universität wien

DISSERTATION

Titel der Dissertation

"On Some Aspects of Random Walks in Random Environments: Trapping, Scaling Limits, Aging, and Percolation"

Verfasser

Tobias Wassmer, MSc

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, im April 2015

Studienkennzahl lt. Studienblatt:	A 796 605 405
Dissertationsgebiet lt. Studienblatt:	Mathematik
Betreuer:	Univ. Prof. Dr. Jiří Černý

Abstract

This thesis consists of three research papers which deal with various aspects of random walks in random environments. In the first part we consider two types of Markov chains that exhibit trapping, and we study associated scaling limits and the phenomenon of aging. In the second part we analyze percolative properties of the set of vertices that are not visited by the simple random walk on a finite graph, the so-called vacant set.

The first paper studies the Metropolis dynamics of the simplest mean-field spin glass model, the Random Energy Model. We show that this dynamics exhibits aging by showing that the properly rescaled time-change process between the Metropolis dynamics and a suitably chosen 'fast' Markov chain converges in distribution to a stable Lévy process. This provides a first proof for aging of a fully 'asymmetric' dynamics on the non-modified Random Energy Model.

The result of the second paper is a complete classification of the possible scaling limits of randomly trapped random walks on $\mathbb{Z}^d$, $d \geq 2$. We show that the possible classes of scaling limits reduce from four in one dimension to only two in higher dimensions. In particular, in the case when the discrete skeleton of the randomly trapped random walk is a simple random walk on $\mathbb{Z}^d$, the scaling limit is either Brownian motion or the Fractional Kinetics process.

In the third paper we show that the vacant set left by the simple random walk on the giant component of a supercritical Erdős-Rényi random graph exhibits a phase transition similar to the classical phase transition of Bernoulli percolation on the complete graph. Moreover, we show that the critical point of this phase transition is closely related to the critical value of random interlacements on the corresponding infinite volume limit, which is a Poisson-Galton-Watson tree.

Zusammenfassung

Diese Dissertation besteht aus drei Forschungsartikeln, welche verschiedene Aspekte von Irrfahrten in zufälligen Umgebungen behandeln. Im ersten Teil betrachten wir zwei Arten von Markovketten die Fallen aufweisen, und wir studieren dazugehörige Grenzprozesse und das Phänomen des Alterns. Im zweiten Teil analysieren wir die perkolativen Eigenschaften der Menge unbesuchter Knoten einer Irrfahrt auf einem endlichen Graphen.

Der erste Artikel ist dem Studium der Metropolis-Dynamik auf dem einfachsten Mean-Field Spin-Glas, dem Random Energy Model, gewidmet. Wir zeigen dass diese Dynamik altert, indem wir zeigen dass der Zeitwechsel-Prozess zwischen der Metropolis-Dynamik und einer passend gewählten 'schnellen' Markovkette in Verteilung zu einem stabilen Lévy-Prozess konvergiert. Dies liefert einen ersten Beweis für Altern einer gänzlich asymmetrischen Dynamik auf dem nicht-modifizierten Random Energy Model.

Das Resultat des zweiten Artikels ist eine Klassifikation aller möglichen Grenzprozesse von Randomly Trapped Random Walks auf $\mathbb{Z}^d$, $d \geq 2$. Wir zeigen, dass die möglichen Klassen von Grenzprozessen sich von vier in einer Dimension auf nur mehr zwei in höheren Dimensionen reduzieren. Insbesondere, wenn das diskrete Skelett des Randomly Trapped Random Walk die einfache Irrfahrt ist, dann ist der Grenzprozess entweder Braunsche Bewegung oder der Fractional Kinetics-Prozess.

Im dritten Artikel zeigen wir, dass die Menge unbesuchter Knoten einer Irrfahrt auf der riesigen Komponente eines Erdős-Rényi Zufallsgraphen einen Phasenübergang durchläuft, ähnlich wie der klassische Phasenübergang in Bernoulli-Perkolation auf dem vollständigen Graphen. Zusätzlich zeigen wir, dass der kritische Punkt dieses Phasenübergangs in engem Bezug steht zum kritischen Wert von Random Interlacements auf dem zugehörigen Grenzobjekt unendlichen Volumens, einem Poisson-Galton-Watson-Baum.

Acknowledgments

My first and most thanks go to my advisor Jiří Černý who gave me the opportunity to pursue my doctoral studies and has supported me during the last two and a half years. A brilliant mathematician and both challenging and inspiring, he has been the best teacher and mentor that I could wish for, and I am very grateful for his patience and all help and advice he gave me.

Next, I would like to thank Gérard Ben Arous and David Croydon for agreeing to referee this work. I would also like to thank Roland Zweimüller for helpful discussions at the start of my second research paper.

Working in the Mathematical Finance and Probability Theory research group at the University of Vienna was a great pleasure. My thanks go to all the people who contributed to the motivating and inspiring environment, in particular to my recent office mates Lingqi, Fabian and Junjian for all the discussions on mathematics' small and life's big questions, and to Astrid for making my life easy on any administrative matter.

A big part of being successful at mathematics is to have good distractions from it, which I was lucky to find in playing and refereeing ice hockey. Many thanks to the Vienna University Tôtônka hockey team for three incredible seasons, in particular to Marc for backing me up not only on but also off the ice. I also thank the referee team for the good times including a few unforgettable moments in the EBEL.

I am grateful to Elisa, Izabela, Fabian, Patrick, and all the part-time members of our small lunch group for sharing the experience of life as a PhD student in Vienna in many culinary excursions.

Special thanks go to my parents who have always believed in me, supported me, and made everything possible. I also thank the rest of my family for encouraging and supporting me.

Finally, I would like to thank Karin, to whom I dedicate this work. Thank you for being who you are and being there for me at any times.

Contents

Bibliographical note

This thesis consists of the following three original research papers:

[ČW15a] J. Černý and T. Wassmer, *Aging of the Metropolis dynamics on the Random Energy Model*, submitted preprint, arXiv:1502.04535, 2015.

[ČW15b] J. Černý and T. Wassmer, *Randomly trapped random walks on* $\mathbb{Z}^d$, Stochastic Process. Appl. **125** (2015), no. 3, 1032–1057. MR 3303967

[Was15] T. Wassmer, *Phase transition for the vacant set left by random walk on the giant component of a random graph*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 756–780. MR 3335024

The papers [ČW15a] and [ČW15b] are presented in the first part of the thesis, in Chapters 2 and 3 respectively. The paper [Was15] is presented in the second part in Chapter 5. The order of presentation of the papers here is in the order of relevance as perceived by the author, which is in opposite chronological order.

Part I

TRAP MODELS AND AGING

1. Introduction to trap models and aging

The first part of this thesis is devoted to the study of certain properties of Markov chains in random environments that exhibit *trapping*. While being interesting objects of their own, these models are mainly motivated by *aging of spin-glasses*, a problem that has been intensively studied in both the physics and the mathematics literature.

In the real world, spin-glasses are materials that consist of a non-magnetic matrix in which somewhat randomly a few magnetic atoms are placed. The magnetic interaction between those impurities depends on the distance between them, their magnetic spins have tendency to be aligned at some distances and anti-aligned at others. When cooled down below their glass transition temperature, spin-glasses relax extremely slowly towards equilibrium, such that on experimental time scales they are actually never in equilibrium. In this out-of-equilibrium regime they exhibit interesting dynamical effects such as *aging*, *rejuvenation*, and *memory*.

The following experiment of [VHO+97] on *thermo-remanent magnetization* illustrates the aging phenomenon. A spin-glass is cooled down below its glass transition temperature T_g and then kept in some external magnetic field for a 'waiting time' t_w. After this time the field is switched off and the remanent magnetization of the spin-glass at an 'observation time' $t_w + t$ is measured. Figure 1 shows the results of such an experiment on the $AgMn_{2.6}$ spin-glass for different waiting times t_w.

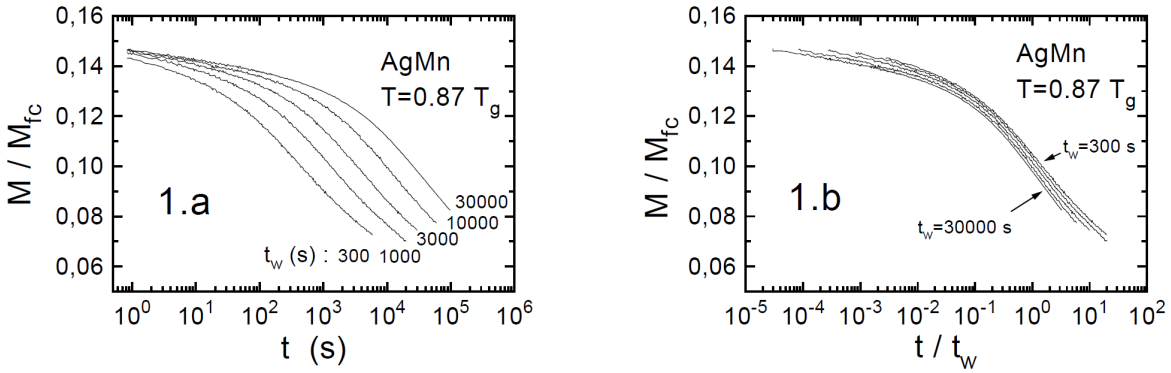

Figure 1: [VHO+97] The thermo-remanent magnetization of the $AgMn_{2.6}$ spin-glass at temperature $T = 9K < 10.4K = T_g$, as a function of t (Figure 1.a) and as a function of t/t_w (Figure 1.b), for waiting times $t_w = 300, 1000, 3000, 10000, 30000s$.

As is evident from Figure 1.a, the thermo-remanent magnetization depends not only on the observation time t but also on the waiting time or 'age' t_w of the system. More

interestingly, Figure 1.b shows that the thermo-remanent magnetization can be approximated by a function of the ratio t/t_w.

Since the first observations of this phenomenon in physical experiments in the 80ies (e.g. [LSNB83, Cha84]), much effort has been put in explaining it, see e.g. [BCKM97] for a review of the physics literature. Among the various approaches there is one that has caught particular attention in the mathematics literature, namely the *Bouchaud trap model*, which describes aging of spin-glasses on a phenomenological level, see e.g. [Ben02, BČ06] for reviews. In Chapter 2 we give a proof for aging of a particular spin-glass model which confirms the predictions of the phenomenological Bouchaud trap model.

Rather independently of the aging problem, trap models inspired by Bouchaud's model evolved as an interesting object of their own, in particular on the integer lattices. In Chapter 3 we consider a certain general trap model called *randomly trapped random walk* and give a classification of all possible scaling limits of the model on the integer lattices of dimensions $d \geq 2$.

In the remainder of this introductory chapter we will outline the history of the analysis of the aging problem and how this lead to the questions that are answered in Chapters 2 and 3. We first describe the challenges in the modeling of spin-glass dynamics in Section 1.1. We then present in Section 1.2 the Bouchaud trap model and in Section 1.3 the approach of proving aging via the *convergence of clock processes*. In Section 1.4 we make an excursion to trap models on integer lattices, where we motivate the model of randomly trapped random walk and present the results of Chapter 3. We describe the results of Chapter 2 in more detail in Section 1.5. Finally, we discuss in Section 1.6 the different methods that were used in the literature to prove convergence of clock processes, and how our methods in Chapters 2 and 3 compare to them.

1.1. Modeling spin-glass dynamics

The task of modeling the dynamics of a spin-glass system can be decomposed to two main levels. The first level is to model the statics of the system by choosing a configuration space and defining a Hamiltonian giving the energy of the configurations. The second level then consists of modeling the dynamics as a process on this state space whose long-time equilibrium should be the equilibrium distribution of the static system, i.e. the *Gibbs measure*.

Assume that there are $N \in \mathbb{N}$ magnetic atoms in the spin-glass, each of them with a spin taking a value of either $+1$ or -1. The natural space of possible spin configurations is thus the N-dimensional hypercube $\mathbb{H}_N = \{-1, 1\}^N$.

As for the Hamiltonian, we will in this thesis exclusively be dealing with *mean-field* spin-glass models. In mean-field models it is assumed that all the magnetic atoms of the spin-glass interact. The randomness of the positions of atoms in the real spin-glass is modeled in terms of random interactions. A prominent example of a mean-field spin-glass is the classical *Sherrington-Kirkpatrick* (SK) model, proposed in [SK75]. In this model the Hamiltonian $H_N(x)$, which gives the energy of every single configuration $x \in \{-1, 1\}^N$, is given by

$$H_N(x) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} J_{ij} x_i x_j.$$

The coupling constants $(J_{ij})_{i,j=1,\dots,N}$ are random, taken as i.i.d. standard Gaussian random variables. The normalization by $\sqrt{N}$ is chosen such that the typical configurations have energy of order N. From the Hamiltonian we obtain the (non-normalized) equilibrium distribution or Gibbs measure of the system at inverse temperature $\beta > 0$,

$$\tau_x = e^{\beta H_N(x)}.$$

A computation shows that the energies $H_N(x)$ of the SK model are themselves centered Gaussian random variables, and their covariance is given by

$$\mathbb{E}[H_N(x)H_N(y)] = NR_N(x,y)^2, \tag{1.1.1}$$

where $R_N(x,y)$ is the normalized overlap on the hypercube, $R_N(x,y) = N^{-1}\sum_{i=1}^N x_i y_i$. It is thus convenient to view the energies of the configurations in the SK model as a Gaussian process indexed on the N-dimensional hypercube with covariance given by (1.1.1).

Though being introduced as a 'solvable model of a spin-glass', it took about 30 years until the statics of the SK model and in particular its Gibbs measure were satisfactorily understood, see e.g. the monographs [Tal11a, Tal11b, Pan13]. In the meantime other 'simpler' models were introduced by somewhat relaxing the dependency structure. The so-called *p-spin SK model* is obtained by changing the power 2 in (1.1.1) to $p \geq 3$. Further, by taking the formal limit $p \to \infty$, the energies $H_N(x)$ become i.i.d. centered Gaussian random variables with variance N. This is the so-called *Random Energy Model* (REM), introduced in [Der80, Der81].

The REM will be the model considered in Chapter 2, we therefore briefly describe its main static features. It is well known (e.g. [Eis83, OP84]) that there is a phase transition for the Gibbs measure of the REM at the critical inverse temperature $\beta_c = \sqrt{2\log 2}$, reflecting the glass transition of real spin-glasses. Roughly said, in the high-temperature regime $\beta < \beta_c$ the Gibbs measure is 'flat', i.e. there are no single configurations contributing considerably to the total mass. In the low-temperature regime $\beta > \beta_c$, on the other hand, a few configurations carry almost all mass of the Gibbs measure, the τ_x's behave like heavy-tailed random variables.

Heuristically, this behavior can be seen by the following computation. For the inverse temperature $\beta > 0$ and an additional parameter $\alpha \in (0,1)$, define the scales

$$r_N = e^{\frac{\alpha^2\beta^2}{2}N} \quad \text{and} \quad g_N = e^{\alpha\beta^2 N}(\alpha\beta\sqrt{2\pi N})^{-\frac{1}{\alpha}}. \tag{1.1.2}$$

Using a standard Gaussian tail approximation it follows that

$$\mathbb{P}[\tau_x > u g_N] = \frac{u^{-\alpha}}{r_N}(1 + o(1)) \quad \text{as } N \to \infty. \tag{1.1.3}$$

The relation (1.1.3) shows that the finitely many states of highest energy that one can find upon observing r_N states have their rescaled Gibbs mass τ_x/g_N behaving like random variables in the domain of attraction of an α-stable law. In particular, choosing $\alpha = \beta_c/\beta$, i.e. $r_N = 2^N$, implies a heavy-tailed behavior with $\alpha < 1$ on the whole hypercube if and only if $\beta > \beta_c$. This explains at least the low temperature phase of the phase transition in the Gibbs measure. The relation (1.1.3) also plays an important role for aging, as we will see later.

On the level of the modeling of dynamics of spin-glasses there is a wide variety of possibilities. We will focus in this thesis on certain usual simplifying assumptions that are made in the mathematics literature.

The first assumption is to consider only Markovian dynamics, which means that the process on the hypercube $\mathbb{H}_N$ can be defined in terms of its transition rates r_{xy} for $x, y \in \mathbb{H}_N$.

The second simplifying assumption is that the spin-glass system only changes one spin at a time. Therefore, the transition rates r_{xy} are non-zero only if x and y are neighbors in the Hamming distance, i.e. they differ in exactly one coordinate, denoted $x \sim y$.

Finally, the equilibrium measure of the Markov process on $\mathbb{H}_N$ should be the equilibrium distribution of the system, the Gibbs measure. To this end it is convenient to choose a Markov process which is reversible with respect to the Gibbs measure, i.e. it satisfies the detailed balance condition

$$\tau_x r_{xy} = \tau_y r_{yx}.$$

The class of models satisfying the above simplifying assumptions is still rather large. We will consider two specific models from the mathematics literature that are 'realistic' in the sense that the dynamics is attracted to 'stable' states, i.e. states of large Gibbs measure. The first one, considered in Chapter 2, is the *Metropolis dynamics*. This is the Markov chain on $\mathbb{H}_N$ with transition rates

$$r_{xy} = \left(1 \wedge \frac{\tau_y}{\tau_x}\right) \mathbf{1}_{\{x \sim y\}}. \tag{1.1.4}$$

This chain chooses its next state with equal probability among more stable neighbors and changes to less stable neighbors with accordingly smaller probability. Note also, as the mean waiting time in some state x is the inverse of the total rate out of x, configurations with very large Gibbs mass (and no equally large neighbor) act as 'traps' for the chain, since the waiting time in these states is much larger than in typical states.

The second model considered in the literature is *Bouchaud's dynamics*, which is the Markov chain with transition rates

$$r_{xy} = \tau_x^{a-1} \tau_y^a \mathbf{1}_{\{x \sim y\}}, \tag{1.1.5}$$

for some asymmetry parameter $a \in [0, 1]$. For $a > 0$ the model is referred to as the *Asymmetric Bouchaud Dynamics* (ABD).

For $a = 0$, where there is no attraction to stable states, Bouchaud's dynamics is also called the *Random Hopping Times* (RHT) dynamics. In that case the transition probabilities to all neighbors of the current state are the same, so that the chain actually performs a simple random walk on the graph with random waiting times that depend only on the Gibbs mass of the current state. This independence of the transition probabilities from the Gibbs measure makes the RHT dynamics particularly tractable, at the expense of not being very 'realistic'. In the same way as for the Metropolis dynamics, states with very large Gibbs mass act as traps for the RHT dynamics.

In what follows we will often refer to the RHT as *symmetric* and to the ABD and Metropolis dynamics as *asymmetric* dynamics.

Having the picture of the 'non-flat' Gibbs measure of the REM at low temperature and the trapping mechanism exhibited by the Metropolis or RHT dynamics in mind, the

following heuristic explanation of the aging behavior arises: On a time scale on which the Markov chain does not reach equilibrium, it does not find the 'deepest' of the traps but evolves among unstable and metastable configurations. The longer the chain is observed, the more stable are the configurations or the deeper are the traps it will find, thus slowing the process down. The stability of the configuration of the system at some age becomes age-dependent, and the older the system, the longer it takes to change its state again, i.e. the longer it takes to 'forget its past', the system 'ages'.

One approach to make this rough description more precise, inspired by physical experiments like the one described at the beginning of this introduction, is to consider the evolution of the system between two large times, the 'waiting time' t_w and the 'observation time' $t_w + t$, and then to analyze appropriate correlation functions $\Pi(t_w, t_w + t)$ of these two time points that measure how much the system forgets its past during that time. Establishing an *aging regime* for the two-point function Π then consists of proving that the limit

$$\lim_{t_w \to \infty} \Pi\big(t_w, t_w(1+\theta)\big) =: \Pi(\theta) \tag{1.1.6}$$

exists and is non-trivial, for either a fixed Gibbs measure $\tau = (\tau_x)_{x \in \mathbb{H}_N}$, or averaged over the randomness of the Hamiltonian.

1.2. Bouchaud's trap model

In the early 90ies physicists introduced several approaches to explain aging of spin-glasses. A prominent example which had a considerable impact in the mathematics literature is the *Bouchaud trap model* [Bou92, BD95]. The main idea in this model is to reduce the state space from the hypercube to the macroscopic network of the metastable states, the 'traps', and investigate on this state space the simplest models for the dynamics and the Hamiltonian, namely the RHT dynamics and an i.i.d. environment for the energies of configurations.

Formally, the Bouchaud trap model is defined on the space $\{1, \dots, n\}$, representing the n 'most important' metastable configurations in the spin-glass. The energies of the configurations are chosen to be i.i.d. exponentially distributed random variables $(E_i)_{i=1,\dots,n}$ with mean one, and the (non-normalized) Gibbs measure of the reduced system at inverse temperature $\beta > 0$ is given by $\tau_i = e^{\beta E_i}$.

The statics of this system are rather simple, since it is immediate that, for $u \geq 1$,

$$\mathbb{P}[\tau_i > u] = u^{-1/\beta}. \tag{1.2.1}$$

This means that in the low-temperature regime $\beta > 1$, the τ_i are independent heavy-tailed random variables in the domain of attraction of a $1/\beta$-stable law, i.e. this Gibbs measure displays a similar behavior as the Gibbs measure of the REM.

The dynamics of the Bouchaud trap model is assumed to evolve on the complete graph on n vertices, $\mathbb{K}_n$. This is reasonable since one is not interested in the behavior of the dynamics outside the 'n most important states', and any transition among them should, at least theoretically, be possible.

The RHT dynamics on $\mathbb{K}_n$ is chosen as the Markov chain $X = (X_t)_{t \geq 0}$ with rates

$$r_{ij} = \frac{1}{n-1}\, \tau_i^{-1}, \quad i \neq j \in \mathbb{K}_n,$$

which means that X performs a simple random walk on $\mathbb{K}_n$ but waits at each site i an exponentially distributed time with mean τ_i.

To describe aging in terms of a two-point correlation function for this simplified system, the question is what two-point function to choose. As observable physical quantities such as the thermo-remanent magnetization are not explicitly describable in the simple model, Bouchaud proposed to analyze abstract two-point functions, for example the probability that the RHT dynamics X on a fixed environment $\tau = (\tau_i)_{i\in\mathbb{K}_n}$ does not change its state during the time interval $[t_w, t_w + t]$,

$$\Pi_n^\tau(t_w, t_w + t) = P_n^\tau[X_{t_w} = X_{t_w+s},\ \forall s < t]. \tag{1.2.2}$$

For this correlation function aging was shown in [BD95] (see also the introduction to [BBG03b]). Namely, using renewal arguments, it was shown that for some time scale $t_w(n)$ growing with the size n of the system, for almost every environment τ,

$$\lim_{n\to\infty} \Pi_n^\tau\big(t_w(n), t_w(n)(1+\theta)\big) = \Pi(\theta). \tag{1.2.3}$$

The function $\Pi(\theta)$ can be given explicitly, cf. (1.3.2). This aging regime for the correlation function (1.2.2) means that the systems is successful in escaping a state in a time scale that is of the same order as the age of the system, i.e. the older the system is, the longer it takes to forget its past.

The aging result (1.2.3) for the Bouchaud trap model can be seen as the first step towards the goal of understanding aging in mean-field spin-glasses. The next aim is to justify the phenomenological simplifications of Bouchaud's trap model and derive similar aging results for models where these simplifications are removed, if necessary one at a time:

- The complete graph $\mathbb{K}_n$ should be replaced by the 'natural' configuration space, the hypercube $\mathbb{H}_N$.
- The i.i.d. energies should be replaced by the correlated energies of e.g. the p-spin SK model.
- The symmetric RHT dynamics should be replaced by one of the 'more realistic' asymmetric dynamics, ABD or Metropolis.

At the moment of writing this thesis, removing all of the three simplifications at once is still an open problem. In Chapter 2 we successfully remove the first and third simplification. We will outline in Section 1.5 what other partial replacements have been successful so far.

1.3. Aging in terms of scaling limits of clock processes

In this section we present the concept that has become the usual approach for proving aging, namely the convergence of clock processes to stable Lévy processes (or *stable subordinators*). This concept was first established in [BČM06, BČ08]. We will motivate it on the simple example of the Bouchaud trap model and explain how it provides aging in terms of convergence of two-point functions as in (1.2.3).

The idea is to express the continuous-time Markov chain, e.g. the RHT dynamics, as a time change of a different process that is easier to analyze. The time-change process is

called *clock process*. As mentioned before, the RHT dynamics is a time change of simple random walk. The clock process of the RHT dynamics is the function $S : \mathbb{N} \to [0, \infty)$, such that $S(k)$ gives the time of the k-th jump of the dynamics.

To be more precise, let $(e_j)_{j\geq 0}$ be independent exponentially distributed random variables with mean one, $\tau = (\tau_i)_{i\in\mathbb{K}_n}$ the non-normalized Gibbs measure of Bouchaud's trap model, and $Y = (Y(j))_{j\geq 0}$ the simple discrete-time random walk on the complete graph $\mathbb{K}_n$. Then the clock process of the RHT dynamics is given by

$$S(k) = \sum_{j=0}^{k-1} \tau_{Y(j)} e_j.$$

On the complete graph $\mathbb{K}_n$ the simple random walk is essentially 'ballistic' on time scales in which it does not discover the whole graph: Since every vertex has $n-1$ neighbors, for large n the probability that once the simple random walk has visited a vertex it will return again to this vertex in 'short time' is very small. Therefore, on such time scales the consecutively found Gibbs masses $\tau_{Y(j)}$ are essentially independent. Moreover, by (1.2.1) they lie in the domain of attraction of a stable law. Ignoring the additional randomness induced by the e_j, it is then not surprising that the properly rescaled process $S(\lfloor tk \rfloor)$ converges to a stable Lévy process.

Indeed, using elementary methods following the general program of [BČ08], it can be shown that for $0 < \kappa < 1$, for almost every environment τ,

$$n^{-\kappa\beta} S(\lfloor tn^\kappa \rfloor) \xrightarrow{n\to\infty} V_{1/\beta}(t), \tag{1.3.1}$$

where $V_{1/\beta}$ is a $1/\beta$-stable Lévy process. The convergence holds in distribution on the space of cadlag functions on $[0, \infty)$ with respect to the standard Skorohod topology. Heuristically, in n^κ steps the simple random walk visits about n^κ states and thus finds deepest traps of depth of order $\tau_x = n^{\kappa\beta}$. The time spent in these deep traps is of order $n^{\kappa\beta}$, so that in the scaling limit they provide the jumps of unit order of the stable process.

The convergence of the clock process can be used to show aging in terms of convergence of two-point functions as in (1.2.3) using the *generalized arcsine law* for stable Lévy processes: As the range of the clock process $S(k)$ are the time points where the RHT dynamics jumps, the probability $\Pi_n^\tau\big(n^{\kappa\beta}, n^{\kappa\beta}(1+\theta)\big)$ that the RHT dynamics does not jump in the time interval $\big[n^{\kappa\beta}, n^{\kappa\beta}(1+\theta)\big]$ can be approximated by the probability that the range of the process $V_{1/\beta}(t)$ does not intersect the time interval $[1, 1+\theta]$. This probability is given by the arcsine law (see e.g. [Ber96, Chapter III.] and the appendix to [BČ06]), and it follows that for almost every environment τ,

$$\lim_{n\to\infty} \Pi_n^\tau\big(n^{\kappa\beta}, n^{\kappa\beta}(1+\theta)\big) = \frac{\sin \alpha\pi}{\pi} \int_0^{1/(1+\theta)} u^{\alpha-1}(1-u)^{-\alpha} du. \tag{1.3.2}$$

For the RHT dynamics on the REM, following the same general program as for the complete graph, it can be shown that for scales g_N, r_N as in (1.1.2), with certain restrictions on the parameters α and β,

$$g_N^{-1} S(\lfloor t r_N \rfloor) \xrightarrow{n\to\infty} V_\alpha(t).$$

Applying the same heuristics as for the complete graph, in $r_N \ll 2^N$ steps the simple random walk on the REM behaves essentially 'ballistic' and finds about r_N states, which

by (1.1.3) implies that the deepest traps observed are of depth of order g_N. Then the time spent in these deep traps is of order g_N, so that aging for the RHT dynamics should occur on the time scale g_N. Via the arcsine law, the convergence of the clock process again provides aging statements for two-point functions as in (1.2.3) on time scales $t_w(N) = g_N$. Note that as in (1.3.1) with $\kappa \in (0,1)$, the time scales g_N can be chosen somewhat freely by adjusting the parameters α and β.

Let us mention here that so-called *extremal aging* occurs when the dynamics is observed on very short time scales, and this can be described by convergence of rescaled clock processes to extremal processes. This has been done not only on the REM [Gü09] but also on the SK and p-spin models [BG12, BGŠ13]. Extremal processes can be interpreted as the limit case of an α-stable process for $\alpha \to 0$.

1.4. Trap models on integer lattices

Besides the hypercube or the complete graph suggested by spin-glass applications, it is natural to consider Bouchaud's trap model also on other graphs such as the integer lattices $\mathbb{Z}^d$, $d \geq 1$. In fact, incorporating knowledge from related areas such as random walk among random conductances, the investigation of trap models on the integer lattices has been, to some extent, more fruitful than on the hypercube, in particular on the level of asymmetric dynamics.

The random environment for trap models on $\mathbb{Z}^d$ is mostly given directly by a collection of i.i.d. random variables $\tau = (\tau_x)_{x\in\mathbb{Z}^d}$ which display a heavy-tailed behavior. The dynamics, in this framework sometimes also called a 'trapped random walk', is then defined as a Markov chain reversible with respect to τ. The RHT dynamics on $\mathbb{Z}^d$ is the Markov chain with transition rates $\frac{1}{2d}\tau_x^{-1}$ from x to nearest-neighbors $y \sim x$. Analogously as in (1.1.4) and (1.1.5), Metropolis dynamics and ABD are defined.

Similarly as for spin-glass models one can then try to establish aging in terms of scaling limits of clock processes and convergence of two-point functions. In addition, the question of scaling limits of the Markov chain itself arises, in particular whether it differs from Brownian motion, the scaling limit of simple random walk.

Aging for the RHT dynamics on $\mathbb{Z}^d$ was studied e.g. in [FIN02, BČ05, BČM06, FM14], and for the ABD on $\mathbb{Z}^d$ in [BČ05, GŠ13]. The scaling limit of the RHT dynamics on $\mathbb{Z}$ was found in [FIN02], and in higher dimensions in [BČ07]. For the ABD on $\mathbb{Z}^d$, $d \geq 2$, the same scaling limit was established in [Mou11, BČ11, Čer11]. An important role in most results is played by the convergence of clock processes to stable Lévy processes. This concept has on $\mathbb{Z}^d$ been successfully extended from the symmetric RHT dynamics to asymmetric dynamics, in particular the ABD.

Let us describe next the different scaling limits of trap models and how they motivated the model that is investigated in Chapter 3.

The scaling limit of the RHT dynamics on $\mathbb{Z}$ is the *Fontes-Isopi-Newman (FIN) singular diffusion* [FIN02]. The FIN diffusion is a Markov process, and it can be expressed as the time change of Brownian motion by a functional of the local time $\ell(t,y)$ of the Brownian motion and a speed measure ρ, a random discrete measure on $\mathbb{R}$ which provides in some sense a random environment. Given ρ, the FIN diffusion is the time change of

Brownian motion by the inverse of

$$S_\rho(t) = \int \ell(t,y) d\rho(y). \tag{1.4.1}$$

As the time-change process depends on the local time of the Brownian motion, it is not independent of the Brownian motion.

In dimensions $d \geq 2$, the scaling limit of the RHT dynamics is completely different. It is the *Fractional Kinetics (FK) process* [BČ07], which is non-Markovian and not given by some random environment. It is the time change of Brownian motion by the inverse of a stable Lévy process which is independent of the Brownian motion.

There is another classical model that admits both Brownian motion and the FK process as scaling limits in all $\mathbb{Z}^d$, $d \geq 1$, namely the *continuous-time random walk* (CTRW) of [MW65]. This is a simple random walk on the integer lattice whose waiting times, in contrary to the RHT dynamics, do not depend on the location. The waiting time at each step is picked independently from some fixed distribution, so that the clock process is in fact a sum of i.i.d. random variables. Its scaling limit is thus either linear or a stable Lévy process, which provides Brownian motion and the FK process as scaling limits of the CTRW.

To better understand the different scaling limit regimes that arise in the above examples, [BCČR14] introduced *randomly trapped random walk* (RTRW). The discrete skeleton of RTRW in [BCČR14] is simple random walk on some graph G, and the waiting times are determined in the following two-step procedure. First, a random environment is given in terms of an i.i.d. collection of probability distributions $(\pi_x)_{x\in G}$ on $(0,\infty)$. The waiting time at some vertex $x \in G$ is then picked from the distribution π_x, independently at every visit to x.

In this way the model contains Bouchaud's RHT dynamics (by taking the π_x's to be exponential distributions with means that are heavy-tailed random variables) as well as the CTRW (by taking all π_x's to be deterministically identical). Moreover, it provides a tool to analyze trap models where the trapping is induced by the geometry of the graph rather than by a random environment, e.g. the simple random walk on the incipient critical Galton-Watson tree projected to the backbone.

The model was defined in [BCČR14] on general graphs and studied in detail on the one-dimensional integer lattice $\mathbb{Z}$. It was shown that on $\mathbb{Z}$ there are in fact four different possibilities for scaling limits:

(i) Brownian motion,
(ii) FK process,
(iii) FIN singular diffusion,
(iv) spatially subordinated Brownian motion.

The new class of *spatially subordinated Brownian motion* is a generalized version of the FIN singular diffusion.

It was conjectured in [BCČR14] that among the four classes above, the latter two cases appear only if the underlying random walk is sufficiently recurrent. In particular, on the integer lattices in dimensions $d \geq 2$ the possible scaling limits should reduce to Brownian motion and the FK process.

In Chapter 3 we prove this conjecture for an even slightly more general version of RTRW. We consider the same environment of random probability distributions, but we

allow the discrete skeleton to be any random walk. The main tool for the analysis of scaling limits of RTRW are scaling limits of the associated clock processes.

In the main result of Chapter 3 we show, under very weak assumptions on the discrete skeleton, that for RTRW on $\mathbb{Z}^d$, $d \geq 2$, the only possible scaling limits of clock processes are a linear deterministic process or a stable Lévy process, cf. Theorem 3.1.1. As a consequence, under the additional assumption that the discrete skeleton itself scales to Brownian motion, the respective scaling limits of the RTRW are Brownian motion or the FK process, cf. Theorem 3.1.2. Moreover, we give sufficient conditions for convergence in both cases in Theorems 3.1.3 and 3.1.4.

Let us note that the model of RTRW does not include ABD or Metropolis dynamics, since it is constructed as the time change of a (not necessarily simple) random walk which is independent of the random trapping environment. It would be interesting to introduce a general model in the spirit of RTRW that allows for dependency between environment and transition probabilities and thus includes asymmetric dynamics such as ABD and Metropolis.

1.5. Aging of the Metropolis dynamics on the REM

Before describing the main result of Chapter 2, let us outline the main steps towards understanding aging of mean-field spin glasses that have been taken so far. Recall that we are aiming at justifying the three simplifications made in Bouchaud's trap model as pointed out at the end of Section 1.2.

The first step in this direction was done in [BBG02, BBG03a, BBG03b], where the authors consider the RHT dynamics on the REM and show, using highly non-trivial renewal methods, that a certain two-point function similar to Π_n^τ as defined in (1.2.2) exhibits an aging regime in the form of (1.2.3), for time scales that are close to the equilibration of the system. The results for aging of the RHT dynamics on the REM were extended in the paper [BČ08], in particular on a broader range of time scales. As described before, this paper also establishes the convergence of clock processes to stable Lévy processes as the main tool to prove aging.

After having verified the predictions of Bouchaud's trap model for RHT dynamics on the REM, i.e. moving on from the complete graph to the 'correct' state space, the hypercube $\mathbb{H}_N$, the next step is to try to remove, one at a time, the remaining simplifications on the two levels of the Hamiltonian and the dynamics.

Replacing the i.i.d. Hamiltonian of the REM by a correlated one was successfully done in [BBČ08] an strengthened in [BG13], where aging was established for the RHT dynamics on the p-spin SK model, $p \geq 3$.

We next focus on the task of replacing the symmetric RHT dynamics by asymmetric dynamics. As mentioned in Section 1.4, a rather complete understanding of aging of the ABD (and implicitly for the Metropolis dynamics) in terms of convergence of two-point functions as well as in terms of convergence of clock processes is available for i.i.d. environments on the integer lattices [BČ05, Mou11, BČ11, Čer11, GŠ13]. Aging of the ABD in the i.i.d. environment of Bouchaud's model on the complete graph is also well understood [Gay12]. Since the aging behavior for asymmetric dynamics on these models is the same as for the RHT dynamics, it is expected that the same aging behavior should occur also on the mean-field spin glass models, at least in the i.i.d. case of the REM.

Though there have been attempts prior to this thesis to describe aging in terms of convergence of clock processes for the ABD and the Metropolis dynamics on the REM, they have been restricted to modified or truncated models:

Mathieu and Mourrat [MM15] consider the ABD on the REM for an asymmetry parameter $a_N \sim \sqrt{\log N/N}$ (cf. (1.1.5)) which tends to zero with the size N of the hypercube. This has the effect that this modified ABD asymptotically recovers certain properties of the RHT dynamics, in the sense that the 'natural accelerated version' of the ABD behaves 'ballistic' like the simple random walk.

In the paper [Gay14], Gayrard investigates the non-modified Metropolis dynamics (1.1.4) but on a truncated REM. In this model all energies E_x for $x \in \mathbb{H}_N$ such that $E_x < u_N$ are set equal to zero, for u_N such that $\mathbb{P}[E_x \geq u_N] = N^{-c}$ for some $c > 3$. Due to this truncation the discrete skeleton of the Metropolis dynamics is simple random walk except on sites with non-zero energy which have at least one other non-zero neighbor. By the choice of the level u_N there is only a sparse set of such sites, consisting of small, disjoint connected components of size smaller than N. Therefore, this model also recovers useful properties of the simple random walk.

The discrete skeleton of the Metropolis dynamics on the non-truncated REM is however far from being a simple random walk. Nevertheless, we give in Chapter 2 a first proof for aging of the Metropolis dynamics on the full REM. The main result in Chapter 2 provides a stable Lévy process as the scaling limit of the clock process between the Metropolis dynamics (1.1.4) and another continuous-time Markov chain on $\mathbb{H}_N$, defined by the rates

$$q_{xy} = \frac{\tau_x \wedge \tau_y}{1 \wedge \tau_x} \mathbf{1}_{\{x \sim y\}}, \qquad x, y \in \mathbb{H}_N.$$

This 'fast chain' is an acceleration of the Metropolis dynamics by the factor $(1 \vee \tau_x)$ at every state x. The clock process giving the time change between the Metropolis dynamics and the accelerated chain is given by

$$S(t) = \int_0^t (1 \vee \tau_{Y_s}) ds.$$

In Theorem 2.1.1 we will show that this clock process, properly rescaled, converges to an α-stable Lévy process V_α. Namely, we show that

$$g_N^{-1} S(t R_N) \xrightarrow{N \to \infty} V_\alpha(t), \tag{1.5.1}$$

where the convergence holds in probability with respect to the environment of the REM, and in distribution with respect to the law of the process for a fixed environment. The topology in which the distributional convergence holds is the Skorohod M_1-topology.

The scale g_N is the same as for the REM, as defined in (1.1.2). An interesting feature of the convergence (1.5.1) is that the scale R_N is random, it depends on the realization of the environment. However, we show that R_N differs from r_N as defined in (1.1.2) eventually only in subexponential prefactors.

Our choice of the accelerated chain differs slightly from the natural choice (taken e.g. in [MM15]), which would be a Markov chain reversible with respect to the uniform measure (somewhat analogously to the simple random walk for the RHT dynamics). However, since the important object for aging in terms of two-point functions is the range of the

limiting stable Lévy process, the exact choice of the accelerated chain is not relevant. The reason for exactly the above choice is that this chain is still uniform on the relevant deep traps, and we are able to control its mixing time to be polynomial in N, cf. the remarks after Theorem 2.1.1.

The convergence (1.5.1) of the rescaled clock process captures the basic aging behavior and confirms the predictions of Bouchaud's trap model on that level. We do however not give an aging statement in terms of a two-point correlation function such as (1.2.3), cf. the discussion of the method in the next section.

1.6. Methods for obtaining convergence of clock processes

To conclude this introductory chapter, we describe the different methods used in the literature to prove aging and scaling limit results in terms of scaling limits of clock processes. In particular, we will outline the novelties in the methods of Chapters 2 and 3.

As already mentioned, the first aging results for the spin-glass models were obtained using renewal arguments [BD95, BBG02, BBG03a, BBG03b]. A clock process was for the first time analyzed in [BČM06] in the context of the RHT on the two-dimensional integer lattice. This paper introduces the so-called *coarse-graining*, a procedure that relies on the fact that, for a fixed environment, the clock process is essentially determined by the visits to certain deep traps. The trajectory of the process is cut into pieces, and it is shown that in each piece there is typically no such deep trap visited. If one is visited, then it is the only one, and in different pieces there are always different deep traps visited. With this, the contributions to the clock process of the consecutively found deep traps can be approximated by an i.i.d. sequence of random variables. The rescaled partial sum process of this sequence is then shown to converge to a stable Lévy process, for almost every realization of the environment.

The coarse-graining procedure was further used in [BČ07] to determine the FK process as the scaling limit of the RHT dynamics on $\mathbb{Z}^d$, $d \geq 2$, and it was extended in [BČ11, Čer11] to analyze asymmetric dynamics on $\mathbb{Z}^d$, $d \geq 2$. For the RHT dynamics, the method relies on fine estimates for simple random walk. On the other hand, the treatment of asymmetric dynamics relies heavily on the knowledge about the *random conductance model* on $\mathbb{Z}^d$, in particular the regime of heavy-tailed conductances [BD10]. As such detailed knowledge is not available for the random conductance model on finite graphs, it is not clear whether the coarse-graining procedure could be successfully adapted to asymmetric dynamics for example on the REM.

The paper [BČ08], which develops the general scheme for proving convergence of clock processes for the RHT dynamics, somewhat formalizes the coarse-graining procedure. It gives abstract conditions on the trapping environment and certain potential theoretic objects related to the simple random walk on a graph that ensure aging in terms of convergence of the rescaled clock process. The program is roughly as follows. The first step is to identify, for a fixed environment, a certain set of vertices which essentially determine the clock process, the deep traps. The contribution of the remaining vertices can be ignored. The second step is then to understand how the deep traps contribute to the clock process. To this end, the simple random walk should hit the deep traps essentially 'uniformly' and in exponentially distributed times. The time spent in deep traps can be controlled by the diagonal Green function.

A different method, introduced in [Gay12], was put forward recently. The clock process is viewed as a partial sum process of dependent random variables, and one then uses abstract conditions due to Durrett and Resnick [DR78] that provide functional limit theorems for such sums. Besides the RHT dynamics [Gay10, BG13, BGŠ13], this method was successfully applied to asymmetric dynamics, for example the ABD on the complete graph [Gay12] and the integer lattices [GŠ13], and recently to the Metropolis dynamics on the truncated REM [Gay14].

In contrary to the methods above which provide *quenched* statements directly, i.e. statements that hold for almost every realization of the environment, another method was introduced in [Mou11] that is particularly suitable for obtaining *annealed* results, i.e. results that hold averaged over the environment. This method relies on the fact that when a new trap is visited, in the annealed setting its depth can be determined at the moment of this first visit. If there are essentially no returns to the relevant deep traps, the consecutively found Gibbs masses of deep traps are trivially an i.i.d. sequence and convergence to a stable process can be deduced. The method was used in [Mou11] to determine the FK process as the scaling limit of the ABD on $\mathbb{Z}^d$ in dimensions $d \geq 5$ with an i.i.d. environment. It was developed further in [MM15] to obtain annealed aging results for the modified ABD with asymmetry parameter tending to zero.

Let us mention the approach of [FM14], which partly inspired the method of Chapter 3. In this paper the authors show aging for RHT-like dynamics in i.i.d. environments, where the discrete skeleton is a not necessarily simple random walk which is independent of the environment. As the one of [Mou11, MM15], the method of [FM14] works in the annealed setting and relies on the fact that the depths of newly discovered traps are i.i.d. random variables. This idea is combined with properties obtained from assumptions on a law of large numbers for the range and slow variation of the escape probability of the discrete skeleton.

The proofs for aging on the one-dimensional integer lattice [FIN02, BČ05] use the property that processes on this discrete one-dimensional space can be expressed as a 'time-scale change of Brownian motion'. The methods of [BCČR14] to classify the scaling limits of RTRW in one dimension build up on these techniques. The RTRW process on the discrete space is expressed as the time change of a simple random walk, where the time-change process is a functional of the local time of the random walk and a certain random 'trap measure'. The convergence of the rescaled RTRW can be deduced from the convergence of the associated trap measures. The limit process is then expressed as the time change of Brownian motion by a time-change process which is a functional of the local time of Brownian motion and the limit measure of the trap measures (as e.g. in (1.4.1) for the FIN diffusion). The classification of possible limit processes thus reduces to a classification of the possible limiting trap measures, which is achieved using the well developed theory on random measures.

Let us now very briefly outline the main ideas and novelties in the methods used in the first part of this thesis, focusing mainly on those ideas that are common for the Chapters 2 and 3.

Since the methods of [BCČR14] to analyze randomly trapped random walks in one dimension heavily rely on the local time of Brownian motion, they are not applicable in dimensions $d \geq 2$. The approach in Chapter 3 is inspired by [FM14] and uses the fact that the range of the random walk and certain related objects satisfy laws of large numbers.

This allows to control the contribution of frequently visited sites and deduce the possible scaling limits of the clock process of RTRW on $\mathbb{Z}^d$, $d \geq 2$, by showing that it must be a process of independent and identically distributed increments.

The sufficiency criterion for convergence of RTRW to Brownian motion is an easy consequence of ergodicity. The sufficiency criterion for convergence of RTRW to the FK process, on the other hand, is rather abstract. It is chosen such that when computing the Laplace transform of one-dimensional marginals of the clock process, it turns out that this Laplace transform converges to the Laplace transform of a stable law.

The main idea in the method of Chapter 2 to show convergence of the clock process of the Metropolis dynamics is inspired by the above approach to give a sufficiency criterion for convergence of RTRW to the FK process.

In the known methods from the literature, the clock process is mostly somehow simplified, reduced, and approximated, and then it is shown that this approximated process converges to a stable Lévy process by e.g. computing the Laplace transform. Taking up the idea from the sufficiency criterion in Chapter 3, we somewhat invert this procedure and directly compute a certain conditional Laplace transform. We then ask what conditions we might need in order to obtain convergence to the Laplace transform of a stable law.

It turns out that basically the only thing needed for this convergence is concentration of a certain 'local time functional'. This of course requires some detailed analysis, but it probably requires less work than would be necessary when applying the known methods, if this is even possible.

The three main technical ingredients for the concentration of the local time functional are the fast mixing of the fast chain, bounds on the mean hitting time of certain deep traps by the fast chain, and the choice of the random scale R_N. Compared to most of the other methods presented before, we do not need any control on the local behavior of the fast chain. We consider this as the main advantage of our method. On the other hand, the lack of local control brings along the disadvantage that it is not possible to deduce any reasonable two-point function statement.

The direct computation of the Laplace transform of the clock process and hence the reduction of the problem to the concentration of the local time functional relies on a specific property of the Metropolis dynamics. We nevertheless believe that the method can be extended to also treat the ABD on the REM. Since the method however exploits the independence structure of the REM, it is not clear whether it can be extended to treat dependent environments.

2. Aging of the Metropolis dynamics on the Random Energy Model

Jiří Černý and Tobias Wassmer

Abstract. We study the Metropolis dynamics of the simplest mean-field spin glass model, the Random Energy Model. We show that this dynamics exhibits aging by showing that the properly rescaled time-change process between the Metropolis dynamics and a suitably chosen 'fast' Markov chain converges in distribution to a stable subordinator. The rescaling might depend on the realization of the environment, but we show that its exponential growth rate is deterministic.

2.1. Introduction

This paper studies the out-of-equilibrium behavior of the *Metropolis dynamics* on the *Random Energy Model* (REM). Our main goal is to answer one of the remaining important open questions in the field, namely whether this dynamics exhibits aging, and, if yes, whether its aging behavior admits the usual description in terms of stable Lévy processes.

Aging is one of the main features appearing in the long-time behavior of complex disordered systems (see e.g. [BCKM97] for a review). It was for the first time observed experimentally in the anomalous relaxation patterns of the residual magnetization of spin-glasses (e.g. [LSNB83, Cha84]). One of the most influential steps in the theoretical modeling of the aging phenomenon is the introduction of the so-called *trap models* by Bouchaud [Bou92] and Bouchaud and Dean [BD95]. These models, while being sufficiently simple to allow analytical treatment, reproduce the characteristic power law decay seen experimentally.

Since then a considerable effort has been made in putting the predictions obtained from the trap models to a solid basis, that is to derive these predictions from an underlying spin-glass dynamics. The first attempt in this direction was made in [BBG02, BBG03a, BBG03b] where it was shown that, for a very particular Glauber-type dynamics, at time scales very close to the equilibration, a well chosen two-point correlation function converges to that given by Bouchaud's trap model.

With the paper [BČ08], where the same type of dynamics was studied in a more general framework and on a broader range of time scales, it emerged that aging establishes itself by the fact that scaling limits of certain additive functionals of Markov chains are stable Lévy processes, and that the convergence of the two-point correlation functions is just a manifestation of the classical arcsine law for stable subordinators.

The Glauber-type dynamics used in those papers, sometimes called *Random Hopping Time* (RHT) dynamics, is however rather simple and is often considered as 'non-realistic', mainly because its transition rates do not take into account the energy of the target state. Its advantage is that it can be expressed as a time change of a simple random walk on the configuration space of the spin-glass, which allows for a certain decoupling of the randomness of the dynamics from the randomness of the Hamiltonian of the spin-glass, making its rigorous studies more tractable.

For more realistic Glauber-type dynamics of spin-glasses, like the so-called *Bouchaud's asymmetric dynamics* or the *Metropolis dynamics*, such decoupling is not possible. As a consequence, these dynamics are far less understood.

Recently, some progress has been achieved in the context of the simplest mean-field spin-glass model, the REM. First, in [MM15], the Bouchaud's asymmetric dynamics have been considered in the regime where the asymmetry parameter tends to zero with the size of the system. Building on the techniques started in [Mou11], this papers confirms the predictions of Bouchaud's trap model in this regime. Second, the Metropolis dynamics have been studied in [Gay14], for a truncated version of the REM, using the techniques developed for the symmetric dynamics in [Gay12, Gay10], again confirming Bouchaud's predictions.

The weak asymmetry assumption of [MM15] and the truncation of [Gay14] have both the same purpose. They aim at overcoming some specific features of the asymmetry and recovering certain features of symmetric dynamics. Our aim in this work is to get rid of this simplifications and treat the non-modified REM with the usual Metropolis dynamics.

Let us also mention that Bouchaud's asymmetric dynamics (and implicitly the Metropolis one) is rather well understood in the context of trap models on $\mathbb{Z}^d$, see [BČ11, Čer11, GŠ13], where it is possible to exploit the connections to the random conductance model with unbounded conductances, [BD10]. Finally, asymmetric dynamics on the complete graph were considered in [Gay12].

Before stating our main result, let us briefly recall the general scheme for proving aging in terms of convergence to stable Lévy processes. The actual spin-glass dynamics, $X = (X_t)_{t\geq 0}$, which is reversible with respect to the Gibbs measure of the Hamiltonian, is compared to another Markov chain $Y = (Y_t)_{t\geq 0}$ on the same space, which is an 'accelerated' version of X and whose stationary measure is uniform. The process Y is typically easier to be understood, e.g. it is a simple random walk for the RHT dynamics, and the original process X can be written as its time change,

$$X(t) = Y(S^{-1}(t)), \tag{2.1.1}$$

for the right continuous inverse S^{-1} of a certain additive functional S of the Markov chain Y, called the *clock process*. The aim is then to show convergence of the properly rescaled clock process S to an increasing stable Lévy process, that is to a stable subordinator.

We now state our main result. We consider the unmodified REM, as introduced in [Der80, Der81]. The state space of this model is the N-dimensional hypercube $\mathbb{H}_N = \{-1, 1\}^N$, and its Hamiltonian is a collection $(E_x)_{x\in\mathbb{H}_N}$ of i.i.d. standard Gaussian random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The non-normalized Gibbs measure $\tau_x = e^{\beta\sqrt{N}E_x}$ at inverse temperature $\beta > 0$ gives the equilibrium distribution of the system.

The Metropolis dynamics on the REM is the continuous-time Markov chain $X = (X_t)_{t\geq 0}$ on $\mathbb{H}_N$ with transition rates

$$r_{xy} = \left(1 \wedge \frac{\tau_y}{\tau_x}\right) \mathbf{1}_{\{x\sim y\}}, \qquad x, y \in \mathbb{H}_N. \tag{2.1.2}$$

Here, $x \sim y$ means that x and y are neighbors on $\mathbb{H}_N$, that is they differ in exactly one coordinate.

As explained above, we will compare the Metropolis chain X with another 'fast' Markov chain $Y = (Y_t)_{t\geq 0}$ with transition rates

$$q_{xy} = \frac{\tau_x \wedge \tau_y}{1 \wedge \tau_x} \mathbf{1}_{\{x\sim y\}}, \qquad x, y \in \mathbb{H}_N. \tag{2.1.3}$$

It can be easily checked using the detailed balance conditions that Y is reversible and that its equilibrium distribution is

$$\nu_x = \frac{1 \wedge \tau_x}{Z_N}, \qquad x \in \mathbb{H}_N,$$

where $Z_N = \sum_{x\in\mathbb{H}_N}(1 \wedge \tau_x)$. Finally, since $r_{xy} = (1 \vee \tau_x)^{-1} q_{xy}$, X can be written as a time change of Y as in (2.1.1) with the clock process S being given by

$$S(t) = \int_0^t (1 \vee \tau_{Y_s}) ds. \tag{2.1.4}$$

For the rest of the paper we only deal with the process Y and the clock process S, the actual Metropolis dynamics X does not appear anymore after this point. For a fixed environment $\tau = (\tau_x)_{x\in\mathbb{H}_N}$, let P_ν^τ denote the law of the process Y started from its stationary distribution ν, and let $D([0,T],\mathbb{R})$ be the space of $\mathbb{R}$-valued cadlag functions on $[0,T]$. We denote by $\beta_c = \sqrt{2\log 2}$ the (static) critical temperature of the REM. Our main result is the following.

Theorem 2.1.1. *Let $\alpha \in (0,1)$ and $\beta > 0$ be such that*

$$\frac{1}{2} < \frac{\alpha^2\beta^2}{\beta_c^2} < 1, \tag{2.1.5}$$

and define

$$g_N = e^{\alpha\beta^2 N}(\alpha\beta\sqrt{2\pi N})^{-\frac{1}{\alpha}}. \tag{2.1.6}$$

Then there are random variables R_N which depend on the environment $(E_x)_{x\in\mathbb{H}_N}$ only, such that for every $T > 0$ the rescaled clock processes

$$S_N(t) = g_N^{-1} S(tR_N), \qquad t \in [0,T],$$

converge in $\mathbb{P}$-probability as $N \to \infty$, in P_ν^τ-distribution on the space $D([0,T],\mathbb{R})$ equipped with the Skorohod M_1-topology, to an α-stable subordinator V_α. The random variables R_N satisfy

$$\lim_{N\to\infty} \frac{\log R_N}{N} = \frac{\alpha^2\beta^2}{2}, \quad \mathbb{P}\text{-a.s.} \tag{2.1.7}$$

Let us make a few remarks on this result.

1. The result of Theorem 2.1.1 confirms that the predictions of Bouchaud's trap model hold for the Metropolis dynamics on the REM, at least at the level of scaling limits of clock processes. It also compares directly to the results obtained for the symmetric (RHT) dynamics in [BČ08]. The scales g_N and R_N are (up to subexponential prefactors) the same as previously, including the condition (2.1.5) or the range of parameters α, β. As in [BČ08], the right inequality in (2.1.5) is completely natural, beyond it Y 'feels' the finiteness of $\mathbb{H}_N$ and aging is not expected to occur. The left inequality in (2.1.5) is technical, it ensures that the relevant deep traps are well separated (cf. Lemma 2.2.1), introducing certain simplifications in the proof. We believe that this bound might be improved to $\alpha^2\beta^2/\beta_c^2 > 0$, by further exploiting our method. Finally, as previously, note that (2.1.5) is satisfied also for $\beta < \beta_c$ for appropriate α, hence aging can occur above the critical temperature.

2. Our choice of the fast chain Y is rather unusual. In view of the previous papers [MM15, BČ11], it would be natural to take instead the chain $\tilde{Y}$ with uniform invariant measure and with transition rates $\tau_x \wedge \tau_y$, that is without the correction $1 \wedge \tau_x$ which appears in (2.1.3). This choice has, however, some deficiencies. On the heuristic level, $\tilde{Y}$ is not an acceleration of X, since it is much slower than X on sites with very small Gibbs measure $\tau_x \ll 1$. These sites, which are irrelevant for the statics, then 'act as traps' on $\tilde{Y}$, making them relevant for the dynamics, which is undesirable. On the technical level, the trapping on sites with small Gibbs measure has the consequence that the mixing time of $\tilde{Y}$ is very large.

Our choice of the fast chain Y runs as fast as X on the sites with small Gibbs measure and thus does not have this deficiency. Moreover, since $\nu_x = Z_N^{-1}$ whenever $E_x \geq 0$, the equilibrium distribution of the fast chain Y is still uniform on the relevant deep traps, so the clock process S retains its usual importance for aging.

Remark also that in order to overcome similar difficulties with the slowly mixing chain $\tilde{Y}$, [MM15] truncate the Hamiltonian of the REM at 0 which effectively sets $\tau_x \geq 1$ for all $x \in \mathbb{H}_N$. We prefer to retain the full REM and use the modified fast chain Y instead. Finally, [Gay14] uses the discrete skeleton of X as the base chain, which has some interesting features but introduces similar undesirable effects.

3. We view Theorem 2.1.1 as an aging statement, without further considering any two-point correlation functions. Actually, it seems hard to derive aging statements for the usual correlation functions from our result without extending the paper considerably. Such derivation usually requires some knowledge of the fast chain Y that goes over the M_1-convergence of the clock processes. This knowledge is typically automatically obtained in the previous approaches. The strength (or the weakness) of our method is that we do not need to obtain such finer knowledge to show the clock process convergence.

4. A rather unusual feature of Theorem 2.1.1 is the fact that the scaling R_N is random, it depends on the random environment. This is again a consequence of our technique. Claim (2.1.7) in Theorem 2.1.1 however shows that at least the exponential growth of R_N is deterministic. The random scale R_N is explicitly defined in (2.2.10). We will see that its definition depends on a somewhat free choice of an auxiliary parameter, but nevertheless the final result does not depend on this parameter. Not only this property makes us

conjecture that R_N should actually satisfy a deterministic law of large numbers,

$$\lim_{N\to\infty} h(N)e^{-\alpha^2\beta^2 N/2}R_N = 1, \qquad \mathbb{P}\text{-a.s.},$$

for some function $h(N)$ growing at most subexponentially.

5. The mode of convergence in Theorem 2.1.1 is not optimal, one would rather like to obtain the convergence in P_ν^τ-distribution for $\mathbb{P}$-almost every environment, which is usually called *quenched* convergence. Actually, Theorem 2.1.1 can be strengthened slightly to a statement which is somewhere between $\mathbb{P}$-a.s. convergence and convergence in $\mathbb{P}$-probability. Namely, the statement holds for a.e. realization of sites with 'small' τ_x, but only in probability over sites with 'large' τ_x, cf. Remark 2.6.4.

6. Our proof of Theorem 2.1.1 strongly exploits the i.i.d. structure of the Hamiltonian of the REM. At present we do not know if it is possible to combine our techniques with those used for the RHT dynamics of the p-spin model in [BBČ08, BG12].

We proceed by commenting on the proof of Theorem 2.1.1, concentrating mainly on its novelties. The general strategy so far to prove such a result has been to first reduce the problem to the clock process restricted to a set of deep traps which govern the behavior of the original clock process. The different methods then all more or less aim at dividing the contribution of consequently found deep traps into essentially i.i.d. blocks. For example in [BČ08] or [BČ11], this is achieved by controlling the hitting probabilities of deep traps, proving that they are hit essentially uniformly in exponentially distributed times, and controlling the time the chain spends at the deep traps by a sharp control of the Green function. Similar rather precise estimates on hitting probabilities and/or Green function are necessary in other approaches. Using this i.i.d. structure, one can then show convergence of the clock process by standard methods, e.g. computing the Laplace transform.

The method used in this paper is slightly inspired by the general approach taken in [FM14] and [ČW15b]. There, models of trapped random walks on $\mathbb{Z}^d$ are considered where few information about the discrete skeleton as well as the waiting times of a continuous-time Markov chain are available, and minimal necessary conditions for convergence of the clock process are found. Taking up this idea, instead of analyzing in detail the behavior of the fast chain Y, we extract the minimal amount of information needed to show convergence of the clock process. In particular, we do not need any exact control of hitting probabilities and Green functions of deep traps, as most previous work did.

The first step in our proof is standard, namely that the main contribution to the clock process comes from a small set of vertices with large Gibbs measure τ_x, the so-called deep traps, and that in fact the clock process of the deep traps converges to a stable subordinator. Denote the set of deep traps by $\mathcal{D}_N$ (see Section 2.2 for details). We will show that the clock process S can be well approximated by the *clock process of the deep traps*

$$S_{\mathcal{D}}(t) = \int_0^t (1 \vee \tau_{Y_s})\mathbf{1}_{\{Y_s\in\mathcal{D}_N\}}ds. \tag{2.1.8}$$

Then it remains to show that in fact $g_N^{-1}S_{\mathcal{D}}(tR_N)$ converges to a stable subordinator.

To this end, we will in some sense invert the standard procedure described above. Instead of approximating the clock process by an i.i.d. block structure and then use

the Laplace transform to show convergence, we will first compute a certain conditional Laplace transform using some special properties of the Metropolis dynamics. Then we analyze what is actually needed in order to show convergence of the unconditional Laplace transform.

A bit more detailed, this will be done as follows. Under condition (2.1.5), the deep traps are almost surely well separated. This fact and the fact that the definition (2.1.3) contains the factor $\tau_x \wedge \tau_y$ imply that the transition rates q_{xy} of the fast chain Y do not depend on the energies E_x of the deep traps, but only on their location. Therefore, one can condition on the location of all traps and the energies E_x of the non-deep traps, which determines the law P_ν^τ of Y, and take the expectation over the energies of the deep traps. We call this a *quasi-annealed* expectation and denote it by $\mathbb{E}_\mathcal{D}$ for the moment. Let $\ell_t(x)$ denote the local time of the fast chain Y (see Section 2.2 for details). As $\mathbb{E}_\mathcal{D}$ is simply an expectation over i.i.d. random variables, the quasi-annealed Laplace transform of the rescaled clock process of the deep traps given Y can be computed. It essentially behaves like

$$\mathbb{E}_\mathcal{D}\big[e^{-\lambda\frac{1}{g_N}S_\mathcal{D}(tR_n)} \mid Y\big] \approx \exp\Big\{ -\mathcal{K}\lambda^\alpha \varepsilon_N \sum_{x\in\mathcal{D}_N} \ell_{tR_N}(x)^\alpha \Big\}. \tag{2.1.9}$$

Here, ε_N is a deterministic sequence tending to 0 as $N \to \infty$. The above approximation shows that the only object related to Y we have to control is the local time functional $\varepsilon_N \sum_{x\in\mathcal{D}_N} \ell_{tR_N}(x)^\alpha$.

We will show that this a priori non-additive functional of Y actually behaves in an additive way, namely that it converges to t as $N \to \infty$, under P_ν^τ for $\mathbb{P}$-a.e. environment τ. For this convergence to hold it is sufficient to have some weak bounds on the mean hitting time of deep traps as well as some control on the mixing of the chain Y together with an appropriate choice of the scale R_N that depends on the environment.

Using standard methods we then strengthen the quasi-annealed convergence to quenched convergence (in the sense of Theorem 2.1.1).

To conclude the introduction, let us comment on how our method might be extended. The key argument in the computation of the quasi-annealed Laplace transform, namely the fact that the chain Y is independent of the depth of the deep traps, seems very specific for the Metropolis dynamics. However, by adapting the method appropriately and using network reduction techniques, we believe that one could also treat Bouchaud's asymmetric dynamics and Metropolis dynamics in the regime where the left-hand side inequality of (2.1.5) fails, i.e. there are neighboring deep traps.

The rest of the paper is structured as follows. Detailed definitions and notations used through the paper are introduced in Section 2.2. In Section 2.3 we analyze the mixing properties of the fast chain Y, which will be crucial at several points later. In Section 2.4 we give bounds on the mean hitting time of deep traps and on the normalizing scale R_N. Using these bounds and the results on the mixing of Y, we show concentration of the local time functional $\varepsilon_N \sum_{x\in\mathcal{D}_N} \ell_{tR_N}(x)^\alpha$ in Section 2.5. We prove convergence of the rescaled clock process of the deep traps in Section 2.6 with the above mentioned computation of the quasi-annealed Laplace transform, using the concentration of the local time functional. Finally, we treat the shallow traps in Section 2.7 by showing that their contribution to the clock process can be neglected. In Appendix 2.A we give the proof of a technical result which is used to bound the expected hitting times in Section 2.4.

2.2. Definitions and notation

In this section we introduce some notation used through the paper and recall a few useful facts. We use $\mathbb{H}_N$ to denote the N-dimensional hypercube $\{-1,1\}^N$ equipped with the usual distance

$$d(x,y) = \frac{1}{2}\sum_{i=1}^{N} |x_i - y_i|,$$

and write $\mathcal{E}_N$ for the set of nearest-neighbor edges $\mathcal{E}_N = \{\{x,y\} : \ d(x,y) = 1\}$.

For given parameters α and β, let

$$\gamma = \frac{\alpha^2\beta^2}{\beta_c^2} \in (1/2, 1), \tag{2.2.1}$$

by condition (2.1.5) in Theorem 2.1.1.

Recall from the introduction that $(E_x : x \in \mathbb{H}_N, N \geq 1)$, is a family of i.i.d. standard Gaussian random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Note that we do not denote the dependence on N explicitly, but we assume that the space $(\Omega, \mathcal{F}, \mathbb{P})$ is the same for all N. For $\beta > 0$ the non-normalized Gibbs factor τ_x is given by $\tau_x = e^{\beta\sqrt{N}E_x}$.

Using the standard Gaussian tail approximation,

$$\mathbb{P}[E_x \geq t] = \frac{1}{t\sqrt{2\pi}}\, e^{-t^2/2}\big(1 + o(1)\big) \quad \text{as } t \to \infty, \tag{2.2.2}$$

we obtain that g_N, as defined in Theorem 2.1.1, satisfies

$$\mathbb{P}[\tau_x > u g_N] = u^{-\alpha} 2^{-\gamma N}\big(1 + o(1)\big).$$

This heuristically important computation explains the appearance of stable laws in the distribution of sums of τ_x: If we observe $2^{\gamma N}$ vertices, then finitely many of them have their rescaled Gibbs measures τ_x/g_N of order unity, and, moreover, those rescaled Gibbs measures behave like random variables in the domain of attraction of an α-stable law.

Recall also that $Y = (Y_t)_{t\geq 0}$ stands for the fast Markov chain whose transition rates q_{xy} are given in (2.1.3), and that $\nu = (\nu_x)_{x\in\mathbb{H}_N}$ denotes the invariant distribution of this chain, $\nu_x = \frac{1\wedge\tau_x}{Z_N}$. For a given environment $\tau = (\tau_x)_{x\in\mathbb{H}_N}$, let P_x^τ and P_ν^τ denote the laws of Y started from a vertex x or from ν respectively, and E_x^τ, E_ν^τ the corresponding expectations.

Note that the normalization factor $Z_N = \sum_{x\in\mathbb{H}_N}(1 \wedge \tau_x)$ satisfies, for every constant $\kappa \in (0, 1/2)$,

$$\kappa 2^N \leq Z_N \leq 2^N \qquad \mathbb{P}\text{-a.s for } N \text{ large enough.} \tag{2.2.3}$$

Indeed, obviously $Z_N \leq 2^N$, and $Z_N \geq \sum_{x\in\mathbb{H}_N} \mathbf{1}_{\{E_x\geq 0\}}$. But $\mathbf{1}_{\{E_x\geq 0\}}$ are i.i.d. Bernoulli random variables, therefore the statement follows immediately by the law of large numbers.

An important role in the study of properties of Y is played by the *conductances* defined by

$$c_{xy} = \nu_x q_{xy} = \frac{\tau_x \wedge \tau_y}{Z_N} \qquad \text{for } x \sim y. \tag{2.2.4}$$

Let θ_s be the left shift on the space of trajectories of Y, that is

$$(\theta_s Y)_t = Y_{s+t}. \tag{2.2.5}$$

Let $H_x = \inf\{t > 0 : \ Y_t = x\}$ be the hitting time of x by Y, J_1 the time of the first jump of Y, and let $H_x^+ = H_x \circ \theta_{J_1} + J_1 = \inf\{t > J_1 : \ Y_t = x\}$ be the return time to x by Y. Similarly define H_A and H_A^+ for a set $A \subset \mathbb{H}_N$. The local time $\ell_t(x)$ of Y is given by

$$\ell_t(x) = \int_0^t \mathbf{1}_{\{Y_s = x\}} ds.$$

Using this notation the clock process S introduced in (2.1.4) can be written as

$$S(t) = \int_0^t (1 \vee \tau_{Y_s}) ds = \sum_{x \in \mathbb{H}_N} \ell_t(x)(1 \vee \tau_x).$$

To define the set of deep traps $\mathcal{D}_N$ and the random scale R_N mentioned in the introduction we introduce a few additional parameters. For $\alpha \in (0,1)$, $\beta > 0$ as in Theorem 2.1.1 and γ as in (2.2.1), we fix γ' and α' such that

$$\frac{1}{2} < \gamma' < \gamma, \quad \text{and} \quad \alpha' = \frac{\beta_c}{\beta}\sqrt{\gamma'}. \tag{2.2.6}$$

An explicit choice of γ' will be made later in Section 2.5. We define the auxiliary scale

$$g'_N = e^{\alpha'\beta^2 N}(\alpha'\beta\sqrt{2\pi N})^{-\frac{1}{\alpha'}},$$

and set

$$\mathcal{D}_N = \{x \in \mathbb{H}_N : \ \tau_x \geq g'_N\}.$$

to be the set of deep traps. By the Gaussian tail approximation (2.2.2) it follows that the density of $\mathcal{D}_N$ satisfies

$$\mathbb{P}[x \in \mathcal{D}_N] = 2^{-\gamma' N}(1 + o(1)). \tag{2.2.7}$$

We quote the following observation on the size and sparseness of $\mathcal{D}_N$. The sparseness will play a key role in our computation of the quasi-annealed Laplace transform in Section 2.6.

Lemma 2.2.1 ([BČ08, Lemma 3.7])**.** *For every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,*

$$|\mathcal{D}_N| 2^{(\gamma'-1)N} \in (1-\varepsilon, 1+\varepsilon). \tag{2.2.8}$$

Moreover, since $\gamma' > 1/2$, there exists $\delta > 0$ such that $\mathbb{P}$-a.s. for N large enough, the separation event

$$\mathscr{S} = \{\min\{d(x,y) : \ x, y \in \mathcal{D}_N\} \geq \delta N\} \tag{2.2.9}$$

holds.

Finally, for the sake of concreteness, let us give the explicit form of the random scale R_N,

$$R_N = 2^{(\gamma-\gamma')N}\left(\sum_{x \in \mathcal{D}_N} \frac{E_x^\tau[\ell_{T_{\text{mix}}}(x)^\alpha]}{E_\nu^\tau[H_x]}\right)^{-1}, \tag{2.2.10}$$

where T_{mix} denotes the mixing time of Y, a randomized stopping time which we will construct in Section 2.3. The reason for this definition will become apparent when we prove

the concentration of the local time functional mentioned in the introduction. Although the definition of R_N seems arbitrary by the somewhat free choice of the parameter γ', Theorem 2.1.1 actually shows that asymptotically R_N will be independent of γ'.

For the rest of the paper, c, c', c'' will always denote positive constants whose values may change from line to line. We will use the notation $g = o(1)$ for a function $g(N)$ that tends to 0 as $N \to \infty$, and $g = O(f)$ for a function $g(N)$ that is asymptotically at most of order $f(N)$, i.e. $\lim_{N\to\infty} |g(N)|/f(N) \leq c$, for some $c > 0$.

2.3. Mixing properties of the fast chain

The fact that the chain Y mixes fast, namely on a scale polynomial in N, plays a crucial role in many of our arguments. In this section we analyze the mixing behavior of Y. We first give a lower bound on the spectral gap λ_Y of Y, which we then use to construct a strong stationary time T_{mix}.

Proposition 2.3.1. *There are constants $\kappa > 0$, $K > 0$, $C_0 > 0$, such that $\mathbb{P}$-a.s. for N large enough,*

$$\lambda_Y \geq \frac{\kappa}{4} N^{-K-1-\beta C_0}.$$

We prove this proposition with help of the Poincaré inequality derived in [DS91]. To state this inequality, let Γ be a complete set of self-avoiding nearest-neighbor paths on $\mathbb{H}_N$, that is for each $x \neq y \in \mathbb{H}_N$ there is exactly one path $\gamma_{xy} \in \Gamma$ connecting x and y. Let $|\gamma|$ be the length of the path γ. By Proposition 1' of [DS91], using also the reversibility of Y and recalling the definition (2.2.4) of the conductances, it follows that

$$\frac{1}{\lambda_Y} \leq \max_{e=\{u,v\}\in\mathcal{E}_N} \left\{ \frac{1}{c_{uv}} \sum_{\substack{\gamma_{xy}\in\Gamma:\\ \gamma_{xy}\ni e}} |\gamma_{xy}| \nu_x \nu_y \right\}. \tag{2.3.1}$$

To minimize the right-hand side of (2.3.1), a special care should be taken of the edges whose conductance $c_{uv} = (\tau_u \wedge \tau_v)/Z_N$ is very small, that is which are incident to vertices with very small τ_u. Those 'bad' edges should be avoided if possible by paths $\gamma \in \Gamma$. They cannot be avoided completely, since Γ should be a complete set of paths. On the other hand, if such edge is the first or the last edge of some path γ_{xy}, its small conductance is canceled by equally small ν_x or ν_y. Therefore, to apply (2.3.1) efficiently, one should find a set of paths Γ such that all paths $\gamma \in \Gamma$ avoid 'bad' vertices, except for vertices at both ends of the paths.

In the context of spin-glass dynamics this method was used before in [FIKP98] to find the spectral gap of the Metropolis dynamics (2.1.2). Using the same approach, that is using the same set of paths Γ as in [FIKP98], we could find a lower bound on the spectral gap of the fast chain Y of leading order $\exp\{-c\sqrt{N \log N}\}$. This turns out to be too small for our purposes, cf. Remark 2.6.4.

In the next lemma we construct a set of paths Γ that avoids more 'bad' vertices, which allows to improve the lower bound on the spectral gap to be polynomial in N. This is possible by using an embedding of $\mathbb{H}_N$ into its subgraph of 'good' vertices, i.e. vertices with not too small τ_x, which is inspired by similar embeddings in [HLN87].

For a nearest-neighbor path $\gamma = \{x_0, \dots, x_n\}$, we call the vertices $x_1, \dots, x_{n-1}$ the interior vertices of γ, and the edges $\{x_i, x_{i+1}\}$, $i = 1, \dots, n-2$, the interior edges of γ.

Lemma 2.3.2. *There is an integer $K > 0$ and a constant $C_0 > 0$, such that $\mathbb{P}$-a.s. for N large enough there exists a complete set of paths Γ, such that the following three properties hold.*

(i) For every path $\gamma \in \Gamma$, every interior edge $e = \{u, v\}$ satisfies

$$Z_N c_{uv} = \tau_u \wedge \tau_v \geq N^{-\beta C_0}.$$

(ii) $|\gamma| \leq 8N$ for all $\gamma \in \Gamma$.
(iii) Every edge $e \in \mathcal{E}_N$ is contained in at most $N^K 2^{N-1}$ paths $\gamma \in \Gamma$.

Proof. For $C_0 > 0$, whose value will be fixed later, we say that $x \in \mathbb{H}_N$ is good if $\tau_x \geq N^{-\beta C_0}$, and it is bad otherwise. To construct the complete set of paths Γ satisfying the required properties, we will use the fact that the set of good vertices is very dense in $\mathbb{H}_N$. In particular, we will show that

$$\mathbb{P}\text{-a.s. for } N \text{ large enough, every } x \in \mathbb{H}_N \text{ has at least } \tfrac{1}{2}C_0\sqrt{N} \text{ good neighbors,} \tag{2.3.2}$$

and

$$\begin{aligned}&\mathbb{P}\text{-a.s. for } N \text{ large enough, for any pair of vertices } x, y \text{ at distance 2 or 3,}\\ &\text{there is a nearest-neighbor path of length at most 7 connecting } x \text{ and } y,\\ &\text{such that all interior vertices of this path are good,}\end{aligned} \tag{2.3.3}$$

To prove these two claims, note first that for any $x \in \mathbb{H}_N$, the probability of being bad is

$$\mathbb{P}\big[\tau_x < N^{-\beta C_0}\big] = \mathbb{P}[E_x < -C_0 N^{-\frac{1}{2}} \log N] = \frac{1}{2} - \int_0^{C_0 N^{-\frac{1}{2}} \log N} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds.$$

For N large enough the integrand is larger than $\frac{1}{2}$, and it follows that

$$\mathbb{P}[x \text{ is bad}] \leq \frac{1}{2}\big(1 - C_0 N^{-\frac{1}{2}} \log N\big) =: \frac{1}{2}(1 - q_N).$$

Hence, the number of bad neighbors of a vertex $x \in \mathbb{H}_N$ is stochastically dominated by a Binomial$\big(N, \frac{1}{2}(1-q_N)\big)$ random variable B. For $\lambda > 0$, the exponential Chebyshev inequality yields

$$\begin{aligned}&\mathbb{P}\big[x \text{ has more than } N - \frac{1}{2}C_0\sqrt{N} \text{ bad neighbors}\big]\\ &\quad\leq \mathbb{P}\big[B \geq N - \tfrac{1}{2}C_0\sqrt{N}\big] = \mathbb{P}\big[e^{\lambda B} \geq e^{\lambda(N - \frac{1}{2}C_0\sqrt{N})}\big]\\ &\quad\leq e^{-\lambda(N-\frac{1}{2}C_0\sqrt{N})}\Big(1 + \frac{1}{2}(1-q_N)(e^\lambda - 1)\Big)^N\\ &\quad= e^{-\lambda(N-\frac{1}{2}C_0\sqrt{N})}\bigg(\frac{e^\lambda}{2}\Big(1 - q_N + e^{-\lambda}(1+q_N)\Big)\bigg)^N\\ &\quad\leq 2^{-N} e^{\frac{\lambda}{2}C_0\sqrt{N}}\Big(\exp\{-q_N + e^{-\lambda}(1+q_N)\}\Big)^N.\end{aligned}$$

Since $q_N \to 0$ as $N \to \infty$, the last term in the parenthesis is bounded by $2e^{-\lambda}$ for N large enough. Inserting q_N and choosing $\lambda = \log N$, the above is bounded by

$$\begin{aligned} &2^{-N} \exp\Big\{\frac{1}{2}C_0\sqrt{N}\log N - C_0\sqrt{N}\log N + 2\Big\} \\ &\qquad \le 2^{-N}\exp\Big\{-\frac{1}{4}C_0\sqrt{N}\log N\Big\}, \end{aligned}$$

for N large enough. With a union bound over all $x \in \mathbb{H}_N$ and using the Borel-Cantelli lemma, (2.3.2) follows.

To prove (2.3.3), we first introduce some notation. For a given vertex x and $\{i_1, \dots, i_k\} \subset \{1, \dots, N\}$, denote by $x^{i_1 \cdots i_k}$ the vertex that differs from x exactly in coordinates $i_1, \dots, i_k$. If two vertices x and y are at distance 2, then $y = x^{kl}$ for some $k, l \in \{1, \dots, N\}$. Then for $\{i, j\} \cap \{k, l\} = \emptyset$ we define the path γ_{xy}^{ij} of length 6 as $\{x, x^i, x^{ij}, x^{ijk}, x^{ijkl} = y^{ij}, y^j, y\}$. Similarly, for x, y with $d(x, y) = 3$, we have $y = x^{klm}$, and for $\{i, j\} \cap \{k, l, m\} = \emptyset$ we define the path γ_{xy}^{ij} of length 7 by $\{x, x^i, x^{ij}, x^{ijk}, x^{ijkl}, x^{ijklm} = y^{ij}, y^j, y\}$. Observe that for fixed x, y with $d(x, y) = 2$ or 3 and for different pairs i, j the innermost 3 or 4 vertices of the paths γ_{xy}^{ij} are disjoint.

We now show that with high probability, for every x, y at distance 2 or 3, we may find i, j such that γ_{xy}^{ij} has only good interior vertices. Fix a pair $x, y \in \mathbb{H}_N$ at distance 2 or 3, and let as above k, l or k, l, m be the coordinates in which x and y differ. Assume for the moment that both x and y have at least $\frac{1}{2}C_0\sqrt{N}$ good neighbors. Then there are at least $\frac{1}{4}C_0^2 N$ pairs i, j such that the vertices x^i and y^j are good. Moreover, since it is a matter of dealing with a constant number of exceptions, we may tacitly assume that $i \neq j$, and $\{i, j\} \cap \{k, l\} = \emptyset$ or $\{i, j\} \cap \{k, l, m\} = \emptyset$, respectively.

The remaining interior vertices $\{x^{ij}, x^{ijk}, x^{ijkl} = y^{ij}\}$ or $\{x^{ij}, x^{ijk}, x^{ijkl}, x^{ijklm} = y^{ij}\}$ are all good with probability strictly larger than $1/2$, so the probability that one or more of these vertices are bad is bounded by $15/16$. Since these 3 or 4 innermost vertices are disjoint for different pairs i, j, by independence the probability that among all $\frac{1}{4}C_0^2 N$ pairs $\{i, j\}$ there is none for which all innermost 3 or 4 vertices of γ_{xy}^{ij} are good is bounded by $(15/16)^{\frac{1}{4}C_0^2 N}$. Hence, for one fixed pair $x, y \in \mathbb{H}_N$ at distance 2 or 3, where both x and y have at least $\frac{1}{2}C_0\sqrt{N}$ good neighbors, the probability that there is no path from x to y of length 6 or 7 with all interior vertices good is bounded by

$$(15/16)^{\frac{1}{4}C_0^2 N}.$$

There are less than $2^N(N^2 + N^3)$ pairs of vertices at distance 2 or 3 respectively, and we know from the proof of (2.3.2) that with probability larger than $1 - e^{-c\sqrt{N}\log N}$ every $x \in \mathbb{H}_N$ has at least $\frac{1}{2}C_0\sqrt{N}$ good neighbors. It follows that the probability that the event in (2.3.3) does not happen is bounded by

$$e^{-c\sqrt{N}\log N} + 2^N(N^2 + N^3)(15/16)^{\frac{1}{4}C_0^2 N}. \tag{2.3.4}$$

Choosing $C_0 > \sqrt{\frac{4\log 2}{\log 15/16}}$ and applying the Borel-Cantelli lemma implies (2.3.3).

We now use the density properties (2.3.2) and (2.3.3) of good vertices to define a (random) mapping from the hypercube to its subgraph of good vertices. Let

$$\mathcal{P}_N = \big\{\{x_0, \dots, x_k\} :\ k \ge 0,\ d(x_i, x_{i-1}) = 1\ \forall\ i = 1, \dots, k\big\}$$

be the set of finite nearest-neighbor paths on $\mathbb{H}_N$, including paths of length zero, which are just single vertices. Define the mapping $\varphi_N : \{\mathbb{H}_N, \mathcal{E}_N\} \to \{\mathbb{H}_N, \mathcal{P}_N\}$ in the following way. For $x \in \mathbb{H}_N$, let

$$\varphi_N(x) = \begin{cases} x, & \text{if } x \text{ is good;} \\ x^i, & \text{if } x \text{ and } x^j,\ j < i, \text{ are bad but } x^i \text{ is good;} \\ x, & \text{if } x \text{ is bad and has no good neighbor.} \end{cases}$$

By (2.3.2), $\mathbb{P}$-a.s. for N large enough the last option will not be used, and therefore φ_N maps all vertices to good vertices. In this case, for two neighboring vertices x, y, their good images $\varphi_N(x)$ and $\varphi_N(y)$ can either coincide, or be at distance 1, 2, or 3.

For an edge $e = \{x, y\} \in \mathcal{E}_N$, let $\varphi_N(e)$ be

- the 'path' $\{\varphi_N(x)\}$, if $\varphi_N(x)$ is good and $\varphi_N(x) = \varphi_N(y)$;
- the path $\{\varphi_N(x), \varphi_N(y)\}$, if both $\varphi_N(x)$ and $\varphi_N(y)$ are good and at distance 1;
- the path $\gamma^{ij}_{\varphi_N(x),\varphi_N(y)}$ with 'minimal' i, j such that all vertices of this path are good, if both $\varphi_N(x)$ and $\varphi_N(y)$ are good with distance 2 or 3 and such path exists;
- the path $\{x, y\}$ in any other case.

From (2.3.2) and (2.3.3) it follows that $\mathbb{P}$-a.s. for N large enough the last option does not occur and φ_N maps all edges to paths that contain only good vertices.

Finally, we extend φ_N to be a map that sends paths to paths. For $\gamma = \{x_0, \dots, x_n\} \in \mathcal{P}_N$ we define $\phi_N(\gamma)$ to be a concatenation of paths $\varphi_N(\{x_{i-1}, x_i\})$, $i = 1, \dots, n$, with possible loops erased by an arbitrary fixed loop-erasure algorithm. Note that ϕ_N can make paths shorter or longer, but by construction, for any path $\gamma \in \mathcal{P}_N$,

$$|\phi_N(\gamma)| \le 7|\gamma|. \tag{2.3.5}$$

We can now construct the random set of paths Γ that satisfies the properties of the lemma. We first define a certain canonical set of paths $\tilde{\Gamma}$, and then use the mapping ϕ_N to construct Γ from $\tilde{\Gamma}$.

For any pair of vertices $x \neq y \in \mathbb{H}_N$, let $\tilde{\gamma}_{xy}$ be the path from x to y obtained by consequently flipping the disagreeing coordinates, starting at coordinate 1. These paths are all of length smaller or equal to N, and the set $\tilde{\Gamma} = \{\tilde{\gamma}_{xy} : \ x \neq y \in \mathbb{H}_N\}$ has the property that any edge e is used by at most 2^{N-1} paths in $\tilde{\Gamma}$. Indeed, if $e = \{u, v\}$, then there is a unique i such that $u_i \neq v_i$. By construction, $e \in \tilde{\gamma}_{xy}$ if

$$\begin{aligned} x &= (x_1, \dots, x_{i-1}, u_i, u_{i+1}, \dots, u_N), \\ y &= (v_1, \dots, v_{i-1}, v_i, y_{i+1}, \dots, y_N). \end{aligned}$$

It follows that a total of $N-1$ coordinates of x and y are unknown, and so the number of possible pairs x, y for paths $\tilde{\gamma}_{xy}$ through e is bounded by 2^{N-1} (cf. [DS91, Example 2.2]).

For any pair $x \neq y \in \mathbb{H}_N$, let the path γ_{xy} in the set Γ be defined by

$$\gamma_{xy} = \begin{cases} \phi_N(\tilde{\gamma}_{xy}), & \text{if } x, y \text{ are good,} \\ \{x\} \circ \phi_N(\tilde{\gamma}_{xy}), & \text{if } x \text{ is bad and } y \text{ is good,} \\ \phi_N(\tilde{\gamma}_{xy}) \circ \{y\}, & \text{if } x \text{ is good and } y \text{ is bad,} \\ \{x\} \circ \phi_N(\tilde{\gamma}_{xy}) \circ \{y\}, & \text{if } x, y \text{ are bad,} \end{cases}$$

where '$\circ$' denotes the path concatenation.

It remains to check that this set of paths Γ indeed satisfies the required properties. First, by construction, Γ is complete, i.e. every path $\gamma_{xy} \in \Gamma$ connects x with y and is nearest-neighbor and self-avoiding. Further, by construction of φ_N and the properties (2.3.2) and (2.3.3), $\mathbb{P}$-a.s. for N large enough, all interior vertices of all $\gamma \in \Gamma$ are good, i.e. (i) is satisfied. Moreover, by (2.3.5) and the construction of the paths $\tilde{\gamma} \in \tilde{\Gamma}$, the paths $\gamma \in \Gamma$ have length at most $7N + 2$, hence (ii) is satisfied for $N \geq 2$. Finally, ϕ_N deforms the paths $\tilde{\gamma} \in \tilde{\Gamma}$ only locally, so that the number of paths in Γ passing through an edge e is bounded by the number of paths in $\tilde{\Gamma}$ passing through the ball of radius 4 around e. But this number is bounded by 2^{N-1} times the number of edges in that ball, which is bounded by N^K for some integer $K > 0$. This proves (iii) and thus finishes the proof of the lemma. □

We can now prove the spectral gap estimate.

Proof of Proposition 2.3.1. $\mathbb{P}$-a.s. for every N large enough we can find a complete set of paths Γ such that (i), (ii) and (iii) of Lemma 2.3.2 and (2.2.3) hold. By (ii), the expression in (2.3.1) over which the maximum is taken is bounded from above by

$$\frac{8N}{Z_N} \frac{1}{\tau_u \wedge \tau_v} \sum_{\gamma_{xy} \ni \{u,v\}} (\tau_x \wedge 1)(\tau_y \wedge 1). \tag{2.3.6}$$

We distinguish three cases for the position of the edge $\{u, v\}$ in a path γ_{xy}.

(1) If $\{u, v\}$ is an interior edge of γ_{xy}, then $\tau_u \wedge \tau_v$ is larger than $N^{-\beta C_0}$ by (i) of Lemma 2.3.2.
(2) If $\{u, v\}$ is at the end of the path γ_{xy}, say at $u = x$, and v is an interior vertex of γ_{xy}, then $\tau_x \wedge \tau_v$ is either larger than $N^{-\beta C_0}$, or it is equal to τ_x in which case it cancels with $\tau_x \wedge 1$. Indeed, if $\tau_x \wedge \tau_v$ was smaller than $N^{-\beta C_0}$ and equal to τ_v, then v would be a bad interior vertex of γ, which contradicts (i) of Lemma 2.3.2.
(3) If γ_{xy} only consists of the single edge $\{x, y\}$, then $\tau_x \wedge \tau_y$ is either larger than 1, or the term $\tau_x \wedge \tau_y$ cancels with the smaller one of $\tau_x \wedge 1$ and $\tau_y \wedge 1$.

It follows that for every edge $\{u, v\}$ the expression (2.3.6) is bounded from above by

$$\frac{8N}{Z_N} N^{\beta C_0} \#\{\text{paths through } e\}.$$

Since, by (iii) of Lemma 2.3.2, the number of paths is bounded by $N^K 2^{N-1}$, and, by (2.2.3), $Z_N \geq \kappa 2^N$, this completes the proof. □

In a next step we construct the mixing time T_{mix} of the fast chain Y. To this end, define the mixing scale

$$m_N = \frac{8}{\kappa} N^{K+3+\beta C_0}. \tag{2.3.7}$$

Then Proposition 2.3.1 reads $\lambda_N \geq 2N^2 m_N^{-1}$.

We assume that our probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is rich enough so that there exist infinitely many independent uniformly on $[0, 1]$ distributed random variables, independent of anything else. A randomized stopping time T is a positive random variable such that the event $\{T \leq t\}$ depends only on $\{Y_s : s \leq t\}$, the environment, and on the values of these additional random variables.

Proposition 2.3.3. *$\mathbb{P}$-a.s. for N large enough, there exists a randomized stopping time T_{mix} with values in $\{m_N, 2m_N, 3m_N, \dots\}$ such that T_{mix} is a strong stationary time for Y, that is for any (possibly random) $Y_0 \in \mathbb{H}_N$,*

(i) $P^\tau_{Y_0}[Y_{T_{\mathrm{mix}}} = y] = \nu_y$,
(ii) for any $k \geq 1$, $P^\tau_{Y_0}[T_{\mathrm{mix}} \geq k m_N] = e^{-(k-1)}$,
(iii) T_{mix} *and* $Y_{T_{\mathrm{mix}}}$ *are independent.*

Proof. This construction follows closely [MM15, Proposition 3.1], with only minor adaptations. Define the following distances from stationarity,

$$\begin{aligned} s(t) &= \min\{s \geq 0 :\ \forall x, y \in \mathbb{H}_N,\ P^\tau_x[Y_t = y] \geq (1-s)\nu(y)\}, \\ \bar{d}(t) &= \max_{x,y \in \mathbb{H}_N} \|P^\tau_x[Y_t \in \cdot\,] - P^\tau_y[Y_t \in \cdot\,]\|_{TV}, \end{aligned}$$

where $\|\cdot\|_{TV}$ denotes the total variation distance. Define the time

$$\mathcal{T} = \inf\{t \geq 0 :\ \bar{d}(t) \leq e^{-1}\}.$$

From [AF02, Lemmas 4.5, 4.6 and 4.23] we know that

$$\begin{aligned} \bar{d}(t) &\leq e^{-\lfloor t/\mathcal{T} \rfloor}, \\ s(2t) &\leq 1 - (1 - \bar{d}(t))^2, \\ \mathcal{T} &\leq \frac{1}{\lambda_Y}\left(1 + \frac{1}{2}\log\frac{1}{\nu^*}\right), \end{aligned} \tag{2.3.8}$$

where $\nu^* = \min_x \nu_x$. Since $\mathbb{P}[\tau_x \leq e^{-N^2}] \leq c e^{-c'N}$, by the Borel-Cantelli lemma, $\mathbb{P}$-a.s. for N large enough, $\log\frac{1}{\nu^*} \leq N^2$. Therefore, by Proposition 2.3.1 and (2.3.8), $\mathbb{P}$-a.s. for N large enough, $\mathcal{T} \leq \frac{1}{2}m_N$, $\bar{d}(\frac{1}{2}m_N) \leq e^{-1}$, and $s(m_N) \leq e^{-1}$, which means that for all $Y_0, y \in \mathbb{H}_N$,

$$P^\tau_{Y_0}[Y_{m_N} = y] \geq (1 - e^{-1})\nu_y.$$

We can now define the strong stationary time T_{mix} with values in $\{m_N, 2m_N, \dots\}$. Let $U_1, U_2, \dots$ be i.i.d. uniformly on $[0,1]$ distributed random variables, independent of anything else. Conditionally on $Y_0 = x$, $Y_{m_N} = y$, let $T_{\mathrm{mix}} = m_N$ if

$$U_1 \leq \frac{(1 - e^{-1})\nu_y}{P^\tau_x[Y_{m_N} = y]} \quad (\leq 1).$$

Otherwise, we define T_{mix} inductively: for every $k \in \mathbb{N}$, conditionally on $T_{\mathrm{mix}} > k m_N$, $Y_{k m_N} = z$ and $Y_{(k+1)m_N} = y$, let $T_{\mathrm{mix}} = (k+1)m_N$ if

$$U_{k+1} \leq \frac{(1 - e^{-1})\nu_y}{P^\tau_z[Y_{m_N} = y]} \quad (\leq 1).$$

By construction, we have for every $x \in \mathbb{H}_N$,

$$P^\tau_x[T_{\mathrm{mix}} = m_N \mid Y_{m_N} = y] = \frac{(1 - e^{-1})\nu_y}{P^\tau_x[Y_{m_N} = y]},$$

and thus

$$P^\tau_{Y_0}[T_{\mathrm{mix}} = m_N,\ Y_{m_N} = y \mid Y_0 = x] = (1 - e^{-1})\nu_y.$$

Similarly, we have

$$P^\tau_{Y_0}[T_{\text{mix}} = (k+1)m_N,\ Y_{(k+1)m_N} = y \mid T_{\text{mix}} > km_N,\ Y_{km_N} = x] = (1-e^{-1})\nu_y.$$

By induction over k, we obtain that for any $k \in \mathbb{N}$ and $y \in \mathbb{H}_N$,

$$P^\tau_{Y_0}[T_{\text{mix}} = km_N,\ Y_{km_N} = y] = e^{-(k-1)}(1-e^{-1})\nu_y,$$

which finishes the proof. □

For future reference we collect here two useful statements that follow directly from the construction of T_{mix}.

Lemma 2.3.4. *For every $t > 0$ and $x \in \mathbb{H}_N$ and every starting distribution ρ,*

$$\begin{aligned} &P^\tau_\rho[Y_t = x | T_{\text{mix}} < t] = \nu_x, \\ &\left|P^\tau_\rho[Y_t = x] - \nu_x\right| \le P^\tau_\rho[T_{\text{mix}} > t] = e^{-\lfloor t/m_N - 1\rfloor}. \end{aligned}$$

2.4. Bounds on mean hitting time and random scale

In this section we prove bounds on the mean hitting time $E^\tau_\nu[H_x]$ of deep traps $x \in \mathcal{D}_N$. As a corollary of the proof we will obtain a useful bound on the Green function in deep traps. The bounds on the mean hitting times will further imply bounds on the random scale R_N, which will imply the claim (2.1.7) of Theorem 2.1.1.

Proposition 2.4.1. *There exists $\delta \in (0, 1/6)$, such that $\mathbb{P}$-a.s. for N large enough,*

$$2^{N-N^{1-\delta}} \le E^\tau_\nu[H_x] \le 2^{N+N^{1-\delta}} \quad \text{for every } x \in \mathcal{D}_N.$$

The proof of Proposition 2.4.1 is split in two parts.

Proof of the upper bound. For the upper bound we use [AF02, Lemma 3.17] which states that

$$E^\tau_\nu[H_x] \le \frac{1-\nu_x}{\lambda_Y \nu_x}.$$

Since $\tau_x \ge 1$ for deep traps $x \in \mathcal{D}_N$, this is smaller than $\frac{Z_N}{\lambda_Y}$, which by Proposition 2.3.1 and (2.2.3) is bounded by $2^{N+N^{1-\delta}}$, $\mathbb{P}$-a.s. for N large enough. □

For the lower bound we will use a version of Proposition 3.2 of [ČTW11] which allows to bound the inverse of the mean hitting time $E^\tau_\nu[H_x]$ in terms of the effective conductance from x to a suitable set B. Recall the definition of the conductances c_{xy} from (2.2.4), and let $c_x = \sum_{y\sim x} c_{xy}$. Following the terminology of [LP14, Chapter 2], we define the effective conductance between a vertex x and a set B as

$$\mathcal{C}(x \to B) = P^\tau_x[H^+_x > H_B]c_x.$$

By Proposition 2.A.1, which is a generalization of [ČTW11, Proposition 3.2] to arbitrary continuous-time finite-state-space Markov chains,

$$\frac{1}{E^\tau_\nu[H_x]} \le \mathcal{C}(x \to B)\nu(B)^{-2}. \tag{2.4.1}$$

To apply this bound effectively, we should find a set B such that $\mathcal{C}(x \to B)$ is small and $\nu(B)$ close to 1. In the next lemma we construct such sets B for every $x \in \mathbb{H}_N$. For these sets we have some control on the conductances connecting B and B^c. Using standard network reduction techniques we can then give a bound on the effective conductance $\mathcal{C}(x \to B)$, which when plugged into (2.4.1) will imply the lower bound on $E_\nu^\tau[H_x]$.

Denote by $B(x,r) = \{y \in \mathbb{H}_N,\ d(x,y) \le r\}$ the ball of radius r around x, and by $\partial B(x,r) = \{y \in \mathbb{H}_N,\ d(x,y) = r\}$ the sphere of radius r.

Lemma 2.4.2. *For every $\delta \in (0, 1/6)$, $\mathbb{P}$-a.s. for N large enough, there exist radii $(\rho_x)_{x\in\mathbb{H}_N}$ satisfying $1 \le \rho_x \le N^{3\delta}$, such that for all $x \in \mathbb{H}_N$ and for all $y \in \partial B(x, \rho_x)$, $\tau_y \le 2^{\frac{1}{2}N^{1-\delta}}$.*

Proof. Fix $\delta \in (0, 1/6)$. We say that a sphere $\partial B(x,r)$ is good if $\tau_y \le 2^{\frac{1}{2}N^{1-\delta}}$ for all $y \in \partial B(x,r)$, otherwise we say that it is bad. Using the Gaussian tail approximation (2.2.2), we get that

$$\mathbb{P}\big[\tau_y > 2^{\frac{1}{2}N^{1-\delta}}\big] \le c e^{-\frac{\log^2 2}{8\beta^2} N^{1-2\delta}}.$$

The size of the sphere $\partial B(x,r)$ is bounded by N^r, hence the probability that the sphere $\partial B(x,r)$ is bad is bounded by

$$N^r \mathbb{P}\big[\tau_y > 2^{\frac{1}{2}N^{1-\delta}}\big] \le c \exp\Big\{r \log N - \frac{\log^2 2}{8\beta^2} N^{1-2\delta}\Big\}.$$

By independence of the τ_x, the probability that for one fixed x all the spheres $\partial B(x,r)$, $r = 1, \dots, N^{3\delta}$, are bad is bounded by

$$\begin{aligned}\prod_{r=1}^{N^{3\delta}} N^r \mathbb{P}\big[\tau_y > 2^{\frac{1}{2}N^{1-\delta}}\big] &\le \Big(N^{N^{3\delta}} \mathbb{P}\big[\tau_y > 2^{\frac{1}{2}N^{1-\delta}}\big]\Big)^{N^{3\delta}} \\ &\le \exp\Big\{N^{3\delta} \log c + N^{6\delta} \log N - \frac{\log^2 2}{8\beta} N^{1+\delta}\Big\}.\end{aligned}$$

Finally, by a union bound, the probability that among all 2^N vertices in $\mathbb{H}_N$ there is one for which all spheres $\partial B(x,r)$, $r = 1, \dots, N^{3\delta}$, are bad is bounded by

$$2^N \Big(N^{N^{3\delta}} \mathbb{P}\big[\tau_y > 2^{\frac{1}{2}N^{1-\delta}}\big]\Big)^{N^{3\delta}} \le \exp\Big\{N^{3\delta} \log c + N^{6\delta} \log N + N \log 2 - \frac{\log^2 2}{8\beta} N^{1+\delta}\Big\}.$$

Since $\delta < 1/6$ this decays faster than exponentially, and so by the Borel-Cantelli lemma the event occurs $\mathbb{P}$-a.s. only for finitely many N, i.e. $\mathbb{P}$-a.s. for N large enough we can find for every $x \in \mathbb{H}_N$ a radius $\rho_x \le N^{3\delta}$ such that the sphere $\partial B(x, \rho_x)$ is good. □

Proof of the lower bound of Proposition 2.4.1. For every $x \in \mathcal{D}_N$ we define the set $A_x = B(x, \rho_x)$ if the radius ρ_x from Lemma 2.4.2 exists, otherwise we take $A_x = \{x\}$. By Lemma 2.4.2 and (2.2.3), $\mathbb{P}$-a.s. for N large enough, for all $x \in \mathcal{D}_N$ all conductances $c_{yz} = (\tau_y \wedge \tau_z)/Z_N$ connecting A_x and A_x^c are smaller than $2^{\frac{1}{2}N^{1-\delta}}/(\kappa 2^N)$.

By the parallel law (cf. [LP14, Chapter 2.3]), the effective conductance between the boundaries of A_x and A_x^c is equal to the sum of all the conductances of edges connecting A_x and A_x^c, and so $\mathbb{P}$-a.s. for N large enough,

$$\mathcal{C}(\partial A_x \to \partial A_x^c) = \sum_{\substack{y \in \partial A_x \\ z \in \partial A_x^c}} c_{yz} \le \kappa^{-1} N^{\rho_x + 1} 2^{\frac{1}{2}N^{1-\delta}} 2^{-N}.$$

By Rayleigh's monotonicity principle (cf. [LP14, Chapter 2.4]), comparing the effective conductances from x to A_x^c before and after setting all the conductances inside A_x to infinity, it follows that

$$\mathcal{C}(x \to A_x^c) \le \mathcal{C}(\partial A_x \to \partial A_x^c) \le \kappa^{-1} N^{\rho_x+1} 2^{\frac{1}{2}N^{1-\delta}} 2^{-N}.$$

Since $\delta < 1/6$ and $\rho_x \le N^{3\delta}$, we have $N^{\rho_x+1} \le 2^{\frac{1}{2}N^{1-\delta}}$ for N large enough, and thus, $\mathbb{P}$-a.s. for N large enough,

$$\mathcal{C}(x \to A_x^c) \le c 2^{-N+N^{1-\delta}}. \tag{2.4.2}$$

Moreover, $\mathbb{P}$-a.s. for N large enough, as $\nu_y = (1 \wedge \tau_y)/Z_N \le 1/Z_N$, using (2.2.3) again,

$$\nu(A_x^c) = 1 - \nu(A_x) \ge 1 - Z_N^{-1}|A_x| \ge 1 - c 2^{-N} N^{N^{3\delta}} \xrightarrow{N\to\infty} 1. \tag{2.4.3}$$

Plugging (2.4.2) and (2.4.3) into (2.4.1) and readjusting δ to accommodate for constants easily yields the required lower bound $E_\nu^\tau[H_x] \ge 2^{N-N^{1-\delta}}$. This completes the proof. □

As a corollary we get a lower bound on $E_x^\tau[\ell_{H_{A_x^c}}(x)]$ for the deep traps $x \in \mathcal{D}_N$.

Corollary 2.4.3. *There are constants $\delta \in (0, 1/6)$ and $c > 0$, such that $\mathbb{P}$-a.s. for N large enough, for all $x \in \mathcal{D}_N$, under P_x^τ the local time of Y in x before leaving A_x, $\ell_{H_{A_x^c}}(x)$, stochastically dominates an exponential random variable with mean $c2^{-N^{1-\delta}}$. In particular, $\mathbb{P}$-a.s. for N large enough,*

$$E_x^\tau\big[\ell_{H_{A_x^c}}(x)\big] \ge c 2^{-N^{1-\delta}}.$$

Proof. The local time at x before hitting A_x^c is an exponential random variable with mean equal to

$$E_x^\tau\big[\#\{\text{visits to } x \text{ before } H_{A_x^c}\}\big] \cdot E_x^\tau[J_1].$$

The expected number of visits before leaving A_x is $P_x^\tau[H_x^+ > H_{A_x^c}]^{-1} = c_x \mathcal{C}(x \to A_x^c)^{-1}$. The mean duration of one visit to x is $E_x^\tau[J_1] = (\sum_{y\sim x} q_{xy})^{-1}$. For the deep traps we have $\tau_x > 1$, therefore $\sum_{y\sim x} q_{xy} = \sum_{y\sim x} c_{xy}/\nu_x = Z_N c_x$. It follows that the local time at x before hitting A_x^c is in fact an exponential random variable with mean $Z_N^{-1}\mathcal{C}(x \to A_x^c)^{-1}$. Using the bounds (2.4.2) and (2.2.3), the claim follows easily. □

As a next consequence we give bounds on the random scale R_N defined in (2.2.10). Note that this lemma also proves the statement (2.1.7) about the asymptotic behavior of R_N in Theorem 2.1.1.

Lemma 2.4.4. *For every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,*

$$2^{(\gamma-\varepsilon)N} \le R_N \le 2^{(\gamma+\varepsilon)N}.$$

Proof. By Proposition 2.3.3, T_{mix}/m_N is a geometric random variable with parameter e^{-1}, and thus $E_x^\tau[\ell_{T_{\text{mix}}}(x)^\alpha] \le E_x^\tau[{T_{\text{mix}}}^\alpha] \le c m_N^\alpha \le e^{\epsilon N}$ by (2.3.7), for every $\epsilon > 0$ and N large enough. Moreover, $|\mathcal{D}_N| \le c' 2^{(1-\gamma')N}$ by (2.2.8). Using the lower bound on $E_\nu^\tau[H_x]$ from Proposition 2.4.1, we obtain that for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,

$$R_N = 2^{(\gamma-\gamma')N} \left(\sum_{x\in\mathcal{D}_N} \frac{E_x^\tau[\ell_{T_{\text{mix}}}(x)^\alpha]}{E_\nu^\tau[H_x]} \right)^{-1} \ge 2^{(\gamma-\varepsilon)N}.$$

For the upper bound we need a lower bound on $E_x^\tau[\ell_{T_{\mathrm{mix}}}(x)^\alpha]$. Recall the sets A_x constructed in the proof of Proposition 2.4.1, and note that

$$E_x^\tau[\ell_{T_{\mathrm{mix}}}(x)^\alpha] \geq E_x^\tau\big[\mathbf{1}_{\{T_{\mathrm{mix}} \geq H_{A_x^c}\}} \ell_{H_{A_x^c}}(x)^\alpha\big]. \tag{2.4.4}$$

By Corollary 2.4.3, $\mathbb{P}$-a.s. for N large enough, the local time at x before hitting A_x^c stochastically dominates an exponential random variable with mean $c2^{-N^{1-\delta}}$, hence

$$P_x^\tau\big[\ell_{H_{A_x^c}}(x) \leq 2^{-2N^{1-\delta}}\big] \leq 1 - e^{-c2^{-N^{1-\delta}}} \leq c2^{-N^{1-\delta}}.$$

Moreover, for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,

$$P_x^\tau[T_{\mathrm{mix}} < H_{A_x^c}] \leq P_x^\tau[Y_{T_{\mathrm{mix}}} \in A_x] = \nu(A_x) \leq \kappa^{-1} 2^{-N} N^{N^{3\delta}} \leq 2^{-\varepsilon N}.$$

Using the last two observations in (2.4.4), $\mathbb{P}$-a.s. for N large enough,

$$\begin{aligned}
E_x^\tau[\ell_{T_{\mathrm{mix}}}(x)^\alpha] &\geq P_x^\tau\big[\{T_{\mathrm{mix}} \geq H_{A_x^c}\} \cap \{\ell_{H_{A_x^c}}(x) \geq 2^{-2N^{1-\delta}}\}\big] \left(2^{-2N^{1-\delta}}\right)^\alpha \\
&\geq 2^{-2\alpha N^{1-\delta}} \left(P_x^\tau\big[\ell_{H_{A_x^c}}(x) \geq 2^{-2N^{1-\delta}}\big] - P_x^\tau\big[\{\ell_{H_{A_x^c}}(x) \geq 2^{-2N^{1-\delta}}\} \cap \{T_{\mathrm{mix}} < H_{A_x^c}\}\big]\right) \\
&\geq 2^{-2\alpha N^{1-\delta}} \left(P_x^\tau\big[\ell_{H_{A_x^c}}(x) \geq 2^{-2N^{1-\delta}}\big] - P_x^\tau\big[T_{\mathrm{mix}} < H_{A_x^c}\big]\right) \\
&\geq 2^{-2\alpha N^{1-\delta}} \left((1 - c'2^{-N^{1-\delta}}) - 2^{-\varepsilon N}\right) \\
&\geq 2^{-\varepsilon N}.
\end{aligned}$$

Combining this with $|\mathcal{D}_N| \geq c2^{(1-\gamma')N}$ by (2.2.8) and the upper bound on $E_\nu^\tau[H_x]$ from Proposition 2.4.1, we obtain the required upper bound on R_N. □

2.5. Concentration of the local time functional

In this section we prove the concentration of the local time functional that appears in the computation of the quasi-annealed Laplace transform of the clock process on the deep traps, as explained in the introduction (cf. (2.1.9)). We denote this functional by

$$L_N(t) = 2^{(\gamma'-\gamma)N} \sum_{x \in \mathcal{D}_N} \ell_{tR_N}(x)^\alpha.$$

So far we had no restriction on the choice of γ' other than $1/2 < \gamma' < \gamma$, see (2.2.6). We now make an explicit choice as follows. Let $\varepsilon_0 = \frac{1}{2}\big((1-\gamma) \wedge (\gamma - \frac{1}{2})\big)$, and define $\gamma' = \gamma - \varepsilon_0$, such that in particular

$$1 - \gamma \geq 2\varepsilon_0, \tag{2.5.1}$$
$$\gamma - \gamma' = \varepsilon_0. \tag{2.5.2}$$

The main result of this section is the following proposition.

Proposition 2.5.1. *For every fixed $t \geq 0$, $\mathbb{P}$-a.s. for N large enough,*

$$P_\nu^\tau\Big[|L_N(t) - t| \geq 2^{-\frac{1}{5}\varepsilon_0 N}\Big] \leq c2^{-\frac{1}{10}\varepsilon_0 N}.$$

Proof. We approximate $L_N(t)$ by the sum of essentially independent random variables as follows. Let $K = \lfloor 2^{\varepsilon_0 N} \rfloor$. For a fixed $t > 0$, define

$$t_k = \frac{tR_N}{K} k, \qquad k = 0, \dots, K.$$

Recall the notation (2.2.5). For every $x \in \mathcal{D}_N$ and $k = 1, \dots, K$, define $H_x^k = t_{k-1} + H_x \circ \theta_{t_{k-1}}$ to be the time of the first visit to x after t_{k-1}, and set

$$\ell_{t,x}^k = \left(\int_{H_x^k \wedge (t_k - 2N^2 m_N)}^{(H_x^k + N^2 m_N) \wedge (t_k - N^2 m_N)} \mathbf{1}_{\{Y_s = x\}} ds \right)^\alpha .$$

The random variable $\ell_{t,x}^k$ gives 'roughly' the α-th power of the time that Y spends in x between t_{k-1} and $t_k - N^2 m_N$, with some suitable truncations. Let further

$$U_N^k(t) = 2^{(\gamma' - \gamma)N} \sum_{x \in \mathcal{D}_N} \ell_{t,x}^k .$$

The next lemma, which we prove later, shows that the sum of the $U_N^k(t)$'s is a good approximation for $L_N(t)$.

Lemma 2.5.2. *For every $t > 0$, $\mathbb{P}$-a.s. for N large enough,*

$$P_\nu^\tau \left[L_N(t) \neq \sum_{k=1}^K U_N^k(t) \right] \leq c 2^{-\frac{1}{2}\varepsilon_0 N}.$$

With Lemma 2.5.2, the proof of the proposition reduces to understanding of the approximating sum $\sum_{k=1}^K U_N^k(t)$. We will compute its expectation and variance under P_ν^τ. In particular, we will show that there is $c < \infty$ such that for every $t > 0$,

$$\left| E_\nu^\tau \Big[\sum_{k=1}^K U_N^k(t) \Big] - t \right| \leq c 2^{-2\varepsilon_0 N}, \qquad \mathbb{P}\text{-a.s. as } N \to \infty, \tag{2.5.3}$$

and

$$\mathrm{Var}_\nu^\tau \left(\sum_{k=1}^K U_N^k(t) \right) \leq c 2^{-\frac{1}{2}\varepsilon_0 N}, \qquad \mathbb{P}\text{-a.s. as } N \to \infty. \tag{2.5.4}$$

The statement of the proposition then follows from Lemma 2.5.2, (2.5.3) and (2.5.4) by routine application of the Chebyshev inequality. Indeed, $\mathbb{P}$-a.s. for N large enough,

$$\begin{aligned}
&P_\nu^\tau \left[|L_N(t) - t| \geq 2^{-\frac{1}{5}\varepsilon_0 N} \right] \\
&\quad \leq P_\nu^\tau \left[L_N(t) \neq \sum_{k=1}^K U_N^k(t) \right] + P_\nu^\tau \left[\left| \sum_{k=1}^K U_N^k(t) - E_\nu^\tau \left[\sum_{k=1}^K U_N^k(t) \right] \right| \geq 2 \cdot 2^{-\frac{1}{5}\varepsilon_0 N} \right] \\
&\quad \leq c 2^{-\frac{1}{2}\varepsilon_0 N} + c' 2^{-\frac{1}{10}\varepsilon_0 N} \leq c'' 2^{-\frac{1}{10}\varepsilon_0 N},
\end{aligned}$$

which is the claim of the proposition.

We proceed by computing the expectation (2.5.3). We will need two lemmas which we show later. The first lemma estimates the probability that a deep trap is visited by the process Y.

Lemma 2.5.3. *For every t_N such that $1 \le t_N \le 2^N$, for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough, for all $x \in \mathcal{D}_N$,*

$$P_\nu^\tau[H_x \le t_N] = \frac{t_N}{E_\nu^\tau[H_x]} + O\big(t_N^2 2^{2(\varepsilon-1)N}\big) + O\big(2^{(\varepsilon-1)N}\big) \le c t_N 2^{(\varepsilon-1)N}.$$

The second lemma then gives the expected contribution of a single $\ell_{t,x}^k$ to $\sum_{k=1}^K U_N^k(t)$.

Lemma 2.5.4. *For every fixed $t > 0$, $k = 1, \dots, K$ and $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough, for all $x \in \mathcal{D}_N$,*

$$E_\nu^\tau\big[\ell_{t,x}^k\big] = \frac{tR_N}{KE_\nu^\tau[H_x]} E_x^\tau\big[\ell_{T_{\mathrm{mix}}}(x)^\alpha\big] + O\big(2^{(2\gamma+3\varepsilon-2\varepsilon_0-2)N}\big).$$

With Lemma 2.5.4 it is easy to compute the expectation (2.5.3). Using that $|\mathcal{D}_N| \le c2^{(1-\gamma')N}$ by (2.2.8), and the definition (2.2.10) of R_N, for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,

$$\begin{aligned} E_\nu^\tau\bigg[\sum_{k=1}^K U_N^k(t)\bigg] &= 2^{(\gamma'-\gamma)N} \sum_{x\in\mathcal{D}_N} \sum_{k=1}^K \bigg(\frac{tR_N}{KE_\nu^\tau[H_x]} E_x^\tau[\ell_{T_{\mathrm{mix}}}(x)^\alpha] + O\big(2^{(2\gamma+3\varepsilon-2\varepsilon_0-2)N}\big)\bigg) \\ &= t + O\big(2^{(\gamma'-\gamma)N} 2^{(1-\gamma')N} 2^{(2\gamma-2+3\varepsilon-\varepsilon_0)N}\big) \\ &= t + O\big(2^{(\gamma-1+3\varepsilon-\varepsilon_0)N}\big). \end{aligned}$$

Choosing $\varepsilon < \varepsilon_0/3$ and recalling (2.5.1) implies (2.5.3).

Next, we estimate the variance (2.5.4). Since ν is the stationary measure for Y, the random variables $U_N^k(t)$, $k = 1, \dots, K$, are identically distributed under P_ν^τ. Hence

$$\mathrm{Var}_\nu^\tau\bigg(\sum_{k=1}^K U_N^k(t)\bigg) = K\,\mathrm{Var}_\nu^\tau\big(U_N^1(t)\big) + 2\sum_{1\le k<j\le K} \mathrm{Cov}_\nu^\tau\big(U_N^k(t), U_N^j(t)\big). \tag{2.5.5}$$

The covariances can be neglected easily. Indeed, since by definition $U_N^k(t)$ depends on the trajectory of Y between times t_{k-1} and $t_k - N^2 m_N$ only, we can use the Markov property at the later time to write

$$\mathrm{Cov}_\nu^\tau\big(U_N^k(t), U_N^j(t)\big) = E_\nu^\tau\big[\big(U_N^k(t) - E_\nu^\tau U_N^k(t)\big) E^\tau[U_N^j(t) - E_\nu^\tau U_N^j(t) | Y_{t_k - N^2 m_N}]\big]. \tag{2.5.6}$$

By Lemma 2.3.4, $\big|P^\tau[Y_{t_k} = y | Y_{t_k - N^2 m_N}] - \nu_y\big| \le e^{-cN^2}$. Using in addition that $U_j^N \le e^{c'N}$ for some sufficiently large c', we see that the inner expectation satisfies

$$\big|E^\tau[U_N^j(t) - E_\nu^\tau U_N^j(t) | Y_{t_k - N^2 m_N}]\big| \le e^{-cN^2/2}.$$

Inserting this inequality back to (2.5.6) and summing over $k < j$ then implies that the second term in (2.5.5) is $O(e^{-cN^2})$ and thus can be neglected when proving (2.5.4).

To control the variance of $U_N^1(t)$ in (2.5.5), it is enough to bound its second moment, which is

$$E_\nu^\tau\big[U_N^1(t)^2\big] = 2^{2(\gamma'-\gamma)N} \bigg(\sum_{x\in\mathcal{D}_N} E_\nu^\tau\big[(\ell_{t,x}^1)^2\big] + \sum_{x\ne y\in\mathcal{D}_N} E_\nu^\tau[\ell_{t,x}^1 \ell_{t,y}^1] \bigg).$$

Since, by definition, $\ell^1_{t,x} \le N^2 m_N$ and $\ell^1_{t,x} \neq 0$ implies $H_x \le tR_N/K$,

$$E^\tau_\nu\big[U^1_N(t)^2\big] \le 2^{2(\gamma'-\gamma)N} N^{4\alpha} m_N^{2\alpha} \bigg(\sum_{x\in\mathcal{D}_N} P^\tau_\nu\bigg[H_x \le \frac{tR_N}{K}\bigg] + \sum_{x\neq y\in\mathcal{D}_N} P^\tau_\nu\bigg[H_x, H_y \le \frac{tR_N}{K}\bigg]\bigg). \tag{2.5.7}$$

By Lemma 2.5.3 and Lemma 2.4.4, for every $\varepsilon > 0$, $\mathbb{P}$-a.s. as $N \to \infty$,

$$P^\tau_\nu\bigg[H_x \le \frac{tR_N}{K}\bigg] \le c2^{(\gamma-1+\varepsilon-\varepsilon_0)N}. \tag{2.5.8}$$

Moreover, by (2.2.8), $|\mathcal{D}_N| \le c2^{(1-\gamma')N}$, and by (2.3.7), $N^{4\alpha} m_N^{2\alpha} \le 2^{\varepsilon N}$, for every $\varepsilon > 0$ and N large enough. It follows that the contribution of the first sum in (2.5.7) to the variance, including the prefactor $K = 2^{\varepsilon_0 N}$ from (2.5.5), can be bounded by

$$c2^{(2(\gamma'-\gamma)+1-\gamma'+\gamma-1+2\varepsilon)N} = c2^{(\gamma'-\gamma+2\varepsilon)N}.$$

By (2.5.2), $\gamma'-\gamma+2\varepsilon \le -\varepsilon_0 + 2\varepsilon < -\frac{1}{2}\varepsilon_0$ for $\varepsilon < \varepsilon_0/4$, and hence this contribution is smaller than $c2^{-\frac{1}{2}\varepsilon_0 N}$ as required for (2.5.4).

For the second summation in (2.5.7) we write

$$P^\tau_\nu\bigg[H_x, H_y \le \frac{tR_N}{K}\bigg] \le P^\tau_\nu\bigg[H_x < H_y \le \frac{tR_N}{K}\bigg] + P^\tau_\nu\bigg[H_y < H_x \le \frac{tR_N}{K}\bigg].$$

By the Markov property, each of these two probabilities can be bounded by

$$\begin{aligned} P^\tau_\nu\bigg[H_x < H_y \le \frac{tR_N}{K}\bigg] &= \int_0^{\frac{tR_N}{K}} P^\tau_\nu[H_x \in du] P^\tau_x\bigg[H_y < \frac{tR_N}{K} - u\bigg] \\ &\le \int_0^{\frac{tR_N}{K}} P^\tau_\nu[H_x \in du] \bigg(P^\tau_x[H_y \le T_{\mathrm{mix}}] + P^\tau_\nu\bigg[H_y \le \frac{tR_N}{K}\bigg]\bigg) \\ &\le P^\tau_\nu\bigg[H_x \le \frac{tR_N}{K}\bigg] \bigg(P^\tau_x[H_y \le T_{\mathrm{mix}}] + P^\tau_\nu\bigg[H_y \le \frac{tR_N}{K}\bigg]\bigg). \end{aligned}$$

Using (2.5.8) and (2.2.8) again, the second sum in (2.5.7) is bounded by

$$c2^{(\gamma-1+\varepsilon-\varepsilon_0)N}\Big(2^{2(1-\gamma')N} 2^{(\gamma-1+\varepsilon-\varepsilon_0)N} + \sum_{x\neq y\in\mathcal{D}_N} P^\tau_x[H_y \le T_{\mathrm{mix}}]\Big). \tag{2.5.9}$$

The first term in the parentheses of (2.5.9) together with the prefactors K from (2.5.5) and $2^{2(\gamma'-\gamma)N} N^{4\alpha} m_N^{2\alpha} \le 2^{(2(\gamma'-\gamma)+\varepsilon)N}$ from (2.5.7), contributes to the variance by at most

$$c2^{(\varepsilon_0+2(\gamma'-\gamma)+\varepsilon+2(1-\gamma')+2(\gamma-1+\varepsilon-\varepsilon_0))N} = c2^{(3\varepsilon-\varepsilon_0)N} \le c2^{-\frac{1}{2}\varepsilon_0 N}$$

if ε is small enough, as required by (2.5.4).

For the second term in the parentheses of (2.5.9) we need the following lemma whose proof is again postponed.

Lemma 2.5.5. *Let $\mathcal{W}_t^x = \sum_{y\in\mathcal{D}_N, y\neq x} \mathbf{1}_{\{H_y\leq t\}}$. Then for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough, for every $x \in \mathcal{D}_N$,*

$$E_x^\tau[\mathcal{W}_{T_{\mathrm{mix}}}^x] \leq 2^{\varepsilon N}.$$

Using Lemma 2.5.5, and including all the prefactors as before, the contribution of the second term in (2.5.9) to the variance (2.5.5) is bounded by

$$c2^{(\varepsilon_0+2(\gamma'-\gamma)+\epsilon+1-\gamma'+\gamma-1+\varepsilon-\varepsilon_0+\varepsilon)N} = c\,2^{(\gamma'-\gamma+3\varepsilon)N} \leq 2^{-\frac{1}{2}\varepsilon_0 N},$$

where for the last inequality we used (2.5.2) again, and choose ε small enough. This completes the proof of (2.5.4) and thus of the proposition. □

We proceed by proving the lemmas used in the above proof.

Proof of Lemma 2.5.3. By [AB92, Theorem 1] the hitting time H_x is approximately exponential in the sense that

$$\left|P_\nu^\tau[H_x > t] - e^{-\frac{t}{E_\nu^\tau[H_x]}}\right| \leq \frac{1}{\lambda_Y E_\nu^\tau[H_x]}.$$

Hence, using Propositions 2.3.1 and 2.4.1 to bound λ_Y and $E_\nu^\tau[H_x]$ respectively, we have for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,

$$\begin{aligned} P_\nu^\tau[H_x \leq t_N] &= (1 - e^{-\frac{t_N}{E_\nu^\tau[H_x]}}) + O\big(2^{(\varepsilon-1)N}\big) \\ &= \frac{t_N}{E_\nu^\tau[H_x]} + O\big(t_N^2 2^{2(\varepsilon-1)N}\big) + O\big(2^{(\varepsilon-1)N}\big). \end{aligned}$$

Finally, if $1 \leq t_N \leq 2^N$ this is bounded by $ct_N 2^{(\varepsilon-1)N}$, which proves the lemma. □

Proof of Lemma 2.5.4. By the strong Markov property and the definition of $\ell_{t,x}^k$,

$$\begin{aligned} E_\nu^\tau\big[\ell_{t,x}^k\big] &\geq P_\nu^\tau\big[H_x \in [t_{k-1}, t_k - 2N^2 m_N]\big] E_x^\tau\big[\ell_{N^2 m_N}(x)^\alpha\big], \\ E_\nu^\tau\big[\ell_{t,x}^k\big] &\leq P_\nu^\tau\big[H_x \in [t_{k-1}, t_k - N^2 m_N]\big] E_x^\tau\big[\ell_{N^2 m_N}(x)^\alpha\big]. \end{aligned} \tag{2.5.10}$$

We will now give approximations of the expressions appearing in (2.5.10).

Observe that for every $s, t > 0$,

$$\ell_t(x)^\alpha \leq \ell_s(x)^\alpha + (\ell_t(x) - \ell_s(x))^\alpha.$$

Using this inequality with $t = N^2 m_N$ and $s = T_{\mathrm{mix}}$ and applying the strong Markov property at T_{mix}, observing that $Y_{T_{\mathrm{mix}}}$ is ν-distributed,

$$E_x^\tau\big[\ell_{N^2 m_N}(x)^\alpha\big] \leq E_x^\tau\big[\ell_{T_{\mathrm{mix}}}(x)^\alpha\big] + E_\nu^\tau\big[\ell_{N^2 m_N}(x)^\alpha\big].$$

By Lemma 2.5.3, using also that by (2.3.7), $\ell_{N^2 m_N}(x)^\alpha \leq N^{2\alpha} m_N^\alpha \leq 2^{\varepsilon N}$ for every $\varepsilon > 0$ and N large enough,

$$E_\nu^\tau\big[\ell_{N^2 m_N}(x)^\alpha\big] \leq P_\nu^\tau\big[H_x \leq N^2 m_N] 2^{\varepsilon N} \leq c2^{(3\varepsilon-1)N}.$$

Hence we obtain the upper bound

$$E_x^\tau\big[\ell_{N^2 m_N}(x)^\alpha\big] \leq E_x^\tau\big[\ell_{T_{\mathrm{mix}}}(x)^\alpha\big] + c2^{(3\varepsilon-1)N}. \tag{2.5.11}$$

For a matching lower bound, note that

$$E_x^\tau\big[\ell_{N^2m_N}(x)^\alpha\big] \geq E_x^\tau\big[\ell_{T_{\rm mix}}(x)^\alpha \mathbf{1}_{\{T_{\rm mix}\leq N^2m_N\}}\big].$$

But from Proposition 2.3.3 it follows that

$$E_x^\tau[\ell_{T_{\rm mix}}(x)^\alpha \mathbf{1}_{\{T_{\rm mix}>N^2m_N\}}] \leq E_x^\tau[{T_{\rm mix}}^\alpha \mathbf{1}_{\{T_{\rm mix}>N^2m_N\}}] \leq \sum_{k=N^2}^\infty (km_N)^\alpha e^{-k} \leq ce^{-c'N^2},$$

so that

$$E_x^\tau\big[\ell_{N^2m_N}(x)^\alpha\big] \geq E_x^\tau\big[\ell_{T_{\rm mix}}(x)^\alpha\big] - ce^{-cN^2}. \tag{2.5.12}$$

Combining (2.5.11) and (2.5.12), we obtain

$$E_x^\tau\big[\ell_{N^2m_N}(x)^\alpha\big] = E_x^\tau\big[\ell_{T_{\rm mix}}(x)^\alpha\big] + O\big(2^{(3\varepsilon-1)N}\big). \tag{2.5.13}$$

Note also that by (2.3.7), for every $\varepsilon > 0$ and N large enough,

$$E_x^\tau[\ell_{T_{\rm mix}}(x)^\alpha] \leq E_x^\tau[{T_{\rm mix}}^\alpha] \leq cm_N^\alpha \leq 2^{\varepsilon N}. \tag{2.5.14}$$

To approximate the probabilities in (2.5.10), we apply Lemma 2.5.3 for $t_N = t_{k-1}$ and $t_N = t_k - iN^2m_N$, for a fixed $t > 0$ and $i = 1, 2$. Using Lemma 2.4.4 to bound R_N and Proposition 2.4.1 to bound $E_\nu^\tau[H_x]$, for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough, for both $i = 1, 2$,

$$P_\nu^\tau\big[H_x \in [t_{k-1}, t_k - iN^2m_N]\big] = \frac{tR_N}{KE_\nu^\tau[H_x]} + O\big(2^{2(\gamma+\varepsilon-\varepsilon_0-1)N}\big) = O\big(2^{(\gamma+\varepsilon-\varepsilon_0-1)N}\big). \tag{2.5.15}$$

Inserting both (2.5.15) and (2.5.13) in (2.5.10), and using (2.5.14), for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,

$$E_\nu^\tau\big[\ell_{t,x}^k\big] = \frac{tR_N}{KE_\nu^\tau[H_x]} E_x^\tau\big[\ell_{T_{\rm mix}}(x)^\alpha\big] + O\big(2^{(2\gamma+3\varepsilon-2\varepsilon_0-2)N}\big).$$

This proves the lemma. □

Proof of Lemma 2.5.2. Note first that

$$\Big\{L_N(t) \neq \sum_{k=1}^K U_N^k(t)\Big\} \subseteq \Big\{\exists x \in \mathcal{D}_N :\ \ell_{tR_N}(x)^\alpha \neq \sum_{k=1}^K \ell_{t,x}^k\Big\}.$$

To control the probability of this event, we introduce some more notation. Set $H_x^{(0)} = 0$, $H_x^{(1)} = H_x$, and for $k \geq 2$ define the time of the 'k-th visit after mixing' inductively as

$$H_x^{(k)} = \inf\{t > T_{\rm mix} \circ \theta_{H_x^{(k-1)}} + H_x^{(k-1)} :\ Y_t = x\}.$$

Let $\mathcal{N}_t^x = \min\{k \geq 0,\ H_x^{(k)} \leq t\}$ be the number of 'visits after mixing' to x before time t. Finally, let $I_k = [t_k - 2N^2m_N, t_k]$. Then

$$\begin{aligned} P_\nu^\tau\Big[\exists x \in \mathcal{D}_N :\ \ell_{tR_N}(x)^\alpha \neq \sum_{k=1}^K \ell_{t,x}^k\Big] \leq{}& P_\nu^\tau\Big[Y_s \in \mathcal{D}_N \text{ for some } s \in \bigcup_{k=1}^K I_k\Big] \\ &+ P_\nu^\tau\big[\exists x \in \mathcal{D}_N :\ \mathcal{N}_{tR_N}^x \geq 2\big] \\ &+ P_\nu^\tau\big[\exists x \in \mathcal{D}_N :\ T_{\rm mix} \circ \theta_{H_x} > N^2m_N\big]. \end{aligned} \tag{2.5.16}$$

We show that each of the three terms on the right-hand side is smaller than $c2^{-\frac{1}{2}\varepsilon_0 N}$, which will prove the lemma.

For the first term in (2.5.16), using the stationarity of ν and the Markov property

$$P_\nu^\tau\Big[Y_s \in \mathcal{D}_N \text{ for some } s \in \bigcup_{k=1}^{K} I_k\Big] \le K \sum_{x\in\mathcal{D}_N} P_\nu^\tau\big[H_x \le 2N^2 m_N\big]. \tag{2.5.17}$$

By Lemma 2.5.3, $\mathbb{P}$-a.s. for all $x \in \mathcal{D}_N$, for $\varepsilon > 0$ small and N large enough,

$$P_\nu^\tau[H_x \le 2N^2 m_N] \le 2^{(\varepsilon-1)N}.$$

Since $|\mathcal{D}_N| \le c2^{(1-\gamma')N}$ by (2.2.8), the right hand side of (2.5.17) is bounded by $c2^{(\varepsilon_0+\varepsilon-\gamma')N}$. Since $\gamma' > 1/2$ and by definition $\varepsilon_0 \le 1/4$, when ε is small enough this is smaller than $c2^{-\frac{1}{2}\varepsilon_0 N}$ as required.

For the second term in (2.5.16), by Lemma 2.5.3 and the strong Markov property at T_{mix}, for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,

$$P_\nu^\tau[H_x^{(2)} \le tR_N] \le P_\nu^\tau[H_x \le tR_N]^2 \le c2^{2(\gamma-1+\varepsilon)N}.$$

Together with (2.2.8) to bound $|\mathcal{D}_N|$, and using (2.5.1) and (2.5.2), $\mathbb{P}$-a.s. for N large enough,

$$\begin{aligned} P_\nu^\tau\big[\exists x \in \mathcal{D}_N :\ \mathcal{N}_{tR_N}^x \ge 2\big] &\le c2^{(1-\gamma')N} 2^{2(\gamma-1+\varepsilon)N} \\ &= c2^{(\gamma-\gamma')N+(\gamma-1)N+\varepsilon N} \\ &\le c2^{(-\varepsilon_0+\varepsilon)N} \le c2^{-\frac{1}{2}\varepsilon_0 N} \end{aligned}$$

as required.

Finally we give a bound on the third term in (2.5.16). By Proposition 2.3.3, $P_x^\tau[T_{\text{mix}} > N^2 m_N] \le e^{-cN^2}$. Thus, with (2.2.8) to bound $|\mathcal{D}_N|$, $\mathbb{P}$-a.s. for N large enough,

$$\begin{aligned} P_\nu^\tau\big[\exists x \in \mathcal{D}_N :\ T_{\text{mix}} \circ \theta_{H_x} > N^2 m_N\big] &\le c2^{(1-\gamma')N} P_x^\tau[T_{\text{mix}} > N^2 m_N] \\ &\le c'2^{-\frac{1}{2}\varepsilon_0 N}. \end{aligned}$$

Together with the previous estimates, this implies that the right-hand side of (2.5.16) is bounded by $c2^{-\frac{1}{2}\varepsilon_0 N}$, and concludes the proof of the lemma. □

Proof of Lemma 2.5.5. Let $\mathcal{H}_0 = 0$ and define recursively for $i \ge 1$

$$\mathcal{H}_i = \inf\{t \ge \mathcal{H}_{i-1} : Y_t \in \mathcal{D}_N \setminus \{Y_{\mathcal{H}_{i-1}}\}\}.$$

By (2.2.9), $\mathbb{P}$-a.s. for N large enough, the vertices in $\mathcal{D}_N$ are at least distance δN from each other. In particular the balls $A_x = B(x, \rho_x)$, $x \in \mathcal{D}_N$, constructed in Lemma 2.4.2 are disjoint. Hence, when on $y \in \mathcal{D}_N$, the random walk Y should first leave A_y in order to visit $\mathcal{D}_N \setminus \{y\}$. The strong Markov property and Corollary 2.4.3 then imply that $\mathcal{H}_i$ stochastically dominates a Gamma random variable with parameters i and $\mu := c2^{N^{1-\delta}}$.

If $\mathcal{W}_t^x \ge i$, then $\mathcal{H}_i \le t$. Hence, for $t \ge \mu$,

$$E_x^\tau\big[\mathcal{W}_t^x\big] = \sum_{i\ge1} P_x^\tau[\mathcal{W}_t^x \ge i] \le \sum_{i\ge1} P_x^\tau[\mathcal{H}_i \le t] \le \sum_{i\ge1} \int_0^t \mu^i u^{i-1} e^{-\mu u} \Gamma(i)^{-1} du = \mu t.$$

It follows that

$$E_x^\tau\big[\mathcal{W}^x_{T_{\mathrm{mix}}}\big] \le E_x^\tau\big[\mathcal{W}^x_{N^2 m_N}\big] + |\mathcal{D}_N| P_x^\tau[T_{\mathrm{mix}} \ge N^2 m_N] \le \mu N^2 m_N + c2^{(\gamma'-1)N} e^{-cN^2} \le 2^{\varepsilon N}$$

by (2.2.8), (2.3.7) and Proposition 2.3.3. This completes the proof. □

For later applications, we state two further consequences of the proof of Lemma 2.5.2.

Lemma 2.5.6. $\mathbb{P}$*-a.s. for* N *large enough,*

$$P_\nu^\tau\big[\exists x \in \mathcal{D}_N :\ \ell_{tR_N}(x) > N^2 m_N\big] \le c2^{-\frac{1}{2}\varepsilon_0 N},$$

and

$$P_\nu^\tau\Big[\big|\{x \in \mathcal{D}_N :\ H_x \le tR_N\}\big| \ge 2^{\frac{3}{2}\varepsilon_0 N}\Big] \le c2^{-\frac{1}{4}\varepsilon_0 N}.$$

Proof. The first claim follows directly from the bounds on the second and third term on the right hand side of (2.5.16) in the proof of Lemma 2.5.2, since the local time in a vertex that is only 'visited once after mixing' is bounded by $T_{\mathrm{mix}} \circ \theta_{H_x}$.

The second assertion can be seen in the following way. Using Lemma 2.5.3 to bound the probability of a single vertex $x \in \mathcal{D}_N$ to be visited before time tR_N and (2.2.8) to bound the size of $\mathcal{D}_N$, for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough,

$$E_\nu^\tau\big[|\{x \in \mathcal{D}_N :\ H_x \le tR_N\}|\big] \le c2^{(1-\gamma')N} 2^{(\gamma-1+\varepsilon)N} \le c2^{(\gamma-\gamma'+\varepsilon)N}.$$

By (2.5.2) this is equal to $c2^{(\varepsilon_0+\varepsilon)N}$, so choosing $\varepsilon < \varepsilon_0/4$ this is smaller than $c2^{\frac{5}{4}\varepsilon_0 N}$. Then by the Markov inequality the probability that there are more than $2^{\frac{3}{2}\varepsilon_0 N}$ vertices visited is smaller than $c2^{-\frac{1}{4}\varepsilon_0 N}$. □

2.6. Clock process of the deep traps

This section contains the main steps leading to the proof of Theorem 2.1.1. Recall from (2.1.8) that the *clock process of deep traps* is given by

$$S_{\mathcal{D}}(t) = \int_0^t (1 \vee \tau_{Y_s}) \mathbf{1}_{\{Y_s \in \mathcal{D}_N\}} ds = \int_0^t \tau_{Y_s} \mathbf{1}_{\{Y_s \in \mathcal{D}_N\}} ds.$$

We now show that $S_{\mathcal{D}}$ converges to a stable process.

Proposition 2.6.1. *Under the assumptions of Theorem 2.1.1, the rescaled clock processes of the deep traps,* $g_N^{-1} S_{\mathcal{D}}(tR_N)$*, converge in* $\mathbb{P}$*-probability as* $N \to \infty$*, in* P_ν^τ *distribution on the space* $D([0,T],\mathbb{R})$ *equipped with the Skorohod* M_1*-topology, to an* α*-stable subordinator* V_α*.*

The proof of Proposition 2.6.1 consists of three steps. In a first step we show convergence in distribution of one-dimensional marginals by showing that the Laplace transform of one-dimensional marginals converges. This step contains, to some extent, the principal insight of this paper and is split in two parts: We first show the quasi-annealed convergence mentioned in the introduction, which is then strengthened to convergence in probability with respect to the environment. The second and third step of the proof of Proposition 2.6.1 are rather standard and deal with the joint convergence of increments and the tightness.

2.6.1. Quasi-annealed convergence

We establish here the connection between the Laplace transform of the clock process of deep traps and the local time functional L_N studied in Section 2.5. The key observation is that the depths of the deep traps are in some sense independent of the fast chain Y, and can be thus averaged out easily.

To formalize this, we introduce a two-step procedure to sample the environment τ. Let $\xi = (\xi_x)_{x\in\mathbb{H}_N}$ be i.i.d. Bernoulli random variables such that, cf. (2.2.7),

$$\mathbb{P}[\xi_x = 1] = 1 - \mathbb{P}[\xi_k = 0] = \mathbb{P}[x \in \mathcal{D}_N] = 2^{-\gamma' N}(1 + o(1)).$$

Further, let $\overline{E} = (\overline{E}_x)_{x\in\mathbb{H}_N}$ be i.i.d. standard Gaussian random variables conditioned to be larger than $\frac{1}{\beta\sqrt{N}} g'_N$, and $\underline{E} = (\underline{E}_x)_{x\in\mathbb{H}_N}$ i.i.d. standard Gaussian random variables conditioned to be smaller than $\frac{1}{\beta\sqrt{N}} g'_N$. The collections ξ, $\overline{E}$ and $\underline{E}$ are mutually independent. The Hamiltonian of the REM can be obtained by setting

$$E_x = \overline{E}_x \mathbf{1}_{\{\xi_x=1\}} + \underline{E}_x \mathbf{1}_{\{\xi_x=0\}}. \tag{2.6.1}$$

From now on we always assume that E_x are given by (2.6.1). Observe that in this procedure the set $\mathcal{D}_N$ coincides with the set $\{x \in \mathbb{H}_N : \xi_x = 1\}$.

We use $\mathcal{G} = \sigma(\xi, \underline{E})$ to denote the σ-algebra generated by the ξ's and $\underline{E}$'s. In particular, the number and positions of deep traps and all the τ_y, $y \notin \mathcal{D}_N$, are $\mathcal{G}$-measurable. The depths of deep traps are however independent of $\mathcal{G}$.

In the next lemma we compute the quasi-annealed Laplace transform of $S_{\mathcal{D}}$. The term *quasi-annealed* refers to the fact that we average over the energies of the deep traps $\overline{E}_x$ (and over the law of the process), but we keep quenched the positions of the deep traps ξ_x and the energies of the remaining traps $\underline{E}_x$.

Lemma 2.6.2. *There is a constant $\mathcal{K} \in (0, \infty)$ such that for every $\lambda > 0$ and $t \geq 0$,*

$$\mathbb{E}\Big[E_\nu^\tau\big[e^{-\frac{\lambda}{g_N} S_{\mathcal{D}}(tR_N)}\big]\Big|\,\mathcal{G}\Big] \xrightarrow{N\to\infty} e^{-\mathcal{K}\lambda^\alpha t}, \qquad \mathbb{P}\text{-a.s.}$$

Proof. Recall the separation event $\mathscr{S}$ defined in (2.2.9). This event depends only on ξ and is therefore $\mathcal{G}$-measurable, and by Lemma 2.2.1 it occurs $\mathbb{P}$-a.s. for N large enough. On $\mathscr{S}$, no deep traps $x \in \mathcal{D}_N$ are neighbors. Since moreover $\tau_x \geq 1$ for $x \in \mathcal{D}_N$, all the transition rates

$$q_{xy}\mathbf{1}_{\mathscr{S}} = \frac{\tau_x \wedge \tau_y}{1 \wedge \tau_x}\mathbf{1}_{\mathscr{S}}, \qquad x, y \in \mathbb{H}_N,$$

are $\mathcal{G}$-measurable. That is, on the event $\mathscr{S}$, the law of the chain Y is in fact $\mathcal{G}$-measurable. Therefore, on $\mathscr{S}$, the order of taking expectations over the depth of the deep traps and the chain Y can be exchanged. Namely, denoting by $\overline{\mathbb{E}}$ the expectation over the random variables $\overline{E}_x$, on $\mathscr{S}$,

$$\begin{aligned}
\mathbb{E}\Big[E_\nu^\tau\big[e^{-\frac{\lambda}{g_N} S_{\mathcal{D}}(tR_N)}\big]\,\Big|\,\mathcal{G}\Big] &= E_\nu^\tau\Big[\overline{\mathbb{E}}\Big[e^{-\frac{\lambda}{g_N} S_{\mathcal{D}}(tR_N)}\Big]\Big] \\
&= E_\nu^\tau\Big[\overline{\mathbb{E}}\Big[\exp\Big\{-\frac{\lambda}{g_N}\int_0^{tR_N} \tau_{Y_s}\mathbf{1}_{\{Y_s\in\mathcal{D}_N\}}ds\Big\}\Big]\Big] \\
&= E_\nu^\tau\Big[\overline{\mathbb{E}}\Big[\exp\Big\{-\frac{\lambda}{g_N}\sum_{x\in\mathcal{D}_N}\ell_{tR_N}(x)\tau_x\Big\}\Big]\Big].
\end{aligned} \tag{2.6.2}$$

We next approximate the inner expectation on the right-hand side of (2.6.2). Since its argument is bounded by one, it will be sufficient to control it on an event of P_ν^τ-probability tending to 1 as $N \to \infty$. Define the event

$$\mathcal{A} = \big\{\text{for all } x \in \mathcal{D}_N,\ \ell_{tR_N}(x) \le N^2 m_N\big\} \cap \Big\{\big|L_N(t) - t\big| \le 2^{-\frac{1}{5}\varepsilon_0 N}\Big\}. \tag{2.6.3}$$

By Proposition 2.5.1 and Lemma 2.5.6, $\mathbb{P}$-a.s. for N large enough, $P_\nu^\tau[\mathcal{A}^c] \le e^{-cN}$.

When performing the inner expectation of (2.6.2), the local times $\ell_{tR_N}(x)$ of Y as well as $\mathcal{D}_N$ are fixed, the expectation is taken only over the energies of the deep traps. By independence of the $\overline{E}_x$ it follows that

$$\begin{aligned}\overline{\mathbb{E}}\left[e^{-\frac{\lambda}{g_N} S_{\mathcal{D}}(tR_N)}\right] &= \prod_{x\in\mathcal{D}_N} \overline{\mathbb{E}}\left[\exp\left\{-\frac{\lambda}{g_N}\ell_{tR_N}(x) e^{\beta\sqrt{N}\,\overline{E}_x}\right\}\right] \\ &= \exp\left\{\sum_{x\in\mathcal{D}_N} \log\overline{\mathbb{E}}\left[\exp\left\{-\frac{\lambda}{g_N}\ell_{tR_N}(x) e^{\beta\sqrt{N}\,\overline{E}_x}\right\}\right]\right\}.\end{aligned} \tag{2.6.4}$$

For $u \in [0, N^2 m_N]$, let

$$\vartheta(u) = 1 - \overline{\mathbb{E}}\left[\exp\left\{-\frac{\lambda}{g_N} u e^{\beta\sqrt{N}\,\overline{E}_x}\right\}\right].$$

Since $\overline{E}_x$ has standard Gaussian distribution conditioned on being larger than $\frac{1}{\beta\sqrt{N}}\log g'_N$, using that by (2.2.7),

$$\mathbb{P}\Big[E_x > \frac{1}{\beta\sqrt{N}}\log g'_N\Big] = \mathbb{P}[x \in \mathcal{D}_N] = 2^{-\gamma' N}(1 + o(1)),$$

it follows that

$$\vartheta(u) = \frac{2^{\gamma' N}}{\sqrt{2\pi}}\,(1 + o(1)) \int_{\frac{1}{\beta\sqrt{N}}\log g'_N}^{\infty} e^{-\frac{s^2}{2}} \left(1 - e^{-\frac{\lambda u}{g_N} e^{\beta\sqrt{N}s}}\right) ds.$$

We use the substitution $s = \frac{1}{\beta\sqrt{N}}(\beta z + \log g_N - \log\lambda - \log u)$. The lower limit of the integral then becomes

$$\frac{1}{\beta}(\log g'_N - \log g_N + \log\lambda + \log u) =: \omega(N).$$

For $u \le N^2 m_N$, ω_N is asymptotically dominated by $\log g'_N - \log g_N \le -cN$, and thus $\lim_{N\to\infty}\omega(N) = -\infty$. After the substitution,

$$\vartheta(u) = \frac{2^{\gamma' N}}{\sqrt{2\pi}}(1 + o(1)) \int_{\omega(N)}^{\infty} e^{-\frac{1}{2\beta^2 N}(\beta z + \log g_N - \log\lambda - \log u)^2} \left(1 - e^{-e^{\beta z}}\right) \frac{1}{\sqrt{N}}\, dz. \tag{2.6.5}$$

For $u \in [0, N^2 m_N]$, using the definition (2.1.6) of g_N, the exponent of the first exponential satisfies

$$\begin{aligned}&-\frac{1}{2\beta^2 N}(\beta z + \log g_N - \log\lambda - \log u)^2 \\ &\quad= -\frac{1}{2\beta^2 N}(\beta z + \alpha\beta^2 N - \frac{1}{\alpha}\log(\alpha\beta\sqrt{2\pi N}) - \log\lambda - \log u)^2 \\ &\quad= -\frac{\alpha^2\beta^2}{2} N + \alpha\log\lambda + \alpha\log u + \log(\alpha\beta\sqrt{2\pi N}) - \alpha\beta z + \mathrm{err}(z) + o(1).\end{aligned} \tag{2.6.6}$$

Here, $o(1)$ is an error independent of the variable z. Note that for the $\log^2 u$ part to be $o(1)$ it is important that m_N defined in (2.3.7) is not too large, see also Remark 2.6.4. The second error term is

$$\mathrm{err}(z) = -\frac{1}{2N}z^2 + \frac{1}{\beta N}z\left(\frac{1}{\alpha}\log(\alpha\beta\sqrt{2\pi N}) + \log\lambda + \log u\right).$$

Observe that $\lim_{N\to\infty}\mathrm{err}(z) = 0$ for every $z \in \mathbb{R}$, and that for every ε there is N_0 large enough, so that for $N \geq N_0$ and all $z \in \mathbb{R}$

$$\mathrm{err}(z) \leq \varepsilon|z|. \tag{2.6.7}$$

Inserting the results of the computation (2.6.6) back into (2.6.5), using that $\alpha^2\beta^2/2 = \gamma\log 2$, we obtain

$$\vartheta(u) = \alpha\beta 2^{(\gamma'-\gamma)N}\lambda^\alpha u^\alpha \int_{\omega(N)}^{\infty} e^{-\alpha\beta z+\mathrm{err}(z)}\left(1-e^{-e^{\beta z}}\right)dz\,(1+o(1)). \tag{2.6.8}$$

We now claim that

$$\int_{\omega(N)}^{\infty} e^{-\alpha\beta z+\mathrm{err}(z)}\left(1-e^{-e^{\beta z}}\right)dz \xrightarrow{N\to\infty} \int_{\mathbb{R}} e^{-\alpha\beta z}\left(1-e^{-e^{\beta z}}\right)dz =: C. \tag{2.6.9}$$

Indeed, the integrand converges point-wise on $\mathbb{R}$ to $e^{-\alpha\beta z}(1-e^{-e^{\beta z}})$ which is integrable if $\alpha < 1$. Moreover, by (2.6.7), the integrand is bounded by $e^{-\alpha\beta z+\varepsilon|z|}(1-e^{-e^{\beta z}})$, which is integrable if we choose $\varepsilon < \beta(1-\alpha)\wedge\alpha\beta$. The claim (2.6.9) follows by the dominated convergence theorem.

We now come back to (2.6.4). Since on $\mathcal{A}$, $\ell_{tR_N}(x) \leq N^2 m_N$ for all $x \in \mathcal{D}_N$, and $\gamma' < \gamma$, we see that $\vartheta(\ell_{tR_N}(x)) = o(1)$ uniformly in $x \in \mathcal{D}_N$ on $\mathcal{A}$. With $\log(1-x) = -x(1+O(x))$ as $x \to 0$ this yields

$$\begin{aligned}\overline{\mathbb{E}}\left[e^{-\frac{\lambda}{g_N}S_{\mathcal{D}}(tR_N)}\right] &= \exp\Big\{\sum_{x\in\mathcal{D}_N}\log\big(1-\vartheta(\ell_{tR_N}(x))\big)\Big\}\\ &= \exp\Big\{-\sum_{x\in\mathcal{D}_N}\vartheta(\ell_{tR_N}(x))(1+o(1))\Big\}.\end{aligned}$$

The inner sum can be easily computed from (2.6.8). Recalling that on $\mathcal{A}$ the local time functional $L_N(t)$ converges, denoting $\mathcal{K} = \alpha\beta C$, we obtain on $\mathcal{A}$,

$$\begin{aligned}\sum_{x\in\mathcal{D}_N}\vartheta(\ell_{tR_N}(x)) &= \alpha\beta C\lambda^\alpha 2^{(\gamma'-\gamma)N}\sum_{x\in\mathcal{D}_N}\ell_{tR_N}(x)^\alpha(1+o(1))\\ &= \alpha\beta C\lambda^\alpha L_N(t)(1+o(1))\\ &= \mathcal{K}\lambda^\alpha t + o(1) \quad \text{as } N\to\infty.\end{aligned} \tag{2.6.10}$$

It follows that on $\mathcal{A}$

$$\overline{\mathbb{E}}\Big[e^{-\frac{\lambda}{g_N}S_{\mathcal{D}}(tR_N)}\Big] = e^{-\mathcal{K}t\lambda^\alpha(1+o(1))} = e^{-\mathcal{K}t\lambda^\alpha} + o(1) \quad \text{as } N\to\infty.$$

Inserting this into (2.6.2), using that $P_\nu^\tau[\mathcal{A}^c] = O(e^{-cN})$, we conclude that, on $\mathscr{S}$, $\mathbb{P}$-a.s. as $N\to\infty$,

$$\mathbb{E}\left[E_\nu^\tau\left[e^{-\frac{\lambda}{g_N}S_{\mathcal{D}}(tR_N)}\right]\,\middle|\,\mathcal{G}\right] = E_\nu^\tau\left[\overline{\mathbb{E}}\left[e^{-\frac{\lambda}{g_N}S_{\mathcal{D}}(tR_N)}\right]\mathbf{1}_{\mathcal{A}}\right] + O(e^{-cN}) = e^{-\mathcal{K}t\lambda^\alpha} + o(1).$$

Since $\mathscr{S}$ occurs $\mathbb{P}$-a.s. for N large enough, this completes the proof. □

2.6.2. Quenched convergence

We strengthen the convergence in Lemma 2.6.2 in the following way.

Lemma 2.6.3. *The one-dimensional marginals of the rescaled clock processes $g_N^{-1} S_{\mathcal{D}}(tR_N)$ converge in $\mathbb{P}$-probability as $N \to \infty$, in P_ν^τ-distribution to an α-stable law, that is for every $t > 0$ and $\lambda > 0$,*

$$E_\nu^\tau \left[e^{-\frac{\lambda}{g_N} S_{\mathcal{D}}(tR_N)} \right] \xrightarrow{N\to\infty} e^{-\mathcal{K}\lambda^\alpha t} \qquad \text{in } \mathbb{P}\text{-probability.}$$

Proof. It will be enough to show that $\mathbb{P}$-a.s. for N large enough,

$$\mathbb{E}\Big[E_\nu^\tau \big[e^{-\frac{\lambda}{g_N} S_{\mathcal{D}}(tR_N)} \big]^2 \Big| \mathcal{G} \Big] = e^{-2\mathcal{K}\lambda^\alpha t} + o(1). \tag{2.6.11}$$

Indeed, if (2.6.11) holds, then the conditional variance

$$\mathrm{Var}\left[E_\nu^\tau \big[e^{-\frac{\lambda}{g_N} S_{\mathcal{D}}(tR_N)} \big] \Big| \mathcal{G} \right] \xrightarrow{N\to\infty} 0, \qquad \mathbb{P}\text{-a.s.},$$

and the claim follows by an application of the Chebyshev inequality and Lemma 2.6.2.

To show (2.6.11), we rewrite

$$\mathbb{E}\Big[E_\nu^\tau \big[e^{-\frac{\lambda}{g_N} S_{\mathcal{D}}(tR_N)} \big]^2 \Big| \mathcal{G} \Big] = \mathbb{E}\Big[\hat{E}_\nu^\tau \big[e^{-\frac{\lambda}{g_N} \sum_{x\in\mathcal{D}_N} (\ell^{(1)}_{tR_N}(x) + \ell^{(2)}_{tR_N}(x))\tau_x} \big] \Big| \mathcal{G} \Big],$$

where $\ell^{(1)}$ and $\ell^{(2)}$ are the local times of two independent Markov chains $Y^{(1)}$ and $Y^{(2)}$, both having law P_ν^τ, and $\hat{E}_\nu^\tau$ is the expectation with respect to the joint law $\hat{P}_\nu^\tau$ of these chains. Again $\mathbb{P}$-a.s. for N large enough the separation event $\mathscr{S}$ holds, and on this event the law $\hat{P}_\nu^\tau$ is $\mathcal{G}$-measurable. Therefore we can exchange the expectations similarly as before. As in Lemma 2.6.2, it will be enough to control the expression on an event of $\hat{P}_\nu^\tau$-probability tending to 1 as $N \to \infty$. We thus set $\hat{\mathcal{A}} = \mathcal{A}^{(1)} \cap \mathcal{A}^{(2)}$ where $\mathcal{A}^{(i)}$ are defined for both chains $Y^{(i)}$ as in (2.6.3). Applying Proposition 2.5.1 and Lemma 2.5.6 for both independent chains, we have that $\mathbb{P}$-a.s. as $N \to \infty$, $\hat{P}_\nu^\tau[\hat{\mathcal{A}}^c] = O(e^{-cN})$.

Let $\mathcal{C}$ be the event that $Y^{(1)}$ and $Y^{(2)}$ visit disjoint sets of deep traps,

$$\mathcal{C} = \Big\{ \{x \in \mathcal{D}_N :\ \ell^{(1)}_{tR_N}(x) > 0\} \cap \{x \in \mathcal{D}_N :\ \ell^{(2)}_{tR_N}(x) > 0\} = \emptyset \Big\}.$$

We claim that $\hat{P}_\nu^\tau[\mathcal{C}^c] = O(e^{-cN})$, $\mathbb{P}$-a.s. as $N \to \infty$. Indeed, by Lemma 2.5.6, with probability larger than $1 - c2^{-\frac{1}{4}\varepsilon_0 N}$, the chain $Y^{(1)}$ visits at most $2^{\frac{3}{2}\varepsilon_0 N}$ different vertices in $\mathcal{D}_N$. By Lemma 2.5.3, each of those vertices has probability smaller than $c2^{(\gamma-1+\varepsilon)N}$ of being hit by $Y^{(2)}$, for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough. Therefore by the choice (2.5.1) of ε_0, $\mathbb{P}$-a.s. for N large enough,

$$\hat{P}_\nu^\tau[\mathcal{C}^c] \le c2^{-\frac{1}{4}\varepsilon_0 N} + 2^{\frac{3}{2}\varepsilon_0 N} c' 2^{(\gamma-1+\varepsilon)N} \le c2^{-\frac{1}{4}\varepsilon_0 N} + c'2^{-\frac{1}{2}\varepsilon_0 N + \varepsilon N},$$

which decays exponentially if $\varepsilon < \varepsilon_0/2$.

Since on $\mathcal{C}$ the τ_x of the vertices $x \in \mathcal{D}_N$ visited by $Y^{(1)}$ and $Y^{(2)}$ are independent, and since the integrand is bounded by 1, we have on the separation event $\mathscr{S}$,

$$\begin{aligned}
&\mathbb{E}\left[E_\nu^\tau\left[e^{-\frac{\lambda}{g_N}S_{\mathcal{D}}(tR_N)}\right]^2 \mid \mathcal{G}\right] \\
&= \hat{E}_\nu^\tau\left[\overline{\mathbb{E}}\left[e^{-\frac{\lambda}{g_N}\sum_{x\in\mathcal{D}_N}(\ell^{(1)}_{tR_N}(x)+\ell^{(2)}_{tR_N}(x))\tau_x}\right]\right] \\
&= \hat{E}_\nu^\tau\left[\overline{\mathbb{E}}\left[e^{-\frac{\lambda}{g_N}\sum_{x\in\mathcal{D}_N}(\ell^{(1)}_{tR_N}(x)+\ell^{(2)}_{tR_N}(x))\tau_x}\right]\mathbf{1}_{\hat{\mathcal{A}}\cap\mathcal{C}}\right] + O(e^{-cN}) \\
&= \hat{E}_\nu^\tau\left[\overline{\mathbb{E}}\left[e^{-\frac{\lambda}{g_N}\sum_{x\in\mathcal{D}_N}\ell^{(1)}_{tR_N}(x)\tau_x}\right]\overline{\mathbb{E}}\left[e^{-\frac{\lambda}{g_N}\sum_{x\in\mathcal{D}_N}\ell^{(2)}_{tR_N}(x)\tau_x}\right]\mathbf{1}_{\hat{\mathcal{A}}\cap\mathcal{C}}\right] + O(e^{-cN}).
\end{aligned}$$

Using the same procedure as in the proof of Lemma 2.6.2, on the event $\hat{\mathcal{A}}$, the two inner expectations, $x \in \mathcal{D}_N$, both converge to

$$\exp\Big\{-\mathcal{K}\lambda^\alpha 2^{(\gamma'-\gamma)N}\sum_{x\in\mathcal{D}_N}\ell^{(i)}_{tR_N}(x)^\alpha\Big\} = \exp\big\{-\mathcal{K}\lambda^\alpha L_N^{(i)}(t)\big\}, \quad i=1,2.$$

Moreover, on $\hat{\mathcal{A}}$, the local time functionals $L_N^{(i)}(t)$ concentrate on t simultaneously. It follows that on $\mathscr{S}$, $\mathbb{P}$-a.s. as $N\to\infty$,

$$\mathbb{E}\left[E_\nu^\tau\left[e^{-\frac{\lambda}{g_N}S_{\mathcal{D}}(tR_N)}\right]^2 \Big| \mathcal{G}\right] = e^{-2\mathcal{K}\lambda^\alpha t} + o(1).$$

Noting again that $\mathscr{S}$ occurs $\mathbb{P}$-a.s. for N large enough, this shows (2.6.11), and hence the lemma. □

Remark 2.6.4. (a) Inspecting the last proof carefully, it follows that Lemma 2.6.3 can be slightly strengthened. Namely, the stated convergence holds a.s. with respect to ξ and $\underline{E}$, and in probability only with respect to $\overline{E}$. The same remark then applies to Theorem 2.1.1.

(b) A closer analysis of the errors made in the computation of the quasi-annealed Laplace transform, in particular in (2.6.6), shows that the error in Lemma 2.6.2 and (2.6.11) is of order $O(N^{-1}\log^2 N)$, where the logarithmic part comes from the $\log^2 u$ part in (2.6.6), u being bounded by $N^2 m_N$, and m_N being polynomial in N. Therefore the variance decay is not enough to apply the Borel-Cantelli lemma and obtain $\mathbb{P}$-a.s. convergence.

(c) Note also that the previous proof, more precisely bounding the $\log^2 u$ part of (2.6.6), requires that $\log(N^2 m_N) \ll N^{1/2}$. This is where our improved techniques to estimate the spectral gap in Proposition 2.3.1 are necessary. As we already remarked, the techniques of [FIKP98] show roughly that $m_N \le e^{\sqrt{N\log N}}$ only, which is not sufficient.

2.6.3. Joint convergence of increments

In the next step, we extend the convergence to joint convergence of increments.

Lemma 2.6.5. *The increments of the rescaled clock processes $g_N^{-1}S_{\mathcal{D}}(tR_N)$ converge jointly in $\mathbb{P}$-probability in P_ν^τ-distribution to the increments of an α-stable subordinator.*

Proof. Fix $k \geq 1$ and $0 = t_0 < t_1 < \cdots < t_k$. We will show that for every $\lambda_1, \ldots, \lambda_k \in (0, \infty)$ and $\mathbb{P}$-a.e. environment τ,

$$\lim_{N\to\infty} E_\nu^\tau \left[e^{-\frac{1}{g_N}\sum_{i=1}^k \lambda_i (S_{\mathcal{D}}(t_i R_N) - S_{\mathcal{D}}(t_{i-1} R_N))} \right] = \lim_{N\to\infty} \prod_{i=1}^k E_\nu^\tau \left[e^{-\frac{\lambda_i}{g_N} S_{\mathcal{D}}((t_i - t_{i-1}) R_N)} \right]. \quad (2.6.12)$$

Then the lemma follows by using the above proved convergence in $\mathbb{P}$-probability in P_ν^τ-distribution of the one-dimensional marginals.

Let $I^i = [t_i R_N - N^2 m_N, t_i R_N]$. For a set $I \subset [0, \infty)$, let $\mathcal{V}(I)$ be the event

$$\mathcal{V}(I) = \{Y_s \notin \mathcal{D}_N \text{ for all } s \in I\}.$$

On the event $\mathcal{V}\left(\cup_{i=1}^k I^i\right)$, for every $i \leq k$,

$$S_{\mathcal{D}}(t_i R_N) - S_{\mathcal{D}}(t_{i-1} R_N) = S_{\mathcal{D}}(t_i R_N - N^2 m_N) - S_{\mathcal{D}}(t_{i-1} R_N). \quad (2.6.13)$$

Moreover, by Lemma 2.5.3, $\mathbb{P}$-a.s. for all $x \in \mathcal{D}_N$, for $\varepsilon > 0$ small and N large enough,

$$P_\nu^\tau [H_x \leq N^2 m_N] \leq 2^{(\varepsilon - 1)N}.$$

By (2.2.8), $|\mathcal{D}_N| \leq c 2^{(1-\gamma')N}$, hence the expected number of vertices $x \in \mathcal{D}_N$ visited in a time-interval of length $N^2 m_N$ is smaller than $c 2^{(\varepsilon - \gamma')N}$, $\mathbb{P}$-a.s. for N large enough. This still holds for a finite union of intervals of length $N^2 m_N$, and so we conclude that by the Markov inequality, $P_\nu^\tau \left[\mathcal{V}\left(\cup_{i=1}^k I^i\right)\right] \to 1$, $\mathbb{P}$-a.s. as $N \to \infty$.

The reason to shorten the time intervals as above is to give the Markov chain Y the time it needs to mix. Define the event

$$\mathcal{M} = \{T_{\text{mix}} \circ \theta_{t_i R_N - N^2 m_N} \leq N^2 m_N \, \forall i = 1, \ldots, k\}.$$

It is easy to see using Proposition 2.3.3 that $P_\nu^\tau[\mathcal{M}] \to 1$, $\mathbb{P}$-a.s. as $N \to \infty$. On the event $\mathcal{M}$ the Markov chain Y always mixes between $t_i R_N - N^2 m_N$ and $t_i R_N$ and thus, by Lemma 2.3.4, for every $i = 1, \ldots, k$ and $y \in \mathbb{H}_N$,

$$P_\nu^\tau [Y_{t_i R_N} = y \mid \mathcal{M}] = \nu_y.$$

Therefore, on $\mathcal{M}$,

$$\left(S_{\mathcal{D}}(t_i R_N - N^2 m_N) - S_{\mathcal{D}}(t_{i-1} R_N)\right)_{i=1,\ldots,k} \stackrel{d}{=} \left(S_{\mathcal{D}}^{(i)}((t_i - t_{i-1}) R_N - N^2 m_N)\right)_{i=1,\ldots,k}, \quad (2.6.14)$$

where the $S_{\mathcal{D}}^{(i)}$ are the clock processes of the deep traps of independent stationary started processes $Y^{(i)}$ having the same law as Y.

Combining observations (2.6.13) and (2.6.14), with the estimates on the probabilities of $\mathcal{V}\left(\cup_{i=1}^k I^i\right)$ and $\mathcal{M}$, since the integrand is bounded by 1, we obtain that $\mathbb{P}$-a.s. as $N \to \infty$,

$$\begin{aligned}
&E_\nu^\tau \left[e^{-\frac{1}{g_N}\sum_{i=1}^k \lambda_i (S_{\mathcal{D}}(t_i R_N) - S_{\mathcal{D}}(t_{i-1} R_N))} \right] \\
&\quad = E_\nu^\tau \left[e^{-\frac{1}{g_N}\sum_{i=1}^k \lambda_i (S_{\mathcal{D}}(t_i R_N - N^2 m_N) - S_{\mathcal{D}}(t_{i-1} R_N))} \mathbf{1}_{\mathcal{V}\left(\cup_{i=1}^k I^i\right) \cap \mathcal{M}} \right] + o(1) \\
&\quad = E_\nu^\tau \left[\prod_{i=1}^k E_\nu^\tau \left[e^{-\frac{\lambda_i}{g_N} S_{\mathcal{D}}^{(i)}((t_i - t_{i-1}) R_N - N^2 m_N)} \right] \mathbf{1}_{\mathcal{V}\left(\cup_{i=1}^k I^i\right) \cap \mathcal{M}} \right] + o(1) \\
&\quad = \prod_{i=1}^k E_\nu^\tau \left[e^{-\frac{\lambda_i}{g_N} S_{\mathcal{D}}^{(i)}((t_i - t_{i-1}) R_N - N^2 m_N)} \right] + o(1).
\end{aligned}$$

Using analogous arguments it can be shown that for every $i = 1, \dots, k$, $\mathbb{P}$-a.s. as $N \to \infty$,

$$E_\nu^\tau \left[e^{-\frac{\lambda_i}{g_N} S_{\mathcal{D}}^{(i)}((t_i - t_{i-1})R_N - N^2 m_N)} \right] = E_\nu^\tau \left[e^{-\frac{\lambda_i}{g_N} S_{\mathcal{D}}^{(i)}((t_i - t_{i-1})R_N)} \right] + o(1).$$

Combining the last two equations proves (2.6.12) and hence the lemma. □

2.6.4. Tightness in the Skorohod topology

The last step in the proof of Proposition 2.6.1 is to show tightness.

Lemma 2.6.6. *The sequence of probability measures $P_\nu^\tau \left[g_N^{-1} S_{\mathcal{D}}(tR_N) \in \cdot \right]$ is $\mathbb{P}$-a.s. tight with respect to the Skorohod M_1-topology on $D([0,T], \mathbb{R})$.*

Proof. The proof is standard but we include it for the sake of completeness. By [Whi02, Theorem 12.12.3], the tightness in the Skorohod M_1-topology on $D([0,T], \mathbb{R})$ is characterized in the following way: For $f \in D([0,T], \mathbb{R})$, $\delta > 0$, $t \in [0,T]$, let

$$w_f(\delta) = \sup \left\{ \inf_{\alpha \in [0,1]} |f(t) - (\alpha f(t_1) + (1-\alpha) f(t_2))| : \ t_1 \le t \le t_2 \le T, \ t_2 - t_1 \le \delta \right\},$$
$$v_f(t, \delta) = \sup \left\{ |f(t_1) - f(t_2)| : \ t_1, t_2 \in [0,T] \cap (t - \delta, t + \delta) \right\}.$$

The sequence of probability measures $P_N = P_\nu^\tau \left[g_N^{-1} S_{\mathcal{D}}(tR_N) \in \cdot \right]$ on $D([0,T], \mathbb{R})$ is tight in the M_1-topology, if

(i) For every $\varepsilon > 0$ there is c such that

$$P_N[f : \ \|f\|_\infty > c] \le \varepsilon, \quad N \ge 1. \tag{2.6.15}$$

(ii) For every $\varepsilon > 0$ and $\eta > 0$, there exist $\delta \in (0,T)$ and N_0 such that

$$P_N[f : \ w_f(\delta) \ge \eta] \le \varepsilon, \quad N \ge N_0, \tag{2.6.16}$$

and

$$P_N[f : \ v_f(0,\delta) \ge \eta] \le \varepsilon \text{ and } P_N[f : \ v_f(T,\delta) \ge \eta] \le \varepsilon, \quad N \ge N_0. \tag{2.6.17}$$

Since the clock processes are increasing, (2.6.15) is equivalent to convergence of the distribution of $g_N^{-1} S_{\mathcal{D}}(TR_N)$, which follows from the convergence of the Laplace transform of the marginal at time T. (2.6.16) is immediate from the fact that the oscillating function $w_f(\delta)$ is always zero since the processes $g_N^{-1} S_{\mathcal{D}}(tR_N)$ are increasing. To check (2.6.17), again by the monotonicity of the $g_N^{-1} S_{\mathcal{D}}(tR_N)$ it is enough to check that for δ small enough and $N \ge N_0$, $P_\nu^\tau[g_N^{-1} S_{\mathcal{D}}(\delta R_N) \ge \eta] \le \varepsilon$. By the convergence of the marginal at time δ, we may take δ such that $\mathbb{P}[V_\alpha(\delta) \ge \eta] \le \frac{\varepsilon}{2}$ and N_0 such that for $N \ge N_0$,

$$\left| P_\nu^\tau \left[\frac{1}{g_N} S_{\mathcal{D}}(\delta R_N) \ge \eta \right] - \mathbb{P}\left[V_\alpha(\delta) \ge \eta \right] \right| \le \frac{\varepsilon}{2}.$$

The reasoning for $v_f(T, \delta)$ is similar. □

2.7. Shallow traps

In this section we show that the convergence of the clock process of the deep traps shown in Section 2.6 is enough for convergence of the clock process itself.

Proposition 2.7.1. *Under the assumptions of Theorem 2.1.1, the clock process of the deep traps approximates the clock process, namely, for every $t \geq 0$,*

$$\frac{1}{g_N}\big(S(tR_N) - S_{\mathcal{D}}(tR_N)\big) \xrightarrow{N\to\infty} 0 \qquad \mathbb{P}\text{-a.s. in } P^\tau_\nu\text{-probability.}$$

Proof. We will split the set of shallow traps $\mathcal{S}_N := \mathbb{H}_N \setminus \mathcal{D}_N$ into two parts and separately deal with the corresponding contributions to the clock process.

We start with 'very shallow traps'. Let $\delta > 0$ be a small constant which will be fixed later and $h_N = e^{\delta\alpha\beta^2 N}$. Define the set of very shallow traps as

$$\overline{\mathcal{S}}_N = \{x \in \mathbb{H}_N : \ \tau_x \leq h_N\}.$$

The contribution of this set to the clock process can easily be neglected as follows. Write

$$E^\tau_\nu\left[\frac{1}{g_N}\int_0^{tR_N} (1 \vee \tau_{Y_s})\mathbf{1}_{\{Y_s \in \overline{\mathcal{S}}_N\}} ds\right] = \frac{1}{g_N}\sum_{x\in\overline{\mathcal{S}}_N} (1 \vee \tau_x) E^\tau_\nu\big[\ell_{tR_N}(x)\big]$$

Note that $E^\tau_\nu[\ell_{tR_N}(x)] = \nu_x tR_N = Z_N^{-1}(1 \wedge \tau_x)tR_N$, and $(1 \vee \tau_x)(1 \wedge \tau_x) = \tau_x \leq h_N$ on $\overline{\mathcal{S}}_N$. With (2.2.3) for Z_N, and Lemma 2.4.4 for R_N, for every $\epsilon > 0$, $\mathbb{P}$-a.s. for N large enough, the right-hand side of the last equation can be bounded from above by

$$g_N^{-1} 2^N h_N Z_N^{-1} tR_N \leq c g_N^{-1} e^{\delta\alpha\beta^2 N} 2^{(\gamma+\epsilon)N}.$$

To obtain exponential decay of this expression, it is enough to take account of the exponential part of g_N, which is $e^{\alpha\beta^2 N}$. Then, up to subexponential factors, using that $\gamma = \frac{\alpha^2\beta^2}{2\log 2}$, the above is bounded by

$$\exp\big\{((\delta-1)\alpha\beta^2 + \frac{1}{2}\alpha^2\beta^2 + \epsilon\log 2)N\big\}.$$

Since $\alpha < 1$, by choosing ϵ and δ small enough this can be made smaller than e^{-cN} for some $c > 0$. Applying the Markov inequality and the Borel-Cantelli lemma,

$$\frac{1}{g_N}\int_0^{tR_N} (1 \vee \tau_{Y_s})\mathbf{1}_{\{Y_s \in \overline{\mathcal{S}}_N\}} ds \xrightarrow{N\to\infty} 0 \qquad \mathbb{P}\text{-a.s. in } P^\tau_\nu\text{-probability.} \tag{2.7.1}$$

To control the contribution of the remaining shallow traps $\mathcal{S}_N \setminus \overline{\mathcal{S}}_N$, we first split this set into slices $\mathcal{S}_N^i$ as follows. Set

$$I_N = \left\lceil \frac{1}{\log 2}(\log g'_N - \log h_N) \right\rceil.$$

Note that by definition of g'_N and h_N, for δ small as fixed above, $I_N = cN + O(1)$ for some $c > 0$. For $i = 1, \dots, I_N$, let

$$\mathcal{S}_N^i = \big\{x \in \mathcal{S}_N \setminus \overline{\mathcal{S}}_N : \ \tau_x \in [2^{-i} g'_N, 2^{-i+1} g'_N)\big\},$$

so that $\mathcal{S}_N \setminus \overline{\mathcal{S}}_N = \cup_{i=1}^{I_N} \mathcal{S}_N^i$.

We next control the sizes of the slices $\mathcal{S}_N^i$. By the tail approximation (2.2.2), for all $i = 1, \dots, I_N$,

$$\begin{aligned} \mathbb{P}[y \in \mathcal{S}_N^i] &\le \mathbb{P}\Big[E_x > \frac{1}{\beta\sqrt{N}}(\log g'_N - i \log 2)\Big] \\ &= f_{N,i}^{(1)} \exp\Big\{ -\frac{1}{2}\alpha'^2 \beta^2 N + \alpha' i \log 2 - f_{N,i}^{(2)} - o(1)\Big\}(1 + o(1)). \end{aligned} \tag{2.7.2}$$

We separately control the two expressions $f_{N,i}^{(1)}$ and $f_{N,i}^{(2)}$. The first one equals

$$f_{N,i}^{(1)} = \frac{\alpha' \beta \sqrt{2\pi N}}{\frac{\sqrt{2\pi}}{\beta\sqrt{N}}(\log g'_N - i \log 2)}.$$

To control this, note that by definition of I_N, for all $i = 1, \dots, I_N$,

$$\log g'_N - i \log 2 \ge \log h_N - \log 2 = \delta \alpha \beta^2 N - \log 2.$$

It follows that, for all $i = 1, \dots, I_N$, $f_{N,i}^{(1)}$ is bounded by some constant $c > 0$, which can be chosen to be independent of i. The second expression to control in (2.7.2) is

$$f_{N,i}^{(2)} = \frac{i^2 \log^2 2}{2\beta^2 N} + \frac{i \log 2}{\alpha' \beta^2 N} \log(\alpha' \beta \sqrt{2\pi N}).$$

This is strictly positive, so it can be omitted in (2.7.2) in order to obtain an upper bound. Using the obtained control on $f_{N,i}^{(1)}$ and $f_{N,i}^{(2)}$ in (2.7.2), as well as the fact that $\gamma' = \frac{\alpha'^2 \beta^2}{2 \log 2}$, we conclude that for all $i = 1, \dots, I_N$,

$$\mathbb{P}[y \in \mathcal{S}_N^i] \le c 2^{-\gamma' N} 2^{\alpha' i}.$$

In particular, the size $|\mathcal{S}_N^i|$ of the i-th slice is dominated by a binomial random variable with parameters $n = 2^N$ and $p = c 2^{\alpha' i} 2^{-\gamma' N}$. Then it follows by the Markov inequality that for every $\epsilon > 0$,

$$\mathbb{P}\big[|\mathcal{S}_N^i| > 2^{\epsilon N} c 2^{\alpha' i} 2^{(1-\gamma')N}\big] \le 2^{-\epsilon N}.$$

Since $I_N = cN + O(1)$, a union bound and the Borel-Cantelli lemma imply that for every $\epsilon > 0$, $\mathbb{P}$-a.s. for N large enough,

$$|\mathcal{S}_N^i| \le 2^{\epsilon_N} c 2^{\alpha' i} 2^{(1-\gamma')N}, \text{ for all } i = 1, \dots, I_N. \tag{2.7.3}$$

Coming back to the contribution of the intermediate traps $\mathcal{S}_N \setminus \overline{\mathcal{S}}_N$ to the clock process, we use as before that $E_\nu^\tau[\ell_{tR_N}(y)] = \nu_y t R_N = \frac{1 \wedge \tau_y}{Z_N} t R_N$, and $(1 \vee \tau_y)(1 \wedge \tau_y) = \tau_y \le 2^{-i+1} g'_N$

on $\mathcal{S}_N^i$. With (2.2.3) for Z_N, Lemma 2.4.4 for R_N, and (2.7.3) for the size of $\mathcal{S}_N^i$, we obtain that for every $\varepsilon > 0$, $\mathbb{P}$-a.s. for N large enough, for all $i = 1, \dots, I_N$,

$$\begin{aligned} E_\nu^\tau\left[\frac{1}{g_N}\int_0^{tR_N}(1 \vee \tau_{Y_s})\mathbf{1}_{\{Y_s \in \mathcal{S}_N^i\}}ds\right] &= \frac{1}{g_N}\sum_{y \in \mathcal{S}_N^i}(1 \vee \tau_y)E_\nu^\tau[\ell_{tR_N}(y)] \\ &\leq g_N^{-1}|\mathcal{S}_N^i|2^{-i+1}g'_N Z_N^{-1} tR_N \\ &\leq c\frac{g'_N}{g_N}2^{(\alpha-1)i}2^{(\gamma-\gamma'+2\varepsilon)N}. \end{aligned}$$

Summing over $i = 1, \dots, I_N$, $\mathbb{P}$-a.s. for N large enough,

$$E_\nu^\tau\left[\frac{1}{g_N}\int_0^{tR_N}(1 \vee \tau_{Y_s})\mathbf{1}_{\{Y_s \in \bigcup_{i=1}^{I_N}\mathcal{S}_N^i\}}ds\right] \leq c'\frac{g'_N}{g_N}2^{(\gamma-\gamma'+2\varepsilon)N}. \tag{2.7.4}$$

We claim that the right hand side of (2.7.4) decays exponentially in N for $\varepsilon > 0$ small enough. To this end, as before, it is enough to take account of the exponential parts in both g_N and g'_N, which contribute to the right hand side of (2.7.4) by

$$e^{(\alpha'-\alpha)\beta^2 N} = 2^{(\sqrt{\gamma'}-\sqrt{\gamma})\frac{2\beta}{\beta_c}N}.$$

Hence, to show the exponential decay on the right hand side of (2.7.4), it is sufficient to prove that we can choose $\varepsilon > 0$ small enough, such that

$$(\sqrt{\gamma'} - \sqrt{\gamma})\frac{2\beta}{\beta_c} + \gamma - \gamma' + 2\varepsilon < 0. \tag{2.7.5}$$

With a first order approximation of the concave function $\sqrt{x}$ at γ,

$$\frac{1}{2\sqrt{\gamma}}(\gamma - \gamma') < \sqrt{\gamma} - \sqrt{\gamma'}.$$

Since, $\frac{1}{2\sqrt{\gamma}} = \frac{\beta_c}{2\alpha\beta} > \frac{\beta_c}{2\beta}$ and $\alpha < 1$, this implies

$$\frac{\beta_c}{2\beta}(\gamma - \gamma') < \sqrt{\gamma} - \sqrt{\gamma'},$$

and (2.7.5) thus holds for $\varepsilon > 0$ small enough. The right hand side of (2.7.4) then decays exponentially, and with Markov inequality we conclude that

$$\frac{1}{g_N}\int_0^{tR_N}(1 \vee \tau_{Y_s})\mathbf{1}_{\{Y_s \in \bigcup_{i=1}^{I_N}\mathcal{S}_N^i\}}ds \xrightarrow{N\to\infty} 0 \quad \mathbb{P}\text{-a.s. in } P_\nu^\tau\text{-probability.}$$

This together with (2.7.1) finishes the proof of the proposition. □

2.8. Conclusion

Theorem 2.1.1 is a direct consequence of Propositions 2.6.1, 2.7.1 and Lemma 2.4.4.

Appendix 2.A. Extremal characterization of mean hitting time

In this appendix we give the proof of the formula (2.4.1) which gives a lower bound on the mean hitting time of a set when starting from stationarity. This formula is a continuous-time version of (a half of) Proposition 3.2 from [ČTW11]. This proposition, as well as the underlying result [AF02, Proposition 3.41], are stated for a continuous-time Markov chain whose waiting times are mean-one exponential random variables. We were not able to find analogous statements for general continuous-time Markov chains in the literature, so we provide short proofs here, for the sake of completeness.

We start by introducing some notation. Let Y be a reversible continuous-time Markov chain on a finite state space $\mathcal{S}$ with transition rates q_{xy} and invariant probability measure ν_x, denote by P_ν and P_x the laws of Y started stationary and from x respectively, and by E_ν, E_x the corresponding expectations. Define the conductances as $c_{xy} = \nu_x q_{xy} = \nu_y q_{yx}$. Let $q_x = \sum_y q_{xy}$ and $c_x = \sum_y c_{xy}$. The transition probability from x to y is $p_{xy} = \frac{q_{xy}}{q_x} = \frac{c_{xy}}{c_x}$. In the same way as in Section 2.2, we define the hitting time H_x and the return time H_x^+ to x by Y, and similarly H_A and H_A^+ for sets $A \subset \mathcal{S}$.

A function g on $\mathcal{S}$ is called *harmonic* in x, if $\sum_y g(y)p_{xy} = g(x)$. For $x \in \mathcal{S}$ and $B \subset \mathcal{S} \setminus \{x\}$, the *equilibrium potential* $g^\star_{x,B}$ is defined as the unique function on $\mathcal{S}$ that is harmonic on $(x \cup B)^c$, 1 on x and 0 on B. It is well known that

$$g^\star_{x,B}(y) = P_y[H_x \le H_B].$$

For a function $g : \mathcal{S} \to \mathbb{R}$, the *Dirichlet form* is defined as

$$D(g,g) = \frac{1}{2}\sum_{z\in\mathcal{S}}\sum_{y\sim z} \nu_z q_{zy}(g(z) - g(y))^2, \tag{2.A.1}$$

where $y \sim z$ means that y and z are neighbors in the sense that $q_{zy} > 0$.

The following proposition is the required generalization of Proposition 3.2 of [ČTW11].

Proposition 2.A.1. *For every $x \in \mathcal{S}$ and $B \subset \mathcal{S} \setminus \{x\}$*

$$\frac{1}{E_\nu[H_x]} \le D(g^\star_{x,B}, g^\star_{x,B})\nu(B)^{-2} = c_x P_x[H_x^+ > H_B]\nu(B)^{-2}. \tag{2.A.2}$$

To prove this proposition we will need a lemma which is a generalization of [AF02, Proposition 3.41] giving the extremal characterization of the mean hitting time.

Lemma 2.A.2. *For every $x \in \mathcal{S}$,*

$$\frac{1}{E_\nu[H_x]} = \inf\left\{ D(g,g) :\ g : \mathcal{S} \to \mathbb{R},\ g(x) = 1, \sum_{y\in\mathcal{S}} \nu_y g(y) = 0 \right\}. \tag{2.A.3}$$

Proof. The proof follows the lines of [AF02] with some minor changes to fit into the setting of general continuous-time chains.

We first show that there is a minimizing function g that equals $g(y) = \frac{Z_{yx}}{Z_{xx}}$, where

$$Z_{yx} = \int_0^\infty \Big(P_y[Y_t = x] - \nu_x \Big)\, dt.$$

To this end, we introduce the Lagrange multiplier γ and consider g as the minimizer of $D(g,g) + \gamma \sum_z \nu_z g(z)$ with $g(x) = 1$. The contribution to this of $g(y)$ for $y \neq x$ is

$$\sum_{z\sim y} \nu_y q_{yz} (g(y) - g(z))^2 + \gamma \nu_y g(y),$$

which is minimized if

$$2 \sum_{z\sim y} \nu_y q_{yz} (g(y) - g(z)) + \gamma \nu_y = 0.$$

From this we get for all $y \in \mathcal{S}$, by introducing the term including the parameter β for the case $y = x$, that

$$g(y) = \sum_{z\sim y} \frac{q_{yz}}{q_y} g(z) - \frac{\gamma}{2} \frac{1}{q_y} + \frac{\beta}{q_y} \mathbf{1}_{\{y=x\}}.$$

Multiplying by q_y and ν_y, and summing over all $y \in \mathcal{S}$,

$$\sum_y \sum_{z\sim y} \nu_y q_{yz} g(y) = \sum_y \sum_{z\sim y} \nu_y q_{yz} g(z) - \frac{\gamma}{2} + \beta \nu_x.$$

By reversibility $\nu_y q_{yz} = \nu_z q_{zy}$, so the term on the left and the first term on the right are identical, which gives $\frac{\gamma}{2} = \beta \nu_x$. Thus there is a minimizing g such that

$$g(y) = \frac{\beta}{q_y} \big(\mathbf{1}_{\{y=x\}} - \nu_x \big) + \sum_{z\sim y} \frac{q_{yz}}{q_y} g(z). \tag{2.A.4}$$

We now show that up to the factor β the function $y \mapsto Z_{yx}$ satisfies the same relation. Indeed, by the strong Markov property at the time J_1 of the first jump of Y, which under P_y is an exponential random variable with mean $\frac{1}{q_y}$,

$$\begin{aligned} Z_{yx} &= \int_0^\infty \left(\int_0^{J_1} \big(\mathbf{1}_{\{y=x\}} - \nu_x \big) dt + \sum_{z\sim y} \frac{q_{yz}}{q_y} \int_0^\infty \big(P_z[Y_t = x] - \nu_x \big) dt \right) dP_y(J_1) \\ &= \frac{1}{q_y} \big(\mathbf{1}_{\{y=x\}} - \nu_x \big) + \sum_{z\sim y} \frac{q_{yz}}{q_y} Z_{zx}. \end{aligned}$$

The function $g(y) = \frac{Z_{yx}}{Z_{xx}}$ thus satisfies the constrains of the variational problem in (2.A.3) and fulfills (2.A.4) with $\beta = 1/Z_{xx}$. It is thus the minimizer of this variational problem.

Moreover, by [AF02, Lemmas 2.11 and 2.12], we have $Z_{xx} = E_\nu[H_x]\nu_x$ and $\nu_x E_y[H_x] = Z_{xx} - Z_{yx}$. Denoting $h(y) = E_y[H_x]$ and using these equalities, we obtain

$$D(g,g) = \frac{1}{E_\nu[H_x]^2} D(h,h) = \frac{1}{E_\nu[H_x]},$$

where for the last equality we used $D(h,h) = E_\nu[H_x]$, by e.g. [AB92, Lemma 6]. This completes the proof. □

With this lemma the proof of Proposition 2.A.1 follows the lines of [ČTW11].

Proof of Proposition 2.A.1. To prove the inequality in (2.A.2), it is sufficient to modify the function $g^\star_{x,B}$ so that it becomes admissible for the variational problem in Lemma 2.A.2. Write $g^\star$ for $g^\star_{x,B}$ and define $\tilde{g}$ on $\mathcal{S}$ as

$$\tilde{g}(z) = \frac{g^\star(z) - \sum_{y\in\mathcal{S}} \nu_y g^\star(y)}{1 - \sum_{y\in\mathcal{S}} \nu_y g^\star(y)}.$$

Then $\tilde{g}$ equals 1 on x and $\sum_{z\in\mathcal{S}} \nu_z \tilde{g}(z) = 0$. Hence, by Lemma 2.A.2,

$$\frac{1}{E_\nu[H_x]} \leq D(\tilde{g},\tilde{g}) = D(g^\star,g^\star)\left(1 - \sum_{y\in\mathcal{S}} \nu_y g^\star(y)\right)^{-2}.$$

But $g^\star$ is non-negative, bounded by 1 and non-zero only on B^c, therefore $\sum_{y\in\mathcal{S}} \nu_y g^\star(y) \leq \nu(B^c)$, the first part of Proposition 2.A.1 follows.

To prove the equality in (2.A.2), we show that

$$D(g^\star_{x,B}, g^\star_{x,B}) = P_x[H_x^+ > H_B]c_x. \tag{2.A.5}$$

Indeed, let again $g^\star = g^\star_{x,B}$. If $g^\star$ is harmonic in z, the second sum in the Dirichlet form (2.A.1) is

$$\sum_{y\sim z} c_{zy}(g^\star(z) - g^\star(y))^2 = \sum_{y\sim z} c_{zy}(g^\star(y)^2 - g^\star(z)^2).$$

This shows that the contribution to the Dirichlet form of every edge that connects two vertices in which $g^\star$ is harmonic or zero vanishes. Therefore $D(g^\star,g^\star)$ reduces to

$$\begin{aligned} D(g^\star,g^\star) &= \frac{1}{2}\left(\sum_{y\sim x} c_{xy}(1-g^\star(y))^2 + \sum_{y\sim x} c_{xy}(1-g^\star(y)^2)\right) \\ &= \sum_{y\sim x} c_{xy}(1-g^\star(y)) \\ &= c_x \sum_{y\sim x} p_{xy} P_y[H_x > H_B] \\ &= c_x P_x[H_x^+ > H_B]. \end{aligned}$$

This proves (2.A.5) and thus the proposition. □

3. Randomly trapped random walks on $\mathbb{Z}^d$

Jiří Černý and Tobias Wassmer

Abstract. We give a complete classification of scaling limits of randomly trapped random walks and associated clock processes on $\mathbb{Z}^d$, $d \geq 2$. Namely, under the hypothesis that the discrete skeleton of the randomly trapped random walk has a slowly varying return probability, we show that the scaling limit of its clock process is either deterministic linearly growing or a stable subordinator. In the case when the discrete skeleton is a simple random walk on $\mathbb{Z}^d$, this implies that the scaling limit of the randomly trapped random walk is either Brownian motion or the Fractional Kinetics process, as conjectured in [BCČR14].

3.1. Introduction

Randomly trapped random walks (RTRWs) were introduced in [BCČR14] for two main reasons. On one hand they generalize several classical models of trapped random walks such as the continuous-time random walk or the symmetric Bouchaud trap model. On the other hand they provide a tool for describing random walks on some classical random structures such as the incipient critical Galton-Watson tree or the invasion percolation cluster on a regular tree.

In [BCČR14] the authors define the RTRW on general graphs and study in depth the model on $\mathbb{Z}$. They give a complete classification of scaling limits, showing that the limit of a RTRW on $\mathbb{Z}$ is one of the following four processes: (i) Brownian motion, (ii) *Fractional Kinetics process*, (iii) *FIN singular diffusion*, or (iv) a new class of processes called *spatially subordinated Brownian motion*. They further give sufficient conditions for convergence to the respective limits and study in detail how the different limits arise. For RTRW on $\mathbb{Z}^d$, $d \geq 2$, they conjectured that only the first two of the above scaling limits are possible, that is RTRW on $\mathbb{Z}^d$ converges after rescaling either to the Brownian motion or to the Fractional Kinetics process. We prove this conjecture here.

Let us briefly introduce the model, its formal definition is given in Section 3.2 below. The RTRW on $\mathbb{Z}^d$ is a particular class of random walk in random environment. Its law is determined by two inputs: (i) its step distribution, that is a probability measure ν on $\mathbb{Z}^d$, and (ii) a probability distribution μ on the space of all probability measures on $(0, \infty)$ characterizing its waiting times. The random environment of the RTRW is given by an i.i.d. collection $\pi = (\pi_x)_{x \in \mathbb{Z}^d}$ of μ-distributed probability measures. For fixed π, the RTRW $X = (X(t))_{t \geq 0}$ is a continuous-time process such that, whenever at vertex x, it stays there a random duration sampled from the distribution π_x and then moves on

according to the transition kernel $\nu(\cdot - x)$. If the process X visits x again at a later time, the duration of this next visit at x is sampled again and independently from the distribution π_x. We always assume that X starts at $0 \in \mathbb{Z}^d$ and use $\mathbb{P}$ for the annealed distribution of the process X.

From the description above it is apparent that the RTRW is a time change of the discrete-time random walk $(Y(n))_{n\geq 0}$ on $\mathbb{Z}^d$ with one-step distribution ν. Formally, X can be written as

$$X(t) = Y(S^{-1}(t)), \tag{3.1.1}$$

where the time-change process $S : \mathbb{N} \to [0, \infty)$, the *clock process*, measures the time needed for a given number of steps of the RTRW and S^{-1} is its right-continuous inverse. In view of (3.1.1) it should not be surprising that the scaling behavior of X is (essentially) determined by the scaling behavior of the clock process.

While we are primarily interested in Y being a simple random walk on $\mathbb{Z}^d$, $d \geq 2$, it does not complicate the proofs to make the following far less restrictive assumption on the random walk Y, that is on the one-step distribution ν: Let $r_n : \mathbb{N} \to [0, 1]$ be the probability that Y does not return to its starting point in n steps,

$$r_n = \mathbb{P}[Y(k) \neq Y(0) \text{ for } k = 1, \dots, n].$$

Assumption A. *The function r_n can be written as $r_n = \frac{1}{\ell^*(n)}$ for a slowly varying function $\ell^* : \mathbb{N} \to [1, \infty)$.*

Assumption A is obviously fulfilled for all transient random walks, where $1/\ell^*(n) \to \gamma$ for some $\gamma \in (0, 1)$, but there are also recurrent walks satisfying it with $\ell^*(n) \to \infty$ as $n \to \infty$. In particular, the classical result of Kesten and Spitzer [KS63, Theorem 3] implies that this assumption holds for all random walks on $d \geq 2$ for which the subgroup of $\mathbb{Z}^d$ generated by the set $\{x : \nu(x) > 0\}$ is d-dimensional (we will call this 'genuinely d-dimensional').

We can state our first main theorem giving the complete classification of the scaling limits of the clock process.

Theorem 3.1.1. *Let $S : \mathbb{N} \to [0, \infty)$ be the clock process of the RTRW. Suppose that Assumption A holds and there is a sequence $a_N \nearrow \infty$ such that for all but countably many $t \in [0, \infty)$*

$$S_N(t) := \frac{1}{a_N} S(\lfloor Nt \rfloor) \xrightarrow{N\to\infty} \mathcal{S}(t) \qquad \text{in } \mathbb{P}\text{-distribution}, \tag{3.1.2}$$

where $\mathcal{S} : [0, \infty) \to [0, \infty)$ is a cadlag process satisfying the non-triviality assumption

$$\limsup_{t\to\infty} \mathcal{S}(t) = \infty \qquad \mathbb{P}\text{-a.s.} \tag{3.1.3}$$

Then one of the following two cases occurs:

(i) The limit clock process is linear, $\mathcal{S}(t) = Mt$ for some constant $M > 0$, and the normalizing sequence satisfies $a_N = N\ell(N)$ for some slowly varying function ℓ.

(ii) The limit clock process is an α-stable subordinator, $\mathcal{S} = V_\alpha$, $\alpha \in (0, 1)$, and the normalizing sequence satisfies $a_N = N^{1/\alpha}\ell(N)$ for some slowly varying function ℓ.

In order to study the scaling limits of the RTRW itself, we need a more restrictive assumption:

Assumption B. *The random walk Y is genuinely d-dimensional. Its one-step distribution ν is centered, $\mathbb{E}[Y(1)] = 0$, and has finite range, $\mathbb{P}[|Y(1)| > C] = 0$ for some $C < \infty$.*

This assumption ensures that the scaling limit of Y is a d-dimensional Brownian motion: There exists a $d \times d$ matrix $\mathcal{A}$ such that

$$Y_N(t) := \frac{1}{\sqrt{N}} \mathcal{A} Y(\lfloor Nt \rfloor) \tag{3.1.4}$$

converges to a standard d-dimensional Brownian motion. Note that by the remark after Assumption A, for $d \geq 2$ Assumption A is implied by Assumption B.

Our second main result classifies the possible scaling limits of RTRW and confirms the conjecture of [BCČR14].

Theorem 3.1.2. *Let $d \geq 2$ and $X : [0, \infty) \to \mathbb{Z}^d$ be the RTRW. Suppose that Assumption B holds and there is a sequence $a_N \nearrow \infty$ such that the processes*

$$X_N(t) := \frac{1}{\sqrt{N}} \mathcal{A} X(a_N t) = Y_N(S_N^{-1}(t)) \tag{3.1.5}$$

converge in $\mathbb{P}$-distribution on the space D^d of cadlag $\mathbb{R}^d$-valued functions equipped with the Skorohod J_1-topology to some process $\mathcal{X} : [0, \infty) \to \mathbb{R}^d$ satisfying the non-triviality assumption

$$\limsup_{t \to \infty} |\mathcal{X}(t)| = \infty \qquad \mathbb{P}\text{-}a.s. \tag{3.1.6}$$

Then one of the following two cases occurs:

- *(i) $a_N = N\ell(N)$ and $\mathcal{X}(t) = B(M^{-1}t)$ for some constant $M > 0$, some slowly varying function ℓ, and a standard d-dimensional Brownian motion B.*
- *(ii) $a_N = N^{1/\alpha}\ell(N)$ for some slowly varying function ℓ and a parameter $\alpha \in (0, 1)$, and $\mathcal{X}(t) = B(V_\alpha^{-1}(t))$, where B is a standard d-dimensional Brownian motion and $V_\alpha^{-1}(t) = \inf\{s \geq 0 : \ V_\alpha(s) > t\}$ is the right-continuous inverse of an α-stable subordinator V_α which is independent of B (i.e. $\mathcal{X}$ is the Fractional Kinetics process).*

Let us make a few remarks about our setting and results. The definition of the RTRW we give here is slightly more general than the one in [BCČR14] since we allow the discrete skeleton to be more general than the simple random walk only. Assumption A on the discrete skeleton is taken from [FM14]. This assumption can be used to show weak laws of large numbers for the range of the random walk and for some related quantities. We would like to point out that the only place in the proof of Theorem 3.1.1 where we use $\mathbb{Z}^d$-specific properties of the random walk is in the derivation of these laws of large numbers. In particular, Theorem 3.1.1 classifying the possible scaling limits of the clock process can be shown to hold for the RTRW on any countable state space where the discrete-time skeleton is a Markov chain satisfying such laws of large numbers for the range and the related quantities.

Our setting generalizes several previous results, let us mention some of them. Mostly, the models studied in the literature involve trapped random walks with some kind of heavy-tailed waiting times, with the aim to show convergence of rescaled clock processes to an α-stable subordinator.

In the so-called *continuous-time random walk* (CTRW), introduced in [MW65], all π_x are deterministically identical heavy-tailed probability distributions, that is for some $\alpha \in (0,1)$ and $c > 0$, and some slowly varying function ℓ,

$$\pi_x[u, \infty) = u^{-\alpha}\ell(u) \text{ as } u \to \infty. \tag{3.1.7}$$

Independently of the nature of the discrete skeleton Y, the clock process is then a sum of i.i.d. heavy-tailed random variables, and it is well known that it converges after normalization to a stable subordinator. The scaling limits of the CTRW were studied in more detail in [MS04].

In the *symmetric Bouchaud trap model* (BTM) the discrete skeleton Y is simple random walk and the π_x are exponential random variables with means m_x that are i.i.d. heavy-tailed random variables satisfying e.g.

$$\mathbb{P}[m_x > u] = cu^{-\alpha}(1 + o(1)) \text{ as } u \to \infty, \tag{3.1.8}$$

The BTM on $\mathbb{Z}^2$ was for the first time studied in [BČM06] where the authors show convergence of the clock process to a stable subordinator and use this to derive aging properties of the model. In [BČ07] it is then shown, in the case of the BTM on $\mathbb{Z}^d$, $d \geq 2$, that the rescaled random walks and clock processes converge jointly to a Brownian motion and a stable subordinator, and therefore the scaling limit of the BTM is the Fractional Kinetics process.

A general model of trapped random walk where the waiting times are exponential with heavy-tailed means as in (3.1.8) is studied in [FM14]. As mentioned above, they consider the discrete skeleton to be an arbitrary random walk on $\mathbb{Z}^d$ satisfying Assumption A. Instead of scaling limits, which require additional restrictions as in our Assumption B, [FM14] focus on the so-called *age process*, which is related to the clock process and describes the 'depth of the trap in which the process stays at a given time'.

Our setting is restricted to the fact that the discrete skeleton Y is independent of the random environment π. There are however interesting models where this is not the case, for example the *asymmetric Bouchaud trap model* (ABTM). In [BČ11] for $d \geq 3$ and in [Mou11] with different methods for $d \geq 5$ it is shown that the scaling limit for ABTM is also Fractional Kinetics. Yet another approach to prove convergence of rescaled clock processes to a stable subordinator is given in [GŠ13], their setting includes the ABTM as a special case.

The majority of the above mentioned previous results are *quenched*, that is the convergence holds for almost every realization of the environment. On the contrary, our results are *annealed*, that is averaged over the environment, but this is not an issue for the classification theorem.

We also believe that when the annealed convergence takes place as in Theorem 3.1.1, then the quenched convergence holds true as well. In high dimensions ($d \geq 5$) this could be proved similarly as in [Mou11], using techniques from [BS02], see also the additional condition in [FM14] under which their annealed result holds quenched. In low dimensions

these methods fail due to many self-intersections of the discrete skeleton. An adaptation of more complicated methods which give the quenched convergence in low dimensions (like the coarse-graining procedure of [BČM06, BČ07] or the techniques of [GŠ13]) to the RTRW seems to be non-trivial and is out of the scope of this paper.

We conclude the introduction by giving sufficient conditions for convergence in both cases of our main theorems. Given the collection of probability measures $\pi = (\pi_x)_{x\in\mathbb{Z}^d}$, let $m_x = \int u\pi_x(du) \in (0,\infty]$ be the mean and $\hat{\pi}_x(\lambda) = \int e^{-\lambda u}\pi_x(du)$ the Laplace transform of π_x. Note that in the next theorem Assumption A is not needed, we only need Y to be non-degenerate.

Theorem 3.1.3. *Let X be RTRW in $d \geq 1$. If $\nu \neq \delta_0$ and the annealed expected waiting time is finite, $\mathbb{E}[m_0] = M < \infty$, then the rescaled clock processes S_N with normalization $a_N = N$ converge in $\mathbb{P}$-distribution on D^1 equipped with the Skorohod J_1-topology to the linear process $\mathcal{S}(t) = Mt$. If in addition Assumption B holds, then the rescaled processes X_N with $a_N = N$ converge in $\mathbb{P}$-distribution on D^d equipped with the Skorohod J_1-topology, and the limit is $\mathcal{X}(t) = B(M^{-1}t)$ as in (i) of Theorem 3.1.2.*

For convergence to Fractional Kinetics we have the following sufficient criterion. In Section 3.5 we will sketch some examples of RTRWs that satisfy this criterion with different functions f.

Theorem 3.1.4. *Let X be RTRW with discrete skeleton Y satisfying Assumption A with slowly varying function ℓ^*. Assume that there is a normalizing sequence $a_N \nearrow \infty$ such that for any positive real number $r > 0$ and a continuous function f,*

$$-\log \mathbb{E}\left[\hat{\pi}_0(\lambda/a_N)^{r\ell^*(N)}\right] = f(r)\lambda^\alpha \frac{\ell^*(N)}{N}(1+o(1)) \text{ as } N \to \infty. \tag{3.1.9}$$

Then the rescaled clock processes S_N with normalization a_N converge in $\mathbb{P}$-distribution on D^1 equipped with the Skorohod M_1-topology to an α-stable subordinator V_α. If in addition Assumption B holds, then the rescaled processes X_N converge in $\mathbb{P}$-distribution on D^1 equipped with the Skorohod J_1-topology, and the limit is the FK process as in (ii) of Theorem 3.1.2.

The rest of this paper is structured as follows. In Section 3.2 we give precise definitions of the model and introduce some notation used through the paper. Theorem 3.1.1 and Theorem 3.1.2 are proved in Sections 3.3 and 3.4 respectively, and Section 3.5 deals with Theorems 3.1.3 and 3.1.4. Finally, in Section 3.6 we prove one technical lemma which is used in the proof of Theorem 3.1.1. In Appendix 3.A we explain how Assumption A on the escape probability implies the laws of large numbers that we mentioned above.

Acknowledgment. The authors would like to thank the referee for carefully reading the manuscript and giving important comments that helped to improve the paper.

3.2. Setting and notation

We start by giving a formal definition of the RTRW. Recall that ν is a probability measure on $\mathbb{Z}^d$ and μ a probability measure on the space of probability measures on $(0,\infty)$. To avoid trivial situations, we assume that $\nu \neq \delta_0$.

Given μ and ν, let $\pi = (\pi_x)_{x\in\mathbb{Z}^d}$ be an i.i.d. sequence of probability measures with marginal μ, and $\xi = (\xi_i)_{i\geq 1}$ an i.i.d. sequence with marginal ν independent of π defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Define

$$Y(n) = \xi_1 + \cdots + \xi_n$$

to be a random walk with step distribution ν and denote by $L(x,n) = \sum_{k=0}^{n} \mathbf{1}_{\{Y(k)=x\}}$ its local time.

Given a realization of π, let further $(\tau_x^i)_{x\in\mathbb{Z}^d, i\geq 1}$ be a collection of independent random variables, independent of ξ, such that every τ_x^i has distribution π_x, defined on the same probability space. The *clock process* of the RTRW, $S : \mathbb{N} \to [0,\infty)$ is then defined by $S(0) = 0$ and

$$S(n) = \sum_{x\in\mathbb{Z}^d} \sum_{i=1}^{L(x,n-1)} \tau_x^i = \sum_{k=0}^{n-1} \tau_{Y(k)}^{L(Y(k),k)} \qquad \text{for } n \geq 1.$$

Finally, we define the RTRW $X = (X(t))_{t\geq 0}$ by

$$X(t) = Y(k) \qquad \text{for } S(k) \leq t < S(k+1),$$

or equivalently

$$X(t) = Y(S^{-1}(t)),$$

where $S^{-1}(t) = \inf\{k \geq 0 : \ S(k) > t\}$ is the right-continuous inverse of S.

Under $\mathbb{P}$, the process X has exactly the law described in the introduction. The random variable τ_x^i denotes the duration of the i-th visit of the vertex x. We refer to $\mathbb{P}$ as *annealed* distribution of X.

We write D^d for the space of the $\mathbb{R}^d$-valued cadlag functions on $[0,\infty)$, and when needed $D^d(J_1)$, $D^d(M_1)$, $D^d(M_1')$ to point out which of Skorohod topologies we use on this space. We refer to [Whi02, Chapter 3.3] for an introduction and [Whi02, Chapters 12–13] for details on these topologies. The less usual M_1'-topology, which plays a role only in Proposition 3.4.1, is a modification of the Skorohod M_1-topology which is convenient for dealing with irregularities at the origin, see [Whi02, Section 13.6.2]. In any case, we will never need to know the actual definitions of these topologies, we only use them when applying results from [Whi02].

It will be useful to introduce the sequence of successive waiting times

$$\tilde{\tau}_k = \tau_{Y(k)}^{L(Y(k),k)}, \qquad k \geq 0.$$

With this notation,

$$S(n) = \sum_{k=0}^{n-1} \tilde{\tau}_k. \tag{3.2.1}$$

We now show that $\tilde{\tau}_k$ is ergodic, which will be used in the proof of Theorem 3.1.3. To this end let $\mathcal{P}'$ be the law on $\Omega' := [0,\infty)^{\mathbb{N}}$ of the sequence $(\tilde{\tau}_k)_{k\geq 0}$ and let θ be the left shift on Ω', $\theta(\tilde{\tau}_1, \tilde{\tau}_2, \dots) = (\tilde{\tau}_2, \tilde{\tau}_3, \dots)$.

Lemma 3.2.1. *The left-shift θ acts ergodically on $(\Omega', \mathcal{P}')$.*

Proof. To show that θ is measure-preserving we follow the environment as 'viewed from the particle'. Namely, let $\Theta : \Omega \to \Omega$ be such that if $\omega' = \Theta(\omega)$, then

$$\begin{aligned}
\xi_i(\omega') &= \xi_{i+1}(\omega), \qquad i \geq 1,\\
\pi_x(\omega') &= \pi_{x+\xi_1(\omega)}(\omega), \qquad x \in \mathbb{Z}^d,\\
\tau^i_x(\omega') &= \begin{cases} \tau^i_{x+\xi_1(\omega)}, & \text{if } x \neq -\xi_1(\omega),\\ \tau^{i+1}_{x+\xi_1(\omega)}, & \text{if } x = -\xi_1(\omega). \end{cases}
\end{aligned}$$

From the independence of ξ from π and τ, and from the i.i.d. properties of π and $\tau^{\cdot}_x$ for every x, it is easy to see that the law of $X \circ \Theta$ agrees with the law of X, that is Θ is $\mathbb{P}$-preserving. Since, in addition, $\tilde{\tau}(\Theta(\omega)) = \theta(\tilde{\tau}(\omega))$ and $\mathcal{P}' = \mathbb{P} \circ \tilde{\tau}^{-1}$, this implies that θ is $\mathcal{P}'$-preserving.

To prove the ergodicity, we show that θ is strongly mixing. To this end it is sufficient to verify that

$$\left|\mathcal{P}'[\theta^{-n}A \cap B] - \mathcal{P}'[A]\mathcal{P}'[B]\right| \xrightarrow{n\to\infty} 0 \tag{3.2.2}$$

for all cylinder sets $A = \{\tilde{\tau}_i \in A_i,\ i \in I\}$, $B = \{\tilde{\tau}_j \in B_j,\ j \in J\}$, where $I, J \subset \mathbb{N}$ are finite sets and $A_i, B_j \subset \mathbb{R}$ are Borel sets, see e.g. [Pet83, Prop 2.5.3]. Fix two such sets A and B and define the event $\mathcal{I}_n(A,B) = \{Y(i+n) = Y(j) \text{ for some } i \in I, j \in J\}$. Denote by $\mathcal{G}^m_k = \sigma(\xi_{k+1}, \xi_{k+2}, \dots, \xi_m)$ the σ-algebra generated by the steps made by the random walk between time k and m, and write $\mathcal{G}_{(n)} = \mathcal{G}_0^{(\max I+n)\vee \max J}$, $\mathcal{G}_{(I,n)} = \mathcal{G}^{\max I+n}_{\min I+n}$, $\mathcal{G}_{(J)} = \mathcal{G}^{\max J}_{\min J}$. By the independence structure of the τ^i_x we have that

$$\mathcal{P}'\left[\theta^{-n}A \cap B \mid \mathcal{G}_{(n)}\right] \mathbf{1}_{\mathcal{I}_n(A,B)^c} = \mathcal{P}'\left[\theta^{-n}A \mid \mathcal{G}_{(n)}\right] \mathcal{P}'\left[B \mid \mathcal{G}_{(n)}\right] \mathbf{1}_{\mathcal{I}_n(A,B)^c}. \tag{3.2.3}$$

Moreover,

$$\begin{aligned}
\mathcal{P}'\left[\theta^{-n}A \mid \mathcal{G}_{(n)}\right] &= \mathcal{P}'\left[\theta^{-n}A \mid \mathcal{G}_{(I,n)}\right],\\
\mathcal{P}'\left[B \mid \mathcal{G}_{(n)}\right] &= \mathcal{P}'\left[B \mid \mathcal{G}_{(J)}\right].
\end{aligned} \tag{3.2.4}$$

By the independence of the ξ_k, as soon as $\max J < \min I + n$ the right hand sides of the above two equations are independent. Denote by $\mathcal{E}'$ the expectation corresponding to $\mathcal{P}'$. Using (3.2.3), (3.2.4) and the independence of the two right hand sides in (3.2.4), and the fact that $\mathcal{P}'[\theta^{-n}A] = \mathcal{P}'[A]$, we have for n large enough,

$$\begin{aligned}
\mathcal{P}'\left[\theta^{-n}A \cap B\right] &= \mathcal{E}'\left[\mathcal{P}'\left[\theta^{-n}A \cap B \mid \mathcal{G}_{(n)}\right]\left(\mathbf{1}_{\mathcal{I}_n(A,B)} + \mathbf{1}_{\mathcal{I}_n(A,B)^c}\right)\right]\\
&= \mathcal{E}'\left[\mathcal{P}'\left[\theta^{-n}A \mid \mathcal{G}_{(n)}\right]\mathcal{P}'\left[B \mid \mathcal{G}_{(n)}\right]\mathbf{1}_{\mathcal{I}_n(A,B)^c}\right] + O\left(\mathcal{P}'[\mathcal{I}_n(A,B)]\right)\\
&= \mathcal{E}'\left[\mathcal{P}'\left[\theta^{-n}A \mid \mathcal{G}_{(I,n)}\right]\mathcal{P}'\left[B \mid \mathcal{G}_{(J)}\right]\right] + O\left(\mathcal{P}'[\mathcal{I}_n(A,B)]\right)\\
&= \mathcal{P}'[\theta^{-n}A]\mathcal{P}'[B] + O\left(\mathcal{P}'[\mathcal{I}_n(A,B)]\right)\\
&= \mathcal{P}'[A]\mathcal{P}'[B] + O\left(\mathcal{P}'[\mathcal{I}_n(A,B)]\right)
\end{aligned} \tag{3.2.5}$$

But

$$\mathcal{P}'[\mathcal{I}_n(A,B)] \leq \sum_{i\in I, j\in J} \mathbb{P}[Y(i+n) = Y(j)],$$

and for n large enough the Markov property for Y implies that $\mathbb{P}[Y(i+n) = Y(j)] = \mathbb{P}[Y(i+n-j) = 0]$, which tends to 0 as $n \to \infty$ for every (non-degenerate) random walk, see e.g. [Spi76, P7.6]. Since I and J are finite, $\mathcal{P}'[\mathcal{I}_n(A,B)] \to 0$ as $N \to \infty$, and thus (3.2.2) follows from (3.2.5). □

3.3. Proof of Theorem 3.1.1

In this section we prove Theorem 3.1.1. In the next two lemmas we study the properties of the limit clock process $\mathcal{S}$.

Lemma 3.3.1. *If the random walk Y satisfies Assumption A and the rescaled clock processes S_N converge to $\mathcal{S}$ in the way as stated in Theorem 3.1.1, then the limit clock process $\mathcal{S}$ has stationary increments and is self-similar with index $\rho > 0$, i.e. $\mathcal{S}(t) \stackrel{d}{=} \lambda^\rho \mathcal{S}(t/\lambda)$. Moreover, the normalizing sequence is of the form $a_N = N^\rho \ell(N)$, for the same $\rho > 0$ and some slowly varying function ℓ.*

Proof. Stationarity of the increments follows immediately from (3.2.1) and the stationarity of the sequence $\tilde{\tau}$ of successive waiting times which was proved in Lemma 3.2.1. To see the self-similarity, fix $\lambda > 0$ and t such that condition (3.1.2) holds for t and t/λ, and $\mathcal{S}(t)$, $\mathcal{S}(\lambda t)$ are not identically zero, which is possible thanks to (3.1.3). Then,

$$\mathcal{S}(t) = \lim_{N\to\infty} \frac{1}{a_N} S(Nt) = \lim_{N\to\infty} \frac{a_{\lambda N}}{a_N} \frac{1}{a_{\lambda N}} S\left(\lambda N \frac{t}{\lambda}\right) \stackrel{d}{=} \mathcal{S}\left(\frac{t}{\lambda}\right) \lim_{N\to\infty} \frac{a_{\lambda N}}{a_N}.$$

Since $\mathcal{S}(t)$ and $\mathcal{S}(t/\lambda)$ are not identically zero, it follows that $\frac{a_{\lambda N}}{a_N}$ must converge to some constant $c(\lambda)$, yielding the scale invariance. Moreover, elementary results of the theory of regularly varying functions imply that $c(\lambda) = \lambda^\rho$ for some $\rho \in \mathbb{R}$, and that a_N is regularly varying of index ρ, that is $a_N = N^\rho \ell(N)$ for some slowly varying function $\ell(N)$. Note that $\rho > 0$ since $\rho = 0$ would imply $\lim_{N\to\infty} \frac{a_{\lambda N}}{a_N} = 1$, hence $\mathcal{S}(t) \stackrel{d}{=} \mathcal{S}(t/\lambda)$, which violates the non-triviality assumption (3.1.3). □

Lemma 3.3.2. *If the random walk Y satisfies Assumption A and the rescaled clock processes S_N converge to $\mathcal{S}$ in the way as stated in Theorem 3.1.1, then the limit clock process $\mathcal{S}$ has independent increments.*

Let us postpone the proof of this lemma and show Theorem 3.1.1 first.

Proof of Theorem 3.1.1. By Lemmas 3.3.1 and 3.3.2, $\mathcal{S}$ has stationary and independent increments and is self-similar with index ρ. From this, the fact that $\mathcal{S} \geq 0$ and the non-triviality assumption (3.1.3) it follows that either $\rho = 1$ and $\mathcal{S}(t) = Mt$ for some $M \in (0, \infty)$, or $\rho > 1$ and $\mathcal{S}$ is an increasing α-stable Lévy process with $\alpha = \rho^{-1} \in (0, 1)$, that is an α-stable subordinator. Lemma 3.3.1 gives the normalizing sequence a_N as claimed. □

In order to show Lemma 3.3.2 we need three technical lemmas which are consequences of laws of large numbers for the range-like objects related to the random walk Y, as mentioned in the introduction.

The first lemma states that for any given times $0 = t_0 < t_1 < \cdots < t_n = t$, the number of vertices visited by the random walk Y in more than one of the time intervals $[\lfloor t_{i-1} N \rfloor, \lfloor t_i N \rfloor - 1]$ is small. To this end, let

$$R(k) = \{Y(0), \ldots, Y(k-1)\}$$

be the range of the random walk Y at time $k-1$, R_N^i be the 'range between $t_{i-1}N$ and $t_i N$',

$$R_N^i = \{Y(k) : \ k = \lfloor N t_{i-1} \rfloor, \ldots, \lfloor N t_i \rfloor - 1\},$$

O_N^i be the set of the points visited only in this time interval,

$$O_N^i = \left\{x \in R_N^i : \ x \notin R_N^j \text{ for all } j \neq i\right\},$$

and M_N^i be the set of points visited in more than one of them, $M_N^i = R_N^i \setminus O_N^i$.

Lemma 3.3.3. *If Y verifies Assumption A, then for any choice of time points $0 = t_0 < t_1 < \cdots < t_n = t$,*

$$\lim_{N\to\infty} |M_N^i| \frac{\ell^*(N)}{N} = 0 \qquad \text{in } \mathbb{P}\text{-probability for all } i = 1, \ldots, n.$$

Proof. The size of the sets O_N^i can be bounded by

$$|R(\lfloor Nt \rfloor)| - \left| \bigcup_{j\neq i} R_N^j \right| \leq |O_N^i| \leq |R(\lfloor Nt_i \rfloor)| - |R(\lfloor Nt_{i-1} \rfloor)|. \tag{3.3.1}$$

Applying the laws of large numbers from Lemma 3.A.1 and the Markov property at times $\lfloor Nt_i \rfloor$, it follows that for every $i = 1, \ldots, n$,

$$|R(\lfloor Nt_i \rfloor)| \frac{\ell^*(N)}{Nt_i} \xrightarrow{N\to\infty} 1, \quad \text{and} \quad \left| \bigcup_{j\neq i} R_N^j \right| \frac{\ell^*(N)}{N(t_n - t_i + t_{i-1})} \xrightarrow{N\to\infty} 1$$

in probability. Inserting this into (3.3.1) yields a law of large numbers for $|O_N^i|$,

$$|O_N^i| \frac{\ell^*(N)}{N(t_i - t_{i-1})} \xrightarrow{N\to\infty} 1$$

in probability. By Lemma 3.A.1 and the Markov property again, $|R_N^i|$ satisfies the same law of large numbers as $|O_N^i|$. Using $|M_N^i| = |R_N^i| - |O_N^i|$ the claim follows. □

The second lemma will help to control the contribution of frequently visited vertices to the clock process. Fix $t > 0$, and for $K > 0$ define the set of 'frequently visited vertices'

$$\mathcal{F}_{N,K} = \left\{x : L(x, \lfloor Nt \rfloor - 1) \geq K\ell^*(N)\right\}. \tag{3.3.2}$$

Let $F_{N,K}$ be the 'number of visits to $\mathcal{F}_{N,K}$'

$$F_{N,K} = \sum_{x\in\mathcal{F}_{N,K}} L(x, \lfloor Nt \rfloor - 1). \tag{3.3.3}$$

Lemma 3.3.4. *If Y verifies Assumption A, then there is a constant $c > 0$ such that for every $\varepsilon > 0$ and fixed $t > 0$*

$$\mathbb{P}[F_{N,K} \geq \epsilon Nt] \leq \epsilon \qquad \text{for all } N \text{ large enough,}$$

with

$$K = K(\epsilon) = -c \log\left(\epsilon^2\right). \tag{3.3.4}$$

Proof. We claim that for ϵ small enough and N large enough,

$$\mathbb{E}\left[F_{N,K}\right] \leq \epsilon^2 Nt. \tag{3.3.5}$$

Applying the Markov inequality then yields the desired result.

To show (3.3.5), let $\psi_k = \mathbf{1}_{\{Y(l)\neq Y(k)\,\forall l<k\}}$ be the indicator of the event that a 'new' vertex is found at time k. Then

$$F_{N,K} = \sum_{k=0}^{\lfloor Nt\rfloor-1} \psi_k L(Y_k, \lfloor Nt\rfloor - 1)\mathbf{1}_{\{L(Y_k,\lfloor Nt\rfloor-1)\geq K\ell^*(N)\}}.$$

Using the Markov property and the fact that $L(Y_k, \lfloor Nt\rfloor - 1)$ is stochastically dominated by $L(0, \lfloor Nt\rfloor - 1)$,

$$\mathbb{E}[F_{N,K}] \leq \mathbb{E}\big[L(0, \lfloor Nt\rfloor - 1)\mathbf{1}_{\{L(0,\lfloor Nt\rfloor-1)\geq K\ell^*(N)\}}\big] \sum_{k=0}^{\lfloor Nt\rfloor-1} \mathbb{E}[\psi_k]. \tag{3.3.6}$$

By (3.A.2), $\sum_{k=0}^{\lfloor Nt\rfloor-1}\mathbb{E}[\psi_k] = \mathbb{E}[|R(\lfloor Nt\rfloor)|] = Nt/\ell^*(\lfloor Nt\rfloor)(1+o(1))$. On the other hand, denoting by $\tilde{H}_0$ the first return time of Y to 0, for every $k \geq 1$

$$\mathbb{P}[L(0, \lfloor Nt\rfloor - 1) \geq k] \leq \big(\mathbb{P}[\tilde{H}_0 \leq \lfloor Nt\rfloor]\big)^{k-1} = (1 - \ell^*(\lfloor Nt\rfloor)^{-1})^{k-1},$$

and thus $L(0, \lfloor Nt\rfloor - 1)$ is stochastically dominated by a geometric random variable with parameter $1/\ell^*(\lfloor Nt\rfloor)$. If G is a geometric variable with parameter p, then for every $M \in \mathbb{N}$,

$$\mathbb{E}[G\mathbf{1}_{\{G\geq M\}}] = (1-p)^{M-1}\Big(M - 1 + \frac{1}{p}\Big).$$

Hence,

$$\begin{aligned}&\mathbb{E}\big[L(0, \lfloor Nt\rfloor - 1)\mathbf{1}_{\{L(0,\lfloor Nt\rfloor-1)\geq K\ell^*(N)\}}\big]\\ &\quad\leq \left(1 - \frac{1}{\ell^*(\lfloor Nt\rfloor)}\right)^{K\ell^*(N)-1}\big(K\ell^*(N) - 1 + \ell^*(\lfloor Nt\rfloor)\big)\end{aligned} \tag{3.3.7}$$

and the claim (3.3.5) follows by inserting K as in (3.3.4), using the slow variation of ℓ^* and combining (3.3.6), (3.3.7). □

The last of the technical lemmas allows to control the influence of an arbitrary subset of waiting times to the sum of all waiting times if the subset is small.

Lemma 3.3.5. *Let $\mathcal{B}_N \subset \{0, 1, \ldots, \lfloor Nt\rfloor-1\}$ be a random set, depending on the trajectory of the random walk Y up to time $\lfloor Nt\rfloor - 1$ only. If Assumption A holds, then for every $t > 0$ and $\delta > 0$,*

$$\lim_{\epsilon\to 0}\lim_{N\to\infty}\mathbb{P}\left[\sum_{k\in\mathcal{B}}\tilde{\tau}_k \geq \delta S(\lfloor Nt\rfloor),\ |\mathcal{B}| \leq \epsilon N\right] = 0.$$

The proof of this lemma is surprisingly lengthy and is therefore postponed to Section 3.6. The main source of complications comes from the fact that we cannot make any assumptions on the moments of the waiting times τ_x^i. It is also essential to use some properties of the random walk Y, as it is easy to construct counterexamples to the lemma when τ_x^i are not summed along the trajectory of Y.

With the above three lemmas we can now show Lemma 3.3.2.

Proof of Lemma 3.3.2. Fix times $0 = t_0 < t_1 < \cdots < t_n = t$. Consider first the following alternative construction of the clock process S. On the same space $(\Omega, \mathcal{F}, \mathbb{P})$, let for every $x \in \mathbb{Z}^d$ independently $(\pi_{x,j})_{j=1,\dots,n}$ be i.i.d. μ-distributed probability measures, and given a realization of these measures, let $(\tau^i_{x,j})_{x\in\mathbb{Z}^d, j,i\geq 1}$ be independent random variables such that every $\tau^i_{x,j}$ has distribution $\pi_{x,j}$. For every vertex $x \in \mathbb{Z}^d$, let $j(x)$ be be such that the first visit to x occurs in the time interval $[\lfloor Nt_{j(x)-1}\rfloor, \lfloor Nt_{j(x)}\rfloor - 1]$. Define a new process $S' : \mathbb{N} \to [0,\infty)$ by

$$S'(k) = \sum_{x\in\mathbb{Z}^d} \sum_{i=1}^{L(x,k-1)} \tau^i_{x,j(x)}.$$

One can think of choosing the distributions π_x at the time of the first visit in x according to the time interval in which this first visit occurs. Constructed in this way, S' has clearly the same distribution as the original clock process S.

We now define an approximation $\tilde{S}$ of S' which collects time $\tau^{L(x,k)}_{x,j}$ whenever at a vertex x at time $k \in [\lfloor Nt_{j-1}\rfloor, \lfloor Nt_j\rfloor - 1]$,

$$\tilde{S}(m) = \sum_{j=1}^{n} \sum_{k=\lfloor Nt_{j-1}\rfloor}^{(m\wedge\lfloor Nt_j\rfloor)-1} \tau^{L(Y(k),k)}_{Y(k),j}. \tag{3.3.8}$$

$\tilde{S}$ can be viewed as the clock for which the whole environment π is being refreshed at all times $\lfloor Nt_j\rfloor$. Therefore, by the independence structure of the $\tau^i_{x,j}$'s, the increments $(\tilde{S}(\lfloor Nt_j\rfloor) - \tilde{S}(\lfloor Nt_{j-1}\rfloor))_{j=1,\dots,n}$ are mutually independent. In addition, for every j, the increment $\tilde{S}(\lfloor Nt_j\rfloor) - \tilde{S}(\lfloor Nt_{j-1}\rfloor)$ is independent of the increments $\{\xi_k : k \notin [\lfloor Nt_{j-1}\rfloor, \lfloor Nt_j\rfloor - 1]\}$ of the random walk Y.

To conclude the proof it is now sufficient to show that for all $j = 1, \dots, n$ and every $\delta > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left[\left|\tilde{S}(\lfloor Nt_j\rfloor) - S'(\lfloor Nt_j\rfloor)\right| > \delta S'(\lfloor Nt_j\rfloor)\right] = 0. \tag{3.3.9}$$

This implies that the limit process $\mathcal{S}$ has independent increments. Indeed, note that (3.3.9) readily implies $\frac{\tilde{S}(\lfloor Nt_j\rfloor)}{S'(\lfloor Nt_j\rfloor)} \to 1$ in $\mathbb{P}$-probability for all j. This means that whenever $\frac{1}{a_N} S'(\lfloor Nt_j\rfloor) \xrightarrow{d} \mathcal{S}(t_j)$, then also $\frac{1}{a_N}\tilde{S}(\lfloor Nt_j\rfloor) \xrightarrow{d} \mathcal{S}'(t_j)$, and therefore the increments $(\mathcal{S}(t_j) - \mathcal{S}(t_{j-1}))_{j=1,\dots,n}$ are independent, whenever (3.1.2) is satisfied for the times t_j. By easy approximation arguments this also holds for the at most countably many t_j's that do not satisfy (3.1.2). Since the times t_j are chosen arbitrarily, it follows that the process $\mathcal{S}$ has independent increments.

In order to show (3.3.9), note that the difference of $\tilde{S}(\lfloor Nt_j\rfloor)$ and $S'(\lfloor Nt_j\rfloor)$ originates in the waiting times in vertices visited in multiple time intervals. Recalling the sets M^j_N

from Lemma 3.3.2,

$$\left|\tilde{S}(\lfloor Nt_j\rfloor) - S'(\lfloor Nt_j\rfloor)\right| \leq \sum_{l=1}^{j} \sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_{x,l}^i.$$

It is therefore sufficient to show that for each $j = 1,\dots,n$ and $1 \leq l \leq j$, and every $\delta > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left[\sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_{x,l}^i \geq \delta S'(\lfloor Nt_j\rfloor)\right] = 0.$$

The probability above is bounded by

$$\mathbb{P}\left[(1+\delta)\left(\sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_{x,l}^i\right) \geq \delta\left(\sum_{x\in R(\lfloor Nt_j\rfloor)\setminus M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_{x,j(x)}^i + \sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_{x,l}^i\right)\right].$$

Note that, by definition of the random variables $\tau_{x,j}^i$, requiring the above probability to tend to 0 as $N\to\infty$ is the same as requiring

$$\lim_{N\to\infty} \mathbb{P}\left[\sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_{x}^i \geq \delta S(\lfloor Nt_j\rfloor)\right] = 0, \tag{3.3.10}$$

for each $j = 1,\dots,n$ and $1 \leq l \leq j$, and every $\delta > 0$, where here S is the original clock process, i.e. the sum of the τ_x^i's which have distributions π_x.

Fix $\epsilon > 0$ small, set K as in (3.3.4), recall the definition of $\mathcal{F}_{N,K}$ from (3.3.2) (with t_j instead of t), and write

$$\begin{aligned}
&\mathbb{P}\left[\sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_x^i \geq \delta S(\lfloor Nt_j\rfloor)\right] \\
&\leq \mathbb{P}\left[\sum_{x\in M_N^l\setminus\mathcal{F}_{N,K}} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_x^i \geq \frac{\delta}{2} S(\lfloor Nt_j\rfloor)\right] + \mathbb{P}\left[\sum_{x\in M_N^l\cap\mathcal{F}_{N,K}} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_x^i \geq \frac{\delta}{2} S(\lfloor Nt_j\rfloor)\right].
\end{aligned} \tag{3.3.11}$$

By Lemma 3.3.3 we can choose N large enough such that $\mathbb{P}[|M_N^l| > \epsilon N/\ell^*(N)] \leq \epsilon$. Then the first term on the right-hand side of (3.3.11) is bounded by

$$\begin{aligned}
&\mathbb{P}\left[\sum_{x\in M_N^l\setminus\mathcal{F}_{N,K}} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_x^i \geq \frac{\delta}{2} S(\lfloor Nt_j\rfloor),\ |M_N^l| \leq \epsilon N/\ell^*(N)\right] + \epsilon \\
&= \mathbb{P}\left[\sum_{k\in\mathcal{B}_1} \tilde{\tau}_k \geq \frac{\delta}{2} S(\lfloor Nt_j\rfloor),\ |M_N^l| \leq \epsilon N/\ell^*(N)\right] + \epsilon.
\end{aligned}$$

Here $\mathcal{B}_1$ is the set of all times where a vertex in $M_N^l \setminus \mathcal{F}_{N,K}$, i.e. with $L(x,\lfloor Nt_j\rfloor - 1) \leq K\ell^*(N)$ is visited. But if $|M_N^l| \leq \epsilon N/\ell^*(N)$, then $|\mathcal{B}_1| \leq \epsilon KN$. Since $\epsilon K \to 0$ as $\epsilon \to 0$ by

the definition of K, we can apply Lemma 3.3.5 to get that the first term on the right-hand side of (3.3.11) converges to 0 when $N \to \infty$ and then $\varepsilon \to 0$.

The second term on the right-hand side of (3.3.11) can be bounded similarly. Recalling $F_{N,K}$ (for t_j) from (3.3.3), it is bounded from above by

$$\mathbb{P}\left[F_{N,K} \geq \epsilon N\right] + \mathbb{P}\left[\sum_{k \in \mathcal{B}_F} \tilde{\tau}_k \geq \frac{\delta}{2} S(\lfloor N t_j \rfloor), F_{N,K} \leq \epsilon N\right].$$

Here $\mathcal{B}_F$ is the set of times where a frequently visited vertex is visited, i.e. $|\mathcal{B}_F| = F_{N,K}$. Applying Lemma 3.3.4 to the first term and Lemma 3.3.5 to the second, this converges to zero as $N \to \infty$ and $\varepsilon \to 0$, and (3.3.10) follows. This finishes the proof of the lemma. □

3.4. Proof of Theorem 3.1.2

The goal of this section is to prove the classification theorem for the RTRW, Theorem 3.1.2. This will be done using Theorem 3.1.1. At first we should however show that the assumptions of Theorem 3.1.2 allow to verify the hypotheses of Theorem 3.1.1.

Proposition 3.4.1 ($d \geq 1$)**.** *Let X_N be as in* (3.1.5)*. Suppose that Assumption B holds and that X_N converge in the sense of Theorem 3.1.2. Then the clock processes S_N, defined as in* (3.1.2)*, converge in $\mathbb{P}$-distribution on $D^d(M_1')$ to some process $\mathcal{S}$. If $\mathcal{S}(0) = 0$, then the convergence holds with respect to the Skorohod M_1-topology.*

We first use this proposition to show Theorem 3.1.2.

Proof of Theorem 3.1.2. By Proposition 3.4.1, S_N converge to some process $\mathcal{S}$ in distribution on $D^d(M_1')$. This convergence implies the convergence of $S_N(t)$ to $\mathcal{S}(t)$ for all but countably many $t \in [0, \infty)$, cf. [Whi02, Theorem 11.6.6 and Corollary 12.2.1]. The non-triviality assumption (3.1.6) implies (3.1.3). We can thus apply Theorem 3.1.1. By this theorem there are only two possibilities, either $\mathcal{S}(t) = Mt$ or $\mathcal{S}(t) = V_\alpha(t)$. Since in both cases $\mathcal{S}(0) = 0$, the convergence of S_N actually holds in the M_1-topology.

The possible limits $\mathcal{S}$ are in the subspace $D^1_{u,\uparrow\uparrow}$ of unbounded strictly increasing functions from $[0, \infty)$ to $\mathbb{R}$, and their inverses are continuous. By [Whi02, Corollary 13.6.4], the inverse map from the space $D^1_{u,\uparrow}(M_1)$ of unbounded non-decreasing functions to $D^1(J_1)$ is continuous at $D^1_{u,\uparrow\uparrow}$, therefore S_N^{-1} converge to $\mathcal{S}^{-1}$ in $\mathbb{P}$-distribution on $D^1(J_1)$. Moreover, the rescaled random walks Y_N converge in $\mathbb{P}$-distribution on $D^d(J_1)$ to a standard d-dimensional Brownian motion B.

To proceed, we need to show that B and the limit clock process $\mathcal{S}$ are independent. This is trivial for the case $\mathcal{S}(t) = Mt$, so we may assume that $\mathcal{S} = V_\alpha$. We will use [Kal02, Lemma 15.6] which applied to our situation states that if $B, \mathcal{S}$ are such that $B(0) = \mathcal{S}(0) = 0$ and the process $(B, \mathcal{S}) \in D^{d+1}$ has independent increments and no fixed jumps, $\mathcal{S}$ is a.s. a step process and $\Delta B \cdot \Delta \mathcal{S} = 0$ a.s., then B and $\mathcal{S}$ are independent. The only assumption that remains to be verified is that $(B, \mathcal{S})$ has jointly independent increments.

For fixed times $0 = t_0 < t_1 < \cdots < t_n = t$, consider the version $\tilde{S}(\lfloor Nt \rfloor)$ from (3.3.8) in the proof of Lemma 3.3.2. We have seen that every increment $\tilde{S}(\lfloor Nt_i \rfloor) - \tilde{S}(\lfloor Nt_{i-1} \rfloor)$ is independent of the increments $\{\xi_k : \ k \notin [\lfloor Nt_{j-1} \rfloor, \lfloor Nt_j \rfloor - 1]\}$ of the random walk

Y. Since there is such a version of $\tilde{S}(\lfloor Nt \rfloor)$ for every choice of times t_j, and every such $\tilde{S}(\lfloor Nt \rfloor)$ converges to $\mathcal{S}$ after normalization, we obtain that for the limit $\mathcal{S}$ every increment $\mathcal{S}(t) - \mathcal{S}(s)$ is independent of $\{B(u) : u \notin [t, s]\}$. Since both B and $\mathcal{S}$ have independent increments, this implies that $(B, \mathcal{S})$ has jointly independent increments. Applying [Kal02, Lemma 15.6] it follows that the two limit processes B and $\mathcal{S}$, and thus also B and $\mathcal{S}^{-1}$ are independent.

It follows that (Y_N, S_N^{-1}) converge in distribution on $D^d(J_1) \times D^1_{u,\uparrow}(J_1)$ to $(B, \mathcal{S}^{-1})$. By [Whi02, Theorem 13.2.2], the composition map from $D^d(J_1) \times D^1_{u,\uparrow}(J_1)$ to $D^d(J_1)$ taking $(y(t), s(t))$ to $y(s(t))$ is continuous at (y, s) if y is continuous and s non-decreasing. From this we conclude that the compositions $X_N(t) = Y_N(S_N^{-1}(t))$ converge in distribution on $D^d(J_1)$ to $B(\mathcal{S}^{-1}(t))$ as required. □

For the proof of Proposition 3.4.1 we will relate the clock process S to the quadratic variation process of the RTRW X and then apply [JS03, Corollary VI.6.29] which states that under some conditions, whenever a sequence of processes converges in distribution, then so does the sequence of their quadratic variations.

We need some definitions first. For a d-dimensional pure-jump process Z, let $Z^{(i)}$ denote the i-th coordinate of Z, and let $\Delta Z^{(i)}(t) = Z^{(i)}(t) - Z^{(i)}(t-)$ be the jump size of $Z^{(i)}$ at time t. The quadratic variation process $[Z, Z]_t$ is a $d \times d$ matrix-valued process, where the (i, j)-th entry is the quadratic covariation of the i-th and j-th coordinate of Z, which is

$$[Z^{(i)}, Z^{(j)}]_t = \sum_{0<s\leq t} \Delta Z^{(i)}(s) \Delta Z^{(j)}(s).$$

We proceed by relating the inverse S_N^{-1} of the clock process to the quadratic variation process of X_N.

Lemma 3.4.2. *Under Assumption B, let $[X_N, X_N]_t$ be the quadratic variation process of X_N, and define $\sigma^2 = \mathbb{E}[|\mathcal{A}\xi_j|^2]$ (recall (3.1.4) and (3.1.5) for the notation). Then for every $t > 0$,*

$$\frac{\operatorname{trace}[X_N, X_N]_t}{\sigma^2 S_N^{-1}(t)} \xrightarrow{N\to\infty} 1 \text{ in } \mathbb{P}\text{-probability.}$$

Proof. Easy computation yields

$$\operatorname{trace}[X_N, X_N]_t = \sum_{i=1}^{d} \sum_{0<s\leq t} (\Delta X_N^{(i)}(s))^2 = \frac{1}{N} \sum_{j \leq S^{-1}(a_N t)} |\mathcal{A}\xi_j|^2.$$

The process S^{-1} has increments of size 1, and since the times between increments are a.s. finite, $S^{-1}(a_N t) \nearrow \infty$ a.s. as $N \to \infty$. Therefore, since $\sigma^2 = \mathbb{E}[|\mathcal{A}\xi_j|^2] < \infty$ by Assumption B, the law of large numbers implies

$$\mathbb{P}\left[\left|\frac{N}{S^{-1}(a_N t)} \operatorname{trace}[X_N, X_N]_t - \sigma^2\right| > \epsilon\right] \longrightarrow 0 \text{ as } N \to \infty \text{ for every } \epsilon > 0.$$

Noting that $\frac{1}{N} S^{-1}(a_N t) = S_N^{-1}(t)$ finishes the proof. □

We now check that the assumptions for [JS03, Corollary VI.6.29] are fulfilled.

Lemma 3.4.3. *If Assumption B holds, then the rescaled processes X_N are local martingales (with respect to the natural filtration $\sigma(X_N(t))$) with bounded increments.*

Proof. The increments are bounded by Assumption B. The local martingale property is unaffected by linear scaling, it is hence sufficient to prove it for the process X.

We show that the sequence of stopping times $\sigma_l = S(l)$, $l \geq 1$, is a localizing sequence for X, i.e. we show that $(X(t \wedge \sigma_l))_{t \geq 0}$ is a martingale for every $l \geq 1$.

We introduce the filtration $\mathcal{F}_t = \sigma(Y(k), S(k) : \ k \leq t)$. Obviously, Y is an $\mathcal{F}$-martingale, and $S^{-1}(t)$ is an $\mathcal{F}$-stopping time for every $t \geq 0$, with $S^{-1}(t) \geq S^{-1}(s)$ for $t \geq s$. The natural filtration for X, $\mathcal{G}_t = \sigma(X(t))$ satisfies $\mathcal{G}_t = \mathcal{F}_{S^{-1}(t)}$ and is right-continuous (see [Kal02, Proposition 7.9]). The sequence of random variables σ_l is indeed an increasing sequence of $\mathcal{G}$-stopping times (σ_l is the time at which the process X jumps for the l-th time). Moreover, by definition $S^{-1}(t \wedge \sigma_l) = S^{-1}(t) \wedge S^{-1}(\sigma_l) \leq S^{-1}(\sigma_l) = l$. Applying Doob's optional sampling theorem (see e.g. [Kal02, Theorem 7.12]) to the discrete-time martingale Y and the bounded stopping time $S^{-1}(t \wedge \sigma_l)$, we obtain

$$\mathbb{E}\left[X(t \wedge \sigma_l) \mid \mathcal{G}_s\right] = \mathbb{E}\left[Y\left(S^{-1}(t \wedge \sigma_l)\right) \mid \mathcal{F}_{S^{-1}(s)}\right] = Y\left(S^{-1}(t \wedge \sigma_l) \wedge S^{-1}(s)\right) = X(s \wedge \sigma_l).$$

This completes the proof. □

We can now prove Proposition 3.4.1.

Proof of Proposition 3.4.1. By Lemma 3.4.3, X_N are local martingales with bounded increments. [JS03, Corollary VI.6.29] then implies that the quadratic variation processes $[X_N, X_N]_t$ converge component-wise on $D^1(J_1)$ to the quadratic variation process $[\mathcal{X}, \mathcal{X}]_t$ of $\mathcal{X}$. Since all jumps of the processes $[X_N^{(i)}, X_N^{(i)}]_t$, $i = 1, \ldots, d$, are positive, [Whi02, Theorem 12.7.3 (continuity of addition at limits with jumps of common sign)] yields that trace$[X_N, X_N]_t$ converges to some non-decreasing process in $D^1(M_1)$. From Lemma 3.4.2 it then follows that the inverses S_N^{-1} of the rescaled clock processes converge to some non-decreasing process $\mathcal{S}^{-1}(t)$ in $D^1(M_1)$.

For non-decreasing functions $x \in D^1$ the right-continuous inverse satisfies $(x^{-1})^{-1} = x$, and thus $S_N = (S_N^{-1})^{-1}$. Hence, by [Whi02, Theorem 13.6.1], which ensures the continuity of the inverse operation, S_N converges to $\mathcal{S}$ in $D^1(M_1)$ provided that $\mathcal{S}(0) = (\mathcal{S}^{-1})^{-1}(0) = 0$.

If we do not know whether $\mathcal{S}(0) = 0$, this theorem does not apply. This issue can be solved by weakening the topology from M_1 to M_1' (see [Whi02, Section 13.6.2] for details). In particular, [Whi02, Theorem 13.6.2] yields that S_N converge to $\mathcal{S}$ in distribution in $D^1(M_1')$. □

3.5. Proofs of sufficiency criteria

Theorem 3.1.3, giving a sufficient criterion for convergence to Brownian motion, is an immediate consequence of the ergodicity of the sequence of successive waiting times.

Proof of Theorem 3.1.3. Consider $\tilde{\tau} = (\tilde{\tau}_k)_{k \geq 0}$ and let $\theta : \ \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ be the left-shift along the sequence, which by Lemma 3.2.1 acts ergodically along $\tilde{\tau}$.

If $\mathbb{E}[\tilde{\tau}_0] = M$ is finite, the function $f(\tilde{\tau}) = \tilde{\tau}_0$ is integrable, and we can apply the ergodic theorem to f to get

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} S(\lfloor Nt \rfloor) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{\lfloor Nt\rfloor -1} \tilde{\tau}_k = \lim_{N\to\infty} t \frac{1}{Nt} \sum_{k=0}^{\lfloor Nt\rfloor -1} f(\theta^k(\tilde{\tau})) \\ = t\mathbb{E}\left[f(\tilde{\tau})\right] = Mt \text{ almost surely.}\end{aligned}$$

Thus we have that the rescaled clock processes S_N converge in distribution on $D^1(J_1)$ to Mt, where the normalization is $a_N = N$. If additionally Assumption B holds, using the same arguments as in the proof of Theorem 3.1.2 we conclude that the X_N converge and the limit $\mathcal{X}$ is as in case (i) of Theorem 3.1.2. □

Before starting the proof of Theorem 3.1.4, which deals with the convergence to the Fractional Kinetics, we briefly sketch some examples that illustrate how different functions f in condition (3.1.9) arise.

First, consider the CTRW defined in (3.1.7). The waiting times τ_x^i of this model lie in the domain of attraction of an α-stable law, that is there is a slowly varying function ℓ_0 (in general different from ℓ of (3.1.7)) such that the sum of N independent waiting times normalized by $a_N = N^{1/\alpha}\ell_0(N)$ converges to an α-stable random variable, see e.g. [Whi02, Theorem 4.5.1]. Thus the quenched Laplace transform (which is deterministic here) satisfies

$$\hat{\pi}_0(\lambda/a_N) = \exp\left\{-c'\lambda^\alpha N^{-1}(1+o(1))\right\} \quad \text{as } N\to\infty$$

for some $c' > 0$. Taking this to the power $r\ell^*(N)$ it follows that the CTRW satisfies condition (3.1.9) with $a_N = N^{1/\alpha}\ell_0(N)$ and $f(r) = r$.

Secondly, consider the following simplified Bouchaud trap model (cf. (3.1.8)). Let $\pi_x = \delta_{\tau_x}$ where the τ_x, $x \in \mathbb{Z}^d$, are heavy-tailed i.i.d. random variables, that is

$$\mathbb{P}[\tau_x > u] = cu^{-\alpha}(1+o(1)) \text{ as } u\to\infty.$$

Then the quenched Laplace transform satisfies

$$\hat{\pi}_0(\lambda/a_N) = \exp\{-\lambda a_N^{-1}\tau_0\}.$$

Taking this to the power $r\ell^*(N)$ and taking the expectation over τ_0, this is the Laplace transform of a random variable in the normal domain of attraction of an α-stable law, evaluated at $r\lambda\ell^*(N)/a_N$. By normal domain of attraction we mean that the sum of N independent such random variables normalized by $c'N^{1/\alpha}$ converges to an α-stable random variable, see e.g. [Whi02, Theorem 4.5.2]. Thus choosing $a_N = c'N^{1/\alpha}\ell^*(N)^{1-1/\alpha}$, the Laplace transform is

$$\begin{aligned}\mathbb{E}\left[\hat{\pi}_0(\lambda/a_N)^{r\ell^*(N)}\right] &= \mathbb{E}\left[\exp\left\{-\frac{\lambda r}{c'\ell^*(N)^{-1/\alpha}}\frac{\tau_x}{N^{1/\alpha}}\right\}\right] \\ &= \exp\left\{-c''\frac{\ell^*(N)}{N}\lambda^\alpha r^\alpha(1+o(1))\right\} \text{ as } N\to\infty.\end{aligned}$$

Condition (3.1.9) is thus satisfied for $f(r) = r^\alpha$.

To see that $f(r)$ can be more than just a power of r, consider the following mixture of the above two models. To this end, let us fix the slowly varying function ℓ of (3.1.7) so that the normalization $a_N = N^{1/\alpha}\ell_0(N)$ of the first example agrees with the normalization $a_N = c'N^{1/\alpha}\ell^*(N)^{1-1/\alpha}$ of the second example. (This is possible e.g. when $1/\ell^*(N) \to \gamma \in (0,1)$, then also ℓ converges to a positive constant, or when $\ell^*(N) \sim c\log^{-1} N$, as is the case for simple random walk on $\mathbb{Z}^2$, then $\ell(N) \sim c'\ell^*(N)^{\alpha-1}$.) The mixture is now defined as follows. For some $p \in (0,1)$, let each π_x with probability p be a heavy-tailed distribution as in (3.1.7), and with probability $1-p$, let π_x be δ_{τ_x} where the τ_x are heavy-tailed random variables with

$$\mathbb{P}[\tau_x > u] = cu^{-\alpha}(1+o(1)) \text{ as } u \to \infty.$$

Then, by combining the arguments above, condition (3.1.9) is satisfied with the normalization $a_N = c'N^{1/\alpha}\ell^*(N)^{1-1/\alpha}$ and $f(r) = pr + (1-p)r^\alpha$.

Proof of Theorem 3.1.4. By Theorem 3.1.1 it is sufficient to show that

$$\lim_{N\to\infty} \mathbb{E}[\exp\{-\lambda S_N(t)\}] = e^{-ct\lambda^\alpha} \tag{3.5.1}$$

for some $c \in (0,\infty)$, this is equivalent to convergence of S_N to an α-stable subordinator. Using the independence of the π_x's, recalling that $\hat\pi_x$ denotes the Laplace transform of π_x, we have

$$\begin{aligned}\mathbb{E}\Big[\exp\Big\{-\frac{\lambda}{a_N}S(\lfloor Nt\rfloor)\Big\}\Big|Y\Big] &= \mathbb{E}\Big[\exp\Big\{-\frac{\lambda}{a_N}\sum_{x\in\mathbb{Z}^d}\sum_{i=1}^{L(x,\lfloor Nt\rfloor-1)}\tau_x^i\Big\}\Big|Y\Big]\\ &= \prod_{x\in\mathbb{Z}^d}\mathbb{E}\Big[\hat\pi_x(\lambda/a_N)^{L(x,\lfloor Nt\rfloor-1)}\Big\}\Big|Y\Big].\end{aligned} \tag{3.5.2}$$

Treating the case when Y is transient first, let $R^k(Nt) = \{x \in \mathbb{Z}^d : L(x,\lfloor Nt\rfloor-1) = k\}$. By Lemma 3.A.1, $|R^k(Nt)|/(Nt) \xrightarrow{N\to\infty} \gamma^2(1-\gamma)^{k-1}$ in probability. Using the translation invariance, the right-hand side of (3.5.2) can be written as

$$\exp\Big\{\sum_{k=1}^{\infty}|R^k(Nt)|\log\mathbb{E}\big[\hat\pi_0(\lambda/a_N)^k\big]\Big\}.$$

For arbitrary $M \in \mathbb{N}$, using the law of large numbers for $|R^k(Nt)|$ and assumption (3.1.9) with the continuity of f,

$$\sum_{k=1}^{M}|R^k(Nt)|\log\mathbb{E}\big[\hat\pi_0(\lambda/a_N)^k\big] \xrightarrow{N\to\infty} -t\lambda^\alpha\sum_{k=1}^{M}f(k\gamma)\gamma(1-\gamma)^{k-1}, \tag{3.5.3}$$

in probability. Applying Jensen's inequality, it is easy to see that $f(k)$ grows at most linearly with k, so the right-hand side of the above expression converges as $M \to \infty$ to a finite value, by assumptions of the theorem. On the other hand, by Jensen's inequality again, for every $\delta > 0$

$$\begin{aligned}&\mathbb{P}\Big[-\sum_{k=M}^{\infty}|R^k(Nt)|\log\mathbb{E}\big[\hat\pi_0(\lambda/a_N)^k\big] \ge \delta\Big]\\ &\quad\le \mathbb{P}\Big[-\log\mathbb{E}\big[\hat\pi_0(\lambda/a_N)\big]\sum_{k=M}^{\infty}|R^k(Nt)|k \ge \delta\Big].\end{aligned} \tag{3.5.4}$$

By the Markov inequality, for $0 < c_1 < -\log(1-\gamma)$, $\mathbb{P}[|R^k(Nt)|/(Nt) \geq e^{-c_1 k}] \leq e^{-c'k}$ uniformly for all $k \geq M$ and N large enough, and thus by a union bound

$$\mathbb{P}\big[\exists k \geq M \text{ such that } |R^k(Nt)|/(Nt) \geq e^{-c_1 k}\big] \leq Ce^{-c'M} \tag{3.5.5}$$

uniformly in N. Using (3.5.5) and the fact that $\log \mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)\big]$ is finite by assumption, it follows that the left-hand side of (3.5.4) converges to 0 in probability when $N \to \infty$ and then $M \to \infty$, and therefore (3.5.3) also holds with $M = \infty$. Using the bounded convergence theorem, it then follows that

$$\begin{aligned}\mathbb{E}\Big[\exp\Big\{-\frac{\lambda}{a_N} S(\lfloor Nt \rfloor)\Big\}\Big] &= \mathbb{E}\Big[\exp\Big\{\sum_{k=1}^{\infty} |R^k(Nt)| \log \mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^k\big]\Big\}\Big] \\ &\xrightarrow{N\to\infty} \exp\Big\{-t\lambda^\alpha \sum_{k=1}^{\infty} f(k\gamma)\gamma(1-\gamma)^{k-1}\Big\},\end{aligned}$$

which proves (3.5.1) in the transient case.

To treat the recurrent case, we fix $\beta > 0$ small and define for $k \geq 1$

$$R^k_\beta(Nt) = \{x \in \mathbb{Z}^d : (k-1)\beta\ell^*(N) < L(x, \lfloor Nt \rfloor - 1) \leq k\beta\ell^*(N)\}.$$

By Lemma 3.A.1, $|R^k_\beta(Nt)|\ell^*(N)/(Nt) \xrightarrow{N\to\infty} e^{-(k-1)\beta} - e^{-k\beta}$ in probability. The right-hand side of (3.5.2) can be bounded from above by

$$\exp\Big\{\sum_{k=1}^{\infty} |R^k_\beta(Nt)| \log \mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^{\beta(k-1)\ell^*(N)}\big]\Big\},$$

and from below by

$$\exp\Big\{\sum_{k=1}^{\infty} |R^k_\beta(Nt)| \log \mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^{\beta k\ell^*(N)}\big]\Big\}.$$

Following the same steps as in the transient case, it can be easily shown that

$$\begin{aligned}&\exp\Big\{-t\lambda^\alpha \sum_{k=1}^{\infty} f(\beta(k-1))\big(e^{-(k-1)\beta} - e^{-k\beta}\big)\Big\} \leq \lim_{N\to\infty} \mathbb{E}\big[e^{-\lambda S_N(t)}\big] \\ &\leq \exp\Big\{-t\lambda^\alpha \sum_{k=1}^{\infty} f(\beta k)\big(e^{-(k-1)\beta} - e^{-k\beta}\big)\Big\}\end{aligned}$$

Since f is a monotone function, the sums in the above expression can be viewed as lower and upper Riemann sums for the integral $\int_0^\infty f(x)e^{-x}\,dx$ to which they tend when $\beta \to 0$. This integral is finite since as argued before f grows at most linearly, and (3.5.1) is proved in the recurrent case. □

3.6. Ignoring small sets

In this section we prove Lemma 3.3.5 which allows us to ignore small sets when dealing with the clock process.

We first assume that the random walk Y is transient, that is $1/\ell^*(n) \to \gamma \in (0,1)$ as $n \to \infty$. We start by noting that for every $x \in R(\lfloor Nt \rfloor)$ and $i \in \{1, \dots, L(x, \lfloor Nt \rfloor - 1)\}$, since $(\tau_x^i)_{i \geq 1}$ are i.i.d.,

$$\mathbb{E}\left[\frac{\tau_x^i}{S(\lfloor Nt \rfloor)} \;\middle|\; Y \right] = \mathbb{E}\left[\frac{\tau_x^1}{S(\lfloor Nt \rfloor)} \;\middle|\; Y \right]. \tag{3.6.1}$$

For fixed $0 \leq l < \lfloor Nt \rfloor$, let $x = Y(l)$ and $i = L(Y(l), l)$, that is $\tilde{\tau}_l = \tau_x^i$. Using (3.6.1) and the fact that $(\tau_x^1)_{x \in \mathbb{Z}^d}$ are i.i.d. under $\mathbb{P}$,

$$\begin{aligned}
\mathbb{E}\left[\frac{\tilde{\tau}_l}{S(\lfloor Nt \rfloor)} \;\middle|\; Y \right] &= \mathbb{E}\left[\frac{\tau_x^i}{\sum_{y \in R(\lfloor Nt \rfloor)} \sum_{j=1}^{L(x, \lfloor Nt \rfloor - 1)} \tau_y^j} \;\middle|\; Y \right] \\
&\leq \mathbb{E}\left[\frac{\tau_x^1}{\sum_{y \in R(\lfloor Nt \rfloor)} \tau_y^1} \;\middle|\; Y \right] \\
&= \frac{1}{|R(\lfloor Nt \rfloor)|} \sum_{z \in R(\lfloor Nt \rfloor)} \mathbb{E}\left[\frac{\tau_z^1}{\sum_{y \in R(\lfloor Nt \rfloor)} \tau_y^1} \;\middle|\; Y \right] \\
&= \frac{1}{|R(\lfloor Nt \rfloor)|}.
\end{aligned} \tag{3.6.2}$$

By the law of large numbers for $R(n)$ (Lemma 3.A.1) in the transient case, there is a constant $C < \infty$ such that for all N large enough

$$\mathbb{P}\left[|R(\lfloor Nt \rfloor)| < CN \right] < \epsilon.$$

Hence, for N large enough,

$$\begin{aligned}
&\mathbb{P}\Big[\sum_{l \in \mathcal{B}} \tilde{\tau}_l \geq \delta S(\lfloor Nt \rfloor),\ |\mathcal{B}| \leq \epsilon N \Big] \\
&\qquad \leq \mathbb{P}\Big[\sum_{l \in \mathcal{B}} \tilde{\tau}_l \geq \delta S(\lfloor Nt \rfloor),\ |\mathcal{B}| \leq \epsilon N,\ |R(\lfloor Nt \rfloor)| \geq CN \Big] + \epsilon.
\end{aligned}$$

Using the Markov inequality and (3.6.2), this is bounded from above by

$$\begin{aligned}
&\leq \frac{1}{\delta} \mathbb{E}\left[\sum_{l \in \mathcal{B}} \mathbb{E}\left[\frac{\tilde{\tau}_l}{S(\lfloor Nt \rfloor)} \;\middle|\; Y \right] \mathbf{1}_{\{|\mathcal{B}| \leq \epsilon N\}} \mathbf{1}_{\{|R(\lfloor Nt \rfloor)| \geq CN\}} \right] + \epsilon. \\
&\leq \frac{1}{\delta} \mathbb{E}\left[\frac{|\mathcal{B}|}{|R(\lfloor Nt \rfloor)|} \mathbf{1}_{\{|\mathcal{B}| \leq \epsilon N\}} \mathbf{1}_{\{|R(\lfloor Nt \rfloor)| \geq CN\}} \right] + \epsilon \\
&\leq \frac{\epsilon}{C\delta} + \epsilon.
\end{aligned}$$

Letting $N \to \infty$ and then $\epsilon \to 0$ completes the proof of the lemma in the transient case.

We now consider the recurrent case. Let $R_{\mathcal{B}} = \{Y(l) : l \in \mathcal{B}\}$, and for $x \in R_{\mathcal{B}}$ let $L_{\mathcal{B}}(x) = |\{l \in \mathcal{B} : Y(l) = x\}|$. Fix some small $\beta > 0$ and let

$$\begin{aligned}
R_{>\beta} &= \{x \in R(\lfloor Nt \rfloor) : L(x, \lfloor Nt \rfloor - 1) > \beta \ell^*(N)\}, \\
R_{\leq \beta} &= \{x \in R(\lfloor Nt \rfloor) : L(x, \lfloor Nt \rfloor - 1) \leq \beta \ell^*(N)\}.
\end{aligned}$$

By Lemma 3.A.1, the sizes of $R_{>\beta}$ and $R_{\leq\beta}$ satisfy weak laws of large numbers with respective averages $Nte^{-\beta}/\ell^*(N)(1+o(1))$ and $Nt(1-e^{-\beta})/\ell^*(N)(1+o(1))$. In particular for $C_\beta = (1-\epsilon)e^{-\beta}t$ and $c_\beta = (1+\epsilon)(1-e^{-\beta})t$, for all N large enough,

$$\mathbb{P}\left[|R_{>\beta}| < C_\beta \frac{N}{\ell^*(N)}\right] + \mathbb{P}\left[|R_{\leq\beta}| > c_\beta \frac{N}{\ell^*(N)}\right] \leq \epsilon.$$

Therefore, for N large enough,

$$\begin{aligned}
&\mathbb{P}\Big[\sum_{l\in\mathcal{B}} \tilde{\tau}_l \geq \delta S(\lfloor Nt\rfloor),\ |\mathcal{B}| \leq \epsilon N\Big] \\
&\leq \mathbb{P}\Big[\sum_{x\in R_\mathcal{B}\cap R_{>\beta}} \sum_{i=1}^{L_\mathcal{B}(x)} \tau_x^i \geq \frac{\delta}{2} S(\lfloor Nt\rfloor),\ |\mathcal{B}| \leq \epsilon N,\ |R_{>\beta}| \geq C_\beta \frac{N}{\log N}\Big] \qquad (3.6.3)\\
&+ \mathbb{P}\Big[\sum_{x\in R_\mathcal{B}\cap R_{\leq\beta}} \sum_{i=1}^{L_\mathcal{B}(x)} \tau_x^i \geq \frac{\delta}{2} S(\lfloor Nt\rfloor),\ |R_{>\beta}| \geq C_\beta \frac{N}{\log N},\ |R_{\leq\beta}| \leq c_\beta \frac{N}{\log N}\Big] + \epsilon. \qquad (3.6.4)
\end{aligned}$$

Using (3.6.1) and the similar reasoning as in the transient case, since $(\sum_{i=1}^{\beta\ell^*(N)} \tau_x^i)_{x\in\mathbb{Z}^d}$ are i.i.d. with respect to the annealed measure and independent of Y, we have for $x \in R_\mathcal{B} \cap R_{>\beta}$,

$$\begin{aligned}
\mathbb{E}\left[\frac{\sum_{i=1}^{L_\mathcal{B}(x)} \tau_x^i}{S(\lfloor Nt\rfloor)} \,\middle|\, Y\right] &= \frac{L_\mathcal{B}(x)}{\beta\ell^*(N)} \mathbb{E}\left[\frac{\sum_{i=1}^{\beta\ell^*(N)} \tau_x^i}{S(\lfloor Nt\rfloor)} \,\middle|\, Y\right] \\
&\leq \frac{L_\mathcal{B}(x)}{\beta\ell^*(N)} \mathbb{E}\left[\frac{\sum_{i=1}^{\beta\ell^*(N)} \tau_x^i}{\sum_{y\in R_{>\beta}} \sum_{i=1}^{\beta\ell^*(N)} \tau_y^i} \,\middle|\, Y\right] \\
&= \frac{L_\mathcal{B}(x)}{|R_{>\beta}|\beta\ell^*(N)}.
\end{aligned}$$

Therefore, using the Markov inequality,

$$\begin{aligned}
(3.6.3) &\leq \frac{2}{\delta}\mathbb{E}\left[\sum_{x\in R_\mathcal{B}\cap R_{>\beta}} \mathbb{E}\left[\frac{\sum_{i=1}^{L_\mathcal{B}(x)} \tau_x^i}{S(\lfloor Nt\rfloor)} \,\middle|\, Y\right] \mathbf{1}_{\{|\mathcal{B}|\leq\epsilon N\}} \mathbf{1}_{\{|R_{>\beta}|\geq C_\beta \frac{N}{\ell^*(N)}\}}\right] \\
&\leq \frac{2}{\delta}\mathbb{E}\left[\sum_{x\in R_\mathcal{B}\cap R_{>\beta}} \frac{L_\mathcal{B}(x)}{|R_{>\beta}|\beta\ell^*(N)} \mathbf{1}_{\{|\mathcal{B}|\leq\epsilon N\}} \mathbf{1}_{\{|R_{>\beta}|\geq C_\beta \frac{N}{\ell^*(N)}\}}\right] \qquad (3.6.5)\\
&\leq \frac{2\epsilon}{\delta\beta C_\beta}.
\end{aligned}$$

where for the last inequality we used the fact that $\sum_x L_\mathcal{B}(x) \leq |\mathcal{B}| \leq \varepsilon N$.

It remains to bound (3.6.4). Using again the fact that $(\sum_{i=1}^{\beta\ell^*(N)} \tau_x^i)_{x\in\mathbb{Z}^d}$ are i.i.d. with

respect to the annealed measure and independent of Y,

$$\begin{aligned}
&\mathbb{P}\Big[\sum_{x\in R_{\mathcal{B}}\cap R_{\le\beta}}\sum_{i=1}^{L_{\mathcal{B}}(x)}\tau_x^i \ge \frac{\delta}{2}S(\lfloor Nt\rfloor)\ \Big|\ Y\Big]\\
&\le \mathbb{P}\Big[\Big(1+\frac{\delta}{2}\Big)\sum_{x\in R_{\mathcal{B}}\cap R_{\le\beta}}\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i \ge \frac{\delta}{2}\sum_{x\in R_{>\beta}\cup(R_{\mathcal{B}}\cap R_{\le\beta})}\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i\ \Big|\ Y\Big]\\
&\le \frac{2+\delta}{\delta}\mathbb{E}\Big[\frac{\sum_{x\in R_{\mathcal{B}}\cap R_{\le\beta}}\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i}{\sum_{x\in R_{>\beta}\cup(R_{\mathcal{B}}\cap R_{\le\beta})}\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i}\ \Big|\ Y\Big]\\
&= \frac{2+\delta}{\delta}\frac{|R_{\mathcal{B}}\cap R_{\le\beta}|}{|R_{>\beta}\cup(R_{\mathcal{B}}\cap R_{\le\beta})|}.
\end{aligned}$$

Therefore,

$$(3.6.4) \le \frac{2+\delta}{\delta}\frac{c_\beta}{C_\beta} = \frac{2+\delta}{\delta}\frac{1+\epsilon}{1-\epsilon}\left(e^\beta-1\right). \tag{3.6.6}$$

Combining (3.6.3)–(3.6.6) and letting $N\to\infty$, then $\epsilon\to 0$ and finally $\beta\to 0$ finishes the proof of the lemma in the recurrent case. □

Appendix 3.A. Laws of large numbers for range-like objects

We prove here that Assumption A implies weak laws of large numbers for several range-related quantities. The proofs are based on the classical paper [DE51], see also [Rév13, Chapter 21].

Recall that

$$R(n) = \{x\in\mathbb{Z}^d:\ L(x,n-1)>0\}$$

is the range of the random walk Y up to time $n-1$. In the recurrent case, i.e. if $\ell^*(n)\to\infty$, define for $k\ge 1$ and $\beta>0$

$$R_\beta^k(n) = \{x\in\mathbb{Z}^d:\ L(x,n-1)\in((k-1),k]\beta\ell^*(n)\}$$

the set of vertices visited $(k-1)\beta\ell^*(n)$ to $k\beta\ell^*$ times up to time $n-1$. In the transient case, if $1/\ell^*(n)\to\gamma\in(0,1)$, let for $k\ge 1$

$$R^k(n) = \{x\in\mathbb{Z}^d:\ L(x,n-1)=k\}$$

the vertices visited exactly k times up to time $n-1$.

We say that a sequence of random variables Z_n satisfies the weak law of large numbers if $Z_n/EZ_n \xrightarrow{n\to\infty} 1$ in probability.

Lemma 3.A.1.

(i) If Assumption A holds, then $|R(n)|$ satisfies the weak law of large numbers with

$$\mathbb{E}[|R(n)|] = \frac{n}{\ell^*(n)}(1+o(1)) \text{ as } n\to\infty.$$

(ii) If in addition $\ell^*(n) \to \infty$ as $n \to \infty$, then $|R^k_\beta(n)|$ satisfies the weak law of large numbers for every $k \geq 1$ and $\beta > 0$, and

$$\mathbb{E}[|R^k_\beta(n)|] = (e^{-(k-1)\beta} - e^{-k\beta})\frac{n}{\ell^*(n)}(1 + o(1)) \text{ as } n \to \infty.$$

(iii) If, on the other hand, $1/\ell^*(n) \to \gamma \in (0, 1)$, then $|R^k(n)|$ satisfies the weak law of large numbers for every $k \geq 1$, and

$$\mathbb{E}[|R^k(n)|] = \gamma^2(1-\gamma)^{k-1}n(1 + o(1)) \text{ as } n \to \infty.$$

Proof. Note that for the simple random walk in $d \geq 3$ and $d = 2$ respectively, part (i) is a classical result from [DE51], part (iii) was hinted at in [ET60, Theorem 12] and proved in [Pit74], whereas part (ii) is a direct consequence of [DE51, Theorem 4] and [Čer07, Theorem 2]. Part (i) above is proved exactly as in [DE51]. We include its proof, since proofs of (ii) and (iii) are its extensions. Let ψ_k be the indicator of the event that a new vertex is found at time k,

$$\psi_k = \mathbf{1}_{\{Y(l) \neq Y(k) \text{ for all } 0 \leq l < k\}},$$

with $\psi_0 = 1$. Recall that ξ_i denote the i.i.d. increments of the random walk Y. Then,

$$\begin{aligned}
\mathbb{E}[\psi_k] &= \mathbb{P}\left[Y(k) \neq Y(k-1),\ Y(k) \neq Y(k-2), \ldots,\ Y(k) \neq Y(0)\right] \\
&= \mathbb{P}\left[\xi_k \neq 0,\ \xi_k + \xi_{k-1} \neq 0, \ldots,\ \xi_k + \cdots + \xi_1 \neq 0\right] \\
&= \mathbb{P}\left[\xi_1 \neq 0,\ \xi_1 + \xi_2 \neq 0, \ldots,\ \xi_1 + \cdots + \xi_k \neq 0\right] \\
&= \mathbb{P}[Y(l) \neq 0 \text{ for } l = 1, \ldots, k] = r_k.
\end{aligned} \tag{3.A.1}$$

For a slowly varying function ℓ, $\sum_{k=1}^n \ell(k) = n\ell(n)(1 + o(1))$ as $n \to \infty$ (see e.g. [Sen76, p. 55]). Therefore, by Assumption A,

$$\mathbb{E}[|R(n)|] = \sum_{k=0}^{n-1} \mathbb{E}[\psi_k] = \frac{n}{\ell^*(n)}(1 + o(1)) \text{ as } n \to \infty. \tag{3.A.2}$$

To prove the weak law of large numbers, we compute the variance. First note that for $i \leq j$, by the Markov property,

$$\begin{aligned}
\mathbb{E}[\psi_i\psi_j] &= \mathbb{E}\left[\mathbf{1}_{\{Y(l) \neq Y(i),\ 0 \leq l < i\}}\mathbf{1}_{\{Y(l) \neq Y(j),\ 0 \leq l < j\}}\right] \\
&\leq \mathbb{E}\left[\mathbf{1}_{\{Y(l) \neq Y(i),\ 0 \leq l < i\}}\mathbf{1}_{\{Y(l) \neq Y(j),\ i \leq l < j\}}\right] = \mathbb{E}[\psi_i]\mathbb{E}[\psi_{j-i}].
\end{aligned} \tag{3.A.3}$$

Then,

$$\begin{aligned}
\operatorname{Var}|R(n)| &= \sum_{0 \leq i,j \leq n-1} \mathbb{E}[\psi_i\psi_j] - \mathbb{E}[\psi_i]\mathbb{E}[\psi_j] \\
&\leq 2\sum_{i=0}^{n-1}\sum_{j=i}^{n-1} \mathbb{E}[\psi_i]\left(\mathbb{E}[\psi_{j-i}] - \mathbb{E}[\psi_j]\right) \\
&\leq 2\sum_{i=0}^{n-1} \mathbb{E}[\psi_i]\left(\max_{k=0,\ldots,n-1}\sum_{j=k}^{n-1} \mathbb{E}[\psi_{j-k}] - \mathbb{E}[\psi_j]\right).
\end{aligned} \tag{3.A.4}$$

By (3.A.1), $\mathbb{E}[\psi_k]$ is non-increasing, therefore the maximum in (3.A.4) is attained in $k = \frac{n}{2}$. The parenthesis in (3.A.4) can then be estimated using elementary properties of slowly varying functions,

$$\begin{aligned}\sum_{j=\frac{n}{2}}^{n-1} \mathbb{E}[\psi_{j-\frac{n}{2}}] - \mathbb{E}[\psi_j] &= \sum_{j=0}^{\frac{n}{2}-1} \frac{1}{\ell^*(j)} - \sum_{j=\frac{n}{2}}^{n-1} \frac{1}{\ell^*(j)} \\ &= \sum_{j=0}^{\frac{n}{2}-1} \frac{1}{\ell^*(j)} - \left(\sum_{j=0}^{n-1} \frac{1}{\ell^*(j)} - \sum_{j=0}^{\frac{n}{2}-1} \frac{1}{\ell^*(j)} \right) \\ &= 2\frac{\frac{n}{2}}{\ell^*(\frac{n}{2})}(1+o(1)) - \frac{n}{\ell^*(n)}(1+o(1)) = \frac{n}{\ell^*(n)} o(1) \text{ as } n \to \infty.\end{aligned}$$

Inserting this into (3.A.4), we obtain

$$\operatorname{Var} |R(n)| \le 2 \sum_{i=0}^{n-1} \mathbb{E}[\psi_i] \frac{n}{\ell^*(n)} o(1) = o\left(\left(\frac{n}{\ell^*(n)} \right)^2 \right) \text{ as } n \to \infty,$$

and the weak law of large numbers for $|R(n)|$ follows by usual arguments.

Before turning to part (ii), we note the following fact on return times. Let as before $H_0^1 = \inf\{i > 0 : \ Y(i) = 0\}$ denote the time of the first return to 0, and $H_0^k = \inf\{i > H_0^{k-1} : \ Y(i) = 0\}$ the time of the k-th return to 0. Let $T_i = H_0^i - H_0^{i-1}$ (with $H_0^0 = 0$) be the successive return times. By the Markov property the $(T_i)_{i \ge 1}$ are i.i.d., and $\mathbb{P}[T_i > n] = r_n = \frac{1}{\ell^*(n)}$ by Assumption A. If $\ell^*(k) \to \infty$, the T_i are a.s. finite and have slowly varying tail. It is well known (e.g. [Dar52, Theorem 3.2]) that for such i.i.d. random variables T_i,

$$\frac{\sum_{i=1}^n T_i}{\max_{i=1}^n T_i} \to 1 \text{ in probability as } n \to \infty. \tag{3.A.5}$$

Since $\ell^*(cn) \sim \ell^*(n)$ as $n \to \infty$,

$$\mathbb{P}\big[\max\{T_i : 1 \le i \le \beta \ell^*(n)\} \le cn \big] = \left(1 - \frac{1}{\ell^*(cn)} \right)^{\beta \ell^*(n)} = e^{-\beta}(1+o(1)) \text{ as } n \to \infty. \tag{3.A.6}$$

From (3.A.5) and (3.A.6) we obtain for every $c > 0$ and $\beta > 0$

$$\mathbb{P}\left[L(0, cn) \ge \beta \ell^*(n) \right] = \mathbb{P}\left[\sum_{i=1}^{\beta \ell^*(n)} T_i \le cn \right] = e^{-\beta}(1+o(1)). \tag{3.A.7}$$

For part (ii) we only prove the statement for $R_\beta(n) = R_\beta^1(n)$, the statement for $k > 1$ follows easily by subtracting the claims with β replaced by βk and $\beta(k-1)$. Consider ψ_k as above, and additionally define functions $\varphi_k = \mathbf{1}_{\{L(Y(k), n-1) \le \beta \ell^*(n)\}}$. Using the Markov property and translation invariance,

$$\begin{aligned}\mathbb{E}[|R_\beta(n)|] = \sum_{k=0}^{n-1} \mathbb{E}[\psi_k \varphi_k] &= \sum_{k=0}^{n-1} \mathbb{E}[\psi_k] \mathbb{P}\left[L(0, n-1-k) \le \beta \ell^*(n) \right] \\ &= \sum_{k=0}^{n-1} \frac{1}{\ell^*(k)} \mathbb{P}\left[\sum_{i=1}^{\beta \ell^*(n)} T_i \ge n-1-k \right].\end{aligned} \tag{3.A.8}$$

If $k \leq (1-\delta)n$ for some $\delta > 0$, then we can apply (3.A.7). Bounding the probability by one in the remaining cases, we see that (3.A.8) is bounded from above by

$$\begin{aligned}\mathbb{E}[|R_\beta(n)|] &\leq \sum_{k=0}^{(1-\delta)n} \frac{1}{\ell^*(k)}(1-e^{-\beta})(1+o_\delta(1)) + \sum_{(1-\delta)n<k<n} \frac{1}{\ell^*(k)} \\ &= \frac{(1-\delta)n}{\ell^*(n)}(1-e^{-\beta})(1+o_\delta(1)) + \frac{\delta n}{\ell^*(n)},\end{aligned}$$

and from below by

$$\mathbb{E}[|R_\beta(n)|] \geq \sum_{k=0}^{(1-\delta)n} \frac{1}{\ell^*(k)}(1-e^{-\beta})(1+o_\delta(1)) = \frac{(1-\delta)n}{\ell^*(n)}(1-e^{-\beta})(1+o_\delta(1)).$$

Sending $\delta \to 0$ proves the statement for $\mathbb{E}[|R_\beta(n)|]$.

To bound the variance, we first note that for $i < i+\delta n \leq j \leq (1-\delta)n$, by the Markov property and using Assumption A and (3.A.7),

$$\begin{aligned}\mathbb{E}\left[\psi_i\varphi_i\psi_j\varphi_j\right] &\leq \mathbb{E}\left[\psi_i \mathbf{1}_{\{L(Y(i),i+\delta n)\leq\beta\ell^*(n)\}}\mathbf{1}_{\{Y(k)\neq Y(j),\ i+\delta n\leq k<j\}}\mathbf{1}_{\{L(Y(j),n)\leq\beta\ell^*(n)\}}\right] \\ &\leq \mathbb{E}[\psi_i]\mathbb{P}\big[L(0,\delta n)\leq\beta\ell^*(n)\big]\mathbb{E}\big[\psi_{j-i-\delta n}\big]\mathbb{P}\big[L(0,\delta n)\leq\beta\ell^*(n)\big] \\ &= \frac{1}{\ell^*(i)}(1-e^{-\beta})\frac{1}{\ell^*(j-i-\delta n)}(1-e^{-\beta})(1+o_\delta(1)). \qquad (3.A.9)\end{aligned}$$

The variance of $|R_\beta(n)|$ is

$$\operatorname{Var}|R_\beta(n)| = 2\sum_{0\leq i\leq j\leq n-1} \mathbb{E}[\psi_i\varphi_i\psi_j\varphi_j] - \mathbb{E}[\psi_i\varphi_i]\mathbb{E}[\psi_j\varphi_j]. \qquad (3.A.10)$$

For $i < i+\delta n \leq j \leq (1-\delta)n$ we can use (3.A.9) and (3.A.8) to get

$$\begin{aligned}&\sum_{i<i+\delta n\leq j\leq(1-\delta)n} \mathbb{E}[\psi_i\xi_i\psi_j\xi_j] - \mathbb{E}[\psi_i\xi_i]\mathbb{E}[\psi_j\xi_j] \\ &\leq (1-e^{-\beta})^2 \sum_{i<i+\delta n\leq j\leq(1-\delta)n} \frac{1}{\ell^*(i)}\left(\frac{1}{\ell^*(j-i-\delta n)}(1+o_\delta(1)) - \frac{1}{\ell^*(j)}(1+o_\delta(1))\right) \\ &= (1-e^{-\beta})^2 \sum_{i=0}^{(1-2\delta)n} \frac{1}{\ell^*(i)}\left(\sum_{j=0}^{(1-2\delta)n-i} \frac{1}{\ell^*(j)}(1+o_\delta(1)) - \sum_{j=i+\delta n}^{(1-\delta)n} \frac{1}{\ell^*(j)}(1+o_\delta(1))\right) \\ &= o_\delta\left(\left(\frac{n}{\ell^*(n)}\right)^2\right). \qquad (3.A.11)\end{aligned}$$

For the remaining i, j, using (3.A.3) we have

$$\begin{aligned}
&\sum_{i=0}^{n-1} \sum_{\substack{i\le j<i+\delta n\\ (1-\delta)n<j<n}} \mathbb{E}[\psi_i\xi_i\psi_j\xi_j] - \mathbb{E}[\psi_i\xi_i]\mathbb{E}[\psi_j\xi_j] \\
&\qquad \le \sum_{i=0}^{n-1} \sum_{\substack{i\le j<i+\delta n\\ (1-\delta)n<j<n}} \mathbb{E}[\psi_i\psi_j] \\
&\qquad \le \sum_{i=0}^{n-1} \sum_{\substack{i\le j<i+\delta n\\ (1-\delta)n<j<n}} \mathbb{E}[\psi_i]\mathbb{E}[\psi_{j-i}] \\
&\qquad \le \sum_{i=0}^{n-1} \frac{1}{\ell^*(i)} \left(\sum_{j=i}^{i+\delta n-1} \frac{1}{\ell^*(j)} + \sum_{j=(1-\delta)n+1}^{n-1} \frac{1}{\ell^*(j)} \right) \\
&\qquad \le 2\frac{\delta n^2}{\left(\ell^*(n)\right)^2}(1+o_\delta(1)).
\end{aligned} \tag{3.A.12}$$

Inserting (3.A.11) and (3.A.12) into (3.A.10) and taking $\delta \to 0$ yields $\operatorname{Var}|R_\beta(n)| = o((\mathbb{E}|R_\beta(n)|)^2)$ and the weak law of large numbers follows.

Finally, part (iii) is proved in the same way as part (ii). The only difference is that instead of using (3.A.7) we note that $L(0,\infty)$ is a geometric random variable with parameter γ, therefore for every $c > 0$,

$$\mathbb{P}\left[L(0,cn) = k\right] = \gamma(1-\gamma)^{k-1}(1+o(1)) \text{ as } n \to \infty.$$

This completes the proof. □

Part II

PERCOLATION OF THE VACANT SET

4. Introduction to percolation of the vacant set

In the second part of the thesis we consider the problem of *fragmentation of a finite graph by random walk*. We are interested in the percolative properties of the *vacant set*, the set of vertices that have not been visited by a random walk up to a certain time. Intuitively, for short times the vacant set will consist of a large connected component that is 'dense' in the graph and possibly some other small components. On the other hand, after a long time the trace of the random walk occupies a large part of the graph and the vacant set should be fragmented into only small connected components. It is thus natural to ask how the transition between these two phases occurs.

This question is, however, only interesting on large graphs, and we will in fact consider asymptotics on a diverging sequence of graphs. Therefore, the precise definition of the model is as follows. Let $G_n = (V_n, E_n)$ be an increasing sequence of possibly random connected simple graphs with finite sets of vertices V_n and edges E_n, such that $|V_n| \to \infty$ as $n \to \infty$. Let $X = (X_k)_{k \geq 0}$ be the simple random walk on G_n, i.e. the discrete-time Markov chain on the state space V_n which chooses its next state uniformly among all neighbors of the current state in the graph. For $u \geq 0$ let

$$\mathcal{V}_n(u) := V_n \setminus \{X_k : \ 0 \leq k \leq u|V_n|\}$$

be the *vacant set of the random walk at level* u, i.e. the set of vertices that have not been visited by the random walk up to time $u|V_n|$. For notational convenience we will use $\mathcal{V}_n(u)$ to denote the set of vertices as well as the subgraph induced by these vertices.

Under some additional assumptions, the vacant set at a time proportional to the size of the graph indeed defines a non-trivial percolation model. Namely, assuming that the graphs G_n are sufficiently fast mixing and that they converge locally (roughly in the sense that every ball of fixed radius in G_n converges to a ball in G, in distribution if the graphs are random) to an infinite graph $G = (V, E)$ which is transient, it can be shown that for every $x \in G_n$ and $u \in (0, \infty)$,

$$P[x \in \mathcal{V}_n(u)] \xrightarrow{n \to \infty} e^{-cu},$$

for an explicit constant $c > 0$ (at least in the case where G is vertex-transitive). This shows that the one-dimensional marginals of the vacant set percolation are non-trivial Bernoulli random variables. The fact that the vacant set is determined by a random walk trajectory, however, introduces a complicated dependency structure, which makes this problem far more involved than independent Bernoulli percolation.

Nevertheless, it is expected that the sizes of connected components of the vacant set, as macroscopic parameters, behave similarly as in independent Bernoulli percolation. We state this as a heuristic conjecture.

Conjecture 4.1. The component structure of the vacant set left by random walk on an increasing sequence of finite graphs satisfying the above assumptions exhibits the following phase transition: There is a critical value $u_\star \in (0, \infty)$, such that for some constant $\kappa > 0$,

- whenever $u < u_\star$, with probability tending to 1 as $n \to \infty$, the vacant set $\mathcal{V}_n(u)$ consists of one unique 'giant' connected component of size of order $|V_n|$, and all other connected components are at most of size of order $\log^\kappa |V_n|$;
- whenever $u > u_\star$, with probability tending to 1 as $n \to \infty$, all connected components of the vacant set $\mathcal{V}_n(u)$ are at most of size of order $\log^\kappa |V_n|$.

In Chapter 5 we will prove Conjecture 4.1 for an increasing sequence of certain random graphs. In this introductory chapter we first recall some facts on percolation theory and introduce the random graph considered in Chapter 5. We then outline in Sections 4.2 and 4.3 the history of the problem of fragmentation of a finite graph by random walk and how the model of *random interlacements* relates to it. Finally, we describe the result and method of Chapter 5 in Section 4.4.

4.1. The Erdős-Rényi phase transition

The nature of the phase transition of Conjecture 4.1 is motivated by the classical phase transition for Bernoulli percolation on the complete graph. Let us recall some facts of this theory. In *Bernoulli edge percolation* on a finite or infinite graph $G = (V, E)$, every edge $e \in E$ is independently set *open* with probability $p \in (0, 1)$ and *closed* otherwise. One is then interested in the geometry of the *open subgraph* which is obtained upon removing all closed edges.

Bernoulli edge percolation on the complete graph on n vertices, $\mathbb{K}_n$, is also known as the *Erdős-Rényi random graph*, cf. the monographs [Bol01, JŁR00, Dur10]. This random graph exhibits the following phase transition:

Theorem 4.1.1 ([ER61])**.** *If the probability of edges being open is $p = \frac{\rho}{n}$ for some $\rho > 0$, then*

- *whenever $\rho < 1$, with probability tending to 1 as $n \to \infty$, all connected components of the open subgraph are at most of size of order* $\log n$*;*
- *whenever $\rho > 1$, with probability tending to 1 as $n \to \infty$, there is one unique 'giant' connected component of the open subgraph of size of order n, and all other connected components are at most of size of order* $\log n$*.*

More is known about the behavior near the critical point $\rho = 1$. Namely, if $p = \frac{1}{n} + \lambda n^{-\frac{4}{3}}$ for some $\lambda \in \mathbb{R}$, then the largest components are of size of order $n^{\frac{2}{3}}$, and the rescaled ordered component sizes converge in distribution to a process equivalent to ordered excursion lengths of a drifted and reflected Brownian motion [Ald97]. This behavior is referred to as the existence of a *critical window*. The width $n^{-\frac{4}{3}}$ of the window is 'correct' in the sense that if $\lambda \to -\infty$, the size of the largest component becomes considerably smaller than $n^{\frac{2}{3}}$, and if $\lambda \to \infty$ it becomes larger than of order $n^{\frac{2}{3}}$ and the second largest component becomes much smaller.

The random graph on which we will consider the vacant set of random walk and prove Conjecture 4.1 is the giant component of the Erdős-Rényi random graph in the

supercritical phase $\rho > 1$. To illustrate some of its features, let us briefly sketch one way how the phase transition of Theorem 4.1.1 can be understood heuristically.

Consider the following exploration of the component containing a given vertex $x \in \mathbb{K}_n$. Since every other vertex is connected to x independently with probability p, the number of neighbors of x in the open subgraph is a Binomial(n, p) random variable, hence for $p = \frac{\rho}{n}$ and large n roughly a Poisson(ρ) random variable. The neighbors of x have a number of additional neighbors in the open subgraph which is Binomial$(n-1, p)$, i.e. still roughly Poisson(ρ). Therefore, as long as one does not go too far from the starting point, the explored connected component looks like a Poisson(ρ)-Galton-Watson tree. This tree grows infinitely large with positive probability whenever $\rho > 1$, which after some additional arguments (*sprinkling*) leads to the existence of the giant component in that case. On the other hand, if $\rho < 1$, then the Poisson(ρ)-Galton-Watson tree goes extinct with probability 1, which rather directly implies that all components remain small.

The two properties of the supercritical phase $\rho > 1$ which will be relevant later are the following. First, the above heuristics explain why the giant component Erdős-Rényi random graph is of size $\xi n + o(n)$, where ξ is the survival probability of a Poisson(ρ)-Galton-Watson tree. Second, one can think of the giant component as looking locally like a Poisson(ρ)-Galton-Watson tree conditioned on survival. We will see in Chapter 5 that this can be made precise by using a coupling of the random graph and a Poisson(ρ)-Galton-Watson tree.

Besides the Erdős-Rényi random graph, i.e. Bernoulli percolation on the complete graph, a phase transition like the one in Theorem 4.1.1 including the critical window was also verified for Bernoulli percolation on other sequences of finite graphs, e.g. on random d-regular graphs [ABS04, NP10] or on finite tori [HvdH07, HvdH11]. Moreover, the same behavior has been shown to appear in other random graph models such as the *random graph with given degree sequence* [MR95, HM12, Rio12].

We have seen that for understanding the Erdős-Rényi random graph it is convenient to consider the local infinite volume limit, which is the Poisson-Galton-Watson tree. Similarly, Bernoulli percolation on $\mathbb{Z}^d$ can be used to understand percolation on finite tori. In particular, the sharpness of a phase transition (disregarding its non-triviality) in the infinite volume limit usually follows from rather soft arguments like the 0-1 law for the Galton-Watson tree or shift-ergodicity for percolation on $\mathbb{Z}^d$. We will come back to this remark later.

4.2. Fragmentation of the torus and random interlacements

Let us come back to Conjecture 4.1. In the probability literature, the problem of fragmentation of a finite graph by random walk was for the first time investigated in [BS08], namely on the discrete torus, i.e. the graph $\mathbb{T}_N^d = (V_N^d, E_N^d)$ with vertex set $V_N^d = (\mathbb{Z}/N\mathbb{Z})^d$ and nearest-neighbor edges E_N^d, in dimensions $d \geq 3$. It was shown that for small values of u and sufficiently high dimensions d, with high probability as N tends to infinity the vacant set $\mathcal{V}_N(u)$ contains a connected component of size of order N^d. This giant component is unique in the sense that there might be other components of size of order N^d, but there is only one containing certain long segments. This result is rather far from the more detailed statement of Conjecture 4.1.

In order to better understand this problem of the vacant set of random walk on the torus, Sznitman introduced the model of *random interlacements* on $\mathbb{Z}^d$ [Szn10] (cf. also [ČT12, DRS14a] for introductions and more references). The occupied or closed vertices in this percolation model are induced by a collection of doubly-infinite nearest-neighbor paths, given by a Poisson point process on the space of these paths modulo time shift. The density of occupied vertices is determined by the intensity parameter $u > 0$ of this Poisson point process. The set of non-occupied vertices as well as the subgraph induced by these vertices is called the *vacant set of random interlacements at intensity u*.

The construction of random interlacements is motivated by the idea to have an infinite volume analogue for the problem of fragmentation of a finite graph by random walk. As remarked at the end of the last section, this can be useful for the analysis of phase transitions. Random interlacements on $\mathbb{Z}^d$ are the local limit of fragmentation by random walk of the discrete torus $\mathbb{T}_N^d$ in the sense that the 'local picture' is the same. Indeed, [Win08] showed that the vacant set of random walk at level u, $\mathcal{V}_N(u)$, converges locally in law as $N \to \infty$ to the vacant set of random interlacements on $\mathbb{Z}^d$ at intensity u.

In [Szn10, SS09] it was shown that random interlacements on $\mathbb{Z}^d$, $d \geq 3$, exhibit a phase transition, similar to Bernoulli percolation. Like in Bernoulli percolation, the existence of this phase transition follows from shift-ergodicity. The non-triviality of the critical intensity $u_\star(d)$ however is highly non-trivial due to the long-range dependency structure of the model.

Theorem 4.2.1 ([Szn10, SS09])**.** *There exists a critical intensity $u_\star(d) \in (0, \infty)$, such that*

- *whenever $u > u_\star(d)$, almost surely all connected components of the vacant set at intensity u are finite;*
- *whenever $u < u_\star(d)$, almost surely there is one unique infinite connected component in the vacant set at intensity u.*

Much more has been proved about random interlacements in the last five years, but there are important open questions related to the problem of fragmentation of the torus by random walk which are still open. Namely, besides $u_\star(d)$ there are at present two other critical parameters for random interlacements on $\mathbb{Z}^d$, $d \geq 3$: $u_1(d)$ and $u_2(d)$ such that $0 < u_1(d) \leq u_\star(d) \leq u_2(d) < \infty$. For all $0 \leq u \leq u_1(d)$, the infinite component of the vacant set is 'locally unique' in the sense that any large connected component in a ball around the origin actually belongs to the infinite component [DRS14b]. For all $u > u_2(d)$, the connectivity function in the vacant set at intensity u has stretched exponential decay [SS10]. It is believed that $u_1(d) = u_\star(d) = u_2(d)$, but a proof of this claim is still missing.

Random interlacements lead to a better understanding of the problem of fragmentation of the torus $\mathbb{T}_N^d$ by random walk. Using a coupling of random interlacements on $\mathbb{Z}^d$ and the random walk on the torus it was possible to extend the results of [BS08]. Namely, in [TW11] it was shown that in all dimensions $d \geq 3$ there are in fact two distinct phases. More precisely, for $u_1(d), u_\star(d), u_2(d)$ as above, it was shown that with high probability as N tends to infinity,

- for $u < u_1(d)$, the largest connected component of $\mathcal{V}_N(u)$ is of size of order N^d, and the second largest connected component of $\mathcal{V}_N(u)$ is at most of size of order $\log^\kappa N$, for some $\kappa > 0$;

- for $u > u_\star(d)$, the largest connected component of $\mathcal{V}_N(u)$ is of size of order $o(N^d)$;
- for $u > u_2(d)$, the largest connected component of $\mathcal{V}_N(u)$ is at most of size of order $\log^\kappa N$, for some $\kappa > 0$.

The first statement was actually shown in [TW11] only for $d \geq 5$ and below a different intensity $\tilde{u}_1(d)$, but it can be extended to $d \geq 3$ using the results of [DRS14b] for random interlacements at levels $u < u_1(d)$.

The above is still not giving the full picture of Conjecture 4.1, i.e. the sharp phase transition at $u = u_\star(d)$. Proving this would require more control on the random interlacements process on $\mathbb{Z}^d$ in both sub- and supercritical phases, in particular it would follow from a proof of $u_1(d) = u_\star(d) = u_2(d)$.

Let us remark also that recently in [ČT14] an improved coupling of random interlacements on $\mathbb{Z}^d$ and the random walk on $\mathbb{T}_N^d$ was used to establish a sharp phase transition at $u = u_\star(d)$, however not for the size of the components but for the diameter of the component of the vacant set containing a given point.

4.3. Locally tree-like and random graphs

The model of random interlacements was extended to arbitrary transient weighted graphs in [Tei09], and conditions involving isoperimetric inequalities were established under which a non-trivial phase transition like in Theorem 4.2.1 holds on such graphs. This provides the infinite volume limit of the problem of fragmentation by random walk also on sequences of finite graphs other than the torus which locally converge to some infinite graph. Motivated by the relation between random interlacements on $\mathbb{Z}^d$ and random walk on the torus $\mathbb{T}_N^d$, Conjecture 4.1 can then be extended by the following.

Conjecture 4.2. The critical level of the phase transition in the component structure of the vacant set of random walk on a sequence of finite graphs should be equal to the critical intensity of random interlacements on the corresponding infinite volume limit.

Both Conjectures 4.1 and 4.2 have indeed been proved for graphs that are locally tree-like. For such graphs, random interlacements on the corresponding infinite volume limit are particularly well understood. Namely, [Tei09] showed that for random interlacements on an infinite tree, the component of the vacant set containing a given point has the same law as a certain branching process. As a consequence, it was shown in [Tei09] that a non-trivial phase transition like in Theorem 4.2.1 holds if the graph is a transient tree with degree bounded and at least 3, endowed with weights bounded above and below. The critical intensity for random interlacements on the infinite d-regular tree $\mathcal{T}_d$, $d \geq 3$, was obtained explicitly.

The good control available for random interlacements on trees was used in [ČTW11] to prove both Conjectures 4.1 and 4.2 on sequences of fast mixing finite graphs whose infinite volume limit is the infinite d-regular tree $\mathcal{T}_d$. More precisely, [ČTW11] proved the phase transition for the vacant set of random walk and identified the critical level with the critical intensity of random interlacements on $\mathcal{T}_d$ for sequences of finite graphs $G_n = (V_n, E_n)$ such that $|V_n| \to \infty$ as $n \to \infty$, and for some $d \geq 3$, $\alpha_1 \in (0, 1)$, $\alpha_2 > 0$,

and all $n \geq 1$,

$$\begin{aligned} &\text{(i)} \quad G_n \text{ is } d\text{-regular;} \\ &\text{(ii)} \quad \text{for any } x \in V_n, \text{ there is at most one cycle contained in the ball of} \\ &\qquad \text{radius } \alpha_1 \log_{d-1} |V_n| \text{ around } x; \\ &\text{(iii)} \quad \text{the spectral gap } \lambda_{G_n} \text{ satisfies } \lambda_{G_n} > \alpha_2 > 0. \end{aligned} \tag{4.3.1}$$

These properties are satisfied e.g. by large girth expanders, and with high probability by large random d-regular graphs.

Another approach to prove Conjecture 4.1 was discovered in [CF11]. This approach requires that the graph under consideration is random, and that a certain local construction of the graph is possible. It was observed by [CF11] that the vacant set left by random walk on such a random graph satisfies a certain 'spatial Markov property'. This has the effect that a phase transition in the random subgraph induced by the vacant set can be deduced directly from the phase transition in the corresponding random graph model.

This idea was used twice in [CF11]. First, a phase transition similar to the one of Conjecture 4.1 without the extension of Conjecture 4.2 was obtained for the vacant set left by random walk on the Erdős-Rényi random graph above the connectivity threshold $p \gg \frac{\log n}{n}$, using that the law of the vacant subgraph is again an Erdős-Rényi random graph which exhibits the classical phase transition of Theorem 4.1.1.

Second, the phase transition for the vacant set on random d-regular graphs was shown independently from [ČTW11]. In that case the vacant subgraph has the law of a random graph with a given degree sequence. For such graphs a phase transition like the one in Theorem 4.1.1 is known to hold, including the existence of a critical window [MR95, HM12, Rio12]. The critical point of the phase transition depends on a single parameter which is computable from the degree sequence of the graph. Note that the phase transition for the vacant set on random d-regular graphs of [CF11] holds *annealed*, i.e. averaged over the randomness in the graph, whereas in [ČTW11] it holds for every realization satisfying assumptions (4.3.1).

Taking up the idea of [CF11], [ČT13] used the detailed results on random graphs with given degree sequence to show the existence of a critical window in the phase transition for the vacant set on random d-regular graphs, by giving sharp estimates on the random degree sequence of the vacant set, incorporating again the relation to random interlacements.

4.4. The vacant set on the giant component

In Chapter 5 we will prove Conjecture 4.1 including Conjecture 4.2 for the random walk on the giant component of the supercritical Erdős-Rényi random graph. That is, we will show the phase transition for the vacant set and relate the critical level to the critical intensity of random interlacements on a Poisson(ρ)-Galton-Watson tree conditioned on non-extinction.

The giant component is a natural candidate for Conjecture 4.1. First, it is fast mixing: the mixing time of simple random walk on the giant component of size ξn is of order $\log^2 n$ [FR08, BKW14]. Second, as described in Section 4.1, the giant component is locally isomorphic to a Poisson(ρ)-Galton-Watson tree conditioned on non-extinction. Thus the

problem of fragmentation by random walk can be related to random interlacements on a Galton-Watson tree.

Besides the fact that the giant component is a natural candidate, the problem of fragmentation by random walk is particularly interesting on this graph. Conjecture 4.1 with the extension of Conjecture 4.2 has before only been proved on regular graphs on a deterministic set of vertices. The non-regularity of the giant component brings up some noteworthy points in the connection to random interlacements on the corresponding infinite volume limit. Moreover, the fact that the giant component consists of a random set of vertices introduces certain technical difficulties.

As for the connection to the infinite volume analogue, random interlacements on Galton-Watson trees were studied in [Tas10]. Relying on the characterization of the vacant set on trees by [Tei09], it was shown that the critical intensity of random interlacements on a Galton-Watson tree conditioned on non-extinction is almost surely constant, i.e. it does not depend on the realization of the tree, that it is non-trivial, and that it is implicitly given as the solution of a certain equation. We will show that the critical level of the phase transition for the vacant set left by random walk on the giant component solves the exact same equation as the critical intensity of random interlacements on a Poisson(ρ)-Galton-Watson tree.

The randomness of the size of the giant component, or rather the randomness of the set of vertices of which it consists, is an obstacle for applying the idea of the spatial Markov property of [CF11]. We overcome this problem by introducing a process that evolves on the whole graph and not only on the giant component and that satisfies the spatial Markov property. By coupling the random walk on the giant component to this process, we are then able to deduce the phase transition of Conjecture 4.1 directly from the classical Erdős-Rényi phase transition of Theorem 4.1.1.

5. Phase transition for the vacant set left by random walk on the giant component of a random graph

Tobias Wassmer

Abstract. We study the simple random walk on the giant component of a supercritical Erdős-Rényi random graph on n vertices, in particular the so-called vacant set at level u, the complement of the trajectory of the random walk run up to a time proportional to u and n. We show that the component structure of the vacant set exhibits a phase transition at a critical parameter $u_\star$: For $u < u_\star$ the vacant set has with high probability a unique giant component of order n and all other components small, of order at most $\log^7 n$, whereas for $u > u_\star$ it has with high probability all components small. Moreover, we show that $u_\star$ coincides with the critical parameter of random interlacements on a Poisson-Galton-Watson tree, which was identified in [Tas10].

5.1. Introduction

Recently, several authors have been studying percolative properties of the *vacant set* left by random walk on finite graphs and the connections of this problem to the *random interlacements* model introduced in [Szn10]. The topic was initiated with the study of random walk on the d-dimensional discrete torus in [BS08], which was further investigated in [TW11]. [ČTW11, ČT13, CF11] studied random walk on the random regular graph, and [CF11] also studied random walk on the Erdős-Rényi random graph above the connectivity threshold.

In this work we consider the supercritical Erdős-Rényi random graph below the connectivity threshold. We prove a phase transition in the component structure of the vacant set left by random walk on the giant component of this graph, and we identify the critical point of this phase transition with the critical parameter of random interlacements on a Poisson-Galton-Watson tree.

We start by introducing some notation to precisely state the result. Let $\mathbb{P}_{n,p}$ be the law of an Erdős-Rényi random graph, i.e. a random graph G such that every possible edge is present independently with probability $p = \frac{\rho}{n}$, defined on the space $\mathcal{G}(n)$ of graphs with vertex set $\{1, 2, ..., n\}$ endowed with the σ-algebra $\mathbb{G}_n$ of all subsets. It is well known that the component structure of G varies with the parameter ρ (see e.g. [ER61, Bol01, JŁR00, Dur10]). We will in this paper consider such a random graph for a fixed constant $\rho > 1$. In this case, with probability tending to 1 as $n \to \infty$, the graph G is supercritical:

There exists a unique largest connected component $\mathcal{C}_1(G)$ of size approximately ξn, the so-called giant component. Here, ξ is the unique solution in $(0,1)$ of $e^{-\rho\xi} = 1-\xi$.

For a graph G on n vertices and its largest connected component $\mathcal{C}_1 = \mathcal{C}_1(G)$ (determined by some arbitrary tie-breaking rule), let $P^{\mathcal{C}_1}$ be the law of the simple discrete-time random walk $(X_k)_{k\geq 0}$ on $\mathcal{C}_1$ started from its stationary distribution, defined on the space $\{1,2,...,n\}^{\mathbb{N}_0}$ of trajectories on n vertices endowed with the cylinder-σ-algebra $\mathbb{F}_n$. Let $\Omega_n = \mathcal{G}(n) \times \{1,2,...,n\}^{\mathbb{N}_0}$ endowed with the product σ-algebra $\mathbb{G}_n \times \mathbb{F}_n$, and define the *annealed* measure by

$$\mathbf{P}_n(A \times B) = \sum_{G \in A} \mathbb{P}_{n,p}(G) P^{\mathcal{C}_1(G)}(B) \qquad \text{for } A \in \mathbb{G}_n,\ B \in \mathbb{F}_n. \tag{5.1.1}$$

On the product space Ω_n we define the *vacant set* of the random walk at level u as

$$\mathcal{V}^u = \mathcal{C}_1 \setminus \{X_k :\ 0 \leq k \leq u\rho(2-\xi)\xi n\}. \tag{5.1.2}$$

We refer to Remark 5.1.2 for an explanation of this somewhat unusual time scaling. Let $\mathcal{C}_1(\mathcal{V}^u)$ and $\mathcal{C}_2(\mathcal{V}^u)$ be the largest and second largest connected components of the subgraph induced by $\mathcal{V}^u$.

Theorem 5.1.1. *The component structure of the subgraph induced by $\mathcal{V}^u$ exhibits a phase transition at a critical value $u_\star$:*

- *For $u < u_\star$, there are positive constants $\zeta(u,\rho) \in (0,1)$, $C < \infty$, such that for every $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbf{P}_n \left[\left| \frac{|\mathcal{C}_1(\mathcal{V}^u)|}{n} - \zeta(u,\rho) \right| \leq \epsilon \right] = 1, \tag{5.1.3}$$

$$\lim_{n\to\infty} \mathbf{P}_n \left[\frac{|\mathcal{C}_2(\mathcal{V}^u)|}{\log^7 n} \leq C \right] = 1. \tag{5.1.4}$$

- *For $u > u_\star$, there is a positive constant $C < \infty$, such that*

$$\lim_{n\to\infty} \mathbf{P}_n \left[\frac{|\mathcal{C}_1(\mathcal{V}^u)|}{\log^7 n} \leq C \right] = 1. \tag{5.1.5}$$

The critical parameter $u_\star$ is the same as the critical parameter of random interlacements on a Poisson(ρ)-Galton-Watson tree conditioned on non-extinction, which is by [Tas10] given as the solution of a certain equation.

We refer to Section 5.2.3 for a short summary of the used results on random interlacements and its critical parameter, and the derivation of the characterizing equation (5.2.15) for $u_\star$. The constant $\zeta(u,\rho)$ is given as the solution of equation (5.5.2).

Theorem 5.1.1 confirms the following general principle: The vacant set of random walk on a sufficiently fast mixing graph exhibits a phase transition and the critical point is related to the critical value of random interlacements on the corresponding infinite volume limit.

This principle has been investigated recently in several other situations. Results that are more detailed than Theorem 5.1.1 are known to hold for random walk on a random d-regular graph on n vertices run up to time un: [ČTW11] and with different methods [CF11]

proved the phase transition in the component structure of the vacant graph, [ČTW11] identified the critical parameter $u_\star$ with the critical value of random interlacements on the infinite d-regular tree, and [ČT13] showed that there is a critical window of width $n^{-\frac{1}{3}}$ around $u_\star$ in which the largest component is of order $n^{\frac{2}{3}}$. [CF11] used their methods to also prove a phase transition for random walk on the Erdős-Rényi random graph above the connectivity threshold ($\rho \gg \log n$). Weaker statements are known for random walk run up to time uN^d on the discrete d-dimensional torus of side-length N, see [BS08] and [TW11]. The statements in this case are proved for u small or large enough respectively, but it is only conjectured that there is indeed a phase transition at a critical parameter $u_\star$ that coincides with the critical value of random interlacements on $\mathbb{Z}^d$ (cf. Conjecture 2.6 in [ČT12]). We believe that in our case, as in [ČT13] for the random regular graph, it should be possible to prove the existence of a critical window around the critical point. We did not further investigate this.

The main difficulties in proving Theorem 5.1.1 compared to previous results are that our graph, i.e. the giant component of an Erdős-Rényi random graph, is of random size and non-regular. The proof consists of three main steps. The key idea of the first step is the following 'spatial Markov property' of random walk on a random graph. Instead of sampling a random graph and performing random walk on the fixed graph, one can consider sites unvisited by the random walk as not yet sampled sites of the random graph. Then the unvisited or vacant part of the graph has the law of some random graph, depending on the random graph model. In the case of a connected Erdős-Rényi random graph the vacant part is again an Erdős-Rényi random graph, this was used to prove the phase transition in [CF11]. In the case of a random regular graph the vacant part is a random graph with a given degree sequence, a well-studied object (see e.g. [HM12]). This was used to prove the phase transition in [CF11] and the critical behavior in [ČT13].

The situation in our case is more involved, because we consider random walk only on the giant component of a not connected Erdős-Rényi random graph. This random walk cannot satisfy such a spatial Markov property, since the graph must be fixed in advance for the giant component to be known. To be able to still use the idea, we introduce in Algorithm 5.4.1 a process $\bar{X} = (\bar{X}_k)_{k\geq 0}$ on an Erdős-Rényi random graph that behaves like a random walk but jumps to another component after having covered a component. In Lemma 5.4.2 we make precise the aforementioned spatial Markov property for this process $\bar{X}$, namely that the vacant graph left by $\bar{X}$ still has the law of an Erdős-Rényi random graph, but with different parameters. The classical results on random graphs imply a phase transition for this vacant graph.

In a second step we translate this phase transition to the vacant graph left by the simple random walk $X = (X_k)_{k\geq 0}$ on the giant component. To this end, we introduce in Proposition 5.4.3 a coupling of X and $\bar{X}$ where the two processes are with high probability identified in a certain time interval. This can be done because the process $\bar{X}$ will typically 'find' the giant component after a short time and then stay on it long enough.

The third step, requiring most of the technical work, is the identification of the critical point of the phase transition. From Lemma 5.4.2 it is clear that the crucial quantity deciding the critical point is the size of the vacant set left by $\bar{X}$. The coupling of X and $\bar{X}$ has the property that the sizes of the vacant sets of X and $\bar{X}$ are closely related (Lemma 5.4.4), which allows to reduce the problem to the investigation of the size of the vacant set left by X. The first part of this paper, Section 5.3, is devoted to this

investigation. In Proposition 5.3.1 we will on one hand compute the expectation of the size of the vacant set left by X, and on the other hand we will show that the size of the vacant set left by X is concentrated around its expectation.

We close the introduction with a remark on the connection to random interlacements and a heuristic explanation of the time scaling $u\rho(2-\xi)\xi n$ that appears in the definition (5.1.2) of $\mathcal{V}^u$. For readers unfamiliar with random interlacements and the notation, we refer to Section 5.2, in particular Section 5.2.3.

Remark 5.1.2. In the giant component $\mathcal{C}_1$ of an Erdős-Rényi random graph the balls $B(x,r)$ around a vertex x with radius r of order $\log n$ typically look like balls around the root $\varnothing$ in a Poisson(ρ)-Galton-Watson tree $\mathcal{T}$ conditioned on non-extinction. One expects that random interlacements on $\mathcal{T}$ give a good description of the trace of random walk on $\mathcal{C}_1$ locally in such balls, where the intensity u of random interlacements is proportional to the running time of the walk. To determine the proportionality factor, we compare the probability that a vertex $x \in \mathcal{C}_1$ has not been visited by the random walk on $\mathcal{C}_1$ up to time t with the probability that the root $\varnothing \in \mathcal{T}$ is in the vacant set of random interlacements on $\mathcal{T}$ at level u.

Note first that the probability that the random walk on $\mathcal{C}_1$ started at x leaves a ball of large radius around x before returning to x is approximately the same as the probability that the random walk on $\mathcal{T}$ started at the root never returns to the root,

$$P_x^{\mathcal{C}_1}[\tilde{H}_x > H_{B(x,r)^c}] \approx P_\varnothing^{\mathcal{T}}[\tilde{H}_\varnothing = \infty]. \tag{5.1.6}$$

The main task of Section 5.3 will be rigorous proof of the following approximation for the random walk on $\mathcal{C}_1$,

$$P^{\mathcal{C}_1}[x \text{ is vacant at time } t] \approx e^{-tP_x^{\mathcal{C}_1}[\tilde{H}_x > H_{B(x,r)^c}]\pi(x)}. \tag{5.1.7}$$

We will also show that the average degree of a vertex in $\mathcal{C}_1$ is $\rho(2-\xi)$, and so the stationary distribution π of the random walk on $\mathcal{C}_1$ is $\pi(x) \approx \frac{\deg(x)}{\rho(2-\xi)\xi n}$. On the other hand, according to [Tei09], the law Q^u of the vacant set of random interlacements on the infinite graph $\mathcal{T}$ at level u satisfies

$$Q^u[\varnothing \text{ is vacant}] = e^{-u \operatorname{cap}_{\mathcal{T}}(\varnothing)}, \tag{5.1.8}$$

where the *capacity* is here $\operatorname{cap}_{\mathcal{T}}(\varnothing) = \deg(\varnothing)P_\varnothing^{\mathcal{T}}[\tilde{H}_\varnothing = \infty]$. As argued above, random interlacements describe the random walk locally, so the probabilities (5.1.7) and (5.1.8) should be approximately equal for the time t corresponding to random interlacements at level u. The approximation of $\pi(x)$ together with (5.1.6) leads to $t = u\rho(2-\xi)\xi n$ if the parameter u in both models should be the same.

Compared to the time scalings uN^d and un in the discussions of random walk on the torus [BS08, TW11] and random regular graphs [ČTW11, ČT13] respectively, where only the size of the graph (in our case the factor ξn) appears in the time scaling, the additional factor $\rho(2-\xi)$ for the average degree might be surprising. It is however only a consequence of how one defines the uniform edge-weight on the underlying graph, which scales the capacity by a constant. For the aforementioned $2d$-regular graphs the weight chosen is $\frac{1}{2d}$. For non-regular graphs it is the canonical choice to define edge-weights as 1, as is done in [Tei09] and [Tas10], and we stick to this definition.

The paper is structured as follows. In Section 5.2 we introduce some further notation and recall some facts on random graphs, random walks, and random interlacements. In Section 5.3 we investigate the size of the vacant set left by the simple random walk X on the giant component. In Section 5.4 we introduce the process $\bar{X}$ and compare it to the random walk X. Finally, we gather all intermediate results to prove Theorem 5.1.1 in Section 5.5.

Acknowledgment. The author would like to thank Jiří Černý for suggesting the problem and for helpful discussions, and the referee for carefully reading the manuscript and giving important comments that helped to improve the paper.

5.2. Notation and preliminaries

We will denote by c, c', c'' positive finite constants with values changing from place to place. ϵ will always denote a small positive constant with value changing from place to place. All these constants may depend on u and ρ, but not on any other object. We will tacitly assume that values like $u\rho(2-\xi)\xi n$, $\log^5 n$, n^ϵ etc. are integers, omitting to take integer parts to ease the notation.

We use the standard o- and O-notation: Given a positive function $g(n)$, a function $f(n)$ is $o(g)$ if $\lim_{n\to\infty} f/g = 0$, and it is $O(g)$ if $\limsup_{n\to\infty} |f|/g < \infty$. We extend this notation to random variables in the following way. For a random variable A_n on a space (Ω_n, Q_n) we use the notation '$A_n = f(n) + o(g)$ Q_n-asymptotically almost surely' meaning '$\forall\ \epsilon > 0$, $Q_n[|A_n - f(n)| \le \epsilon g(n)] \to 1$ as $n \to \infty$', and '$A_n = O(g)$ Q_n-asymptotically almost surely' meaning '$\exists\ C > 0$ such that $Q_n[|A_n| \le Cg(n)] \to 1$ as $n \to \infty$'.

5.2.1. (Random) graphs

For a non-oriented graph we use the notation G to denote the set of vertices in the graph as well as the graph itself, consisting of vertex-set and edge-set. For vertices $x, y \in G$, $x \sim y$ means that x and y are neighbors, i.e. $\{x, y\}$ is an edge of G. We denote by $\deg(x)$ the number of neighbors of x in G, and by $\Delta_G = \max_{x\in G} \deg(x)$ the maximum degree. By $\mathrm{dist}(x, y)$ we denote the usual graph distance, and for $r \in \mathbb{N}$, $B(x, r)$ is the set of vertices y with $\mathrm{dist}(x, y) \le r$. For a subset $A \subset G$, denote its complement $A^c = G \setminus A$ and its (interior) boundary $\partial A = \{x \in A :\ \exists y \in A^c,\ x \sim y\}$.

We denote by $\mathcal{C}_i(G)$ the i-th largest connected component of a graph G. If there are equally large components, we order these arbitrarily. The subgraph induced by a vertex-set $V \subset G$ is defined as the graph with vertices V and edges $\{x, y\}$ if and only if $x, y \in V$ and $x \sim y$ in G. Again we use the notation $\mathcal{C}_i(G)$ for the set of vertices as well as for the induced subgraph. Usually (but not necessarily) $\mathcal{C}_1 = \mathcal{C}_1(G)$ will be the unique giant component. A graph or graph component is called *simple* if it is connected and has at most one cycle, i.e. the number of edges is at most equal to the number of vertices.

Recall from the introduction that $\mathbb{P}_{n,p}$ denotes the law of an Erdős-Rényi random graph, i.e. a random graph on n vertices such that every edge is present independently with probability $p = \frac{\rho}{n}$. Let $\mathbb{E}_{n,p}$ be the corresponding expectation. An event is said to hold *asymptotically almost surely* (a.a.s.) if it holds with probability tending to 1 as $n \to \infty$ (cf. the above defined o- and O-notation). Throughout this work $\rho > 1$ is a fixed

constant. It is well known that the following properties then hold $\mathbb{P}_{n,p}$-a.a.s.

$$\text{The graph } G \text{ has a unique giant component } \mathcal{C}_1 \text{ of size } |\mathcal{C}_1| \text{ satisfying } ||\mathcal{C}_1| - \xi n| \leq n^{3/4}, \text{ where } \xi \text{ is the unique solution in } (0,1) \text{ of } e^{-\rho\xi} = 1 - \xi. \text{ All other components are simple and of size smaller than } C \log n, \text{ for some fixed constant } C. \tag{5.2.1}$$

$$\text{The spectral gap } \lambda_{\mathcal{C}_1} \text{ of the random walk on the giant component (cf. (5.2.12)) satisfies } \lambda_{\mathcal{C}_1} \geq \tfrac{c}{\log^2 n} \text{ for some fixed constant } c. \tag{5.2.2}$$

$$\text{The maximum degree } \Delta_G \text{ satisfies } \Delta_G \leq \log n. \tag{5.2.3}$$

(5.2.1) and (5.2.3) are classical results (see e.g. [ER61, Bol01, JŁR00, Dur10]), and (5.2.2) follows from [LPW09, Theorem 12.4] with the $O(\log^2 n)$ bound on the mixing time of the random walk on the giant component proved in [BKW14]. We use the terminology *typical graphs* for graphs G on n vertices satisfying (5.2.1), (5.2.2) and (5.2.3). We will usually prove our statements for typical graphs only, since we are interested in a.a.s.-behavior.

For a quantitative version of the first statement in (5.2.1) see [vdH08, Theorem 4.8], which states that

$$\mathbb{P}_{n,p}\left[||\mathcal{C}_1| - \xi n| > n^{3/4}\right] \leq cn^{-c'}. \tag{5.2.4}$$

The choice of the constant 3/4 is arbitrary.

We will also need a quantitative version of (5.2.3), we therefore briefly present a proof. Fix a vertex $x \in G$ and denote all other vertices by y_i, $i = 1, ..., n-1$. Let $\mathcal{E}_i = \mathbf{1}_{\{\{x,y_i\} \text{ is an edge}\}}$. Then the $\mathcal{E}_i$ are i.i.d. Bernoulli(p) random variables, $\deg(x) = \sum_{i=1}^{n-1} \mathcal{E}_i$, and for any fixed $\alpha > 0$ by the exponential Chebyshev inequality,

$$\mathbb{P}_{n,p}[\deg(x) > \log n] \leq n^{-\alpha}\mathbb{E}_{n,p}\left[e^{\alpha \sum \mathcal{E}_i}\right] = n^{-\alpha}\left(1 + \frac{\rho}{n}(e^\alpha - 1)\right)^{n-1} \leq cn^{-\alpha},$$

where the constant c depends on α. We choose $\alpha = 4$, this will be suitable for our purposes. Then a union bound implies

$$\mathbb{P}_{n,p}[\Delta_G > \log n] \leq n\mathbb{P}_{n,p}[\deg(x) > \log n] \leq cn^{1-\alpha} = cn^{-3}. \tag{5.2.5}$$

5.2.2. Random walks

Let $P^{\mathcal{C}_1}$ be the law and $E^{\mathcal{C}_1}$ the corresponding expectation of the simple discrete-time random walk $X = (X_k)_{k\geq 0}$ on the component $\mathcal{C}_1$ started stationary, i.e. the law of the Markov chain with state space $\mathcal{C}_1$, transition probabilities $p_{xy} = \frac{1}{\deg(x)}\mathbf{1}_{\{x\sim y\}}$ and $X_0 \sim \pi$, where π is the stationary distribution, $\pi(x) = \frac{\deg(x)}{\sum_{y\in\mathcal{C}_1}\deg(y)}$. (5.2.1) and the a.a.s. upper bound (5.2.3) on the maximum degree Δ_G imply the following bounds on π. $\mathbb{P}_{n,p}$-a.a.s.

$$\pi(x) = \frac{\deg(x)}{\sum_{v\in\mathcal{C}_1}\deg(v)} \leq \frac{c\log n}{n}, \tag{5.2.6}$$

$$\pi(x) = \frac{\deg(x)}{\sum_{v\in\mathcal{C}_1}\deg(v)} \geq \frac{c}{n\log n}. \tag{5.2.7}$$

For real numbers $0 \leq s \leq r$ denote by $X_{[s,r]} = \{X_k : s \leq k \leq r\}$ the set of vertices visited by X between times s and r. We let the random walk X run up to time t and

denote by $\mathcal{V}(t) = \mathcal{C}_1 \setminus X_{[0,t]}$ the vacant set left by the random walk at time t, and again we use the notation $\mathcal{V}(t)$ to also denote the subgraph of $\mathcal{C}_1$ induced by these vertices. As defined in (5.1.2), we will use the short notation $\mathcal{V}^u$ for $\mathcal{V}(u\rho(2-\xi)\xi n)$.

We will, where it is clear in the context, drop the superscript from $P^{\mathcal{C}_1}$ and $E^{\mathcal{C}_1}$. The notation P_x is then used to denote the law of the random walk on $\mathcal{C}_1$ started at vertex x, E_x is the corresponding expectation. For a set $A \subset \mathcal{C}_1$ we denote by

$$H_A = \inf\{t \geq 0 : \ X_t \in A\}, \qquad \tilde{H}_A = \inf\{t \geq 1 : \ X_t \in A\}$$

the entrance time and hitting time respectively of A, and we write H_x and $\tilde{H}_x$ if $A = \{x\}$. From [AB92, Lemma 2] or [AF02, Proposition 3.21] together with (5.2.6) we get the following bound on $E[H_x]$. $\mathbb{P}_{n,p}$-a.a.s. for all $x \in \mathcal{C}_1$,

$$E[H_x] \geq \frac{(1-\pi(x))^2}{\pi(x)} \geq \frac{cn}{\log n}. \tag{5.2.8}$$

For all real valued functions f and g on $\mathcal{C}_1$ define the *Dirichlet form*

$$\mathcal{D}(f,g) = \frac{1}{2} \sum_{x,y \in \mathcal{C}_1} (f(x) - f(y))(g(x) - g(y))\pi(x)p_{xy}. \tag{5.2.9}$$

A function f on $\mathcal{C}_1$ is *harmonic* on $A \subset \mathcal{C}_1$ if $\sum_y p_{xy} f(y) = f(x)$ for $x \in A$. For $x \in \mathcal{C}_1$ and $r \in \mathbb{N}$ define the *equilibrium potential* $g^\star : \mathcal{C}_1 \to \mathbb{R}$ as the unique function harmonic on $B(x,r) \setminus \{x\}$, 1 on $\{x\}$ and 0 on $B(x,r)^c$. The dependence of $g^\star$ on x and r is kept implicit. Then it is well known that

$$g^\star(y) = P_y\big[H_x < H_{B(x,r)^c}\big], \tag{5.2.10}$$

$$\mathcal{D}(g^\star, g^\star) = P_x\big[\tilde{H}_x > H_{B(x,r)^c}\big]\pi(x). \tag{5.2.11}$$

The spectral gap of the random walk on $\mathcal{C}_1$ is given by

$$\lambda_{\mathcal{C}_1} = \min\{\mathcal{D}(f,f) : \ \pi(f^2) = 1, \ \pi(f) = 0\}. \tag{5.2.12}$$

The relevance of the bound (5.2.2) on $\lambda_{\mathcal{C}_1}$ is in the speed of mixing of the random walk on $\mathcal{C}_1$. From [LPW09, Theorem 12.3 and Lemma 6.13] it follows that for all $t \in \mathbb{N}$

$$\max_{x,y \in \mathcal{C}_1} |P_x[X_t = y] - \pi(y)| \leq \frac{1}{\min_{z \in \mathcal{C}_1} \pi(z)} e^{-\lambda_{\mathcal{C}_1} t}. \tag{5.2.13}$$

5.2.3. Random interlacements

Random interlacements were introduced in [Szn10] on $\mathbb{Z}^d$ as a model to describe the local structure of the trace of a random walk on a large discrete torus, and in [Tei09] the model was generalized to arbitrary transient graphs. It is a special dependent site percolation model where the occupied vertices on a graph are constructed as the trace left by a Poisson point process on the space of doubly-infinite trajectories modulo time shift. The density of this Poisson point process is determined by a parameter $u > 0$. The critical value $u_\star$ is the infimum over the u for which almost surely all connected components of non-occupied vertices are finite.

In [Tas10] it is shown that for random Galton-Watson trees the critical value $u_\star$ is almost surely constant with respect to the tree measure and is implicitly given as the solution of a certain equation. Except for the identification of the critical parameter of Theorem 5.1.1 with this $u_\star$ as the solution of the same equation, we will not use any results on random interlacements. We refer to the lecture notes [ČT12] for an introduction to random interlacements and many more references.

We quote the result from [Tas10] to derive the characterizing equation for $u_\star$ in the case of a Poisson-Galton-Watson tree. This requires some more notation. Denote by $\mathbb{P}_\mathcal{T}$ the law of the supercritical Poisson(ρ)-Galton-Watson rooted tree conditioned on non-extinction and by $\mathbb{E}_\mathcal{T}$ the corresponding conditional expectation. Let $f(s) = e^{\rho(s-1)}$ be the probability generating function of the Poisson(ρ) distribution, and denote by q the extinction probability of a (unconditioned) Poisson(ρ)-Galton-Watson tree. It is well known that q is the unique solution in $(0,1)$ of the equation $f(s) = s$, and hence $q = 1-\xi$, where ξ is as in (5.2.1). Let

$$\tilde{f}(s) = \frac{f((1-q)s+q) - q}{1-q}. \tag{5.2.14}$$

This is in fact the probability generating function of the offspring in the subtree of vertices with infinite line of descent (see e.g. [LP14, Proposition 5.26]).

Consider the simple discrete-time random walk $(X_k)_{k\geq 0}$ on the rooted tree $\mathcal{T}$ started at the root $\varnothing$, whose law we denote by $P_\varnothing^\mathcal{T}$, and let $\tilde{H}_\varnothing = \inf\{t \geq 1 : X_t = \varnothing\}$ be the hitting time of the root. Define the *capacity* of the root by $\mathrm{cap}_\mathcal{T}(\varnothing) = \deg(\varnothing) P_\varnothing^\mathcal{T}[\tilde{H}_\varnothing = \infty]$.

By [Tas10, Theorem 1], the critical parameter $u_\star$ of random interlacements on the Galton-Watson tree conditioned on non-extinction is $\mathbb{P}_\mathcal{T}$-a.s. constant and given as the unique solution in $(0,\infty)$ of the equation

$$\left(\tilde{f}^{-1}\right)'\left(\mathbb{E}_\mathcal{T}\left[e^{-u\,\mathrm{cap}_\mathcal{T}(\varnothing)}\right]\right) = 1.$$

In particular for the Poisson(ρ)-Galton-Watson tree,

$$\left(\tilde{f}^{-1}\right)'(t) = \frac{1}{\rho\xi t + \rho(1-\xi)},$$

and $u_\star$ is the solution of

$$\rho\xi\mathbb{E}_\mathcal{T}\left[e^{-u\,\mathrm{cap}_\mathcal{T}(\varnothing)}\right] + \rho(1-\xi) = 1. \tag{5.2.15}$$

5.3. Size of the vacant set

In this section we investigate the size of the vacant set $\mathcal{V}^u$ left by the random walk X on the giant component $\mathcal{C}_1$. As already mentioned we omit the superscripts from $P^{\mathcal{C}_1}$ and $E^{\mathcal{C}_1}$. Recall the definition (5.1.1) of the annealed measure $\mathbf{P}_n$.

Proposition 5.3.1.

(1) $E[|\mathcal{V}^u|]$ can asymptotically be approximated in terms of a Poisson(ρ)-Galton-Watson tree conditioned on non-extinction:

$$E[|\mathcal{V}^u|] = \xi n \mathbb{E}_{\mathcal{T}}\left[e^{-u\,\mathrm{cap}_{\mathcal{T}}(\varnothing)}\right] + o(n) \qquad \mathbb{P}_{n,p}\text{-a.a.s.}$$

(2) The random variable $|\mathcal{V}^u|$ is concentrated around its mean:

$$|\mathcal{V}^u| = E[|\mathcal{V}^u|] + o(n) \qquad \mathbf{P}_n\text{-a.a.s.}$$

5.3.1. Expectation of the size of the vacant set

The proof of part (1) of Proposition 5.3.1 is split up into several steps. We first quote and extend [JLT14, Proposition 11.2]. It formalizes the well known fact that an Erdős-Rényi random graph locally looks like a Galton-Watson tree. Here, by locally we mean balls of radius of order $\log n$. More precisely, fix some $\gamma > 0$ such that $6\gamma \log \rho < 1$, and set

$$r = \gamma \log n. \tag{5.3.1}$$

For a graph G, a vertex $x \in G$ and a tree $\mathcal{T}$ with root $\varnothing$, define the event

$$\mathcal{I}_x(G,\mathcal{T}) = \left\{ \begin{matrix} B(x,r+1) \subset G \text{ is isomorphic to } B(\varnothing,r+1) \subset \mathcal{T}, \\ \text{with the isomorphism sending } x \text{ to } \varnothing \end{matrix} \right\}. \tag{5.3.2}$$

Denote by $\mathbb{P}^0_{\mathcal{T}}$ the law of the unconditioned Poisson(ρ)-Galton-Watson tree $\mathcal{T}$, and by $\{|\mathcal{T}| < \infty\}$, $\{|\mathcal{T}| = \infty\}$ the events of extinction and non-extinction respectively of the tree $\mathcal{T}$.

Proposition 5.3.2.

(1) Given an arbitrary fixed vertex $x \in \{1,2,...,n\}$, there is a coupling Q_x of G under $\mathbb{P}_{n,p}$ and a tree $\mathcal{T}$ under $\mathbb{P}^0_{\mathcal{T}}$, such that for n large enough

$$Q_x\left[\mathcal{I}_x(G,\mathcal{T})\right] \geq 1 - cn^{3\gamma\log\rho - 1}. \tag{5.3.3}$$

For n large enough, this coupling satisfies

$$Q_x[x \in \mathcal{C}_1,\ |\mathcal{T}| < \infty] \leq cn^{-c'}, \tag{5.3.4}$$

$$Q_x[x \notin \mathcal{C}_1,\ |\mathcal{T}| = \infty] \leq cn^{-c'}. \tag{5.3.5}$$

(2) For an arbitrary point $x \in G$, with r as in (5.3.1),

$$\mathbb{P}_{n,p}\left[|B(x,r)| \geq n^{3\gamma\log\rho}\right] \leq cn^{3\gamma\log\rho - 1}. \tag{5.3.6}$$

(3) Given two arbitrary fixed vertices $x \neq y$, there is a coupling $Q_{x,y}$ of G under $\mathbb{P}_{n,p}$ and two trees $\mathcal{T}_x$ and $\mathcal{T}_y$, each having law $\mathbb{P}^0_{\mathcal{T}}$, such that $\mathcal{T}_x$ and $\mathcal{T}_y$ are independent and for n large enough

$$Q_{x,y}\left[\mathcal{I}_x(G,\mathcal{T}_x) \text{ and } \mathcal{I}_y(G,\mathcal{T}_y)\right] \geq 1 - cn^{6\gamma\log\rho - 1}, \tag{5.3.7}$$

and statements (5.3.4) and (5.3.5) hold under $Q_{x,y}$ for x, $\mathcal{T}_x$ and y, $\mathcal{T}_y$ respectively.

Proof. (5.3.3) is, up to the enlargement of the radius by 1, the statement of [JLT14, Proposition 11.2], and (5.3.6) is [JLT14, Corollary 11.3]. Note that, in contrary to the actual statement, [JLT14, Proposition 11.2] is proved for an a priori fixed vertex and not a randomly chosen one.

For part (1) it remains to show the properties (5.3.4) and (5.3.5). For simplicity write $B_x = B(x,r) \subset G$ and $B_\varnothing = B(\varnothing, r) \subset \mathcal{T}$. Denote by $\{z \leftrightarrow B_z^c\}$ the event that z is connected to the complement of B_z, or equivalently that ∂B_z is non-empty, and by $\{z \not\leftrightarrow B_z^c\}$ its complement. To prove (5.3.4), we first claim that

$$\mathbb{P}_{n,p}[x \in \mathcal{C}_1,\ x \not\leftrightarrow B_x^c] \leq cn^{-c'}. \tag{5.3.8}$$

To see this, note that if $x \in \mathcal{C}_1$ and $x \not\leftrightarrow B_x^c$, then $B_x = \mathcal{C}_1$. But by (5.3.6), B_x is unlikely to be large: For every small $\epsilon > 0$, $\mathbb{P}_{n,p}[|B_x| \geq n^{1-\epsilon}] \leq cn^{-c'}$. However, if B_x is smaller than $n^{1-\epsilon}$ and $B_x = \mathcal{C}_1$, then $\mathcal{C}_1$ is smaller than $n^{1-\epsilon}$, but this happens with probability smaller than $cn^{-c'}$ by (5.2.4), and (5.3.8) follows.

Note that if the coupling succeeds, i.e. the balls of radius $r+1$ are isomorphic, then $\{x \leftrightarrow B_x^c\} = \{\varnothing \leftrightarrow B_\varnothing^c\}$. This happens with probability $\geq 1 - cn^{-c'}$ by (5.3.3), so together with (5.3.8),

$$\begin{aligned} Q_x[x \in \mathcal{C}_1,\ |\mathcal{T}| < \infty] &\leq Q_x[x \leftrightarrow B_x^c,\ |\mathcal{T}| < \infty] + cn^{-c'} \\ &\leq Q_x[\varnothing \leftrightarrow B_\varnothing^c,\ |\mathcal{T}| < \infty] + cn^{-c'} \\ &= \mathbb{P}_{\mathcal{T}}^0[\varnothing \leftrightarrow B_\varnothing^c,\ |\mathcal{T}| < \infty] + cn^{-c'}. \end{aligned}$$

The tree $\mathcal{T}$ conditioned on extinction has the law of a subcritical Galton-Watson tree with mean offspring number $m < 1$ (see e.g. [LP14, Proposition 5.26]). If q is the extinction probability and Z_k denotes the size of the k-th generation of the tree, we can use the Markov inequality to get

$$\begin{aligned} \mathbb{P}_{\mathcal{T}}^0[\varnothing \leftrightarrow B_\varnothing^c,\ |\mathcal{T}| < \infty] &= \mathbb{P}_{\mathcal{T}}^0\left[Z_r \geq 1 \;\middle|\; |\mathcal{T}| < \infty\right] q \\ &\leq \mathbb{E}_{\mathcal{T}}^0\left[Z_r \;\middle|\; |\mathcal{T}| < \infty\right] q = qm^{\gamma \log n} = cn^{-c'}, \end{aligned}$$

which proves (5.3.4).

For (5.3.5), let C_x be the component of G containing x. Let $M > 0$ be such that $M\gamma > (\rho - 1 - \log \rho)^{-1}$. Then, by e.g. [Dur10, Theorem 2.6.4], $\mathbb{P}_{n,p}[x \notin \mathcal{C}_1,\ |C_x| > M\gamma \log n] \leq cn^{-c'}$. Using this on the first line and (5.3.3) on the second, it follows that

$$\begin{aligned} Q_x[x \notin \mathcal{C}_1,\ |\mathcal{T}| = \infty] &\leq Q_x[|C_x| \leq M\gamma \log n,\ |\mathcal{T}| = \infty] + cn^{-c'} \\ &\leq Q_x[|B_\varnothing| \leq M\gamma \log n,\ |\mathcal{T}| = \infty] + cn^{-c'}. \end{aligned}$$

To bound this latter probability that the ball of radius $r = \gamma \log n$ in a surviving Poisson(ρ)-Galton-Watson tree is smaller than Mr, let again Z_r be the size of the r-th generation and denote by $Z_r^\star$ the number of particles in the r-th generation with infinite line of descent. Then

$$\begin{aligned} Q_x[|B_\varnothing| \leq Mr,\ |\mathcal{T}| = \infty] &\leq \mathbb{P}_{\mathcal{T}}^0[Z_r \leq Mr \mid |\mathcal{T}| = \infty]\mathbb{P}_{\mathcal{T}}^0[|\mathcal{T}| = \infty] \\ &\leq \mathbb{P}_{\mathcal{T}}^0[Z_r^\star \leq Mr \mid |\mathcal{T}| = \infty]\xi. \end{aligned}$$

By e.g. [LP14, Proposition 5.26] or [AN72, Theorem I.12.1]

$$\mathbb{P}^0_{\mathcal{T}}[Z^\star_r \le Mr \mid |\mathcal{T}| = \infty] = \tilde{\mathbb{P}}_{\mathcal{T}}[\tilde{Z}_r \le Mr],$$

where $\tilde{Z}_r$ under $\tilde{\mathbb{P}}_{\mathcal{T}}$ is the r-th generation size of a Galton-Watson tree with offspring distribution defined by the probability generating function $\tilde{f}$ as in (5.2.14), a tree with extinction probability $\tilde{q} = 0$. Let $\kappa = \tilde{f}'(0) = f'(q)$. Since f, the probability generating function of Poisson(ρ), is strictly convex and increasing, and by definition of $q = 1-\xi$, we have $0 < \kappa < 1$. Let $\tilde{f}_r$ be the r-th iterate of $\tilde{f}$, which is in fact the probability generating function of $\tilde{Z}_r$. From [AN72, Corollary I.11.1] we know that

$$\lim_{r\to\infty} \kappa^{-r} \tilde{f}_r(s) = Q(s) \in (0,\infty) \quad \text{exists for } 0 \le s < 1.$$

It follows that

$$\tilde{f}_r(s) \le (Q(s) + \epsilon)\kappa^r$$

for $r \ge r_0(s,\epsilon)$. Using this, for any $\lambda > 0$ we obtain for $r \ge r_0(e^{-\lambda},\epsilon)$

$$\begin{aligned}\tilde{\mathbb{P}}_{\mathcal{T}}[\tilde{Z}_r \le Mr] &\le \tilde{\mathbb{P}}_{\mathcal{T}}[e^{-\lambda\tilde{Z}_r} \ge e^{-\lambda Mr}] \le e^{\lambda Mr}\tilde{f}_r(e^{-\lambda}) \\ &\le (Q(s)+\epsilon)e^{\lambda Mr + r\log\kappa}.\end{aligned}$$

By choosing $\lambda < -\frac{\log\kappa}{M}$ we can make this smaller than $ce^{-c'r}$, and (5.3.5) follows since $r = \gamma \log n$. This finishes the proof of part (1) of the proposition.

We now prove part (3). Define the coupling $Q_{x,y}$ as follows. By using part (1) of the proposition, we can find a coupling of two independent graphs G_x and G_y, both with vertex set $x, y, 3, ..., n$, and two independent Poisson(ρ)-Galton-Watson trees $\mathcal{T}_x$ and $\mathcal{T}_y$, such that with probability larger than $1-2cn^{3\gamma\log\rho-1}$ both $\mathcal{I}_x(G_x,\mathcal{T}_x)$ and $\mathcal{I}_y(G_y,\mathcal{T}_y)$ hold.

We then construct a graph G with the same vertex set $x, y, 3, ..., n$ in the following way. We first explore the ball $B(x, r+1) \subset G$ by determining the state of all possible edges with at least one adjacent vertex in $B(x,r) \subset G_x$ according to their state in G_x, i.e. setting them present or absent. In a second step we determine the ball $B(y, r+1) \subset G$ in the same way by G_y, only that we do not change the state of already determined edges. The remaining edges in G are set present independently with probability p and absent otherwise.

By construction this graph G has law $\mathbb{P}_{n,p}$. If both $\mathcal{I}_x(G_x,\mathcal{T}_x)$ and $\mathcal{I}_y(G_y,\mathcal{T}_y)$ hold and there is no collision in the second step, i.e. we never want to set an edge present that is already set absent or vice versa, then both $\mathcal{I}_x(G,\mathcal{T}_x)$ and $\mathcal{I}_y(G,\mathcal{T}_y)$ hold, and the coupling succeeds. It thus remains to bound the probability of such a collision.

Note that if there is a collision, then the sets of vertices $B(x, r+1)$ and $B(y, r+1)$ must have non-empty intersection: If $B(x,r+1)\cap B(y,r+1) = \emptyset$, the only edges possibly causing a collision are edges $\{u, v\}$ with $u \in B(x,r)$ and $v \in B(y,r)$, but these edges must be set absent by both G_x and G_y, or else $u \in B(y, r+1)$ or $v \in B(x, r+1)$.

The sets $B(x,r+1)$ and $B(y,r+1)$ are smaller than $n^{3\gamma\log\rho}$ with probability larger than $1 - cn^{3\gamma\log\rho-1}$ by (5.3.6), and they are by construction random subsets of $\{x, y, 3, ..., n\}$. But the probability that two random subsets of $\{x, y, 3, ..., n\}$ of size k intersect is smaller than $\frac{k^2}{n}$, so the probability of a collision is smaller than

$$Q_{x,y}[B(x,r+1)\cap B(y,r+1) \ne \emptyset] \le 2cn^{3\gamma\log\rho-1} + \frac{1}{n}n^{6\gamma\log\rho} \le cn^{6\gamma\log\rho-1}.$$

This proves (5.3.7). By construction it is clear that statements (5.3.4) and (5.3.5) hold analogously under $Q_{x,y}$. □

We will denote by $\mathbb{E}_{Q_x}$ and $\mathbb{E}_{Q_{x,y}}$ the expectations corresponding to the couplings Q_x and $Q_{x,y}$. For easier use later we now define some events and estimate their probabilities. Let $\mathcal{B}_x$ on the space of the coupling Q_x be the event

$$\mathcal{B}_x = \mathcal{I}_x(G,\mathcal{T}) \cap \Big(\{x \in \mathcal{C}_1,\ |\mathcal{T}| = \infty\} \cup \{x \notin \mathcal{C}_1,\ |\mathcal{T}| < \infty\}\Big), \tag{5.3.9}$$

This event can canonically also be defined on the space of the coupling $Q_{x,y}$ when replacing $\mathcal{T}$ by $\mathcal{T}_x$. Then define on the space of $Q_{x,y}$ the event

$$\mathcal{B}_{x,y} = \mathcal{B}_x \cap \mathcal{B}_y. \tag{5.3.10}$$

From Proposition 5.3.2 it is immediate that

$$Q_x[\mathcal{B}_x] \geq 1 - cn^{-c'}, \tag{5.3.11}$$

$$Q_{x,y}[\mathcal{B}_{x,y}] \geq 1 - cn^{-c'}. \tag{5.3.12}$$

On the space of the coupling Q_x, and similarly on the space of $Q_{x,y}$, we further define the event

$$\{x \text{ good}\} = \{x \in \mathcal{C}_1\} \cap \{|\mathcal{T}| = \infty\} \cap \mathcal{I}_x(G,\mathcal{T}) = \mathcal{B}_x \cap \{x \in \mathcal{C}_1\}. \tag{5.3.13}$$

Since $\mathbb{P}^0_{\mathcal{T}}[|\mathcal{T}| = \infty] = \xi$ and $\mathbf{1}_{\{x \text{ good}\}} = \mathbf{1}_{\{|\mathcal{T}|=\infty\}} - \mathbf{1}_{\{|\mathcal{T}|=\infty,\ x\notin\mathcal{C}_1\}} - \mathbf{1}_{\{|\mathcal{T}|=\infty,\ x\in\mathcal{C}_1,\ \mathcal{I}_x(G,\mathcal{T})^c\}}$, it follows with (5.3.3) and (5.3.5) that

$$Q_x[x \text{ good}] = \xi + o(1) \text{ as } n \to \infty.$$

Note that the probability of x being good is bounded away from zero, so every graph property holding $\mathbb{P}_{n,p}$-a.a.s., as well as every property of a ball of radius r in a Galton-Watson tree holding $\mathbb{P}^0_{\mathcal{T}}$-a.a.s. as $r \to \infty$ will also hold $Q_x[\ \cdot\ |\ x \text{ good}]$-a.a.s.

As a first application of Proposition 5.3.2 we prove a law of large numbers for the sum of degrees of vertices in the giant component, which leads to an approximation of the stationary measure π. This result may be well known, we did however not find it in the literature. The technique of the proof will be used again later.

Lemma 5.3.3.

$$\sum_{x\in\mathcal{C}_1} \deg(x) = \sum_{x\in G} \mathbf{1}_{\{x\in\mathcal{C}_1\}} \deg(x) = \rho(2-\xi)\xi n + o(n) \qquad \mathbb{P}_{n,p}\text{-a.a.s.}$$

Proof. Every vertex in the random graph G has Binomial$(n-1, \frac{\rho}{n})$ neighbors, but on $\mathcal{C}_1$ their degree is above average and there is some dependency. For $x \in G$ denote

$$Z_x = \mathbf{1}_{\{x\in\mathcal{C}_1\}} \deg(x),$$
$$\tilde{Z}_x = \mathbf{1}_{\{|\mathcal{T}|=\infty\}} \deg(\varnothing),$$

where the tree $\mathcal{T}$ is defined by the coupling Q_x from Proposition 5.3.2, and $\varnothing$ is the root of $\mathcal{T}$. We will approximate $\mathbb{E}_{n,p}[Z_x] = \mathbb{E}_{Q_x}[Z_x]$ by $\mathbb{E}_{Q_x}[\tilde{Z}_x]$ and show that the sum of the Z_x is concentrated around its expectation using the second moment method.

Let us first compute the expectation of $\tilde{Z}_x$. Recall that $\mathbb{P}_{\mathcal{T}}$ denotes the law of the Poisson(ρ)-Galton-Watson tree conditioned on non-extinction, and $\mathbb{E}_{\mathcal{T}}$ the corresponding conditional expectation. Then

$$\mathbb{E}_{Q_x}\left[\tilde{Z}_x\right] = \mathbb{E}^0_{\mathcal{T}}\left[\deg(\varnothing) \;\middle|\; |\mathcal{T}| = \infty\right] \mathbb{P}^0_{\mathcal{T}}\left[|\mathcal{T}| = \infty\right] = \mathbb{E}_{\mathcal{T}}\left[\deg(\varnothing)\right] \xi. \tag{5.3.14}$$

Using the same technique as in the proof of [LP14, Proposition 5.26], it is straightforward to see that the expected offspring in a Galton-Watson tree conditioned on non-extinction is

$$\mathbb{E}_{\mathcal{T}}[\deg(\varnothing)] = \frac{1}{1-q}(f'(1) - qf'(q)),$$

where f is the probability generating function of the offspring distribution. Here, the offspring is Poisson(ρ), so $q = 1 - \xi$, $f'(1) = \rho$ and $f'(q) = \rho(1-\xi)$, which leads to

$$\mathbb{E}_{\mathcal{T}}\left[\deg(\varnothing)\right] = \frac{1}{\xi}(\rho - \rho(1-\xi)^2) = \rho(2-\xi). \tag{5.3.15}$$

We now approximate $\mathbb{E}_{Q_x}[Z_x]$ by $\mathbb{E}_{Q_x}[\tilde{Z}_x]$. Because $\tilde{Z}_x$ is unbounded, we will truncate it by $\log n$. By definition $\tilde{Z}_x$ is stochastically dominated by a Poisson(ρ) random variable Λ, in particular it has finite mean, and therefore $\mathbb{E}_{Q_x}[\tilde{Z}_x \mathbf{1}_{\{\tilde{Z}_x < \log n\}}] \nearrow \mathbb{E}_{Q_x}[\tilde{Z}_x]$ as $n \to \infty$. Using $E[e^{t\Lambda}] = e^{\rho(e^t-1)}$ we have $P[\Lambda \geq \log n] = P[e^{t\Lambda} \geq n^t] \leq e^{\rho(e^t-1)} n^{-t} = cn^{-c'}$. It follows that

$$\begin{aligned}\mathbb{E}_{Q_x}[\tilde{Z}_x \wedge \log n] &= \mathbb{E}_{Q_x}[\tilde{Z}_x \mathbf{1}_{\{\tilde{Z}_x < \log n\}}] + \log n Q_x[\tilde{Z}_x \geq \log n] \\ &= \mathbb{E}_{Q_x}[\tilde{Z}_x] + o(1) \text{ as } n \to \infty.\end{aligned}$$

Recall from (5.3.9) the definition of the event $\mathcal{B}_x$, on which $Z_x = \tilde{Z}_x$, and $Z_x = \tilde{Z}_x \wedge \log n$ if $\Delta_G \leq \log n$. With (5.3.11) and (5.2.5) we can bound

$$\left|\mathbb{E}_{Q_x}[Z_x] - \mathbb{E}_{Q_x}[\tilde{Z}_x \wedge \log n]\right| \leq nQ_x\left[\Delta_G > \log n\right] + \log n Q_x\left[\mathcal{B}_x^c\right] \leq cn^{-c'}. \tag{5.3.16}$$

With (5.3.14) and (5.3.15) it follows that

$$\mathbb{E}_{n,p}\left[\sum_{x \in G} Z_x\right] = n\mathbb{E}_{Q_x}\left[Z_x\right] = \rho(2-\xi)\xi n + o(n) \text{ as } n \to \infty.$$

It remains to show that the sum of the Z_x is concentrated. Take $x \neq y$ arbitrary vertices in G and consider the coupling $Q_{x,y}$ from Proposition 5.3.2. Recall from (5.3.10) the definition of the event $\mathcal{B}_{x,y}$. On $\mathcal{B}_{x,y}$ we have $Z_x = \tilde{Z}_x$ and $Z_y = \tilde{Z}_y$, so with (5.3.12) and (5.2.5) we get

$$\begin{aligned}&\left|\mathbb{E}_{Q_{x,y}}[Z_x Z_y] - \mathbb{E}_{Q_{x,y}}\left[(\tilde{Z}_x \wedge \log n)(\tilde{Z}_y \wedge \log n)\right]\right| \\ &\qquad \leq n^2 Q_{x,y}\left[\Delta_G > \log n\right] + \log^2 n Q_{x,y}\left[\mathcal{B}_{x,y}^c\right] \leq cn^{-c'}.\end{aligned} \tag{5.3.17}$$

The trees $\mathcal{T}_x$ and $\mathcal{T}_y$ are independent, so $\tilde{Z}_x \wedge \log n$ and $\tilde{Z}_y \wedge \log n$ are independent. Therefore, from (5.3.16) and (5.3.17) we conclude that for two arbitrary vertices $x \neq y$,

$$\mathbb{E}_{n,p}[Z_x Z_y] = \mathbb{E}_{n,p}[Z_x]\mathbb{E}_{n,p}[Z_y] + o(1) \text{ as } n \to \infty.$$

Denote $Z = \sum_{x\in G} Z_x$. It follows from the above, together with (5.2.5), that

$$\begin{aligned}\mathbb{E}_{n,p}\left[Z^2\right] &= \sum_{x\in G}\mathbb{E}_{n,p}[Z_x^2] + \sum_{x\neq y}\left(\mathbb{E}_{n,p}[Z_x]\mathbb{E}_{n,p}[Z_y] + o(1)\right)\\ &= O(n\log^2 n) + O(n^3)\mathbb{P}_{n,p}[\Delta_G > \log n] + \mathbb{E}_{n,p}\left[Z\right]^2 - n\mathbb{E}_{n,p}[Z_x]^2 + o(n^2)\\ &= \mathbb{E}_{n,p}\left[Z\right]^2 + o(n^2) \text{ as } n\to\infty.\end{aligned}$$

Thus $\operatorname{Var} Z = o(n^2)$ and the Chebyshev inequality implies for any $\epsilon > 0$

$$\mathbb{P}_{n,p}\left[|Z - \mathbb{E}_{n,p}[Z]| > \epsilon n\right] = o(1) \text{ as } n\to\infty.$$

This finishes the proof of the lemma. □

We proceed with the proof of part (1) of Proposition 5.3.1, i.e. the computation of $E[|\mathcal{V}^u|]$. First observe that

$$E[|\mathcal{V}^u|] = \sum_{x\in\mathcal{C}_1} P[x \text{ is vacant at time } u\rho(2-\xi)\xi n] = \sum_{x\in\mathcal{C}_1} P[H_x > u\rho(2-\xi)\xi n].$$

The task is therefore to approximate the probabilities $P[H_x > u\rho(2-\xi)\xi n]$.

Assume that the random walk X is the discrete skeleton of a simple continuous-time random walk X^c, i.e. the times between jumps of X^c are i.i.d. Exponential(1). Denote by H_x^c the entrance time of x for this continuous-time walk and by S_k the time of the k-th jump. It is clear that $E[S_k] = k$ and $E[H_x^c] = E[H_x]$. From [AB92] or [AF02, Proposition 3.23] we know that the distribution of the entrance time of such a continuous-time walk can be approximated by an exponential distribution, namely for all $t > 0$

$$\left|P[H_x^c > t] - e^{-\frac{t}{E[H_x]}}\right| \leq \frac{1}{\lambda_{\mathcal{C}_1} E[H_x]}. \tag{5.3.18}$$

If $k = k(n) \to \infty$ as $n\to\infty$, by the law of large numbers $P[|S_k - k| > \epsilon k] = o(1)$ as $n\to\infty$ for all $\epsilon > 0$. This implies

$$\begin{aligned}P[H_x > k] = P[H_x^c > S_k] &= P[H_x^c > S_k,\ S_k \geq (1-\epsilon)k] + P[H_x^c > S_k,\ S_k < (1-\epsilon)k]\\ &\leq P[H_x^c > (1-\epsilon)k] + o(1) \text{ as } n\to\infty \text{ for all } \epsilon > 0,\end{aligned}$$

and similarly

$$P[H_x > k] \geq P[H_x^c > (1+\epsilon)k] + o(1) \text{ as } n\to\infty \text{ for all } \epsilon > 0.$$

We obtain $P[H_x > k] = P[H_x^c > k] + o(1)$ as $n\to\infty$, and together with the bounds (5.2.2) for $\lambda_{\mathcal{C}_1}$ and (5.2.8) for $E[H_x]$ it follows from (5.3.18) that $\mathbb{P}_{n,p}$-a.a.s.

$$\left|P[H_x > u\rho(2-\xi)\xi n] - e^{-\frac{u\rho(2-\xi)\xi n}{E[H_x]}}\right| = o(1). \tag{5.3.19}$$

Approximating the probabilities $P[H_x > u\rho(2-\xi)\xi n]$ therefore reduces to the investigation of $E[H_x]$. We will use Proposition 3.2 from [ČTW11], which states that $E[H_x]$ can be approximated in terms of the Dirichlet form of the equilibrium potential $g^\star$ (cf. (5.2.10) and (5.2.11)).

Proposition 5.3.4 ([ČTW11, Proposition 3.2])**.**

$$\mathcal{D}(g^\star, g^\star)\left(1 - 2\sup_{y\in B(x,r)^c}|f^\star(y)|\right) \le \frac{1}{E[H_x]} \le \mathcal{D}(g^\star, g^\star)\frac{1}{\pi(B(x,r)^c)^2}, \tag{5.3.20}$$

where $f^\star(y) = 1 - \frac{E_y[H_x]}{E[H_x]}$.

To use this result, we need to control the function $f^\star$. To this end, we give in the next lemma a bound on the probability that the random walk on $\mathcal{C}_1$ started outside $B(x,r)$ hits x before some time T. Recall the coupling Q_x from Proposition 5.3.2, the definition (5.3.13) of the event $\{x \text{ good}\}$, and the definition (5.3.1) of the radius r.

Lemma 5.3.5. *There is a constant c, such that, for $T \in \mathbb{N}$ possibly depending on n,*

$$Q_x\left[\sup_{y\in B(x,r)^c} P_y[H_x \le T] \le Te^{-cr} \;\middle|\; x \text{ good}\right] \to 1 \text{ as } n \to \infty.$$

Proof. For x good let $\mathcal{T}$ be the infinite Poisson(ρ)-Galton-Watson tree defined by the coupling Q_x to which the neighborhood of x is isomorphic. Let $P_w^{\mathcal{T}}$ be the law of the simple random walk on the tree $\mathcal{T}$ started at $w \in \mathcal{T}$. To bound the escape probability of random walk on a Galton-Watson tree we use [JLT14, Proposition 11.5], which states that

$$\sup_{w\in\partial B(\varnothing,r)} P_w^{\mathcal{T}}[H_\varnothing < \infty] \le e^{-cr} \qquad \mathbb{P}_{\mathcal{T}}^0\text{-a.a.s. as } r \to \infty.$$

Since $P_w^{\mathcal{T}}[H_\varnothing < \infty] \ge P_w^{\mathcal{T}}[H_\varnothing < H_{B(\varnothing,r)^c}]$, this implies

$$\sup_{w\in\partial B(\varnothing,r)} P_w^{\mathcal{T}}[H_\varnothing < H_{B(\varnothing,r)^c}] \le e^{-cr} \qquad \mathbb{P}_{\mathcal{T}}^0\text{-a.a.s. as } r \to \infty.$$

As argued before, since $Q_x[x \text{ good}]$ is bounded away from zero, this also holds $Q_x[\,\cdot \mid x \text{ good}]$-a.a.s. For x good, $P_w^{\mathcal{T}}[H_\varnothing < H_{B(\varnothing,r)^c}] = P_z[H_x < H_{B(x,r)^c}]$, where $z \in \partial B(x,r)$ is the image of w under the isomorphism between $B(x,r+1) \subset G$ and $B(\varnothing,r+1) \subset \mathcal{T}$. It follows that

$$Q_x\left[\sup_{z\in\partial B(x,r)} P_z\left[H_x < H_{B(x,r)^c}\right] \le e^{-cr} \;\middle|\; x \text{ good}\right] \to 1 \text{ as } n \to \infty.$$

On the way from $y \in B(x,r)^c$ to x, the random walk on $\mathcal{C}_1$ must visit some $z \in \partial B(x,r)$. From there it either reaches x or leaves $B(x,r)$ again. The probability of the first event is $Q_x[\,\cdot \mid x \text{ good}]$-a.a.s. bounded by e^{-cr}, and if the second event occurs, we can repeat the previous reasoning. But in time T, this procedure can be repeated at most T times, leading to the required bound on $P_y[H_x \le T]$. □

With Lemma 5.3.5 we can give a bound on $\sup_{y\in B(x,r)^c}|f^\star(y)|$ on the left hand side of (5.3.20).

Lemma 5.3.6. *There are constants c, c', such that*

$$Q_x\left[\sup_{y\in B(x,r)^c}\left|1 - \frac{E_y[H_x]}{E[H_x]}\right| \le cn^{-c'} \;\middle|\; x \text{ good}\right] \to 1 \text{ as } n \to \infty. \tag{5.3.21}$$

Proof. Note first that by the general $O(k^3)$-bound on the expected cover time C_G of a graph G on k vertices (see e.g. [AKL$^+$79]), we have

$$\sup_{z\in\mathcal{C}_1} E_z[H_x] \le C_{\mathcal{C}_1} \le n^3. \tag{5.3.22}$$

Before considering the expectation of H_x with the random walk started from $y \in B(x,r)^c$, we consider the expectation of H_x starting from X_T for some time T where the walk is well mixed. Set $T = \log^4 n$. With (5.2.13), (5.2.2), (5.2.7) and (5.3.22) we get $\mathbb{P}_{n,p}$-a.a.s. for all $z \in \mathcal{C}_1$

$$\begin{aligned}\big|E_z[E_{X_T}[H_x]] - E[H_x]\big| &\le \sum_{z'\in\mathcal{C}_1} \big|P_z[X_T = z'] - \pi(z')\big| E_{z'}[H_x] \\ &\le \sum_{z'\in\mathcal{C}_1} \frac{1}{\min_{v\in\mathcal{C}_1}\pi(v)} e^{-\lambda_{\mathcal{C}_1} T} E_{z'}[H_x] \\ &\le cn^5 \log n e^{-c'\log^2 n} \le cn^{-c'}.\end{aligned} \tag{5.3.23}$$

By the Markov property at time T and using (5.3.23), $\mathbb{P}_{n,p}$-a.a.s.

$$E_z[H_x] \le T + E_z[E_{X_T}[H_x]] \le T + E[H_x] + cn^{-c'}. \tag{5.3.24}$$

With (5.2.8) it follows that $\mathbb{P}_{n,p}$-a.a.s. for all $z \in \mathcal{C}_1$

$$\frac{E_z[H_x]}{E[H_x]} - 1 \le (T + cn^{-c'})\frac{1}{E[H_x]} \le cn^{-c'}. \tag{5.3.25}$$

Since everything holding $\mathbb{P}_{n,p}$-a.a.s. also holds $Q_x[\ \cdot\ |\ x \text{ good}]$-a.a.s., (5.3.25) is enough for one side of (5.3.21).

For the other side take now $y \in B(x,r)^c$ and apply the Markov property at time T, use (5.3.23) on the first line and (5.3.24) for the supremum on the second line to get $\mathbb{P}_{n,p}$-a.a.s.

$$\begin{aligned}E_y[H_x] &\ge E_y[\mathbf{1}_{\{H_x>T\}} E_{X_T}[H_x]] = E_y[E_{X_T}[H_x]] - E_y[\mathbf{1}_{\{H_x\le T\}} E_{X_T}[H_x]] \\ &\ge E[H_x] - cn^{-c'} - P_y[H_x \le T] \sup_{z\in\mathcal{C}_1} E_z[H_x] \\ &\ge E[H_x] - 2cn^{-c'} - P_y[H_x \le T](T + E[H_x]).\end{aligned}$$

This holds $\mathbb{P}_{n,p}$-a.s.s., so as argued before it also holds $Q_x[\ \cdot\ |\ x \text{ good}]$-a.a.s. With the bound (5.2.8) and using Lemma 5.3.5, where we note that $e^{-cr} = n^{-c'}$ by (5.3.1), it follows that $Q_x[\ \cdot\ |\ x \text{ good}]$-a.a.s.

$$\frac{E_y[H_x]}{E[H_x]} - 1 \ge -cn^{-1-c'}\log n - \log^4 n e^{-c''r}\left(\frac{c'''\log^5 n}{n} + 1\right) \ge -cn^{-c'}.$$

Together with (5.3.25) this proves the lemma. □

Applying Lemma 5.3.6 in (5.3.20) and using Lemma 5.3.3, we obtain the following approximation of the probabilities $P[H_x > u\rho(2-\xi)\xi n]$.

Lemma 5.3.7. *For any fixed $u > 0$ and every $\epsilon > 0$,*

$$Q_x\left[\left|P\left[H_x > u\rho(2-\xi)\xi n\right] - e^{-uP_\varnothing^{\mathcal{T}}\left[\tilde{H}_\varnothing > H_{B(\varnothing,r)^c}\right]\deg(\varnothing)}\right| \le \epsilon \;\middle|\; x \text{ good}\right] \to 1 \text{ as } n \to \infty.$$

Proof. First recall (5.2.11) and use (5.2.6) to get $\mathbb{P}_{n,p}$-a.a.s.

$$\mathcal{D}(g^\star, g^\star) = P_x\left[\tilde{H}_x > H_{B(x,r)^c}\right]\pi(x) \le \frac{c\log n}{n}. \tag{5.3.26}$$

For the left hand approximation in (5.3.20), Lemma 5.3.6 and (5.3.26) imply that $Q_x[\ \cdot\ |\ x \text{ good}]$-a.a.s.

$$\frac{1}{E[H_x]} \ge \mathcal{D}(g^\star, g^\star) - cn^{-1-c'}. \tag{5.3.27}$$

For the right hand approximation in (5.3.20), first recall that by (5.3.6) $\mathbb{P}_{n,p}$-a.a.s., $|B(x,r)| \le n^{1-\epsilon}$ for some $\epsilon > 0$. Together with (5.2.6) we get $\mathbb{P}_{n,p}$-a.a.s.

$$\pi(B(x,r)^c) \ge 1 - |B(x,r)| \max_{v\in\mathcal{C}_1}\pi(v) \ge 1 - cn^{-\epsilon}\log n.$$

Using this and (5.3.26) in (5.3.20) yields $\mathbb{P}_{n,p}$-a.a.s.

$$\begin{aligned}\frac{1}{E[H_x]} &\le \mathcal{D}(g^\star, g^\star)\frac{1}{(1-cn^{-\epsilon}\log n)^2} \le \mathcal{D}(g^\star, g^\star)\left(1 + cn^{-\epsilon}\log n\right)\\ &\le \mathcal{D}(g^\star, g^\star) + cn^{-1-\epsilon}\log^2 n \le \mathcal{D}(g^\star, g^\star) + cn^{-1-c'}.\end{aligned} \tag{5.3.28}$$

Combining (5.3.27) and (5.3.28) we obtain that $Q_x[\ \cdot\ |\ x \text{ good}]$-a.a.s.

$$e^{-\frac{u\rho(2-\xi)\xi n}{E[H_x]}} = e^{-u\rho(2-\xi)\xi n(\mathcal{D}(g^\star,g^\star)+o(n^{-1}))} = e^{-u\rho(2-\xi)\xi n\mathcal{D}(g^\star,g^\star)} + o(1).$$

Together with (5.3.19) it follows that

$$Q_x\left[\left|P\left[H_x > u\rho(2-\xi)\xi n\right] - e^{-u\rho(2-\xi)\xi n\mathcal{D}(g^\star,g^\star)}\right| \le \epsilon \;\middle|\; x \text{ good}\right] \to 1 \text{ as } n\to\infty. \tag{5.3.29}$$

Lemma 5.3.3 implies that $\mathbb{P}_{n,p}$-a.a.s. for $x \in \mathcal{C}_1$, $\pi(x) = \frac{\deg(x)}{\rho(2-\xi)\xi n}(1+o(1))$. Recalling (5.2.11), this implies that $\mathbb{P}_{n,p}$-a.a.s.

$$u\rho(2-\xi)\xi n\mathcal{D}(g^\star, g^\star) = uP_x\left[\tilde{H}_x > H_{B(x,r)^c}\right]\deg(x) + o(1).$$

Using this in (5.3.29), and noting that if x is good,

$$e^{-uP_x\left[\tilde{H}_x > H_{B(x,r)^c}\right]\deg(x)} = e^{-uP_\varnothing^{\mathcal{T}}\left[\tilde{H}_\varnothing > H_{B(\varnothing,r)^c}\right]\deg(\varnothing)},$$

finishes the proof of the lemma. □

Proof of part (1) *of Proposition 5.3.1.* We use the same technique as in the proof of Lemma 5.3.3: We compute the expectation of $E[|\mathcal{V}^u|]$ under $\mathbb{P}_{n,p}$ and then show that $E[|\mathcal{V}^u|]$ is concentrated. Define the random variables

$$\begin{aligned}W_x &= \mathbf{1}_{\{x\in\mathcal{C}_1\}}P[H_x > u\rho(2-\xi)\xi n],\\ \tilde{W}_x &= \mathbf{1}_{\{|\mathcal{T}|=\infty\}}e^{-uP_\varnothing^{\mathcal{T}}\left[\tilde{H}_\varnothing > H_{B(\varnothing,r)^c}\right]\deg(\varnothing)},\end{aligned}$$

where the tree $\mathcal{T}$ is defined by the coupling Q_x from Proposition 5.3.2, and $\varnothing$ is the root of $\mathcal{T}$.

Let us first compute the expectation of $\tilde{W}_x$ as $n \to \infty$. Since $r \to \infty$ as $n \to \infty$, and the tree $\mathcal{T}$ has law $\mathbb{P}^0_{\mathcal{T}}$,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}_{Q_x}\big[\tilde{W}_x\big] &= \lim_{n\to\infty} \mathbb{E}^0_{\mathcal{T}}\left[e^{-uP^{\mathcal{T}}_{\varnothing}\left[\tilde{H}_{\varnothing} > H_{B(\varnothing,r)^c}\right]\deg(\varnothing)} \;\middle|\; |\mathcal{T}| = \infty\right] \mathbb{P}^0_{\mathcal{T}}\left[|\mathcal{T}| = \infty\right] \\ &= \mathbb{E}_{\mathcal{T}}\left[e^{-u\,\mathrm{cap}_{\mathcal{T}}(\varnothing)}\right]\xi.\end{aligned} \tag{5.3.30}$$

For $\epsilon > 0$, define on the space of the coupling Q_x the event

$$\mathcal{A}_{x,\epsilon} = \{|W_x - \tilde{W}_x| \le \epsilon\}. \tag{5.3.31}$$

By definitions (5.3.9) and (5.3.13) of the events $\mathcal{B}_x$ and $\{x \text{ good}\}$, on $\mathcal{B}_x$ either $W_x = \tilde{W}_x = 0$ or x is good, i.e. $\mathcal{A}^c_{x,\epsilon} \cap \mathcal{B}_x = \mathcal{A}^c_{x,\epsilon} \cap \{x \text{ good}\}$. With Lemma 5.3.7 and (5.3.11) it follows that

$$\begin{aligned}Q_x[\mathcal{A}^c_{x,\epsilon}] &\le Q_x\left[\mathcal{A}^c_{x,\epsilon},\ \mathcal{B}_x\right] + Q_x\left[\mathcal{B}^c_x\right] \\ &\le Q_x\left[\mathcal{A}^c_{x,\epsilon} \;\middle|\; x \text{ good}\right] Q_x\left[x \text{ good}\right] + Q_x\left[\mathcal{B}^c_x\right] = o(1) \text{ as } n \to \infty.\end{aligned} \tag{5.3.32}$$

Since W_x and $\tilde{W}_x$ are bounded by 1, this implies

$$\left|\mathbb{E}_{Q_x}[W_x] - \mathbb{E}_{Q_x}[\tilde{W}_x]\right| \le \epsilon + Q_x[\mathcal{A}^c_{x,\epsilon}] \text{ for any } \epsilon > 0,$$

and thus

$$\mathbb{E}_{Q_x}[W_x] = \mathbb{E}_{Q_x}[\tilde{W}_x] + o(1) \text{ as } n \to \infty. \tag{5.3.33}$$

With (5.3.30) we conclude that

$$\mathbb{E}_{n,p}\big[E[|\mathcal{V}^u|]\big] = \mathbb{E}_{n,p}\left[\sum_{x\in G} W_x\right] = n\mathbb{E}_{Q_x}[W_x] = \xi n \mathbb{E}_{\mathcal{T}}\left[e^{-u\,\mathrm{cap}_{\mathcal{T}}(\varnothing)}\right] + o(n) \text{ as } n \to \infty.$$

For the concentration of $E[|\mathcal{V}^u|]$ we use again the second moment method. Consider the coupling $Q_{x,y}$ from Proposition 5.3.2 for two fixed vertices $x \neq y$. The random variable $\tilde{W}_z$ as well as the event $\mathcal{A}_{z,\epsilon}$ for $z \in \{x, y\}$ are canonically also defined on the space of $Q_{x,y}$ when replacing $\mathcal{T}$ by $\mathcal{T}_z$ in the definition of $\tilde{W}_z$. Let $\mathcal{A}_{x,y,\epsilon} = \mathcal{A}_{x,\epsilon} \cap \mathcal{A}_{y,\epsilon}$, and recall the definition (5.3.10) of the set $\mathcal{B}_{x,y}$, on which either $W_z = \tilde{W}_z = 0$ or z is good, for both $z \in \{x, y\}$. Note that the statement of Lemma 5.3.7 also holds on the space of $Q_{x,y}$ when replacing $\mathcal{T}$ by $\mathcal{T}_z$ for both $z \in \{x, y\}$ respectively. As in (5.3.32), with Lemma 5.3.7 and (5.3.12) we obtain

$$\begin{aligned}Q_{x,y}[\mathcal{A}^c_{x,y,\epsilon}] &\le Q_{x,y}\left[\mathcal{A}^c_{x,\epsilon},\ \mathcal{B}_{x,y}\right] + Q_{x,y}\left[\mathcal{A}^c_{y,\epsilon},\ \mathcal{B}_{x,y}\right] + Q_{x,y}\left[\mathcal{B}^c_{x,y}\right] \\ &\le Q_{x,y}\left[\mathcal{A}^c_{x,\epsilon} \;\middle|\; x \text{ good}\right] + Q_{x,y}\left[\mathcal{A}^c_{y,\epsilon} \;\middle|\; y \text{ good}\right] + Q_{x,y}\left[\mathcal{B}^c_{x,y}\right] = o(1) \text{ as } n \to \infty.\end{aligned}$$

Since the W_z and $\tilde{W}_z$ are bounded by 1, it follows that

$$\left|\mathbb{E}_{Q_{x,y}}[W_x W_y] - \mathbb{E}_{Q_{x,y}}[\tilde{W}_x \tilde{W}_y]\right| \le \epsilon + Q_{x,y}[\mathcal{A}^c_{x,y,\epsilon}] \text{ for any } 1 > \epsilon > 0,$$

and thus

$$\mathbb{E}_{Q_{x,y}}[W_x W_y] = \mathbb{E}_{Q_{x,y}}[\tilde{W}_x \tilde{W}_y] + o(1) \text{ as } n \to \infty. \tag{5.3.34}$$

The trees $\mathcal{T}_x$ and $\mathcal{T}_y$ are independent, so the random variables $\tilde{W}_x$ and $\tilde{W}_y$ are independent. Therefore, (5.3.33) and (5.3.34) imply that for arbitrary vertices $x \neq y$

$$\mathbb{E}_{n,p}[W_x W_y] = \mathbb{E}_{n,p}[W_x]\mathbb{E}_{n,p}[W_y] + o(1) \text{ as } n \to \infty.$$

Recall that $E[|\mathcal{V}^u|] = \sum_{x \in G} W_x$. By the boundedness of the W_x, it follows directly from the above that

$$\mathbb{E}_{n,p}\left[E[|\mathcal{V}^u|]^2\right] = \mathbb{E}_{n,p}\left[E[|\mathcal{V}^u|]\right]^2 + o(n^2) \text{ as } n \to \infty.$$

Thus $\operatorname{Var} E[|\mathcal{V}^u|] = o(n^2)$ and the Chebyshev inequality implies for any $\epsilon > 0$

$$\mathbb{P}_{n,p}\left[|E[|\mathcal{V}^u|] - \mathbb{E}_{n,p}\left[E[|\mathcal{V}^u|]\right]| > \epsilon n\right] = o(1) \text{ as } n \to \infty.$$

This finishes the proof of the first part of Proposition 5.3.1. □

5.3.2. Concentration of the size of the vacant set

To prove part (2) of Proposition 5.3.1, we use similar techniques as in [ČTW11] and [ČT13]. We define a sequence of i.i.d. stationary started random walk trajectories of length n^δ and glue them together at the endpoints to obtain a trajectory which is, by the fast mixing of the random walk, in distribution close to the random walk on $\mathcal{C}_1$ but has a different dependency structure, which allows to apply the following concentration result by [McD98].

Theorem 5.3.8 ([McD98, Theorem 3.7])**.** *Let $W = (W_1, ..., W_M)$ be a family of random variables W_k taking values in a set $\mathcal{A}_k$, and let f be a bounded real-valued function on $\prod \mathcal{A}_k$. Let μ denote the mean of $f(W)$. Define*

$$r_k(y_1, ..., y_{k-1}) = \sup_{y, y' \in \mathcal{A}_k} \Big| E\left[f(W) \;\middle|\; W_k = y,\ W_i = y_i\ \forall i < k\right] - E\left[f(W) \;\middle|\; W_k = y',\ W_i = y_i\ \forall i < k\right] \Big|,$$

and let

$$R^2 = \sup_{y_1, ..., y_{M-1}} \sum_{k=1}^{M} r_k^2(y_1, ..., y_{k-1}).$$

Then for any $t \geq 0$,

$$P\left[|f(W) - \mu| \geq t\right] \leq 2e^{-\frac{t^2}{R^2}}.$$

Let us define precisely the above mentioned approximation of the random walk. Denote by P_x^L the restriction of P_x to $\mathcal{C}_1^{L+1}$, i.e. the law of the trajectory $(X_0, ..., X_L)$ and by $P_{x,z}^L$ the law of the random walk bridge, that is P_x^L conditioned on $X_L = z$. Fix $\delta > 0$ and let $L = n^\delta$. For a given typical random graph G define on an auxiliary probability space $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{P})$ the i.i.d. random variables $(Z^i)_{i \geq 0}$ as vertices of $\mathcal{C}_1$ chosen according to the stationary measure π. Given the collection (Z^i), let $(Y^i)_{i \geq 1}$ be conditionally independent elements of $\mathcal{C}_1^{L+1}$ such that each $(Y_k^i)_{k=0,...,L}$ is distributed according to the random walk bridge P_{Z^{i-1}, Z^i}^L. We define the concatenation of the Y^i as

$$\mathcal{X}_t = Y_{t-(i-1)L}^i, \qquad \text{when } (i-1)L \leq t < iL.$$

Denote by $\mathcal{P}^u$ the law of $\mathcal{X}$ on $\mathcal{C}_1^{u\rho(2-\xi)\xi n+1}$ and write P^u for $P^{u\rho(2-\xi)\xi n}$, that is P restricted to $\mathcal{C}_1^{u\rho(2-\xi)\xi n+1}$. The next lemma shows that $\mathcal{P}^u$ approximates P^u well if L is large enough.

Lemma 5.3.9. *$\mathbb{P}_{n,p}$-a.a.s. the measures $\mathcal{P}^u$ and P^u are equivalent, and for n large enough and constants c, c' depending on δ,*

$$\left|\frac{dP^u}{d\mathcal{P}^u} - 1\right| \leq ce^{-c'n^{\frac{\delta}{2}}}.$$

Proof. Let u' be the smallest number greater or equal to u such that $u'\rho(2-\xi)\xi n$ is an integer multiple of L and set $m = \frac{u'\rho(2-\xi)\xi n}{L}$. Since $\mathcal{P}^u$ and P^u are the restrictions of $\mathcal{P}^{u'}$ and $P^{u'}$ to $\mathcal{C}_1^{u\rho(2-\xi)\xi n+1}$, it is sufficient to prove the lemma for $\mathcal{P}^{u'}$ and $P^{u'}$. Let A be any measurable subset of $\mathcal{C}_1^{u\rho(2-\xi)\xi n+1}$. Then by the Markov property

$$\begin{aligned} P^{u'}[A] &= \sum_{x_0,\dots,x_m \in \mathcal{C}_1} P^{u'}\left[A \;\middle|\; X_{iL} = x_i,\ 0 \leq i \leq m\right] P^{u'}\left[X_{iL} = x_i,\ 0 \leq i \leq m\right] \\ &= \sum_{x_0,\dots,x_m \in \mathcal{C}_1} P^{u'}\left[A \;\middle|\; X_{iL} = x_i,\ 0 \leq i \leq m\right] \pi(x_0) \prod_{k=0}^{m} P_{x_k}^L[X_L = x_{k+1}]. \end{aligned} \tag{5.3.35}$$

Next, note that $\mathcal{P}^{u'}\left[X_{iL} = x_i,\ 0 \leq i \leq m\right] = 0$ if and only if $P^{u'}\big[X_{iL} = x_i,\ 0 \leq i \leq m\big] = 0$: One can always choose the m x_i's, but there might not be any way to connect them by random walk bridges, whence the probability is zero. In this case, there is also no random walk trajectory going through this points. On the other hand, when there is no such trajectory, there are also no bridges.

From this and the construction of the measure $\mathcal{P}$ it follows that, whenever this is well-defined,

$$\begin{aligned} \mathcal{P}^{u'}\left[A \;\middle|\; X_{iL} = x_i,\ 0 \leq i \leq m\right] &= P^{u'}\left[A \;\middle|\; X_{iL} = x_i,\ 0 \leq i \leq m\right], \\ \mathcal{P}^{u'}\left[X_{iL} = x_i,\ 0 \leq i \leq m\right] &= \prod_{k=0}^{m} \pi(x_k). \end{aligned} \tag{5.3.36}$$

Comparing (5.3.35) and (5.3.36), it remains to control the ratio $\frac{P_x^L[X_L=y]}{\pi(y)}$. We use (5.2.13), (5.2.7) and (5.2.2) to get $\mathbb{P}_{n,p}$-a.a.s.

$$\left|\frac{P_x^L[X_L = y]}{\pi(y)} - 1\right| \leq \frac{1}{(\min_{z\in\mathcal{C}_1}\pi(z))^2} e^{-\lambda_{\mathcal{C}_1} L} \leq cn^2 \log^2 n e^{-\frac{c'}{\log^2 n}L}.$$

With $\frac{n^\delta}{\log^2 n} \geq cn^{\frac{\delta}{2}}$ for n large enough it follows that $\mathbb{P}_{n,p}$-a.a.s.

$$\left(1 - cn^2\log^2 n e^{-c'n^{\frac{\delta}{2}}}\right)^m \leq \frac{P^{u'}[A]}{\mathcal{P}^{u'}[A]} \leq \left(1 + cn^2\log^2 n e^{-c'n^{\frac{\delta}{2}}}\right)^m,$$

and hence $\mathbb{P}_{n,p}$-a.a.s. $\mathcal{P}^{u'}$ and $P^{u'}$ are equivalent, and the lemma follows by changing constants to accommodate the terms polynomial in n and $\log n$. □

Proof of part (2) *of Proposition 5.3.1.* We show that for any $\delta > 0$,

$$P\left[\left||\mathcal{V}^u| - E[|\mathcal{V}^u|]\right| \geq n^{\frac{1}{2}+\delta}\right] \leq ce^{-c'n^{\frac{\delta}{2}}} \qquad \mathbb{P}_{n,p}\text{-a.a.s.}, \tag{5.3.37}$$

which implies the statement of the proposition.

Set $m = \left\lfloor \frac{u\rho(2-\xi)\xi n}{L} \right\rfloor$ and $u' = \frac{mL}{\rho(2-\xi)\xi n}$. Then $u\rho(2-\xi)\xi n - u'\rho(2-\xi)\xi n \leq L$, and so $\left||\mathcal{V}^u| - |\mathcal{V}^{u'}|\right| \leq L$. It follows that for n large enough

$$\begin{aligned} P\left[\left||\mathcal{V}^u| - E[|\mathcal{V}^u|]\right| \geq n^{\frac{1}{2}+\delta}\right] &\leq P\left[\left||\mathcal{V}^{u'}| - E[|\mathcal{V}^{u'}|]\right| \geq n^{\frac{1}{2}+\delta} - 2L\right] \\ &\leq P\left[\left||\mathcal{V}^{u'}| - E[|\mathcal{V}^{u'}|]\right| \geq \frac{1}{2}n^{\frac{1}{2}+\delta}\right]. \end{aligned} \tag{5.3.38}$$

Let $\mathcal{U}^{u'} = \mathcal{C}_1 \setminus \mathcal{X}_{[0,mL]}$ be the vacant set left by the concatenation $\mathcal{X}$, and denote by $\mathcal{E}$ the expectation corresponding to $\mathcal{P}$. Lemma 5.3.9 implies that $\mathbb{P}_{n,p}$-a.a.s.

$$\begin{aligned} \left|P\big[\mathcal{V}^{u'} \in \ \cdot\ \big] - \mathcal{P}\big[\mathcal{U}^{u'} \in \ \cdot\ \big]\right| &\leq ce^{-c'n^{\frac{\delta}{2}}}, \\ \left|E[|\mathcal{V}^{u'}|] - \mathcal{E}[|\mathcal{U}^{u'}|]\right| &\leq cne^{-c'n^{\frac{\delta}{2}}} \leq \frac{1}{4}n^{\frac{1}{2}+\delta}. \end{aligned}$$

From this we obtain that $\mathbb{P}_{n,p}$-a.a.s.

$$P\left[\left||\mathcal{V}^{u'}| - E[|\mathcal{V}^{u'}|]\right| \geq \frac{1}{2}n^{\frac{1}{2}+\delta}\right] \leq \mathcal{P}\left[\left||\mathcal{U}^{u'}| - \mathcal{E}[|\mathcal{U}^{u'}|]\right| \geq \frac{1}{4}n^{\frac{1}{2}+\delta}\right] + ce^{-c'n^{\frac{\delta}{2}}}. \tag{5.3.39}$$

We now apply Theorem 5.3.8 with $M = m$, $\mathcal{A}_k = \mathcal{C}_1^{L+1}$, $W_k = Y^k$ and $f(W) = |\mathcal{U}^{u'}|$. We claim that

$$\begin{aligned} &r_k(y_1, ..., y_{k-1}) \\ &= \sup_{y,y' \in \mathcal{A}_k} \left|E\left[|\mathcal{U}^{u'}| \mid Y^k = y,\ Y^i = y_i\ \forall i < k\right] - E\left[|\mathcal{U}^{u'}| \mid Y^k = y',\ Y^i = y_i\ \forall i < k\right]\right| \\ &\leq 2L. \end{aligned}$$

Indeed, when conditioning additionally on $Y^{k+2}, ..., Y^m$, the only two different segments Y^k and Y^{k+1} can change the size of the vacant set by at most the length of two segments, and the claim follows by integrating over all possible $Y^{k+2}, ..., Y^m$.

Then $R^2 \leq m(2L)^2 \leq \frac{u\rho(2-\xi)\xi n}{L} 4L^2 = cn^{1+\delta}$, and Theorem 5.3.8 implies

$$\mathcal{P}\left[\left||\mathcal{U}^{u'}| - \mathcal{E}[|\mathcal{U}^{u'}|]\right| \geq \frac{1}{4}n^{\frac{1}{2}+\delta}\right] \leq 2e^{-2\frac{\frac{1}{16}n^{1+2\delta}}{cn^{1+\delta}}} = ce^{-c'n^{\delta}}.$$

This together with (5.3.38) and (5.3.39) proves (5.3.37) and hence part (2) of Proposition 5.3.1. □

5.4. Coupling of processes

In this section we introduce a process $\bar{X}$ which satisfies the spatial Markov property described in the introduction. We derive a phase transition in the vacant set of this process, and we compare it with the simple random walk X on the giant component.

Consider the following algorithm defined on an auxiliary probability space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$ which builds an element of $\Omega_n = \mathcal{G}(n) \times \{1, 2, ..., n\}^{\mathbb{N}_0}$, that is a graph on n vertices and a random walk-like process on this graph. All the random choices made in the algorithm are independent variables defined on $\tilde{\Omega}$.

Algorithm 5.4.1. *At the beginning all n vertices are unvisited, and all $\binom{n}{2}$ possible edges are unexplored. When the algorithm (or the so defined process) passes an unvisited vertex, this vertex is marked visited. Edges adjacent to the vertex will be explored and become either open or closed. After being explored, the state of an edge does not change.*

(1) Start at time 0 with a uniformly chosen vertex v_0 among all n vertices, mark it visited.

(2) Being at time $k \geq 0$ with current vertex v_k, check first if there are any unvisited vertices left:

- *If there are, let any unexplored edge adjacent to v_k be explored and marked open with probability $p = \frac{\rho}{n}$ and closed otherwise. All vertices w such that the edge $\{v_k, w\}$ is open are called neighbors of v_k.*
- *If there are no unvisited vertices left, let $\{v_l\}_{l>k}$ be uniformly at random chosen vertices and terminate the algorithm (this choice of continuation of the process v_k is totally arbitrary and does not influence the reasoning below).*

(3) If v_k has at least one neighbor, and if there are any unvisited vertices adjacent to explored edges, choose vertex v_{k+1} uniformly among all neighbors of v_k and mark v_{k+1} visited, go to step (2) and proceed with current vertex v_{k+1}.

(4) If v_k has no neighbors or if there are no unvisited vertices adjacent to explored edges, the current component is entirely covered. Then choose vertex v_{k+1} uniformly among all n vertices, mark it visited, go to step (2) and proceed with current vertex v_{k+1}.

By construction, the law of the graph explored by this algorithm (edges present if they are marked open) is $\mathbb{P}_{n,p}$. Let $\bar{X}$ be the process defined by $\bar{X}_k = v_k$.

It will be helpful to have two different points of view on Algorithm 5.4.1. The first is to look at the picture at the end of the algorithm: There is a graph G and a trajectory of $\bar{X}$ covering all the vertices of the graph. Using this point of view, denote by $\bar{P}^G$ the law on $(\{1, 2, ..., n\}^{\mathbb{N}_0}, \mathbb{F}_n)$ of the process $\bar{X}$ under $\tilde{P}$ conditioned on the event that the graph explored by the algorithm is $G \in \mathcal{G}(n)$ (i.e. conditioned on the random choices in Algorithm 5.4.1 that determine the states of edges, but not on the random choices that determine the trajectory of $\bar{X}$). Under $\bar{P}^G$, the process $\bar{X}$ is, between two occurrences of step (4) of the algorithm, a simple random walk on the currently explored component, started with uniform distribution on this component. Define on $\Omega_n = \mathcal{G}(n) \times \{1, 2, ..., n\}^{\mathbb{N}_0}$ the annealed measure (cf. (5.1.1)) by

$$\bar{\mathbf{P}}_n(A \times B) = \sum_{G \in A} \mathbb{P}_{n,p}(G) \bar{P}^G(B) \qquad \text{for } A \in \mathbb{G}_n,\ B \in \mathbb{F}_n.$$

The second point of view is to look at Algorithm 5.4.1 as building the graph G on-the-go. Having this in mind, the next lemma, which is crucial for the proof of Theorem 5.1.1, is straightforward (cf. [CF11, Lemma 6] for a similar statement). Let $\bar{\mathcal{V}}(t) = G \setminus \bar{X}_{[0,t]}$ be the vacant set left by the process $\bar{X}$ at time t, defined on $(\Omega_n, \bar{\mathbf{P}}_n)$. Once again we use the same notation $\bar{\mathcal{V}}(t)$ for the set of vertices as well as the induced subgraph of G.

Lemma 5.4.2. *Under $\bar{\mathbf{P}}_n$ conditioned on $|\bar{\mathcal{V}}(t)| = N$ the graph $\bar{\mathcal{V}}(t)$ has marginal law $\mathbb{P}_{N,p}$.*

Proof. By construction of Algorithm 5.4.1, the vacant graph $\bar{\mathcal{V}}(t)$ consists of the $|\bar{\mathcal{V}}(t)|$ unvisited vertices at time t. Edges possibly connecting $\bar{\mathcal{V}}(t)$ and the already visited vertices as well as all edges possibly connecting two already visited vertices are explored. So the edges eligible to be edges of $\bar{\mathcal{V}}(t)$ are exactly all unexplored edges at time t. Because their state has not yet been decided by the algorithm, all these edges are open with probability $\frac{\rho}{n}$, independently of what happened before, independently of each other. Therefore, the vacant graph $\bar{\mathcal{V}}(t)$ is a standard Erdős-Rényi random graph on $N = |\bar{\mathcal{V}}(t)|$ vertices, every edge present with probability $p = \frac{\rho}{n}$, and hence it has law $\mathbb{P}_{N,p}$. □

From Lemma 5.4.2 and the classical results on random graphs it follows directly that the component structure of the vacant graph $\bar{\mathcal{V}}(t)$ exhibits a phase transition at the time t for which $|\bar{\mathcal{V}}(t)|\frac{\rho}{n} = 1$. To translate this phase transition to the simple random walk X on the giant component $\mathcal{C}_1(G)$, we need to couple X to the process $\bar{X}$. We do this by first giving a coupling of X and $\bar{X}$ under $P^{\mathcal{C}_1}$ and $\bar{P}^G$ respectively on a fixed typical graph G. In Section 5.5 we will extend this coupling to an 'annealed coupling' of X and $\bar{X}$ under $\mathbf{P}_n$ and $\bar{\mathbf{P}}_n$ respectively.

Proposition 5.4.3. *For n large enough, for every fixed typical graph $G \in \mathcal{G}(n)$ there exists a coupling Q^G of $\bar{X}$ under $\bar{P}^G$ and X under $P^{\mathcal{C}_1(G)}$ such that*

$$Q^G\left[\{X_k = \bar{X}_{k+2\log^5 n} \text{ for all } k = 0, 1, ..., u\rho(2-\xi)\xi n\}^c\right] \leq \frac{c}{n^{c'}}.$$

Proof. We first show that $\bar{X}$ typically is on the largest component $\mathcal{C}_1$ at time $\log^5 n$, that it mixes quickly, and then stays on $\mathcal{C}_1$ until time $u\rho(2-\xi)\xi n + 2\log^5 n$. This will allow us to identify X with $\bar{X}$ in this time interval on an event of high probability.

Let G be the typical graph (i.e. a graph satisfying (5.2.1), (5.2.2) and (5.2.3)) explored by Algorithm 5.4.1 and $\mathcal{C}_1$ its giant component, i.e. we look at the picture after completion of the algorithm. Define the probability distribution $\bar{\pi}$ on G as the distribution of $\bar{X}_{2\log^5 n}$, and view the stationary distribution π of the random walk on $\mathcal{C}_1$ as a distribution on the whole graph G by setting $\pi \equiv 0$ on $G \setminus \mathcal{C}_1$. Denote by $||\cdot||_{\mathrm{TV}}$ the total variation norm. Define $\tau = \min\{t \geq \log^5 n : \text{step (4) of Algorithm 5.4.1 is performed}\}$. τ is the first time after $\log^5 n$ where $\bar{X}$ does not behave like a random walk. We show that for n large enough the following properties hold for a typical graph G:

$$\bar{P}^G[\bar{X}_{\log^5 n} \notin \mathcal{C}_1] \leq \frac{c}{n^{1+c'}}, \tag{5.4.1}$$

$$\bar{P}^G[\tau \leq u\rho(2-\xi)\xi n + 2\log^5 n] \leq \frac{c}{n^{1+c'}}, \tag{5.4.2}$$

$$\|\bar{\pi} - \pi\|_{\mathrm{TV}} \leq \frac{c}{n^{c'}}. \tag{5.4.3}$$

Since G is typical, there is a giant component of size $||\mathcal{C}_1| - \xi n| \leq n^{3/4}$, and all other components are simple (i.e. they have at most as many edges as vertices) and of size smaller than $C\log n$. For (5.4.1), since for n large enough the random walk cannot cover $\mathcal{C}_1$ in $\log^4 n$ steps,

$$\bar{P}^G[\bar{X}_{\log^4 n} \notin \mathcal{C}_1] \leq \bar{P}^G[\bar{X} \text{ starts on a small component and stays on small components for time longer than } \log^4 n]. \tag{5.4.4}$$

Let N_s be the number of small components that $\bar{X}$ visits before reaching the giant component. By construction and since by (5.2.1) $|\mathcal{C}_1| \geq (\xi - \epsilon)n$ for some $\epsilon > 0$, N_s is stochastically dominated by a Geometric$(\xi - \epsilon)$ random variable, in particular it has a finite mean. Then by the Markov inequality

$$\bar{P}^G[N_s \geq \log n] \leq \frac{c}{\log n}. \tag{5.4.5}$$

Let $C_s^{(i)}$ be the cover time of the i-th small component covered by $\bar{X}$. The expected cover time of a graph on k vertices and m edges is bounded by $2m(k-1)$ (see e.g. [AKL$^+$79]), so the expected cover time $\bar{E}^G[C_s^{(i)}]$ of a simple component of size smaller than $C \log n$ is bounded by $C' \log^2 n$. The Markov inequality implies

$$\bar{P}^G \left[\sum_{i=1}^{\log n} C_s^{(i)} \geq \log^4 n \right] \leq \frac{\log n \bar{E}^G[C_s^{(i)}]}{\log^4 n} \leq \frac{c}{\log n}. \tag{5.4.6}$$

From (5.4.5) and (5.4.6) it follows that the probability on the right hand side of (5.4.4) is smaller than $\frac{c}{\log n}$. Given $\bar{X}$ has not found $\mathcal{C}_1$ after $\log^4 n$ steps, some small components are partly or entirely covered, but one can use the same line of arguments as above for the next $\log^4 n$ steps to get

$$\bar{P}^G \left[\bar{X}_{2\log^4 n} \notin \mathcal{C}_1 \mid \bar{X}_{\log^4 n} \notin \mathcal{C}_1 \right] \leq \bar{P}^G \left[\bar{X}_{\log^4 n} \notin \mathcal{C}_1 \right].$$

Using this, we have

$$\begin{aligned} \bar{P}^G \left[\bar{X}_{2\log^4 n} \notin \mathcal{C}_1 \right] &= \bar{P}^G \left[\bar{X}_{2\log^4 n} \notin \mathcal{C}_1 \mid \bar{X}_{\log^4 n} \notin \mathcal{C}_1 \right] \bar{P}^G \left[\bar{X}_{\log^4 n} \notin \mathcal{C}_1 \right] \\ &\leq \bar{P}^G \left[\bar{X}_{\log^4 n} \notin \mathcal{C}_1 \right]^2. \end{aligned}$$

Since $\bar{X}$ cannot cover $\mathcal{C}_1$ in $\log^5 n$ steps we can iterate the above $\log n$ times, then

$$\bar{P}^G \left[\bar{X}_{\log^5 n} \notin \mathcal{C}_1 \right] \leq \left(\frac{c}{\log n} \right)^{\log n} \leq \frac{c}{n^{1+c'}},$$

which proves (5.4.1).

To prove (5.4.2) first note that $P^{\mathcal{C}_1}[\ \cdot\ ] = \sum_{z \in \mathcal{C}_1} \pi(z) P_z^{\mathcal{C}_1}[\ \cdot\ ]$. With (5.2.6) it follows that

$$\sup_{z \in \mathcal{C}_1} P_z^{\mathcal{C}_1}[\ \cdot\ ] \leq \frac{1}{\min_{z \in \mathcal{C}_1} \pi(z)} P^{\mathcal{C}_1}[\ \cdot\ ] \leq cn \log n P^{\mathcal{C}_1}[\ \cdot\ ]. \tag{5.4.7}$$

Using (5.4.1) we have

$$\begin{aligned} &\bar{P}^G[\tau \leq u\rho(2-\xi)\xi n + 2\log^5 n] \\ &\quad \leq \sup_{z \in \mathcal{C}_1} P_z^{\mathcal{C}_1}[\text{cover time of } \mathcal{C}_1 \text{ is smaller than } u\rho(2-\xi)\xi n + 2\log^5 n] + \bar{P}^G[\bar{X}_{\log^5 n} \notin \mathcal{C}_1] \\ &\quad \leq \sup_{z \in \mathcal{C}_1} P^{\mathcal{C}_1}[\text{vacant set } \mathcal{V}(u\rho(2-\xi)\xi n + 2\log^5 n) \text{ is empty}] + \frac{c}{n^{1+c'}}. \end{aligned}$$

Since adding a trajectory of length $2\log^5 n$ can decrease the size of the vacant set by at most $2\log^5 n = o(n)$, it follows that asymptotically $|\mathcal{V}(u\rho(2-\xi)\xi n + 2\log^5 n)| =$

$|\mathcal{V}(u\rho(2-\xi)\xi n)| + o(n)$. Using (5.4.7), from (5.3.37) and part (1) of Proposition 5.3.1 it follows that for a typical graph and ϵ small enough

$$\sup_{z\in\mathcal{C}_1} P_z^{\mathcal{C}_1}[|\mathcal{V}(u\rho(2-\xi)\xi n)| < \epsilon n] \le cn\log n P^{\mathcal{C}_1}[|\mathcal{V}(u\rho(2-\xi)\xi n)| < \epsilon n] \le c' n\log n e^{-c'' n^{\frac{\delta}{2}}},$$

where $\delta > 0$ is the parameter defining the length of the random walk bridges in Section 5.3.2. For any choice of δ we can find constants such that the above expression is smaller than $\frac{c}{n^{1+c'}}$, and (5.4.2) follows.

For the proof of (5.4.3) let $P_\mu^{\mathcal{C}_1}$ denote the law of the random walk on $\mathcal{C}_1$ started at initial distribution μ. When $\bar{X}$ is on $\mathcal{C}_1$ at time $\log^5 n$, it has then some distribution μ and it cannot cover $\mathcal{C}_1$ in time $\log^5 n$. Using (5.2.13), we thus get for every $y \in \mathcal{C}_1$

$$\begin{aligned}\left|\bar{P}^G[\bar{X}_{2\log^5 n} = y] - \pi(y)\right| &\le \bar{P}^G\left[\bar{X}_{\log^5 n} \notin \mathcal{C}_1\right] + \sup_\mu \left|P_\mu^{\mathcal{C}_1}[X_{\log^5 n} = y] - \pi(y)\right| \\ &\le \bar{P}^G\left[\bar{X}_{\log^5 n} \notin \mathcal{C}_1\right] + \frac{1}{\min_{v\in\mathcal{C}_1}\pi(v)} e^{-\lambda_{\mathcal{C}_1}\log^5 n}.\end{aligned}$$

With (5.4.1), (5.2.2) and (5.2.7), it follows for every $y \in \mathcal{C}_1$

$$\left|\bar{P}^G[\bar{X}_{2\log^5 n} = y] - \pi(y)\right| \le \frac{c}{n^{1+c'}}. \tag{5.4.8}$$

We set $\pi \equiv 0$ on $G\setminus\mathcal{C}_1$, and by (5.4.1) and (5.4.2), $\bar{P}^G[\bar{X}_{2\log^5 n} = y] \le \frac{c}{n^{1+c'}}$ for $y \in G\setminus\mathcal{C}_1$. Thus (5.4.8) holds for all $y \in G$. (5.4.3) follows from (5.4.8) since by e.g. [LPW09, Proposition 4.2] we have $\|\bar{\pi} - \pi\|_{\mathrm{TV}} \le n \max_{y\in G}\left|\bar{P}^G[\bar{X}_{2\log^5 n} = y] - \pi(y)\right|$.

We can now define the coupling of X under $P^{\mathcal{C}_1}$ and $\bar{X}$ under $\bar{P}^G$. Consider again the (possibly enlarged) auxiliary probability space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$, on which originally $\bar{X}$ was defined. On this auxiliary space we define a random variable Y on G with distribution π. Y depends on the graph G (i.e. it depends on the random choices in Algorithm 5.4.1 that determine the states of edges), and it may depend on the random choices that determine the trajectory of $\bar{X}$ up to time $2\log^5 n$, but it is independent of all the random choices that determine the trajectory of $\bar{X}$ at times $2\log^5 n + k$, $k \ge 1$. By e.g. [LPW09, Proposition 4.7] we can choose Y such that $\tilde{P}[\bar{X}_{2\log^5 n} \ne Y] = \|\bar{\pi} - \pi\|_{\mathrm{TV}}$. By (5.4.3) it follows that

$$\tilde{P}[\bar{X}_{2\log^5 n} \ne Y] \le \frac{c}{n^{c'}}. \tag{5.4.9}$$

Moreover, we define on $\tilde{\Omega}$ a collection $\tilde{X}^z$, $z \in G$, of independent simple random walks on G started at z, independent of $\bar{X}$ and Y (i.e. depending only on the random choices in Algorithm 5.4.1 that determine the states of edges, but independent of the random choices that determine the trajectory of $\bar{X}$).

Define the process X using $\bar{X}$, Y and $\tilde{X}^z$ as follows,

$$\begin{aligned}&X_k = \bar{X}_{k+2\log^5 n} \text{ for } 0 \le k \le \tau, \\ &X_k = \tilde{X}_k^{\bar{X}_\tau} \text{ for } k > \tau,\end{aligned}\Big\} \quad \text{if } \bar{X}_{\log^5 n} \in \mathcal{C}_1 \text{ and } Y = \bar{X}_{2\log^5 n},$$

$$X_k = \tilde{X}_k^Y \text{ for } k \ge 0, \qquad \text{if } \bar{X}_{\log^5 n} \in \mathcal{C}_1 \text{ and } Y \ne \bar{X}_{2\log^5 n}, \tag{5.4.10}$$

$$X_k = \tilde{X}_k^Y \text{ for } k \ge 0, \qquad \text{if } \bar{X}_{\log^5 n} \notin \mathcal{C}_1.$$

Let Q^G denote the joint law of $\bar{X}$ and X on $\{1,2,...,n\}^{2\mathbb{N}_0}$. Since Y has distribution π and $\bar{X}$ behaves like a random walk between occurrences of step (4) of Algorithm 5.4.1, in any case X is a simple random walk on $\mathcal{C}_1$ started stationary, so Q^G is indeed a coupling of simple random walk on the giant component and the process $\bar{X}$ from Algorithm 5.4.1 with marginal laws $P^{\mathcal{C}_1}$ and $\bar{P}^G$ respectively.

By (5.4.9), (5.4.1) and (5.4.2) the first case of the coupling (5.4.10) happens with probability $\geq 1 - \frac{c}{n^{c'}}$, and by (5.4.2) also $\tau > u\rho(2-\xi)\xi n + 2\log^5 n$ with high probability, and the statement of the proposition follows. □

The coupling (5.4.10) defined in the proof of Proposition 5.4.3 will allow us to deduce the phase transition in the vacant set left by X from the phase transition in the vacant set left by $\bar{X}$. To apply the results of Section 5.3, we have to find the relation between the sizes of these vacant sets. This relation is given by the next lemma. Denote $\bar{\mathcal{V}}^u = \bar{\mathcal{V}}(u\rho(2-\xi)\xi n + 2\log^5 n) = G \setminus \bar{X}_{[0,u\rho(2-\xi)\xi n + 2\log^5 n]}$ and as before $\mathcal{V}^u = \mathcal{V}(u\rho(2-\xi)\xi n) = \mathcal{C}_1 \setminus X_{[0,u\rho(2-\xi)\xi n]}$.

Lemma 5.4.4. *For a sequence of typical graphs G and any fixed $u > 0$, with respect to the corresponding sequence of couplings Q^G, the random variables $|\bar{\mathcal{V}}^u|$ and $|\mathcal{V}^u|$ satisfy*

$$|\bar{\mathcal{V}}^u| = |\mathcal{V}^u| + (1-\xi)n + o(n) \quad Q^G\text{-a.a.s.}$$

Proof. Denote $\bar{\mathcal{W}}^u = G \setminus \bar{X}_{[2\log^5 n, u\rho(2-\xi)\xi n + 2\log^5 n]}$. Then $\big||\bar{\mathcal{V}}^u| - |\bar{\mathcal{W}}^u|\big| \leq 2\log^5 n$, and for any $\epsilon > 0$, $\epsilon n - 2\log^5 n \geq \frac{\epsilon}{2} n$ for n large enough, thus

$$Q^G\left[\big||\bar{\mathcal{V}}^u| - |\mathcal{V}^u| - (1-\xi)n\big| > \epsilon n\right] \leq Q^G\left[\big||\bar{\mathcal{W}}^u| - |\mathcal{V}^u| - (1-\xi)n\big| > \frac{\epsilon}{2} n\right].$$

By Proposition 5.4.3, Q^G-a.a.s. the sets $\bar{\mathcal{W}}^u$ and $\mathcal{V}^u$ differ only by the small components of the graph G, i.e. $\bar{\mathcal{W}}^u = \mathcal{V}^u \cup \bigcup_{i\geq 2} \mathcal{C}_i(G)$. By (5.2.1), in a typical graph G the total size of small components satisfies $\big|\bigcup_{i\geq 2} \mathcal{C}_i(G) - (1-\xi)n\big| \leq \frac{\epsilon}{2} n$ for n large enough. Therefore, for every $\epsilon > 0$,

$$Q^G\left[\big||\bar{\mathcal{W}}^u| - |\mathcal{V}^u| - (1-\xi)n\big| > \frac{\epsilon}{2} n\right] \leq Q^G\left[\bar{\mathcal{W}}^u \neq \mathcal{V}^u \cup \bigcup_{i\geq 2} \mathcal{C}_i(G)\right] \to 0 \text{ as } n \to \infty.$$

This proves the lemma. □

5.5. Proof of main result

We first extend the coupling Q^G that was defined for typical graphs in Proposition 5.4.3. Let Q^G for a non-typical graph G be the joint law on $\{1,2,...,n\}^{2\mathbb{N}_0}$ of two independent processes X and $\bar{X}$ under $P^{\mathcal{C}_1(G)}$ and $\bar{P}^G$ respectively. We define the annealed coupling measure $\mathbf{Q}_n$ on the space $\Omega'_n = \mathcal{G}(n) \times \{1,2,...,n\}^{2\mathbb{N}_0}$ with the canonical coordinates G, $\bar{X}$, X as

$$\mathbf{Q}_n(A \times B) = \sum_{G \in A} \mathbb{P}_{n,p}(G) Q^G(B),$$

where $A \in \mathbb{G}_n$ and $B = B_1 \times B_2$ with $B_i \in \mathbb{F}_n$ for $i = 1, 2$ (cf. (5.1.1) for the definition of the σ-algebras $\mathbb{G}_n$ and $\mathbb{F}_n$). Then $\mathbf{Q}_n$ is a coupling of the two processes X and $\bar{X}$, where X has marginal law $\mathbf{P}_n$ and $\bar{X}$ has marginal law $\bar{\mathbf{P}}_n$, and since every G is $\mathbb{P}_{n,p}$-a.a.s. a typical graph the statements of Proposition 5.4.3 and Lemma 5.4.4 hold $\mathbf{Q}_n$-a.a.s.

Proof of Theorem 5.1.1. For the proof we use the annealed coupling $\mathbf{Q}_n$ of X and $\bar{X}$. As a direct consequence of Proposition 5.3.1 and Lemma 5.4.4 we obtain that

$$|\bar{\mathcal{V}}^u| = \xi n \mathbb{E}_{\mathcal{T}}\left[e^{-u\,\mathrm{cap}_{\mathcal{T}}(\varnothing)}\right] + (1-\xi)n + o(n) \qquad \mathbf{Q}_n\text{-a.a.s.}$$

It follows from Lemma 5.4.2 and the classical results on random graphs that the graph $\bar{\mathcal{V}}^u = G \setminus \bar{X}_{[0,u\rho(2-\xi)\xi n + 2\log^5 n]}$ exhibits a phase transition at the value u such that $\lim_{n\to\infty} |\bar{\mathcal{V}}^u|\frac{\rho}{n} = 1$. This value is the solution $u_\star$ of the equation

$$\rho\xi\mathbb{E}_{\mathcal{T}}\left[e^{-u\,\mathrm{cap}_{\mathcal{T}}(\varnothing)}\right] + \rho(1-\xi) = 1. \tag{5.5.1}$$

$\bar{\mathcal{V}}^u$ has therefore $\mathbf{Q}_n$-a.a.s. a unique giant component $\mathcal{C}_1(\bar{\mathcal{V}}^u)$ of size $\zeta(u,\rho)n + o(n)$ and all other components of size smaller than $\bar{C}\log n$ if $u < u_\star$, and it has $\mathbf{Q}_n$-a.a.s. all components of size smaller than $\bar{C}\log n$ for $u > u_\star$, where $\bar{C} > 0$ is some fixed constant. For $u < u_\star$, the constant $\zeta(u,\rho)$ is given as the unique solution in $(0,1)$ of the equation

$$\exp\left\{-\zeta\left(\rho\xi\mathbb{E}_{\mathcal{T}}\left[e^{-u\,\mathrm{cap}_{\mathcal{T}}(\varnothing)}\right] + \rho(1-\xi)\right)\right\} = 1-\zeta. \tag{5.5.2}$$

It remains to translate this phase transition to the vacant graph $\mathcal{V}^u$ of the random walk on the giant component.

Let us first translate the phase transition to the subgraph induced by the slightly enlarged set $\bar{\mathcal{V}}^u \cup \bar{X}_{[0,2\log^5 n]}$. Adding one vertex of degree d in G to the graph $\bar{\mathcal{V}}^u$ can merge at most d components of $\bar{\mathcal{V}}^u$. By (5.2.3) the degree d is $\mathbf{Q}_n$-a.a.s. bounded by $\log n$, so adding the vertices of $\bar{X}_{[0,2\log^5 n]}$ can $\mathbf{Q}_n$-a.a.s. merge at most $2\log^6 n$ components. It follows that $\mathbf{Q}_n$-a.a.s., by adding $\bar{X}_{[0,2\log^5 n]}$ to $\bar{\mathcal{V}}^u$, any component of size smaller than $\bar{C}\log n$ in $\bar{\mathcal{V}}^u$ can either merge with the giant component if there is one, or it can become a component of size at most $2\bar{C}\log^7 n$. Also, in the supercritical phase the giant component can $\mathbf{Q}_n$-a.a.s. grow by at most $2\bar{C}\log^7 n = o(n)$. Therefore, the graph induced by $\bar{\mathcal{V}}^u \cup \bar{X}_{[0,2\log^5 n]}$ exhibits a phase transition at $u_\star$ with the same size $\zeta(u,\rho)n + o(n)$ of the giant component for $u < u_\star$, and with the bound $2\bar{C}\log^7 n$ for the size of the second largest component for $u < u_\star$ and the largest component for $u > u_\star$.

Recall that $\bar{\mathcal{W}}^u$ denotes the set $G \setminus \bar{X}_{[2\log^5 n, u\rho(2-\xi)\xi n + 2\log^5 n]}$ as well as the induced subgraph. We have the following inclusions of sets and induced subgraphs in G,

$$\bar{\mathcal{V}}^u \subset \bar{\mathcal{W}}^u \subset \bar{\mathcal{V}}^u \cup \bar{X}_{[0,2\log^5 n]}.$$

Consider the vacant set $\mathcal{V}^u \subset \mathcal{C}_1$ of the random walk X on the giant component. By Proposition 5.4.3 and since every graph is $\mathbb{P}_{n,p}$-a.a.s. a typical graph, we have $\mathbf{Q}_n$-a.a.s. $\bar{\mathcal{W}}^u = \mathcal{V}^u \cup \bigcup_{i\geq 2}\mathcal{C}_i(G)$. It follows that

$$\bar{\mathcal{V}}^u \subset \mathcal{V}^u \cup \bigcup_{i\geq 2}\mathcal{C}_i(G) \quad \mathbf{Q}_n\text{-a.a.s.} \tag{5.5.3}$$

$$\mathcal{V}^u \subset \bar{\mathcal{V}}^u \cup \bar{X}_{[0,2\log^5 n]} \quad \mathbf{Q}_n\text{-a.a.s.} \tag{5.5.4}$$

Note that $\mathbf{Q}_n$-a.a.s. the union $\bigcup_{i\geq 2}\mathcal{C}_i(G)$ of all components of G except the largest are exactly all small components of size smaller than $C\log n$. From this and (5.5.3) it follows that $|\mathcal{C}_1(\mathcal{V}^u)|$ is $\mathbf{Q}_n$-a.a.s. bounded from below by $|\mathcal{C}_1(\bar{\mathcal{V}}^u)|$ whenever $|\mathcal{C}_1(\bar{\mathcal{V}}^u)|$ is larger than of order $\log n$. From (5.5.4) it follows that $|\mathcal{C}_1(\mathcal{V}^u)|$ is Q_n-a.a.s. bounded from

above by $\left|\mathcal{C}_1\left(\bar{\mathcal{V}}^u \cup \bar{X}_{[0,2\log^5 n]}\right)\right|$. The respective phase transitions in $\bar{\mathcal{V}}^u$ and $\bar{\mathcal{V}}^u \cup \bar{X}_{[0,2\log^5 n]}$ thus immediately imply the statements (5.1.3) and (5.1.5) of Theorem 5.1.1.

To prove (5.1.4), i.e. the uniqueness of the giant component in the supercritical phase, fix $u < u_\star$ and let $\mathcal{L}_n$ be the event that there are two distinct components $\mathcal{C}_a$ and $\mathcal{C}_b$ in $\mathcal{V}^u$ both of size strictly larger than $2\bar{C}\log^7 n$, with $\bar{C}$ as defined below (5.5.1). We show that $\mathbf{Q}_n[\mathcal{L}_n] \to 0$ as $n \to \infty$, which proves (5.1.4). First note that if $\mathcal{L}_n$ happens, then either $\mathcal{C}_a \cap \bar{\mathcal{V}}^u$ and $\mathcal{C}_b \cap \bar{\mathcal{V}}^u$ are distinct components in $\bar{\mathcal{V}}^u$ or the inclusion in (5.5.3) does not hold, which is unlikely, so

$$\mathbf{Q}_n[\mathcal{L}_n] \le \mathbf{Q}_n\left[\mathcal{L}_n,\ \mathcal{C}_a \cap \bar{\mathcal{V}}^u \text{ and } \mathcal{C}_b \cap \bar{\mathcal{V}}^u \text{ are distinct components in } \bar{\mathcal{V}}^u\right] + o(1) \text{ as } n \to \infty.$$

But if $\mathcal{C}_a \cap \bar{\mathcal{V}}^u$ and $\mathcal{C}_b \cap \bar{\mathcal{V}}^u$ are distinct components in $\bar{\mathcal{V}}^u$, at least one of $\mathcal{C}_a \cap \bar{\mathcal{V}}^u$ or $\mathcal{C}_b \cap \bar{\mathcal{V}}^u$ is subset of $\bigcup_{i\ge 2}\mathcal{C}_i(\bar{\mathcal{V}}^u)$, which is a union of components that are $\mathbf{Q}_n$-a.a.s. all of size smaller than $\bar{C}\log n$. On the other hand by (5.5.4), $\mathcal{C}_a \subset \left(\mathcal{C}_a \cap \bar{\mathcal{V}}^u\right) \cup \bar{X}_{[0,2\log^5 n]}$, and as discussed before this last union cannot be larger than $2\bar{C}\log^7 n$ if $\mathcal{C}_a \cap \bar{\mathcal{V}}^u$ consists only of components of size smaller than $\bar{C}\log n$. Thus

$$\begin{aligned}
\mathbf{Q}_n[\mathcal{L}_n] &\le \mathbf{Q}_n\left[\mathcal{L}_n,\ \mathcal{C}_a \cap \bar{\mathcal{V}}^u \text{ or } \mathcal{C}_b \cap \bar{\mathcal{V}}^u \text{ is subset of } \textstyle\bigcup_{i\ge 2}\mathcal{C}_i(\bar{\mathcal{V}}^u)\right] + o(1)\\
&\le \mathbf{Q}_n\left[\text{at least one of the } \mathcal{C}_i(\bar{\mathcal{V}}^u),\ i \ge 2, \text{ is larger than } \bar{C}\log n\right] + o(1)\\
&= o(1) \text{ as } n \to \infty.
\end{aligned}$$

This proves (5.1.4).

To see that the critical parameter $u_\star$ coincides with the critical parameter $u_\star$ of random interlacements on a Poisson(ρ)-Galton-Watson tree conditioned on non-extinction, it suffices to notice that the characterizing equations (5.5.1) and (5.2.15) of these two parameters are the same. □

Bibliography

[AB92] D. J. Aldous and M. Brown, *Inequalities for rare events in time-reversible Markov chains. I*, Stochastic inequalities (Seattle, WA, 1991), IMS Lecture Notes Monogr. Ser., vol. 22, Inst. Math. Statist., Hayward, CA, 1992, pp. 1–16. MR 1228050

[ABS04] N. Alon, I. Benjamini, and A. Stacey, *Percolation on finite graphs and isoperimetric inequalities*, Ann. Probab. **32** (2004), no. 3A, 1727–1745. MR 2073175

[AF02] D. J. Aldous and J. A. Fill, *Reversible markov chains and random walks on graphs*, 2002, Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[AKL+79] R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovász, and C. Rackoff, *Random walks, universal traversal sequences, and the complexity of maze problems*, 20th Annual Symposium on Foundations of Computer Science (San Juan, Puerto Rico, 1979), IEEE, New York, 1979, pp. 218–223. MR 598110

[Ald97] D. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, Ann. Probab. **25** (1997), no. 2, 812–854. MR 1434128

[AN72] K. B. Athreya and P. E. Ney, *Branching processes*, Springer-Verlag, New York, 1972, Die Grundlehren der mathematischen Wissenschaften, Band 196. MR 0373040

[BBČ08] G. Ben Arous, A. Bovier, and J. Černý, *Universality of the REM for dynamics of mean-field spin glasses*, Comm. Math. Phys. **282** (2008), no. 3, 663–695. MR 2426140

[BBG02] G. Ben Arous, A. Bovier, and V. Gayrard, *Aging in the random energy model*, Physical Review Letters **88** (2002), no. 8, 087201.

[BBG03a] G. Ben Arous, A. Bovier, and V. Gayrard, *Glauber dynamics of the random energy model. I. Metastable motion on the extreme states*, Comm. Math. Phys. **235** (2003), no. 3, 379–425. MR 1974509

[BBG03b] G. Ben Arous, A. Bovier, and V. Gayrard, *Glauber dynamics of the random energy model. II. Aging below the critical temperature*, Comm. Math. Phys. **236** (2003), no. 1, 1–54. MR 1977880

[BČ05] G. Ben Arous and J. Černý, *Bouchaud's model exhibits two different aging regimes in dimension one*, Ann. Appl. Probab. **15** (2005), no. 2, 1161–1192. MR 2134101

[BČ06] G. Ben Arous and J. Černý, *Dynamics of trap models*, Mathematical statistical physics, Elsevier B. V., Amsterdam, 2006, pp. 331–394. MR 2581889

[BČ07] G. Ben Arous and J. Černý, *Scaling limit for trap models on $\mathbb{Z}^d$*, Ann. Probab. **35** (2007), no. 6, 2356–2384. MR 2353391

[BČ08] G. Ben Arous and J. Černý, *The arcsine law as a universal aging scheme for trap models*, Comm. Pure Appl. Math. **61** (2008), no. 3, 289–329. MR 2376843

[BČ11] M. T. Barlow and J. Černý, *Convergence to fractional kinetics for random walks associated with unbounded conductances*, Probab. Theory Related Fields **149** (2011), no. 3-4, 639–673. MR 2776627

[BCČR14] G. Ben Arous, M. Cabezas, J. Černý, and R. Royfman, *Randomly Trapped Random Walks*, to appear in Ann. Probab., arXiv:1302.7227, 2014.

[BCKM97] J.-P. Bouchaud, L. F. Cugliandolo, J. Kurchan, and M. Mézard, *Out of equilibrium dynamics in spin-glasses and other glassy systems*, Spin Glasses and Random Fields, World Scientific, 1997, pp. 161–223.

[BČM06] G. Ben Arous, J. Černý, and T. Mountford, *Aging in two-dimensional Bouchaud's model*, Probab. Theory Related Fields **134** (2006), no. 1, 1–43. MR 2221784

[BD95] J.-P. Bouchaud and D.S. Dean, *Aging on parisi's tree*, J. Phys. I France **5** (1995), no. 3, 265–286.

[BD10] M. T. Barlow and J.-D. Deuschel, *Invariance principle for the random conductance model with unbounded conductances*, Ann. Probab. **38** (2010), no. 1, 234–276. MR 2599199

[Ben02] G. Ben Arous, *Aging and spin-glass dynamics*, Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002), Higher Ed. Press, Beijing, 2002, pp. 3–14. MR 1957514

[Ber96] J. Bertoin, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996. MR 1406564

[BG12] G. Ben Arous and O. Gün, *Universality and extremal aging for dynamics of spin glasses on subexponential time scales*, Comm. Pure Appl. Math. **65** (2012), no. 1, 77–127. MR 2846638

[BG13] A. Bovier and V. Gayrard, *Convergence of clock processes in random environments and ageing in the p-spin SK model*, Ann. Probab. **41** (2013), no. 2, 817–847. MR 3077527

[BGŠ13] A. Bovier, V. Gayrard, and A. Švejda, *Convergence to extremal processes in random environments and extremal ageing in SK models*, Probab. Theory Related Fields **157** (2013), no. 1-2, 251–283. MR 3101847

[BKW14] I. Benjamini, G. Kozma, and N. Wormald, *The mixing time of the giant component of a random graph*, Random Structures Algorithms **45** (2014), no. 3, 383–407. MR 3252922

[Bol01] B Bollobás, *Random graphs*, second ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001. MR 1864966

[Bou92] J.-P. Bouchaud, *Weak ergodicity breaking and aging in disordered systems*, J. Phys. I France **2** (1992), no. 9, 1705–1713.

[BS02] E. Bolthausen and A.-S. Sznitman, *On the static and dynamic points of view for certain random walks in random environment*, Methods Appl. Anal. **9** (2002), no. 3, 345–375, Special issue dedicated to Daniel W. Stroock and Srinivasa S. R. Varadhan on the occasion of their 60th birthday. MR 2023130

[BS08] I. Benjamini and A.-S. Sznitman, *Giant component and vacant set for random walk on a discrete torus*, J. Eur. Math. Soc. (JEMS) **10** (2008), no. 1, 133–172. MR 2349899

[Čer07] J. Černý, *Moments and distribution of the local time of a two-dimensional random walk*, Stochastic Process. Appl. **117** (2007), no. 2, 262–270. MR 2290196

[Čer11] J. Černý, *On two-dimensional random walk among heavy-tailed conductances*, Electron. J. Probab. **16** (2011), 293–313. MR 2771138

[CF11] C. Cooper and A. Frieze, *Component structure of the vacant set induced by a random walk on a random graph*, Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (Philadelphia, PA), SIAM, 2011, pp. 1211–1221. MR 2858394

[Cha84] R. V. Chamberlin, *Time decay of the thermoremanent magnetization in spin-glasses as a function of the time spent in the field-cooled state*, Phys. Rev. B **30** (1984), no. 9, 5393–5395.

[ČT12] J. Černý and A. Teixeira, *From random walk trajectories to random interlacements*, Ensaios Matemáticos [Mathematical Surveys], vol. 23, Sociedade Brasileira de Matemática, Rio de Janeiro, 2012. MR 3014964

[ČT13] J. Černý and A. Teixeira, *Critical window for the vacant set left by random walk on random regular graphs*, Random Structures Algorithms **43** (2013), no. 3, 313–337. MR 3094422

[ČT14] J. Černý and A. Teixeira, *Random walks on torus and random interlacements: Macroscopic coupling and phase transition*, arXiv:1411.7795, 2014.

[ČTW11] J. Černý, A. Teixeira, and D. Windisch, *Giant vacant component left by a random walk in a random d-regular graph*, Ann. Inst. Henri Poincaré Probab. Stat. **47** (2011), no. 4, 929–968. MR 2884219

[ČW15a] J. Černý and T. Wassmer, *Aging of the Metropolis dynamics on the Random Energy Model*, submitted preprint, arXiv:1502.04535, 2015.

[ČW15b] J. Černý and T. Wassmer, *Randomly trapped random walks on* $\mathbb{Z}^d$, Stochastic Process. Appl. **125** (2015), no. 3, 1032–1057. MR 3303967

[Dar52] D. A. Darling, *The influence of the maximum term in the addition of independent random variables*, Trans. Amer. Math. Soc. **73** (1952), 95–107. MR 0048726

[DE51] A. Dvoretzky and P. Erdős, *Some problems on random walk in space*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, University of California Press, Berkeley and Los Angeles, 1951, pp. 353–367. MR 0047272

[Der80] B. Derrida, *Random-energy model: limit of a family of disordered models*, Phys. Rev. Lett. **45** (1980), no. 2, 79–82. MR 575260

[Der81] B. Derrida, *Random-energy model: an exactly solvable model of disordered systems*, Phys. Rev. B (3) **24** (1981), no. 5, 2613–2626. MR 627810

[DR78] R. Durrett and S. I. Resnick, *Functional limit theorems for dependent variables*, Ann. Probab. **6** (1978), no. 5, 829–846. MR 503954

[DRS14a] A. Drewitz, B. Ráth, and A. Sapozhnikov, *An introduction to random interlacements*, Springer Briefs in Mathematics, Springer, Cham, 2014. MR 3308116

[DRS14b] A. Drewitz, B. Ráth, and A. Sapozhnikov, *Local percolative properties of the vacant set of random interlacements with small intensity*, Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 4, 1165–1197. MR 3269990

[DS91] P. Diaconis and D. Stroock, *Geometric bounds for eigenvalues of Markov chains*, Ann. Appl. Probab. **1** (1991), no. 1, 36–61. MR 1097463

[Dur10] R. Durrett, *Random graph dynamics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2010. MR 2656427

[Eis83] T. Eisele, *On a third-order phase transition*, Comm. Math. Phys. **90** (1983), no. 1, 125–159. MR 714615

[ER61] P. Erdős and A. Rényi, *On the evolution of random graphs*, Bull. Inst. Internat. Statist. **38** (1961), 343–347. MR 0148055

[ET60] P. Erdős and S. J. Taylor, *Some problems concerning the structure of random walk paths*, Acta Math. Acad. Sci. Hungar. **11** (1960), 137–162. (unbound insert). MR 0121870

[FIKP98] L. R. G. Fontes, M. Isopi, Y. Kohayakawa, and P. Picco, *The spectral gap of the REM under Metropolis dynamics*, Ann. Appl. Probab. **8** (1998), no. 3, 917–943. MR 1627811

[FIN02] L. R. G. Fontes, M. Isopi, and C. M. Newman, *Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension*, Ann. Probab. **30** (2002), no. 2, 579–604. MR 1905852

[FM14] L. R. G. Fontes and P. Mathieu, *On the dynamics of trap models in* $\mathbb{Z}^d$, Proc. Lond. Math. Soc. (3) **108** (2014), no. 6, 1562–1592. MR 3218319

[FR08] N. Fountoulakis and B. A. Reed, *The evolution of the mixing rate of a simple random walk on the giant component of a random graph*, Random Structures Algorithms **33** (2008), no. 1, 68–86. MR 2428978

[Gü09] O. Gün, *Universality of transient dynamics and aging for spin glasses*, PhD thesis, New York University, 2009. MR 2717915

[Gay10] V. Gayrard, *Aging in reversible dynamics of disordered systems. II. Emergence of the arcsine law in the random hopping time dynamics of the REM*, arXiv:1008.3849, 2010.

[Gay12] V. Gayrard, *Convergence of clock process in random environments and aging in Bouchaud's asymmetric trap model on the complete graph*, Electron. J. Probab. **17** (2012), no. 58, 33. MR 2959064

[Gay14] V. Gayrard, *Convergence of clock processes and aging in Metropolis dynamics of a truncated REM*, arXiv:1402.0388, 2014.

[GŠ13] V. Gayrard and A. Švejda, *Convergence of clock processes on infinite graphs and aging in Bouchaud's asymmetric trap model on* $\mathbb{Z}^d$, arXiv:1309.3066, 2013.

[HLN87] J. Hastad, T. Leighton, and M. Newman, *Reconfiguring a hypercube in the presence of faults*, Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '87, ACM, 1987, pp. 274–284.

[HM12] H. Hatami and M. Molloy, *The scaling window for a random graph with a given degree sequence*, Random Structures Algorithms **41** (2012), no. 1, 99–123. MR 2943428

[HvdH07] M. Heydenreich and R. van der Hofstad, *Random graph asymptotics on high-dimensional tori*, Comm. Math. Phys. **270** (2007), no. 2, 335–358. MR 2276449

[HvdH11] M. Heydenreich and R. van der Hofstad, *Random graph asymptotics on high-dimensional tori II: volume, diameter and mixing time*, Probab. Theory Related Fields **149** (2011), no. 3-4, 397–415. MR 2776620

[JŁR00] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. MR 1782847

[JLT14] M. Jara, C. Landim, and A. Teixeira, *Universality of trap models in the ergodic time scale*, Ann. Probab. **42** (2014), no. 6, 2497–2557. MR 3265173

[JS03] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR 1943877

[Kal02] O. Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR 1876169

[KS63] H. Kesten and F. Spitzer, *Ratio theorems for random walks. I*, J. Analyse Math. **11** (1963), 285–322. MR 0162279

[LP14] R. Lyons and Y. Peres, *Probability on trees and networks*, Cambridge University Press. In preparation, current version available at `http://mypage.iu.edu/~rdlyons/`, 2014.

[LPW09] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009, With a chapter by James G. Propp and David B. Wilson. MR 2466937

[LSNB83] L. Lundgren, P. Svedlindh, P. Nordblad, and O. Beckman, *Dynamics of the relaxation-time spectrum in a CuMn spin-glass*, Phys. Rev. Lett. **51** (1983), no. 10, 911–914.

[McD98] C. McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, Algorithms Combin., vol. 16, Springer, Berlin, 1998, pp. 195–248. MR 1678578

[MM15] P. Mathieu and J.-C. Mourrat, *Aging of asymmetric dynamics on the random energy model*, Probab. Theory Related Fields **161** (2015), no. 1-2, 351–427. MR 3304755

[Mou11] J.-C. Mourrat, *Scaling limit of the random walk among random traps on* $\mathbb{Z}^d$, Ann. Inst. Henri Poincaré Probab. Stat. **47** (2011), no. 3, 813–849. MR 2841076

[MR95] M. Molloy and B. Reed, *A critical point for random graphs with a given degree sequence*, Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993), vol. 6, 1995, pp. 161–179. MR 1370952

[MS04] M. M. Meerschaert and H.-P. Scheffler, *Limit theorems for continuous-time random walks with infinite mean waiting times*, J. Appl. Probab. **41** (2004), no. 3, 623–638. MR 2074812

[MW65] E. W. Montroll and G. H. Weiss, *Random walks on lattices. II*, J. Mathematical Phys. **6** (1965), 167–181. MR 0172344

[NP10] A. Nachmias and Y. Peres, *Critical percolation on random regular graphs*, Random Structures Algorithms **36** (2010), no. 2, 111–148. MR 2583058

[OP84] E. Olivieri and P. Picco, *On the existence of thermodynamics for the random energy model*, Comm. Math. Phys. **96** (1984), no. 1, 125–144. MR 765963

[Pan13] D. Panchenko, *The Sherrington-Kirkpatrick model*, Springer Monographs in Mathematics, Springer, New York, 2013. MR 3052333

[Pet83] K. Petersen, *Ergodic theory*, Cambridge Studies in Advanced Mathematics, vol. 2, Cambridge University Press, Cambridge, 1983. MR 833286

[Pit74] J. H. Pitt, *Multiple points of transient random walks*, Proc. Amer. Math. Soc. **43** (1974), 195–199. MR 0386021

[Rév13] P. Révész, *Random walk in random and non-random environments*, third ed., World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2013. MR 3060348

[Rio12] O. Riordan, *The phase transition in the configuration model*, Combin. Probab. Comput. **21** (2012), no. 1-2, 265–299. MR 2900063

[Sen76] E. Seneta, *Regularly varying functions*, Lecture Notes in Mathematics, Vol. 508, Springer-Verlag, Berlin-New York, 1976. MR 0453936

[SK75] D. Sherrington and S. Kirkpatrick, *Solvable model of a spin-glass*, Phys. Rev. Let. **35** (1975), 1792–1796.

[Spi76] F. Spitzer, *Principles of random walk*, second ed., Springer-Verlag, New York-Heidelberg, 1976, Graduate Texts in Mathematics, Vol. 34. MR 0388547

[SS09] V. Sidoravicius and A.-S. Sznitman, *Percolation for the vacant set of random interlacements*, Comm. Pure Appl. Math. **62** (2009), no. 6, 831–858. MR 2512613

[SS10] V. Sidoravicius and A.-S. Sznitman, *Connectivity bounds for the vacant set of random interlacements*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 4, 976–990. MR 2744881

[Szn10] A.-S. Sznitman, *Vacant set of random interlacements and percolation*, Ann. of Math. (2) **171** (2010), no. 3, 2039–2087. MR 2680403

[Tal11a] M. Talagrand, *Mean field models for spin glasses. Volume I*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 54, Springer-Verlag, Berlin, 2011, Basic examples. MR 2731561

[Tal11b] M. Talagrand, *Mean field models for spin glasses. Volume II*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 55, Springer, Heidelberg, 2011, Advanced replica-symmetry and low temperature. MR 3024566

[Tas10] M. Tassy, *Random interlacements on Galton-Watson trees*, Electron. Commun. Probab. **15** (2010), 562–571. MR 2737713

[Tei09] A. Teixeira, *Interlacement percolation on transient weighted graphs*, Electron. J. Probab. **14** (2009), no. 54, 1604–1628. MR 2525105

[TW11] A. Teixeira and D. Windisch, *On the fragmentation of a torus by random walk*, Comm. Pure Appl. Math. **64** (2011), no. 12, 1599–1646. MR 2838338

[vdH08] R. van der Hofstad, *Random graphs and complex networks*, Lecture notes in preparation, `http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf`, 2008.

[VHO+97] E. Vincent, J. Hammann, M. Ocio, J.-P. Bouchaud, and L. F. Cugliandolo, *Slow dynamics and aging in spin glasses*, Complex Behavior of Glassy Systems (M. Rubi, ed.), Lecture notes in Physics, vol. 492, Springer-Verlag, Berlin, 1997, (arXiv:cond-mat/9607224), pp. 184–219.

[Was15] T. Wassmer, *Phase transition for the vacant set left by random walk on the giant component of a random graph*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 756–780. MR 3335024

[Whi02] W. Whitt, *Stochastic-process limits*, Springer Series in Operations Research, Springer-Verlag, New York, 2002, An introduction to stochastic-process limits and their application to queues. MR 1876437

[Win08] D. Windisch, *Random walk on a discrete torus and random interlacements*, Electron. Commun. Probab. **13** (2008), 140–150. MR 2386070

Tobias Wassmer – Curriculum Vitae

Born: 21 February 1986

Nationality: Swiss

Languages: German (native)– English – French – Italian

Education

2012–2015 PhD in Mathematics, University of Vienna. Advisor: Jiří Černý

2010–2011 MSc in Mathematics, Eidgenössische Technische Hochschule Zürich. Thesis: *Large deviations for random matrices*, Advisor: Jiří Černý

2006–2010 BSc in Mathematics, Eidgenössische Technische Hochschule Zürich. Thesis: *On the supercritical Galton Watson branching process*, Advisor: Alain-Sol Sznitman

2001–2005 Eidgenössische Maturität, Neue Kantonsschule Aarau

Employement

2012–2015 University Assistant PraeDoc, University of Vienna

2011–2012 Teacher for Mathematics, Alte Kantonsschule Aarau