compact. Almost surely, Ξ_t converges to Ξ_0 as $t \downarrow 0$.*
(b) The conclusion of part (a) holds for the set-valued coalescing stable process on $\mathbb{T}$.

Theorem 3.5.3 (Instantaneous local finiteness)**.** *Let Ξ be the set-valued coalescing stable process on $\mathbb{R}$ with Ξ_0 a possibly unbounded closed set. Almost surely, Ξ_t is a locally finite set for all $t \geq 0$.*

We now proceed to establish hitting time estimates and maximal inequalities for stable processes that are analogous to those established for Brownian motions on the finite and infinite gaskets in Section 3.3. With these in hand, the proofs of Theorem 3.5.1 and Theorem 3.5.2 follow along similar, but simpler, lines to those in the proofs of the corresponding results for the gasket (Theorem 3.4.1 and Theorem 3.4.2), and so we omit them. However, the proof of Theorem 3.5.3 is rather different from that of its gasket counterpart (Theorem 3.4.3), and so we provide the details at the end of this section.

Lemma 3.5.4. *Let $Z = X' - X''$ where X' and X'' are two independent copies of X, so that Z is a symmetric stable process with index α. For any $0 < \delta < \beta$,*

$$\mathbf{P}^z\{Z_t = 0 \ \text{ for some } t \in (\delta, \beta)\} > 0.$$

Proof. The proof follows from [22, Theorem 16] which says that the single points are not essentially polar for the process Z, the fact that Z has a continuous symmetric transition density with respect to Lebesgue measure, and the Markov property of the Z. □

It is well-know that symmetric stable process Z on $\mathbb{R}$ hits points (see, for example, [22, Chapter VIII, Lemma 13]). Thus there exists a $0 < \beta < \infty$ so that

$$0 < \mathbf{P}^1\{Z_t = 0 \text{ for some } t \in (0, \beta)\} =: \underline{p} \text{ (say)}.$$

By scaling,

$$\mathbf{P}^\varepsilon\{Z_t = 0 \text{ for some } t \in (0, \beta\varepsilon^\alpha)\} = \underline{p}.$$

Lemma 3.5.5. *Suppose that X' and X'' are two independent stable processes on $\mathbb{R}$ starting at x' and x''. For any $\varepsilon > 0$,*

$$\inf_{|x'-x''| \leq \varepsilon} \mathbf{P}\{X'_t = X''_t \text{ for some } t \in (0, \beta\varepsilon^\alpha)\} = \underline{p}.$$

Since $X'_t = X''_t$ always implies that $\exp(2\pi i X'_t) = \exp(2\pi i X''_t)$ (but converse is not true), we have the following corollary of the above lemma.

Corollary 3.5.6. *If Y' and Y'' are two independent stable processes on $\mathbb{T}$ starting at y' and y'', then for any $\varepsilon > 0$*

$$\inf_{|y'-y''|\le 2\pi\varepsilon} \mathbf{P}\{Y'_t = Y''_t \text{ for some } t \in (0, \beta\varepsilon^\alpha)\} \ge \underline{p}.$$

Lemma 3.5.7 ([22])**.** *Suppose that X is an α-stable process on the real line. There exists a constant $C > 0$ such that*

$$\mathbf{P}^0\left\{\sup_{0\le s\le 1} |X_s| > u\right\} \le Cu^{-\alpha}, \quad u \in \mathbb{R}_+.$$

Corollary 3.5.8. *(a) Let $X^1, X^2, \ldots, X^n$ be independent stable processes of index $\alpha > 1$ on $\mathbb{R}$ starting from $x^1, x^2, \ldots, x^n$ respectively. Then for each $x \in \mathbb{R}_+$ and $t > 0$,*

$$\mathbf{P}\Big\{\sup_{0\le s\le t} |X^i_s - x^i| > u \text{ for some } 1 \le i \le n\Big\} \le Cntu^{-\alpha}.$$

(b) The same bound holds for n independent stable processes on $\mathbb{T}$ when $u < \pi$.

Again we set $\Gamma := 1/(1 - \underline{p}/5) > 1$. Fix $(\alpha - 1)/2 < \eta < \alpha - 1$ and define $h := 1 - (1 + \eta)/\alpha > 0$. Recall the definitions of τ^A_m and $\mathcal{R}(A; I)$.

Lemma 3.5.9. *Fix $0 < \varepsilon \le 1/2$.*
(a) There is a constant $C_1 = C_1(\varepsilon)$ such that Ξ be a set-valued coalescing stable process in $\mathbb{R}$ with $\Xi_0 = A$, then

$$\mathbf{P}\left\{\tau^A_{\lceil n\Gamma^{-1}\rceil} > \beta(2\ell/n)^\alpha \text{ or } \mathcal{R}(A, [0, \tau_{\lceil n\Gamma^{-1}\rceil}]) \not\subseteq A^{\varepsilon\ell n^{-h}}\right\} \le C_1 n^{-\eta}, \tag{3.12}$$

where $n = \#A$ and $\ell/2$ is the diameter of A.
(b) Let Ξ be the set-valued coalescing process in $\mathbb{T}$ with $\Xi_0 = A$, where A has cardinality n. Then there exists constant $C_1 = C_1(\varepsilon)$, independent of A, such that

$$\mathbf{P}\left\{\tau^A_{\lceil n\Gamma^{-1}\rceil} > \beta(2/n)^\alpha \text{ or } \mathcal{R}(A, [0, \tau_{\lceil n\Gamma^{-1}\rceil}]) \not\subseteq A^{\varepsilon n^{-h}}\right\} \le C_1 n^{-\eta}.$$

Proof. (a) Note that $A^{\varepsilon\ell} \subseteq [a - \ell/2, a + \ell/2]$ for some $a \in \mathbb{R}$, and this interval can be divided into $n/2$ subintervals of length $2\ell/n$. We follow closely the proof of Lemma 3.4.5. By considering a suitable partial coalescing particle system consisting of at least $n/4$ pairs of particles where a pair can only coalesce if they have started from the same subinterval, we have that the number of surviving particles in the original coalescing system is at most $\lceil \Gamma^{-1} n \rceil$ within time $t_n := \beta(2\ell/n)^\alpha$ with error probability bounded by $\exp(-C'_1 n)$.

By Corollary 3.5.8, the maximum displacement of n independent stable particles

on $\mathbb{R}$ within time t_n is at most

$$\varepsilon(t_n)^{1/\alpha} n^{(1+\eta)/\alpha} = 2\beta^{1/\alpha}\varepsilon\ell n^{-1+(1+\eta)/\alpha} = 2\beta^{1/\alpha}\varepsilon\ell n^{-h}$$

with error probability at most $c_2 n^{-\eta}$.

(b) The proof for part (b) is similar. □

Using strong Markov property and Lemma 3.5.9 repetitively as we did in the proof of Lemma 3.4.6, we can obtain the following lemma. We omit the details.

Lemma 3.5.10. *Let $0 < \varepsilon \le 1/2, \ell > 0$ be given. Let $\nu_i := \varepsilon\gamma^{-hi}$ and $\eta_i := \beta 2^\alpha \Gamma^{-\alpha i}$.*
(a) Given a finite set $A \subset \mathbb{R}$, let Ξ denote the set-valued coalescing stable process in $\mathbb{R}$ with $\Xi_0 = A$. Then, there exist constants $C_2 = C_2(\varepsilon)$ such that

$$\mathbf{P}\Big\{\tau^A_{\lceil\Gamma^k\rceil} > \ell^\alpha \sum_{i=k+1}^{m} \eta_i \text{ or } \mathcal{R}(A; [0, \tau^A_{\lceil\Gamma^k\rceil}]) \not\subseteq (A)^{\ell\sum_{i=k+1}^m \nu_i}\Big\} \le C_2\Gamma^{-\eta k},$$

uniformly over all sets A such that $A \subseteq [a - \ell/4, a + \ell/4]$ for some $a \in \mathbb{R}$ and $\#A = \lceil\Gamma^m\rceil$.
(b) Given a finite set $A \subset \mathbb{T}$, let Ξ denote the set-valued coalescing stable process in $\mathbb{T}$ with $\Xi_0 = A$. Then, there exist constants $C_2 = C_2(\varepsilon)$ such that

$$\mathbf{P}\Big\{\tau^A_{\lceil\Gamma^k\rceil} > \sum_{i=k+1}^{m} \eta_i \text{ or } \Xi_{\tau^A_{\lceil\Gamma^k\rceil}} \not\subseteq (A)^{\sum_{i=k+1}^m \nu_i}\Big\} \le C_2\Gamma^{-\eta k},$$

uniformly over all sets $A \subseteq \mathbb{T}$ such that $\#A = \lceil\Gamma^m\rceil$.

Proof of Theorem 3.5.3. By scaling, it is enough to show that for each $0 < t_1 < t_2 < \infty$, almost surely, the set $\Xi_t \cap [-1, 1]$ is finite for each $t \in [t_1, t_2]$. Set $d := 2/\eta$. For $r \ge 1$, define

$$J_{r,1} := \Big[-\sum_{j=1}^{r} j^d, -\sum_{j=1}^{r-1} j^d\Big) \quad \text{and} \quad J_{r,2} := \Big[\sum_{j=1}^{r-1} j^d, \sum_{j=1}^{r} j^d\Big).$$

Then the collection $\{J_{r,i}\}_{r\ge1,i=1,2}$ forms a partition of the real line into bounded sets. Note that $\inf_{x\in[-1,1],y\in J_{r,i}} |x - y| \asymp r^{d+1}$ as $r \to \infty$.

Let D be a countable dense subset of Q. Run a partial coalescing system starting from D such that two particles coalesce if and only if they collide and both belonged initially to the same $J_{r,i}$. Let $(\Xi^{r,i}_t)_{t\ge0}$ denote the set-valued coalescing process consisting of the (possibly empty) subset of the particles starting from $D \cap J_{r,i}$. By arguing similarly as in the proof of Theorem 3.4.3, it suffices to prove that the set $[-1, 1] \cap \Xi^{r,i}_t$ is empty for all $t \in [t_1, t_2]$ for all but finitely many pairs (r, i) almost surely.

Fix a pair (r,i). Find $\varepsilon > 0$ such that $\sum_{i=0}^{\infty} \nu_i \le 1/2$ which implies that $(J_{r,i})^{\sum_{i=0}^{\infty}\nu_i} \subseteq (J_{r,i})^{r^d}$. Let $A_1 \subseteq A_2 \subseteq \dots$ be an increasing sequence of finite sets such that for $\bigcup_m A_m = D \cap J_{r,i}$. Let $\widetilde{\Xi}^m$ be a coalescing set-valued stable processes such that $\widetilde{\Xi}^m_0 = A_m$ and couple these processes together so that $\widetilde{\Xi}^1_t \subseteq \widetilde{\Xi}^2_t \subseteq \dots \subseteq \Xi^{r,i}_t$. Set $b = b(r) := (2/\eta)\lceil \log_\Gamma r \rceil$. Note that by Lemma 3.5.10, Corollary 3.5.8, and the fact that there exists $c_1 > 0$ such that for all r sufficiently large

$$\inf_{x\in[-1,1], y\in J_{r,i}} |x-y| - r^d \ge c r^{d+1},$$

we can write

$$\begin{aligned}
&\mathbf{P}\Big\{\Xi^{r,i}_t \cap [-1,1] \neq \emptyset \text{ for some } t \in [t_1,t_2]\Big\} \\
&\quad = \lim_{m\to\infty} \mathbf{P}\Big\{\widetilde{\Xi}^m_t \cap [-1,1] \neq \emptyset \text{ for some } t \in [t_1,t_2]\Big\} \\
&\quad \le \limsup_{m\to\infty} \mathbf{P}\Big\{\tau^{A_m}_{\lceil \Gamma^b\rceil} > \sum_{i=b+1}^{\infty} \eta_i \text{ or } \widetilde{\Xi}^m_{\tau^{A_m}_{\lceil\Gamma^b\rceil}} \not\subseteq (J_{r,i})^{r^d} \text{ or max displacement} \\
&\qquad \text{of the remaining } \lceil\Gamma^b\rceil \text{ coalescing particles in } [\tau^{A_m}_{\lceil\Gamma^b\rceil}, t_2] > c r^{d+1}\Big\} \\
&\quad \le \limsup_{m\to\infty} \mathbf{P}\Big\{\tau^{A_m}_{\lceil \Gamma^b\rceil} > \sum_{i=r+1}^{\infty} \eta_i \text{ or } \widetilde{\Xi}^m_{\tau^{A_m}_{\lceil\Gamma^b\rceil}} \not\subseteq (J_{r,i})^{r^d}\Big\} \\
&\qquad + \mathbf{P}\Big\{\text{max displacement of } \lceil\Gamma^b\rceil \text{ independent particles in } [0,t_2] > c r^{d+1}\Big\} \\
&\le C_2 \Gamma^{-\eta b} + C_3 \lceil \Gamma^b\rceil c r^{-\alpha(d+1)} \le C_2' r^{-2} + C_3' r^{-\alpha}
\end{aligned}$$

for suitable constants $C_2', C_3' > 0$. The proof now follows from the Borel-Cantelli lemma. $\square$

Chapter 4

Geometric Influences

4.1 Definitions and Main Results

Definition 4.1.1 ([19]). *Let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function. The influence of the i-th coordinate on f is*

$$I_i(f) := \mathbf{P}\{x \in \{0,1\}^n : f(x) \neq f(x \oplus e_i)\},$$

where $x \oplus e_i$ denotes the point obtained from x by replacing x_i by $1 - x_i$ and leaving the other coordinates unchanged. Here $\mathbf{P}$ denotes the product Bernoulli measure on $\{0,1\}^n$.

In [43], Bourgain *et. al.* generalized of the above definition of influences for the boolean functions defined on $\mathbb{R}^n$ equipped with any product measure $\nu^{\otimes n}$. In what follows we always identity a boolean function $f = 1_A$ with the set A.

Definition 4.1.2 ([43]). *Let $A \subseteq \mathbb{R}^n$ be a Borel measurable set. Then the influence of the ith coordinate on A is defined as*

$$I_i(A) := \nu^{\otimes n}\{x \in \mathbb{R}^n : 1_{A_i^x} \text{ is not a constant function }\},$$

where

$$A_i^x := \{y \in \mathbb{R} : (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \in A\}$$

is the restriction of A along the fiber of $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ in the i-th direction.

While KKL bound continues to hold for this definition [43], this definition is somewhat artificial. For example, let $A_1 = \{|x_1| \leq 1, |x_2| \leq 1\}$ and $A_2 = \{|x_1| \leq 1, |x_2| \leq 100\}$ be two subsets of $\mathbb{R}^2$ with product Gaussian measures. Clearly, the set A_2 essentially does not depend on the second coordinate x_2 where the set A_1 does. But according to the above definition, the influences of the second coordinate on both A_1 and A_2 are same.

Mossel *et. al.* introduced [110] a more reasonable definition of influences of the subsets in $\mathbb{R}^n$.

Definition 4.1.3 (Variance Influence, [110])**.** *Let $A \subseteq \mathbb{R}^n$ be a Borel measurable set. Then the variance influence of the ith coordinate on A is defined as*

$$I_i^{\mathrm{var}}(A) := \mathbb{E}_{\nu^{\otimes n}}[\mathrm{Var}_{x_i}(1_{A_i^x})].$$

But there is nontrivial KKL bound for this definition.

We define influences of a subset A as follows:

Definition 4.1.4 (Geometric influence)**.** *Let ν be a probability measure on $\mathbb{R}$. Given a Borel-measurable set $A \subseteq \mathbb{R}$, its lower Minkowski content $m_\nu(A)$ is defined as*

$$m_\nu(A) := \liminf_{r \downarrow 0} \frac{\nu(A + [-r, r]) - \nu(A)}{r}.$$

Consider the product measure $\nu^{\otimes n}$ on $\mathbb{R}^n$. Then for any Borel-measurable set $A \subseteq \mathbb{R}^n$, the geometric influence *of the i-th coordinate on A is*

$$I_i^{\mathcal{G}}(A) := \mathbb{E}_x[m_{\nu_i}(A_i^x)].$$

In order to make the measure we take the influence with respect to clear, we sometimes denote the influence as $I_i^{\mathcal{G}}(A)\Big|_\nu$.

The geometric meaning of the influence is that for a monotone set A, the sum of influences of A is equal to the size of its boundary with respect to a uniform enlargement, that was studied in e.g., [27, 28, 16].

Proposition 4.1.1. *Let ν be a probability measure on $\mathbb{R}$ with C^1 density λ and cumulative distribution function Λ. Assume further that $\lambda(z) > 0$ for all $z \in \mathbb{R}$, that $\lim_{|z| \to \infty} \lambda(z) = 0$, and that λ' is bounded. Let $A \subset \mathbb{R}^n$ be a monotone set. Then*

$$\lim_{r \downarrow 0} \frac{\nu^{\otimes n}(A + [-r, r]^n) - \nu^{\otimes n}(A)}{r} = \sum_{i=1}^n I_i^{\mathcal{G}}(A).$$

We show that for the Gaussian measure on $\mathbb{R}^n$, the geometric influences satisfy the following analogue of the KKL theorem:

Theorem 4.1.2. *Consider the product spaces $\mathbb{R}^n$ endowed with the product Gaussian measure $\mu^{\otimes n}$. Then for any Borel-measurable set $A \subset \mathbb{R}^n$ with $\mu^{\otimes n}(A) = t$ there exists $1 \leq i \leq n$ such that*

$$I_i^{\mathcal{G}}(A) \geq ct(1 - t)\frac{\sqrt{\log n}}{n},$$

where $c > 0$ is a universal constant.

The result extends to a larger set of log-concave measures called Boltzmann measures (see Definition 4.3.2), and is tight up to the constant factor. The proof uses the relation between geometric influences and the h-influences defined in [95], combined with isoperimetric estimates for the underlying probability measures.

Using the same methods, we obtain analogues of Talagrand's bound on the vector of influences [135], and of Friedgut's theorem stating that a function with a low sum of influences essentially depends on a few coordinates [76].

Theorem 4.1.3. *Consider the product spaces $\mathbb{R}^n$ endowed with the product Gaussian measure $\mu^{\otimes n}$. For any Borel-measurable set $A \subset \mathbb{R}^n$, we have:*

1. *If $\mu^{\otimes n}(A) = t$, then*

$$\sum_{i=1}^{n} \frac{I_i^{\mathcal{G}}(A)}{\sqrt{-\log I_i^{\mathcal{G}}(A)}} \geq c_1 t(1-t),$$

2. *If A is monotone and $\sum_{i=1}^n I_i^{\mathcal{G}}(A)\sqrt{-\log I_i^{\mathcal{G}}(A)} = s$, then there exists a set $B \subset \mathbb{R}^n$ such that 1_B is determined by at most $\exp(c_2 s/\epsilon)$ coordinates and $\mu^{\otimes n}(A \triangle B) \leq \epsilon$,*

where c_1 and c_2 are universal constants.

We also show that the geometric influences can be used in Russo-type formulas for location families.

Proposition 4.1.4. *Let ν be a probability measure on $\mathbb{R}$ with continuous density λ and cumulative distribution function Λ. Let $\{\nu_\alpha : \alpha \in \mathbb{R}\}$ denote a family of probability measures which is obtained by translating ν, that is, ν_α has a density λ_α satisfying $\lambda_\alpha(x) = \lambda(x - \alpha)$.*

Assume that λ is bounded and satisfies $\lambda(z) > 0$ on (κ_L, κ_R), the interior of the support of ν. Let A be an increasing subset of $\mathbb{R}^n$. Then the function $\alpha \to {\nu_\alpha}^{\otimes n}(A)$ is differentiable and its derivative is given by

$$\frac{d{\nu_\alpha}^{\otimes n}(A)}{d\alpha} = \sum_{i=1}^{n} I_i^{\mathcal{G}}(A),$$

where the influences are taken w.r.t. the measure $\nu_\alpha^{\otimes n}$.

Theorem 4.1.2 and Proposition 4.1.4 can be combined to get the following corollary which is the Gaussian analogue of the sharp threshold result obtained by Friedgut and

Kalai [77] for the product Bernoulli measure on the hypercube. We call a set transitive if it is invariant under the action of some transitive subgroup of the permutation group S_n.

Corollary 4.1.5. *Let μ_α denote the Gaussian measure on the real line with mean α and variance 1. Let $A \subset \mathbb{R}^n$ be an increasing transitive set. For any $\delta > 0$, denote by $\alpha_A(\delta)$ the unique value of α such that $\mu_\alpha^{\otimes n}(A) = \delta$. Then for any $\epsilon > 0$,*

$$\alpha_A(1-\epsilon) - \alpha_A(\epsilon) \leq c \log(1/2\epsilon)/\sqrt{\log n},$$

where c is a universal constant.

We now use the geometric influences to obtain an isoperimetric result for the Gaussian measure on $\mathbb{R}^n$:

Theorem 4.1.6. *Consider the product spaces $\mathbb{R}^n$ endowed with the product Gaussian measure $\mu^{\otimes n}$. Then for any transitive Borel-measurable set $A \subset \mathbb{R}^n$ we have*

$$\liminf_{r \downarrow 0} \frac{\mu^{\otimes n}(A + [-r, r]^n) - \mu^{\otimes n}(A)}{r} \geq ct(1-t)\sqrt{\log n},$$

where $t = \mu^{\otimes n}(A)$ and $c > 0$ is a universal constant .

This result also extends to all Boltzmann measures.

Since the Gaussian measure is rotation invariant, it is natural to consider the influence sum of rotations of sets. Of particular interest are families of sets that are closed under rotations. In Section 4.5 we study the effect of *rotations* on the geometric influences, and show that under mild regularity condition of being in a certain class $\mathcal{J}_n$ (see Definition 4.5.1), the sum of geometric influences of a convex set can be increased up to $\Omega(\sqrt{n})$ by (a random) orthogonal rotation:

Theorem 4.1.7. *Consider the product Gaussian measure $\mu^{\otimes n}$ on $\mathbb{R}^n$. For any convex set $A \in \mathcal{J}_n$ with $\mu^{\otimes n}(A) = t$, there exists an orthogonal transformation g on $\mathbb{R}^n$ such that*

$$\sum_{i=1}^n I_i^{\mathcal{G}}(g(A)) \geq ct(1-t)\sqrt{-\log(t(1-t))} \times \sqrt{n},$$

where $c > 0$ is a universal constant. Moreover,

$$\mathbb{E}_{M \sim \nu}\left[\sum_{i=1}^n I_i^{\mathcal{G}}(M(A))\right] \geq c\sqrt{n}t(1-t)\sqrt{-\log(t(1-t))},$$

where M is drawn according to the Haar measure ν over the orthogonal group of rotations.

4.1.1 Organization

The rest of the chapter is organized as follows: In Section 4.2 we prove Proposition 4.1.1 and Proposition 4.1.4, thus establishing the geometric meaning of the new definition. In Section 4.3 we discuss the relation between the geometric influences and the h-influences, and prove Theorem 4.1.2. In Section 4.4 we apply Theorem 4.1.2 to establish a lower bound on the size of the boundary of transitive sets with respect to uniform enlargement, proving Theorem 4.1.6. Finally, in Section 4.5 we study the effect of *rotations* on the geometric influences. We conclude the introduction with a brief statistical application of the results established here.

4.1.2 A statistical application

Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d. $\mathrm{N}(\theta, 1)$. Suppose we want to test the hypothesis: $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1 \ \ (\theta_1 > \theta_0)$ with level significance at most $\beta \in (0, 1/2)$.

The remarkable classical result by Neyman and Pearson [113] says that the most powerful test for the above problem is based on the sample average $\bar{Z}_n = n^{-1} \sum_{i=1}^{n} Z_i$ and the critical region of the test is given by $\mathcal{C}_{\mathrm{mp}} = \{\bar{Z}_n > K\}$ where the constant K is chosen is such that $\mathbf{P}_{\theta_0}\{\mathcal{C}_{\mathrm{mp}}\} = \beta$. It can be easily checked that to achieve power at least $1 - \beta$ for this test, we need the parameters θ_0 and θ_1 to be separated by at least $|\theta_1 - \theta_0| > C(\beta)/\sqrt{n}$ for some appropriate constant $C(\beta)$.

Consider the following setup where the test statistics is given by $f(Z_1, \ldots, Z_n)$ where $f : \mathbb{R}^n \to \mathbb{R}$ is a measurable function which is non-degenerate, transitive and non-decreasing in each of its coordinates. The transitivity of f ensures equal weight is given to each data point while constructing the test statistics and the monotonicity of f implies that the distribution of f depends on θ in a monotone fashion. Note that we do not assume any smoothness property of f. In general the test statistics $f(Z_1, \ldots, Z_n)$, in contrast to the sample average which is a sufficient statistics for this problem, may be resulted from an 'inefficient compression' of the data and we have only access to the compressed data.

As f is non-decreasing, in this case the critical region would be of the form $\mathcal{C} = \{f(Z_1, \ldots, Z_n) > K\}$ where the threshold K is chosen so that $\mathbf{P}_{\theta_0}\{\mathcal{C}\} = \beta$.

Note that the region $\mathcal{C}$ satisfies

(i) $\mathbf{P}_{\theta_0}\{\mathcal{C}\} = \beta$.

(ii) $\mathcal{C}$ is transitive,

(iii) $\mathcal{C}$ is an increasing set.

Clearly, the most powerful test belongs to this class but in general a test of above type can be of much less power. An interesting open question will be to find the worst test (that is, having lowest power) among all tests satisfying (i), (ii) & (iii). Intuitively if θ_1 and θ_0 are far apart, even a very weak test can detect the difference

between the null and the alternative. Corollary 4.1.5 gives us a quantitative estimate of how far apart the parameters need to be so that we can safely distinguish them no matter what test we use. Indeed any test satisfying (i), (ii) and (iii) still has power at least $1-\beta$ as long as $|\theta_1 - \theta_0| > c\log(1/2\beta)/\sqrt{\log n}$ for some absolute constant c.

For the test $\{\max_i Z_i > K\}$, the dependence on n in the above bound is tight up to constant factors.

We briefly note that the statistical reasoning introduced here may be combined with Theorem 2.1 in [77]. Thus a similar statement holds when $Z_1, Z_2, \ldots, Z_n$ are i.i.d. Bernoulli(p) and we want to test the hypothesis: $H_0 : p = p_0$ vs $H_1 : p = p_1$ $(1 > p_1 > p_0 > 0)$. In this case, the power of any test satisfying (i), (ii) and (iii) is at least $1-\beta$ as long as $|p_1 - p_0| > c\log(1/2\beta)/\log n$ for some absolute constant c.

4.2 Boundary Under Uniform Enlargement and Derivatives

In this section we provide the geometric interpretation of the influence. We begin by proving Proposition 4.1.1.

4.2.1 Proof of Proposition 4.1.1

In our proof we use the following simple lemma:

Lemma 4.2.1. *Let λ be as given in Proposition 4.1.1. Given $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ such that for all $x, y \in \mathbb{R}$,*

$$|\lambda(x) - \lambda(y)| \le C_\varepsilon |\Lambda(x) - \Lambda(y)| + \varepsilon/4.$$

Proof. Since $\lim_{|z|\to\infty} \lambda(z) = 0$, there exist $0 < z_2 < z_1$ such that $\sup_{|z|\ge z_1} \lambda(z) \le \varepsilon/8$ and $\sup_{|z|\ge z_2} \lambda(z) \le \varepsilon/4$. We consider several cases.

1. **Case I.** $|x| > z_2, |y| > z_2$. In this case, by the choice of z_2, we have $|\lambda(x) - \lambda(y)| \le \varepsilon/4$.

2. **Case II.** $|x| \le z_1, |y| \le z_1$**.** Since the function λ'/λ is continuous, there exists K such that $|\lambda'(z)|/\lambda(z) \le K$ for all $|z| \le z_1$. Hence,

$$|\lambda(x) - \lambda(y)| = |\int_y^x \lambda'(z)dz| \le K|\int_y^x \lambda(z)dz| = K|\Lambda(x) - \Lambda(y)|.$$

3. **Case III a.** $x > z_1, |y| < z_2$**.** In this case,

$$|\lambda(x) - \lambda(y)| \le 2\|\lambda\|_\infty \le \frac{2\|\lambda\|_\infty}{\Lambda(z_1) - \Lambda(z_2)}(\Lambda(x) - \Lambda(y)).$$

4. **Case III b.** $x < -z_1, |y| < z_2$. Similarly,

$$|\lambda(x) - \lambda(y)| \leq 2\|\lambda\|_\infty \leq \frac{2\|\lambda\|_\infty}{\Lambda(-z_2) - \Lambda(-z_1)}(\Lambda(y) - \Lambda(x)).$$

This completes the proof of the lemma, by taking

$$C_\varepsilon = \max\left(K, \frac{2\|\lambda\|_\infty}{\Lambda(z_1) - \Lambda(z_2)}, \frac{2\|\lambda\|_\infty}{\Lambda(-z_2) - \Lambda(-z_1)}\right).$$

□

Now we are ready to present the proof of Proposition 4.1.1.

Proof. Without loss of generality, assume that A is decreasing. Thus, $\nu^{\otimes n}(A + [-r, r]^n) = \nu^{\otimes n}(A + [0, r]^n)$. We decompose $\nu^{\otimes n}(A + [0, r]^n) - \nu^{\otimes n}(A)$ as

$$\sum_{i=1}^{n} \nu^{\otimes n}(A + [0, r]^i \times \{0\}^{n-i}) - \nu^{\otimes n}(A + [0, r]^{i-1} \times \{0\}^{n-i+1}). \tag{4.1}$$

It follows immediately from (4.1) that it is sufficient to show that given $\varepsilon > 0$, there exists $\delta > 0$ such that for all $1 \leq i \leq n$ and for all $0 < r < \delta$,

$$\left|\frac{\nu^{\otimes n}(A + [0, r]^{i-1} \times [0, r] \times \{0\}^{n-i}) - \nu^{\otimes n}(A + [0, r]^{i-1} \times \{0\}^{n-i+1})}{r} - I_i^{\mathcal{G}}(A)\right| \leq \varepsilon. \tag{4.2}$$

For a fixed i, define

$$B_r^i = A + [0, r]^{i-1} \times \{0\}^{n-i+1}.$$

Obviously, B_r^i is a decreasing set. Note that $A + [0, r]^{i-1} \times [0, r] \times \{0\}^{n-i} = B_r^i + \{0\}^{i-1} \times [0, r] \times \{0\}^{n-i}$. Hence, Equation (4.2) can be rewritten as

$$\left|\frac{\nu^{\otimes n}(B_r^i + \{0\}^{i-1} \times [0, r] \times \{0\}^{n-i}) - \nu^{\otimes n}(B_r^i)}{r} - I_i^{\mathcal{G}}(A)\right| \leq \varepsilon. \tag{4.3}$$

For any decreasing set $D \subset \mathbb{R}^n$ and for any $x \in \mathbb{R}^n$, define

$$t_i(D; x) := \sup\{y : y \in D_i^x\} \in [-\infty, \infty],$$

with the convention that the supremum of the empty set is $-\infty$. We use two simple observations:

1. For any decreasing set D (and in particular, for A and for B_r^i), it is clear that $\nu^{\otimes n}(D) = \mathbb{E}_x \Lambda(t_i(D; x))$.

2. For a decreasing set D, we have $I_i^{\mathcal{G}}(D) = \mathbb{E}_x\, \lambda(t_i(D;x))$. This follows from a known property of the lower Minkowski content: In the case when ν has a continuous density λ and L is a semi-infinite ray, that is, $L = [\ell, \infty)$ or $L = (-\infty, \ell]$, we have $m_\nu(L) = \lambda(\ell)$.

We further observe that

$$\left| \frac{\nu^{\otimes n}(B_r^i + \{0\}^{i-1} \times [0,r] \times \{0\}^{n-i}) - \nu^{\otimes n}(B_r^i)}{r} - \mathbb{E}_x\, \lambda(t_i(B_r^i;x)) \right| \leq r||\lambda'||_\infty. \quad (4.4)$$

Indeed, by Observation (1), the l.h.s. of (4.4) is equal to

$$\left| \mathbb{E}_x \left[\frac{\Lambda(t_i(B_r^i;x) + r) - \Lambda(t_i(B_r^i;x))}{r} - \lambda(t_i(B_r^i;x)) \right] \right|. \quad (4.5)$$

By the Mean Value Theorem, there exists $h \in [0,r]$ such that

$$\frac{\Lambda(t_i(B_r^i;x) + r) - \Lambda(t_i(B_r^i;x))}{r} = \lambda(t_i(B_r^i;x) + h),$$

and thus,

$$(4.5) = \left| \mathbb{E}_x[\lambda(t_i(B_r^i;x) + h) - \lambda(t_i(B_r^i;x))] \right| \leq r||\lambda'||_\infty.$$

Combining Equations (4.3) and (4.4), and ensuring that $r < \varepsilon/(2||\lambda'||_\infty)$, it is sufficient to show that

$$\left| \mathbb{E}_x\, \lambda(t_i(B_r^i;x)) - I_i^{\mathcal{G}}(A) \right| \leq \varepsilon/2,$$

and by Observation (2), this is equivalent to

$$\left| \mathbb{E}_x\, \lambda(t_i(B_r^i;x)) - \mathbb{E}_x\, \lambda(t_i(A;x)) \right| \leq \varepsilon/2. \quad (4.6)$$

By Lemma 4.2.1 and Observation (1), we have

$$\begin{aligned} |\, \mathbb{E}_x\, \lambda(t_i(B_r^i;x)) - \mathbb{E}_x\, \lambda(t_i(A;x))| &\leq C_\varepsilon\, \mathbb{E}_x\, |\Lambda(t_i(B_r^i;x)) - \Lambda(t_i(A;x))| + \varepsilon/4 \\ &= C_\varepsilon\, \mathbb{E}_x \left(\Lambda(t_i(B_r^i;x)) - \Lambda(t_i(A;x)) \right) + \varepsilon/4 \\ &= C_\varepsilon(\nu^{\otimes n}(B_r^i) - \nu^{\otimes n}(A)) + \varepsilon/4. \end{aligned}$$

It thus remains to show that there exists $\delta > 0$ sufficiently small such that for all $0 < r < \delta$,

$$\nu^{\otimes n}(B_r^i) - \nu^{\otimes n}(A) \leq \frac{\varepsilon}{4C_\varepsilon}. \quad (4.7)$$

We can write

$$\nu^{\otimes n}(B_r^i) - \nu^{\otimes n}(A) = \sum_{j=1}^{i-1} \Big(\nu^{\otimes n}(A + [0,r]^j \times \{0\}^{n-j}) - \nu^{\otimes n}(A + [0,r]^{j-1} \times \{0\}^{n-j+1}) \Big),$$

and thus it is sufficient to find $\delta > 0$ such that for all $0 < r < \delta$ and for all $1 \le j \le i-1$,

$$\nu^{\otimes n}(A + [0,r]^j \times \{0\}^{n-j}) - \nu^{\otimes n}(A + [0,r]^{j-1} \times \{0\}^{n-j+1}) \le \frac{\varepsilon}{4nC_\varepsilon}.$$

Since for any decreasing $D \subset \mathbb{R}^n$,

$$|\nu^{\otimes n}(D + \{0\}^{j-1} \times [0,r] \times \{0\}^{n-j}) - \nu^{\otimes n}(D)| \le \|\lambda\|_\infty r,$$

we can choose $\delta = \min\{\frac{\varepsilon}{4nC_\varepsilon \|\lambda\|_\infty}, \frac{\varepsilon}{2||\lambda'||_\infty}\}$. This completes the proof. □

Remark 10. *We note that the same proof (with minor modifications) holds for any convex set A. The only non-obvious change is noting that the Minkowski content of a segment $[a,b]$ is $m_\nu([a,b]) = \lambda(a) + \lambda(b)$, where λ is the density of the measure ν. On the other hand, it is clear that the statement of Proposition 4.1.1 does not hold for general measurable sets. For example, if $A = \mathbb{Q}^n$ where $\mathbb{Q}$ is the set of rational numbers, then the L^∞-boundary of A is ∞, while the sum of geometric influences of A is zero. It seems an interesting question to determine to which classes of measurable sets Proposition 4.1.1 applies.*

4.2.2 Proof of Proposition 4.1.4

Define a function $\Pi : \mathbb{R}^n \to [0,\infty)$ by

$$\Pi(\alpha_1, \ldots, \alpha_n) = \nu_{\alpha_1} \otimes \ldots \otimes \nu_{\alpha_n}(A).$$

The partial derivative of Π w.r.t. the i-th coordinate can be written as

$$\frac{\partial \Pi(\alpha_1, \ldots, \alpha_n)}{\partial \alpha_i} = \lim_{r \downarrow 0} \frac{\mathbb{E}_x \, \nu_{\alpha_i + r}(A_i^x) - \mathbb{E}_x \, \nu_{\alpha_i}(A_i^x)}{r}. \tag{4.8}$$

For $x \in \mathbb{R}^n$, define

$$s_i(A;x) := \inf\{y : y \in A_i^x\} \in [-\infty, \infty].$$

Since A is monotone increasing, for any $x \in \mathbb{R}^n$ we have

$$\begin{aligned} \frac{\nu_{\alpha_i + r}(A_i^x) - \nu_{\alpha_i}(A_i^x)}{r} &= \frac{\nu_{\alpha_i + r}([s_i(A;x), \infty)) - \nu_{\alpha_i}([s_i(A;x), \infty))}{r} \\ &= \frac{1}{r} \int_{s_i(A;x) - r}^{s_i(A;x)} \lambda_{\alpha_i}(z) dz, \end{aligned} \tag{4.9}$$

and by the Fundamental Theorem of Calculus, this expression converges to $\lambda_{\alpha_i}(s_i(A;x))$ as $r \to 0$. Moreover, (4.9) is uniformly bounded by $\|\lambda_{\alpha_i}\|_\infty = \|\lambda\|_\infty$. Therefore, by the Dominated Convergence Theorem, it follows that the first order partial derivatives of Π exist and are given by

$$\frac{\partial \Pi(\alpha_1, \ldots, \alpha_n)}{\partial \alpha_i} = \mathbb{E}_{x \sim \nu_{\alpha_1} \otimes \cdots \otimes \nu_{\alpha_n}} \lambda_{\alpha_i}(s_i(A;x)) = I_i^{\mathcal{G}}(A),$$

where the influence is w.r.t. the measure $\nu_{\alpha_1} \otimes \cdots \otimes \nu_{\alpha_n}$. (For the last equality, see Observation (2) in the proof of Proposition 4.1.1 above. Here we use the convention that $\lambda_{\alpha_i}(-\infty) = \lambda_{\alpha_i}(\infty) = 0$).

Hence, by the chain rule, it is sufficient to check that all the partial derivatives of Π are continuous at $(\alpha, \ldots, \alpha)$. Without loss of generality, we assume that $\alpha = 0$. Note that

$$\mathbb{E}_{x \sim \nu_{\alpha_1} \otimes \cdots \otimes \nu_{\alpha_n}} \lambda_{\alpha_i}(s_i(A;x)) = \mathbb{E}_{x \sim \nu \otimes \cdots \otimes \nu} \left(\prod_{j=1}^n \frac{\lambda_{\alpha_j}(x_j)}{\lambda(x_j)} \right) \lambda_{\alpha_i}(s_i(A;x)). \tag{4.10}$$

For each $x \in \mathbb{R}^n$,

$$\prod_{j=1}^n \frac{\lambda_{\alpha_j}(x_j)}{\lambda(x_j)} \lambda_{\alpha_i}(s_i(A;x)) \to \prod_{j=1}^n \lambda(s_i(A;x)) \tag{4.11}$$

as $\max |\alpha_i| \to 0$. Hence, the continuity of the partial derivatives would follow from the Dominated Convergence Theorem if (4.11) was uniformly bounded. In order to obtain such bound, we consider a compact subset.

There exist $\kappa_L < K_L < K_R < \kappa_R$ and $\delta > 0$ such that $\nu([K_L + \delta, K_R - \delta]) \geq 1 - \varepsilon$. Let $c := \min_{z \in [K_L, K_R]} \lambda(z)$. Note that $c > 0$. If $|\alpha_j| \leq \delta$ for all j, then

$$\left| (4.10) - \mathbb{E}_{x \sim \nu \otimes \cdots \otimes \nu} \left(\prod_{j=1}^n \frac{\lambda_{\alpha_j}(x_j)}{\lambda(x_j)} 1_{\{K_L \leq x_j \leq K_R\}} \right) \lambda_{\alpha_i}(s_i(A;x)) \right| \leq \varepsilon \cdot n \cdot \|\lambda\|_\infty. \tag{4.12}$$

Indeed, denoting $S = \{x \in \mathbb{R}^n : \exists j, x_j \notin [K_L, K_R]\}$ and using Equation (4.10), we have

$$(4.12) = \left| \mathbb{E}_{x \sim \nu_{\alpha_1} \otimes \cdots \otimes \nu_{\alpha_n}} 1_S \lambda_{\alpha_i}(s_i(A;x)) \right| \leq ||\lambda||_\infty \, \mathbb{E}_{x \sim \nu_{\alpha_1} \otimes \cdots \otimes \nu_{\alpha_n}} 1_S \leq \varepsilon \cdot n \cdot \|\lambda\|_\infty,$$

where the last inequality is a union bound using the choice of K_L and K_R.

Similarly, by a union bound we have

$$\left| \mathbb{E}_{x \sim \nu \otimes \cdots \otimes \nu} \lambda(s_i(A;x)) - \mathbb{E}_{x \sim \nu \otimes \cdots \otimes \nu} 1_{\{K_L \leq x_j \leq K_R \ \forall j\}} \lambda(s_i(A;x)) \right| \leq \varepsilon \cdot n \cdot \|\lambda\|_\infty. \tag{4.13}$$

Combining (4.12) with (4.13), it is sufficient to prove that

$$\mathbb{E}_{x\sim\nu\otimes\cdots\otimes\nu}\prod_{j=1}^{n}\frac{\lambda_{\alpha_j}(x_j)}{\lambda(x_j)}1_{\{K_L\le x_j\le K_R\}}\lambda_{\alpha_i}(s_i(A;x))\to$$
$$\mathbb{E}_{x\sim\nu\otimes\cdots\otimes\nu}\prod_{j=1}^{n}1_{\{K_L\le x_j\le K_R\}}\lambda(s_i(A;x)).$$

This indeed follows from the Dominated Convergence Theorem, since for each $x\in\mathbb{R}^n$,

$$\prod_{j=1}^{n}\frac{\lambda_{\alpha_j}(x_j)}{\lambda(x_j)}1_{\{K_L\le x_j\le K_R\}}\lambda_{\alpha_i}(s_i(A;x))\to\prod_{j=1}^{n}1_{\{K_L\le x_j\le K_R\}}\lambda(s_i(A;x))$$

as $\max|\alpha_i|\to 0$ and is uniformly bounded by $c^{-n}\|\lambda\|_\infty^{n+1}$. This completes the proof. □

4.3 Relation to h-influences and a general lower bound on geometric influences

In this section we analyze the geometric influences by reduction to problems concerning h-influences introduced in a recent paper by the first author [95]. First we describe and extend the results on h-influences, and then we show their relation to geometric influences.

4.3.1 h-Influences

Definition 4.3.1 (h-influences, [95])**.** *Let $h:[0,1]\to[0,\infty)$ be a measurable function. For a measurable subset A of X^n equipped with a product measure $\nu^{\otimes n}$, the h-influence of the i-th coordinate on A is*

$$I_i^h(A):=\mathbb{E}_x[h(\nu(A_i^x))].$$

The two main results concerning h-influences are a monotonization lemma and an analogue of the KKL theorem.

Lemma 4.3.1 ([95])**.** *Let $h:[0,1]\to[0,1]$ be a concave continuous function. For every Borel measurable set $A\subseteq[0,1]^n$, there exists a monotone set $B\subseteq[0,1]^n$ such that:*

1. *$u^{\otimes n}(A)=u^{\otimes n}(B)$.*
2. *For all $1\le i\le n$, we have $I_i^h(A)\ge I_i^h(B)$.*

Theorem 4.3.2 ([95]). *Denote the entropy function as* $\mathrm{Ent}(x) := -x\log x - (1-x)\log(1-x)$. *Consider the product space* $[0,1]^n$, *endowed with the product Lebesgue measure* $u^{\otimes n}$. *Let* $h : [0,1] \to [0,1]$ *such that* $h(x) \geq \mathrm{Ent}(x)$ *for all* $0 \leq x \leq 1$. *Then for every measurable set* $A \subseteq [0,1]^n$ *with* $u^{\otimes n}(A) = t$, *there exists* $1 \leq i \leq n$ *such that the* h*-influence of the* i*-th coordinate on* A *satisfies*

$$I_i^h(A) \geq ct(1-t)\log n/n,$$

where $c > 0$ *is a universal constant.*

Other results on h-influences which we shall use later include analogues of several theorems concerning influences on the discrete cube: Talagrand's lower bound on the vector of influences [135], a variant of the KKL theorem for functions with low influences [77], and Friedgut's theorem asserting that a function with a low influence sum essentially depends on a few coordinates [76].

In the application to geometric influences we would like to use h-influences for certain functions h that do not dominate the entropy function. In order to overcome this problem, we use the following lemma, that allows to relate general h-influences to the Entropy-influence (i.e., h-influence for $h(x) = \mathrm{Ent}(x)$).

Lemma 4.3.3. *Consider the product space* $(\mathbb{R}^n, \nu^{\otimes n})$, *where* ν *has a continuous cumulative distribution function* Λ. *Let* $h : [0,1] \to [0,\infty)$, *and let* $A \subseteq \mathbb{R}^n$ *be a Borel-measurable set. For all* $1 \leq i \leq n$,

$$I_i^h(A) \geq \frac{1}{2}\delta \cdot I_i^{\mathrm{Ent}}(A), \tag{4.14}$$

where

$$\delta = \delta(A,i) = \inf_{x \in [\vartheta(I_i^{\mathrm{Ent}}(A)/2), 1-\vartheta(I_i^{\mathrm{Ent}}(A)/2)]} \frac{h(x)}{\mathrm{Ent}(x)}, \tag{4.15}$$

and $\vartheta(y) = y/(-2\log y)$.

Proof. Set $f = 1_A$. Let u be the Lebesgue measure on $[0,1]$. Define $g(x_1,\ldots,x_n) := f(\Lambda^{-1}(x_1),\ldots,\Lambda^{-1}(x_n))$ and write B for the set $\{x \in \mathbb{R}^n : g(x) = 1\}$. Since $\Lambda^{-1}(u) \stackrel{d}{=} \nu$, the set B satisfies $u^{\otimes n}(B) = \nu^{\otimes n}(A) = t$ and

$$I_i^h(B)\big|_{u^{\otimes n}} = I_i^h(A)\big|_{\nu^{\otimes n}} \qquad \text{for each } \ 1 \leq i \leq n.$$

Denote $\alpha := \mathrm{Ent}^{-1}(I_i^{\mathrm{Ent}}(A)/2)$. It is clear that for any $x \notin [\alpha, 1-\alpha]$,

$$\mathrm{Ent}(x) \leq \mathrm{Ent}\left(\mathrm{Ent}^{-1}(I_i^{\mathrm{Ent}}(A)/2)\right) = I_i^{\mathrm{Ent}}(A)/2,$$

and thus,

$$\begin{aligned}\mathbb{E}_x\left[\mathrm{Ent}(u(B_i^x))1_{\{u(B_i^x)\in[\alpha,1-\alpha]\}}\right] &= I_i^{\mathrm{Ent}}(B)\big|_{u^{\otimes n}} - \mathbb{E}_x\left[\mathrm{Ent}(u(B_i^x))1_{\{u(B_i^x)\notin[\alpha,1-\alpha]\}}\right] \\ &\geq I_i^{\mathrm{Ent}}(A)/2.\end{aligned}$$

Therefore, by (4.15),

$$\begin{aligned}I_i^h(A)\big|_{\nu^{\otimes n}} = I_i^h(B)\big|_{u^{\otimes n}} &\geq \mathbb{E}_x\left[h(u(B_i^x))1_{\{u(B_i^x)\in[\alpha,1-\alpha]\}}\right] \\ &\geq \left(\inf_{x\in[\mathrm{Ent}^{-1}(I_i^{\mathrm{Ent}}(A)/2),1-\mathrm{Ent}^{-1}(I_i^{\mathrm{Ent}}(A)/2)]}\frac{h(x)}{\mathrm{Ent}(x)}\right) I_i^{\mathrm{Ent}}(A)/2 \\ &\geq \delta\cdot I_i^{\mathrm{Ent}}(A)/2,\end{aligned}$$

where the last step follows from the fact that $\vartheta(x) \leq \mathrm{Ent}^{-1}(x)$ for $x \leq 1/2$ which is easy to verify. □

4.3.2 Relation between geometric influences and h-influences for log-concave measures

It is straightforward to check the following relation between the geometric influences and the h-influences for monotone sets. The proof follows immediately from Observation (2) in the proof of Proposition 4.1.1.

Lemma 4.3.4. *Consider the product space $(\mathbb{R}^n, \nu^{\otimes n})$ where ν has a continuous density λ. Let Λ denote the cumulative distribution function of ν. Then for any monotone set $A \subseteq \mathbb{R}^n$,*

$$I_i^{\mathcal{G}}(A) = I_i^h(A) \qquad \forall 1 \leq i \leq n,$$

where $h(t) = \lambda(\Lambda^{-1}(t))$ when A is decreasing and $h(t) = \lambda(\Lambda^{-1}(1-t))$ when A is increasing. Here Λ^{-1} denotes the unique inverse of the function Λ.

Using Lemma 4.3.1 and Lemma 4.3.4, we can obtain a monotonization lemma for geometric influences that holds if the underlying measure has a log-concave density. In order to show this, we use the following isoperimetric inequality satisfied by log-concave distributions (see, for example, [27]).

Theorem 4.3.5. *Let ν have a log-concave density λ and let Λ be the corresponding cumulative distribution function. Denote the (unique) inverse of the function Λ by Λ^{-1}. Fix any $t \in (0,1)$. Then in the class of all Borel-measurable sets of ν-measure t, the extremal sets are intervals of the form $(-\infty, a]$ or $[a, \infty)$ for some $a \in \mathbb{R}$. That is, for $t \in (0,1)$ and for every Borel-measurable set $A \subseteq R$ with $\nu(A) = t$,*

$$\nu(A + [-r, r]) \geq \min\left\{\Lambda(\Lambda^{-1}(t) + r), 1 - \Lambda(\Lambda^{-1}(1-t) - r)\right\} \qquad \forall r > 0. \quad (4.16)$$

If λ is symmetric (around the median), then the above expression is simplified to

$$\nu(A + [-r, r]) \geq \Lambda(\Lambda^{-1}(t) + r) \qquad \forall r > 0. \tag{4.17}$$

Now we are ready to present the monotonization lemma.

Lemma 4.3.6. *Consider the product measure $\nu^{\otimes n}$ on $\mathbb{R}^n$ where ν is a probability distribution with a continuous symmetric log-concave density λ satisfying $\lim_{|z|\to\infty} \lambda(z) = 0$. Then for any Borel set $A \subset \mathbb{R}^n$,*

(i) $I_i^{\mathcal{G}}(A) \geq I_i^h(A)$ for all $1 \leq i \leq n$, where $h(t) = \lambda(\Lambda^{-1}(t))$.

(ii) There exists an increasing set B such that $\nu^{\otimes n}(B) = \nu^{\otimes n}(A)$ and

$$I_i^{\mathcal{G}}(B) \leq I_i^{\mathcal{G}}(A) \qquad \text{for all } \ 1 \leq i \leq n.$$

Proof. Let Λ be the cumulative distribution of ν. Fix $x \in \mathbb{R}^n$. By Theorem 4.3.5, we have, for all $r > 0$,

$$\frac{\nu(A_i^x + [-r, r]) - \nu(A_i^x)}{r} \geq \frac{\Lambda(\Lambda^{-1}(\nu(A_i^x)) + r) - \Lambda(\Lambda^{-1}(\nu(A_i^x)))}{r}.$$

Taking limit of the both sides as $r \to 0^+$, we obtain

$$m_\nu(A_i^x) \geq \lambda(\Lambda^{-1}(\nu(A_i^x))) = h(\nu(A_i^x)),$$

which implies the first part of the lemma.

For a proof of the second part, we start by noting that the assumptions on ν imply that h is concave and continuous. Thus we can invoke Lemma 4.3.1 to find an increasing set B such that $\nu^{\otimes n}(B) = \nu^{\otimes n}(A)$ and $I_i^h(B) \leq I_i^h(A)$ for all $1 \leq i \leq n$. By the first part of the lemma, $I_i^h(A) \leq I_i^{\mathcal{G}}(A)$ for all $1 \leq i \leq n$. On the other hand, it follows from Lemma 4.3.4 that $I_i^{\mathcal{G}}(B) = I_i^h(B)$ for all $1 \leq i \leq n$. Hence,

$$I_i^{\mathcal{G}}(B) = I_i^h(B) \leq I_i^h(A) \leq I_i^{\mathcal{G}}(A),$$

as asserted. □

To keep our exposition simple, we will restrict our attention to an important family of log-concave distributions known as Boltzmann measures for the rest of the section. We mention in passing that some of the techniques that we are going to develop can be applied to other log-concave measures with suitable isoperimetric properties.

4.3.3 Lower bounds on geometric influences for Boltzmann measures

Definition 4.3.2 (Boltzmann Measure)**.** *The density of the Boltzmann measure μ_ρ with parameter $\rho \geq 1$ is given by*

$$\phi_\rho(x) := \frac{1}{2\Gamma(1+1/\rho)} e^{-|x|^\rho} dx, \quad x \in \mathbb{R}.$$

Note that $\rho = 2$ corresponds to the Gaussian measure with variance $1/2$ while $\rho = 1$ gives the two-sided exponential measure.

The following estimates on the tail probability of Boltzmann measures are well-known and easy to verify.

Lemma 4.3.7. *Let Φ_ρ denote the cumulative distribution function of the Boltzmann distribution with parameter ρ. Then for $z > 0$, we have*

$$\frac{1}{2\rho\Gamma(1+1/\rho)} \left(\frac{1}{z^{\rho-1}} - \frac{\rho-1}{z^\rho} \right) e^{-|z|^\rho} \leq 1 - \Phi_\rho(z) \leq \frac{1}{2\rho\Gamma(1+1/\rho)} \frac{1}{z^{\rho-1}} e^{-|z|^\rho}.$$

In particular,

$$\phi_\rho(\Phi_\rho^{-1}(x)) \asymp x(1-x)(-\log(x(1-x)))^{(\rho-1)/\rho} \tag{4.18}$$

for x close to zero or one.

It follows from Lemma 4.3.6(i) and Lemma 4.3.7 that for Boltzmann measures, the geometric influences lie between previously studied h-influences. On the one hand, they are greater than Variance influences (i.e., h-influences with $h(t) = t(1-t)$), that were studied in, e.g., [83, 110]. On the other hand, for monotone sets they are smaller than the Entropy-influences.

It is well-known that there is no analogue of the KKL influence bound for the Variance-influence, and a tight lower bound on the maximal Variance-influence is the trivial bound:

$$\max_{1 \leq i \leq n} I_i^{\mathrm{Var}}(A) \geq ct(1-t)/n,$$

where t is the measure of the set A. This inequality is an immediate corollary of the Efron-Stein inequality (see, e.g., [129]). On the other hand, the analogue of the KKL bound proved in [95] holds only for h-influences with $h(t) \geq \mathrm{Ent}(t)$. In order to show KKL-type lower bounds for geometric influences, we use the following two results.

The first result is a dimension-free isoperimetric inequality for the Boltzmann measures.

Lemma 4.3.8 ([16])**.** *Fix $\rho > 1$ and let μ_ρ denote the Boltzmann measure with parameter ρ. Then there exists a constant $k = k(\rho) > 0$ such that for any $n \geq 1$ and*

any measurable $A \in \mathbb{R}^n$, we have

$$\mu_\rho^{\otimes n}(A + [-r, r]^n) \geq \mu_\rho\{(-\infty, \Phi_\rho^{-1}(t) + kr]\}, \quad t = \mu_\rho^{\otimes n}(A).$$

The second key ingredient is a simple corollary of Lemma 4.3.3.

Lemma 4.3.9. *Consider the product spaces $(\mathbb{R}^n, \mu_\rho^{\otimes n})$, where μ_ρ denotes the Boltzmann measure with parameter $\rho > 1$. For any $A \subset \mathbb{R}^n$ and for all $1 \leq i \leq n$,*

$$I_i^{\mathcal{G}}(A) \geq c I_i^{\mathrm{Ent}}(A) \left(-\log(I_i^{\mathrm{Ent}}(A))\right)^{-1/\rho},$$

where $c = c(\rho) > 0$ is a universal constant.

Proof. In view of Lemma 4.3.6, it is sufficient to prove that

$$I_i^{h}(A) \geq c I_i^{\mathrm{Ent}}(A) \left(-\log(I_i^{\mathrm{Ent}}(A))\right)^{-1/\rho},$$

for $h(x) := \phi_\rho(\Phi_\rho^{-1}(x))$. This indeed follows immediately from Lemma 4.3.3 using the estimate on $h(x)$ given in equation (4.18). □

Now we are ready to prove the KKL-type lower bounds. We start with an analogue of the KKL theorem [91].

Theorem 4.3.10. *Consider the product spaces $(\mathbb{R}^n, \mu_\rho^{\otimes n})$, where μ_ρ denotes the Boltzmann measure with parameter $\rho > 1$. There exists a constant $c = c(\rho) > 0$ such that for all $n \geq 1$ and for any Borel-measurable set $A \subset \mathbb{R}^n$ with $\nu^{\otimes n}(A) = t$, we have*

$$\max_{1 \leq i \leq n} I_i^{\mathcal{G}}(A) \geq c t(1-t) \frac{(\log n)^{1-1/\rho}}{n}.$$

Proof. The proof is divided into two cases, according to $\nu^{\otimes n}(A) = t$. If $t(1-t)$ is not very small, the proof uses Lemma 4.3.6 and Lemma 4.3.9. If $t(1-t)$ is very small, the proof relies on Lemma 4.3.6 and Lemma 4.3.8. We note that the same division into cases appears in the proof of the KKL theorem [91]: the core of the proof is the case where $t(1-t)$ is not "too small", and the other case follows immediately from the Edge Isoperimetric Inequality on the cube.

Case A: $t(1-t) > n^{-1}$. By Theorem 4.3.2, there exists $1 \leq i \leq n$, such that

$$I_i^{\mathrm{Ent}}(A) \geq c t(1-t) \frac{\log n}{n}.$$

Since $t(1-t) > 1/n$, it follows from Lemma 4.3.9 that

$$I_i^{\mathcal{G}}(A) \geq c I_i^{\mathrm{Ent}}(A) \left(-\log(I_i^{\mathrm{Ent}}(A))\right)^{-1/\rho} \geq c' t(1-t) \frac{\log n}{n} \cdot (\log n)^{-1/\rho},$$

where c' is a universal constant, as asserted.

Case B: $t(1-t) \leq n^{-1}$. In view of Lemma 4.3.6, we can assume w.l.o.g. that the set A is increasing. In that case, by Proposition 4.1.1, we have

$$\sum_{i=1}^{n} I_i^{\mathcal{G}}(A) = \liminf_{r \to 0+} \frac{\mu_\rho^{\otimes n}(A + [-r,r]^n) - \mu_\rho^{\otimes n}(A)}{r}.$$

By Lemma 4.3.8,

$$\liminf_{r \to 0+} \frac{\mu_\rho^{\otimes n}(A + [-r,r]^n) - \mu_\rho^{\otimes n}(A)}{r} \geq k\phi_\rho(\Phi_\rho^{-1}(t)). \tag{4.19}$$

Since in this case $t(1-t) \leq n^{-1}$, it follows from Lemma 4.3.7 that

$$\sum_{i=1}^{n} I_i^{\mathcal{G}}(A) \geq k\phi_\rho(\Phi_\rho^{-1}(t)) \geq k' t(1-t)(\log n)^{(\rho-1)/\rho},$$

for some constant $k'(\rho) > 0$. This completes the proof. □

Theorem 4.1.2 is an immediate consequence of Theorem 4.3.10. The derivation of Corollary 4.1.5 from Theorem 4.1.2 and Proposition 4.1.4 is exactly the same as the proof of Theorem 2.1 in [77] (which is the analogous result for Bernoulli measures on the discrete cube), and thus is omitted here.

We conclude this section with several analogues of results for influences on the discrete cube. In the theorem below, Part (1) corresponds to Talagrand's lower bound on the vector of influences [135], Part (2) corresponds to a variant of the KKL theorem for functions with low influences established in [77], Part (3) corresponds to Friedgut's characterization of functions with a low influence sum [76], and Part (4) corresponds to Hatami's characterization of functions with a low influence sum in the continuous case [83]. Statements (1), (3), and (4) of the theorem follow immediately using Lemma 4.3.9 from the corresponding statements for the Entropy-influence proved in [95], and Statement (2) is an immediate corollary of Statement (1).

Theorem 4.3.11. *Consider the product spaces $(\mathbb{R}^n, \mu_\rho^{\otimes n})$, where μ_ρ denotes the Boltzmann measure with parameter $\rho > 1$. For all $n \geq 1$, for any Borel-measurable set $A \subset \mathbb{R}^n$, and for all $\alpha > 0$, we have:*

1. *If $\mu_\rho^{\otimes n}(A) = t$, then*

$$\sum_{i=1}^{n} \frac{I_i^{\mathcal{G}}(A)}{(-\log I_i^{\mathcal{G}}(A))^{1-1/\rho}} \geq c_1 t(1-t),$$

2. *If* $\mu_\rho^{\otimes n}(A) = t$ *and* $\max_{1\le i\le n} I_i^{\mathcal{G}}(A) \le \alpha$, *then*

$$\sum_{i=1}^{n} I_i^{\mathcal{G}}(A) \ge c_1 t(1-t)(-\log\alpha)^{1-1/\rho},$$

3. *If* A *is monotone and* $\sum_{i=1}^n I_i^{\mathcal{G}}(A)(-\log I_i^{\mathcal{G}}(A))^{1/\rho} = s$, *then there exists a set* $B \subset \mathbb{R}^n$ *such that* 1_B *is determined by at most* $\exp(c_2 s/\epsilon)$ *coordinates and* $\mu_\rho^{\otimes n}(A \triangle B) \le \epsilon$,

4. *If* $\sum_{i=1}^n I_i^{\mathcal{G}}(A)(-\log I_i^{\mathcal{G}}(A))^{1/\rho} = s$, *then there exists a set* $B \subset \mathbb{R}^n$ *such that* 1_B *can be represented by a decision tree of depth at most* $\exp(c_3 s/\epsilon^2)$ *and* $\mu_\rho^{\otimes n}(A \triangle B) \le \epsilon$,

where c_1, c_2, *and* c_3 *are positive constants which depend only on* ρ.

See, e.g., [83] for the definition of a decision tree. Theorem 4.1.3 is a special case of Statements (1) and (3) of Theorem 4.3.11 obtained for $\rho = 2$.

4.3.4 A remark on geometric influences for more general measures

It's worth mentioning that Theorem 4.3.10 and 4.3.11 hold for any measure ν on $\mathbb{R}$ which is absolutely continuous with respect to the lebesgue measure and there exist constants $\rho \ge 1, a > 0$ such that the isoperimetric function $\mathcal{I}_\nu$ of ν satisfies

$$\mathcal{I}_\nu(t) \ge a\min(t, 1-t)(-\log\min(t, 1-t))^{1-1/\rho}, \quad t \in [0,1].$$

The proofs are exactly similar to those given for Boltzmann measures except the following remarks.

Lemma 4.3.6(i) now holds with $h(t) = \mathcal{I}_\nu(t)$. Lemma 4.3.6(ii) does not hold in general but this is not a problem since for the proof of Theorem 4.3.10 we only need the first part of the lemma. Instead of Lemma 4.3.8, we now use the the following dimension-free isoperimetric inequality (see [16]) of the product measure $\nu^{\otimes n}$: for all $n \ge 1$ and $A \subseteq \mathbb{R}^n$ measurable,

$$\liminf_{r\to 0+} \frac{\nu^{\otimes n}(A + [-r,r]^n) - \nu^{\otimes n}(A)}{r} \ge \frac{a}{K}\min(t, 1-t)(-\log\min(t, 1-t))^{1-1/\rho},$$

where $t = \nu^{\otimes n}(A)$ and $K > 0$ is a universal constant.

4.4 Boundaries of transitive sets under uniform enlargement

It follows from the classical Gaussian isoperimetric inequality by Tsirelson and Sudakov [133], Borell [40] (see also [27]) that for the Gaussian case, in any dimension, the half spaces are extremal under uniform enlargement, which implies that the boundary measure of any measurable set $A \subset \mathbb{R}^n$ with $\Phi^{\otimes n}(A) = t$ obeys the following lower bound:

$$\liminf_{r \downarrow 0} \frac{\Phi^{\otimes n}(A + [-r, r]^n) - \Phi^{\otimes n}(A)}{r} \geq \phi(\Phi^{-1}(t)), \tag{4.20}$$

and the bound is achieved when A is a half-space.

In this section we consider the same isoperimetric problem under an additional symmetry condition:

Find a lower bound on the boundary measure (under uniform enlargement) of sets in $\mathbb{R}^n$ that are transitive.

The invariance under permutation condition rules out candidates like the half-spaces and one might expect that under this assumption, a set should have "large" boundary. This intuition is confirmed by Theorem 4.1.6. In this section we prove a stronger version of this theorem that holds for all Boltzmann measures.

Theorem 4.4.1. *Consider the product spaces $(\mathbb{R}^n, \mu_\rho^{\otimes n})$, where μ_ρ denotes the Boltzmann measure with parameter $\rho > 1$. There exists a constant $c = c(\rho) > 0$ such that the following holds for all $n \geq 1$:*

For any transitive Borel-measurable set $A \subset \mathbb{R}^n$, we have

$$\liminf_{r \downarrow 0} \frac{\mu_\rho^{\otimes n}(A + [-r, r]^n) - \mu_\rho^{\otimes n}(A)}{r} \geq ct(1-t)(\log n)^{1-1/\rho},$$

where $t = \mu_\rho^{\otimes n}(A)$.

The transitivity assumption on A implies that Theorem 4.4.1 is an immediate consequence of Theorem 4.3.10, once we establish the following lemma.

Lemma 4.4.2. *Let λ be a continuous symmetric log-concave density on $\mathbb{R}$. Let A be any Borel-measurable subset of $\mathbb{R}^n$. Then*

$$\liminf_{r \downarrow 0} \frac{\nu^{\otimes n}(A + [-r, r]^n) - \nu^{\otimes n}(A)}{r} \geq \sum_{i=1}^{n} I_i^h(A),$$

where $h(x) = \lambda(\Lambda^{-1}(x))$ for all $x \in [0, 1]$.

Proof. The proof is similar to the proof of Proposition 4.1.1. For all $1 \le i \le n$, define

$$B_r^i = A + [-r, r]^{i-1} \times \{0\}^{n-i+1}.$$

Like in the proof of Proposition 4.1.1, it is sufficient to show that for each i,

$$\liminf_{r \downarrow 0} \frac{\nu^{\otimes n}(B_r^i + \{0\}^{i-1} \times [-r, r] \times \{0\}^{n-i}) - \nu^{\otimes n}(B_r^i)}{r} \ge I_i^h(A). \tag{4.21}$$

Note that for all $x \in \mathbb{R}^n$, both $\nu^{\otimes n}(B_r^i)$ and $\nu((B_r^i)_i^x)$ are increasing as functions of r, and thus, they tend to some limit as $r \searrow 0$. Furthermore, we can assume that $\nu^{\otimes n}(\bar{A} \setminus A) = 0$, since otherwise,

$$\liminf_{r \downarrow 0} \frac{\nu^{\otimes n}(A + [-r, r]^n) - \nu^{\otimes n}(A)}{r} \ge \liminf_{r \downarrow 0} \frac{\nu^{\otimes n}(\bar{A} \setminus A)}{r} \to \infty.$$

Therefore,

$$\nu^{\otimes n}(B_r^i) \searrow \nu^{\otimes n}(\bar{A}) = \nu^{\otimes n}(A),$$

and

$$\nu((B_r^i)_i^x) \searrow \nu(A_i^x), \tag{4.22}$$

for almost every $x \in \mathbb{R}^n$ (w.r.t. the measure $\nu^{\otimes n}$).

Now observe that by the one-dimensional isoperimetric inequality for symmetric log-concave distributions (Theorem 4.3.5),

$$\begin{aligned} \nu^{\otimes n}(B_r^i + \{0\}^{i-1} \times [-r, r] \times \{0\}^{n-i}) &= \mathbb{E}_x\, \nu((B_r^i)_i^x + [-r, r]) \\ &\ge \mathbb{E}_x\, \Lambda(\Lambda^{-1}(\nu((B_r^i)_i^x)) + r). \end{aligned}$$

Therefore, using the Mean Value Theorem like in the proof of Lemma 4.1.1, we get

$$\begin{aligned} \liminf_{r \downarrow 0} &\frac{\nu^{\otimes n}(B_r + \{0\}^{i-1} \times [-r, r] \times \{0\}^{n-i}) - \nu^{\otimes n}(B_r)}{r} \\ &\ge \liminf_{r \downarrow 0} \mathbb{E}_x \inf_{z \in [\Lambda^{-1}(\nu((B_r^i)_i^x)), \Lambda^{-1}(\nu((B_r^i)_i^x)) + r]} \lambda(z). \end{aligned} \tag{4.23}$$

Finally, by (4.22), for almost every $x \in \mathbb{R}^n$,

$$\lim_{r \downarrow 0} \inf_{z \in [\Lambda^{-1}(\nu((B_r^i)_i^x)), \Lambda^{-1}(\nu((B_r^i)_i^x)) + r]} \lambda(z) = \lambda(\Lambda^{-1}(\nu(A_i^x))),$$

and thus, by the Dominated Convergence Theorem,

$$\liminf_{r \downarrow 0} \mathbb{E}_x \inf_{z \in [\Lambda^{-1}(\nu((B_r^i)_i^x)), \Lambda^{-1}(\nu((B_r^i)_i^x)) + r]} \lambda(z) = \mathbb{E}_x\, \lambda(\Lambda^{-1}(\nu(A_i^x))) = I_i^h(A).$$

This completes the proof of the lemma, and thus also the proof of Theorem 4.4.1. □

4.4.1 Tightness of Theorems 4.3.10, 4.3.11, and 4.4.1

We conclude this section with showing that Theorems 4.3.10, 4.3.11, and 4.4.1 are tight (up to constant factors) among sets with constant measure, which we set for convenience to be 1/2. We demonstrate this by choosing an appropriate sequence of '*one-sided boxes*'.

Proposition 4.4.3. *Consider the product spaces $(\mathbb{R}^n, \mu_\rho^{\otimes n})$, where μ_ρ denotes the Boltzmann measure with parameter $\rho \geq 1$. Let $B_n := (-\infty, a_n]^n$ where a_n is chosen such that $\Phi_\rho(a_n)^n = 1/2$. Then there exists a constant $c = c(\rho)$ such that*

$$I_i^{\mathcal{G}}(B_n) \leq c \cdot \frac{(\log n)^{1-1/\rho}}{n},$$

for all $1 \leq i \leq n$.

Proof. Fix an i. By elementary calculation,

$$I_i^{\mathcal{G}}(B_n) = \Phi_\rho(a_n)^{n-1}\phi_\rho(a_n) = (1/2)^{(n-1)/n}\phi_\rho(a_n).$$

Note that $1 - \Phi_\rho(a_n) \asymp n^{-1}$, and thus, by Lemma 4.3.7, $a_n \asymp (\log n)^{1/\rho}$. Furthermore, since by Lemma 4.3.7, $\phi_\rho(z) \asymp z^{\rho-1}(1 - \Phi_\rho(z))$ for large z, we have $I_i^{\mathcal{G}}(B_n) \asymp n^{-1}(\log n)^{1-1/\rho}$, as asserted. □

The tightness of Theorem 4.3.10 and Theorem 4.3.11 (1) follows immediately from Proposition 4.4.3. The tightness of Theorem 4.4.1 follows using Proposition 4.1.1 since B is monotone. The tightness of Theorem 4.3.11 (2) and the tightness in s in Theorem 4.3.11 (3) and Theorem 4.3.11 (4) follows by considering the subset $B_k \times \mathbb{R}^{n-k} \subset \mathbb{R}^n$.

4.5 Geometric influences under rotation

Consider the product Gaussian measure $\mu^{\otimes n}$ on $\mathbb{R}^n$. In Section 4.3 we obtained lower bounds on the sum of geometric influences, and in particular we showed that for a transitive set $A \subset \mathbb{R}^n$, the sum is at least $\Omega(t(1-t)\sqrt{\log n})$, where $t = \mu^{\otimes n}(A)$.

In this section we consider a different symmetry group, the group of rotations of $\mathbb{R}^n$. The interest in this group comes from the fact that the Gaussian measure is invariant under rotations while the influence sum is not.

Indeed, a half space of measure 1/2 may have influence sum as small as of order 1 when it is aligned with one of the axis and as large as of order $\sqrt{n}$ when it is aligned with the diagonal direction $(1, 1, \ldots, 1)$.

In this section we show that under some mild conditions (that do not contain any invariance assumption), rotation allows to increase the sum of geometric influences up to $\Omega(t(1-t)\sqrt{-\log(t(1-t))}\sqrt{n})$. The dependence on n in this lower bound is tight for several examples, including half-spaces and L^2-balls. We note that on the other extreme, rotation cannot decrease the sum of geometric influences below $\Omega\left(t(1-t)\sqrt{-\log(t(1-t))}\right)$, as follows from a combination of Proposition 4.1.1, Lemma 4.3.6(ii) and the isoperimetric inequality (4.20).

Definition 4.5.1. *Let $B(x,r) := \{y \in \mathbb{R}^n : \|y-x\|_2 < r\}$ be the open ball in $\mathbb{R}^n$ with center at x and radius r and let $\bar{B}(x,r)$ be the corresponding closed ball. For $\varepsilon > 0$ and $A \subseteq R^n$, define*

$$A_\varepsilon := \{x \in A : \bar{B}(x,\varepsilon) \cap A^c = \emptyset\}, \qquad \text{and} \qquad A^\varepsilon := \{x \in \mathbb{R}^n : B(x,\varepsilon) \cap A \neq \emptyset\}.$$

Finally, denote by $\mathcal{J}_n$ the collection of all measurable sets $B \subseteq \mathbb{R}^n$ for which there exists $\delta > 0$ such that for all $0 < \varepsilon < \delta$, we have

$$(B_\varepsilon)^{2\varepsilon} \supseteq B. \tag{4.24}$$

The crucial ingredient in the proof Theorem 4.1.7 is a lemma asserting that under the conditions of the theorem, an enlargement of A by a *random* rotation of the cube $[-r,r]^n$ increases $\mu^{\otimes n}(A)$ significantly.

Notation 4.5.1. *Let $O = O(n,\mathbb{R})$ be the set of all orthogonal transformations on $\mathbb{R}^n$, and let ν be the (unique) Haar measure on O. Denote by M a random element of O distributed according to the measure ν.*

Lemma 4.5.2. *There exists a constant $K > 0$ such that for any $A \in \mathcal{J}_n$, we have*

$$\mathbb{E}_{M\sim\nu}\left[\mu^{\otimes n}(A + M^{-1}(Kn^{-1/2}[-r,r]^n))\right] \geq \mu^{\otimes n}(A) + \frac{1}{2}\mu^{\otimes n}(A^{r/3} \setminus A),$$

for all sufficiently small $r > 0$ (depending on A).

First we show that Lemma 4.5.2 implies Theorem 4.1.7.

Proof. Note that for any $g \in O$, $g(A)$ is convex, and that $\mu^{\otimes n}$ is invariant under g. Thus by Remark 10 after Proposition 4.1.1, we have for any convex set A,

$$\begin{aligned}\sum_{i=1}^n I_i^{\mathcal{G}}(g(A)) &= \lim_{r\to 0+} \frac{\mu^{\otimes n}(g(A) + [-r,r]^n) - \mu^{\otimes n}(g(A))}{r} \\ &= \lim_{r\to 0+} \frac{\mu^{\otimes n}(A + g^{-1}([-r,r]^n)) - \mu^{\otimes n}(A)}{r}.\end{aligned}$$

Furthermore, note that for any $g \in O$,

$$\lim_{r\to 0+} \frac{\mu^{\otimes n}(A + g^{-1}([-r,r]^n)) - \mu^{\otimes n}(A)}{r} \leq \lim_{r\to 0+} \frac{\mu^{\otimes n}(A + \sqrt{n}[-r,r]^n) - \mu^{\otimes n}(A)}{r}$$
$$= \sqrt{n} \times \sum_{i=1}^{n} I_i^{\mathcal{G}}(A).$$

Therefore, by the Dominated Convergence Theorem,

$$\mathbb{E}_{M\sim\nu}\left[\sum_{i=1}^{n} I_i^{\mathcal{G}}(M(A))\right] = \lim_{r\to 0+} \frac{\mathbb{E}_{M\sim\nu}\left[\mu^{\otimes n}(A + M^{-1}([-r,r]^n))\right] - \mu^{\otimes n}(A)}{r}. \quad (4.25)$$

By Lemma 4.5.2, we have (for a sufficiently small r):

$$\mathbb{E}_{M\sim\nu}\left[\mu^{\otimes n}(A + M^{-1}(Kn^{-1/2}[-r,r]^n))\right] - \mu^{\otimes n}(A) \geq \frac{1}{2}\mu^{\otimes n}(A^{r/3} \setminus A).$$

By the standard Gaussian isoperimetric inequality,

$$\mu^{\otimes n}(A^{r/3} \setminus A) \geq \mu((-\infty, \Phi^{-1}(t) + r/3]).$$

Substituting into equation (4.25), we get

$$\mathbb{E}_{M\sim\nu}\left[\sum_{i=1}^{n} I_i^{\mathcal{G}}(M(A))\right] \geq \limsup_{r\to 0+} \frac{\mu((-\infty, \Phi^{-1}(t) + K^{-1}n^{1/2}r/3])}{2r}$$
$$\geq \frac{\sqrt{n}}{6K}\phi(\Phi^{-1}(t)) \geq c\sqrt{n}t(1-t)\sqrt{-\log(t(1-t))},$$

for some constant $c > 0$. Thus, there exists at least one orthogonal transformation $g \in O$ such that

$$\sum_{i=1}^{n} I_i^{\mathcal{G}}(g(A)) \geq ct(1-t)\sqrt{-\log(t(1-t))} \times \sqrt{n},$$

as asserted. □

Now we present the proof of Lemma 4.5.2.

Proof. By Fubini's theorem, we have

$$\mathbb{E}_{M\sim\nu}\left[\mu^{\otimes n}(A + M^{-1}(Kn^{-1/2}[-r,r]^n))\right]$$
$$= \mathbb{E}_{M\sim\nu}\left[\mu^{\otimes n}\{x \in \mathbb{R}^n : x = y + z, y \in A, z \in M^{-1}(Kn^{-1/2}[-r,r]^n)\}\right]$$
$$= \mathbb{E}_{x\sim\mu^{\otimes n}}\left[\nu\{g \in O : x = y + z, y \in A, z \in g^{-1}(Kn^{-1/2}[-r,r]^n)\}\right]. \quad (4.26)$$

Since each $x \in A$ can be trivially represented as $y + z$ with $y = x \in A, z = 0 \in g^{-1}(Kn^{-1/2}[-r, r]^n)$ for any $g \in O$, the assertion of the lemma would follow immediately from equation (4.26) once we show that for all $x \in A^{r/3} \setminus A$,

$$\nu\{g \in O : x = y + z, y \in A, z \in g^{-1}(Kn^{-1/2}[-r, r]^n)\} \geq 1/2. \tag{4.27}$$

Since $A \in \mathcal{J}_n$, we can choose r sufficiently small such that $A \subset (A_{r/3})^{2r/3}$, and thus $A^{r/3} \subset (A_{r/3})^r$. Therefore, for any $x \in A^{r/3} \setminus A$, there exists $y \in A_{r/3}$, such that $||x-y||_2 < r$. If there exists $y' \in B(y, r/3)$ such that $x-y' \in g^{-1}(Kn^{-1/2}[-r, r]^n)$, then x can be represented as $y'+(x-y')$, as required in the left hand side of equation (4.27). Therefore, it is sufficient to prove the following claim:

Claim 4.5.3. *For any $x, y \in \mathbb{R}^n$ such that $||x - y||_2 < r$,*

$$\nu\Big\{g \in O : \exists\ y' \in B(y, r/3) \text{ such that } x - y' \in g^{-1}(Kn^{-1/2}[-r, r]^n)\Big\} \geq 1/2.$$

Proof of the claim. Fix $x, y \in \mathbb{R}^n$ such that $||x - y||_2 < r$. We have

$$\begin{aligned}
&\Big\{g \in O : \exists\ y' \in B(y, r/3) \text{ such that } x - y' \in g^{-1}(Kn^{-1/2}[-r, r]^n)\Big\} \\
&= \Big\{g \in O : \exists\ y' \in B(y, r/3) \text{ such that } g(x - y') \in Kn^{-1/2}[-r, r]^n\Big\} \\
&= \Big\{g \in O : \exists\ y'' \in B(0, r/3) \text{ such that } g(x - y) - y'' \in Kn^{-1/2}[-r, r]^n\Big\} \\
&= \Big\{g \in O : \inf_{y'' \in B(0,r/3)} ||g(x - y) - y''||_\infty \leq Kn^{-1/2}r\Big\}.
\end{aligned}$$

Note that

$$\nu\Big\{g \in O : \inf_{y'' \in B(0,r/3)} ||g(x - y) - y''||_\infty \leq Kn^{-1/2}r\Big\} \tag{4.28}$$

is invariant under rotation of the vector $(x - y)$, and in particular,

$$(4.28) = \nu\Big\{g \in O : \inf_{y'' \in B(0,r/3)} ||g(||x - y||_2 \times e_1) - y''||_\infty \leq Kn^{-1/2}r\Big\},$$

where $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^n$ is the unit vector along the first coordinate axis.

A well-known property of the Haar measure says that if $M \in O$ is distributed according to ν, then any column of M is distributed like a normalized vector of independent standard Gaussians. That is,

$$M_{column} \sim \frac{Z}{\|Z\|_2},$$

where $Z = (Z_1, \ldots, Z_n)$ is a random n-vector with i.i.d. standard Gaussian entries.

Thus, $M(||x-y||_2 \times e_1)$ is distributed like $||x-y||_2 \times Z/||Z||_2$. Therefore, we have

$$\begin{aligned}(4.28) &= \mathbf{P}_{Z\sim\mu^{\otimes n}}\Big(\inf_{y''\in B(0,r/3)} || \ ||x-y||_2 \times \frac{Z}{||Z||_2} - y''||_\infty \le Kn^{-1/2}r\Big) \\ &\ge \mathbf{P}_{Z\sim\mu^{\otimes n}}\Big(\inf_{y'''\in B(0,1/3)} ||\frac{Z}{||Z||_2} - y'''||_\infty \le Kn^{-1/2}\Big).\end{aligned}$$

Note that if $Z \in \mathbb{R}^n$ satisfies

$$\frac{\sum_i Z_i^2 1_{|Z_i|/||Z||_2 > Kn^{-1/2}}}{||Z||_2^2} < 1/9,$$

then the vector y''' defined by $y'''_i = (Z_i \cdot 1_{|Z_i|/||Z||_2>Kn^{-1/2}})/||Z||_2$ satisfies

$$y''' \in B(0,1/3) \qquad \text{and} \qquad ||\frac{Z}{||Z||_2} - y'''||_\infty \le Kn^{-1/2}.$$

Hence,

$$\begin{aligned}(4.28) &\ge \mathbf{P}_{Z\sim\mu^{\otimes n}}\Big(\inf_{y'''\in B(0,1/3)} ||\frac{Z}{||Z||_2} - y'''||_\infty \le Kn^{-1/2}\Big) \\ &\ge \mathbf{P}_{Z\sim\mu^{\otimes n}}\Big(\frac{\sum_i Z_i^2 1_{|Z_i|/||Z||_2>Kn^{-1/2}}}{||Z||_2^2} < 1/9\Big).\end{aligned}$$

Finally, by the Markov inequality,

$$\mathbf{P}_{Z\sim\mu^{\otimes n}}\left[\sum_{i:|Z_i|>K/2} Z_i^2 \ge \frac{n}{36}\right] \le \frac{n \times \left[\mathbb{E}\, Z_1^2 1_{\{|Z_1|>K/2\}}\right]}{n/36} \le 1/4$$

for sufficiently large $K > 0$, and by the concentration of norm of a Gaussian vector, $\mathbf{P}[\|Z\|_2 > \sqrt{n}/2] \ge 3/4$. Therefore,

$$\begin{aligned}(4.28) &\ge \mathbf{P}_{Z\sim\mu^{\otimes n}}\Big(\frac{\sum_i Z_i^2 1_{|Z_i|/||Z||_2>Kn^{-1/2}}}{||Z||_2^2} < 1/9\Big) \\ &\ge \mathbf{P}_{Z\sim\mu^{\otimes n}}\left[\left(\sum_{i:|Z_i|>K/2} Z_i^2 \le \frac{n}{36}\right) \wedge (\|Z\|_2 > \sqrt{n}/2)\right] \\ &\ge 3/4 + 3/4 - 1 = 1/2.\end{aligned}$$

This completes the proof of the claim and of Lemma 4.5.2. □

Intuitively, the condition $A \in \mathcal{J}_n$ means that the boundary of A is "sufficiently smooth". One can easily check that if $A \in \mathcal{J}_n$, then the boundary of A is a porous

set and thus has Hausdorff dimension strictly less than n (see [141] and references therein to know more about porous sets). However, this condition is far from being sufficient. Here we give a sufficient condition for a set to belong to $\mathcal{J}_n$ in terms of smoothness of its boundary.

Definition 4.5.2. *Let $A \subset \mathbb{R}^n$ be a measurable set. We write $\partial A \in C^1$ and say that the boundary of A is of class C^1 if for any point $z \in \partial A$, there exists $r = r(z) > 0$ and a one-to-one mapping ψ of $B(z,r)$ onto an open set $D = D \subseteq \mathbb{R}^n$ such that:*

- $\psi \in C^1(\bar{B}(z,r))$ *and* $\psi^{-1} \in C^1(\bar{D})$,
- $\psi(B(z,r) \cap \partial A) = D \cap \{x \in \mathbb{R}^n : x_1 = 0\}$,
- $\psi(B(z,r) \cap \mathrm{int}(A)) \subseteq (0,\infty) \times \mathbb{R}^{n-1}$.

Proposition 4.5.4. *Let $A \subset \mathbb{R}^n$ be a bounded set with $\partial A \in C^1$. Then $A \in \mathcal{J}_n$.*

Proof. Suppose on the contrary that $A \notin \mathcal{J}_n$. Then there exists a sequence $\{x^m\}_{m=1}^{\infty}$ such that $x^m \in A$ but $x^m \notin (A_{1/m})^{2/m}$. Since A is bounded, the sequence contains a subsequence $\{x^{m_k}\}$ converging to a point x^0. Clearly, $x^0 \in \partial A$.

Since $\partial A \in C^1$, we can define a new set of local coordinates $(y_1, y_2, \ldots, y_n)$ (also denoted by (y_1, y'), where $y' \in \mathbb{R}^{n-1}$), such that:

1. The point x^0 is the origin with respect to the y-coordinates,
2. There exists an open neighborhood $(-\delta_0, \delta_0) \times U \subseteq \mathbb{R} \times \mathbb{R}^{n-1}$ containing the origin and a continuously differentiable function $f : U \to \mathbb{R}_+$, such that in the y-coordinates,

$$\partial A \cap [(-\delta_0, \delta_0) \times U] = \{(f(y'), y') : y' \in U\},$$

and

$$\mathrm{int}A \cap [(-\delta_0, \delta_0) \times U] = \{(y_1, y') : y' \in U, f(y') < y_1 < \delta_0\}. \tag{4.29}$$

By the construction of the new coordinates, $f(y') \geq 0$ for all $y' \in U$, and $f(0) := f(0,0,\ldots,0) = 0$. Since $f \in C^1(U)$, it follows that $\nabla f(0) = 0$. Hence, by the continuity of the partial derivatives of f, there exists $r_0 > 0$ such that $\|\nabla f(y')\|_\infty \leq 1/(3\sqrt{n})$ for all $y' \in B_{n-1}(0, r_0) \subseteq U$.

Let $y^m = (y_1^m, (y^m)')$ be the representation of the point x^m in the y-coordinates. Find m large enough such that $1/m < \min\{\delta_0/10, r_0/10\}$, and y^m lies within $A \cap [0, \delta_0/2] \times B_{n-1}(0, r_0/2)$. Define

$$z = (z_1, z_2, \ldots, z_n) = y^m + (1.5m^{-1}, 0, \ldots, 0, 0).$$

We claim that $B(z, 1/m) \subseteq A$. This would be a contradiction to the hypothesis $y^m \notin (A_{1/m})^{2/m}$.

Note that by the choice of m, we have $z \in A$, and moreover,

$$\begin{aligned}\operatorname{dist}(z, \partial A) &\geq \operatorname{dist}\left(z, \partial A \cap [(-\delta_0, \delta_0) \times B_{n-1}(0, r_0)]\right) \\ &= \inf_{y' \in B_{n-1}(0,r_0)} \|\left(y_1^m + 1.5m^{-1}, (y^m)'\right) - (f(y'), y')\|_2. \end{aligned} \tag{4.30}$$

We would like to show that if $||(y^m)' - y'||_2$ is "small" then $|y_1^m + 1.5m^{-1} - f(y')|$ is "big", and thus in total, the right hand side of equation (4.30) cannot be "too small".

Define $w_1 := y_1^m + 1.5m^{-1} - f((y^m)')$. Note that since $y^m \in A$, it follows from equation (4.29) that $w_1 \geq 1.5m^{-1}$. By the Mean Value Theorem, for each $y' \in B_{n-1}(0, r_0)$,

$$\begin{aligned}|f((y^m)') - f(y')| &\leq \left(\sup_{y'' \in B_{n-1}(0,r_0)} \|\nabla f(y'')\|_\infty\right) \|(y^m)' - y'\|_1 \\ &\leq \frac{\|(y^m)' - y'\|_1}{3\sqrt{n}} \leq \frac{\|(y^m)' - y'\|_2}{3},\end{aligned}$$

and thus,

$$|y_1^m + 1.5m^{-1} - f(y')| = |w_1 - (f(y') - f((y^m)'))| \geq 1.5m^{-1} - \frac{\|(y^m)' - y'\|_2}{3}.$$

Consequently, if $\|(y^m)' - y'\|_2 \geq 4.5m^{-1}$, then

$$\|\left(y_1^m + 1.5m^{-1}, (y^m)'\right) - (f(y'), y')\|_2 \geq \|(y^m)' - y'\|_2 \geq 4.5m^{-1},$$

and if $\|(y^m)' - y'\|_2 < 4.5m^{-1}$, then

$$\begin{aligned}\|\left(y_1^m + 1.5m^{-1}, (y^m)'\right) - (f(y'), y')\|_2 &\geq \sqrt{\|(y^m)' - y'\|_2^2 + \left(1.5m^{-1} - \frac{\|(y^m)' - y'\|_2}{3}\right)^2} \\ &= \min_{0 \leq s < 4.5m^{-1}} \sqrt{s^2 + (1.5m^{-1} - s/3)^2} = \sqrt{\frac{81}{40}} m^{-1}.\end{aligned}$$

Combining the two cases, we get

$$\begin{aligned}\operatorname{dist}(z, \partial A) &\geq \inf_{y' \in B_{n-1}(0,r_0)} \|\left(y_1^m + 1.5m^{-1}, (y^m)'\right) - (f(y'), y')\|_2 \\ &\geq \min\left(4.5m^{-1}, \sqrt{\frac{81}{40}} m^{-1}\right) > 1/m.\end{aligned}$$

This completes the proof. □

If the condition $A \in \mathcal{J}_n$ is removed, we can prove only a weaker lower bound on the maximal sum of geometric influences that can be obtained by rotation.

Proposition 4.5.5. *Consider the product Gaussian measure $\mu^{\otimes n}$ on $\mathbb{R}^n$. For any convex set A with $\mu^{\otimes n}(A) = t$, there exists an orthogonal transformation g on $\mathbb{R}^n$ such that*

$$\sum_{i=1}^{n} I_i^{\mathcal{G}}(g(A)) \geq ct(1-t)\sqrt{-\log(t(1-t))}\frac{\sqrt{n}}{\sqrt{\log n}},$$

where $c > 0$ is a universal constant.

The proof of Proposition 4.5.5 uses a weaker variant of Lemma 4.5.2:

Lemma 4.5.6. *Let M be as defined in Notation 4.5.1. There exists a constant $K > 0$ such that for any $A \subset \mathbb{R}^n$ and for any $r > 0$, we have*

$$\mathbb{E}_{M\sim\nu}\left[\mu^{\otimes n}(A + M^{-1}(K\sqrt{\log n}\cdot n^{-1/2}[-r,r]^n))\right] \geq \mu^{\otimes n}(A) + \frac{1}{2}\mu^{\otimes n}(A^r \setminus A).$$

Proof of the lemma. By Fubini's theorem, we have

$$\begin{aligned}
&\mathbb{E}_{M\sim\nu}\left[\mu^{\otimes n}(A + M^{-1}(K\sqrt{\log n}\cdot n^{-1/2}[-r,r]^n))\right] \\
&= \mathbb{E}_{x\sim\mu^{\otimes n}}\left[\nu\{g \in O : x \in A + g^{-1}(K\sqrt{\log n}\cdot n^{-1/2}[-r,r]^n)\}\right].
\end{aligned}$$

Thus, it is sufficient to prove that for any $x \in A^r \setminus A$,

$$\nu\{g \in O : x \in A + g^{-1}(K\sqrt{\log n}\cdot n^{-1/2}[-r,r]^n)\} \geq 1/2.$$

Equivalently, it is sufficient to prove that for any $x \in B(0,r)$,

$$\nu\{g \in O : x \in g^{-1}(K\sqrt{\log n}\cdot n^{-1/2}[-r,r]^n)\} \geq 1/2.$$

We can assume w.l.o.g. that $x = r'\cdot e_1$ for some $r' < r$. By the argument used in the proof of Lemma 4.5.2, if $M \in O$ is distributed according to ν, then $M(r'\cdot e_1)$ is distributed like $r'\cdot Z/||Z||_2$, where $Z = (Z_1,\ldots,Z_n)$ is a random n-vector with i.i.d. standard Gaussian entries. Hence,

$$\begin{aligned}
&\nu\{g \in O : x \in g^{-1}(K\sqrt{\log n}\cdot n^{-1/2}[-r,r]^n)\} \\
&= \mathbf{P}_{Z\sim\mu^{\otimes n}}\left(||r'\frac{Z}{||Z||_2}||_\infty \leq K\sqrt{\log n}\cdot n^{-1/2}r\right) \\
&\geq \mathbf{P}_{Z\sim\mu^{\otimes n}}\left(||\frac{Z}{||Z||_2}||_\infty \leq K\sqrt{\log n}\cdot n^{-1/2}\right) \\
&\geq \mathbf{P}_{Z\sim\mu^{\otimes n}}\left[\left(||Z||_\infty \leq K\sqrt{\log n}/2\right) \wedge (\|Z\|_2 \geq \sqrt{n}/2)\right]. \qquad (4.31)
\end{aligned}$$

We have

$$\mathbf{P}_{Z\sim\mu^{\otimes n}}(||Z||_\infty \le K\sqrt{\log n}/2) \ge 1 - n\mathbf{P}(Z_i > K\sqrt{\log n}/2) \ge 1 - \frac{n}{\sqrt{2\pi}}\cdot n^{-K^2/8} \ge 3/4,$$

for a sufficiently big K. Therefore,

$$(4.31) \ge 3/4 + 3/4 - 1 = 1/2,$$

and this completes the proof of the lemma. $\square$

The derivation of Proposition 4.5.5 from Lemma 4.5.6 is the same as the derivation of Theorem 4.1.7 from Lemma 4.5.2.

Note that the convexity assumption on A is used only to apply Proposition 4.1.1 that relates the sum of influences to the size of the boundary w.r.t. uniform enlargement. Thus, our argument also shows that for *any* measurable set A with $\mu^{\otimes n}(A) = t$, there exists an orthogonal transformation g on $\mathbb{R}^n$ such that

$$\lim_{r\to 0+} \frac{\mu^{\otimes n}(g(A) + [-r,r]^n) - \mu^{\otimes n}(g(A))}{r} \ge ct(1-t)\sqrt{-\log(t(1-t))}\frac{\sqrt{n}}{\sqrt{\log n}},$$

where $c > 0$ is a universal constant.

Finally, we note that apparently the assertion of Proposition 4.5.5 is not optimal, and the lower bound asserted in Theorem 4.1.7 should hold for general convex sets.

Bibliography

[1] R. Albert and A.L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, 2002.

[2] David Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1(2):228–266, 1991.

[3] David Aldous. The continuum random tree I. *Ann. Probab.*, 19:1–28, 1991.

[4] David Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

[5] David Aldous and J. Michael Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.

[6] David Aldous and J. Michael Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.

[7] David Aldous and J. Michael Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.

[8] David J. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.*, 3(4):450–465, 1990.

[9] Omer Angel and Oded Schramm. Uniform infinite planar triangulations. *Comm. Math. Phys.*, 241(2-3):191–213, 2003.

[10] R. Arratia. Coalescing brownian motions on the line. Ph.D. Thesis, 1979.

[11] R. Arratia. Coalescing brownian motions and the voter model on $\mathbb{Z}$. Unpublished partial manuscript. Available from rarratia@math.usc.edu., 1981.

[12] Z. Bai and J.W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*. Science Press, 2006.

[13] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4):1119–1178, 1999.

[14] Martin T. Barlow. Diffusions on fractals. In *Lectures on probability theory and statistics (Saint-Flour, 1995)*, volume 1690 of *Lecture Notes in Math.*, pages 1–121. Springer, Berlin, 1998.

[15] M.T. Barlow and E.A. Perkins. Brownian motion on the Sierpinski gasket. *Probability Theory and Related Fields*, 79(4):543–623, 1988.

[16] F. Barthe. Infinite dimensional isoperimetric inequalities in product spaces with the supremum distance. *Journal of Theoretical Probability*, 17(2):293–308, 2004.

[17] M. Bauer and O. Golinelli. On the kernel of tree incidence matrices. *J. Integer Seq.*, 3(1):Article 00.1.4, 1 HTML document (electronic), 2000.

[18] M. Bauer and O. Golinelli. Random incidence matrices: Moments of the spectral density. *Journal of Statistical Physics*, 103(1):301–337, 2001.

[19] M. Ben-Or and N. Linial. Collective coin flipping. *Randomness and Computation*, pages 91–115, 1990.

[20] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.

[21] J. Bertoin. *Levy processes*. Cambridge University Press, 1998.

[22] Jean Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[23] Shankar Bhamidi. Universal techniques for preferential attachment: Local and global analysis, 2007. Preprint available at `http://www.stat.berkeley.edu/users/shanky/preferent.pdf`.

[24] Norman Biggs. *Algebraic graph theory*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 1993.

[25] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[26] Pavel Bleher and Alexander Its, editors. *Random Matrix Models and Their Applications*. Cambridge University Press, 2001.

[27] S. Bobkov. Extremal properties of half-spaces for log-concave distributions. *The Annals of Probability*, pages 35–48, 1996.

[28] S.G. Bobkov. Isoperimetric problem for uniform enlargement. *Studia Math*, 123(1):81–95, 1997.

[29] Béla Bollobás. *Modern graph theory*, volume 184 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[30] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[31] Béla Bollobás and Graham Brightwell. Box-spaces and random partial orders. *Trans. Amer. Math. Soc.*, 324(1):59–72, 1991.

[32] Béla Bollobás and Graham Brightwell. The height of a random partial order: concentration of measure. *Ann. Appl. Probab.*, 2(4):1009–1018, 1992.

[33] Béla Bollobás and Vladimir Nikiforov. Graphs and Hermitian matrices: eigenvalue interlacing. *Discrete Math.*, 289(1-3):119–127, 2004.

[34] Béla Bollobás and Oliver M. Riordan. Mathematical results on scale-free random graphs. In *Handbook of graphs and networks*, pages 1–34. Wiley-VCH, Weinheim, 2003.

[35] Béla Bollobás and Peter Winkler. The longest chain among random points in Euclidean space. *Proc. Amer. Math. Soc.*, 103(2):347–353, 1988.

[36] J. A. Bondy and U. S. R. Murty. *Graph theory with applications*. American Elsevier Publishing Co., Inc., New York, 1976.

[37] C. Bordenave and M. Lelarge. Resolvent of large random graphs, 2007. `arXiv:0801.0155v2`.

[38] C. Bordenave and M. Lelarge. The rank of diluted random graphs, 2009. `arXiv:arXiv:0907.4244`.

[39] Charles Bordenave. Eigenvalues of euclidean random matrices, 2006.

[40] C. Borell. The Brunn-Minkowski inequality in gauss space. *Inventiones Mathematicae*, 30(2):207–216, 1975.

[41] Phillip Botti and Russell Merris. Almost all trees share a complete set of immanantal polynomials. *J. Graph Theory*, 17(4):467–476, 1993.

[42] Phillip Botti and Russell Merris. Almost all trees share a complete set of immanantal polynomials. *J. Graph Theory*, 17(4):467–476, 1993.

[43] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, and N. Linial. The influence of variables in product spaces. *Israel Journal of Mathematics*, 77(1):55–64, 1992.

[44] Graham Brightwell. Models of random partial orders. In *Surveys in combinatorics, 1993 (Keele)*, volume 187 of *London Math. Soc. Lecture Note Ser.*, pages 53–83. Cambridge Univ. Press, Cambridge, 1993.

[45] Graham Brightwell and Sarah Goodall. The number of partial orders of fixed width. *Order*, 13(4):315–337, 1996.

[46] A. Broder. Generating random spanning trees. In *Proc. 30'th IEEE Symp. Found. Comp. Sci.*, pages 442–447, 1989.

[47] M. Camarri and J. Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electron. J. Probab*, 5(2):1–18, 2000.

[48] F. Chung, L. Lu, and V. Vu. Eigenvalues of random power law graphs. *Annals of Combinatorics*, 7(1):21–33, 2003.

[49] F. Chung, L. Lu, and V. Vu. Spectra of random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 100(11):6313–6318, 2003.

[50] Fan R. K. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1997.

[51] C. Cooper and A. Frieze. A general model of web graphs. *Random Structures and Algorithms*, 22(3):311–335, 2003.

[52] D. Cvetković, P. Rowlinson, and S. Simić. *Eigenspaces of graphs*, volume 66 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1997.

[53] Dragoš M. Cvetković, Michael Doob, Ivan Gutman, and Aleksandar Torgašev. *Recent results in the theory of graph spectra*, volume 36 of *Annals of Discrete Mathematics*. North-Holland Publishing Co., Amsterdam, 1988.

[54] Dragoš M. Cvetković, Michael Doob, and Horst Sachs. *Spectra of graphs*. Johann Ambrosius Barth, Heidelberg, third edition, 1995. Theory and applications.

[55] C. Daskalakis, R.M. Karp, E. Mossel, S. Riesenfeld, and E. Verbin. Sorting and selection in posets. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 392–401. Society for Industrial and Applied Mathematics, 2009.

[56] Percy Deift. *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*. Courant Lecture Notes. American Mathematical Society, 2000.

[57] Deepak Dhar. Entropy and phase transitions in partially ordered sets. *J. Math. Phys.*, 19(8):1711–1713, 1978.

[58] Persi Diaconis and Nathaniel Thiem. Supercharacter formulas for pattern groups. *Trans. Amer. Math. Soc.*, 361(7):3501–3533, 2009.

[59] Peter Donnelly, Steven N. Evans, Klaus Fleischmann, Thomas G. Kurtz, and Xiaowen Zhou. Continuum-sites stepping-stone models, coalescing exchangeable partitions and random trees. *Ann. Probab.*, 28(3):1063–1110, 2000.

[60] Joseph L. Doob. *Classical potential theory and its probabilistic counterpart*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1984 edition.

[61] Richard Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, second edition, 1996.

[62] Rick Durrett. *Random graph dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2007.

[63] Alexei Ermakov, Bálint Tóth, and Wendelin Werner. On some annihilating and coalescing systems. *J. Statist. Phys.*, 91(5-6):845–870, 1998.

[64] S. N. Evangelou and E. N. Economou. Spectral density singularities, level statistics, and localization in a sparse random matrix ensemble. *Phys. Rev. Lett.*, 68(3):361–364, Jan 1992.

[65] Steven N. Evans. Coalescing Markov labelled partitions and a continuous sites genetics model with infinitely many types. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(3):339–358, 1997.

[66] Steven N. Evans and Klaus Fleischmann. Cluster formation in a stepping-stone model with continuous, hierarchically structured sites. *Ann. Probab.*, 24(4):1926–1952, 1996.

[67] U. Faigle and Gy. Turán. Sorting and recognition problems for ordered sets. *SIAM J. Comput.*, 17(1):100–113, 1988.

[68] William Feller. *An introduction to probability theory and its applications. Vol. I*. Third edition. John Wiley & Sons Inc., New York, 1968.

[69] M. A. Fiol. Eigenvalue interlacing and weight parameters of graphs. *Linear Algebra Appl.*, 290(1-3):275–301, 1999.

[70] Abraham Flaxman, Alan Frieze, and Trevor Fenner. High degree vertices and eigenvalues in the preferential attachment graph. In *Approximation, randomization, and combinatorial optimization*, volume 2764 of *Lecture Notes in Comput. Sci.*, pages 264–274. Springer, Berlin, 2003.

[71] Abraham Flaxman, Alan Frieze, and Trevor Fenner. High degree vertices and eigenvalues in the preferential attachment graph. *Internet Math.*, 2(1):1–19, 2005.

[72] L. R. G. Fontes, M. Isopi, C. M. Newman, and K. Ravishankar. The Brownian web: characterization and convergence. *Ann. Probab.*, 32(4):2857–2883, 2004.

[73] Peter Forrester. Log-gases and random matrices, 2009. Preprint of book available at `http://www.ms.unimelb.edu.au/ matpjf/matpjf.html`.

[74] E. Friedgut. Influences in product spaces: KKL and BKKKL revisited. *Combinatorics, Probability and Computing*, 13(01):17–29, 2004.

[75] E. Friedgut and G. Kalai. Every monotone graph property has a sharp threshold. *Proceedings of the American Mathematical Society*, 124(10):2993–3002, 1996.

[76] Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998.

[77] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 124(10):2993–3002, 1996.

[78] Chris Godsil and Gordon Royle. *Algebraic graph theory*, volume 207 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2001.

[79] G. R. Grimmett. Random labelled trees and their branching networks. *J. Austral. Math. Soc. Ser. A*, 30(2):229–237, 1980/81.

[80] Alice Guionnet. *Large Random Matrices: Lectures on Macroscopic Asymptotics: École d'Été de Probabilités de Saint-Flour XXXVI 2006*. Lecture Notes in Mathematics. Springer, 2009.

[81] Willem H. Haemers. Interlacing eigenvalues and graphs. *Linear Algebra Appl.*, 226/228:593–616, 1995.

[82] B. M. Hambly and T. Kumagai. Diffusion processes on fractal fields: heat kernel estimates and large deviations. *Probab. Theory Related Fields*, 127(3):305–352, 2003.

[83] H. Hatami. Decision trees and influences of variables over product probability spaces. *Combinatorics, Probability and Computing*, 18(03):357–369, 2009.

[84] Tim Hobson and Roger Tribe. On the duality between coalescing Brownian particles and the heat equation driven by Fisher-Wright noise. *Electron. Comm. Probab.*, 10:136–145 (electronic), 2005.

[85] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[86] Chris Howitt and Jon Warren. Dynamics for the Brownian web and the erosion flow. *Stochastic Process. Appl.*, 119(6):2028–2051, 2009.

[87] Kiyosi Itô and Henry P. McKean, Jr. *Diffusion processes and their sample paths*. Springer-Verlag, Berlin, 1974. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 125.

[88] Peter Jagers. General branching processes as Markov fields. *Stochastic Process. Appl.*, 32(2):183–212, 1989.

[89] Peter Jagers and Olle Nerman. The growth and composition of branching populations. *Adv. in Appl. Probab.*, 16(2):221–259, 1984.

[90] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[91] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In *Proceedings of 29th IEEE Symp. Foundations of Computer Science (FOCS*, 1988.

[92] W. D. Kaigh. An invariance principle for random walk conditioned by a late return to zero. *Ann. Probability*, 4(1):115–121, 1976.

[93] G. Kalai and S. Safra. Threshold phenomena and influence. *Computational Complexity and Statistical Physics*, pages 25–60, 2006.

[94] R. Karp and M. Sipser. Maximum matchings in sparse random graphs. *22nd Annual Symposium on Foundations of Computer Science*, pages 364–375, 1981.

[95] N. Keller. On the influences of variables on boolean functions in product spaces. Available at `http://front.math.ucdavis.edu/0905.4216`.

[96] N. Keller. Lower bound on the correlation between monotone families in the average case. *Adv. in Appl. Math.*, 43(1):31–45, 2009.

[97] BL Kleitman et al. A phase transition on partial orders. *Physica A: Statistical and Theoretical Physics*, 96(1-2):254–259, 1979.

[98] D. J. Kleitman and B. L. Rothschild. Asymptotic enumeration of partial orders on a finite set. *Trans. Amer. Math. Soc.*, 205:205–220, 1975.

[99] Achim Klenke. Different clustering regimes in systems of hierarchically interacting diffusions. *Ann. Probab.*, 24(2):660–697, 1996.

[100] Takashi Kumagai and Karl-Theodor Sturm. Construction of diffusion processes on fractals, d-sets, and general metric measure spaces. *J. Math. Kyoto Univ.*, 45(2):307–327, 2005.

[101] Yves Le Jan and Olivier Raimond. Flows, coalescence and noise. *Ann. Probab.*, 32(2):1247–1315, 2004.

[102] Tom Lindstrøm. Brownian motion on nested fractals. *Mem. Amer. Math. Soc.*, 83(420):iv+128, 1990.

[103] G.A. Margulis. Probabilistic characteristics of graphs with large connectivity. *Problemy Peredachi Informatsii*, 10(2):101–108, 1974.

[104] Frederick A. Matsen and Steven N. Evans. Ubiquity of synonymity: almost all large binary trees are not uniquely identified by their spectra or their immanantal polynomials, 2005.

[105] Frederick A. Matsen and Steven N. Evans. Ubiquity of synonymity: almost all large binary trees are not uniquely identified by their spectra or their immanantal polynomials. U.C. Berkeley Department of Statistics Technical Report No. 698, 2006.

[106] Madan Lal Mehta. *Random Matrices*, volume 142 of *Pure and Applied Mathematics*. Academic Press, Third edition, 2004.

[107] Madan Lal Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[108] M. Mihail and C.H. Papadimitriou. On the eigenvalue power law. *Proceedings of RANDOM*, pages 254–262, 2002.

[109] Tamás F. Móri. The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.*, 14(3):339–348, 2005.

[110] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. To appear in Ann. Math., 2009.

[111] Ranjiva Munasinghe, R. Rajesh, Roger Tribe, and Oleg Zaboronski. Multiscaling of the n-point density function for coalescing Brownian motions. *Comm. Math. Phys.*, 268(3):717–725, 2006.

[112] Olle Nerman and Peter Jagers. The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch. Verw. Gebiete*, 65(3):445–460, 1984.

[113] J. Neyman and ES Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A*, pages 289–337, 1933.

[114] Hans Jürgen Prömel, Angelika Steger, and Anusch Taraz. Phase transitions in the evolution of partial orders. *J. Combin. Theory Ser. A*, 94(2):230–275, 2001.

[115] Oscar Rojo. On the spectra of certain rooted trees. *Linear Algebra Appl.*, 414(1):218–243, 2006.

[116] Oscar Rojo. The spectra of some trees and bounds for the largest eigenvalue of any tree. *Linear Algebra Appl.*, 414(1):199–217, 2006.

[117] Oscar Rojo. The spectra of a graph obtained from copies of a generalized Bethe tree. *Linear Algebra Appl.*, 420(2-3):490–507, 2007.

[118] Oscar Rojo. Spectra of weighted generalized Bethe trees joined at the root. *Linear Algebra Appl.*, 428(11-12):2961–2979, 2008.

[119] Oscar Rojo and María Robbiano. An explicit formula for eigenvalues of Bethe trees and upper bounds on the largest eigenvalue of any tree. *Linear Algebra Appl.*, 427(1):138–150, 2007.

[120] Oscar Rojo and María Robbiano. On the spectra of some weighted rooted trees and applications. *Linear Algebra Appl.*, 420(2-3):310–328, 2007.

[121] Oscar Rojo and Ricardo Soto. The spectra of the adjacency matrix and Laplacian matrix for some balanced trees. *Linear Algebra Appl.*, 403:97–117, 2005.

[122] Anna Rudas, Bálint Tóth, and Benedek Valkó. Random trees and general branching processes. *Random Structures Algorithms*, 31(2):186–202, 2007.

[123] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[124] L. Russo. An approximate zero-one law. *Probability Theory and Related Fields*, 61(1):129–139, 1982.

[125] Allen J. Schwenk. Almost all trees are cospectral. In *New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich., 1971)*, pages 275–307. Academic Press, New York, 1973.

[126] Robert T. Smythe and Hosam M. Mahmoud. A survey of recursive trees. *Teor. Ĭmovīr. Mat. Stat.*, (51):1–29, 1994.

[127] Florin Soucaliuc, Bálint Tóth, and Wendelin Werner. Reflection and coalescence between independent one-dimensional Brownian paths. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(4):509–545, 2000.

[128] Florin Soucaliuc and Wendelin Werner. A note on reflecting Brownian motions. *Electron. Comm. Probab.*, 7:117–122 (electronic), 2002.

[129] J. Michael Steele. An Efron-Stein inequality for nonsymmetric statistics. *Ann. Statist.*, 14(2):753–758, 1986.

[130] J. Michael Steele. *Probability theory and combinatorial optimization*, volume 69 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[131] K. Strimmer and V. Moulton. Likelihood analysis of phylogenetic networks using directed graphical models. *Molecular Biology and Evolution*, 17(6):875, 2000.

[132] K. Strimmer, C. Wiuf, and V. Moulton. Recombination analysis using directed graphical models. *Molecular Biology and Evolution*, 18(1):97, 2001.

[133] V. N. Sudakov and B. S. Tirelson. Extremal properties of half-spaces for spherically invariant measures. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 41:14–24, 165, 1974.

[134] Rongfeng Sun and Jan M. Swart. The Brownian net. *Ann. Probab.*, 36(3):1153–1208, 2008.

[135] Michel Talagrand. On Russo's approximate zero-one law. *Ann. Probab.*, 22(3):1576–1587, 1994.

[136] Bálint Tóth and Wendelin Werner. The true self-repelling motion. *Probab. Theory Related Fields*, 111(3):375–452, 1998.

[137] Boris Tsirelson. Scaling limit, noise, stability. In *Lectures on probability theory and statistics*, volume 1840 of *Lecture Notes in Math.*, pages 1–106. Springer, Berlin, 2004.

[138] M. Tulino and S. Verdu. *Random Matrix Theory and Wireless Communications*. Foundations and Trends in Communications and Information Theory. Now Publishers, 2004.

[139] Peter Winkler. Random orders. *Order*, 1(4):317–331, 1985.

[140] Jie Xiong and Xiaowen Zhou. On the duality between coalescing Brownian motions. *Canad. J. Math.*, 57(1):204–224, 2005.

[141] L. Zajíček. Porosity and σ-porosity. *Real Anal. Exchange*, 13(2):314–350, 1987/88.

[142] Xiaowen Zhou. Clustering behavior of a continuous-sites stepping-stone model with Brownian migration. *Electron. J. Probab.*, 8:no. 11, 15 pp. (electronic), 2003.

[143] Xiaowen Zhou. Stepping-stone model with circular Brownian migration. *Canad. Math. Bull.*, 51(1):146–160, 2008.

[144] V.M. Zolotarev. Lévy metric. In Michiel Hazewinkel, editor, *Encyclopaedia of Mathematics*. Kluwer Academic Publishers, 2001.