Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

On some random forests with determinantal roots

Luca Avena[1], Alexandre Gaudillière[2]

submitted: December 3, 2013

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: luca.avena@wias-berlin.de

[2] Aix Marseille Université
CNRS, LATP UMR 7353
39 rue Joliot-Curie
13013 Marseille
France
E-Mail: gaudilli@cmi.univ-mrs.fr

No. 1881
Berlin 2013

2010 *Mathematics Subject Classification.* Primary: 05C81, 60J20, 15A15, secondary: 4215A18, 05C85.

Key words and phrases. Finite networks, spanning forests, determinantal processes, random sets, hitting times, local equilibria, Wilson's algorithm, random partitions, coalescence and fragmentation.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

Consider a finite weighted oriented graph. We study a probability measure on the set of spanning rooted oriented forests on the graph. We prove that the set of roots sampled from this measure is a determinantal process, characterized by a possibly non-symmetric kernel with complex eigenvalues. We then derive several results relating this measure to the Markov process associated with the starting graph, to the spectrum of its generator and to hitting times of subsets of the graph. In particular, the mean hitting time of the set of roots turns out to be independent of the starting point, conditioning or not to a given number of roots. Wilson's algorithm provides a way to sample this measure and, in absence of complex eigenvalues of the generator, we explain how to get samples with a number of roots approximating a prescribed integer. We also exploit the properties of this measure to give some probabilistic insight into the proof of an algebraic result due to Micchelli and Willoughby [13]. Further, we present two different related coalescence and fragmentation processes.

1 A random set problem and a forest measure

1.1 "Well distributed points" in a given graph

Let us consider the following problem. We have a square chessboard with sides of length $2l$ and a simple random walk on it. More precisely, think the chessboard as the square lattice box $\mathcal{X} := \{1, \cdots, 2l\}^2$ with the simple random walk X on $\mathcal{X}$. Denote by T_R the hitting time of a set $R \subset \mathcal{X}$ for the walk X and by P_x the law of X starting from $x \in \mathcal{X}$. Can you find a probability law $\mathbb{P}$ on the subsets R of $\mathcal{X}$ with cardinality $|R| = |\mathcal{X}|/2$ such that

$$\mathbb{E}\left[E_x\left[T_R\right]\right] \text{ does not depend on } x \text{ ?}$$

In other words, can you sample $2l^2$ "well distributed points" among the $4l^2$ points of $\mathcal{X}$? In this example, a possible simple answer is the following. Take R to be the set of either white or black squares of the chessboard with probability $1/2$.

One could then raise the following questions:

- What if the random subsets R are required to have any other cardinality $|R| = m \leq 4l^2$?
- What if, instead of the chessboard $\mathcal{X}$, we consider any other finite weighted (possibly oriented) graph?

We notice that the case $m = |\mathcal{X}|$ is trivial, and that in the case $m = 1$, it is known that it suffices to choose the point in $\mathcal{X}$ according to the stationary measure of the walk (see e.g. [11]). One of the main goals of this paper is to answer these questions for $1 < m < |\mathcal{X}|$.

In the sequel, we work on a finite oriented weighted graph with the natural irreducible Markov chain associated with it. We study a certain probability measure on the set of spanning rooted oriented forests on this graph. It will turn out that the set of roots of the forest sampled from this measure, with conditioning on the number of roots, provides a solution to this problem in full generality. As far as

practical sampling issues are concerned, by using an algorithm due to Wilson and Propp [19, 15] we can sample this measure without conditioning. Furthermore, under the assumption that the generator of the random walk associated with the starting weighted graph has only real eigenvalues, we explain how to get a sample with an approximate prescribed number m of roots within an error of order $\sqrt{m}$. In Section 1.2 below, we introduce the main framework and notation, and in Section 1.3 we describe the structure of the paper and the results we derive.

1.2 Forest measures

Let X be a Markov process on a finite state space $\mathcal{X}$, with $|\mathcal{X}| = n$. Assume X is irreducible with generator given by

$$(Lf)(x) = \sum_{y \in \mathcal{X}} w(x,y)[f(y) - f(x)], \quad x \in \mathcal{X}, \tag{1.1}$$

with $f : \mathcal{X} \to \mathbb{R}$ arbitrary and $\{w(x,y) \in [0,\infty) : (x,y) \in \mathcal{X} \times \mathcal{X}\}$ a given collection of non-negative transition rates. Note that such a Markov process has variable speed depending on the current state, namely, if the chain is at position x, the next jump will be performed after an exponential time of rate

$$w(x) = \sum_{y \in \mathcal{X} \setminus \{x\}} w(x,y) < \infty. \tag{1.2}$$

The collections of rates induces a structure of oriented weighted graph on $\mathcal{X}$. In fact, consider the set of oriented edges

$$\mathcal{E} := \{(x,y) \in \mathcal{X} \times \mathcal{X} : \ x \neq y \text{ and } w(x,y) > 0\}, \tag{1.3}$$

then $\mathcal{G} := (\mathcal{X}, \mathcal{E})$ is a weighted oriented graph.

The main object of our investigation is a measure on the spanning rooted oriented forests in $\mathcal{E}$. Let $\mathcal{F}$ denote the set of spanning forests constituted by oriented rooted trees (oriented towards their roots). For $\phi \in \mathcal{F}$, define the weight of a such a forest by

$$w(\phi) := \prod_{e \in \phi} w(e). \tag{1.4}$$

Definition 1.1. (Standard forest measure) *Denote by $\rho(\phi) \subset \mathcal{X}$ the set of roots of a forest $\phi \in \mathcal{F}$, i.e., the set of points which are the roots of the trees constituting ϕ. Fix $q > 0$ and define on $\mathcal{F}$ the measure w_q given by*

$$w_q(\phi) := q^{|\rho(\phi)|} \prod_{e \in \phi} w(e) = q^{|\rho(\phi)|} w(\phi). \tag{1.5}$$

By normalizing it via the partition function:

$$Z(q) := \sum_{\phi \in \mathcal{F}} w_q(\phi), \tag{1.6}$$

we denote the resulting probability measure by

$$\nu_q(\phi) := \frac{w_q(\phi)}{Z(q)}, \quad \phi \in \mathcal{F}. \tag{1.7}$$

We call standard measure*,* standard partition function *and* standard probability measure*, the objects defined by equations* (1.5)*,* (1.6)*, and* (1.7)*, respectively.*

Remarks: There are obvious similarities between the weights appearing in equation (1.5) and those of Fortuyn-Kasteleyn model. We stress here the main differences. FK-percolation is defined on spanning graphs that are not required to be forests. However, in the zero limit of its main parameters, properly rescaled, the model does converge to a measure on spanning forests (see e.g. [7]). Nevertheless, our forests are rooted and this extra structure introduces an entropic factor in comparison (by projection on unrooted forests) with this zero limit of FK-percolation.

We note that the weights in (1.4) are those associated by Freidlin and Wentzell with the so called "W-graphs" [6] and we will recover some of their results in this paper. Our standard forest measure has also been studied in various works. For example to sample points from the stationary measure of the random walk [15], or to study in [8] recurrent configurations of Abelian dissipative sandpile introduced in [18]. This spanning forest measure and other associated objects we will discuss later are actually a variation on the theme of uniform spanning trees and loop-erased random walks. We refer to [10] and references therein for the vast literature on the subject.

As it will become clear, the measure in Definition 1.1 and the associated partition function encode several information related to the chain X. At the occurrence, we will derive several results related with slightly more general measures and partition functions. To this aim, we introduce here some further notation. Let us first introduce a natural generalization of the measure in Definition 1.1. Let $\mathcal{M}_{n\times n}$ be the space of square matrices of size n. Given a collection of extra weights $\{q(x) \in [0,\infty] : x \in \mathcal{X}\}$, let Q be the diagonal matrix in $\mathcal{M}_{n\times n}$ defined by $(Q)_{i,i} := q(x_i)$ with $x_i \in \mathcal{X}$ for $i = 1,\dots,n$. We anticipate that these extra weights will be interpreted as killing rates for the chain X. Set $S := \{x \in \mathcal{X} : \ q(x) = +\infty\}$ and define the measure w_Q by

$$w_Q(\phi) := w(\phi) \prod_{x\in\rho(\phi)\setminus S} q(x)\mathbb{1}_{\{S\subset\rho(\phi)\}}. \tag{1.8}$$

By assuming that there is at least one $x \in \mathcal{X}$ with $q(x) > 0$ we can turn w_Q into a probability measure on $\mathcal{F}$ by normalizing it via the partition function:

$$Z_Q := \sum_{\phi\in\mathcal{F}} w_Q(\phi), \tag{1.9}$$

and we denote the resulting probability measure by

$$\nu_Q(\phi) := \frac{w_Q(\phi)}{Z_Q}, \quad \phi \in \mathcal{F}. \tag{1.10}$$

This is the general form of the probability measure at the core of our investigation.

When answering the questions raised in Section 1.1, we need the following special case of the generalized measure in (1.8). For a given subset $R \subset \mathcal{X}$, suppose that the collection of extra weights $\{q(x) \in [0,\infty], x \in \mathcal{X}\}$ is such that

$$q(x) = \begin{cases} +\infty & \text{if } x \in R, \\ q \geq 0 \ (q > 0 \text{ if } R = \emptyset) & \text{otherwise.} \end{cases} \tag{1.11}$$

In this case, $S = R$, and we write

$$w_{q,R}(\phi) := q^{|\rho(\phi)\setminus R|} w(\phi)\mathbb{1}_{\{R\subset\rho(\phi)\}} \quad \text{and} \quad Z_R(q) := \sum_{\phi\in\mathcal{F}} w_{q,R}(\phi), \tag{1.12}$$

for the measure and the partition function, $w_Q(\cdot)$ and Z_Q, respectively. In particular, for $q = 0$, $w_{0,R}(\phi) = w(\phi)\mathbb{1}_{\{\rho(\phi)=R\}}$ and $Z_R(0) = \sum_{\phi:\rho(\phi)=R} w(\phi)$. Note further that, when $R = \emptyset$, we recover the *standard* measure and partition function, $w_{q,\emptyset}(\cdot) = w_q(\cdot)$ and $Z_\emptyset(q) = Z(q)$.

In the sequel we will denote by Φ a random variable on a probability space $(\Omega, \mathbb{P})$ with values in $\mathcal{F}$ and law ν_Q. To avoid ambiguities, when needed, we write $\mathbb{P}_Q$, $\mathbb{P}_{q,R}$ or $\mathbb{P}_q$ in place of $\mathbb{P}$.

1.3 Results and structure of the paper

1.3.1 Main results

Our analysis of the forest measures introduced above will lead us to several results. Before describing them and the organization of the paper, we emphasize herein what we consider as our three main results.

First, in Theorem 3.4 we prove that the set of roots $\rho(\phi)$ sampled from the standard measure in (1.5) is a determinantal process. More precisely, denoting by $K(x,y) = P_x(X(T_q) = y)$, for $x, y \in \mathcal{X}$, the transition probabilities of the Markov chain X in (1.1) observed at independent exponentially distributed times of parameter q, we have that

$$\mathbb{P}_q\left(A \subset \rho\left(\Phi\right)\right) = \det_A(K), \quad A \subseteq \mathcal{X},$$

with $\det_A(K)$ being the determinant of K restricted to A (see Section 1.3.3 below for the notation). This echoes Burton and Pemantle transfer current theorem ([2], Thm 1.1). Our kernel K is however not required to be reversible (and with possibly complex eigenvalues), and we present a direct proof not relying on transfer currents.

The second result is an answer to the questions in the introductory Section 1.1. In fact, in Theorem 3.5 we give a formula for the hitting times of random sets constituted by the roots of our standard random forests, with or without conditioning on having a fixed number of roots. In particular, such a formula is independent of the starting point x. While Wilson and Propp algorithm gives a way to sample the unconditioned measure, in Section 3.2 we explain how to obtain, in absence of complex eigenvalues, a sample with approximately m roots, with an error of order $\sqrt{m}$, for any $m \leq n$.

Our third result concerns a non-trivial algebraic statement on symmetric matrices derived in [13], Thm 3.2 therein. This theorem has been used in some recent probabilistic works [5, 14] to study absorption times of reversible Markov chains on general finite graphs, as a key tool to define, in such a general setting, the local equilibria introduced in [3]. In [14], the author motivates the importance of having a probabilistic interpretation of Micchelli and Willoughby algebraic result. In Section 3.3 we restate their result and give, by means of our standard forest measure, probabilistic insights into their proof.

1.3.2 Organization of the paper and results description

The rest of the paper is organized as follows.

Background material: Section 2 is a warm-up section where we provide some known background material. In Section 2.1 we prove in a slightly different way a result originally due to Marchal [12] on loop-erased trajectories, Proposition 2.1. In Section 2.2 we recall Wilson's algorithm, Definition 2.2, and as a corollary of Proposition 2.1, following [12, 15], we show how to sample our unconditioned measures, Corollary 2.3.

Results: Section 3 is the bulk of this work where we derive our results. Each subsection presents some new results related with our forest measures. Section 3.1 presents the first analysis on the forest measures, in Theorem 3.1 and Corollary 3.2 therein we show some connection of these measures with the spectrum of the chain X. In Theorem 3.4 we prove that the set of roots is a determinantal process. In Section 3.2 we prove Theorem 3.5 on the hitting times of the roots set, answering the questions raised in Section 1.1. Section 3.3 concerns a partial probabilistic interpretation of the proof and result by Micchelli and Willoughby [13]. In Section 3.4 we mention a couple of coalescence and fragmentation processes associated with our measures. One of them is obtained by coupling together all the standard forest measures for different values of q. The focus of Section 3.5 is on the 'rooted

partitions' induced by our spanning forests, Proposition 3.11, and on the cumulants, or truncated correlation functions, of the roots process, Section 3.5.2.

Appendix: Appendix A is devoted to technical known results used along the paper, which we derive in our context in order to have a self-contained work. In Appendix A.1, we recall what is the Schur complement for block matrices and its probabilistic interpretation, Proposition A.1. In Appendix A.2, we give different proofs of two lemmas from Freidlin and Wentzell (Lemmas 3.2, 3.3 in [6]) on hitting distribution and times of subsets of the graph, again by means of our forest measure. These lemmas are used in Section 3.2. Appendix A.3 concerns the notion of divided differences which are used in Section 3.3. We state three equivalent definitions and prove a related lemma.

To conclude this introductory section and to simplify the reading, we fix here some notation. Further notation will be introduced at the occurrence.

1.3.3 Main notation

- *Set theory*
  - **Spaces:** $\mathcal{X}$ will be our reference state space for the Markov process X. In the sequel, we work with extensions of $\mathcal{X}$ which will be denoted by $\bar{\mathcal{X}}$ or with more general spaces denoted by $\mathcal{Y}$.
  - **Subsets:** the symbols $\subset$ and $\subsetneq$ will be used as inclusion and strict inclusion, respectively. Subsets will be generally denoted by capital letters: A, B, R, S.
  - **Complement:** unless specified, we always deal with subsets of the space $\mathcal{X}$. For a given subset $A \subset \mathcal{X}$, we denote by $A^c := \mathcal{X} \setminus A$ the complement of A in $\mathcal{X}$.
- *Graphs*
  - **Edges**: $\mathcal{E}$ introduced in (1.3) stands for the set of oriented edges on $\mathcal{X}$.
  - **Extreme points of oriented edges:** for an oriented edge $e = (x, y) \in \mathcal{X} \times \mathcal{X}$, we denote the starting and the ending points of e by $\underline{e} = x$ and $\overline{e} = y$, and by $w(e) := w(x, y)$ the weight associated to it.
  - **Forest space:** $\mathcal{F}$ denotes the set of spanning rooted oriented forests on $\mathcal{X}$.
  - **A given forest:** elements of $\mathcal{F}$ are denoted by $\phi \in \mathcal{F}$, which can also be seen as subsets of $\mathcal{E}$.
  - **Roots:** given $\phi \in \mathcal{F}$, $\rho(\phi)$ denotes the set of the roots of the trees constituting ϕ.
  - **Tree associated with a given vertex:** given x in $\mathcal{X}$ and $\phi \in \mathcal{F}$ we denote by $\tau_\phi(x)$ the unique tree in ϕ that covers x. We write $\rho(\tau_\phi(x)) = \{x\}$ when $\tau_\phi(x)$ is rooted at x.
- *Matrix*
  - **Square matrices space:** the space of square matrices with n columns and n rows is denoted by $\mathcal{M}_{n\times n}$.
  - **Restriction of a matrix:** given a matrix $M \in \mathcal{M}_{n\times n}$ with elements x_i's in $\mathcal{X}$, for any subset A of $\mathcal{X}$, $[M]_A$ stands for the matrix M restricted to its elements doubly indexed in A.
  - **Determinants:** $\det_A(M)$ will denote the determinant of the matrix obtained from M by removing all the lines and columns with indexes outside $A \subset \mathcal{X}$, i.e., $\det_A(M) = \det([M]_A)$.

- *Random walks*
  - **Markov processes and discrete skeletons:** X and Y will denote Markov processes on the spaces $\mathcal{X}$ and $\mathcal{Y}$, respectively. $\hat{X}$ and $\hat{Y}$ will denote some associated discrete-time Markov chains.
  - **Generators:** L and $\mathcal{L}$ denote the generators of the Markov processes X and Y respectively. For a given subset A of $\mathcal{X}$, L_A stands for the generator of the Markov process X *restricted* to the component A. While $\mathcal{L}^A$ denotes the *trace* of the process Y on some fixed subset A of $\mathcal{Y}$.

2 Background material

2.1 On loop-erased trajectories

We introduce here a slightly more general setting than in Section 1.2.

Let Y be a Markov process on a finite state space $\mathcal{Y}$. Assume Y is irreducible with generator given by

$$(\mathcal{L}f)(y) = \sum_{z\in\mathcal{Y}} \alpha(y,z)[f(z) - f(y)], \quad y \in \mathcal{Y}, \tag{2.1}$$

with $f : \mathcal{Y} \to \mathbb{R}$ arbitrary and $\{\alpha(y,z) \in [0,+\infty] : (y,z) \in \mathcal{Y}\times\mathcal{Y}\}$ a given collection of non-negative transition rates. Let

$$\alpha(y) := \sum_{z\in\mathcal{Y}\setminus\{y\}} \alpha(y,z) \in [0,\infty]. \tag{2.2}$$

Let B be a subset of $\mathcal{Y}$ such that

$$\{y \in \mathcal{Y} : \alpha(y) = +\infty\} \subset B \subseteq \mathcal{Y}, \tag{2.3}$$

so that, $\alpha(y) < \infty$ for any $y \in B^c := \mathcal{Y} \setminus B$, the complement of B in $\mathcal{Y}$.

Denote by $\gamma(B,l) = (y_0,\dots,y_l)$ a self-avoiding path of $l+1$ points and length l such that $y_i \in B^c$ for $i = 0,\dots,l-1$ and $y_l \in B$. For $y_0 \in B^c$, let P_{y_0} the law of the random walk Y when starting from y_0. Denote by Γ_B a random trajectory obtained from Y under P_{y_0} as follows. Stop the walk Y when it enters the set B for the first time and erase all its loops. After this procedure, we are left with a self-avoiding trajectory of variable length. In the next proposition we compute the probability that Γ_B equals a given trajectory $\gamma(B,l)$ of length l. To simplify the proof, we use the discrete skeleton of the Markov process Y on B^c. This justify the following definitions.

Set

$$\bar{\alpha} := \max_{y\in B^c} \alpha(y) < \infty, \tag{2.4}$$

and let $\hat{P} \in M_{m\times m}$, with $m = |Y|$, be the transition matrix identified by the entries

$$\hat{p}(y,z) := \begin{cases} \delta_{y,z} & \text{if } y \in B \\ \alpha(y,z)/\bar{\alpha} & \text{if } z \neq y \text{ and } y \in B^c, \\ 1 - \displaystyle\sum_{x\in\mathcal{Y}\setminus\{y\}} \hat{p}(y,x) & \text{if } y = z \in B^c. \end{cases} \tag{2.5}$$

Such a matrix $\hat{P}$ is a Markovian transition matrix for a discrete-time random walk $\hat{Y}$ on $\mathcal{Y}$. In particular, for an arbitrary function f, by construction we have that

$$(\mathcal{L}f)(y) = (\bar{\alpha}(\hat{P} - \mathbb{1})f)(y), \text{ for all } y \in B^c. \tag{2.6}$$

We are in shape to prove the claimed proposition, by using a nice independence argument we learned from Laurent Tournier [17].

Proposition 2.1. (Marchal [12]) *Consider the random walk Y on $\mathcal{Y}$ with generator $\mathcal{L}$ as in* (2.1)*. Fix $y_0 \in B^c$. Then, under P_{y_0},*

$$P_{y_0}\left(\Gamma_B = \gamma(B,l)\right) = \prod_{i=0}^{l-1} \alpha(y_i, y_{i+1}) \frac{\det_{B^c\setminus\{y_0,\ldots,y_{l-1}\}}(-\mathcal{L})}{\det_{B^c}(-\mathcal{L})}, \tag{2.7}$$

with the matrix notation according to Section 1.3.3 and $\gamma(B,l) = (y_0,\ldots,y_l)$ a self-avoiding path of length l such that $y_i \in B^c$ for $i = 0,\ldots,l-1$ and $y_l \in B$.

Proof. For the discrete chain $\hat{Y}$, let $\hat{T}^+_{y_0}$ and $\hat{T}_B$ be the first return time to y_0 and the hitting time of B, respectively. More precisely, $\hat{T}^+_{y_0} := \inf\{k \geq 1 : \hat{Y}_k = y_0\}$ and $\hat{T}_B := \inf\{k \geq 0 : \hat{Y}_k \in B\}$. Note that by definition of Γ_B, we have that

$$P_{y_0}\left(\Gamma_B = \gamma(B,l) | \hat{T}^+_{y_0} < \hat{T}_B\right) = P_{y_0}\left(\Gamma_B = \gamma(B,l)\right).$$

As a consequence, we can write:

$$\begin{aligned} P_{y_0}\left(\Gamma_B = \gamma(B,l)\right) &= P_{y_0}\left(\Gamma_B = \gamma(B,l) | \hat{T}^+_{y_0} < \hat{T}_B\right) P_{y_0}\left(\hat{T}^+_{y_0} < \hat{T}_B\right) \\ &\quad + P_{y_0}\left(\Gamma_B = \gamma(B,l), \hat{T}^+_{y_0} > \hat{T}_B\right) \\ &= P_{y_0}\left(\Gamma_B = \gamma(B,l)\right) P_{y_0}\left(\hat{T}^+_{y_0} < \hat{T}_B\right) \\ &\quad + P_{y_0}\left(\Gamma_B = \gamma(B,l), \hat{T}^+_{y_0} > \hat{T}_B\right). \end{aligned} \tag{2.8}$$

It follows from (2.8) that

$$P_{y_0}\left(\Gamma_B = \gamma(B,l)\right) = \frac{P_{y_0}\left(\Gamma_B = \gamma(B,l), \hat{T}^+_{y_0} > \hat{T}_B\right)}{P_{y_0}\left(\hat{T}^+_{y_0} > \hat{T}_B\right)}. \tag{2.9}$$

Denote by $l_{y_0}(\hat{T}_B)$ the local time $\hat{Y}$ spends at y_0 before entering B, and by $[\hat{P}]_{B^c}$ the matrix $\hat{P}$ restricted to B^c, then

$$\begin{aligned} \frac{1}{P_{y_0}\left(\hat{T}^+_{y_0} > \hat{T}_B\right)} &= E_{y_0}[l_{y_0}(\hat{T}_B)] = \sum_{k\geq 0} [\hat{P}]^k_{B^c}(y_0,y_0) \\ &= ([\mathbb{1} - \hat{P}]_{B^c})^{-1}(y_0,y_0) = \frac{\det_{B^c\setminus\{y_0\}}(\mathbb{1}-\hat{P})}{\det_{B^c}(\mathbb{1}-\hat{P})}, \end{aligned} \tag{2.10}$$

where the last equality follows by Cramer's formula for an inverse matrix.

On the other hand, for the numerator in the r.h.s. of equation (2.9), we can write

$$P_{y_0}\left(\Gamma_B = (y_0,\ldots,y_l), \hat{T}^+_{y_0} > \hat{T}_B\right) = \hat{p}(y_0,y_1) P_{y_1}\left(\Gamma_{B\cup\{y_0\}} = (y_1,\ldots,y_l)\right). \tag{2.11}$$

By plugging (2.10) and (2.11) into (2.9), and iterating, we have that

$$\begin{aligned}
P_{y_0}\left(\Gamma_B = \gamma(B,l)\right) =& \hat{p}(y_0,y_1) P_{y_1}\left(\Gamma_{B\cup\{y_0\}} = (y_1,\dots,y_l)\right) \frac{\det_{B^c\setminus\{y_0\}}(\mathbb{1}-\hat{P})}{\det_{B^c}(\mathbb{1}-\hat{P})} \\
=& \cdots = \prod_{i=0}^{l-1} \hat{p}(y_i,y_{i+1}) \frac{\det_{B^c\setminus\{y_0,\dots,y_{l-1}\}}(\mathbb{1}-\hat{P})}{\det_{B^c}(\mathbb{1}-\hat{P})} \\
=& \bar{\alpha}^{-l} \prod_{i=0}^{l-1} \alpha(y_i,y_{i+1}) \frac{\det_{B^c\setminus\{y_0,\dots,y_{l-1}\}}(-\mathcal{L}/\bar{\alpha})}{\det_{B^c}(-\mathcal{L}/\bar{\alpha})} \\
=& \prod_{i=0}^{l-1} \alpha(y_i,y_{i+1}) \frac{\det_{B^c\setminus\{y_0,\dots,y_{l-1}\}}(-\mathcal{L})}{\det_{B^c}(-\mathcal{L})}.
\end{aligned}$$

□

2.2 Wilson's algorithm

We introduce here the algorithm due to Wilson and Propp [15] which allow us to sample the measure (1.10). First, we extend the Markov process X defined through (1.1) on $\mathcal{X}$ to a Markov process $\bar{X}$ on the space $\bar{\mathcal{X}} = \mathcal{X} \cup \{\Delta\}$ by interpreting Δ as an absorbing state and by adding some killing rates. Consider the space $\mathcal{X}$, with $|\mathcal{X}| = n$. Assume a collection of killing rates $\{q(x) \in [0,\infty] : x \in \mathcal{X}\}$ is given, and let Q be the diagonal matrix in $\mathcal{M}_{n\times n}$ defined by $(Q)_{i,i} := q(x_i)$ with $x_i \in \mathcal{X}$ for $i = 1,\dots,n$. Consider the Markov process $\bar{X}$ on the finite state space $\bar{\mathcal{X}}$ with generator given by

$$(\mathcal{L}f)(x) = \begin{cases} (Lf)(x) + q(x)[f(\Delta) - f(x)], & \text{if } x \in \mathcal{X}, \\ 0, & \text{if } x = \Delta. \end{cases} \tag{2.12}$$

with $f : \bar{\mathcal{X}} \to \mathbb{R}$ arbitrary, and L defined in (1.1). In particular, the matrix $-\mathcal{L}$ associated with the generator in (2.12) satisfies

$$[-\mathcal{L}]_{\mathcal{X}} = Q - L. \tag{2.13}$$

Next, we describe the algorithm. For any $A \subset \bar{\mathcal{X}}$, note that, due to irreducibility, $T_A := \inf\{t \geq 0 : \bar{X}_t \in A\}$ is a.s. finite.

Definition 2.2. (Wilson's algorithm)

1 *Start the process $\bar{X}$ from any point $x_1 \in \mathcal{X}$ until it reaches the absorbing state Δ.*

2 *Erase all the loops of the trajectory described by $\bar{X}$ up to time T_Δ. Call $\gamma_1(\Delta)$ this self-avoiding trajectory. ($\gamma_1(\Delta)$ is such that $\Delta \in \gamma_1(\Delta)$.)*

3 *If $\gamma_1(\Delta)$ covers the whole $\bar{\mathcal{X}}$ stop, else pick any point $x_2 \in \mathcal{X} \setminus \{\gamma_1(\Delta)\}$, with $\{\gamma_1(\Delta)\}$ denoting the set of points covered by $\gamma_1(\Delta)$. Start the process $\bar{X}$ from x_2 until it hits the set $\{\gamma_1(\Delta)\}$.*

4 *Erase all the loops of the trajectory described by $\bar{X}$ starting from x_2 up to time $T_{\{\gamma_1(\Delta)\}}$. Call $\gamma_2(\Delta)$ this self-avoiding trajectory.*

5 *If $\cup_{i=1,2}\{\gamma_i(\Delta)\} = \bar{\mathcal{X}}$ stop, else pick any point $x_3 \in \mathcal{X} \setminus \cup_{i=1,2}\{\gamma_i(\Delta)\}$. Start the process $\bar{X}$ from x_2 until it hits the set $\cup_{i=1,2}\{\gamma_i(\Delta)\}$.*

6 *Iterate until $\bar{\mathcal{X}}$ is covered.*

Denote by $\mathcal{T}_{\bar{\mathcal{X}}}$ the set of spanning oriented trees on $\bar{\mathcal{X}}$ rooted at Δ. This algorithm produces in finite time an element τ of $\mathcal{T}_{\bar{\mathcal{X}}}$. As a corollary of Proposition 2.1, we can easily compute the probability that the algorithm produces a given $\tau \in \mathcal{T}_{\bar{\mathcal{X}}}$.

Corollary 2.3. *Fix a tree $\tau \in \mathcal{T}_{\bar{\mathcal{X}}}$. Let $\partial\rho(\tau) := \{e \in \tau : \overline{e} \in \{\Delta\} \cup S\}$ be the set of edges in τ connected to the root Δ (possibly via the edges with infinite rate from S to Δ). Denote by $\mathbb{P}(Wil = \tau)$ the probability that Wilson's algorithm produces the tree τ. Then*

$$\mathbb{P}(Wil = \tau) = \frac{\left[\prod_{e \in \partial\rho(\tau)} q(\underline{e})\right]\left[\prod_{e \in \tau \setminus \partial\rho(\tau)} w(e)\right]}{\det_{\mathcal{X}\setminus S}(-\mathcal{L})}. \tag{2.14}$$

Proof. Recall the notation in Proposition 2.1. Set $\mathcal{Y} = \bar{\mathcal{X}} = \mathcal{X} \cup \{\Delta\}$ and $Y = \bar{X}$. Start with $B = \{\Delta\} \cup S$. By the definition of the algorithm, the proof follows by iterating the formula in equation (2.7) where at each iteration we set the right B according to the given tree τ. Whatever the choice of starting points in Wilson algorithm we get the same result. □

We conclude this section by observing that there exists a natural bijection between $\mathcal{F}$ and $\mathcal{T}_{\bar{\mathcal{X}}}$. Indeed, given $\phi \in \mathcal{F}$, let $\tau(\phi)$ be the unique element in $\mathcal{T}_{\bar{\mathcal{X}}}$ obtained from ϕ by adding all the edges connecting the roots in ϕ to Δ, i.e. add all edges e such that $\underline{e} \in \rho(\phi)$ and $\overline{e} = \Delta$. Vice versa, given $\tau \in \mathcal{T}_{\bar{\mathcal{X}}}$, by removing all edges $e \in \partial\rho(\tau)$ we can identify a unique element $\phi \in \mathcal{F}$. This simple observation together with Corollary 2.3 allow us to sample the measure in (1.10) using Wilson's algorithm.

3 Results

This section is the bulk of the paper containing our main theorems. Each subsection is devoted to some new results relating the Markov chain X to the law of the forest at the core of our investigation.

3.1 The forest and its roots

We start here to analyze the measure introduced in (1.10) on the space $\mathcal{F}$ of spanning rooted oriented forests on $\mathcal{X}$. We compute the partition function, we identify the distribution of the number of roots in the standard case and we prove that the roots process is a determinantal one.

Theorem 3.1. (Partition function and spectrum) *Assume a collection of killing rates $\{q(x) \in [0,\infty] : x \in \mathcal{X}\}$ is given, and let Q be the diagonal matrix in $\mathcal{M}_{n\times n}$ defined by $(Q)_{i,i} := q(x_i)$ with $x_i \in \mathcal{X}$ for $i = 1, \dots, n$. Let ν_Q the probability measure on $\mathcal{F}$ in* (1.10)*. Then*

$$Z_Q = \det_{\mathcal{X}\setminus S}(Q - L), \tag{3.1}$$

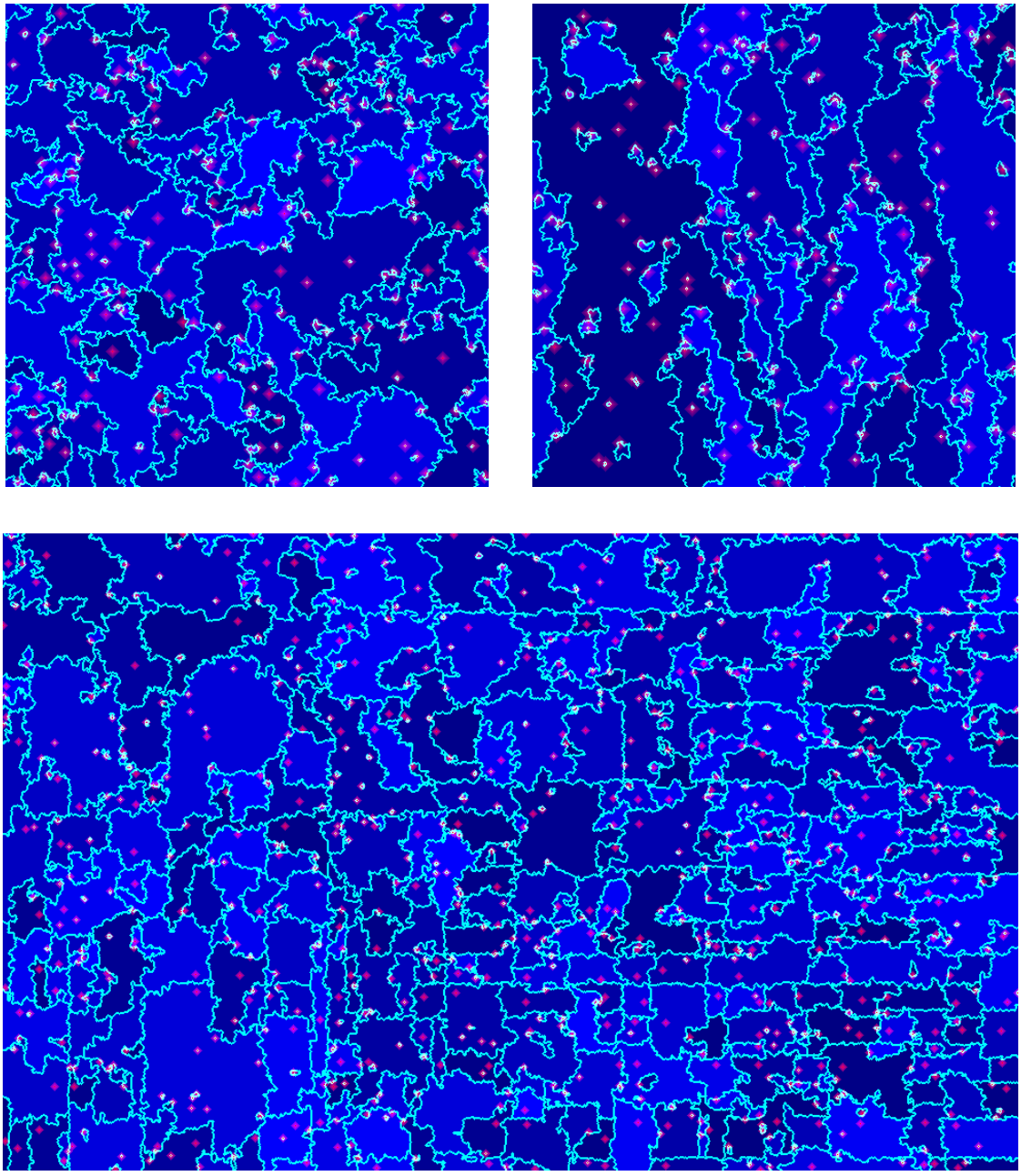
and, recalling the notation from Section 2.2,

$$\nu_Q(\phi) = \mathbb{P}(Wil = \tau(\phi)). \tag{3.2}$$

In the case $q(x) \equiv q > 0$, we recover the standard probability measure in (1.7)*, and the standard partition function in* (1.6) *satisfies*

$$Z(q) = \chi_L(q), \tag{3.3}$$

Figure 1: Samples from ν_q for $q = .001$ on the two-dimensional 512×512 torus with uniform rates equal to 1 between nearest neighbours for the first picture and, for the second picture, with an additional northward drift, such that $w(x, y) = 1.2$ if y is the northern nearest neighbour of x and $w(x, y) = 1$ otherwise. The third picture is a sample from ν_q for the same $q = .001$ on the 987×610 rectangular grid and for Metropolis random walk in a Browian sheet potential V, i.e., nearest neighbour rates are given by $w(x, y) = \exp\{-\beta[V(y) - V(x)]_+\}$ with $\beta = .04$ and V is the restriction to the grid of a Brownian sheet with 0 value on the west and north sides of the box. In each picture different blue levels are assigned to points in different trees, cyan lines separate neighbouring trees, and the forest roots are at the centers of red diamonds.



where χ_L is the characteristic polynomial of L, i.e.

$$Z(q) = \prod_{i=0}^{n-1}(q+\lambda_i) = q\prod_{i=1}^{n-1}(q+\lambda_i), \tag{3.4}$$

where the λ_i's are the eigenvalues of $-L$ ordered by non-decreasing real part. When $R \neq \emptyset$ and $q \geq 0$, we have that

$$Z_R(q) = \chi_{[L]_{R^c}}(q) \tag{3.5}$$

with $\chi_{[L]_{R^c}}$ being the characteristic polynomial of the sub-Markovian generator of the process killed in R.

Remark: This kind of results goes back to Kirchhoff [9]. Here we include the non-reversible case and stress the dependence in q.

Proof. As observed in the previous section, for each forest $\phi \in \mathcal{F}$ there is a unique $\tau(\phi) \in \mathcal{T}_{\bar{\mathcal{X}}}$. By Corollary 2.3 we then have that

$$\mathbb{P}(Wil = \tau(\phi)) = \frac{\left[\prod_{e\in\partial\rho(\tau)} q(\underline{e})\right]\left[\prod_{e\in\tau\setminus\partial\rho(\tau)} w(e)\right]}{\det_{\mathcal{X}\setminus S}(-\mathcal{L})} = \frac{w_Q(\phi)}{\det_{\mathcal{X}\setminus S}(Q-L)}.$$

By summing over all $\phi \in \mathcal{F}$ we immediately get (3.1). In fact,

$$Z_Q := \sum_{\phi\in\mathcal{F}} w_Q(\phi) = \det_{\mathcal{X}\setminus\mathcal{S}}(Q-L)\sum_{\phi\in\mathcal{F}}\mathbb{P}(Wil=\tau(\phi)) = \det_{\mathcal{X}\setminus\mathcal{S}}(Q-L).$$

Moreover, if $q(x) \equiv q > 0$, then $Q = q\mathbb{1}$ and $Z(q) = \det[q\mathbb{1} - L] = \chi_L(q)$. □

In the standard case, an immediate consequence of Theorem 3.1 is a characterization of the law of the cardinality of the set of roots.

Corollary 3.2. (Root number distribution) *Assume the standard case $q(x) \equiv q > 0$ and that L has real spectrum. Let N be a sum of n independent Bernoulli random variables with parameters $q/(q+\lambda_i)$. Under the probability measure $\mathbb{P}_q$, the random variable $|\rho(\Phi)|$ counting the number of roots (or equivalently, of trees) in Φ has the same law as N.*

Proof. Observe that the coefficient of degree k in

$$Z(q) = \sum_{\phi} q^{|\rho(\phi)|}w(\phi)$$

is the total weight of the set of forest with exactly k roots. Since $Z(q) = \prod_i(q+\lambda_i)$ we get,

$$\begin{aligned}
\mathbb{P}_q\left(|\rho(\Phi)| = k\right) =& \frac{1}{Z(q)}\sum_{\phi\in\mathcal{F}:|\rho(\phi)|=k} w_q(\phi) = \left(\prod_{i=0}^{n-1}(q+\lambda_i)\right)^{-1} q^k \sum_{I\in\mathcal{P}[n-k]}\prod_{i\in I}\lambda_i \\
&= \sum_{I\in\mathcal{P}[n-k]}\left[\prod_{i\in I}\left(\frac{\lambda_i}{q+\lambda_i}\right)\right]\left[\prod_{j\notin I}\left(\frac{q}{q+\lambda_j}\right)\right] \\
&= \sum_{J\in\mathcal{P}[k]}\left[\prod_{j\in J}\left(\frac{q}{q+\lambda_j}\right)\right]\left[\prod_{i\notin J}\left(1-\frac{q}{q+\lambda_i}\right)\right],
\end{aligned}$$

where $\mathcal{P}[k]$ stands for the set of all possible k elements of the set $\{0,1,\ldots,n-1\}$. □

Remark: When the spectrum of L does contain a non-real part, one can still compute the law of $|\rho(\Phi)|$ and get the same algebraic expressions in terms of the eigenvalues. One can also compute momenta by differentiating with respect to q the logarithm of the partition function. In particular, the mean value and the variance are given by

$$\mathbb{E}_q[|\rho(\Phi)|] = \sum_{i=0}^{n-1} \frac{q}{q+\lambda_i}, \tag{3.6}$$

$$\mathbb{V}\mathrm{ar}_q(|\rho(\Phi)|) = \sum_{i=0}^{n-1} \frac{q}{q+\lambda_i} - \left(\frac{q}{q+\lambda_i}\right)^2. \tag{3.7}$$

We note however that the contribution of the imaginary part of the eigenvalues can make uneasy the comparison between variance and mean value, at least for small values of q. This is the reason why, when dealing with the question of getting samples with a number of roots that approximates a given $m \leq n$, we will restrict ourselves to the real spectrum case.

Next, we prove that under $\mathbb{P}_Q$ the set of roots $\rho(\Phi) \setminus S$ is a determinantal process as suggested after [4] by the previous result. This is the content of Theorem 3.4 for which we will present an algebraic and a probabilistic proof.

Let us first show a simple lemma. Consider the Markov process $\bar{X}$ on $\bar{\mathcal{X}} = \mathcal{X} \cup \{\Delta\}$ defined via its generator in (2.12). Let T_Δ be the hitting time of $\{\Delta\}$. Denote by $\bar{X}(T_{\Delta,-})$ the last point visited by the process $\bar{X}$ before time T_Δ.

Let K be the transition matrix defined by

$$K(x,y) := P_x(\bar{X}(T_{\Delta,-}) = y), \quad \text{for } x, y \in \mathcal{X} \setminus S. \tag{3.8}$$

This transition kernel can also be expressed in terms of Green's functions:

Lemma 3.3. *For $x, y \in \mathcal{X} \setminus S$, let*

$$G_\Delta(x,y) := E_x[l_y(T_\Delta)] = (Q-L)^{-1}(x,y) \tag{3.9}$$

be the Green's function of the process $\bar{X}$ on the subspace $\mathcal{X} \setminus S$. Then

$$K(x,y) = G_\Delta(x,y)q(y). \tag{3.10}$$

Proof. We use the notation of Section 2.1 to work in time. Set $\mathcal{Y} = \bar{\mathcal{X}}$, $Y = \bar{X}$, $B = \{\Delta\}$ and write

$$\begin{aligned} K(x,y) = P_x(\bar{X}(T_{\Delta,-}) = y) &= \sum_{k\geq 1} P_x(\hat{Y}(k-1) = y, \hat{T}_\Delta = k) \\ &= \sum_{k\geq 1} P_x(\hat{Y}(k-1) = y, \hat{T}_\Delta \leq k-1)\hat{p}(y,\Delta) = E_x[l_y(\hat{T}_\Delta)]\frac{q(y)}{\bar{\alpha}} = G_\Delta(x,y)q(y). \end{aligned}$$

□

Note that, in the standard case $q(x) \equiv q > 0$, we can write

$$K(x,y) = P_x(X(T_q) = y), \quad \text{for } x, y \in \mathcal{X}, \tag{3.11}$$

with T_q being an independent exponential random variable of parameter q. In particular, if $G_q(x,y) := E_x[l_y(T_q)]$ is the Green's function up to time T_q, then

$$K(x,y) = qG_q(x,y) = q(q\mathbb{1} - L)^{-1}(x,y). \tag{3.12}$$

Theorem 3.4. (Determinantal roots) *Under ν_Q with finite killing rates (i.e. $S = \emptyset$), the set of roots in ϕ is a determinantal process with kernel K defined in* (3.8)*. Equivalently, for any $A \subset \mathcal{X}$:*

$$\mathbb{P}_Q\left(A \subset \rho\left(\Phi\right)\right) = \det_A(K).$$

More generally, if $S \neq \emptyset$, $\rho(\Phi) \setminus S$ is still a determinantal process on $\mathcal{X} \setminus S$ with kernel K, that is:

$$\mathbb{P}_Q\left(A \subset \rho\left(\Phi\right) \setminus S\right) = \det_A(K), \quad \text{for any } A \subseteq \mathcal{X} \setminus S. \tag{3.13}$$

Proof. **(Algebraic proof of Thm 3.4)** Assume first $S = \emptyset$. Consider a set $A \subset \mathcal{X}$ with $|A| = r$ of the form $A = \{x_1, x_2, \ldots, x_r\}$. By choosing the different points in A as starting point at each iteration in Wilson's algorithm, by (2.7), we get

$$\begin{aligned}\mathbb{P}_Q\left(\{x_1, x_2, \cdots, x_r\} \subset \rho\left(\Phi\right)\right) &= q(x_1)\frac{\det_{\{x_1\}^c}(Q-L)}{\det_{\mathcal{X}}(Q-L)} q(x_2)\frac{\det_{\{x_1,x_2\}^c}(Q-L)}{\det_{\{x_1\}^c}(Q-L)} \times \cdots \\ &\cdots \times q(x_r)\frac{\det_{A^c}(Q-L)}{\det_{\{x_1,\ldots,x_{r-1}\}^c}(Q-L)} = \left[\prod_{i=1}^{r} q(x_i)\right]\frac{\det_{A^c}(Q-L)}{\det_{\mathcal{X}}(Q-L)} \\ &= \frac{\det_{A^c}(Q-L)}{\det_{\mathcal{X}}(Q-L)}\det_A(Q).\end{aligned} \tag{3.14}$$

In case $A = \mathcal{X}$, the claim is straightforward since equation (3.14) reads

$$\mathbb{P}_Q\left(\rho\left(\Phi\right) = \mathcal{X}\right) = \frac{\det(Q)}{\det(Q-L)} = \det\left((Q-L)^{-1}\right)\det(Q) = \det\left((Q-L)^{-1}Q\right),$$

and the r.h.s. equals $\det(K)$ due to Lemma 3.3.

In case $A \subsetneq \mathcal{X}$, we can use the Schur complement (see (A.15) in Appendix A.2) to show that

$$\frac{\det_{A^c}(Q-L)}{\det_{\mathcal{X}}(Q-L)} = \det_A\left((Q-L)^{-1}\right).$$

Therefore, from equation (3.14), we have that

$$\mathbb{P}_Q\left(\{x_1, x_2, \cdots, x_r\} \subset \rho\left(\Phi\right)\right) = \det_A((Q-L)^{-1})\det_A(Q) = \det_A((Q-L)^{-1}Q) = \det_A(K),$$

where the second to the last equality is justified because Q is a diagonal matrix and the last equality follows from (3.10).

When $S \neq \emptyset$, the proof is the same. We just have to subtract from the various considered sets. □

Proof. **(Probabilistic proof of Thm 3.4)** To avoid a heavy notation, we consider only the case $S = \emptyset$. Starting from the Markov process X and the killing rates $q(x)$, we construct two different absorbing states Δ_A and Δ_{A^c} accessible from the set A and A^c, respectively. Set $\mathcal{Y} = \mathcal{X} \cup \{\Delta_A, \Delta_{A^c}\}$. Let Y be the Markov process with generator

$$\begin{aligned}(\mathcal{L}f)(x) = (Lf)(x) &+ q(x)\mathbb{1}_{\{x\in A\}}[f(\Delta_A) - f(x)] \\ &+ q(x)\mathbb{1}_{\{x\in A^c\}}[f(\Delta_{A^c}) - f(x)], \quad \text{if } x \in \mathcal{X},\end{aligned} \tag{3.15}$$

and $(\mathcal{L}f)(x) = 0$ if $x \in \{\Delta_A, \Delta_{A^c}\}$ with $f : \mathcal{Y} \to \mathbb{R}$ and L as in (1.1).

Next, consider the subspace $\bar{\bar{A}} := A \cup \{\Delta_A, \Delta_{A^c}\} \subset \mathcal{Y}$. Let $Y^{\bar{\bar{A}}}$ be the Markov process with state space $\bar{\bar{A}}$ obtained as the trace of the process Y on $\bar{\bar{A}}$. Let us remark two features of Wilson's algorithm. First, Wilson's algorithm can be extended to the case of a state space with more than one

absorbing state. In this case it produces a rooted spanning forest instead of a tree. Second, Wilson's algorithm is uniquely determined once we fix a state space with some absorbing set and a Markov generator. These observations justify the following definitions. Let D be the set of ending points of the edges starting from A after running Wilson algorithm on $\bar{\mathcal{X}}$ with absorbing state Δ and generator $\mathcal{L}$. Similarly let D' associated in the same way with A, when Wilson's algorithm is run on $\bar{\bar{A}}$ with absorbing set $\{\Delta_A, \Delta_{A^c}\}$ and generator $\mathcal{L}^{\bar{\bar{A}}}$. Observe at this point that

$$\mathbb{P}(A \subset \rho(\Phi)) = \mathbb{P}(D = \{\Delta\}) = \mathbb{P}(D' = \{\Delta_A\}),$$

and, by using Proposition 2.1, compute

$$\begin{aligned}\mathbb{P}\left(A \subset \rho\left(\Phi\right)\right) = \mathbb{P}(D' = \{\Delta_A\}) &= \frac{\prod_{a\in A} q(a)}{\det_{\bar{\bar{A}}\setminus\{\Delta_A,\Delta_{A^c}\}}(-\mathcal{L}^{\bar{\bar{A}}})} \\ &= \frac{\det_A(Q)}{\det_A(-\mathcal{L}^{\bar{\bar{A}}})} = \det\left(\left(\left[-\mathcal{L}^{\bar{\bar{A}}}\right]_A\right)^{-1}\right)\det_A(Q) \\ &= \det_A\left(G^{\bar{\bar{A}}}_{\{\Delta_A,\Delta_{A^c}\}}\right)\det_A(Q),\end{aligned} \tag{3.16}$$

where $G^{\bar{\bar{A}}}_{\{\Delta_A,\Delta_{A^c}\}}(x,y)$ denotes the Green's function of the process $Y^{\bar{\bar{A}}}$ stopped when entering the absorbing states $\{\Delta_A, \Delta_{A^c}\}$. Note now that for $x, y \in A$,

$$G^{\bar{\bar{A}}}_{\{\Delta_A,\Delta_{A^c}\}}(x,y) = G^{\mathcal{Y}}_{\{\Delta_A,\Delta_{A^c}\}}(x,y) = G^{\bar{\mathcal{X}}}_{\{\Delta\}}(x,y) \tag{3.17}$$

with $\bar{\mathcal{X}} = \mathcal{X} \cup \{\Delta\}$ and $G^{\bar{\mathcal{X}}}_{\{\Delta\}}$ being the Green's function of the process $\bar{X}$ in (2.12). Finally, since $G^{\bar{\mathcal{X}}}_{\{\Delta\}}(x,y) = (Q-L)^{-1}(x,y)$ for $x, y \in \mathcal{X}$, the claim follows by combining equations (3.16) and (3.17). □

Figure 2: On the one-dimensional torus our determinantal process is distributed like a simple gas of particles with a repulsive "nearest-neighbour two-body interaction" that is quite easily detected by comparison with a Bernoulli product measure. Here are samples of our roots set (first line of each picture) and of the product measure (second line) with exactly 6, 12, 18, 24, 30, 36, 42, 48, 54 and 60 roots or points, respectively.

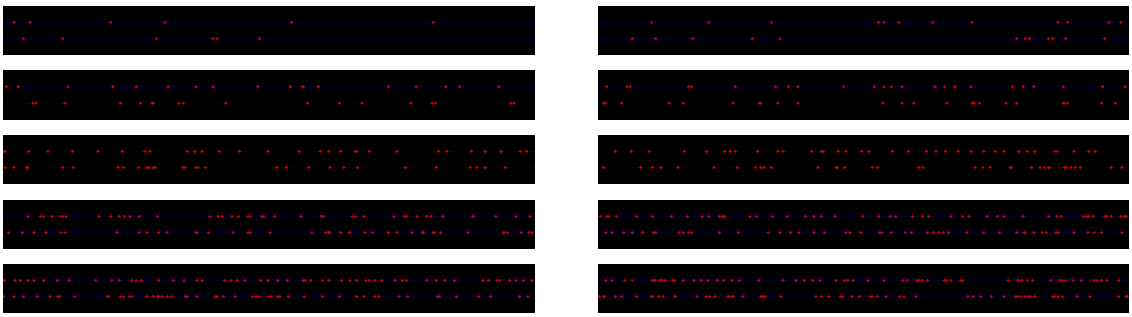
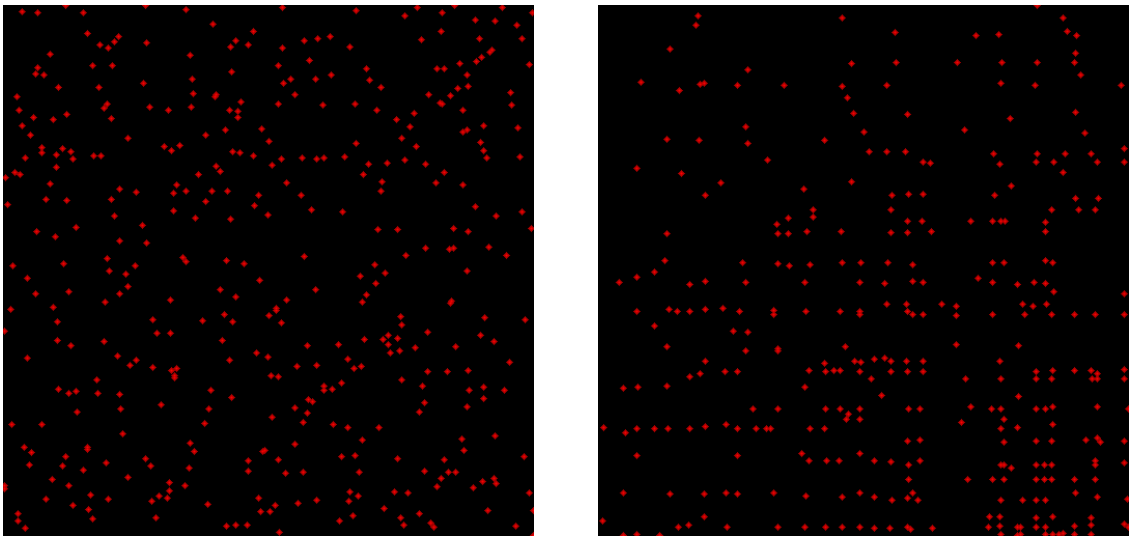


Figure 3: On the two dimensional torus, the difference between the law of our roots set and that of a Bernoulli product measure is far too subtle to be detected on a single sample. Walking away from translation invariant models, we can find much more correlations. Here are pictures of a sample on the two-dimensional torus with uniform rates equal to 1 and $q = .002$ (on the left), and of a sample associated with the Metropolis random walk on the square grid in a Brownian sheet potential with inverse temperature $\beta = .16$ and extinction rate $q = 10^{-4}$ (on the right).



3.2 Hitting times of the set of roots

In this section we answer the question raised in the introductory Section 1.1. To this end, we focus on hitting times of a given subset $R \subset \mathcal{X}$, i.e., $T_R := \inf\{t \geq 0 : X_t \in R\}$. In Appendix A.2, Lemma A.3, we derive, in terms of our measures, a result originally due to Freidlin and Wentzell, see Lemma 3.3 in [6], which allows to give a formula for $E_x[T_R]$. In this section, with the help of Lemma A.3, we compute the expectation $E_x[T_R]$ when R is the set of roots sampled from ν_q, or $\mathbb{P}_q(\cdot||\rho(\Phi)| = m)$ for any given m.

Theorem 3.5. (Hitting-time formulas) *Let ν_Q be the forest probability measure defined in* (1.10)*. Then, for any $x \in \mathcal{X}$,*

$$\mathbb{E}_Q\left[E_x\left[T_{\rho(\Phi)}\right]\right] = \sum_{y\in\mathcal{X}} \frac{1}{q(y)}\left[\mathbb{P}_Q\left(\rho\left(\tau_\Phi(x)\right) = \{y\}\right) - \mathbb{P}_Q(\rho(\Phi) = \{y\})\right], \tag{3.18}$$

with $\tau_\Phi(x)$ being the unique tree in Φ containing x. In the standard case, $q(x) \equiv q > 0$, equation (3.18) *reduces to*

$$\mathbb{E}_q\left[E_x\left[T_{\rho(\Phi)}\right]\right] = \frac{1}{q}\left[1 - \mathbb{P}_q\left(|\rho(\Phi)| = 1\right)\right] = \frac{1}{q}\left(1 - \prod_{i=1}^{n-1}\frac{\lambda_i}{q+\lambda_i}\right). \tag{3.19}$$

Moreover, for $m < n$ and any $x \in \mathcal{X}$,

$$\mathbb{E}_q\left[E_x\left[T_{\rho(\Phi)}\right] \mid |\rho(\Phi)| = m\right] = \frac{\mathbb{P}_q(|\rho(\Phi)| = m+1)}{q\mathbb{P}_q(|\rho(\Phi)| = m)}. \tag{3.20}$$

Remark: Notice, that the r.h.s. in (3.20) can simply be expressed as the ratio between two successive coefficients of the characteristic polynomial of the generator L.

Proof. Observe first that under ν_Q, the probability of sampling a given set $R \subset \mathcal{X}$ as set of roots can be written as

$$\mathbb{P}_Q(\rho(\Phi) = R) = \frac{\sum_{\phi:\rho(\phi)=R} w_Q(\phi)}{Z_Q} = \frac{Z_R(0)}{Z_Q}\det{}_{R\setminus S}(Q), \tag{3.21}$$

with $Z_R(0)$ as in (1.12). By using equation (3.21) together with (A.17), we have that

$$\begin{aligned}
\mathbb{E}_Q[E_x[T_{\rho(\Phi)}]] &= \sum_{R\neq\emptyset} \mathbb{P}_Q(\rho(\Phi) = R) E_x[T_R] \\
&= \sum_{R\neq\emptyset:\ R\supset S} \frac{\det_{R\setminus S}(Q)}{Z_Q} \sum_{y\notin R} \sum_{\substack{\phi:\rho(\tau_\phi(x))=\{y\},\\ \rho(\phi)=R\cup\{y\}}} w(\phi) \\
&= \frac{1}{Z_Q} \sum_{y\notin S} \sum_{\substack{R\neq\emptyset:\ R\supset S\\ R\not\ni y}} \sum_{\substack{\phi:\rho(\tau_\phi(x))=\{y\},\\ \rho(\phi)=R\cup\{y\}}} \frac{w_Q(\phi)}{q(y)} \\
&= \frac{1}{Z_Q} \sum_{y\notin S} \left[\sum_{\phi:\rho(\tau_\phi(x))=\{y\}} \frac{w_Q(\phi)}{q(y)} - \sum_{\phi:\rho(\phi)=\{y\}} \frac{w_Q(\phi)}{q(y)} \right] \\
&= \sum_{y\notin S} \frac{1}{q(y)} \left[\mathbb{P}_Q(\rho(\tau_\phi(x)) = \{y\}) - \mathbb{P}_Q(\rho(\Phi) = \{y\}) \right].
\end{aligned} \tag{3.22}$$

The restriction of summing over $y \notin S$ can be removed, since $1/q(y) = 0$ for $y \in S$. Hence (3.18) holds and (3.19) readily follows when $q(x) \equiv q > 0$. The proof of (3.20) follows by an analogous computation. □

Note that the r.h.s. of (3.19) and (3.20) is uniform in the starting point x. This latter observation allows to answer the questions in the introduction. In fact, no matter of the geometry of the graph and the weights we are considering, we can take random subsets $R = \rho(\Phi)$ with law $\mathbb{P}_q(\cdot||\rho(\Phi)| = m)$ and the formula in equation (3.20) says that the hitting times do not dependent on the starting point x.

To practically obtain a sample from ν_q with approximately $m \leq n$ roots when L has only real eigenvalues, one can use Wilson's algorithm and play with the parameter q as follows (see Figure 4). If q is such that

$$\sum_{i=0}^{n-1} \frac{q}{q+\lambda_i} = m, \tag{3.23}$$

one has an expected number of m roots with fluctuations of order $\sqrt{m}$ or smaller (see Corollary 3.2). In principle one should compute the eigenvalues of L, which is in general difficult for large n, and then solve equation (3.23) in q. To overcome this obstacle, a possible alternative procedure is the following:

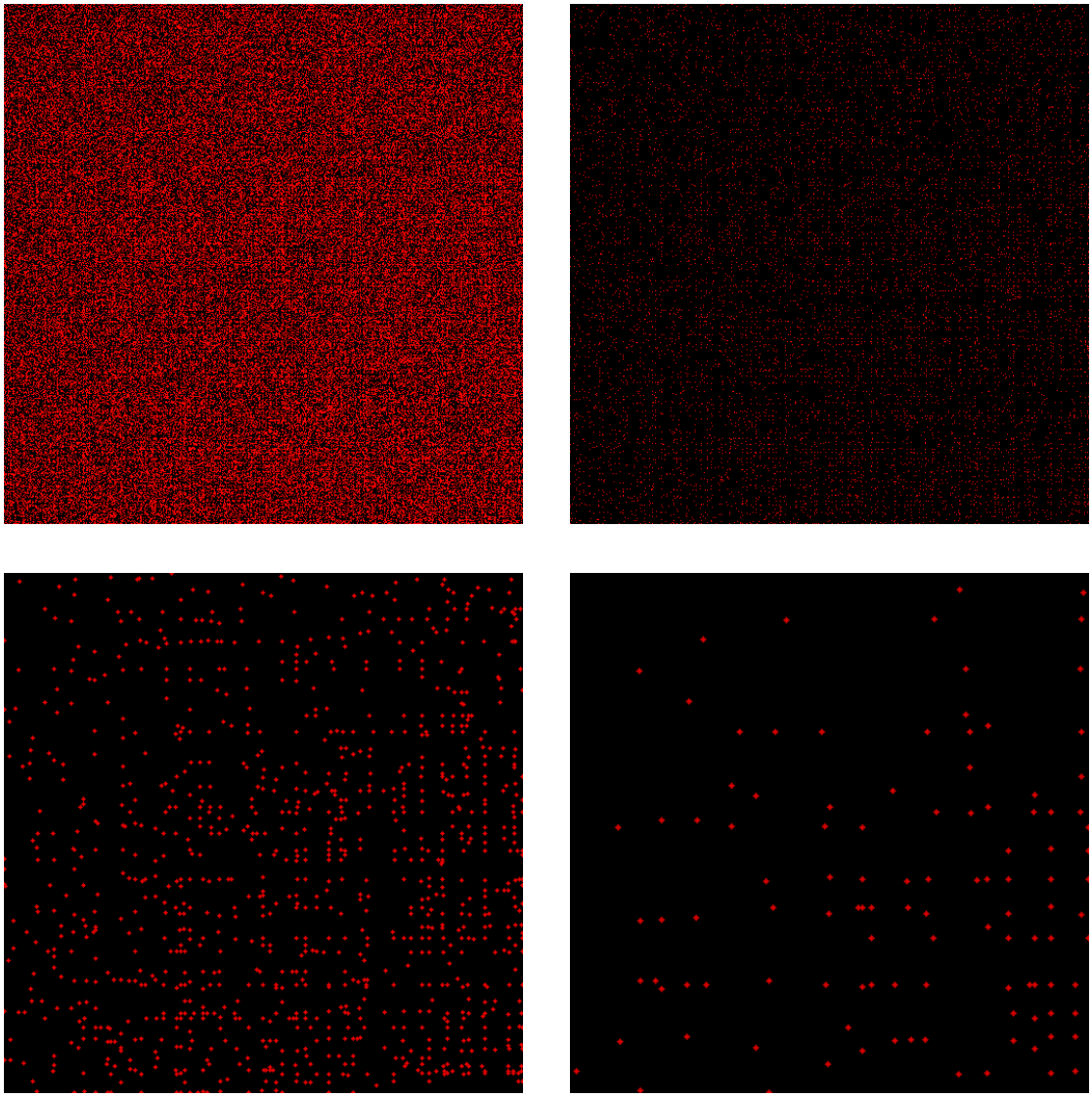
1 Start with any positive q and run Wilson's algorithm with this parameter to get a sample from ν_q with a certain number r of roots.

2 Replace q by $q * m/r$ and run again Wilson's algorithm with this new parameter to get a new sample with another number of roots, say r again.

3 Iterate the previous step until a sample with r roots satisfying $m - 2\sqrt{m} \leq r \leq m + 2\sqrt{m}$ is obtained.

As a matter of fact, $q \mapsto qm/\sum_i q/(q+\lambda_i)$ rapidly converges to the solution of (3.23), hence the algorithm rapidly reaches an end. Since we believe this procedure to be quite far from an optimal one, we are only sketchy on this point.

Figure 4: We sampled approximatively 100000, 10000, 1000 and 100 roots on the 512×512 grid for the random walk in a Brownian sheet potential with inverse temperature $\beta = .04$ by following the procedure described in Section 3.2. We obtained 100443, 10032, 1042 and 111 roots in 8, 6, 6 and 8 iterations, respectively.



We conclude this section with an estimate on the mean hitting times above when taking the maximum over the starting points x.

Proposition 3.6. (Max hitting-time estimates) *Under the standard measure ν_q, the following bound holds*

$$\mathbb{E}_q\left[\max_{x\in\mathcal{X}} E_x\left[T_{\rho(\Phi)}\right]\right] \leq \frac{1}{q}\left[\mathbb{E}_q\left[|\rho(\Phi)|\right] - \mathbb{P}_q(|\rho(\Phi)|=1)\right] = \frac{1}{q}\left(\sum_{i=0}^{n-1}\frac{q}{q+\lambda_i} - \prod_{i=1}^{n-1}\frac{\lambda_i}{q+\lambda_i}\right). \tag{3.24}$$

Furthermore, for any $k \leq n$, we have the estimate

$$\mathbb{E}_q \left[\max_{x \in \mathcal{X}} E_x \left[T_{\rho(\Phi)} \right] \mid |\rho(\Phi)| = m \right] \leq \frac{(m+1)\mathbb{P}_q \left(|\rho(\phi)| = m+1 \right)}{q \mathbb{P}_q \left(|\rho(\phi)| = m \right)}. \tag{3.25}$$

Proof. As we remarked before, Wilson's algorithm works also when considering more than one absorbing state. Denote by T_R^{Wil} the running time of Wilson's algorithm (i.e. the total running time of the loop erased random walks needed to cover the whole graph) when the absorbing states form a non-empty subset R of $\mathcal{X}$ (this amounts to sample $\nu_{0,R}$). It can be shown (see e.g. [12], Proposition 1) that the mean running time can be expressed in spectral terms as $\sum_{i=0}^{|R|-1} \frac{1}{\lambda_{i,R}}$, with $\lambda_{i,R}$ being the eigenvalues of the operator $[L]_{R^c}$, the sub-Markovian generator associated with the process absorbed in R[1]. Note at this point that we can overestimate the l.h.s. of (3.24) by the expectation of $T^{Wil}_{\rho(\Phi)}$. Hence, using (3.5) and looking at the coefficient of degree 1 in $Z_R(q)$,

$$\begin{aligned}
\mathbb{E}_q \left[\max_{x \in \mathcal{X}} E_x \left[T_{\rho(\Phi)} \right] \right] \leq \mathbb{E}_q \left[T^{Wil}_{\rho(\Phi)} \right] &= \mathbb{E}_q \left[\sum_{i=0}^{|\rho(\Phi)|-1} \frac{1}{\lambda_{i,\rho(\Phi)}} \right] = \sum_{R \neq \emptyset} \frac{Z_R(0)}{Z(q)} q^{|R|} \sum_{i=0}^{|R|-1} \frac{1}{\lambda_{i,R}} \\
&= \frac{1}{Z(q)} \sum_{R \neq \emptyset} q^{|R|} \sum_{\phi : \rho(\phi) \supset R, |\rho(\Phi)| = |R|+1} \frac{w_{q,R}(\phi)}{q} \\
&= \frac{1}{qZ(q)} \sum_{k=1}^{n} \sum_{R : |R|=k} \sum_{\phi : \rho(\phi) \supset R, |\rho(\phi)| = k+1} q^{|R|} w_{q,R}(\phi) \\
&= \frac{1}{qZ(q)} \sum_{k=1}^{n} \sum_{\phi : |\rho(\phi)| = k+1} w_q(\phi)(k+1) \\
&= \frac{1}{q} \sum_{\phi : |\rho(\phi)| \geq 2} \frac{w_q(\phi)|\rho(\phi)|}{Z(q)},
\end{aligned} \tag{3.26}$$

and the latter equals the r.h.s. of (3.24). The bounds in (3.25) follows by a similar argument. □

Remark: We expect these estimates to be good either when q is very small or very large, or when m is close to 1 or n. Improving them in the intermediate regime seems a challenging problem we were not able to solve.

3.3 Re-reading Micchelli-Willoughby proof

Throughout this section we work with the Markov process X on $\mathcal{X}$ in (1.1), *under the assumption that X is reversible with respect to some probability measure μ on $\mathcal{X}$*, i.e., L is a self-adjoint operator in $l^2(\mu)$ endowed with the inner product

$$\langle f, g \rangle_\mu := \sum_{x \in \mathcal{X}} \mu(x) f(x) g(x). \tag{3.27}$$

For $R \subsetneq \mathcal{X}$, possibly $R = \emptyset$, we turn the points of R into "absorbing points" by adding infinite weight edges towards a cemetery Δ and we assume the resulting Markov process with sub-Markovian

[1] This running time is actually independent of the obtained sample and its law is the same as that of a sum of independent exponential variables with parameters $\lambda_{i,R}$ when these eigenvalues are real. The same holds in the case of complex eigenvalues by defining "the sum of exponential variables" through its Laplace transform and the same algebraic formula as in the real case.

generator $[L]_{R^c}$ to be irreducible. We denote by $\lambda_{0,R} < \lambda_{1,R} \leq \cdots \leq \lambda_{l-1,R}$, with $l = n - |R|$, the eigenvalues of $[-L]_{R^c}$, and following [3, 5, 14], we define, for each x in R^c, a sequence of *local equilibria* by setting

$$\nu^x_{l-1} = \delta_x, \tag{3.28}$$

$$\nu^x_{k-1} = \nu^x_k \frac{[L]_{R^c} + \lambda_{k,R}}{\lambda_{k,R}}, \quad 1 \leq k \leq l-1. \tag{3.29}$$

Theorem 3.2 in [13] is a statement on symmetric matrices that in our setting can be stated as follows.

Theorem 3.7. (Micchelli and Willoughby [13]) *Fix an arbitrary $x \in \mathcal{X}$, for all $k < l$, ν^x_k in equation* (3.28) *is a non-negative measure.*

In this section we give a proof of this theorem following the key steps of Micchelli and Willoughby's algebraic proof, however, unlike the original proof, we develop probabilistic or combinatorial arguments. Before starting the proof we note, following [14], that equation (3.29) can be rewritten as

$$\nu^x_k [L]_{R^c} = \lambda_{k,R}(\nu_{k-1} - \nu_k), \tag{3.30}$$

which gives the following interpretation. The process leaves the measure, or "state", ν^x_k at rate $\lambda_{k,R}$ to be absorbed in R or to decay into ν^x_{k-1}. Provided that ν^x_k and ν^x_{k-1} are indeed non-negative measures, this can be turned into a rigorous mathematical statement [12]. Then, by looking at the different decay times up to an exponential time T_q that is independent from the process, and by observing that, by Hamilton-Cayley theorem, the process leaves the state ν^x_0 at rate $\lambda_{0,R}$ only to be absorbed in R, we get, for all x and y in R^c,

$$\begin{aligned} P_x(X(T_q \wedge T_R) = y) = &\frac{q}{q+\lambda_{l-1,R}}\nu^x_{l-1}(y) + \frac{\lambda_{l-1,R}}{q+\lambda_{l-1,R}}\frac{q}{q+\lambda_{l-2,R}}\nu^x_{l-2}(y) \\ &+ \cdots + \frac{\lambda_{l-1,R}}{q+\lambda_{l-1,R}} \cdots \frac{\lambda_{2,R}}{q+\lambda_{2,R}}\frac{q}{q+\lambda_{1,R}}\nu^x_1(y) \\ &+ \frac{\lambda_{l-1,R}}{q+\lambda_{l-1,R}} \cdots \frac{\lambda_{1,R}}{q+\lambda_{1,R}}\frac{q}{q+\lambda_{0,R}}\nu^x_0(y). \end{aligned} \tag{3.31}$$

The left hand side in (3.31) is the probability to have $\rho(\tau_\phi(x)) = \{y\}$ when ϕ is sampled from $\nu_{q,R}$. Then, multiplying by $Z_R(q) = \prod_i (q + \lambda_{i,R})$ (recall (3.5)), dividing by q, and denoting the result by $W_R(q)(x,y)$, we can rewrite (3.31) as

$$W_R(q)(x,y) := Z_R(q)[q\mathbb{1} - L]^{-1}_{R^c}(x,y) \tag{3.32}$$

$$= \frac{1}{q} \sum_{\substack{\phi:\, \rho(\tau_\phi(x))=\{y\} \\ \rho(\phi)\supseteq R}} q^{|\rho(\phi)|-|R|} w(\phi) \tag{3.33}$$

$$\begin{aligned} = &(q+\lambda_{0,R})\cdots(q+\lambda_{l-2,R})\nu^x_{l-1}(y) \\ &+(q+\lambda_{0,R})\cdots(q+\lambda_{l-3,R})\lambda_{l-1,R}\nu^x_{l-2}(y) \\ &+\cdots+(q+\lambda_{0,R})\lambda_{l-1,R}\lambda_{l-2,R}\cdots\lambda_{2,R}\nu^x_1(y) \\ &+\lambda_{n-1,R}\cdots\lambda_{1,R}\nu^x_0(y). \end{aligned} \tag{3.34}$$

Next, by density and continuity, we can restrict ourselves to the case of distinct eigenvalues, and equation (3.34) suggests, for $k < l$, the following relation for the divided differences (see Definition A.23 in Appendix A.3) of W_R:

$$W_R[-\lambda_{0,R}, \ldots, -\lambda_{k,R}](x,\cdot) = \lambda_{l-1,R}\cdots\lambda_{k+1,R}\nu^x_k = \delta_x \prod_{i=k+1}^{l-1}([L]_{R^c} + \lambda_{i,R}), \tag{3.35}$$

that is

$$W_R[-\lambda_{0,R},\dots,-\lambda_{k,R}] = \prod_{i=k+1}^{l-1} ([L]_{R^c} + \lambda_{i,R}). \tag{3.36}$$

We can now start to follow the main steps of Micchelli and Willoughby's proof.

Step 1: Checking equation (3.36).
We simply use Definition A.22 and spectral decomposition. With μ_i being the right eigenvector associated with $\lambda_{i,R}$, and, for any measure ν, $\langle \mu_i, \nu\rangle = \sum_{x\notin R} \mu_i(x)\nu(x)/\mu(x)$, we have, recalling (3.32), for any q,

$$\nu W_R(q) = \sum_{i=1}^{n-1} \langle \mu_i, \nu\rangle \prod_{i\neq j} (q+\lambda_{i,R})\mu_j. \tag{3.37}$$

This gives

$$\begin{aligned}
\nu W_R[-\lambda_{0,R}\dots,-\lambda_{k,R}] &= \sum_{r=0}^{k} \frac{\nu W_R(-\lambda_{r,R})}{\prod_{m\neq r}(\lambda_{m,R}-\lambda_{r,R})} = \sum_{r=0}^{k}\sum_{j=1}^{l-1} \langle \mu_j,\nu\rangle \frac{\prod_{i\neq j}(\lambda_{i,R}-\lambda_{r,R})\mu_j}{\prod_{m\neq r}(\lambda_{m,R}-\lambda_{r,R})} \\
&= \sum_{r=0}^{k} \langle \mu_r,\nu\rangle \frac{\prod_{i\neq j}(\lambda_{i,R}-\lambda_{r,R})\mu_r}{\prod_{m\neq r}(\lambda_{m,R}-\lambda_{r,R})} = \sum_{r=0}^{k} \langle \mu_r,\nu\rangle \prod_{i=k+1}^{l-1} (\lambda_{i,R}-\lambda_{r,R})\mu_r \\
&= \sum_{r=0}^{l-1} \langle \mu_r,\nu\rangle \prod_{i=k+1}^{l-1} (\lambda_{i,R}-\lambda_{r,R})\mu_r = \nu \prod_{i=k+1}^{l-1} ([L]_{R^c}+\lambda_{i,R})
\end{aligned} \tag{3.38}$$

and equation (3.36) readily follows.

Step 2: A combinatorial identity.
The key point of the proof lies in the following lemma.

Lemma 3.8. *For any $x \neq y$ in $\mathcal{X}\setminus R$,*

$$W_R(q)(x,y) = w(x,y)Z_{R\cup\{x,y\}}(q) + \sum_{z,z'\in\mathcal{X}\setminus(R\cup\{x,y\})} w(x,z)W_{R\cup\{x,y\}}(q)(z,z')w(z',y) \tag{3.39}$$

In addition one has

$$W_R(q)(x,x) = Z_{R\cup\{x\}}(q). \tag{3.40}$$

Proof. Let us first consider the case $x \neq y$. Due to (3.33), we have that

$$W_R(q)(x,y) = \sum_{\substack{\phi:\,\rho(\tau_\phi(x))=\{y\}\\ \rho(\phi)\supseteq R}} q^{|\rho(\phi)|-1-|R|} w(\phi) \tag{3.41}$$

We also have

$$Z_{R\cup\{x,y\}}(q) = \sum_{\phi':\rho(\phi')\supseteq R\cup\{x,y\}} q^{|\rho(\phi')|-2-|R|} w(\phi') \tag{3.42}$$

and

$$W_{R\cup\{x,y\}}(q)(z,z') = \frac{1}{q} \sum_{\substack{\phi'':\,\rho(\tau_{\phi''}(z))=\{z'\}\\ \rho(\phi)\supseteq R\cup\{x,y\}}} q^{|\rho(\phi)|-2-|R|} w(\phi) = \sum_{\substack{\phi'':\,\rho(\tau_{\phi''}(z))=\{z'\}\\ \rho(\phi)\supseteq R\cup\{x,y\}}} q^{|\rho(\phi)|-3-|R|} w(\phi). \tag{3.43}$$

Next, define for each ϕ in (3.41), $\phi' := \phi \setminus \{(x,y)\}$ if (x,y) belongs to ϕ, and $\phi'' := \phi \setminus \{(x,z);(z',y)\}$ if x is connected in ϕ to y through z and z' (possibly with $z = z'$) in such a way that (x,z) and (z',y) belong to ϕ. Finally, by observing that $|\rho(\phi')| = |\rho(\phi)| + 1$ and $|\rho(\phi'')| = |\rho(\phi)| + 2$, (3.39) is obtained from (3.41), (3.42) and (3.43). Then, equation (3.40) follows from (3.41) for $y = x \notin R$. □

Step 3: Conclusion by induction with Cauchy interlacement theorem.
For $l \geq 0$, let $\mathcal{P}[l]$ be the following statement:

> For all $R \subset \mathcal{X}$ such that $|R^c| = l$, for all $L \geq l$, for all $\xi_0 > \xi_1 > \cdots > \xi_L$ such that $\xi_i \geq -\lambda_{i,R}$ for all $i < l$, for all $k \leq L$, and for all $x, y \notin R$:
>
> $$W_R[\xi_0, \ldots, \xi_k](x,y) \geq 0. \tag{3.44}$$

We can proceed inductively to show that $\mathcal{P}[l]$ holds.

For $l = 0, 1$, the claim is obvious. Fix $l \geq 2$. In the case $x = y$, the inductive hypothesis is unnecessary. Indeed, from (3.40), one has

$$W_R[\xi_0, \ldots, \xi_k](x,x) = Z_{R\cup\{x\}}[\xi_0, \ldots, \xi_k]. \tag{3.45}$$

Then note that, by Cauchy interlacement theorem, $\xi_i \geq -\lambda_{i,R}$ implies that $\xi_i \geq -\lambda_{i,R\cup\{x\}}$ for $i < l-1$, and hence, by Lemma A.7, we get

$$W_R[\xi_0, \ldots, \xi_k](x,x) \geq 0. \tag{3.46}$$

When $x \neq y$, $\mathcal{P}[l]$ follows in the same way by using (3.39) and the inductive hypothesis.

We can finally conclude the proof of Theorem 3.7. It suffices to apply the claim with $\xi_i = -\lambda_{i,R}$ for all $i < l$ and recall (3.36) or (3.35).

3.4 Coalescence and fragmentation processes

In this section we present two coalescence and fragmentation processes closely related with our forest measures.

3.4.1 Coupling the forest measures for different values of q.

To build the first process, we couple all the ν_q's together for different q's. This coupling can be seen as a coalescence and fragmentation process when q decreases to 0 and $t = 1/q$ is thought as time.

We make use of Wilson's original representation of his algorithm with "site-indexed random paths". Assume that, to each site of the graph, is attached an infinite list/collection of arrows pointing towards one neighbour, and that these arrows are independently distributed according to the discrete skeleton transition probabilities. In other words, an arrow, pointing towards the neighbour y of a site x, appears at each level in the list associated with x with probability $w(x,y)/\bar{w}$ (in this context, we set $w(x,x) = \bar{w} - \sum_{y\neq x} w(x,y)$, and consider x itself as one of its possible neighbours). Imagine that each list of arrows attached to a site is piled down in such a way that it make sense to talk of an infinite stack with an arrow on the top of this stack. By using this representation, one can generate the random walk on the graph as follows. At each jump time, the random walk steps to the neighbour pointed by the arrow on the top of the stack where the walker is sitting, and the top arrow is erased from the stack.

To describe Wilson's algorithm one has to introduce a further ingredient: pointers to the absorbing state Δ in each stack. Such a pointer should appear with probability $q/(\bar{w}+q)$ at each level in the list.

One way to introduce it is by generating independent uniform random variables U together with each original arrow in the stack, and by replacing the latter by a pointer to the absorbing state whenever $U < q/(\bar{w} + q)$.

A possible description of Wilson's algorithm is then the following.

i. Start with a particle on each site, particles and sites will be divided into active and frozen particles or sites and, at the beginning, all sites and particles are declared to be active.

ii. Choose a particle among all the active ones and look at the arrow at the top of the stack it is seated on. Call x the site where the particle is seated.

   - If the arrow is the pointer towards Δ, declare the particle to be frozen and site x as well.
   - If the arrow points towards another site $y \neq x$, remove the particle and keep the arrow.
   - If the arrow points to x itself, remove the arrow.

iii. Once again, choose a particle among all the active ones, look at the arrow on the top of the stack it is seated on, and call x the site where the particle is seated.

   - If the arrow points to Δ, the particle is declared to be frozen, and so are declared x and all the sites eventually leading to x by following discovered top pile arrows paths.
   - If the arrow points to a frozen particle or site, remove the particle, keep the arrow, and freeze site x as well as any site eventually leading to x by following discovered top pile arrows paths.
   - If the arrow points to an active, then there are two possibilities. By following the arrows at the top of the stacks, we either reach a different active particle, or runs in a loop back to x. In the former case, remove the chosen particle from site x and keep the discovered arrow. In the latter, erase all the arrows along the loop and set an active particle on each site of the loop. Note that this last case includes the possibility for the discovered arrow of pointing to x itself, in which case, we just have to remove the discovered arrow.

iv. Iterate the previous steps up to exhaustion of the active particles.

The crucial observation is that, no matter of the choice of the particles at the beginning of the steps, when this algorithm stops, the same arrows are erased and the same spanning forest with a frozen particle at each root is obtained. In particular, by choosing at each step the last encountered active particle, or the same as in the previous step when we just erased a loop, we perform a simple loop-erased random walk up to freezing.

Since ν_q is sampled in this way, and the same uniform variables U can be used for each q, this provides a coupling among all the ν_q's. By means of this coupling, one can actually sample $\nu_{q'}$ starting from a sample of ν_q for $q' < q$. Let us now explain this fact. Note first that, running this algorithm for sampling $\nu_{q'}$, one can reach at some point the final configuration obtained for ν_q with the only difference that some frozen particles of the final configuration obtained with parameter q can still be active at this intermediate step of the algorithm run with q'. It suffices, indeed, to choose the sequence of active particles in the same way with both parameters. This is possible since each pointer to Δ in the stacks with parameter q' is associated with a pointer to Δ at the same level in the stacks with parameter q. Thus we just have to replace some frozen particle of the configuration sampled with ν_q and continue the algorithm with parameter q' to sample $\nu_{q'}$. To do so, it suffices to declare active each frozen particle at site x with probability

$$p = P\left(U > \frac{q'}{\bar{w} + q'} \,\middle|\, U < \frac{q}{\bar{w} + q}\right) = \frac{\bar{w}(q - q')}{q(q' + \bar{w})},$$

then, at the top of the stack set an arrow that points toward y with probability $w(x,y)/\bar{w}$, and to keep x frozen with probability $1-p$.

When q continuously decreases, we obtain a coalescence-fragmentation process ξ such that $\xi(t)$ is distributed according to $\nu_{1/t}$ at all time t, and in which each tree can fragment and partially coalesce with the other trees of the forest. When a root of a tree turns active, the tree is eventually fragmented into a forest, some trees of which being possibly "graftedön the previous frozen trees. It is worth noting that, by (3.6), the mean number of trees is decreasing along this coalescence process ξ.

The previous observations show that we can sample the "finite dimensional distributions" of the process, i.e. the law of $(\xi(t_1),\xi(t_2),\dots,\xi(t_k))$ for any choice of $0<t_1<t_2<\cdots<t_k$. We can actually sample whole trajectories $(\xi(t))_{0\le t\le T}$ for any finite T. In fact, note first that at each time $t=1/q$, the next frozen particle (or root) becoming active is uniformly distributed among all the roots, and the time σ when it "wakes upïs such that the variable

$$V := \frac{1/\sigma}{\bar{w}+1/\sigma} \tag{3.47}$$

has the law of the maximum of m independent uniform variables on $[0, q/(\bar{w}+q))$, with m being the number of roots at time t. Since, for all $v<q/(\bar{w}+q)$,

$$\mathbb{P}(V<v) = \left(\frac{v}{q/(\bar{w}+q)}\right)^m = \left(\frac{v(\bar{w}+q)}{q}\right)^m, \tag{3.48}$$

V has the same law as $qU^{1/m}/(\bar{w}+q)$, with U uniform on $[0,1)$. Using (3.47) we can then sample σ by setting

$$\sigma = \frac{\bar{w}+q-qU^{1/m}}{q\bar{w}U^{1/m}} = t\,\frac{\bar{w}+(1-U^{1/m})/t}{\bar{w}U^{1/m}}. \tag{3.49}$$

Summing up, in order to sample the whole trajectory it suffices to proceed as follows once ξ is sampled at a given jump time t:

- Choose uniformly a root x.
- Sample the next jump time σ from a uniform random variable U on $[0,1)$, by using (3.49).
- Restart the algorithm with parameter $1/\sigma$ by declaring active the particle in x and putting an arrow to y with probability $w(x,y)/\bar{w}$.

We conclude by observing that ξ "crossesälmost surely all the manifolds $\mathcal{F}_m := \{\phi\in\mathcal{F}:\ |\rho(\phi)|=m\}$. Indeed, by (3.6), it starts from $\mathcal{F}_n$ and reaches $\mathcal{F}_1$ almost surely, and, each time the number of roots decreases, it does so by only 1 unit: when the "activated"tree fragments into trees that coalesce only with the previously frozen ones. With $T_m=\min\{t\ge 0:\ \xi(t)\in\mathcal{F}_m\}$, it is also simple to sample $\xi(T_m)$. We note however that $\xi(T_m)$ is not distributed according to $\mathbb{P}(\Phi\in\cdot\,|\,|\rho(\Phi)|=m)$.

3.4.2 Forest measures as invariant measures for coalescence and fragmentation processes

The other dynamics we want to mention is a simple variant of the tree random walk introduced in [1] to prove the so called Markov chain tree theorem. For fixed q, the dynamics we now present is another coalescence and fragmentation process for which the standard measure ν_q is the stationary probability measure.

Remind that for a given forest $\phi\in\mathcal{F}$ and $x\in\mathcal{X}$, we denote by $\tau_\phi(x)$ the unique tree in ϕ containing x, and note that if $\underline{e}\in\rho(\phi)$ then $e\notin\phi$. Our dynamics can then be defined as follows. 1

Figure 5: Snapshots of the coalescence and fragmentation process in Section 3.4.1 associated with a simple random walk on the two-dimensional torus of size 512×512 at times $t = 1/q$ equal to 0, .5, 2, 8, 32, 128, 512, . . . , 524288. Roots are red, non-root vertices at the border of trees are cyan, other vertices are in blue.

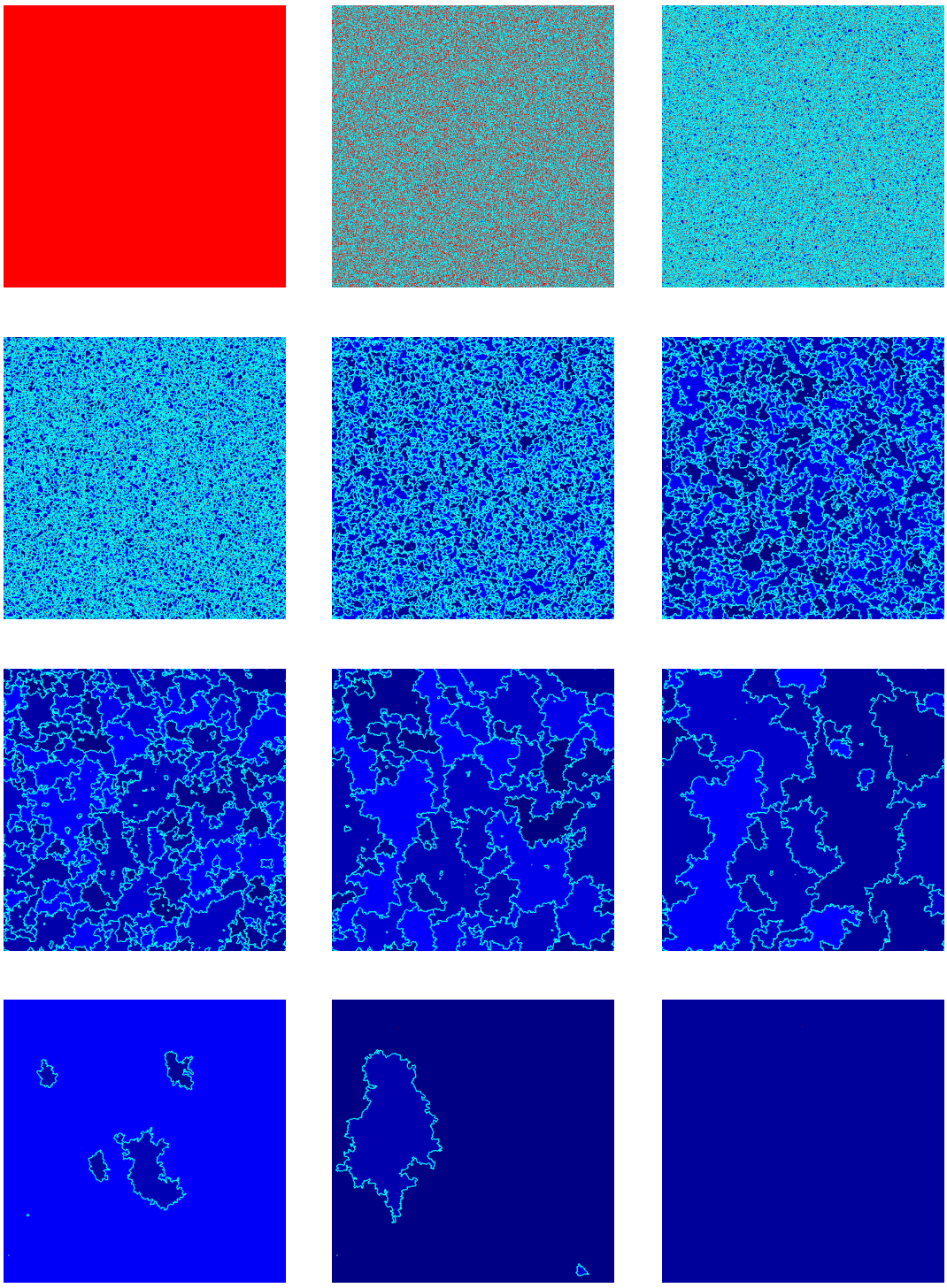


Figure 6: Snapshots at times $t = 1/q$ equal to 0, .5, 2, 8, 32, 128, 512, . . . , 524288 of the coalescence and fragmentation process on the square grid of size 512×512 for the random walk in a Brownian sheet potential with inverse temperature $\beta = .16$.

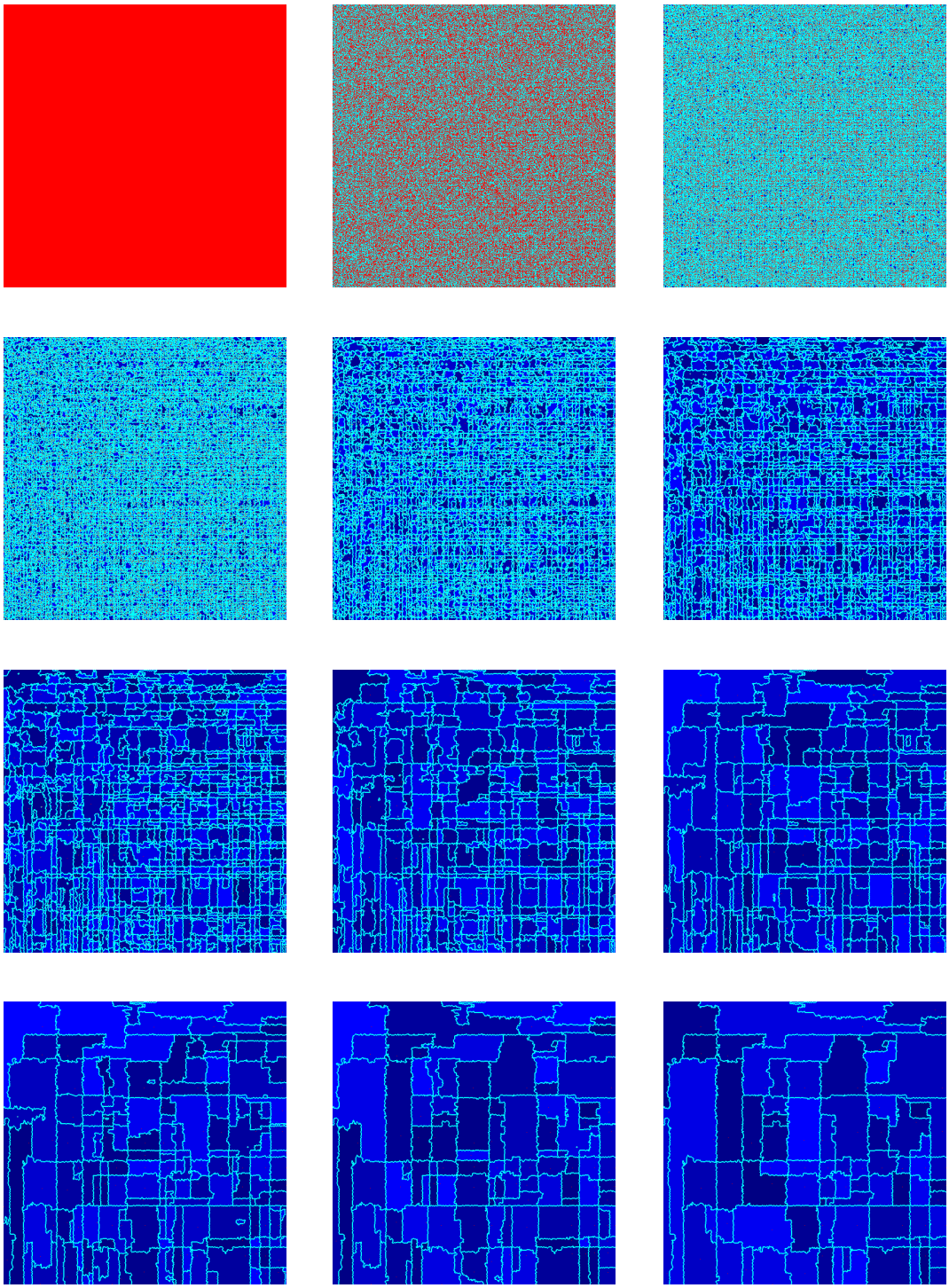
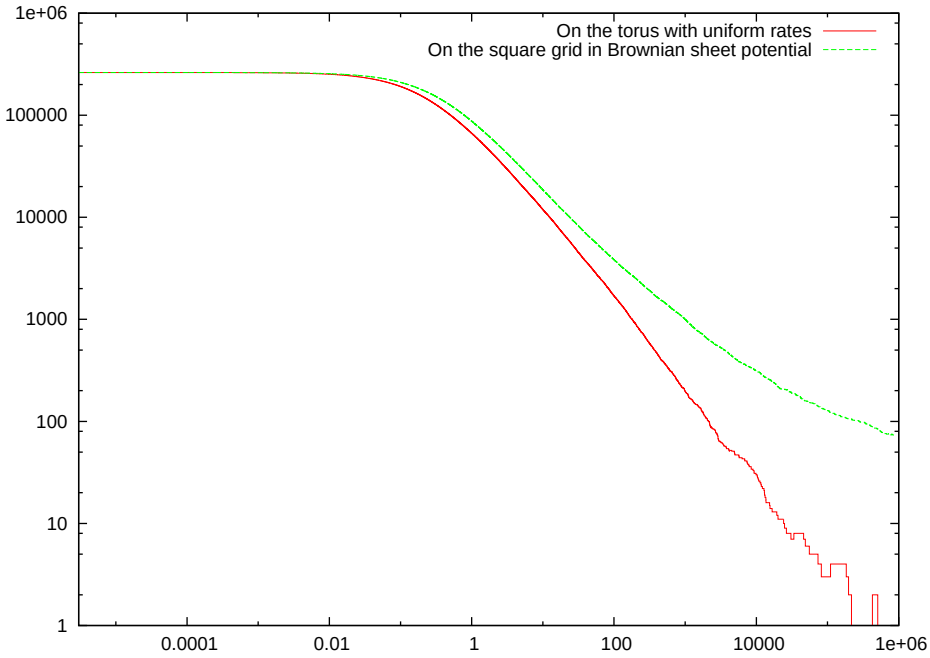


Figure 7: Numbers of trees as a function of time for the coalescence and fragmentation processes in Figures 5 and 6 in loglog scale.



Definition 3.9. (Forest Dynamics) *Fix $q \in [0, \infty)$. Let ψ be the Markov process with state space $\mathcal{F}$ characterized by the following generator acting on functions $f : \mathcal{F} \to \mathbb{R}$:*

$$(\mathcal{G}f)(\phi) = \sum_{e \in \mathcal{E}} \gamma(\phi, e)[f(\phi^e) - f(\phi)], \tag{3.50}$$

where the transition rate $\gamma(\phi, e)$ and the new state ϕ^e are defined as follows:

1. *If $\underline{e} \in \rho(\phi)$ and $\overline{e} \notin \tau_\phi(\underline{e})$, then $\gamma(\phi, e) = w(\underline{e}, \overline{e})$ and $\phi^e = \phi \cup \{e\}$.*
2. *If $\underline{e} \in \rho(\phi)$ and $\overline{e} \in \tau_\phi(\underline{e})$, then $\gamma(\phi, e) = w(\underline{e}, \overline{e})$ and $\phi^e = \phi \cup \{e\} \setminus \{e'\}$, with e' being the unique edge in ϕ such that $\underline{e'} = \overline{e}$.*
3. *If $e \in \phi$, then $\gamma(\phi, e) = q$ and $\phi^e = \phi \setminus \{e\}$.*
4. *$\gamma(\phi, b) = 0$ else.*

The rules corresponding to 1, 2 and 3 can be rephrased by saying that we *add, swap* and *remove* a bond from the forest ϕ, respectively. Notice that such a dynamics induces a non-conservative dynamics on the set of roots. In particular, when transition 1 occurs, $|\rho(\phi^e)| = |\rho(\phi)| - 1$ and two trees merge into one. On the other hand, when transition 3 occurs, $|\rho(\phi^e)| = |\rho(\phi)| + 1$ and the tree containing e is fragmented into two trees where the new appearing root is at $\underline{e}$. Transitions of type 2 produce a rearranging in one of the tree. They leave invariant the cardinality of the set of roots but the location of the root in the modified tree is relocated at the vertex $\underline{e'} = \overline{e}$.

Proposition 3.10. (Invariance) *For all $q > 0$ the measure w_q in* (1.8) *is invariant for $\mathcal{G}$ in* (3.50)*.*

Proof. We need to show that for any $\phi' \in \mathcal{F}$

$$\sum_{\phi \in \mathcal{F}} \sum_{e \in \mathcal{E} : \phi^e = \phi'} w_q(\phi)\gamma(\phi, e) = \sum_{e \in \mathcal{E}} w_q(\phi')\gamma(\phi', e). \tag{3.51}$$

Indeed, (3.51) amounts to say that for any $f : \mathcal{F} \to \mathbb{R}$

$$\begin{aligned}\sum_{\phi\in\mathcal{F}} w_q(\phi)(\mathcal{G}f)(\phi) &= \sum_{\phi\in\mathcal{F}} w_q(\phi)\sum_{e\in\mathcal{E}}\gamma(\phi,e)[f(\phi^e)-f(\phi)]\\ &= \sum_{\phi'\in\mathcal{F}}\sum_{\phi\in\mathcal{F}} w_q(\phi)\sum_{e\in\mathcal{E}:\phi^e=\phi'}\gamma(\phi,e)f(\phi^e) - \sum_{\phi'\in\mathcal{F}} w_q(\phi')\sum_{e\in\mathcal{E}}\gamma(\phi',e)f(\phi')\\ &= \sum_{\phi'\in\mathcal{F}}\Big[\sum_{\phi\in\mathcal{F}}\sum_{e\in\mathcal{E}:\phi^e=\phi'} w_q(\phi)\gamma(\phi,e) - \sum_{e\in\mathcal{E}} w_q(\phi')\gamma(\phi',e)\Big]f(\phi') = 0.\end{aligned}$$

For a given $\phi' \in \mathcal{F}$, we start by splitting the l.h.s. of (3.51) in the three terms corresponding to the different transitions allowed by the dynamics in Definition 3.9 whenever $\gamma(\phi, e) > 0$.

$$\begin{aligned}\sum_{\phi\in\mathcal{F}}\sum_{e\in\mathcal{E}:\phi^e=\phi'} w_q(\phi)\gamma(\phi,e) &= \sum_{\phi\subsetneq\phi'}\sum_{e\in\mathcal{E}:\phi^e=\phi'} w_q(\phi)\gamma(\phi,e) + \sum_{\phi\supsetneq\phi'}\sum_{e\in\mathcal{E}:\phi^e=\phi'} w_q(\phi)\gamma(\phi,e)\\ &\quad + \sum_{\phi:|\rho(\phi)|=|\rho(\phi')|}\sum_{e\in\mathcal{E}:\phi^e=\phi'} w_q(\phi)\gamma(\phi,e) = I + II + III.\end{aligned}$$

We can rewrite

$$I = \sum_{e\in\phi'} w_q(\phi'\setminus\{e\})\gamma(\phi'\setminus\{e\},e) = \sum_{e\in\phi'} w_q(\phi')\gamma(\phi',e), \tag{3.52}$$

$$II = \sum_{\underline{e}\in\rho(\phi'),\overline{e}\notin\tau_{\phi'}(\underline{e})} w_q(\phi'\cup\{e\})\gamma(\phi'\cup\{e\},e) = \sum_{\underline{e}\in\rho(\phi'),\overline{e}\notin\tau_{\phi'}(\underline{e})} w_q(\phi')\gamma(\phi',e), \tag{3.53}$$

and, denoting, for each e' such that $\underline{e'} \in \rho(\phi')$ and $\overline{e'} \in \tau_{\phi'}(\underline{e'})$, by e the unique bond in the only one cycle of $\phi' \cup \{e'\}$ such that $\overline{e} \in \rho(\phi')$, with $\phi = \phi' \setminus \{e\} \cup \{e'\}$,

$$III = \sum_{\underline{e'}\in\rho(\phi'),\overline{e'}\in\tau_{\phi'}(\underline{e'})} w_q(\phi)\gamma(\phi,e) = \sum_{\underline{e'}\in\rho(\phi'),\overline{e'}\in\tau_{\phi'}(\underline{e'})} w_q(\phi')\gamma(\phi',e'). \tag{3.54}$$

Summing I, II and III together we then get (3.51). □

Remark: When $q = 0$ we recover the Anantharam and Tsoukas dynamics and the proof of the Markov chain tree theorem. Indeed, the standard forest measure restricted to spanning trees is the invariant measure of the dynamics. Starting with a single tree, its roots follows a Markovian evolution with generator L, so that, at equilibrium, in the long time limit we have

$$\mu(x) = \frac{\sum_{\phi:|\rho(\phi)|=1} w(\phi)\mathbb{1}_{\{\rho(\phi)=\{x\}\}}}{\sum_{\phi:|\rho(\phi)|=1} w(\phi)}, \tag{3.55}$$

with μ being the stationary distribution associated to L.

3.5 Two last observations on the roots distribution

We conclude with two last results concerning static properties of the standard measure ν_q. The first one concerns the "rooted partition" inherited from a forest sampled from ν_q. We show, as a consequence of the Markov chain tree theorem, that the roots are distributed according to the restricted measures associated with the sampled partition, these restricted measures being the equilibrium measures of the restricted dynamics inside each component of the partition. Our last result concerns the determinantal process $\rho(\Phi)$: its cumulants are given by a nice formula it is worth to notice.

3.5.1 On the rooted partition sampled from the standard measure

For ϕ in $\mathcal{F}$ with roots $x_1, \ldots, x_m$, let us denote by $\mathcal{P}(\phi) = \{A_1, \ldots, A_m\}$ the partition of $\mathcal{X}$ where each component A_i is the set of sites spanned by $\tau_\phi(x_i)$. Since each component of the partition comes with a special point corresponding to a root, we call *rooted partition* the pair $(\mathcal{P}(\phi), \rho(\phi))$ We note that for each A in $\mathcal{P}(\phi)$, the *restricted dynamics* with generator

$$(L_A f)(x) = \sum_{y \in A} w(x,y)[f(y) - f(x)], \quad f : A \to \mathbb{R}, \quad x \in A, \tag{3.56}$$

has only one irreducible component since each x in A is connected with the root. As a consequence the restricted dynamics has a unique equilibrium measure which we call *restricted measure* μ_A. Note that when L has a *reversible* equilibrium $\mu = \mu_{\mathcal{X}}$, then μ_A is nothing but the equilibrium measure μ conditioned on A, i.e. $\mu_A = \mu(\cdot|A)$.

Proposition 3.11. (Roots at restricted equilibrium) *Fix $m \in \{1, \ldots, n\}$, then*

$$\mathbb{P}_q\Big(\rho(\Phi) = \{x_1, \cdots, x_m\}\Big|\mathcal{P}(\Phi) = \{A_1, \ldots, A_m\}\Big) = \prod_{i=1}^m \mu_{A_i}(x_i), \tag{3.57}$$

for any partition $\{A_1, \ldots, A_m\}$ of $\mathcal{X}$ and any $x_i \in A_i$, for $i = 1, \ldots, m$.

Proof. For each i in $\{1, \ldots, m\}$, let us call $\mathcal{T}_i$ the set of spanning trees of A_i. For each τ_i in $\mathcal{T}_i$, define

$$w_i(\tau_i) = \prod_{e \in \tau_i} w(e), \tag{3.58}$$

and for y_i in A_i, write $\rho(\tau_i) := \{y_i\}$, if y_i is the root of the tree $\tau_i \in \mathcal{T}_i$. Compute

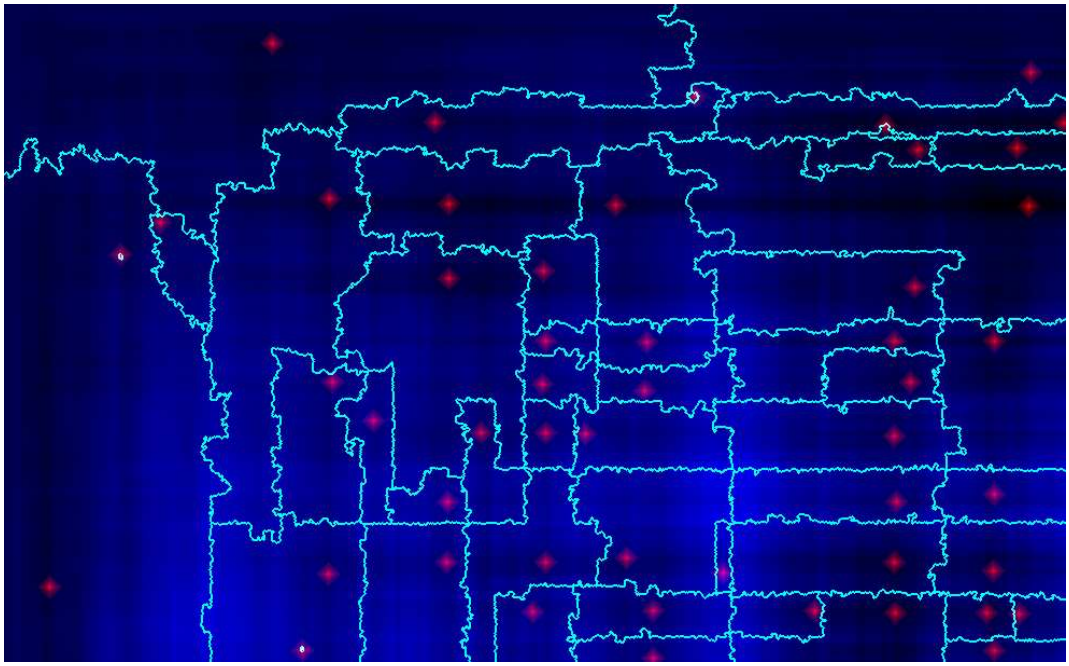
$$\begin{aligned}
&\mathbb{P}_q\Big(\rho(\phi) = \{x_1, \cdots, x_m\}\Big|\mathcal{P}(\phi) = \{A_1, \ldots, A_m\}\Big) \\
&= \frac{\mathbb{P}_q\Big(\rho(\phi) = \{x_1, \cdots, x_m\}, \mathcal{P}(\phi) = \{A_1, \ldots, A_m\}\Big)}{\mathbb{P}_q\Big(\mathcal{P}(\phi) = \{A_1, \ldots, A_m\}\Big)} \\
&= \frac{q^m \sum_{\tau_1 \in \mathcal{T}_1} \cdots \sum_{\tau_m \in \mathcal{T}_m} \prod_{i=1}^m w_i(\tau_i) \mathbb{1}_{\{\rho(\tau_i) = \{x_i\}\}}}{q^m \sum_{\tau_1 \in \mathcal{T}_1} \cdots \sum_{\tau_m \in \mathcal{T}_m} \prod_{i=1}^m w_i(\tau_i)} \\
&= \prod_{i=1}^m \frac{\sum_{\tau_i \in \mathcal{T}_i} w_i(\tau_i) \mathbb{1}_{\{\rho(\tau_i) = \{x_i\}\}}}{\sum_{\tau_i \in \mathcal{T}_i} w_i(\tau_i)} \\
&= \prod_{i=1}^m \mu_{A_i}(x_i),
\end{aligned}$$

where the last equality follows by (3.55) applied to the restricted dynamics. □

Remark: When X is reversible with respect to a measure μ, this gives a way to build the associated Gaussian free field with mass $m = \sqrt{q}$, that is the Gaussian process $\xi = (\xi_x)_{x \in \mathcal{X}}$ with covariance matrix

$$\Gamma = \left(\frac{G_q(x,y)}{\mu(y)}\right)_{x,y \in \mathcal{X}} = \left(\frac{K(x,y)}{q\mu(y)}\right)_{x,y \in \mathcal{X}},$$

Figure 8: A rooted partition with 50 roots (at the center of red diamonds) sampled for the Metropolis random walk in a Brownian sheet potential on the 987×610 grid and with inverse temperature $\beta = .06$. Blue levels depend on the potential: the lower the potential, the darker the blue. We see that each root is distributed according to the restricted equilibrium of its own piece of the partition. See also the first two pictures in Figure 1.



by successive sampling of the standard measure ν_q. Start from independent centered random variables ζ_x with variance $\mu(x)$, $x \in \mathcal{X}$, sample ϕ according to ν_q, call $A(x)$ the set of vertices of $\tau_\phi(x)$ and set

$$\tilde{\xi}_x = \frac{1}{\mu(A(x))} \sum_{y \in A(x)} \frac{\zeta_y}{\sqrt{q}}.$$

Then the random field $\tilde{\xi}$ has zero mean and covariance matrix Γ and the rescaled partial sum $\sum_{i=1}^n \tilde{\xi}^i/\sqrt{n}$, with $\tilde{\xi}^1$, $\tilde{\xi}^2$, ... independent copies of $\tilde{\xi}$, converges in law to ξ as n goes to infinity. Indeed, for each x and y in $\mathcal{X}$, $\tilde{\xi}_x$ and $\tilde{\xi}_y$ are centered and one computes

$$\mathbb{E}\left[\tilde{\xi}_x \tilde{\xi}_y\right] = \sum_{A \subset \mathcal{X}} \mathbb{P}\left(A(x) = A(y) = A\right) \frac{1}{\mu(A)^2} \sum_{z \in A} \frac{\mu(z)}{q} = \sum_{A \ni x,y} \mathbb{P}\left(A(x) = A(y) = A\right) \frac{1}{q\mu(A)}. \tag{3.59}$$

On the other hand, following Wilson algorithm,

$$\begin{aligned} K(x,y) &= \mathbb{P}\left(y \in \rho(\Phi), A(x) = A(y)\right) \\ &= \sum_{A \ni x,y} \mathbb{P}\left(A(x) = A(y) = A\right) \mathbb{P}\left(y \in \rho(\Phi) \middle| A(x) = A(y) = A\right) \\ &= \sum_{A \ni x,y} \mathbb{P}\left(A(x) = A(y) = A\right) \frac{\mu(y)}{\mu(A)}, \end{aligned}$$

by combining it with (3.59), we conclude that

$$\mathbb{E}\left[\tilde{\xi}_x \tilde{\xi}_y\right] = \frac{K(x,y)}{q\mu(y)}. \tag{3.60}$$

3.5.2 Cumulants

Let us associate with our random forests Φ with law ν_q, the random variables

$$\eta_x = \mathbb{1}_{\{x \in \rho(\Phi)\}}, \quad x \in \mathcal{X}, \tag{3.61}$$

note that they completely describe the roots process. For $A = \{x_1, \dots, x_k\} \subset \mathcal{X}$ with distinct x_i's, the cumulants of these random variables are defined by

$$\kappa_A(\eta) = \kappa(\eta_{x_1}, \dots, \eta_{x_k}) = \frac{\partial^k}{\partial\lambda_1 \dots \partial\lambda_k} \ln \mathbb{E}\left[\exp\left\{\sum_{i=1}^{k} \lambda_i \eta_{x_i}\right\}\right]\Bigg|_{\lambda=0}. \tag{3.62}$$

These quantities are the so-called truncated correlation functions, that can also be recursively defined by

$$\mathbb{E}\left[\prod_{x\in A} \eta_x\right] = \sum_{\Pi \in \mathcal{P}_A} \prod_{B\in\Pi} \kappa_B(\eta), \tag{3.63}$$

where $\mathcal{P}_A$ stands for the set of partitions of A.

The determinantal nature of the roots process makes its cumulants easy to compute. With $A \subset \mathcal{X}$ and $\mathcal{S}_A$ being the permutation group on A, one has

$$\mathbb{E}\left[\prod_{x\in A} \eta_x\right] = \mathbb{P}\left(A \subset \rho(\Phi)\right) = \det_A(K) = \sum_{\sigma\in\mathcal{S}_A} (-1)^{\mathrm{sgn}(\sigma)} \prod_{x\in A} K(x, \sigma(x)). \tag{3.64}$$

Making a cycle decomposition of each permutation in this sum and denoting by $\mathcal{C}_B$ the set of long cycles on $B \subset A$, i.e. the set of cycles of length $|B|$ in B, after some simple algebra, we get

$$\mathbb{E}\left[\prod_{x\in A} \eta_x\right] = \sum_{\Pi\in\mathcal{P}_A} \prod_{B\in\Pi} \sum_{\sigma\in\mathcal{C}_B} (-1)^{|B|-1} \prod_{x\in B} K(x, \sigma(x)). \tag{3.65}$$

This identifies our cumulants through (3.63) and gives the following lemma.

Lemma 3.12. *For all $A \subset \mathcal{X}$*

$$\kappa_A(\eta) = (-1)^{|A|-1} \sum_{\sigma\in\mathcal{C}_A} \prod_{x\in A} P_x(X(T_q) = \sigma(x)), \tag{3.66}$$

where $\mathcal{C}_A$ stands for the set of cycles of length $|A|$ in A.

Remark: We cannot help making the following observation. In the case of uniformly equal weights between nearest neighbours, for large q, $(-1)^{|A|-1}\kappa_A(\eta)$ behaves like the natural low temperature partition function associated with an embedded travelling salesman problem. In this regime, on the one hand, Wilson's algorithm quickly provides perfect samples of the roots process and, on the other hand, the cumulant is the expected value of some observable for the system made of n independent copies of $\rho(\Phi)$ [16]. This suggests that one could find a practical way to estimate this low temperature partition function and then solve the travelling salesman problem. Unfortunately, the corresponding observable has an exponentially small probability to be different from 0 and consequently, it is in reality impossible to estimate its mean in this way.

A Appendix

A.1 Schur complement and trace process

Assume we have a Markov process Y on a finite state space $\mathcal{Y}$, with generator $\mathcal{L}$ given by

$$(\mathcal{L}f)(y) = \sum_{z\in\mathcal{Y}} \alpha(y,z)[f(z) - f(y)], \quad y \in \mathcal{Y}, \tag{A.1}$$

with $f : \mathcal{Y} \to \mathbb{R}$ arbitrary and $\{\alpha(y,z) \in [0,+\infty) : (y,z) \in \mathcal{Y}\times\mathcal{Y}\}$ a given collection of non-negative and finite transition rates. As in Section 2.1, let us denote by $\hat{Y}$ the discrete skeleton of the random walk Y with Markovian transition matrix $\hat{P}$.

Fix a subset $A \subset \mathcal{Y}$ with $|A| = k$ and consider a new Markov chain $\hat{Y}^A$ with state space A obtained as the trace of the process $\hat{Y}$ on A. In other words, $\hat{Y}^A$ is the random walk with transition matrix $\hat{P}^A \in \mathcal{M}_{k\times k}$, with entries

$$\hat{p}^A(x,y) = P_x(\hat{Y}(\hat{T}_A^+) = y), \quad \text{for } x,y \in A, \tag{A.2}$$

where $\hat{T}_A^+$ denotes the first return time in A of the chain $\hat{Y}$.

Back to the continuous-time setting, denote by Y^A be the continuous-time version of $\hat{Y}^A$ with jump times given by exponential random variables of parameter

$$\bar{\alpha} := \max_{y\in\mathcal{Y}} \alpha(y) < \infty, \tag{A.3}$$

i.e. the process with generator

$$\mathcal{L}^A := \overline{\alpha}(\hat{P}^A - \mathbb{1}_A). \tag{A.4}$$

Equivalently, Y^A is the trace of the Markov process Y on A, namely, the process obtained by following the trajectory of Y at infinite velocity outside A and without speeding up inside A.

Proposition A.1. (Schur complement and trace process) *Given the Markov process Y on $\mathcal{Y}$ with generator $\mathcal{L}$, fix a subset $A \subset \mathcal{Y}$ with $|A| = k$. Let $\mathcal{L}^A$ be the generator of the Markov process Y^A obtained as the trace of the process Y on A. Then, $\mathcal{L}^A$ is the Schur complement of $[\mathcal{L}]_{A^c}$ in $\mathcal{L}$, i.e.*

$$\mathcal{L}^A = [\mathcal{L}]_A - [\mathcal{L}]_{A,A^c}[\mathcal{L}]_{A^c}^{-1}[\mathcal{L}]_{A^c,A}, \tag{A.5}$$

with $[\mathcal{L}]_A$ and $[\mathcal{L}]_{A^c}$ being the sub-Markovian generator of the process killed outside A and A^c, respectively. In other words, $[\mathcal{L}]_A$ and $[\mathcal{L}]_{A^c}$ represent the operators obtained from the matrix representation of $\mathcal{L}$, by deleting the row and the columns indexed by sites outside A and A^c, respectively. Whereas $[\mathcal{L}]_{A,A^c}$ and $[\mathcal{L}]_{A^c,A}$ are the operators obtained from the matrix representation of $\mathcal{L}$, by keeping only those rates from A to A^c and A^c to A, respectively.

Proof. Denote by $[\hat{P}]_A \in \mathcal{M}_{k\times k}$, the sub-Markovian matrix $\hat{P}$ restricted to A. Note that $[\hat{P}]_A$ is different from $\hat{P}^A$. Due to (A.2), for any $x,y \in A$, we can write

$$\begin{aligned}\hat{p}^A(x,y) &= \hat{p}(x,y) + \sum_{z,z'\in A^c} \hat{p}(x,z)\left(\sum_{k\geq 0}[\hat{P}]_{A^c}^k(z,z')\right)\hat{p}(z',y) \\ &= [\hat{P}]_A(x,y) + \sum_{z,z'\in A^c} \hat{p}(x,z)\left(\mathbb{1}_{A^c} - [\hat{P}]_{A^c}\right)^{-1}(z,z')\hat{p}(z',y).\end{aligned} \tag{A.6}$$

Subtracting $\mathbb{1}_A$ on both side of (A.6), we obtain that

$$\hat{P}^A - \mathbb{1}_A = [\hat{P} - \mathbb{1}]_A - [\hat{P} - \mathbb{1}]_{A,A^c}\left([\hat{P}]_{A^c} - \mathbb{1}_{A^c}\right)^{-1}[\hat{P} - \mathbb{1}]_{A^c,A}. \tag{A.7}$$

We then get our result by multiplying both sides by $\bar{\alpha}$. □

When $\mathcal{Y}$ contains an absorbing set B and $A \subset B^c$ we can do the same computation with the sub-Markovian generator $[\mathcal{L}]_{B^c}$ in place of $\mathcal{L}$. For any x and y in A the mean local time in y starting from x and before hitting B is the same for Y and the trace process Y^A, i.e.

$$G_B(x,y) = G_B^A(x,y), \tag{A.8}$$

that is

$$\left[[\mathcal{L}]_{B^c}^{-1}\right]_A = \left([\mathcal{L}]_{B^c}^A\right)^{-1}. \tag{A.9}$$

More generally, one has the following definition and properties.

Definition A.2. (Schur complement) *Let M be a 2×2 block matrix in $\mathcal{M}_{n\times n}$ of the form*

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \tag{A.10}$$

where $A \in \mathcal{M}_{k\times k}$ and $D \in \mathcal{M}_{(n-k)\times(n-k)}$, for some $k \leq n$. The Schur complement of D in M is defined as the matrix

$$S_M(D) := A - BD^{-1}C. \tag{A.11}$$

One can then check:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_k & BD^{-1} \\ 0 & I_{n-k} \end{bmatrix} \begin{bmatrix} S_M(D) & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I_k & 0 \\ D^{-1}C & I_{n-k} \end{bmatrix}. \tag{A.12}$$

It follows that

$$\det(M) = \det(D)\det\left(S_M(D)\right) \tag{A.13}$$

and, as a generalization of (A.9),

$$M^{-1} = \begin{bmatrix} S_M(D)^{-1} & -S_M(D)BD^{-1} \\ -D^{-1}CS_M(D) & D^{-1} + D^{-1}CS_M(D)BD^{-1} \end{bmatrix}. \tag{A.14}$$

In particular, from (A.14) and (A.13),

$$\det{}_I\left(M^{-1}\right) = \det\left(S_M(D)^{-1}\right) = \det(S_M(D))^{-1} = \frac{\det_{I^c}(M)}{\det(M)} \tag{A.15}$$

with $I = \{1;\cdots;k\}$ and $I^c = \{k+1;\cdots;n\}$. This relation is used in the algebraic proof of Theorem 3.4.

A.2 Lemma on hitting distributions and times

In this section we use our forest measure analysis to prove two formulas on the hitting distribution and the expectation of hitting times of a given subset of the given graph. This result is originally due to Freidlin and Wentzell, see Lemmas 3.2 and 3.3 in [6].

Lemma A.3. (Freidlin and Wentzell [6]) *Fix a non-empty subset R of $\mathcal{X}$. Recall the notation in* (1.12)*. Consider the Markov process X on $\mathcal{X}$ identified by* (1.1)*, and let T_R be the hitting time of the set R. Then, for any $x \in R^c$ and $y \in R$*

$$P_x(X(T_R) = y) = \frac{1}{Z_R(0)} \sum_{\substack{\phi:\rho(\phi)=R \\ \rho(\tau_x)=\{y\}}} w(\phi), \tag{A.16}$$

and, for any $x \in R^c$,

$$E_x[T_R] = \frac{1}{Z_R(0)} \sum_{y\notin R} \sum_{\substack{\phi:\rho(\phi)=R\cup\{y\} \\ \rho(\tau_x)=\{y\}}} w_(\phi). \tag{A.17}$$

Proof. Consider the extended space $\bar{\mathcal{X}}$ and the extended Markov process $\bar{X}$ as in (2.12), with the killing rates as in (1.11) for $q = 0$. Equation (A.16) is simply obtained by considering Wilson's algorithm started from x.

To prove (A.17) we will use the discrete skeleton of Section (2.1) with $\mathcal{Y} = \bar{\mathcal{X}}$ and $B = R$. Let

$$G_R(x, y) = E_x[l_y(T_R)] \quad \text{and} \quad \hat{G}_R(x, y) = E_x[l_y(\hat{T}_R)], \quad x, y \in R^c, \tag{A.18}$$

be the continuous and the discrete-time Green's functions before hitting the set R, respectively.

Since

$$E_x[T_R] = \sum_{y \notin R} G_R(x, y),$$

it suffices to show that

$$G_R(x, y) = \frac{1}{Z_R(0)} \sum_{\substack{\phi: \rho(\phi) = R \cup \{y\} \\ \rho(\tau_x) = \{y\}}} w(\phi). \tag{A.19}$$

Since $\hat{G}_R(x, y) = P_x(\hat{T}_y < \hat{T}_R) G_R(y, y) = P_x(\hat{T}_y < \hat{T}_R)/P_y(\hat{T}_y^+ > \hat{T}_R)$, with $\hat{T}_y^+$ being the return time to y,

$$\begin{aligned} G_R(x, y) = \frac{1}{\bar{\alpha}} \hat{G}_R(x, y) &= \frac{P_x(T_y < T_R)}{\overline{\alpha} P_y(\hat{T}_y^+ > \hat{T}_R)} = \frac{P_x(T_y < T_R)}{\overline{\alpha} \sum_{z \neq y} \hat{p}(y, z) P_z(\hat{T}_y > \hat{T}_R)} \\ &= \frac{P_x(T_y < T_R)}{\overline{\alpha} \sum_{z \neq y} \hat{p}(y, z) P_z(T_y > T_R)} = \frac{P_x(T_y < T_R)}{\sum_{z \neq y} \alpha(y, z) \left[1 - P_z(T_y < T_R)\right]}. \end{aligned} \tag{A.20}$$

Observe that $P_x(T_y < T_R) = P_x(X(T_{R \cup \{y\}}) = y)$ for any $x, y \in R^c$, thus using (A.16):

$$\begin{aligned} G_R(x, y) &= \frac{P_x(X(T_{R \cup \{y\}}) = y)}{\sum_{z \neq y} \alpha(y, z) \left[1 - P_z(X(T_{R \cup \{y\}}) = y)\right]} \\ &= \frac{\frac{1}{Z_{R \cup \{y\}}(0)} \sum_{\substack{\phi: \rho(\phi) = R \cup \{y\} \\ \rho(\tau_x) = \{y\}}} w(\phi)}{\frac{1}{Z_{R \cup \{y\}}(0)} \sum_{z \neq y} \alpha(y, z) \sum_{\substack{\phi: \rho(\phi) = R \cup \{y\} \\ \rho(\tau_x) \neq \{y\}}} w(\phi)} \\ &= \frac{1}{Z_R(0)} \sum_{\substack{\phi: \rho(\phi) = R \cup \{y\} \\ \rho(\tau_x) = \{y\}}} w(\phi). \end{aligned}$$

□

A.3 Divided differences

In this appendix, we recall three equivalent definitions of the notion of divided differences of a real function. We further give a lemma due to Micchelli and Willoughby for which we provide an alternative elementary proof which plays with these different definitions. This result is used in Section 3.3.

Definition A.4. (Divided differences 1) *We call divided difference of a function f at the distinct points $x_0, x_1, \ldots, x_k$, the quantity $f[x_0, \cdots, x_k]$ recursively defined via*

$$f[x_0, \cdots, x_k] = \frac{f[x_1, \cdots, x_k] - f[x_0, \cdots, x_{k-1}]}{x_k - x_0}, \tag{A.21}$$

with

$$f[x_i] = f(x_i).$$

From this definition, we see that the divided differences at k points of a function f can be seen as a sort of k-th discrete derivative. One then show by induction

Definition A.5. (Divided differences 2) *For any function f and distinct points $x_0, x_1, \ldots, x_{k-1}, x_k$,*

$$f[x_0, \cdots, x_k] = \sum_{i=0}^{k} \frac{f(x_i)}{\prod_{j \neq i}(x_i - x_j)}. \tag{A.22}$$

Note in this second definition, that $f[x_0, \cdots, x_k]$ is independent of the order of the x_i's. From (A.22) one can then check

Definition A.6. (Divided differences 3) *For any function f and distinct points $x_0, x_1, \ldots, x_{k-1}, x_k$,*

$$\begin{aligned} Q(x) =& f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &+ \ldots + f[x_0, \cdots, x_k](x - x_0) \ldots (x - x_{k-1}) \end{aligned} \tag{A.23}$$

is the unique polynomial of degree k with $Q(x_i) = f(x_i)$, for $i = 0, \cdots, k$.

Lemma A.7. (Micchelli and Willoughby [13]) *Consider a polynomial of degree n of the form*

$$f(x) = \prod_{i=0}^{n-1}(x - \alpha_i),$$

with n distinct real zeros α_i in decreasing order: $\alpha_0 > \alpha_1 > \cdots > \alpha_{n-1}$. Let $\beta_0 > \beta_1 > \cdots > \beta_N$ with $N \geq n$ and such that

$$\beta_i \geq \alpha_i, \quad \textit{for all } i < n. \tag{A.24}$$

Then, for any $k \leq N$,

$$f[\beta_0, \beta_1, \ldots, \beta_k] \geq 0.$$

Proof. We prove the following statement by induction on $r = n - k$:

$$\text{“For any } \beta_0 > \beta_1 > \cdots > \beta_N \text{ satisfying (A.24), } f[\beta_0, \ldots, \beta_k] \geq 0.\text{”} \tag{A.25}$$

Since f is a polynomial of degree n, (A.25) follows from Definition A.23 as soon as $r < 0$. Also, since the dominant coefficient of f is 1, the same argument shows $f[\beta_0, \ldots, \beta_n] = 1$ and the claim holds for $r = 0$. Fix now $r > 0$, i.e. $k < n$, and $\beta_0 > \cdots > \beta_N$ satisfying (A.24). If $\beta_0 \neq \alpha_0$ then

$$\begin{aligned} \frac{f[\beta_0, \alpha_1, \ldots, \alpha_k] - f[\alpha_0, \alpha_1, \ldots, \alpha_k]}{\beta_0 - \alpha_0} &= \frac{f[\beta_0, \alpha_1, \ldots, \alpha_k] - f[\alpha_1, \ldots, \alpha_k, \alpha_0]}{\beta_0 - \alpha_0} && \text{(A.26)} \\ &= f[\beta_0, \alpha_1, \ldots, \alpha_k, \alpha_0] && \text{(A.27)} \\ &= f[\beta_0, \alpha_0, \alpha_1, \ldots, \alpha_k]. && \text{(A.28)} \end{aligned}$$

By Definition A.22 we have $f[\alpha_0, \ldots, \alpha_k] = 0$ and the numerator in the l.h.s. of (A.26) is merely equal to $f[\beta_0, \alpha_1, \ldots, \alpha_k]$. The denominator is positive from (A.24) and so is the r.h.s. of (A.28) by induction hypothesis. It follows that

$$f[\beta_0, \alpha_1, \ldots, \alpha_k] \geq 0 \tag{A.29}$$

and the same is true when $\beta_0 = \alpha_0$. If $\beta_1 \neq \alpha_1$, we compute

$$\frac{f[\beta_0, \beta_1, \alpha_2, \ldots, \alpha_k] - f[\beta_0, \alpha_1, \ldots, \alpha_k]}{\beta_1 - \alpha_1} = f[\beta_0, \beta_1, \alpha_1, \ldots, \alpha_k], \tag{A.30}$$

get then in the same way

$$f[\beta_0, \beta_1, \alpha_2, \ldots, \alpha_k] \geq 0 \tag{A.31}$$

and the same is true when $\beta_1 = \alpha_1$. Proceeding similarly we eventually obtain that

$$f[\beta_0, \ldots, \beta_k] \geq 0. \tag{A.32}$$

□

Acknowledgements: A.G. thanks Laurent Tournier for the many fruitful discussions on the matter, and, in particular, the nice independence argument we presented in Section 2. He also thanks Pablo Ferrari, Inès Armendariz, Pablo Groisman, Leonardo Rolla and Matthieu Jonckheere for their hospitality in Buenos Aires where all this work started.

References

[1] V. Anantharam and P. Tsoucas, A proof of the Markov chain tree theorem, *Statist. Probab. Lett.*, 8(2), 189–192 (1989).

[2] R. Burton and R. Pemantle, Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances, *Ann. Probab.* 21, 1329–1371 (1993).

[3] P. Diaconis and L. Miclo, On times to quasi-stationarity for birth and death processes, *J. Theoret. Probab.* 22(3), 558–586 (2009).

[4] J. B. Hough, M. Krishnapur, Y. Peres and B. Virag, Determinantal processes and independence, *Probab. Surv.* 3, 206-âĂŞ229 (2006).

[5] J. Fill, On hitting times and fastest strong stationary times for skip-free and more general chains, *J. Theor. Probab.* 22, 587–600 (2009).

[6] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, Grundlehren der Mathematischen Wissenschaften 260, (Second edition ed.). New York: Springer-Verlag (1998).

[7] G. Grimmett, *The random-cluster model*, Grundlehren der mathematischen Wissenschaften 333, Springer (2006).

[8] A. A. Járai, F. Redig and E. Saada, Zero dissipation limit in the Abelian sandpile model, *Preprint: arXiv:0906.3128v3* (2011).

[9] G. Kirchhoff, Über di Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvnischer Ströme geführt wird, *Ann. Phys. chem* 72, 497–508 (1847).

[10] Y. Le Jan, *Markov paths, loops and fields*, vol. 2026 of Lecture Notes in Mathematics. Springer, Heidelberg, 2011. Lectures from the 38th Probability Summer School held in Saint-Flour, 2008, Ecole d'Été de Probabilités de Saint-Flour.

[11] D. A. Levin, Y. Peres and E. L. Wilmer, *Markov chains and mixing times*, American Mathematical Society (2008).

[12] P. Marchal, Loop-erased random walks, spanning trees and Hamiltonian cycles, *Elect. Comm. Probab.* 5 , 39–50 (2000).

[13] C. A. Micchelli and R. A. Willoughby, On functions which preserve the class of Stieltjes matrices, *Lin. Alg. Appl.* 23, 141–156 (1979).

[14] L. Miclo, On absorption times and Dirichlet eigenvalues, *ESAIM: Probability and Statistics* 14, 117–150 (2010).

[15] J. Propp and D. Wilson, How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph, *Journal of Algorithms* 27, 170–217 (1998).

[16] G. S. Sylvester, Representations and inequalities for Ising model Ursell functions, *Comm. Math. Phys.* 42, 209–220 (1975).

[17] L. Tournier, private communication.

[18] V. T. Tsuchiya and M. Katori, Proof of breaking of self-organized criticality in a nonconservative abelian sandpile model, *Phys. Rev. E* 61, 1183–1188 (2000).

[19] D. Wilson, Generating random spanning trees more quickly than the cover time, *Proceedings of the twenty-eighth annual acm symposium on the theory of computing*, 296–303 (1996).