SCALING LIMITS FOR THE CRITICAL LEVEL-SET PERCOLATION OF THE GAUSSIAN FREE FIELD ON REGULAR TREES

JIŘÍ ČERNÝ, RAMON LOCHER

ABSTRACT. We continue the study of the level-set percolation of the discrete Gaussian free field (GFF) on regular trees in the critical regime, initiated in [ČL25]. First, we derive a sharp asymptotic estimate for the probability that the connected component of the critical level set containing the root of the tree reaches generation n. In particular, we show that the one-arm exponent satisfies $\rho = 1$. Next, we establish a Yaglom-type limit theorem for the values of the GFF at generation n within this component. Finally, we show that, after a correct rescaling, this component conditioned on reaching generation n converges, as $n \to \infty$, to Aldous' continuum random tree.

1. INTRODUCTION

The Gaussian free field's level-set percolation, especially on $\mathbb{Z}^d$, is a significant model in percolation theory that is characterised by its long-range dependencies. Initial investigations into this model trace back to the 1980s, with pioneering studies [BLM87, MS83, LS86]. Over the last decade, renewed interest has been ignited by the findings in [RS13], demonstrating that on $\mathbb{Z}^d$, the model undergoes a distinctive percolative phase transition at a critical threshold $h^* = h^*(d)$ for any dimension $d \geq 3$. Follow-up research, including papers [DRS14, PR15, DPR18, Szn19, CN20, GRS22, PS22], has provided a comprehensive understanding of the model's behaviour in both subcritical and supercritical phases, often making use of additional natural critical points in order to work in strongly sub-/super-critical regime. Notably, [DCGRS23] confirmed the alignment of these critical points with h^*, indicating a precise phase transition.

In this paper, we continue the study of level-set percolation of the discrete Gaussian free field (GFF) on regular trees in the critical regime, building on the work initiated in [ČL25]. Our main contributions are threefold: First, we derive a sharp asymptotic estimate for the one-arm probability. Second, we establish a Yaglom-type limit theorem for the field at vertices located at distance n from the root. Finally, we show that the connected component of the critical level set, conditioned to be large, converges to Aldous' Continuum Random Tree. These results are stated precisely in Theorems 2.1, 2.2, and 2.3, respectively.

The considered model was initially studied in [Szn16] where the critical value h^* was identified in terms of the largest eigenvalue of a specific integral operator. Additionally, a comparison with random interlacements was employed to establish bounds on h^*, notably demonstrating that $0 < h^* < \infty$. Subsequently, in [AČ20], sub- and supercritical phases of the model were studied in detail. The main results of this paper include the continuity of the percolation probability outside the critical level h^*, and accurate estimates on the size of connected components of the level sets in both phases. Later, in our previous paper [ČL25], properties of the critical and the near-critical model were investigated. It was proved there that there is no percolation at the critical point h^*, and that the percolation probability is continuous also at this point, with a precise asymptotics formula for this probability in the regime $h \uparrow h^*$. Further, we provided rather precise estimates on the tail of the size of the connected component at criticality. Here, we complement these findings with further results concerning the model at the critical level h^*.

As in [Szn16, AČ20, ČL25], we will strongly rely on the fact that the model admits a representation as a branching process with an uncountable and unbounded type space. For the critical single-type Galton-Watson process, analogous versions of Theorems 2.1 and 2.2 go back to Kolmogorov [Kol38] and Yaglom [Yag47]. In the more general case of branching processes with a finite type space, similar results are also well known. General scaling limits in the spirit of Theorem 2.3 were first introduced for critical Galton-Watson processes in [Ald91], and later extended to branching processes with finite or countably infinite types [Mie08, dR17]. More recently, [Pow19] shows a similar scaling limit result in the setting of critical branching diffusion on bounded domains.

Recent advances in branching process theory have extended these classical results in several important directions. [CTJP24] obtained refined convergence rates for Yaglom limits in varying environments, providing Wasserstein metric bounds that may also be useful for analyses in our unbounded type-space setting. [BDIM23] established Yaglom-type theorems for branching processes in sparse random environments, an intermediate framework that connects the classical Galton-Watson model with fully random environments. [BFRS24] proved a Yaglom-type theorem for near-critical branching processes in random environments and further showed that, under survival conditioning, the genealogical structure of the population at a fixed time horizon converges to a time-changed Brownian coalescent point process.

Due to the nature of the branching process appearing in our model (in particular because its type space is uncountable and unbounded), no previous results are directly applicable to our setting. In this article, we adapt and extend the strategy of [Pow19], addressing the fundamental challenge of unbounded type spaces.

2. Model and results

We start with the definition of the model. Let $\mathbb{T}$ be the infinite $(d+1)$-regular tree, $d \geq 2$, rooted at an arbitrary fixed vertex $o \in \mathbb{T}$, endowed with the usual graph distance $d(\cdot,\cdot)$. On $\mathbb{T}$, we consider the Gaussian free field $\varphi = (\varphi_v)_{v\in\mathbb{T}}$ which is a centred Gaussian process whose covariance function agrees with the Green function of the simple random walk on $\mathbb{T}$ (see (3.2) for the precise definition). We use P to denote the law of this process on $\mathbb{R}^{\mathbb{T}}$. For $x \in \mathbb{R}$, we write P_x for the conditional distribution of φ given that $\varphi_o = x$,

$$P_x[\,\cdot\,] := P[\,\cdot \mid \varphi_o = x]. \tag{2.1}$$

(For an explicit construction of P_x, see (3.4) and the paragraph below it.) Furthermore, let $\bar{o} \in \mathbb{T}$ be an arbitrary fixed neighbour of the root o, and define the forward tree $\mathbb{T}^+$ by

$$\mathbb{T}^+ := \{v \in \mathbb{T} : \bar{o} \text{ is not contained in the geodesic path from } o \text{ to } v\}. \tag{2.2}$$

We analyse the percolation properties of the (super-)level sets of φ above level $h \in \mathbb{R}$,

$$E_\varphi^h := \{v \in \mathbb{T} : \varphi_v \geq h\}. \tag{2.3}$$

In particular, we are interested in the connected component of this set containing the root o,

$$\mathcal{C}_o^h := \{v \in \mathbb{T} : v \text{ is connected to } o \text{ in } E_\varphi^h\}, \quad h \in \mathbb{R}. \tag{2.4}$$

The critical height h^* of the level-set percolation is defined by

$$h^* = h^*(d) := \inf\big\{h \in \mathbb{R} : P[|\mathcal{C}_o^h| = \infty] = 0\big\}. \tag{2.5}$$

It is well known that h^* is non-trivial and strictly positive (see [Szn16, Corollary 4.5]). Moreover, as proved in [Szn16], h^* can be characterized with the help of the operator norms of a certain family of non-negative operators $(L_h)_{h\in\mathbb{R}}$ acting on the space $L^2(\nu)$, where ν is a centred Gaussian measure with variance $\sigma_\nu^2 = d/(d-1)$. We provide more details on this characterization in Section 3 below. Here, we only define λ_h to be the largest eigenvalue of L_h and χ_h the corresponding normed eigenfunction, and recall that h^* is the unique solution to

$$\lambda_{h^*} = 1. \tag{2.6}$$

The scalar product on $L^2(\nu)$ will be denoted by $\langle\cdot,\cdot\rangle$.

We use N_n^h to denote the set of vertices in $\mathcal{C}_o^h$ that are at a distance n from the root,

$$N_n^h = \{v \in \mathcal{C}_o^h : d(v,o) = n\}. \tag{2.7}$$

and set

$$N_n^{h,+} = N_n^h \cap \mathbb{T}^+. \tag{2.8}$$

Since we almost exclusively deal with the critical case, we abbreviate $\chi \coloneqq \chi_{h^*}$, $\mathcal{C}_o = \mathcal{C}_o^{h^*}$, $L \coloneqq L_{h^*}$, $N_n \coloneqq N_n^{h^*}$ and $N_n^+ \coloneqq N_n^{h^*,+}$.

Our first result describes the exact asymptotic behaviour of the probabilities (conditional and unconditional) that N_n and N_n^+ are non-empty, that is, that $\mathcal{C}_o$ has diameter at least n.

Theorem 2.1. *For every $x \ge h^*$, as $n \to \infty$,*

$$P_x[N_n^+ \neq \emptyset] = C_1 \chi(x) n^{-1}(1+o(1)), \tag{2.9}$$

$$P[N_n^+ \neq \emptyset] = C_1 \langle 1,\chi\rangle n^{-1}(1+o(1)), \tag{2.10}$$

where

$$C_1 = \frac{2d}{d-1}\frac{1}{\langle \chi^2,\chi\rangle}. \tag{2.11}$$

If the event $\{N_n^+ \neq \emptyset\}$ is replaced by $\{N_n \neq \emptyset\}$, the same results hold with C_1 replaced by $\widetilde{C}_1 = C_1(d+1)/d$.

In particular, Theorem 2.1 proves that the one-arm exponent of the critical level set, defined by $\rho = -\lim_{n\to\infty} \log n/\log(P[N_n \neq \emptyset])$ (see, e.g., [Gri99, Section 9.1]) satisfies

$$\rho = 1. \tag{2.12}$$

This complements the values of two other important critical exponents for our model given in [ČL25, (2.22)], where it was shown that $\delta = 2$ and $\beta = 1$.

The one-arm probability and the associated critical exponent is an actively studied quantity in several prominent percolation models. Notably, in the context of level-set percolation of the GFF on the metric graph of $\mathbb{Z}^d$, recent work has led to a detailed understanding of the one-arm probability across all dimensions. This includes the derivation of bounds for $d = 3$ in [DPR25], for $d > 6$ in [CD25], and for the intermediate regime $3 \le d \le 6$ in [CD24].

Our remaining main results consider the critical component conditioned on being large, more precisely conditioned on the rare event $\{N_n^+ \neq \emptyset\}$. The first such result is a Yaglom-type limit theorem for the GFF restricted to N_n^+.

Theorem 2.2. *For $f \in L^2(\nu)$, $n \geq 1$, and $x \geq h^*$, let $Z_n^{f,x}$, resp. Z_n^f, be a random variable distributed as $n^{-1} \sum_{v \in N_n^+} f(\varphi_v)$ under the conditional measure $P_x[\,\cdot\,|N_n^+ \neq \emptyset]$, resp. $P[\,\cdot\,|N_n^+ \neq \emptyset]$. Let further Z be an exponential random variable with mean one. Then, with C_1 as in (2.11),*

$$\lim_{n\to\infty} Z_n^{f,x} = \lim_{n\to\infty} Z_n^f = C_1^{-1}\langle \chi, f\rangle Z \qquad \text{in distribution.} \tag{2.13}$$

In the case of the critical single-type Galton-Watson process, results analogous to Theorem 2.2 trace back to the classical work of Yaglom [Yag47]. For more general branching processes with a finite type space, similar theorems are also well established; see, for instance, [Mod71, Theorem 10.1] and other foundational texts in the branching process literature. In recent years, Yaglom-type limits have been proven in various extended settings. These include critical non-local branching Markov processes [HHKW22], branching Brownian motion with absorption [MS22], and branching diffusions in bounded domains [Pow19]. Additionally, [GLLP22] establishes a Yaglom-type result for critical branching processes in a random Markovian environment with finite state space.

Our third result concerns a scaling limit for the critical component $\mathcal{C}_o \cap \mathbb{T}^+$, viewed as a metric space, under the conditional law $P_x[\,\cdot\,|N_n^+ \neq \emptyset]$. To this end, let $(T_{n,x}, d_{n,x})$ be a random compact metric space whose law coincides with that of $(\mathcal{C}_o \cap \mathbb{T}^+, n^{-1}d)$ under $P_x[\,\cdot\,|N_n^+ \neq \emptyset]$ (recall that d denotes the distance on $\mathbb{T}$). We show that the sequence $(T_{n,x}, d_{n,x})$ converges in distribution to a conditioned Brownian continuum random tree $(T_{\mathbf{e}}, d_{\mathbf{e}})$, whose contour function $\mathbf{e}$ is a Brownian excursion conditioned to reach height at least 1.

Theorem 2.3. *For every $x \geq h^*$, as $n \to \infty$,*

$$(T_{n,x}, d_{n,x}) \to (T_{\mathbf{e}}, d_{\mathbf{e}}) \tag{2.14}$$

in distribution, with respect to the Gromov-Hausdorff topology.

General scaling limits of critical Galton-Watson processes, in the spirit of Theorem 2.3, were first introduced in [Ald91]. A corresponding result for critical multi-type processes with finitely many types was established in [Mie08], and later extended to processes with a countably infinite type space in [dR17]. More recently, [CKKM24] extended the classical continuous random tree convergence to Galton-Watson trees evolving in a random environment, where each generation has a random offspring distribution with mean one and finite expected variance. In the context of critical branching diffusions, [Pow19] proves an invariance principle under the assumptions of a bounded domain, finite second moment of the offspring distribution, and an elliptic diffusion generator. However, for branching diffusions in general (unbounded) domains, analogous results are not yet available (see [Pow19, Question 1.8]).

Remark 2.4. Theorems 2.2 and 2.3 hold without any further change if the conditioning therein is changed from $P_x[\,\cdot\,|N_n^+ \neq \emptyset]$ to $P_x[\,\cdot\,|N_n \neq \emptyset]$. For the sake of brevity, we refrain from providing detailed proofs of these results.

We briefly discuss the organisation of this article. In Section 3, we collect relevant background on the GFF on regular trees along with the framework of branching processes with spines. Section 4 presents the proof of Theorem 2.1. In Section 5, we prove Theorem 2.2 and, in a dedicated subsection, establish additional results for the model conditioned on the event $\{N_n^+ \neq \emptyset\}$. Section 6 introduces an auxiliary martingale S_n and establishes scaling limit results for this martingale and some related processes. Section 7 focuses on the "height process" H_n, defined via the distance to the origin in a depth-first traversal of $\mathcal{C}_o \cap \mathbb{T}^+$, and investigates its connection to the martingale S_n. Finally, Section 8 concludes the article with the proof of Theorem 2.3, which combines topological arguments with the results from Sections 6 and 7.

3. Notation and useful results

In this section we introduce the notation used throughout the paper and recall some known facts about the level set percolation of the Gaussian free field on trees. We then briefly present the formalism of branching processes with spines and apply it to our model.

As already stated in the introduction, we use $\mathbb{T}$ to denote the $(d+1)$-regular tree, $d \geq 2$, that is an infinite tree whose every vertex has exactly $d+1$ neighbours. For two vertices $v, w \in \mathbb{T}$ we use $d(v,w)$ to denote the usual graph distance. The tree is rooted at an arbitrary fixed vertex $o \in \mathbb{T}$, $\bar{o} \in \mathbb{T}$ denotes a fixed neighbour of o, and $\mathbb{T}^+$ stands for the forward tree, see (2.2). We set $|v| = d(o,v)$ and write

$$S_n = \{v \in \mathbb{T} : |v| = n\}, \qquad S_n^+ = S_n \cap \mathbb{T}^+ \tag{3.1}$$

for the spheres with radius n centred at o. For every $v \in \mathbb{T} \setminus \{o\}$ we use $p(v)$ to denote its parent in $\mathbb{T}$, that is the only vertex on the geodesic path from v to o with $|p(v)| = |v| - 1$. We write $\mathrm{desc}(v)$ for the set of direct descendants of v, and $\mathrm{sib}(v) = \mathrm{desc}(p(v))$ for the set of its siblings, including itself. Finally, if w is an ancestor of v, that is w lies on the geodesics from o to v, we write $w \preceq v$.

Throughout the paper we use the usual notation for the asymptotic relation of two functions f and g: We will write $f(s) \sim g(s)$ as $s \to \infty$ if $\lim_{s\to\infty} f(s)/g(s) = 1$, $f(s) = o(g(s))$ as $s \to \infty$ if $\lim_{s\to\infty} |f(s)|/g(s) = 0$, and $f(s) = O(g(s))$ as $s \to \infty$ if $\limsup_{s\to\infty} |f(s)|/g(s) < \infty$. We use $c, c', c_1, \dots$ to denote finite positive constants whose value may change from place to place and which can only depend on d. The dependence of these constants on additional parameters is explicitly mentioned.

3.1. Properties of the GFF.

We consider the Gaussian free field $\varphi = (\varphi_v)_{v\in\mathbb{T}}$ which is the centred Gaussian process on $\mathbb{T}$ whose covariance function is the Green function of the simple random walk on $\mathbb{T}$,

$$E[\varphi_v \varphi_w] = g(v,w) \coloneqq \frac{1}{d+1} \mathbb{E}_v \Big[\sum_{k=0}^{\infty} 1_{X_k = w} \Big], \qquad v, w \in \mathbb{T}, \tag{3.2}$$

where $\mathbb{E}_v$ stands for the expectation with respect to the simple random walk $(X_k)_{k\geq 0}$ on $\mathbb{T}$ starting at $v \in \mathbb{T}$.

We frequently use the fact that the Gaussian free field on $\mathbb{T}$ can be viewed as a multi-type branching process with a continuous type space (see [Szn16, Section 3] and [AČ20, Section 2.1]). To this end, we define

$$\sigma_\nu^2 \coloneqq \frac{d}{d-1} \quad \text{and} \quad \sigma_Y^2 \coloneqq \frac{d+1}{d}, \tag{3.3}$$

and let $(Y_v)_{v\in\mathbb{T}}$ be a collection of independent centred Gaussian random variables on some auxiliary probability space such that $Y_o \sim \mathcal{N}(0, \sigma_\nu^2)$ and $Y_v \sim \mathcal{N}(0, \sigma_Y^2)$ for $v \neq o$. We then define another field $\widetilde{\varphi}$ on $\mathbb{T}$ by

(3.4)
(a) $\widetilde{\varphi}_o \coloneqq Y_o$,
(b) for $v \neq o$ we recursively set $\widetilde{\varphi}_v \coloneqq d^{-1} \widetilde{\varphi}_{p(v)} + Y_v$.

As explained, e.g., in [AČ20, (2.9)], the law of $(\widetilde{\varphi}_v)_{v\in\mathbb{T}}$ agrees with the law P of the Gaussian free field φ. Therefore, we will always assume that the considered Gaussian free field is constructed in this way and will not distinguish between φ, and $\widetilde{\varphi}$ from now on.

Representation (3.4) of φ can be used to give a concrete construction for the conditional probability P_x introduced in (2.1): It suffices to replace (a) in (3.4) by $\widetilde{\varphi}_o = x$. In addition, (3.4) can directly be used to construct a monotone coupling of P_x and P_y. As a consequence:

$$\text{If } x < y\text{, then } P_y \text{ stochastically dominates } P_x, \tag{3.5}$$

that is $E_x[f(\varphi)] \le E_y[f(\varphi)]$ for every bounded increasing function $f : \mathbb{R}^{\mathbb{T}} \to \mathbb{R}$.

From the construction (3.4) it follows that the GFF on $\mathbb{T}$ can be viewed as a multi-type branching process where the type corresponds to the value of the field φ. The type of the initial individual o of this branching process is distributed as Y_o. Every individual v in this branching process then has d descendants ($d + 1$ if $v = o$) whose types are independently given by $d^{-1}\varphi_v + Y$, with $Y \sim N(0, \sigma_Y^2)$. The branching process point of view can be adapted to the connected component $\mathcal{C}_o^h$ (defined in (2.4)) by considering the same multi-type branching process but killing immediately all individuals with type smaller than h (and not allowing them to have descendants themselves).

We now recall in more detail the spectral machinery introduced in [Szn16] in order to characterise the critical value h^*. Let ν be a centred Gaussian measure on $\mathbb{R}$ with variance σ_ν^2 (as defined in (3.3)), and let Y be a centred Gaussian random variable with variance σ_Y^2. The expectation with respect to this random variable is denoted by E_Y. We consider the Hilbert space $L^2(\nu) := L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \nu)$, and for every $h \in \mathbb{R}$, define the operator L_h acting on $L^2(\nu)$ by

$$\begin{aligned} L_h[f](x) &:= 1_{[h,\infty)}(x)\, d\, E_Y\Big[1_{[h,\infty)}\Big(Y + \frac{x}{d}\Big) f\Big(Y + \frac{x}{d}\Big)\Big] \\ &= 1_{[h,\infty)}(x)\, d \int_{[h,\infty)} f(y)\rho_Y\Big(y - \frac{x}{d}\Big)\, \mathrm{d}y, \end{aligned} \tag{3.6}$$

where ρ_Y denotes the density of Y. From the branching process construction (3.4) of φ, it follows that

$$L_h[f](x) = E_x\Big[\sum_{v\in N_1^{h,+}} f(\varphi_v)\Big], \tag{3.7}$$

where $N_n^{h,+}$ is defined in (2.8). Iterating this expression one also obtains

$$L_h^n[f](x) = E_x\Big[\sum_{v\in N_n^{h,+}} f(\varphi_v)\Big], \qquad n \in \mathbb{N}. \tag{3.8}$$

Finally, we let λ_h stand for the operator norm of L_h in $L^2(\nu)$,

$$\lambda_h := \|L_h\|_{L^2(\nu)\to L^2(\nu)}. \tag{3.9}$$

The following proposition summarises some known properties of the operator L_h as well as the connection between L_h and the critical height h^*.

Proposition 3.1 ([Szn16] Propositions 3.1, 3.3, Corollary 4.5)**.** *For all $h \in \mathbb{R}$, L_h is a self-adjoint non-negative Hilbert-Schmidt operator on $L^2(\nu)$, λ_h is a simple eigenvalue of L_h, and there exists a unique $\chi_h \ge 0$ with unit $L^2(\nu)$-norm, which is continuous, strictly positive on $[h, \infty)$, vanishing on $(-\infty, h)$, and such that*

$$L_h[\chi_h] = \lambda_h \chi_h. \tag{3.10}$$

Additionally, the map $h \mapsto \lambda_h$ is a decreasing homeomorphism from $\mathbb{R}$ to $(0, d)$ and h^* is the unique value in $\mathbb{R}$ such that $\lambda_{h^*} = 1$. Finally, for every $d \geq 2$,

$$0 < h^* < \infty. \tag{3.11}$$

Combining Proposition 3.1 with (3.8) gives that for every $n \in \mathbb{N}$,

$$E_x\Big[\sum_{w \in N_n^{h,+}} \chi_h(\varphi_w)\Big] = \lambda_h^n \chi_h(x). \tag{3.12}$$

We will require estimates on the norms of $L_h[f]$, which follow from the hypercontractivity of the Ornstein-Uhlenbeck semigroup. For part (a) of the following proposition, we refer to (3.14) in [Szn16] or (2.14) in [AČ20]. Part (b) then follows directly from part (a) in combination with generalized Hölder's inequality.

Proposition 3.2. *(a) For every $f \in L^2(\nu)$, $h \in \mathbb{R}$, $1 < p < \infty$ and $q \leq (p-1)d^2 + 1$,*

$$\|L_h[f]\|_{L^q(\nu)} \leq d\|f\|_{L^p(\nu)}. \tag{3.13}$$

(b) For every $1 \leq k \leq (d^2+1)/2$ and $f_1, \dots, f_k \in L^2(\nu)$

$$\Big\|\prod_{i=1}^k L_h[f_i]\Big\|_{L^2(\nu)} \leq d^k \prod_{i=1}^k \|f_i\|_{L^2(\nu)}. \tag{3.14}$$

The next proposition recalls known properties of the critical component $\mathcal{C}_o$ from [ČL25].

Proposition 3.3 (Theorem 2.1 and Theorem 2.3 in [ČL25])**.** *There is $C \in (0, \infty)$ such that for every $x \in \mathbb{R}$, as $n \to \infty$,*

$$P_x[|\mathcal{C}_o \cap \mathbb{T}^+| > n] = C\chi(x)n^{-1/2}(1 + o(1)). \tag{3.15}$$

In particular, for every $x \in \mathbb{R}$,

$$P_x[|\mathcal{C}_o \cap \mathbb{T}^+| = \infty] = 0. \tag{3.16}$$

We will also need the following three properties of χ, the first one is proved in Remark 2.5 of [ČL25], the second is a direct consequence of Proposition 3.1 in [AČ20], and the last one is proved in Appendix A:

$$x \mapsto \chi(x) \text{ is non-decreasing,} \tag{3.17}$$

$$c_1 x \leq \chi(x) \leq c_2 x \quad \text{for all } x \geq h^* \text{ and some } c_1, c_2 \in (0, \infty), \tag{3.18}$$

$$x \mapsto \chi(x) \text{ is Lipschitz on } [h^*, \infty). \tag{3.19}$$

3.2. **Branching processes with spines.** We now recall the machinery of branching processes with spines which is frequently used in the theory of branching processes, and specialize it to our model, in order to study the critical component $\mathcal{C}_o$. We then state many-to-few formulas that express certain moments for the original branching process in terms of the dynamics along the spines, see Proposition 3.5 below. Later, in Section 5.1, we will see that the processes with spines can be used to describe the distribution of $\mathcal{C}_o$ conditioned on being infinite. The content of this section is mostly based on [HR17].

The main idea of the machinery is to designate several lines of descent in the branching process, called spines. These spines are then used to introduce a certain change of measure, under which the vertices on the spine exhibit modified branching behaviour, while the non-spine vertices behave as in the original process.

For the construction, we need some notation. For $x \in \mathbb{R}$ and $k \in \mathbb{N}$, we first introduce measures P_x^k under which the behaviour of the field is the same as under P_x, but in addition there are k distinguished spines. Formally, the measure P_x^k is the distribution of a $(\mathbb{R} \cup \{\dagger\}) \times \{0, 1, \dots, k\}$-valued stochastic process $(\varphi_v, l_v)_{v \in \mathbb{T}^+}$. This process assigns to every $v \in \mathbb{T}^+$ a field value $\varphi_v \in \mathbb{R} \cup \{\dagger\}$ (where $\dagger$ is a cemetery state) and a number $l_v \in \{0, \dots, k\}$ that represents the number of spine marks on v. Under P_x^k , the law of $(\varphi_v)_{v \in \mathbb{T}^+}$ is a straightforward modification of the original measure P_x, the only difference is that we set $\varphi_v = \dagger$ for every $v \notin \mathcal{C}_o$. The random variables $(l_v)_{v \in \mathbb{T}^+}$ are independent of the field values $(\varphi_v)_{v \in \mathbb{T}^+}$. The root node o has exactly k marks, $l_o = k$. The remaining random variables $(l_v)_{v \neq o}$ are constructed recursively as follows: If a node $v \in \mathbb{T}^+$ carries j marks, then each of its j marks 'moves' to one of its d direct descendants independently uniformly at random. (Note that nodes in the cemetery state can carry marks.) As consequence, under P_x^k, in every generation n, there are exactly k marks present, that is $\sum_{v \in S_n^+} l_v = k$. We use P_x^k also for the corresponding expectations.

For $i = 1, \dots, k$, we denote by σ_n^i the node that carries the i-th mark in generation n (that is $|\sigma_n^i| = n$), and set $\xi_n^i = \varphi_{\sigma_n^i}$ to be its type. We also define $\operatorname{skel}(n) = \{v \in \mathbb{T}^+ : |v| \leq n, l_v \geq 1\}$ to be the set of nodes of generation at most n having at least one mark. We let $\mathcal{F}_n$ stand for the natural filtration of the branching process, and $\mathcal{F}_n^k$ for the filtration containing in addition the information about the k spine marks,

$$\mathcal{F}_n = \sigma(\varphi_v : v \in \mathbb{T}^+, |v| \leq n) \quad \text{and} \quad \mathcal{F}_n^k = \sigma(\varphi_v, l_v : v \in \mathbb{T}^+, |v| \leq n). \tag{3.20}$$

Any $f : \mathbb{R} \to \mathbb{R}$ is extended to $\mathbb{R} \cup \{\dagger\}$ by setting $f(\dagger) = 0$. Then, by definition of P_x^k,

$$E_x\Big[\sum_{v \in N_n^+} f(\varphi_v)\Big] = P_x^k\Big[\sum_{v \in S_n^+} f(\varphi_v)\Big]. \tag{3.21}$$

We now define another measure Q_x^k, where the nodes without a spine mark behave as under P_x^k but the nodes with a mark have a modified branching behaviour: Under Q_x^k the movement of the marks, and thus the distribution of $(l_v)_{v \in \mathbb{T}^+}$, is exactly the same as under P_x^k: If a node v carries k marks, each of the marks is given to one of its d direct descendants independently uniformly at random. To describe the distribution of the field φ under Q_x^k, we first define a transition kernel (recall ρ_Y from (3.6))

$$\mathcal{K}(x, \mathrm{d}y) = d \frac{\chi(y)}{\chi(x)} \rho_Y\Big(y - \frac{x}{d}\Big)\, \mathrm{d}y, \qquad x \geq h^*, y \in \mathbb{R}. \tag{3.22}$$

Note that, by (3.6) and Proposition 3.1, for every $x \geq h^*$, $\mathcal{K}(x, \cdot)$ is a probability measure with support $[h^*, \infty)$. Conditionally on the marks $(l_v)_{v \in \mathbb{T}^+}$, the field $(\varphi_v)_{v \in \mathbb{T}^+}$ under Q_x^k is recursively constructed by

(3.23)
- (a) $\varphi_o := x$,
- (b) If $v \neq o$ and $l_v = 0$, then $\varphi_v = d^{-1}\varphi_{p(v)} + Y_v$ (as under P_x^k).
- (c) If $v \neq o$ and $l_v \geq 1$, then φ_v is $\mathcal{K}(\varphi_{p(v)}, \cdot)$-distributed, independently of previous randomness.

To simplify notation, we write Q_x instead of Q_x^1; in this case we also set $\sigma_n = \sigma_n^1$ and $\xi_n = \xi_n^1$.

Note that, unlike under P_x^k, under the measure Q_x^k the nodes in the cemetery state cannot carry any mark. Consequently, Q_x^k-a.s., there are nodes not in the cemetery state in every generation.

By construction, under Q_x^k, the process $(\xi_n^i)_{n \in \mathbb{N}}$ recording the value of the field along the i-th spine is a Markov chain with transition kernel $\mathcal{K}$. This chain never enters the cemetery state. The following lemma determines its invariant distribution.

Lemma 3.4. *The Markov chain $(\xi_n)_{n\in\mathbb{N}}$ with the transition kernel $\mathcal{K}$ has a unique invariant distribution π given by*

$$\pi(\mathrm{d}x) = \chi(x)^2\nu(\mathrm{d}x). \tag{3.24}$$

Proof. We first show that π is invariant for $\mathcal{K}$. Comparing (3.6) and (3.22) yields that $\mathcal{K}(x,A) = L[1_A\chi](x)/\chi(x)$ for every $x \ge h^*$. Therefore, writing the action of $\mathcal{K}$ on π as inner product on $L^2(\nu)$, using that L is self-adjoint and χ is its eigenfunction with eigenvalue 1 (see Proposition 3.1),

$$(\pi\mathcal{K})(A) = \langle \mathcal{K}(\cdot,A),\chi^2\rangle_\nu = \langle L[1_A\chi],\chi\rangle_\nu = \langle 1_A\chi, L[\chi]\rangle_\nu = \langle 1_A\chi,\chi\rangle_\nu = \pi(A), \tag{3.25}$$

which shows that π is an invariant measure. The uniqueness follows from the irreducibility of $(\xi_n)_{n\in\mathbb{N}}$. □

Next, we state several moment formulas that are frequently used throughout the paper. Such formulas are well understood in the theory of branching processes, see, e.g., [HR17] and references therein. The proof of the following proposition is based on Lemma 8 of that paper and can be found in Appendix B.

Proposition 3.5. *For all functions $f,g : [h^*,\infty) \to \mathbb{R}$ for which the expectations below are well defined,*

$$E_x\Big[\sum_{v\in N_n^+} f(\varphi_v)\Big] = Q_x\Big[f(\xi_n)\frac{\chi(x)}{\chi(\xi_n)}\Big], \tag{3.26}$$

$$\begin{aligned} E_x\Big[\sum_{v,w\in N_n^+} f(\varphi_v)g(\varphi_w)\Big] &= \chi(x)\frac{d-1}{d}\sum_{k=0}^{n-1} Q_x\Big[\chi(\xi_k)Q_{\xi_k}\Big[\frac{f(\xi_{n-k})}{\chi(\xi_{n-k})}\Big]Q_{\xi_k}\Big[\frac{g(\xi_{n-k})}{\chi(\xi_{n-k})}\Big]\Big] \\ &\quad + \chi(x)Q_x\Big[\frac{f(\xi_n)g(\xi_n)}{\chi(\xi_n)}\Big]. \end{aligned} \tag{3.27}$$

3.3. **Asymptotic behaviour of the moments.** The main result of this section is Proposition 3.8 giving precise asymptotic estimates on the quantities appearing in (3.26), (3.27). To prove them, we could, in principle, use this proposition together with the known results on the convergence of Markov chains. However, it is easier, and for our purposes slightly more practical, to use formula (3.8) together with L^2-estimates on the operator L.

Since L is a self-adjoint Hilbert-Schmidt operator (see Proposition 3.1), $L^2(\nu)$ has an orthonormal basis consisting of the eigenfunctions $\{e_k\}_{k\ge1}$ of L corresponding to the eigenvalues $\{\lambda_k\}_{k\ge1}$. By Proposition 3.1 we may assume that $1 = \lambda_1 > |\lambda_2| \ge |\lambda_3| \ge \dots$, and $e_1 = \chi$. We set $\gamma = |\lambda_2| \in (0,1)$. By decomposing $f \in L^2(\nu)$ as

$$f = \sum_{k\ge1}\langle e_k,f\rangle e_k = \langle\chi,f\rangle\chi + \sum_{k\ge2}\langle e_k,f\rangle e_k =: \langle\chi,f\rangle\chi + \beta[f], \tag{3.28}$$

for every $n\in\mathbb{N}$, it holds

$$L^n[f] = \sum_{k\ge1}\lambda_k^n\langle e_k,f\rangle e_k = \langle\chi,f\rangle\chi + \sum_{k\ge2}\lambda_k^n\langle e_k,f\rangle e_k = \langle\chi,f\rangle\chi + L^n[\beta[f]], \tag{3.29}$$

with

$$\|L^n[\beta[f]]\| \le \gamma^n\|\beta[f]\| \le \gamma^n\|f\|. \tag{3.30}$$

The following simple lemma will later be used to deduce the pointwise convergence from the $L^2(\nu)$-one.

Lemma 3.6. *There is a function $q : [h^*, \infty) \to [0, \infty)$ such that for every $x \ge h^*$ and $f \in L^2(\nu)$*

$$|L[f](x)| \le \|f\| q(x). \tag{3.31}$$

Proof. Using the definition (3.6) of L and the Cauchy–Schwarz inequality,

$$\begin{aligned} |L[f](x)| &\le d \int_{\mathbb{R}} |f(y)| \rho_Y\Big(y - \frac{x}{d}\Big)\,\mathrm{d}y = d \int_{\mathbb{R}} |f(y)| \frac{\rho_Y(y - x/d)}{\rho_\nu(y)} \nu(\mathrm{d}y) \\ &\le d \|f\|_{L^2(\nu)} \Big\| \frac{\rho_Y(\cdot - x/d)}{\rho_\nu(\cdot)} \Big\|_{L^2(\nu)} =: \|f\|_{L^2(\nu)} q(x). \end{aligned} \tag{3.32}$$

By (3.3), $\sigma_\nu > \sigma_Y$. Using this one can easily check by a direct computation that $q(x) < \infty$ for every $x \ge h^*$. ☐

Remark 3.7. The same computation also shows that $q(x) \in L^2(\nu)$, but we will not use this fact.

We can now give the asymptotic estimates on moments appearing in Proposition 3.5.

Proposition 3.8. *(a) For every $f \in L^2(\nu)$, $x \ge h^*$, and $n \in \mathbb{N}$,*

$$E_x\Big[\sum_{v \in N_n^+} f(\varphi_v)\Big] = \chi(x)\langle \chi, f\rangle + \varepsilon_n^f(x), \tag{3.33}$$

where the error term ε_n^f satisfies (with q as in Lemma 3.6)

$$\|\varepsilon_n^f\| \le \gamma^n \|f\| \qquad \text{and} \qquad |\varepsilon_n^f(x)| \le \gamma^{n-1}\|f\| q(x). \tag{3.34}$$

(b) For every $f, g \in L^2(\nu)$, $x \ge h^*$, and $n \in \mathbb{N}$

$$\begin{aligned} E_x\Big[\sum_{v,w \in N_n^+} f(\varphi_v) g(\varphi_w)\Big] &- E_x\Big[\sum_{v \in N_n^+} f(\varphi_v) g(\varphi_v)\Big] \\ &= \chi(x) \frac{d-1}{d} \langle \chi, f\rangle\langle \chi, g\rangle\langle \chi^2, \chi\rangle n + \varepsilon_n^{f,g}(x), \end{aligned} \tag{3.35}$$

where the error term $\varepsilon_n^{f,g}$ satisfies

$$\|\varepsilon_n^{f,g}\| \le C\|f\|\|g\| \qquad \text{and} \qquad \varepsilon_n^{f,g}(x) \le \|f\|\|g\|\tilde{q}(x) \tag{3.36}$$

with $\tilde{q}(x) \le C(q(x) + \chi(x)) < \infty$.

Proof. (a) By (3.8), the left-hand side of (3.33) equals $L^n[f]$. Therefore, by (3.29), $\varepsilon_n^f = L^n[\beta[f]]$. (3.30) then implies the first claim in (3.34). The second one follows from the first one and Lemma 3.6.

(b) Combining (3.27) with (3.26) and (3.8), we obtain that

$$E_x\Big[\sum_{v,w \in N_n^+} f(\varphi_v) g(\varphi_w)\Big] = \frac{d-1}{d} \sum_{k=0}^{n-1} L^k\Big[L^{n-k}[f] L^{n-k}[g]\Big](x) + L^n[fg](x). \tag{3.37}$$

By (3.33), $L^{n-k}[f]L^{n-k}[g] = (\langle \chi, f\rangle \chi + \varepsilon_{n-k}^f)(\langle \chi, g\rangle\chi + \varepsilon_{n-k}^g)$ and thus, again by (3.33),

$$\begin{aligned} L^k\Big[L^{n-k}[f]L^{n-k}[g]\Big] &= \langle \chi, L^{n-k}[f]L^{n-k}[g]\rangle \chi + \varepsilon_k^{L^{n-k}[f]L^{n-k}[g]} \\ &= \Big(\langle \chi, f\rangle\langle \chi, g\rangle\langle \chi, \chi^2\rangle + \langle \chi, f\rangle\langle \chi, \chi\varepsilon_{n-k}^g\rangle + \langle \chi, g\rangle\langle \chi, \chi\varepsilon_{n-k}^f\rangle \\ &\quad + \langle \chi, \varepsilon_{n-k}^f \varepsilon_{n-k}^g\rangle\Big)\chi + \varepsilon_k^{L^{n-k}[f]L^{n-k}[g]}. \end{aligned} \tag{3.38}$$

Note also that the $L^n[fg](x)$ summand in (3.37) equals $E_x[\sum_{v\in N_n^+} f(\varphi_v)g(\varphi_v)]$. Therefore, the error term in (3.35) satisfies

$$\varepsilon_n^{f,g}(x) = \frac{d-1}{d}\sum_{k=0}^{n-1}\Big(\big(\langle\chi,f\rangle\langle\chi,\chi\varepsilon^g_{n-k}\rangle + \langle\chi,g\rangle\langle\chi,\chi\varepsilon^f_{n-k}\rangle + \langle\chi,\varepsilon^f_{n-k}\varepsilon^g_{n-k}\rangle\big)\chi(x) + \varepsilon_k^{L^{n-k}[f]L^{n-k}[g]}(x)\Big). \tag{3.39}$$

To bound this expression, note that $|\langle\chi,f\rangle| \le \|f\|$ and $|\langle\chi,\chi\varepsilon^g_{n-k}\rangle| \le \|\chi^2\|\|\varepsilon^g_{n-k}\| \le C\gamma^{n-k}\|g\|$, by part (a). Therefore, the first summand in (3.39) satisfies $|\langle\chi,f\rangle\langle\chi,\chi\varepsilon^g_{n-k}\rangle| \le C\gamma^{n-k}\|f\|\|g\|$. Analogously, $|\langle\chi,g\rangle\langle\chi,\chi\varepsilon^f_{n-k}\rangle| \le C\gamma^{n-k}\|f\|\|g\|$. To estimate the third summand, we observe that by (3.33) $\varepsilon^f_{n-k} = L^{n-k}[\beta[f]]$. Therefore, using Proposition 3.2(b), $\|\varepsilon^f_{n-k}\varepsilon^g_{n-k}\| \le d^2\|\varepsilon^f_{n-k-1}\|\|\varepsilon^g_{n-k-1}\|$ and thus

$$\langle\chi,\varepsilon^f_{n-k}\varepsilon^g_{n-k}\rangle \le d^2\|\varepsilon^f_{n-k-1}\|\|\varepsilon^g_{n-k-1}\| \le d^2\gamma^{2(n-k-1)}\|f\|\|g\|. \tag{3.40}$$

Finally, for the fourth summand, using again Proposition 3.2(b) and part (a), since the operator norm of L equals 1,

$$\begin{aligned}\|\varepsilon_k^{L^{n-k}[f]L^{n-k}[g]}\| &\le \gamma^k\|L^{n-k}[f]L^{n-k}[g]\|\\ &\le d^2\gamma^k\|L^{n-k-1}[f]\|\|L^{n-k-1}[g]\| \le C\gamma^k\|f\|\|g\|.\end{aligned} \tag{3.41}$$

Taking the norm in (3.39) and making use of the above estimates yields

$$\|\varepsilon_n^{f,g}\| \le C\|f\|\|g\|\sum_{k=0}^{n-1}(\gamma^{n-k} + \gamma^{2(n-k-1)} + \gamma^k) < C\|f\|\|g\|, \tag{3.42}$$

since $\gamma < 1$. This establishes the bound on $\|\varepsilon_n^{f,g}\|$.

To establish the claimed pointwise bound for $|\varepsilon_n^{f,g}(x)|$, observe that by part (a), we have $\big|\varepsilon_k^{L^{n-k}[f]L^{n-k}[g]}(x)\big| \le \gamma^k\|L^{n-k}[f]L^{n-k}[g]\|q(x) \le C\gamma^k\|f\|\|g\|q(x)$. Combining this with (3.39) and the previously derived bounds on the inner products appearing there, we easily conclude. ☐

4. Proof of Theorem 2.1

In this section we prove Theorem 2.1 which describes the tail behaviour of the diameter of the critical cluster. We also show several estimates that will later be used in the proof of Theorem 2.2.

Let $\mathbb{F}$ be the set of all non-increasing functions $f : \mathbb{R} \to [0,1)$. For every $f \in \mathbb{F}$ and $n \in \mathbb{N}$, we introduce

$$u_n^f(x) := \begin{cases} E_x\big[1_{\{N_n^+\neq\emptyset\}}\big(1 - \prod_{u\in N_n^+} f(\varphi_u)\big)\big], & \text{for } x \ge h^*,\\ 0, & \text{for } x < h^*.\end{cases} \tag{4.1}$$

Note that $u_n^f(x) \in [0,1]$ and by (3.5) it is increasing in x. Taking $f \equiv 0$,

$$u_n^0(x) = P_x[N_n^+ \neq \emptyset], \tag{4.2}$$

which explains the relevance of this definition for the proof of Theorem 2.1.

We will use the following inequality, which often allows us to consider the special case $f \equiv 0$ only: For every $f \in \mathbb{F}$

$$u_n^0(x)(1 - f(h^*)) \le u_n^f(x) \le u_n^0(x). \tag{4.3}$$

Indeed, the second inequality follows directly from the definition of u_n^f, since $f \geq 0$. To see the first one, note that since f is non-increasing, $\prod_{v \in N_n^+} f(\varphi_v) \leq f(h^*)$ when $N_n^+ \neq \emptyset$, and thus $E_x[1_{\{N_n^+ \neq \emptyset\}}(1 - \prod_{v \in N_n^+} f(\varphi_v))] \geq E_x[1_{\{N_n^+ \neq \emptyset\}}(1 - f(h^*))] = (1 - f(h^*))u_n^0(x)$.

Similarly to (3.28) we decompose u_n^f as

$$u_n^f = \langle \chi, u_n^f \rangle \chi + \beta[u_n^f] \tag{4.4}$$

and define $a_n^f \coloneqq \langle \chi, u_n^f \rangle$ and $b_n^f \coloneqq \|\beta[u_n^f]\|$, so that

$$\|u_n^f\|^2 = (a_n^f)^2 + (b_n^f)^2. \tag{4.5}$$

Due to (4.2), in order to show Theorem 2.1, we need precise asymptotic estimates on a_n^f and b_n^f. These will be proved step by step in the following several lemmas. The theorem is then shown at the end of the section.

By Proposition 3.3, $P_x[|\mathcal{C}_o \cap \mathbb{T}^+| = \infty] = 0$. Therefore, $P_x[N_n^+ \neq \emptyset] \to 0$ as $n \to \infty$. As a consequence, using also (4.3), for every $f \in \mathbb{F}$,

$$u_n^f(x) \to 0 \text{ as } n \to \infty, \text{ pointwise and thus in } L^2(\nu). \tag{4.6}$$

As a consequence, for every $f \in \mathbb{F}$,

$$\lim_{n\to\infty} a_n^f = \lim_{n\to\infty} b_n^f = 0. \tag{4.7}$$

The following lemma provides a basic recursive relation for u_n^f, based on the branching process representation. This relation will be important to obtain a more precise description of the asymptotic behaviour of a_n^f and b_n^f as $n \to \infty$.

Lemma 4.1. *For every non-increasing $f : \mathbb{R} \to [0, 1)$ and $n \in \mathbb{N}$,*

$$u_n^f(x) = 1 - \Big(1 - \frac{1}{d} L[u_{n-1}^f](x)\Big)^d \quad \text{for } x \in \mathbb{R}. \tag{4.8}$$

As a consequence,

$$u_n^f(x) = L[u_{n-1}^f](x) - g(L[u_{n-1}^f](x)), \tag{4.9}$$

with $g(x) = 1 - x - (1 - d^{-1}x)^d$. This function satisfies, for $c_1, c_2 \in (0, \infty)$,

$$c_1 x^2 \leq g(x) \leq c_2 x^2 \quad \text{for all } x \in [0, 1]. \tag{4.10}$$

Proof. For $x < h^*$, both sides of (4.8) are trivially zero by definition of u_n^f and L. For $x \geq h^*$, by the branching process construction of φ, recalling $\mathcal{F}_n$ from (3.20), for $n_0 < n$,

$$\begin{aligned} E_x\Big[\prod_{v \in N_n^+} f(\varphi_v)\Big|\mathcal{F}_{n_0}\Big] &= \prod_{v \in N_{n_0}^+} E_{\varphi_v}\Big[\prod_{w \in N_{n-n_0}^+} f(\varphi_w)\Big] \\ &= \prod_{v \in N_{n_0}^+} \bigg(1 - E_{\varphi_v}\Big[1 - \prod_{w \in N_{n-n_0}^+} f(\varphi_w)\Big]\bigg) \\ &= \prod_{v \in N_{n_0}^+} \big(1 - u_{n-n_0}^f(\varphi_v)\big). \end{aligned} \tag{4.11}$$

Therefore, setting $n_0 = 1$,

$$u_n^f(x) = E_x\Big[1 - E_x\Big[\prod_{v \in N_n^+} f(\varphi_v)\Big|\mathcal{F}_1\Big]\Big] = 1 - E_x\Big[\prod_{v \in N_1^+} \big(1 - u_{n-1}^f(\varphi_v)\big)\Big]. \tag{4.12}$$

Using the conditional independence of the $(\varphi_v : v \in S_1^+)$ given $\varphi_o = x$, and the fact that $u_n^f(x) = 0$ for $x < h^*$,

$$E_x\Big[\prod_{v\in N_1^+}\big(1-u_{n-1}^f(\varphi_v)\big)\Big] = E_x\Big[\prod_{v\in S_1^+}\big(1-u_{n-1}^f(\varphi_v)\big)\Big] = \prod_{v\in S_1^+} E_x[1-u_{n-1}^f(\varphi_v)]. \tag{4.13}$$

Using (3.7), $E_x[u_{n-1}^f(\varphi_v)] = d^{-1}L[u_{n-1}^f](x)$ for every $v \in S_1^+$. Together with (4.12) and (4.13), this proves (4.8) and (4.9). Inequality (4.10) is proved in [ČL25, Lemma 5.3]. □

The following lemma provides a rough lower bound for u_n^0 and a_n^0.

Lemma 4.2. *There exists a constant $c > 0$ such that for all $n \ge 1$*

$$\|u_n^0\| \ge a_n^0 \ge cn^{-1}. \tag{4.14}$$

Proof. The statement will follow if we show

$$0 \le g(x)n^{-1} \le u_n^0(x) \qquad \text{for all } n \ge 1 \tag{4.15}$$

for some non-trivial function $g : \mathbb{R} \to [0,1]$. To show (4.15) we use the estimate on the volume of $\mathcal{C}_o$ from Proposition 3.3. By this proposition and (3.5) there is a positive constant c_1 such that, for all $m \ge 1$ and $x \ge h^*$,

$$P_x[|\mathcal{C}_o \cap \mathbb{T}^+| > m] \ge c_1 m^{-1/2}. \tag{4.16}$$

Since $|\mathcal{C}_o \cap \mathbb{T}^+| = \sum_{k\ge 0}|N_k^+|$, this implies for all $m, n \ge 1$

$$c_1 m^{-1/2} \le P_x\Big[\sum_{k\ge 0}|N_k^+| > m, N_n^+ = \emptyset\Big] + P_x\Big[\sum_{k\ge 0}|N_k^+| > m, N_n^+ \ne \emptyset\Big] \tag{4.17}$$

$$\le P_x\Big[\sum_{k=0}^{n-1}|N_k^+| > m\Big] + P_x[N_n^+ \ne \emptyset]. \tag{4.18}$$

By the Markov inequality, $P_x[\sum_{k=0}^{n-1}|N_k^+| > m] \le \frac{1}{m}\sum_{k=0}^{n-1}E_x[|N_k^+|]$, where, by Proposition 3.8, $E_x[|N_k^+|] \le c_2(x)$ for some x-dependent constant $c_2(x) > 0$. Thus,

$$c_1 m^{-1/2} \le c_2(x)\frac{n}{m} + P_x[N_n^+ \ne \emptyset]. \tag{4.19}$$

Recalling (4.2) and choosing $m = \lfloor \delta^2 n^2 \rfloor$ with $\delta = 2c_1c_2(x)^{-1}$, it follows that

$$u_n^0(x) \ge \frac{c_1^2}{2c_2(x)}n^{-1} \quad \text{for all } n \ge 1, \tag{4.20}$$

showing (4.15) and thus the lemma. □

The next lemma gives upper bounds on a_n^f and b_n^f.

Lemma 4.3. *There is $c < \infty$ such that for every $f \in \mathbb{F}$ and $n \ge 1$*

$$a_n^f \le cn^{-1} \qquad \text{and} \qquad b_n^f \le cn^{-2}. \tag{4.21}$$

Proof. To prove the first statement in (4.21) we recall the recursive relation (4.9) and project it on $\operatorname{span}\{\chi\}$. Using $\langle \chi, L[u_n^f]\rangle = \langle \chi, u_n^f\rangle = a_n^f$, and the lower bound on g from (4.10), this yields

$$a_{n+1}^f = \langle \chi, u_{n+1}^f\rangle = \langle \chi, L[u_n^f]\rangle - \langle \chi, g(L[u_n^f])\rangle \le a_n^f - c_1\langle \chi, L[u_n^f]^2\rangle. \tag{4.22}$$

By Proposition 3.1 and (3.17), $\chi(x) \ge c > 0$ for all $x \in [h^*, \infty)$. Hence, $\langle \chi, L[u_n^f]^2\rangle \ge c\langle 1, L[u_n^f]^2\rangle = c\|L[u_n^f]\|^2 \ge c(a_n^f)^2$. Applied to (4.22), this gives

$$a_{n+1}^f \le a_n^f - c(a_n^f)^2. \tag{4.23}$$

This implies that a_n^f is decreasing in n and, after rearranging, also

$$c \le \frac{a_n^f - a_{n+1}^f}{(a_n^f)^2} \le \frac{a_n^f - a_{n+1}^f}{a_{n+1}^f a_n^f} = \frac{1}{a_{n+1}^f} - \frac{1}{a_n^f}. \tag{4.24}$$

Summing this over n running from 0 to $n-1$ yields

$$cn \le \sum_{k=0}^{n-1} \Big(\frac{1}{a_{k+1}^f} - \frac{1}{a_k^f} \Big) = \frac{1}{a_n^f} - \frac{1}{a_0^f} \le \frac{1}{a_n^f}, \tag{4.25}$$

proving the first part of (4.21).

To prove the second part we project (4.9) onto the orthogonal complement of span $\{\chi\}$ and take norms. With the triangle inequality, this gives

$$b_{n+1}^f \le \|\beta[L[u_n^f]]\| + \|\beta[g(L[u_n^f])]\|. \tag{4.26}$$

By (3.30), $\|\beta[L[u_n^f]]\| = \|L[\beta[u_n^f]]\| \le \gamma \|\beta[u_n^f]\| = \gamma b_n^f$. By the upper bound of g from (4.10) and Proposition 3.2, since β is a projection, $\|\beta[g(L[u_n^f])]\| \le \|g(L[u_n^f])\| \le c_2 \|L[u_n^f]^2\| \le c\|u_n^f\|^2$. Applied to (4.26), this gives

$$b_{n+1}^f \le \gamma b_n^f + c\|u_n^f\|^2 = \gamma b_n^f + c(a_n^f)^2 + c(b_n^f)^2. \tag{4.27}$$

Taking now $f = 0$, since $b_n^0 \to 0$ as $n \to \infty$, there is $\gamma' \in (\gamma, 1)$ and $n' < \infty$ such that $\gamma b_n^0 + c(b_n^0)^2 \le \gamma' b_n^0$ for all $n \ge n'$. Together with the first part of (4.21), this implies

$$b_{n+1}^0 \le \gamma' b_n^0 + cn^{-2} \quad \text{for } n \ge n'. \tag{4.28}$$

Applying this recursively yields

$$b_{n+1}^0 \le \gamma'\big(\gamma' b_{n-1}^0 + c(n-1)^{-2}\big) + cn^{-2} \le \cdots \le (\gamma')^{n+1-n'} b_{n'}^0 + c \sum_{k=0}^{n-n'} (\gamma')^k (n-k)^{-2}. \tag{4.29}$$

For $n \ge 2n'$, by splitting the sum at $k = n/2$, using also that $b_n^0 \le 1$ for all n, this can be bounded by $c(\gamma')^{n/2} + c(n/2)^{-2} \le cn^{-2}$, proving the second half of (4.21) for $f = 0$.

For a general $f \in \mathbb{F}$, by (4.3) and the already proven statements, $\|u_n^f\|^2 \le \|u_n^0\|^2 = (a_n^0)^2 + (b_n^0)^2 \le cn^{-2}$. Inserting this into (4.27) gives

$$b_{n+1}^f \le \gamma b_n^f + cn^{-2}, \tag{4.30}$$

which looks like (4.28). The second part of (4.21) for general $f \in \mathbb{F}$ then follows by the same arguments as for $f = 0$. □

When $f \in \mathbb{F}$ is fixed, Lemmas 4.2 and 4.3 show that $\lim_{n\to\infty} u_n^f/\|u_n^f\| = \chi$ in $L^2(\nu)$ and $c < n\|u_n^f\| < c'$. In the special case $f = 0$, this is almost sufficient to prove Theorem 2.1, we only need to improve the estimate on $\|u_n^f\|$. In contrast, in the proof of Theorem 2.2 we will need to consider functions f varying with n. There our estimates are not sufficient, since the lower bound in Lemma 4.2 holds only for $f = 0$ and cannot be true uniformly over $f \in \mathbb{F}$. We now provide tools allowing us to deal with this case as well.

As it turns out, it is enough to prove the uniformity over a certain family of non-increasing functions. Specifically, let

$$\hat{\mathbb{F}} = \{f_\lambda : \lambda \in [0, \infty)\} \subset \mathbb{F}, \tag{4.31}$$

where

$$f_0 \equiv 0 \quad \text{and} \quad f_\lambda(x) := \exp\Big(- \frac{\chi(x)}{\lambda} \Big) \quad \text{for } \lambda > 0. \tag{4.32}$$

To simplify the notation, we define (recall (4.1), (4.4))

$$u_n^\lambda = u_n^{f_\lambda}, \qquad a_n^\lambda = a_n^{f_\lambda} = \langle \chi, u_n^\lambda \rangle, \qquad b_n^\lambda = b_n^{f_\lambda} = \|\beta[u_n^\lambda]\|. \tag{4.33}$$

The first preliminary step in proving the uniformity over $\hat{\mathbb{F}}$ is the following lemma. In the special case $\lambda = 0$, this already follows from Lemmas 4.2 and 4.3.

Lemma 4.4. *There is a constant $c < \infty$ so that*

$$b_n^\lambda \le c a_n^\lambda \quad \text{and} \quad \|u_n^\lambda\| \le c a_n^\lambda \quad \text{for all } n \in \mathbb{N}_0 \text{ and } \lambda \ge 0. \tag{4.34}$$

Proof. As the case $\lambda = 0$ already follows from Lemmas 4.2 and 4.3, it is enough to show (4.34) with $\lambda \ge 0$ replaced by $\lambda > 0$.

By (4.27) from the proof of Lemma 4.3,

$$b_{n+1}^\lambda \le \gamma b_n^\lambda + c\|u_n^\lambda\|^2, \tag{4.35}$$

with c that is uniform over $f \in \mathbb{F}$. An iterative application of this inequality yields

$$b_{n+1}^\lambda \le \gamma^n b_0^\lambda + c \sum_{l=0}^{n} \gamma^{n-l} \|u_l^\lambda\|^2. \tag{4.36}$$

To continue, we argue that

$$\|u_l^\lambda\| \le \lambda^{-1} \quad \text{for every } l \in \mathbb{N}_0,\ \lambda > 0. \tag{4.37}$$

Indeed, by Lemma 4.1, $u_{n+1}^\lambda(x) \le L[u_n^\lambda](x)$. Applying this recursively, taking norms, and using that the operator norm of L is one, we obtain

$$\|u_l^\lambda\| \le \|L^l[u_0^\lambda]\| \le \|u_0^\lambda\|. \tag{4.38}$$

Moreover, since by definition $0 \le u_0^\lambda = 1 - \exp(-\lambda^{-1}\chi) \le \lambda^{-1}\chi$, we know that $\|u_0^\lambda\| \le \lambda^{-1}\|\chi\| = \lambda^{-1}$. Together with (4.38), this shows (4.37).

Since also $b_0^\lambda \le \|u_0^\lambda\| \le \lambda^{-1}$ and $\gamma < 1$, (4.36) and (4.37) imply

$$b_{n+1}^\lambda \le \gamma^n \lambda^{-1} + c\lambda^{-2}. \tag{4.39}$$

On the other hand, by Lemma 4.3, $b_n^\lambda \le cn^{-2}$, and thus

$$b_n^\lambda \le c \min\Big(\gamma^n \lambda^{-1} + \lambda^{-2}, \frac{1}{n^2}\Big). \tag{4.40}$$

Since $1 - f_\lambda(h^*) \ge c\lambda^{-1}$, (4.3) implies that

$$a_n^\lambda = \langle u_n^\lambda, \chi \rangle \ge (1 - f_\lambda(h^*))\langle u_n^0, \chi \rangle \ge c\lambda^{-1} a_n^0 \ge c\lambda^{-1} n^{-1}, \tag{4.41}$$

where the last inequality follows from Lemma 4.2. Therefore,

$$\frac{b_n^\lambda}{a_n^\lambda} \le \frac{c \min\big(\gamma^n \lambda^{-1} + \lambda^{-2}, n^{-2}\big)}{\lambda^{-1} n^{-1}} \le c \min\Big(c\gamma^n n + \frac{n}{\lambda}, \frac{\lambda}{n}\Big) \le c\gamma^n n + c, \tag{4.42}$$

which is bounded from above by some constant c independent of λ or n. This proves the first statement. The second follows from the first one and $\|u_n^\lambda\|^2 = (a_n^\lambda)^2 + (b_n^\lambda)^2$. $\square$

We now improve the result of Lemma 4.4 and show that $b_n^\lambda = o(a_n^\lambda)$, uniformly in λ.

Proposition 4.5. *Uniformly over $\lambda \ge 0$,*

$$\lim_{n\to\infty} \frac{u_n^\lambda}{\langle \chi, u_n^\lambda \rangle} = \chi \quad \text{in } L^2(\nu). \tag{4.43}$$

Proof. The proof consists of two steps. First we show that, for a suitable choice of $n_0(n)$, which slowly diverges with n,

$$\frac{u_n^\lambda}{\langle \chi, u_{n-n_0(n)}^\lambda\rangle} \to \chi \quad \text{in } L^2(\nu) \text{ as } n\to\infty, \text{ uniformly in } \lambda \geq 0. \tag{4.44}$$

In a second step, we then show that $\langle \chi, u_{n-n_0(n)}^\lambda\rangle / \langle \chi, u_n^\lambda\rangle \to 1$ as $n\to\infty$ uniformly in $\lambda\geq 0$, implying the claim of the proposition.

By (4.11) from the proof of Lemma 4.1,

$$\begin{aligned} u_n^\lambda(x) &= E_x\Big[E_x\Big[1-\prod_{v\in N_n^+} f_\lambda(\varphi_v)\Big|\mathcal{F}_{n_0}\Big]\Big] = E_x\Big[1-\prod_{v\in N_{n_0}^+}\big(1-u_{n-n_0}^\lambda(\varphi_v)\big)\Big] \\ &= E_x\Big[\sum_{v\in N_{n_0}^+} u_{n-n_0}^\lambda(\varphi_v)\Big] - E_x\Big[\prod_{v\in N_{n_0}^+}\big(1-u_{n-n_0}^\lambda(\varphi_v)\big) - 1 + \sum_{v\in N_{n_0}^+} u_{n-n_0}^\lambda(\varphi_v)\Big]. \end{aligned} \tag{4.45}$$

By Proposition 3.8(a), the first expectation on the right-hand side satisfies

$$E_x\Big[\sum_{u\in N_{n_0}^+} u_{n-n_0}^\lambda(\varphi_u)\Big] = \chi(x)\langle\chi, u_{n-n_0}^\lambda\rangle + \varepsilon_{n_0}^{\lambda,n-n_0}(x) \tag{4.46}$$

with $\|\varepsilon_{n_0}^{\lambda,n-n_0}\| \leq \gamma^{n_0}\|u_{n-n_0}^\lambda\|$.

To estimate the second expectation we need the following inequality: For any finite index set I and $a_i\in[0,1]$, $i\in I$,

$$0\leq \prod_{i\in I}(1-a_i) - 1 + \sum_{i\in I} a_i \leq \frac{1}{2}\sum_{i\neq j\in I} a_i a_j. \tag{4.47}$$

This inequality follows easily from Bonferroni inequalities for independent events A_i with $P(A_i)=a_i$. Indeed,

$$1-\prod_{i\in I}(1-a_i) = P(\cup_{i\in I}A_i) \leq \sum_{i\in I} P(A_i) = \sum_{i\in I} a_i \tag{4.48}$$

implies the lower bound in (4.47), and

$$1-\prod_{i\in I}(1-a_i) = P(\cup_{i\in I}A_i) \geq \sum_{i\in I} P(A_i) - \frac{1}{2}\sum_{i\neq j\in I} P(A_i\cap A_j) = \sum_{i\in I} a_i - \frac{1}{2}\sum_{i\neq j\in I} a_i a_j \tag{4.49}$$

implies the upper bound. Inequality (4.47) and Proposition 3.8(b) imply

$$\begin{aligned} 0 &\leq E_x\Big[\prod_{u\in N_{n_0}^+}\big(1-u_{n-n_0}^\lambda(\varphi_u)\big) - 1 + \sum_{u\in N_{n_0}^+} u_{n-n_0}^\lambda(\varphi_u)\Big] \\ &\leq c\chi(x)\langle\chi, u_{n-n_0}^\lambda\rangle^2\langle\chi,\chi^2\rangle n_0 + \bar\varepsilon_{n_0}^{\lambda,n-n_0}(x), \end{aligned} \tag{4.50}$$

with $\|\bar\varepsilon_{n_0}^{\lambda,n-n_0}\| \leq C\|u_{n-n_0}^\lambda\|^2$.

Dividing (4.45) by $\langle \chi, u_{n-n_0}^\lambda \rangle$ and combining it with (4.46) and (4.50), we obtain

$$\begin{aligned} &\left\| \frac{u_n^\lambda}{\langle \chi, u_{n-n_0}^\lambda \rangle} - \chi \right\| \\ &\leq \frac{\|\varepsilon_{n_0}^{\lambda, n-n_0}\|}{\langle \chi, u_{n-n_0}^\lambda \rangle} + \frac{c\|\chi \langle \chi, u_{n-n_0}^\lambda \rangle^2 \langle \chi, \chi^2 \rangle n_0\|}{\langle \chi, u_{n-n_0}^\lambda \rangle} + \frac{\|\tilde{\varepsilon}_{n_0}^{\lambda, n-n_0}(x)\|}{\langle \chi, u_{n-n_0}^\lambda \rangle} \\ &\leq \gamma^{n_0} \frac{\|u_{n-n_0}^\lambda\|}{\langle \chi, u_{n-n_0}^\lambda \rangle} + c\langle \chi, \chi^2 \rangle \langle \chi, u_{n-n_0}^\lambda \rangle n_0 + c \frac{\|u_{n-n_0}^\lambda\|}{\langle \chi, u_{n-n_0}^\lambda \rangle} \|u_{n-n_0}^\lambda\| \\ &\leq c\gamma^{n_0} + \frac{c\langle \chi, \chi^2 \rangle n_0}{n - n_0} + \frac{c}{n - n_0}, \end{aligned} \tag{4.51}$$

where in the last inequality we used that $\|u_n^\lambda\| / \langle \chi, u_n^\lambda \rangle$ is uniformly bounded by Lemma 4.4, and both $\langle \chi, u_n^\lambda \rangle$ and $\|u_n^\lambda\|$ are uniformly bounded above by cn^{-1} by Lemma 4.3. This establishes (4.44) with, for example, $n_0(n) = \lfloor \log(n) \rfloor$.

Next we show that (4.51) also implies

$$\left| \frac{\langle \chi, u_n^\lambda \rangle}{\langle \chi, u_{n-n_0(n)}^\lambda \rangle} - 1 \right| \to 0 \quad \text{as } n \to \infty, \text{ uniformly in } \lambda \geq 0. \tag{4.52}$$

To see this, we use that $1 = \langle \chi, \chi \rangle$ and $|\langle \chi, f \rangle| \leq \|f\|$ to get

$$\begin{aligned} \left| \frac{\langle \chi, u_n^\lambda \rangle}{\langle \chi, u_{n-n_0(n)}^\lambda \rangle} - 1 \right| &= \left| \frac{\langle \chi, u_n^\lambda \rangle}{\langle \chi, u_{n-n_0(n)}^\lambda \rangle} - \langle \chi, \chi \rangle \right| = \left| \left\langle \chi, \frac{u_n^\lambda}{\langle \chi, u_{n-n_0(n)}^\lambda \rangle} - \chi \right\rangle \right| \\ &\leq \left\| \frac{u_n^\lambda}{\langle \chi, u_{n-n_0(n)}^\lambda \rangle} - \chi \right\|, \end{aligned} \tag{4.53}$$

which by (4.51) converges to zero uniformly in λ. This shows (4.52), which together with (4.44) finishes the proof. □

Using the hypercontractivity of L from Proposition 3.2, it is easy to slightly improve Proposition 4.5. This will be useful to ensure that certain products still lie in $L^2(\nu)$.

Lemma 4.6. *Uniformly over $\lambda \geq 0$,*

$$\lim_{n \to \infty} \frac{u_n^\lambda}{\langle \chi, u_n^\lambda \rangle} = \chi \quad \textit{in } L^{5/2}(\nu) \textit{ and pointwise over } [h^*, \infty). \tag{4.54}$$

Proof. We will first show (4.54). We set $p = 5/2$. By Lemma 4.1, $u_n^\lambda = L[u_{n-1}^\lambda] + \sum_{l=2}^d c_l L[u_{n-1}^\lambda]^l$. Therefore, also using that $L[\chi] = \chi$,

$$\left\| \frac{u_n^\lambda}{\langle \chi, u_n^\lambda \rangle} - \chi \right\|_{L^p(\nu)} \leq \left\| L\left[\frac{u_{n-1}^\lambda}{\langle \chi, u_n^\lambda \rangle} - \chi \right] \right\|_{L^p(\nu)} + \frac{c}{\langle \chi, u_n^\lambda \rangle} \sum_{l=2}^d \|L[u_{n-1}^\lambda]^l\|_{L^p(\nu)}. \tag{4.55}$$

Since $0 \leq u_n^\lambda \leq 1$, it holds that $0 \leq L[u_n^\lambda] \leq d$. Therefore,

$$\|L[u_{n-1}^\lambda]^l\|_{L^p(\nu)} \leq d^{l-2} \|L[u_{n-1}^\lambda]^2\|_{L^p(\nu)} = d^{l-2} \|L[u_{n-1}^\lambda]\|_{L^{2p}(\nu)}^2 \leq c\|u_{n-1}^\lambda\|_{L^2(\nu)}^2, \tag{4.56}$$

where in the last inequality we used Proposition 3.2(a) with $q = 2p = 5$ and $p = 2$. Therefore, using also (4.52) (with $n_0 = 1$), Lemma 4.4, and the fact that $\|u_n^\lambda\| \leq \|u_n^0\| \to 0$ as $n \to \infty$, we get

$$\frac{\|L[u_{n-1}^\lambda]^l\|_{L^p(\nu)}}{\langle \chi, u_n^\lambda \rangle} \leq \frac{\|u_{n-1}^\lambda\|_{L^2(\nu)}^2}{c\langle \chi, u_{n-1}^\lambda \rangle} \to 0 \quad \text{as } n \to \infty \text{ uniformly in } \lambda \geq 0. \tag{4.57}$$

Using Proposition 3.2(a) again for the remaining term on the right-hand side of (4.55),

$$\Big\| L\Big[\frac{u^\lambda_{n-1}}{\langle\chi,u^\lambda_n\rangle}-\chi\Big]\Big\|_{L^p(\nu)} \le d\Big\|\frac{u^\lambda_{n-1}}{\langle\chi,u^\lambda_n\rangle}-\chi\Big\|_{L^2(\nu)} \to 0 \quad \text{as } n\to\infty \text{ uniformly in } \lambda\ge 0. \tag{4.58}$$

Here, the convergence follows from Proposition 4.5. Combining (4.55) with (4.57) and (4.58) shows (4.54).

We now show the pointwise convergence. By similar steps as in (4.55), using Lemma 3.6, for $x \ge h^*$,

$$\begin{aligned}\Big|\frac{u^\lambda_n(x)}{\langle\chi,u^\lambda_n\rangle}-\chi(x)\Big| &\le \Big|L\Big[\frac{u^\lambda_{n-1}}{\langle\chi,u^\lambda_n\rangle}-\chi\Big](x)\Big| + \frac{1}{\langle\chi,u^\lambda_n\rangle}\sum_{l=2}^d |L[u^\lambda_{n-1}](x)|^l \\ &\le \Big\|\frac{u^\lambda_{n-1}}{\langle\chi,u^\lambda_n\rangle}-\chi\Big\| q(x) + \frac{1}{\langle\chi,u^\lambda_n\rangle}\sum_{l=2}^d \|u^\lambda_{n-1}\|^l q(x)^l.\end{aligned} \tag{4.59}$$

Using (4.57) and (4.58), it is immediate that the right-hand side of (4.59) converges to 0 as $n\to\infty$, uniformly in $\lambda\ge 0$. This shows the pointwise convergence and finishes the proof. □

The following lemma provides the final ingredient for the proof of Theorem 2.1. Its result allows us to determine the exact asymptotic behaviour of a^λ_n as $n\to\infty$.

Lemma 4.7. *Let C_1 be as in (2.11). Then,*

$$\lim_{n\to\infty}\frac{1}{n}\Big(\frac{1}{a^\lambda_n}-\frac{1}{a^\lambda_0}\Big) = C_1^{-1} \quad \textit{uniformly in } \lambda\ge 0. \tag{4.60}$$

Proof. We show that

$$\lim_{n\to\infty}\Big(\frac{1}{a^\lambda_n}-\frac{1}{a^\lambda_{n-1}}\Big) = \lim_{n\to\infty}\frac{a^\lambda_{n-1}-a^\lambda_n}{a^\lambda_{n-1}a^\lambda_n} = C_1^{-1} \tag{4.61}$$

uniformly in $\lambda\ge 0$. The claim of the lemma then follows from (4.61) by telescoping the difference therein.

By Lemma 4.1 again, this time writing the prefactor of the quadratic term explicitly,

$$a^\lambda_n = \langle\chi,u^\lambda_n\rangle = \langle\chi,L[u^\lambda_{n-1}]\rangle - \binom{d}{2}\frac{1}{d^2}\langle\chi,L[u^\lambda_{n-1}]^2\rangle - \sum_{l=3}^d c_l\langle\chi,L[u^\lambda_{n-1}]^l\rangle. \tag{4.62}$$

Using that L is self-adjoint and $L[\chi]=\chi$, we get $\langle\chi,L[u^\lambda_{n-1}]\rangle = \langle\chi,u^\lambda_{n-1}\rangle = a^\lambda_{n-1}$. Hence, after dividing by $(a^\lambda_{n-1})^2$, (4.62) implies

$$\begin{aligned}\frac{a^\lambda_{n-1}-a^\lambda_n}{(a^\lambda_{n-1})^2} &= \binom{d}{2}d^{-2}\Big\langle\chi,\frac{L[u^\lambda_{n-1}]^2}{(a^\lambda_{n-1})^2}\Big\rangle + \sum_{l=3}^d c_l\Big\langle\chi,\frac{L[u^\lambda_{n-1}]^l}{(a^\lambda_{n-1})^2}\Big\rangle \\ &= \binom{d}{2}d^{-2}\Big\langle\chi,L\Big[\frac{u^\lambda_{n-1}}{a^\lambda_{n-1}}\Big]^2\Big\rangle + \sum_{l=3}^d c_l\Big\langle\chi,L\Big[\frac{u^\lambda_{n-1}}{a^\lambda_{n-1}}\Big]^2 L[u^\lambda_{n-1}]^{l-2}\Big\rangle.\end{aligned} \tag{4.63}$$

By Propositions 4.5 and 3.2, $L[u^\lambda_{n-1}/a^\lambda_{n-1}]^2 \to \chi^2$ in $L^2(\nu)$, uniformly in λ. By (4.6) and (4.3), $0\le u^\lambda_n \le u^0_n \to 0$ in $L^2(\nu)$ as $n\to\infty$. Therefore, using that $L[u^\lambda_{n-1}]\le d$ and the boundedness of L, $L[u^\lambda_{n-1}]^{l-2} \le d^{l-3}L[u^\lambda_{n-1}] \to 0$ in $L^2(\nu)$ uniformly in λ. Using this in (4.63) then yields

$$\lim_{n\to\infty} \frac{a^\lambda_{n-1} - a^\lambda_n}{(a^\lambda_{n-1})^2} = \binom{d}{2} d^{-2}\langle \chi, \chi^2\rangle = C_1^{-1} \quad \text{uniformly in } \lambda \ge 0. \tag{4.64}$$

From this (4.61) follows, if we show

$$\lim_{n\to\infty} \frac{a^\lambda_n}{a^\lambda_{n-1}} = 1 \quad \text{uniformly in } \lambda \ge 0. \tag{4.65}$$

To see this, note that by multiplying (4.64) by a^λ_{n-1},

$$1 - a^\lambda_n / a^\lambda_{n-1} = a^\lambda_{n-1} C_1^{-1} + o(a^\lambda_{n-1}) = o(1), \tag{4.66}$$

where the $o(1)$ term is independent of λ since $0 \le a^\lambda_n \le a^0_n \to 0$ for $n \to \infty$. This shows (4.65), and together with (4.64) completes the proof of (4.61). □

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. Recall from (4.2) that $u^0_n(x) = P_x[N^+_n \ne \emptyset]$. Therefore, by Proposition 4.5 and Lemma 4.6,

$$\lim_{n\to\infty} \frac{P_\cdot[N^+_n \ne \emptyset]}{\langle \chi, u^0_n\rangle} = \chi \quad \text{in } L^2(\nu) \text{ and pointwise.} \tag{4.67}$$

Moreover, by Lemma 4.7,

$$\langle \chi, u^0_n \rangle = a^0_n = C_1 n^{-1}(1 + o(1)). \tag{4.68}$$

Combining these two facts one obtains (2.9), and by integrating over $\varphi(o)$ which is ν-distributed also (2.10).

To show the last statement of the theorem, let $x_1, \dots, x_{d+1} = \bar o$ be the neighbours of o in $\mathbb{T}$, and let $A_{n,i}$ be the event "the subtree of x_i intersects N_n". By the branching process construction, for every $n \ge 1$, the events $A_{n,1}, \dots, A_{n,d+1}$ are independent and have the same probability $p = p(n, x)$. Further, $P_x[N_n \ne \emptyset] = P_x[\cup_{i=1}^{d+1} A_{n,i}] = 1 - (1-p)^{d+1}$ and similarly $P_x[N^+_n \ne \emptyset] = 1 - (1-p)^d$ and thus

$$P_x[N_n \ne \emptyset] = 1 - (1 - P_x[N^+_n \ne \emptyset])^{\frac{d+1}{d}} = \frac{d+1}{d} P_x[N^+_n \ne \emptyset] + O(P_x[N^+_n \ne \emptyset]^2). \tag{4.69}$$

This directly implies the statement for $P_x[N_n \ne \emptyset]$. The statement for $P[N_n \ne \emptyset]$ is again obtained by integration over φ_o. □

5. Proof of Theorem 2.2

The goal of this section is twofold: We show Theorem 2.2, and then, in Section 5.1, provide further results concerning the behaviour of $\mathcal{C}_o$ conditioned on $\{N^+_n \ne \emptyset\}$. These results are, more or less, direct consequences of Theorems 2.1 and 2.2, and will later be used to show Theorem 2.3.

We start with a preparatory lemma which is a special case of Theorem 2.2 for functions f orthogonal to χ.

Lemma 5.1. *For every $f \in L^2(\nu)$ such that $\langle \chi, f\rangle = 0$ there is a sequence $\varepsilon^f_n \in L^2(\nu)$ converging to zero in $L^2(\nu)$ and pointwise, such that for every $\delta > 0$, $x \ge h^*$, and $n \ge 1$,*

$$P_x\Big[\Big|\frac{1}{n}\sum_{v\in N^+_n} f(\varphi_v)\Big| > \delta \,\Big|\, N^+_n \ne \emptyset\Big] \le c\delta^{-2}\varepsilon^f_n(x). \tag{5.1}$$

Proof. By the conditional Markov inequality

$$P_x\Big[\Big|\frac{1}{n}\sum_{v\in N_n^+} f(\varphi_v)\Big| > \delta\Big|N_n^+\neq\emptyset\Big] \le \delta^{-2}E_x\Big[\Big(\frac{1}{n}\sum_{v\in N_n^+} f(\varphi_v)\Big)^2\Big|N_n^+\neq\emptyset\Big]. \tag{5.2}$$

The conditional expectation on the right-hand side satisfies

$$E_x\Big[\Big(\frac{1}{n}\sum_{v\in N_n^+} f(\varphi_v)\Big)^2\Big|N_n^+\neq\emptyset\Big] = \Big(\frac{1}{nP_x[N_n^+\neq\emptyset]}\Big)\Big(\frac{1}{n}E_x\Big[\Big(\sum_{v\in N_n^+} f(\varphi_v)\Big)^2\Big]\Big). \tag{5.3}$$

By Proposition 3.8(a,b), since $\langle\chi, f\rangle = 0$,

$$\frac{1}{n}E_x\Big[\Big(\sum_{v\in N_n^+} f(\varphi_v)\Big)^2\Big] = \varepsilon_n^f(x), \tag{5.4}$$

with $\varepsilon_n^f \to 0$ in L^2 and pointwise. Further, by the stochastic domination (3.5), using Theorem 2.1, $(nP_x[N_n^+\neq\emptyset])^{-1} \le (nP_{h^*}[N_n^+\neq\emptyset])^{-1} \le c$ and the statement of the lemma follows. □

Proof of Theorem 2.2. We first consider the special case $f=\chi$; the general case will then easily follow using Lemma 5.1.

We start by showing that for every $\alpha>0$,

$$E\big[e^{-\alpha Z_n^{\chi,x}}\big] = \frac{C_1}{C_1+\alpha} + \varepsilon_n^\alpha(x), \tag{5.5}$$

with $\varepsilon_n^\alpha \to 0$ pointwise and in $L^2(\nu)$. Using the definitions (4.1), (4.32) of u_n^f and f_λ, setting as before $u_n^\lambda = u_n^{f_\lambda}$, and recalling (4.2), we obtain

$$E\big[e^{-\alpha Z_n^{\chi,x}}\big] = E_x\Big[\exp\Big(-\frac{\alpha}{n}\sum_{v\in N_n^+}\chi(\varphi_v)\Big)\Big|N_n^+\neq\emptyset\Big] = 1-\frac{u_n^{n/\alpha}(x)}{u_n^0(x)}. \tag{5.6}$$

By Lemma 4.6 and (4.68),

$$u_n^0(x) = Cn^{-1}\chi(x)(1+\varepsilon(x)), \tag{5.7}$$

where $\varepsilon\to 0$ pointwise. Since Lemma 4.6 holds uniformly in λ, we can apply it with $\lambda = n/\alpha$ to obtain

$$u_n^{n/\alpha}(x) = a_n^{n/\alpha}\chi(x)(1+\tilde\varepsilon(x)), \tag{5.8}$$

again with $\tilde\varepsilon\to 0$ pointwise. Moreover, by Lemma 4.7, again using the uniformity in λ,

$$(na_n^{n/\alpha})^{-1} = (na_0^{n/\alpha})^{-1} + C_1^{-1} + o(1) \quad \text{as } n\to\infty. \tag{5.9}$$

By the definitions of a_n^λ, u_n^λ and f_λ, and by the monotone convergence theorem,

$$\begin{aligned} na_0^{n/\alpha} &= n\langle\chi, u_0^{n/\alpha}\rangle = n\big\langle\chi, E_\cdot\big[1-f_{n/\alpha}(\varphi_o)\big]\big\rangle \\ &= n\langle\chi, 1-f_{n/\alpha}\rangle = \big\langle\chi, n\big(1-e^{-\alpha\chi(\cdot)/n}\big)\big\rangle \xrightarrow{n\to\infty} \alpha\langle\chi,\chi\rangle = \alpha. \end{aligned} \tag{5.10}$$

Inserting this into (5.9) yields

$$a_n^{n/\alpha} = \frac{\alpha C_1(1+o(1))}{n(\alpha+C_1)} \quad \text{as } n\to\infty. \tag{5.11}$$

Combining (5.6)–(5.8), (5.11) gives

$$1-E\big[e^{-\alpha Z_n^{\chi,x}}\big] = \frac{\alpha}{\alpha+C_1}(1+o(1)), \tag{5.12}$$

which shows the pointwise convergence in (5.5). The $L^2(\nu)$-convergence follows by the dominated convergence theorem, since the left-hand side of (5.5) is bounded by 1.

By Lévy's continuity theorem for the Laplace transform, (5.5) implies the statement (2.13) of the theorem in the case $Z_n^{f,x}$ with $f = \chi$. For general $f \in L^2(\nu)$ we write $f = \langle \chi, f\rangle \chi + \beta[f]$ as usual. Then $Z_n^{f,x} = \langle \chi, f\rangle Z_n^{\chi,x} + Z_n^{\beta[f],x}$. The statement for $Z_n^{f,x}$ then directly follows, since the first summand converges to an exponential random variable with mean $C_1^{-1}\langle \chi, f\rangle$, by the first step of the proof, and the second summand converges to 0 in probability, by Lemma 5.1, since $\langle \chi, \beta[f]\rangle = 0$.

We now show the statement of the theorem for Z_n^f. Let ν_n be the law of φ_o under $P[\,\cdot\,|N_n^+ \neq \emptyset]$. By integrating (5.5) over $\nu_n(\mathrm{d}x)$

$$E\big[e^{-Z_n^\chi}\big] = \frac{C_1}{C_1 + \alpha} + \int \varepsilon_n^\alpha(x)\nu_n(\mathrm{d}x). \tag{5.13}$$

We need to show that the last integral is $o(1)$. Observe that $\nu_n(\mathrm{d}x) = P[\varphi_o \in \mathrm{d}x|N_n^+ \neq \emptyset] = P[\varphi_o \in \mathrm{d}x, N_n^+ \neq \emptyset]P[N_n^+ \neq \emptyset]^{-1} = P_x[N_n^+ \neq \emptyset]P[N_n^+ \neq \emptyset]^{-1}\nu(\mathrm{d}x)$. Therefore, using the Cauchy–Schwarz inequality and that $\varepsilon_n^\alpha \to 0$ in $L^2(\nu)$, the integral in (5.13) is bounded above by

$$o(1)\Big(\int \frac{P_x[N_n^+ \neq \emptyset]^2}{P[N_n^+ \neq \emptyset]^2}\nu(\mathrm{d}x)\Big)^{1/2}. \tag{5.14}$$

By (4.67), $P_x[N_n^+ \neq \emptyset]P[N_n^+ \neq \emptyset]^{-1} = \chi(x)\langle 1, \chi\rangle^{-1} + \varepsilon_n(x)$, with $\varepsilon_n(x) \to 0$ in $L^2(\nu)$ as $n \to \infty$. This implies that the integral in (5.14) is $O(1)$ and together with (5.13) and again Lévy's continuity theorem shows statement (2.13) for Z_n^f with $f = \chi$. For Z_n^f and general $f \in L^2(\nu)$ the statement then follows by integrating (5.1) over $\nu_n(\mathrm{d}x)$ and applying similar arguments as above. □

5.1. **Further results for the conditioned model.** The following few results which all concern the model conditioned to $\{N_n^+ \neq \emptyset\}$ will later be used in the proof of Theorem 2.3. The first one explains the role of the measure Q_x introduced using the spine construction in Section 3.2. Similar results are well established in branching processes literature, see for instance [HH07, Theorem 5] or [CR90, Theorem 4]. Proposition 1.5 in [Pow19] can be seen as the analogous result for branching diffusion on bounded domains.

Proposition 5.2. *For all $K \in \mathbb{N}$, $x \geq h^*$ and $B \in \mathcal{F}_K$ (see (3.20))*

$$\lim_{n\to\infty} P_x[B|N_n^+ \neq \emptyset] = Q_x[B]. \tag{5.15}$$

Proof. We follow similar steps as in the proof of [Pow19, Proposition 1.5]. For every $n \geq K$,

$$P_x[B|N_n^+ \neq \emptyset] = E_x\Big[\frac{1_B P_x[N_n^+ \neq \emptyset|\mathcal{F}_K]}{P_x[N_n^+ \neq \emptyset]}\Big]. \tag{5.16}$$

Writing $N_K^+ = \{v_1, \dots, v_{|N_K^+|}\}$ and defining the events $A_i = \{v_i$ is root of a subtree in $\mathcal{C}_o$ with height of at least $n - K\}$,

$$\{N_n^+ \neq \emptyset\} = \bigcup_{k=1}^{|N_K^+|} A_k = \bigcup_{k=1}^{|N_K^+|} \Big(A_k \cap \bigcap_{j<k} A_j^c\Big), \tag{5.17}$$

where the last union is disjoint. Since the A_i are independent conditionally on $\mathcal{F}_K$,

$$P_x[N_n^+ \neq \emptyset | \mathcal{F}_K] = \sum_{i=1}^{|N_K^+|} P_{\varphi_{v_i}}[N_{n-K}^+ \neq \emptyset] \prod_{j<i} P_{\varphi_{v_j}}[N_{n-K}^+ = \emptyset]. \tag{5.18}$$

By Theorem 2.1, $P_{\varphi_{v_i}}[N_{n-K}^+ \neq \emptyset] \sim C_1(n-K)^{-1}\chi(\varphi_{v_i}) \to 0$ and $P_x[N_n^+ \neq \emptyset] \sim C_1 n^{-1}\chi(x)$ as $n \to \infty$. Therefore,

$$\lim_{n\to\infty} \frac{1_B P_x[N_n^+ \neq \emptyset|\mathcal{F}_K]}{P_x[N_n^+ \neq \emptyset]} = 1_B \frac{\sum_{v\in N_K} \chi(\varphi_v)}{\chi(x)}, \qquad P_x\text{-a.s.} \tag{5.19}$$

In order to insert this into (5.16), we use the generalised dominated convergence theorem. To this end we bound the fraction on the left-hand side of (5.19) by a function g_n satisfying $g_n \to g$ in $L^1(P_x)$ and P_x-a.s.: Using (5.17) and Theorem 2.1 (or (4.67), (4.68) from its proof),

$$P_x[N_n^+ \neq \emptyset|\mathcal{F}_K] \le \sum_{v\in N_K^+} P_{\varphi_v}[N_{n-K}^+ \neq \emptyset] = \sum_{v\in N_K^+} C_1(n-K)^{-1}(\chi(\varphi_v) + \bar\varepsilon_{n-K}(\varphi_v)) \tag{5.20}$$

with $\bar\varepsilon_n \to 0$ pointwise and in $L^2(\nu)$ as $n \to \infty$. For the denominator, again by Theorem 2.1, $P_x[N_n^+ \neq \emptyset] \ge P_{h^*}[N_n^+ \neq \emptyset] \ge cn^{-1}$ for some constant c. Thus, for $n > K$,

$$\begin{aligned}\frac{1_B P_x[N_n^+ \neq \emptyset|\mathcal{F}_K]}{P_x[N_n^+ \neq \emptyset]} &\le \frac{cn}{n-K} \sum_{v\in N_K^+} (\chi(\varphi_v) + \bar\varepsilon_{n-K}(\varphi_v)) \\ &\le c' \sum_{v\in N_K^+} (\chi(\varphi_v) + \bar\varepsilon_{n-K}(\varphi_v)) =: g_n,\end{aligned} \tag{5.21}$$

which converges a.s. to $g := c' \sum_{v\in N_K^+} \chi(\varphi_v)$. Moreover, by Proposition 3.8,

$$E_x[|g_n - g|] \le E_x\Big[c' \sum_{v\in N_K^+} |\bar\varepsilon_{n-K}(\varphi_v)|\Big] = \langle \chi, |\bar\varepsilon_{n-K}|\rangle \chi(x) + \varepsilon_K^{|\bar\varepsilon_{n-K}|}(x). \tag{5.22}$$

Since $\|\bar\varepsilon_n\| \to 0$, the bounds established in (3.34) imply that the right-hand side of (5.22) converges to 0, that is $g_n \to g$ in $L^1(P^x)$. Therefore, by the generalised dominated convergence theorem, using (5.16), (5.19),

$$\lim_{n\to\infty} P_x[B|N_n^+ \neq \emptyset] = E_x\Big[1_B \frac{\sum_{v\in N_K}\chi(\varphi_v)}{\chi(x)}\Big] = Q_x[B], \tag{5.23}$$

where the last equality follows from Lemma B.2 with $k = 1$ and $Y(v_1) = 1_B\chi(\varphi_{v_1})$. □

The following lemma and its corollary follow almost directly from the results established in the proof of Theorem 2.2.

Lemma 5.3. *For every $f \in L^2(\nu)$, $x \ge h^*$ and $\delta > 0$,*

$$\lim_{n\to\infty} P_x\left[\left|\frac{\sum_{v\in N_n^+} f(\varphi_v)}{\sum_{v\in N_n^+}\chi(\varphi_v)} - \langle\chi, f\rangle\right| > \delta \,\middle|\, N_n^+ \neq \emptyset\right] = 0. \tag{5.24}$$

Proof. Observe that

$$\frac{\sum_{v\in N_n^+} f(\varphi_v)}{\sum_{v\in N_n^+}\chi(\varphi_v)} - \langle\chi, f\rangle = \frac{\sum_{v\in N_n^+}(f(\varphi_v) - \langle\chi, f\rangle\chi(\varphi_v))}{\sum_{v\in N_n^+}\chi(\varphi_v)} = \frac{n^{-1}\sum_{v\in N_n^+}\beta[f](\varphi_v)}{n^{-1}\sum_{v\in N_n^+}\chi(\varphi_v)}. \tag{5.25}$$

For any $A, B \in \mathbb{R}$ and $\varepsilon, \delta > 0$, $\{|A/B| > \varepsilon\} \subset \{|A| > \delta\} \cup \{|B| < \delta/\varepsilon\}$, which together with (5.25) allows us to bound the left-hand side of (5.24) by

$$P_x\Big[\Big|n^{-1}\sum_{v\in N_n^+}\beta[f](\varphi_v)\Big| > \bar\delta_n \,\Big|\, N_n^+\neq\emptyset\Big] + P_x\Big[n^{-1}\sum_{v\in N_n^+}\chi(\varphi_v) < \bar\delta_n/\delta \,\Big|\, N_n^+\neq\emptyset\Big]. \tag{5.26}$$

We now show that there is a sequence $\bar\delta_n \to 0$ such that both summands converge to zero. By Lemma 5.1, the first summand is bounded by $\bar\delta_n^{-2}\bar\varepsilon_n^f(x)$, where the $\bar\varepsilon_n^f$-term is independent of $\bar\delta_n$ and converges to zero. Therefore, if $\bar\delta_n \to 0$ sufficiently slowly, then also $\bar\delta_n^{-2}\bar\varepsilon_n^f(x) \to 0$. For the second summand, fix $\varepsilon > 0$. Then for n large enough so that $\bar\delta_n/\delta \le \varepsilon$,

$$P_x\Big[n^{-1}\sum_{v\in N_n^+}\chi(\varphi_v) < \bar\delta_n/\delta \,\Big|\, N_n^+\neq\emptyset\Big] \le P_x\Big[n^{-1}\sum_{v\in N_n^+}\chi(\varphi_v) < \varepsilon \,\Big|\, N_n^+\neq\emptyset\Big]. \tag{5.27}$$

By Theorem 2.2, the right-hand side converges to $P(Z < C_1\varepsilon)$ which can be made arbitrarily small by choosing ε small. This implies that the second summand in (5.26) converges to zero and completes the proof. □

Corollary 5.4. *For every $x \ge h^*$, $\delta > 0$ and $f, g \in L^2(\nu)$ with $g > 0$,*

$$\lim_{n\to\infty} P_x\Big[\Big|\frac{\sum_{v\in N_n^+} f(\varphi_v)}{\sum_{v\in N_n^+} g(\varphi_v)} - \frac{\langle\chi, f\rangle}{\langle\chi, g\rangle}\Big| > \delta \,\Big|\, N_n^+\neq\emptyset\Big] = 0. \tag{5.28}$$

In particular, setting $g \equiv 1$,

$$\lim_{n\to\infty} P_x\Big[\Big|\frac{1}{|N_n^+|}\sum_{v\in N_n^+} f(\varphi_v) - \frac{\langle\chi, f\rangle}{\langle\chi, 1\rangle}\Big| > \delta \,\Big|\, N_n^+\neq\emptyset\Big] = 0. \tag{5.29}$$

Proof. It is easy to see that there is $\delta' = \delta'(\delta, f, g) > 0$ such that the event on the left-hand side of (5.28) is contained in the union of $\{|\sum f(\varphi_v)/\sum\chi(\varphi_v) - \langle\chi, f\rangle| \ge \delta'\}$ and $\{|\sum g(\varphi_v)/\sum\chi(\varphi_v) - \langle\chi, g\rangle| \ge \delta'\}$. The statement then follows by Lemma 5.3. □

The final result of this section is the following technical lemma which can be seen as a somewhat stronger version of (5.29) in Corollary 5.4. It is tailored to be used in the proof of Proposition 6.7.

Lemma 5.5. *Write $N_n^+ = \{v_1, v_2, \dots, v_{|N_n^+|}\}$ and set $N_{n,M}^+ = \{v_1, v_2, \dots, v_M\}$. For $\delta, \rho > 0$ and $f : \mathbb{R} \to \mathbb{R}$ bounded, define the event*

$$B_n(f, \delta, \rho) := \{\rho n \le |N_n^+|\} \cap \Big(\bigcup_{\rho n \le M \le |N_n^+|}\Big\{\Big|\frac{\sum_{v\in N_{n,M}^+} f(\varphi_v)}{M} - \frac{\langle\chi, f\rangle}{\langle\chi, 1\rangle}\Big| > \delta\Big\}\Big). \tag{5.30}$$

Then for every $x \ge h^*$,

$$\lim_{n\to\infty} P_x\big[B_n(f, \delta, \rho)\big| N_n^+\neq\emptyset\big] = 0. \tag{5.31}$$

Proof. We start by outlining the strategy of the proof. We will define events $A_n = A_n(n_0)$ (see (5.34)) and $A'_n = A'_n(n_0, q)$ (see (5.40)), so that when $n_0 = n_0(n)$ and $q = q(n)$ are chosen correctly,

$$\begin{aligned} &\lim_{n\to\infty} P_x\big[A_n(n_0(n))\big| N_n^+\neq\emptyset\big] = 0,\\ &\lim_{n\to\infty} P_x\big[A'_n(n_0(n), q(n))\big| N_n^+\neq\emptyset\big] = 0,\\ &B_n(f, \delta, \rho) \subseteq A_n(n_0(n)) \cup A'_n(n_0(n), q(n)) \quad \text{for } n \text{ large enough.} \end{aligned} \tag{5.32}$$

The statement of the lemma follows directly from these claims by a union bound.

To define $A_n = A_n(n_0)$, we fix $0 \le n_0 \le n$, write $N^+_{n_0} = \{w_1, \dots, w_{|N^+_{n_0}|}\}$, and set $N^i_n = \{v \in N^+_n : w_i \text{ is ancestor of } v\}$. For $1 \le i \le |N^+_{n_0}|$, we set

$$m_i = \begin{cases} |N^i_n|^{-1} \sum_{v \in N^i_n} f(\varphi_v), & \text{if } |N^i_n| > 0, \\ \langle \chi, f\rangle / \langle \chi, 1\rangle, & \text{otherwise,} \end{cases} \tag{5.33}$$

and define the events

$$A^i_n = \left\{ \left| m_i - \frac{\langle \chi, f\rangle}{\langle \chi, 1\rangle} \right| > \frac{\delta}{2} \right\} \quad \text{and} \quad A_n = A_n(n_0) = \bigcup_{i=1}^{|N^+_{n_0}|} A^i_n. \tag{5.34}$$

We now prove an upper bound on $P[A_n | N^+_n \neq \emptyset]$. By a union bound, using that $A^i_n \subset \{N^i_n \neq \emptyset\} \subset \{N^+_n \neq \emptyset\}$,

$$\begin{aligned} P_x[A_n | N^+_n \neq \emptyset] &= P_x[N^+_n \neq \emptyset]^{-1} P_x[A_n \cap \{N^+_n \neq \emptyset\}] \\ &\le P_x[N^+_n \neq \emptyset]^{-1} E_x\Bigg[\sum_{i=1}^{|N^+_{n_0}|} P_x[A^i_n \cap \{N^i_n \neq \emptyset\} | \mathcal{F}_{n_0}] \Bigg]. \end{aligned} \tag{5.35}$$

Using the branching process properties of the GFF,

$$\begin{aligned} &P_x[A^i_n \cap \{N^i_n \neq \emptyset\} | \mathcal{F}_{n_0}] \\ &= P_{\varphi_{w_i}}\Bigg[\Bigg| \frac{1}{|N^+_{n-n_0}|} \sum_{v \in N^+_{n-n_0}} f(\varphi_v) - \frac{\langle \chi, f\rangle}{\langle \chi, 1\rangle} \Bigg| > \frac{\delta}{2} \,\Bigg|\, N^+_{n-n_0} \neq \emptyset \Bigg] P_{\varphi_{w_i}}[N^+_{n-n_0} \neq \emptyset]. \end{aligned} \tag{5.36}$$

By Corollary 5.4(b), the first probability on the right-hand side is bounded by some $\varepsilon_{n-n_0}(\varphi_{w_i})$ satisfying $1 \ge \varepsilon_n \to 0$ pointwise, and thus in any $L^q(\nu)$ by the bounded convergence theorem. By Proposition 4.5 and Lemmas 4.6, 4.7, the second probability equals

$$P_{\varphi_{w_i}}[N^+_{n-n_0} \neq \emptyset] = C_1 (n-n_0)^{-1} \big(\chi(\varphi_{w_i}) + \bar{\varepsilon}_{n-n_0}(\varphi_{w_i}) \big), \tag{5.37}$$

where $\bar{\varepsilon} \to 0$ pointwise and in $L^{5/2}(\nu)$. Combining these statements with (5.35) and (5.36), and using then Proposition 3.8 in the numerator, we obtain

$$\begin{aligned} P_x[A_n | \{N^+_n \neq \emptyset\}] &\le \frac{E_x\Big[\sum_{w \in N^+_{n_0}} \varepsilon_{n-n_0}(\varphi_w) C_1 (n-n_0)^{-1} (\chi(\varphi_w) + \bar{\varepsilon}_{n-n_0}(\varphi_w)) \Big]}{C_1 n^{-1} (\chi(x) + \bar{\varepsilon}_n(x))} \\ &= \frac{c(x) n}{n - n_0} \langle \chi, \varepsilon_{n-n_0}(\chi + \bar{\varepsilon}_{n-n_0}) \rangle \chi(x) + \tilde{\varepsilon}_{n_0}(x), \end{aligned} \tag{5.38}$$

where $\tilde{\varepsilon}_n \to 0$ pointwise. Finally, using Hölder's inequality on the inner product, since $\chi, \chi^2 \in L^2(\nu)$ (by (3.18)), $\varepsilon_n \to 0$ in $L^5(\nu)$ and $\bar{\varepsilon} \to 0$ in $L^{5/2}(\nu)$, it follows that for every $x \ge h^*$,

$$\text{if } n_0 \to \infty \text{ and } n - n_0 \to \infty, \text{ then } P_x[A_n(n_0) | \{N^+_n \neq \emptyset\}] \to 0. \tag{5.39}$$

We now turn to the events A'_n. For given $n_0 \le n$ and q, let

$$A'_n = A'_n(n_0, q) = \bigcup_{i=1}^{|N^+_{n_0}|} \{ |N^i_n| > q \}. \tag{5.40}$$

To bound the probability $P[A'_n|N_n^+ \neq \emptyset]$, we write

$$P_x[A'_n \cap \{N_n^+ \neq \emptyset\}] \le P_x[A'_n] \le E_x\Big[\sum_{i=1}^{|N_{n_0}^+|} P_{\varphi_{w_i}}[|N_{n-n_0}^+| > q]\Big]. \tag{5.41}$$

By the Markov inequality and Proposition 3.8(b) (with $f = g = 1$),

$$P_{\varphi_{w_i}}[|N_{n-n_0}^+| > q] \le q^{-2} E_{\varphi_{w_i}}[|N_{n-n_0}^+|^2] = Cq^{-2}(\chi(\varphi_{w_i})(n-n_0) + \varepsilon_{n-n_0}(\varphi_{w_i})), \tag{5.42}$$

where $\|\varepsilon_{n-n_0}\| \le c$. Thus, by Proposition 3.8(a),

$$\begin{aligned} P_x[A'_n \cap \{N_n^+ \neq \emptyset\}] &\le q^{-2} E_x\Big[\sum_{v \in N_{n_0}^+} (c\chi(\varphi_v)(n-n_0) + \varepsilon_{n-n_0}(\varphi_v))\Big] \\ &\le cq^{-2}\big((n-n_0+1)\chi(x) + \varepsilon'_{n_0}(x)\big). \end{aligned} \tag{5.43}$$

with $\varepsilon'_n \to 0$ pointwise. Together with $P_x[N_n^+ \neq \emptyset] \ge cn^{-1}$, by Theorem 2.1, this implies

$$P_x[A'_n|N_n^+ \neq \emptyset] \le \frac{n}{q^2}\Big((n-n_0+C)\chi(x) + \varepsilon'_{n_0}(x)\Big). \tag{5.44}$$

We will now give sufficient conditions on the functions $n_0(n)$ and $q(n)$ so that for large enough n,

$$B_n(f,\delta,\rho) \subset A_n \cup A'_n, \text{ or equivalently } B_n(f,\delta,\rho)^c \supset (A_n)^c \cap (A'_n)^c. \tag{5.45}$$

By definition, $\{\rho n \le |N_n^+|\}^c \subseteq B_n(f,\delta,\rho)^c$. Therefore, (5.45) is implied by

$$(A_n)^c \cap (A'_n)^c \cap \{\rho n \le |N_n^+|\} \subseteq B_n(f,\delta,\rho)^c. \tag{5.46}$$

To prove this, we assume that $(A_n)^c \cap (A'_n)^c \cap \{\rho n \le |N_n^+|\}$ holds. For $M \le |N_n^+|$, let

$$k(M) := \inf\Big\{\sum_{i=1}^{\ell} |N_n^i| : \ell \in \{0,\dots,|N_{n_0}^+|\} \text{ such that } \sum_{i=1}^{\ell}|N_n^i| \ge M\Big\} \ge M. \tag{5.47}$$

On $(A_n)^c$, $\big||N_n^i|^{-1}\sum_{v\in N_n^i} f(\varphi_v) - \langle\chi,f\rangle/\langle\chi,1\rangle\big| \le \delta/2$ for all $i = 1,\dots,|N_{n_0}^+|$. Therefore, for all M with $\rho n \le M \le |N_n^+|$

$$\left|\frac{\sum_{v\in N_{n,k(M)}^+} f(\varphi_v)}{k(M)} - \frac{\langle\chi,f\rangle}{\langle\chi,1\rangle}\right| \le \frac{\delta}{2}. \tag{5.48}$$

Writing $\sum_{v\in N_{n,M}^+} f(\varphi_v) = \sum_{v\in N_{n,k(M)}^+} f(\varphi_v) - \sum_{v\in N_{n,k(M)}^+\setminus N_{n,M}^+} f(\varphi_v)$, and using that the absolute value of the second sum is bounded by $(k(M)-M)\sup|f|$, we obtain that for every $M \ge \rho n$

$$\begin{aligned} &\left|\frac{\sum_{v\in N_{n,M}^+} f(\varphi_v)}{M} - \frac{\sum_{v\in N_{n,k(M)}^+} f(\varphi_v)}{k(M)}\right| \\ &= \left|\frac{\sum_{v\in N_{n,k(M)}^+} f(\varphi_v)}{M} - \frac{\sum_{v\in N_{n,k(M)}^+\setminus N_{n,M}^+} f(\varphi_v)}{M} - \frac{\sum_{v\in N_{n,k(M)}^+} f(\varphi_v)}{k(M)}\right| \\ &\le \left|\sum_{v\in N_{n,k(M)}^+} f(\varphi_v)\Big(\frac{1}{M} - \frac{1}{k(M)}\Big)\right| + \frac{k(M)-M}{M}\sup|f| \\ &= \frac{k(M)-M}{M}\left(\left|\frac{\sum_{v\in N_{n,k(M)}^+} f(\varphi_v)}{k(M)}\right| + \sup|f|\right) \le c(f,\delta)\frac{q}{\rho n}, \end{aligned} \tag{5.49}$$

where in the last inequality we used (5.48) and the fact that $k(M) - M \le q$ on $(A'_n)^c$. If the right-hand side of (5.49) is smaller than $\delta/2$, then together with (5.48), this implies that $B_n(f,\delta,\rho)^c$ holds, implying (5.46) and thus (5.45).

To finish the proof of (5.32), we must choose $n_0 = n_0(n)$ and $q = q(n)$ so that (5.39) applies and the left-hand sides of (5.43), (5.49) tend to zero. This is easily done by setting, e.g., $q(n) = n^{3/4}$ and $n_0 = n - n^{1/4}$. □

6. The S_n martingale

The remaining three sections of this paper are dedicated to proving Theorem 2.3. In this section, we will introduce a martingale based on the depth-first traversal of a sequence of copies of $\mathcal{C}_o \cap \mathbb{T}^+$, and prove that it satisfies an invariance principle, see Proposition 6.1. This martingale can be seen as an analogue to the Lukasiewicz path used to study critical Galton-Watson trees. The second part of the section then demonstrates another two scaling limit results, Proposition 6.8 and Proposition 6.10 which are both consequences of Proposition 6.1.

To define the martingale, we consider an i.i.d. sequence $\mathbf{T} = ((T^1,\varphi^1),(T^2,\varphi^2),\dots)$ where every (T^i,φ^i) is distributed as $(\mathcal{C}_o \cap \mathbb{T}^+, \varphi|_{\mathcal{C}_o\cap\mathbb{T}^+})$ under P_x. To keep the notation simple, we keep using P_x for the probability measure associated with the whole sequence $\mathbf{T}$, and for $v \in T^i$ we write φ_v instead of φ^i_v. We further use o^i to denote the root of T^i, and set $N_n^{+,i} = \{v \in T^i : d(o^i, v) = n\}$. Note that $|N_n^{+,i}|$ has the same distribution as $|N_n^+|$ which was studied in detail in the previous sections. In particular, we know that all T^i are a.s. finite.

Throughout this section, we use the notation for the parent $p(v)$, direct descendants $\mathrm{desc}(v)$, and siblings $\mathrm{sib}(v)$ of $v \in T^i$ relative to the rooted tree T^i (not $\mathbb{T}$), $w \preceq v$ means that w is an ancestor of v, cf. below (3.1).

We now describe the depth-first traversal $\mathbf{v} = (v_1, v_2, \dots)$ of $\mathbf{T}$. It starts at the root of T^1, that is $v_1 = o^1$, and then explores the tree T^1 in a depth-first manner. After visiting all vertices of T^1, it proceeds to o^2, explores T^2 in a depth-first manner, and so forth. v_i denotes the i-th vertex visited during this traversal. The notation $v <_{\mathbf{v}} w$ indicates that v precedes w in $\mathbf{v}$ (with $\le_{\mathbf{v}}$ representing the reflexive version). For any $v \in \cup_i T^i$ we write $\Lambda(v)$ to denote the index of the tree which v belongs to (that is $v \in T^{\Lambda(v)}$). We define

$$\Lambda_n = \Lambda(v_n) \quad \text{and} \quad H_n = |v_n| = d(v_n, o^{\Lambda_n}), \tag{6.1}$$

that is Λ_n is the index of the tree which is explored at step n, and H_n is the "height" of the n-th explored vertex. For every $v \in \cup_i T^i$, we define the set $Y(v)$ as the union of all siblings of some ancestor of v that appear later in $\mathbf{v}$ than this ancestor itself, that is,

$$Y(v) = \bigcup_{w \preceq v} \{u \in \mathrm{sib}(w) : w <_{\mathbf{v}} u\}. \tag{6.2}$$

With this notation we can introduce the key object of this section, the process

$$S_n = \chi(\varphi_{v_n}) - \sum_{i \le \Lambda_n} \chi(\varphi_{o^i}) + \sum_{w \in Y(v_n)} \chi(\varphi_w), \qquad n \in \mathbb{N}, \tag{6.3}$$

which is adapted to the filtration

$$\mathcal{H}_n = \sigma\Big((v,\varphi_v) : v \in \bigcup_{i=1}^n \mathrm{sib}(v_i)\Big), \qquad n \in \mathbb{N}. \tag{6.4}$$

The next proposition shows that S_n converges to a Brownian motion with variance

$$\sigma^2 = \langle \chi, \mathcal{V} \rangle / \langle \chi, 1 \rangle, \tag{6.5}$$

where

$$\mathcal{V}(x) := P_x\Big[\Big(\sum_{w\in N_1^+} \chi(\varphi_w)\Big)^2\Big] - P_x\Big[\sum_{w\in N_1^+} \chi(\varphi_w)\Big]^2 = \mathrm{Var}_x\Big(\sum_{w\in N_1^+} \chi(\varphi_w)\Big). \tag{6.6}$$

Here and below $(B_t)_{t\ge 0}$ denotes the standard Brownian motion.

Proposition 6.1. *For every $x \ge h^*$, in P_x-distribution w.r.t. the Skorokhod topology,*

$$\lim_{n\to\infty} \Big(\frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}\Big)_{t\ge 0} = (\sigma B_t)_{t\ge 0}. \tag{6.7}$$

We need a few preparatory steps to show this proposition. We start by proving that S is a martingale.

Lemma 6.2. *The process S is a $\mathcal{H}$-martingale under P_x.*

Proof. We first show that

$$S_{k+1} - S_k = -\chi(\varphi_{v_k}) + \sum_{w\in \mathrm{desc}(v_k)} \chi(\varphi_w). \tag{6.8}$$

To this end we need to distinguish between several possible scenarios:

(1) If v_k is not a leaf, that is $\mathrm{desc}(v_k) \ne \emptyset$, then v_{k+1} is the first child of v_k, and thus $Y(v_{k+1}) = (\mathrm{desc}(v_k) \setminus \{v_{k+1}\}) \cup Y(v_k)$ where the union is disjoint. From this (6.8) follows.

(2) If v_k is a leaf, that is $\mathrm{desc}(v_k) = \emptyset$, then there are three possible cases for v_{k+1}: Either it is the next sibling (with respect to $<_{\mathbf{v}}$) of v_k, or the next sibling of some ancestor of v_k, or it is the root of the next tree in the tree sequence. In the first two cases, $Y(v_k) = Y(v_{k+1}) \cup \{v_{k+1}\}$ where the union is disjoint, and thus $S_{k+1} - S_k = -\chi(\varphi_{v_k})$ and thus (6.8) holds. In the last case, when v_{k+1} is the root of the next tree, both sets $Y(v_k)$ and $Y(v_{k+1})$ must be empty and $\Lambda_{k+1} = \Lambda_k + 1$, which implies (6.8) also in this case.

From (6.8) and the branching process properties of φ it follows that

$$\begin{aligned} E_x[S_{k+1} - S_k | \mathcal{H}_k] &= E_x\Big[-\chi(\varphi_{v_k}) + \sum_{w\in \mathrm{desc}(v_k)} \chi(\varphi_w)\Big|\mathcal{H}_k\Big] \\ &= -\chi(\varphi_{v_k}) + E_{\varphi_{v_k}}\Big[\sum_{w\in N_1^+} \chi(\varphi_w)\Big] = -\chi(\varphi_{v_k}) + L\chi(\varphi_{v_k}) = 0, \end{aligned} \tag{6.9}$$

since χ is the eigenfunction of L (see (3.12)). This finishes the proof. □

The next four simple lemmas will be used to control the quadratic variation of S.

Lemma 6.3. *For every $k \ge 1$, with $\mathcal{V}$ as in (6.6),*

$$E_x[(S_{k+1} - S_k)^2 | \mathcal{H}_k] = \mathcal{V}(\varphi_{v_k}). \tag{6.10}$$

Proof. By (6.8) the conditional expectation $E_x[(S_{k+1} - S_k)^2 | \mathcal{H}_k]$ equals

$$\begin{aligned} &\chi(\varphi_{v_k})^2 - 2\chi(\varphi_{v_k}) E_x\Big[\sum_{w\in \mathrm{desc}(v_k)} \chi(\varphi_w)\Big|\mathcal{H}_k\Big] + E_x\Big[\Big(\sum_{w\in \mathrm{desc}(v_k)} \chi(\varphi_w)\Big)^2\Big|\mathcal{H}_k\Big] \\ &= \chi(\varphi_{v_k})^2 - 2\chi(\varphi_{v_k}) E_{\varphi_{v_k}}\Big[\sum_{w\in N_1^+} \chi(\varphi_w)\Big] + E_{\varphi_{v_k}}\Big[\Big(\sum_{w\in N_1^+} \chi(\varphi_w)\Big)^2\Big]. \end{aligned} \tag{6.11}$$

From this the statement follows using (3.12) again. □

Lemma 6.4. *There is $c < \infty$ such that $\mathcal{V}(x) < c$ for all $x \geq h^*$.*

Proof. Let $\{w_1, \dots, w_d\}$ be the children of the root in $\mathbb{T}^+$. Since $(\varphi_{w_i})_{i=1}^d$ are independent under P_x and $\chi(x) = 0$ for $x < h^*$, $\mathrm{Var}_x(\sum_{v \in N_1^+} \chi(\varphi_v)) = \mathrm{Var}_x(\sum_{i=1}^d \chi(\varphi_{w_i})) = \sum_{i=1}^d \mathrm{Var}_x(\chi(\varphi_{w_i})) = d\,\mathrm{Var}_Y(\chi(Y + x/d))$, where $Y \sim \mathcal{N}(0, \sigma_Y^2)$, see (3.4). Since χ is Lipschitz on $[h^*, \infty)$ by (3.19), the statement follows. □

To state the next result, let U_n be a random variable distributed uniformly on $\{1, \dots, n\}$, defined on the same probability space as the sequence $\mathbf{T}$, independent of $\mathbf{T}$. We write

$$H_n^* = H_{U_n} \quad \text{and} \quad \Lambda_n^* = \Lambda_{U_n}, \tag{6.12}$$

for the height and the tree index of a vertex chosen uniformly amongst the first n explored vertices.

Lemma 6.5. *For all $x \geq h^*$,*

$$\lim_{u \to \infty} \sup_{n \geq 1} P_x[H_n^* \geq u\sqrt{n}] = 0. \tag{6.13}$$

Proof. By Theorem 2.1 and Proposition 3.3, there is $c(x) < \infty$ such that for all $n \geq 1$,

$$P_x[N_n^+ \neq \emptyset] \leq \frac{c(x)}{n} \qquad \text{and} \qquad E_x[\Lambda_n] \leq c(x)\sqrt{n}. \tag{6.14}$$

To see the second inequality in (6.14), we note that by the independence of the trees T^i and Proposition 3.3, $P_x[\Lambda_n \geq k] \leq P_x[|T^i| < n, i = 1, \dots, k-1] = P_x[|T^1| < n]^{k-1} \leq (1 - c(x)n^{-1/2})^{k-1}$ and thus $E_x[\Lambda_n] = \sum_{k=1}^\infty P_x[\Lambda_n \geq k] \leq c(x)^{-1} n^{1/2}$ as claimed.

Since $\{H_n^* \geq u\sqrt{n}\} \subseteq \{\exists i \leq \Lambda_n : N_{\lceil u\sqrt{n} \rceil}^{+,i} \neq \emptyset\}$, by a union bound,

$$P_x[H_n^* \geq u\sqrt{n}] \leq \sum_{i \geq 1} P_x\big[i \leq \Lambda_n, N_{\lceil u\sqrt{n} \rceil}^{+,i} \neq \emptyset\big]. \tag{6.15}$$

The events $\{i \leq \Lambda_n\}$ and $\{|N_{\lceil u\sqrt{n} \rceil}^{+,i}| > 0\}$ are independent. Hence, using the first half of (6.14), this is bounded by $c(x)u^{-1}n^{-1/2}\sum_{i \geq 1} P_x[i \leq \Lambda_n]$. Since $\sum_{i \geq 1} P_x[i \leq \Lambda_n] = E_x[\Lambda_n] \leq c(x)\sqrt{n}$, by the second half of (6.14), the claim follows. □

Lemma 6.6. *For every $\delta > 0$ and $x \geq h^*$ there exists a sequence $R(n)$ such that $\lim_{n \to \infty} R(n) = \infty$ and*

$$\sup_{n \geq 1} P_x[H_n^* \leq R(n)] \leq \delta. \tag{6.16}$$

Proof. We will show that for every $K < \infty$

$$\lim_{n \to \infty} P_x[H_n^* \leq K] = 0, \tag{6.17}$$

which implies the statement of the lemma.

For fixed $K > 0$, by definition of H_n^*,

$$P_x[H_n^* \leq K] = n^{-1} \sum_{l=1}^n P_x[H_l \leq K] = n^{-1} E_x\Big[\sum_{h=0}^K \sum_{l=1}^n 1_{H_l = h}\Big]. \tag{6.18}$$

Since $\sum_{l=1}^n 1_{H_l = h} \leq \sum_{i=1}^{\Lambda_n} |N_h^{+,i}|$ and thus $\sum_{h=0}^K \sum_{l=1}^n 1_{H_l = h} \leq \sum_{i=1}^{\Lambda_n} \sum_{h=0}^K |N_h^{+,i}|$. Moreover, $(\sum_{h=0}^K |N_h^{+,i}|)_{i=1}^{\Lambda_n}$ are i.i.d. and with finite mean. Noting that Λ_n is a stopping time with respect to $\mathcal{G}_i = \sigma(T^j : j \le i)$ and using Wald's equation (see e.g. [Dur19, Theorem 2.6.2]),

$$P_x[H_n^* \le K] \le \frac{1}{n} E_x\Big[\sum_{i=1}^{\Lambda_n}\sum_{h=0}^{K} |N_h^{+,i}|\Big] = \frac{1}{n} E_x[\Lambda_n] E_x\Big[\sum_{h=0}^{K} |N_h^{+,1}|\Big], \tag{6.19}$$

which by (6.14) converges to 0 as $n \to \infty$. □

The following law of large numbers will be important for the proof of Proposition 6.1.

Proposition 6.7. *Let f be a bounded function and $m_n^f := n^{-1}\sum_{i=1}^n f(\varphi_{v_i})$. Then*

$$\lim_{n\to\infty} m_n^f = m_\infty^f := \frac{\langle \chi, f\rangle}{\langle \chi, 1\rangle}, \quad \textit{in } P_x\textit{-probability.} \tag{6.20}$$

Proof. Let

$$N_{k,n}^{+,i} = N_k^{+,i} \cap \{v_1, \dots, v_n\} \tag{6.21}$$

be the part of $N_k^{+,i}$ traversed in the first n steps, and let $N_n^* = N_{H_n^*,n}^{+,\Lambda_n^*}$. We set

$$m_n^* = |N_n^*|^{-1} \sum_{v \in N_n^*} f(\varphi_v). \tag{6.22}$$

Denoting by $\mathcal{F}_{\mathbf{T}}$ the σ-algebra generated by the sequence $\mathbf{T} = (T_i, \varphi^i)_{i\ge 1}$, it holds that

$$m_n^f = E_x[m_n^* | \mathcal{F}_{\mathbf{T}}] \tag{6.23}$$

(effectively, the expectation here is over U_n only). To show the proposition it will thus be sufficient to show that

$$\lim_{n\to\infty} m_n^* = m_\infty^f \quad \text{in } P_x\text{-probability.} \tag{6.24}$$

To see that this is indeed enough, note that since f is bounded m_n^* is dominated by a constant. Hence, by the dominated convergence theorem, $m_n^* \to m_\infty^f$ also in $L^2(P_x)$. By (6.23), since the conditional expectation is a contraction on $L^2(P_x)$, $m_n^f \to m_\infty^f$ in $L^2(P_x)$, and thus in P_x-probability as claimed.

To prove (6.24), we fix $\varepsilon, \delta > 0$ and set

$$A_n = \{|m_n^* - m_\infty^f| > \varepsilon\}. \tag{6.25}$$

We then fix $C < \infty$ such that $P_x[H_n^* \ge Cn^{1/2}] \le \delta/3$ for all $n \ge 1$, which is possible by Lemma 6.5, and take $R(n)$ as in Lemma 6.6, so that $P_x[H_n^* \le R(n)] \le \delta/3$. Setting $B_n = \{R(n) \le H_n^* \le Cn^{1/2}\}$, it follows that for all $n \ge 1$,

$$P_x[A_n] \le 2\delta/3 + P_x[A_n \cap B_n] \tag{6.26}$$

We further set $m^i_{h,n} = |N^{+,i}_{h,n}|^{-1} \sum_{v\in N^{+,i}_{h,n}} f(\varphi_v)$ and $A^i_{h,n} = \{|m^i_{h,n} - m^f_\infty| > \varepsilon\}$. Using the definitions of U_n and m^*_n, and then decomposing by possible values of H_k, Λ_k, we obtain

$$\begin{aligned} P_x[A_n \cap B_n] &= \sum_{k=1}^{n} P_x[A_n \cap B_n | U_n = k] P_x[U_n = k] \\ &= \sum_{k=1}^{n} \frac{1}{n} P_x[A^{\Lambda_k}_{H_k,n}, R(n) \le H_k \le Cn^{1/2}] \\ &= \frac{1}{n} \sum_{l=R(n)}^{Cn^{1/2}} \sum_{i=1}^{n} E_x\Big[1_{A^i_{l,n}} \sum_{k=1}^{n} 1_{\{H_k=l,\Lambda_k=i\}}\Big]. \\ &= \frac{1}{n} \sum_{l=R(n)}^{Cn^{1/2}} \sum_{i=1}^{n} E_x\big[1_{A^i_{l,n}} |N^{+,i}_{l,n}|\big], \end{aligned} \tag{6.27}$$

where in the last step we used $\sum_{k=1}^n 1_{\{H_k=l,\Lambda_k=i\}} = |N^{+,i}_{l,n}|$. To analyse the expectation on the right-hand side, we define σ-algebras $\mathcal{G}_i = \sigma(T^1, T^2, \dots, T^i)$ and set $\tau_i = |T^1| + \dots + |T^{i-1}|$ (which is $\mathcal{G}_{i-1}$-measurable). Then, using that $(T^i, \varphi^i)_{i\ge1}$ is an i.i.d. sequence,

$$\begin{aligned} E_x\big[1_{A^i_{l,n}} |N^{+,i}_{l,n}| \big| \mathcal{G}_{i-1}\big] &= 1_{\{\tau_i \le n\}} E_x\big[1_{A^1_{l,n-\tau_i}} \big|N^{+,1}_{l,n-\tau_i}\big|\big] \\ &= 1_{\{\tau_i \le n\}} E_x\big[1_{A_{l,n-\tau_i}} |N^+_{l,n-\tau_i}| \big| N^+_l \ne \emptyset\big] P_x[N^+_l \ne \emptyset], \end{aligned} \tag{6.28}$$

where on the last line we omitted the superscript '1' since $N^{+,1}_{l,k}$ has the same distribution as $N^+_{l,k}$. By Theorem 2.1, Proposition 3.8 and (6.14) there is a $c = c(x)$ such that

$$P_x\big[N^+_l \ne \emptyset\big] \le cl^{-1}, \qquad E_x\big[\big|N^+_l\big|^2 \,\big|\, N^+_l \ne \emptyset\big]^{1/2} \le cl, \qquad E_x\big[\Lambda_n\big] \le cn^{1/2}. \tag{6.29}$$

Further, for $B_n(f, \varepsilon, \rho)$ as in Lemma 5.5, it holds that $A_{l,k} \cap \{|N^+_{l,k}| \ge l\rho\} \subseteq B_l(f, \varepsilon, \rho)$. Therefore, choosing $\rho = \delta/(6Cc^2)$ and decomposing on whether $|N^+_{l,n-\tau_i}|$ is larger than $l\delta/(6Cc^2)$, we obtain

$$\begin{aligned} &E_x\big[1_{A^1_{l,n-\tau_i}} |N^{+,1}_{l,n-\tau_i}| \big| N^{+,1}_l \ne \emptyset\big] \\ &\le \frac{l\delta}{6Cc^2} + E_x\big[1_{B_n(f,\varepsilon,\rho)} |N^+_{l,n-\tau_i}| \big| N^+_l \ne \emptyset\big] \\ &\le \frac{l\delta}{6Cc^2} + E_x\big[|N^+_l|^2 \big| N^+_l \ne \emptyset\big]^{1/2} P_x\big[B_l(f, \varepsilon, \rho) \big| N^+_l \ne \emptyset\big]^{1/2}, \end{aligned} \tag{6.30}$$

where in the last step we used the Cauchy–Schwarz inequality and $|N^+_{l,n-\tau_i}| \le |N^+_l|$. Combining (6.28) and (6.30) with the first two claims in (6.29) we obtain

$$E_x\big[1_{A^i_{l,n}} |N^{+,i}_{l,n}|\big] \le \Big(\frac{\delta}{6Cc} + c^2 P_x\big[B_l(f, \varepsilon, \rho) \,\big|\, N^+_l \ne \emptyset\big]^{1/2}\Big) P_x[\tau_i \le n]. \tag{6.31}$$

Inserting this into (6.26), (6.27), we obtain that $P_x[A_n]$ is bounded by

$$\frac{2\delta}{3} + \frac{1}{n}\Big(\frac{\delta}{6Cc} + c^2 \sup_{l \ge R(n)} P_x\big[B_l(f, \varepsilon, \rho) \,\big|\, N^+_l \ne \emptyset\big]^{1/2}\Big) \sum_{l=R(n)}^{\lfloor C\sqrt{n}\rfloor} \sum_{i=1}^{n} P_x[\Lambda_n \ge i], \tag{6.32}$$

Since $R(n)$ diverges, Lemma 5.5 implies that the supremum in this formula tends to zero as $n \to \infty$. By the last property in (6.29), $\sum_{l=R(n)}^{\lfloor C\sqrt{n}\rfloor} \sum_{i=1}^{n} P_x[\Lambda_n \ge i] \le C\sqrt{n} E_x[\Lambda_n] \le Ccn$.

Therefore

$$P_x[A_n] \le \frac{2\delta}{3} + \Big(\frac{\delta}{6} + Cc^3 o(1)\Big) \le \delta \tag{6.33}$$

for n large enough. Since ε and δ are arbitrary, this shows (6.24) and with the initial comment completes the proof. □

We are now ready to prove Proposition 6.1.

Proof of Proposition 6.1. We apply a martingale functional central limit theorem, see e.g. [Dur19, Theorem 8.2.8]. To check its assumptions we need to show that

(a) $\lim_{n\to\infty} n^{-1}\sum_{k=1}^n E_x[(S_k - S_{k-1})^2|\mathcal{H}_{k-1}] = \frac{\langle\chi,\mathcal{V}\rangle}{\langle\chi,1\rangle}$ in P_x-probability,

(b) $\lim_{n\to\infty} n^{-1}\sum_{k=1}^n E_x[(S_k - S_{k-1})^2 1_{\{|S_k - S_{k-1}|>\varepsilon\sqrt{n}\}}] = 0$.

Condition (a) follows from Lemma 6.3 and Proposition 6.7, using that $\mathcal{V}$ is bounded by Lemma 6.4.

We now show (b). By the Cauchy–Schwarz inequality, $E_x[(S_k-S_{k-1})^2 1_{\{|S_k-S_{k-1}|>\varepsilon\sqrt{n}\}}] \le E_x[(S_k - S_{k-1})^4]^{1/2} P_x[(S_k - S_{k-1})^2 > \varepsilon^2 n]^{1/2}$. By Lemmas 6.3, 6.4, $E_x[(S_k - S_{k-1})^2] \le c$ uniformly in k and thus $P_x[(S_k-S_{k-1})^2 > \varepsilon^2 n] \le c\varepsilon^{-2}n^{-1}$. Thus, to prove (b) it is enough to show

$$E_x[(S_k - S_{k-1})^4] \le c \tag{6.34}$$

uniformly in k. By (6.8), $S_k - S_{k-1} = \sum_{v\in\text{desc}(v_{k-1})}\chi(\varphi_v) - \chi(\varphi_{v_{k-1}})$, and thus

$$\begin{aligned} E_x[(S_k - S_{k-1})^4|\mathcal{H}_{k-1}] &= E_{\varphi_{v_{k-1}}}\Big[\Big(\sum_{v\in N_1^+}\chi(\varphi_v) - \chi(\varphi_{v_{k-1}})\Big)^4\Big] \\ &= E_{\varphi_{v_{k-1}}}\Big[\Big(\sum_{v\in N_1^+}\chi(\varphi_v) - E_{\varphi_{v_{k-1}}}\big[\sum_{v\in N_1^+}\chi(\varphi_v)\big]\Big)^4\Big], \end{aligned} \tag{6.35}$$

where we used that $\chi(\varphi_{v_{k-1}}) = E_{\varphi_{v_{k-1}}}[\sum_{v\in N_1^+}\chi(\varphi_v)]$ in the last line. We write $N_1^+ = \{w_1,\dots,w_d\}$ and follow similar arguments as in the proof of Lemma 6.4, using that $\chi(x) = 0$ for $x < h^*$ and exploiting the independence of the φ_{w_i}, to obtain that (6.35) is bounded above by $c_d E_{\varphi_{v_{k-1}}}[(\chi(\varphi_{w_1}) - E_{\varphi_{v_{k-1}}}[\chi(\varphi_{w_1})])^4]$. Since χ is Lipschitz on $[h^*,\infty)$, $\chi(\varphi_{w_1}) - E_x[\chi(\varphi_{w_1})]$ is sub-Gaussian, and thus $E_x[(\chi(\varphi_{w_1}) - E_x[\chi(\varphi_{w_1})])^4] \le c$, where the constant c only depends on d and the Lipschitz constant of χ but not on x. Statement (6.34) then follows by taking the expectation of (6.35). □

6.1. Further invariance principles.

We now prove several further convergence results that are a consequence of Proposition 6.1 and which will be useful later.

Proposition 6.8. *Let B_t be a standard Brownian motion, let L_t^0 be its local time at 0. Then, as $n\to\infty$, in P_x-distribution with respect to the Skorokhod topology,*

$$\Big(\frac{\sum_{w\in Y(v_{\lfloor nt\rfloor})}\chi(\varphi_w)}{\sqrt{n}}, \frac{\Lambda(v_{\lfloor nt\rfloor})}{\sqrt{n}}\Big)_{t\ge 0} \to \Big(\sigma|B_t|, \frac{\sigma}{\chi(x)}L_t^0\Big)_{t\ge 0}. \tag{6.36}$$

To show Proposition 6.8 from Proposition 6.1, we need the following lemma. It states that (under scaling) the first summand in the definition (6.3) of S_n is negligible.

Lemma 6.9. *In P_x-distribution with respect to the Skorokhod topology,*

$$\lim_{n\to\infty}\Big(\frac{\chi(\varphi_{v_{\lfloor nt\rfloor}})}{\sqrt{n}}\Big)_{t\ge 0} = 0. \tag{6.37}$$

Proof. As χ grows at most linearly (see (3.18)), it is enough to show that for any fixed $t > 0$ and $\varepsilon > 0$,

$$\lim_{n\to\infty} P_x\Big[\max_{i\le nt} \varphi_{v_i}/\sqrt{n} > \varepsilon\Big] = 0. \tag{6.38}$$

Let $H(T^i) = \max_{v\in T^i} H(v)$ denote the height of tree T^i. For every $j, k \in \mathbb{N}$, the probability in (6.38) is bounded from above by

$$\begin{aligned} &P_x\Big[\max_{i\le nt} \varphi_{v_i} > \varepsilon\sqrt{n},\ \Lambda_{\lfloor nt\rfloor} \le j,\ H(T_i) \le k \text{ for } i = 1,\dots,j\Big] \\ &\quad + P_x[\exists i \in \{1,\dots,j\} : H(T^i) > k,\ \Lambda_{\lfloor nt\rfloor} \le j] + P_x[\Lambda_{\lfloor nt\rfloor} > j]. \end{aligned} \tag{6.39}$$

To show (6.38), we thus need to choose $j_n \to \infty$ and $k_n \to \infty$ so that all three summands converge to zero.

We start with the first summand. Restricted to $\{\Lambda_{\lfloor nt\rfloor} \le j\}$ and $\{H(T^i) \le k$ for $i = 1,\dots,j\}$, $\max_{k\le nt} \varphi_{v_k}$ is dominated by the maximum of all φ_v with v such that $\Lambda(v) \le j$ and $H(v) \le k$. Considering not only the maximum over the connected components of the level set, but over the first k generations in j copies of the whole forward tree $\mathbb{T}^+$, this is dominated by the maximum of jd^k non-negatively correlated Gaussian random variables with mean at most x and bounded variance. By Gaussian comparison techniques, the mean of this maximum is of order $x + c\sqrt{\log jd^k} \le c\sqrt{k\log j}$ for j, k large enough. Thus, by the Markov inequality,

$$P_x\Big[\max_{i\le nt} \varphi_{v_i} > \varepsilon\sqrt{n},\ \Lambda_{\lfloor nt\rfloor} \le j,\ H(T_i) \le k \text{ for } i = 1,\dots,j\Big] \le \frac{C}{\varepsilon n^{1/2}}\sqrt{k\log(j)}. \tag{6.40}$$

By a union bound, using Theorem 2.1, the second summand in (6.39) satisfies

$$P_x[\exists i \in \{1,\dots,j\} : H(T^i) > k,\ \Lambda_{\lfloor nt\rfloor} \le j] \le c(x)j/k. \tag{6.41}$$

Finally, by the Markov inequality and (6.14), the third summand can be bounded by

$$P_x[\Lambda_{\lfloor nt\rfloor} > j] \le c(x)\sqrt{tn}/j. \tag{6.42}$$

Setting now, e.g., $j_n = n^{2/3}$ and $k_n = n^{5/6}$, all three summands in (6.39) converge to 0 as required. □

Proof of Proposition 6.8. We use the continuous mapping theorem together with the already established convergence statements. By Proposition 6.1, $n^{-1/2}S_{\lfloor n\cdot\rfloor} \to \sigma B_\cdot$, and by Lemma 6.9, $n^{-1/2}\chi(\varphi_{v_{\lfloor n\cdot\rfloor}}) \to 0$ as $n \to \infty$, in P_x-distribution, in the Skorokhod topology. Therefore

$$n^{-1/2}(S_{\lfloor n\cdot\rfloor} - \chi(\varphi_{v_{\lfloor n\cdot\rfloor}})) = n^{-1/2}\Big(\sum_{w\in Y(v_{\lfloor n\cdot\rfloor})} \chi(\varphi_w) - \sum_{i\le\Lambda_{\lfloor n\cdot\rfloor}} \chi(\varphi_{o^i})\Big) \tag{6.43}$$

also converges to $\sigma B_\cdot$ as $n \to \infty$. Next, note that

$$\inf_{k\le n}\{S_k - \chi(\varphi_{v_k})\} = \inf_{k\le n}\Big\{\sum_{w\in Y(v_k)} \chi(\varphi_w) - \sum_{i\le\Lambda_k} \chi(\varphi_{o^i})\Big\} = -\sum_{i\le\Lambda_n} \chi(\varphi_{o^i}). \tag{6.44}$$

Therefore, setting $\underline{B}_t = \inf_{s\le t} B_s$ and using the continuous mapping theorem with the map $g(X_t) = (X_t - \inf_{s\le t} X_s, -\inf_{s\le t} X_s)$,

$$n^{-1/2}\Big(\sum_{w\in Y(v_{\lfloor n\cdot\rfloor})} \chi(\varphi_w), \sum_{i\le\Lambda_{\lfloor n\cdot\rfloor}} \chi(\varphi_{o^i})\Big) \to \big(\sigma(B_\cdot - \underline{B}_\cdot), -\sigma\underline{B}_\cdot\big) \tag{6.45}$$

in distribution as $n \to \infty$. By Lévy's theorem (see, e.g., Theorem VI.2.3 in [RY99]), the right-hand side of (6.45) has the same distribution as $(\sigma|B_\cdot|, \sigma L^0_\cdot(B))$. Together with the fact that under P_x, $\sum_{i \le \Lambda_{\lfloor nt \rfloor}} \chi(\varphi_{o^i}) = \chi(x)\Lambda_{\lfloor nt \rfloor}$, this implies the proposition. □

Another consequence of Proposition 6.1 is the following scaling limit result for S_n conditioned to reach a certain height on the first tree T_1. To this end we define

$$\bar{S}_n = \begin{cases} \sum_{w \in Y(v_n)} \chi(\varphi_w) & \text{if } n \le |T^1|, \\ 0 & \text{if } n > |T^1|. \end{cases} \tag{6.46}$$

Note that, up to an additive correction $\chi(\varphi_{v_n}) - \chi(x)$, $\bar{S}_\cdot$ is equal to $S_\cdot$ restricted to T_1.

Proposition 6.10. *For $y > 0$, let $(\mathbf{e}^{\ge y/\sigma})_{t \ge 0}$ be a Brownian excursion conditioned to reach at least height y/σ. Then, in distribution under $P_x[\,\cdot\,|\sup_k \bar{S}_k \ge \sqrt{n}y]$, with respect to the Skorokhod topology,*

$$\lim_{n \to \infty} \big(n^{-1/2} \bar{S}_{\lfloor nt \rfloor}\big)_{t \ge 0} = (\sigma \mathbf{e}^{\ge y/\sigma})_{t \ge 0}. \tag{6.47}$$

We omit the proof of this proposition as it would be identical to the proof of Proposition 6.13 in [Pow19], which itself follows [DLG02, Proposition 2.5.2].

7. Relation of S_n and H_n

The aim of this section will be to establish a connection between the martingale S_n and the height process H_n (see (6.3) and (6.1) for definitions). This will be useful in the proof of Theorem 2.3 in Section 8. Throughout the whole section, we only consider the field on the first tree $(T, \varphi) = (T^1, \varphi^1)$ in the infinite tree sequence $\mathbf{T} = ((T^1, \varphi^1), (T^2, \varphi^2), \dots)$ introduced in Section 6. As a consequence, we only consider $\bar{S}$ introduced in (6.46) instead of S. We will see that, approximately, $\bar{S}(v) \approx H(v)/C_1$ with C_1 as in (2.11).

Motivated by this, for $\eta > 0$ we say that $v \in T$ is *η-bad* if

$$\Big|\frac{\bar{S}(v)}{H(v)} - C_1^{-1}\Big| > \eta. \tag{7.1}$$

Fixing in addition $R > 0$, we say that $v \in T$ is *(η, R)-bad* if there exists a $w \prec v$ such that $H(w) \ge R$ and w is η-bad. This means that v is (η, R)-good (i.e. not (η, R)-bad) if all its ancestors in generations at least R are η-good. We set

$$N_n^{(\eta,R)} = \{v \in N_n^+ :\ v \text{ is } (\eta, R)\text{-bad}\}. \tag{7.2}$$

The first main result of this section shows that this set is relatively small.

Proposition 7.1. *For every $\varepsilon > 0$ and $x \ge h^*$,*

$$\lim_{R \to \infty} \sup_{n \ge R} P_x\Big[\frac{|N_n^{(\eta,R)}|}{|N_n^+|} > \varepsilon \,\Big|\, N_n^+ \ne \emptyset\Big] = 0. \tag{7.3}$$

Proof. The proof follows a strategy similar to the proof of Proposition 6.17 in [Pow19], with adaptations that are necessary in order to handle the unbounded domain in our setting. For $\varepsilon > 0$, $R > 0$ and $n \ge 0$ we define the event

$$E^\varepsilon_{R,n} = \Big\{\frac{\sum_{v \in N_n^+} \chi(\varphi_v) 1_{\{v \text{ is } (\eta, R)\text{-bad}\}}}{\sum_{v \in N_n^+} \chi(\varphi_v)} > \varepsilon\Big\}. \tag{7.4}$$

We first claim that it suffices to show that

$$\lim_{R \to \infty} \sup_{n \ge R} P_x[E^\varepsilon_{R,n} | N_n^+ \ne \emptyset] = 0. \tag{7.5}$$

To see that this indeed implies the lemma, we write

$$\frac{|N_n^{(\eta,R)}|}{|N_n^+|} = \frac{|N_n^{(\eta,R)}|}{\sum_{v\in N_n^+}\chi(\varphi_v)}\cdot\frac{\sum_{v\in N_n^+}\chi(\varphi_v)}{|N_n^+|}. \tag{7.6}$$

Therefore, for every $\delta>0$, the probability in (7.3) is bounded from above by

$$P_x\Big[\frac{|N_n^{(\eta,R)}|}{\sum_{v\in N_n^+}\chi(\varphi_v)}>\delta\,\Big|\,N_n^+\neq\emptyset\Big]+P_x\Big[\frac{|N_n^+|}{\sum_{v\in N_n^+}\chi(\varphi_v)}<\delta/\varepsilon\,\Big|\,N_n^+\neq\emptyset\Big]. \tag{7.7}$$

By Lemma 5.3 (with $f=1$), the second probability tends to 0 as $n\to\infty$ if $\delta=\delta(\varepsilon)$ is small enough. Since $\chi(x)\geq\chi(h^*)>0$ for every $x\geq h^*$, it holds that $|N_n^{(\eta,R)}|\leq c\sum_{v\in N_n^+}\chi(\varphi_v)1_{\{v\text{ is }(\eta,R)\text{-bad}\}}$, and thus

$$P_x\Big[\frac{|N_n^{+,(\eta,R)}|}{\sum_{v\in N_n^+}\chi(\varphi_v)}>\delta\Big|N_n^+\neq\emptyset\Big]\leq P_x\big[E_{R,n}^{\delta/c}|N_n^+\neq\emptyset\big], \tag{7.8}$$

which together with (7.5) implies the claim of the lemma.

We now prove (7.5). By first using the Markov inequality, and then Lemma B.2 with $k=1$ and $Y(v_1)=\chi(\varphi_{v_1})1_{\{v_1\text{ is }(\eta,R)\text{-bad}\}}1_{\{N_n^+\neq\emptyset\}}/(\sum_{w\in N_n^+}\chi(\varphi_w))$, it holds

$$\begin{aligned}P_x[E_{R,n}^{\varepsilon}|N_n^+\neq\emptyset]&\leq\varepsilon^{-1}E_x\Big[\frac{\sum_{v\in N_n^+}\chi(\varphi_v)1_{\{v\text{ is }(\eta,R)\text{-bad}\}}}{\sum_{v\in N_n^+}\chi(\varphi_v)}\Big|N_n^+\neq\emptyset\Big]\\&=\varepsilon^{-1}E_x\Big[\frac{1_{\{N_n^+\neq\emptyset\}}\sum_{v\in N_n^+}\chi(\varphi_v)1_{\{v\text{ is }(\eta,R)\text{-bad}\}}}{\sum_{v\in N_n^+}\chi(\varphi_v)}\Big]P_x[N_n^+\neq\emptyset]^{-1}\\&=\varepsilon^{-1}Q_x\Big[\frac{\chi(x)P_x[N_n^+\neq\emptyset]^{-1}}{\sum_{v\in N_n^+}\chi(\varphi_v)}1_{\{\sigma_n\text{ is }(\eta,R)\text{-bad}\}}\Big],\end{aligned} \tag{7.9}$$

where, as defined in Section 3, σ_n is the vertex on the spine in the n-th generation. To continue, we first show the following two claims:

(i) $\sup_{n\geq R}Q_x[\sigma_n\text{ is }(\eta,R)\text{-bad}]\to0$ as $R\to\infty$.

(ii) Set $Z_n=\frac{\chi(x)P_x[N_n^+\neq\emptyset]^{-1}}{\sum_{v\in N_n^+}\chi(\varphi_v)}$. Then for all $\delta>0$, there exist $R',K>0$ such that $Q_x[Z_n1_{\{Z_n>K\}}]\leq\delta$ for all $n\geq R'$.

Claim (i) will follow from the ergodic behaviour of the field along the spine $(\sigma_n)_{n\geq0}$ under Q_x. Let $\text{sib}^<(v)=\{w\in\text{sib}(v):v<_{\mathbf{v}}w\}$. Then $\bar S(\sigma_n)=\sum_{k\leq n}\sum_{v\in\text{sib}^<(\sigma_k)}\chi(\varphi_v)$. Recall that $\xi_k:=\varphi(\sigma_k)$. By conditioning on ξ_{k-1}, due to the fact that under Q_x the spine mark at generation k is uniformly distributed on descendants of its position at generation $k-1$ and that non-spine vertices behave as under P_x,

$$\begin{aligned}Q_x\Big[\sum_{v\in\text{sib}^<(\sigma_k)}\chi(\varphi_v)\Big]&=Q_x\Big[Q_{\xi_{k-1}}\Big[\sum_{v\in\text{sib}^<(\sigma_1)}\chi(\varphi_v)\Big]\Big]\\&=Q_x\Big[\frac{1}{2}Q_{\xi_{k-1}}\Big[\sum_{v\in N_1^+\setminus\{\sigma_1\}}\chi(\varphi_v)\Big]\Big]\\&=Q_x\Big[\frac{d-1}{2d}P_{\xi_{k-1}}\Big[\sum_{v\in N_1^+}\chi(\varphi_v)\Big]\Big]=\frac{d-1}{2d}Q_x[\chi(\xi_{k-1})].\end{aligned} \tag{7.10}$$

By Lemma 3.4 the invariant measure of the Markov chain $(\xi_k)_{k\geq 0}$ under Q_x is given by $\chi(y)^2\nu(\mathrm{d}y)$, so the last expression converges to C_1^{-1} as $k\to\infty$. Claim (i) then follows from an ergodicity argument.

To see (ii), note that again by Lemma B.2,

$$\begin{aligned} Q_x[Z_n 1_{\{Z_n>K\}}] &= \frac{P_x[1_{\{Z_n>K\}}1_{\{N_n^+\neq\emptyset\}}]}{P[N_n^+\neq\emptyset]} = P_x[Z_n>K|N_n^+\neq\emptyset]. \\ &\leq P_x\Big[\frac{\sum_{v\in N_n^+}\chi(\varphi_v)}{n}<\frac{1}{Kc}\Big|N_n^+\neq\emptyset\Big], \end{aligned} \tag{7.11}$$

where in the last inequality we used the fact that $\chi(x)P_x[N_n^+\neq\emptyset]^{-1}\geq c$, by Theorem 2.1. By Theorem 2.2, this converges to the probability that an exponentially distributed random variable is smaller than $(Kc_x)^{-1}$. It is thus straightforward to find K and R' such that the last probability is smaller than δ for all $n\geq R'$, proving (ii).

Returning back to (7.9), its right-hand side is bounded by

$$\begin{aligned} &Q_x[Z_n 1_{\{\sigma_n \text{ is } (\eta,R)\text{-bad}\}}] \\ &\quad = Q_x[Z_n 1_{\{Z_n>K\}}1_{\{\sigma_n \text{ is } (\eta,R)\text{-bad}\}}] + Q[Z_n 1_{\{Z_n\leq K\}}1_{\{\sigma_n \text{ is } (\eta,R)\text{-bad}\}}] \\ &\quad \leq Q_x[Z_n 1_{\{Z_n>K\}}] + KQ[\sigma_n \text{ is } (\eta,R)\text{-bad}], \end{aligned} \tag{7.12}$$

which together with (i) and (ii) implies (7.5) and concludes the proof. □

To state the second main result of this section, we introduce (enlarging the probability space if necessary) random variables $\tau^1,\dots,\tau^k$ which under P_x, conditionally on T, are independent and uniformly distributed on $\{1,2,\dots,|T|\}$. Using these random variables, we define two $k\times k$ random matrices

$$\begin{aligned} (D_n^{\bar S})_{i,j} &= n^{-1}\big(\bar S(v_{\tau^i})+\bar S(v_{\tau^j})-2\bar S(v_{\tau^i}\wedge v_{\tau^j})\big) \\ (D_n^{H})_{i,j} &= n^{-1}\big(H(v_{\tau^i})+H(v_{\tau^j})-2H(v_{\tau^i}\wedge v_{\tau^j})\big), \end{aligned} \tag{7.13}$$

where, as usual, $(v_1,v_2,\dots)$ denotes the depth-first traversal of T, and $v\wedge w$ denotes the most recent common ancestor of v and w in T.

Proposition 7.2. *For every $k\geq 1$ and $\varepsilon>0$,*

$$\lim_{n\to\infty} P_x\big[\|C_1^{-1}D_n^H - D_n^{\bar S}\|>\varepsilon\big|N_n^+\neq\emptyset\big]=0, \tag{7.14}$$

where $\|\cdot\|$ denotes the Frobenius norm of $k\times k$ matrices.

To prove this proposition, we need two lemmas. The first one estimates from below the size of T conditionally on $\{N_n^+\neq\emptyset\}$.

Lemma 7.3. *For every $x\geq h^*$,*

$$\lim_{q\to 0}\sup_{n\geq 0} P_x\big[|T|\leq qn^2|N_n^+\neq\emptyset\big]=0. \tag{7.15}$$

Proof. It will be sufficient to show that for every $\delta>0$ there exist q and n_0 such that for every $n\geq n_0$,

$$P_x[|T|\leq qn^2|N_n^+\neq\emptyset]\leq\delta. \tag{7.16}$$

Indeed, since there are only finitely many $n<n_0$, by decreasing the value of q, the inequality (7.16) can be made true for all $n\geq 0$. This then directly implies the claim of the lemma.

To show (7.16), observe that for every $\eta > 0$

$$\begin{aligned} P_x[|T| \le qn^2 | N_n^+ \neq \emptyset] \\ \le P_x\big[|T| \le qn^2, |N_n^+| \ge \eta n | N_n^+ \neq \emptyset\big] + P_x\big[|N_n^+| < \eta n | N_n^+ \neq \emptyset\big]. \end{aligned} \tag{7.17}$$

By Theorem 2.2, there is $\eta > 0$ small such that the second probability on the right-hand side of (7.17) is bounded by $\delta/2$ for all n large enough. It is thus sufficient to show that for every $\delta > 0$ and $\eta > 0$ there is n_0 so that for all $n \ge n_0$

$$P_x\big[|T| \le qn^2, |N_n^+| \ge \eta n \big| N_n^+ \neq \emptyset\big] \le \delta/2. \tag{7.18}$$

Note first that

$$P_x\big[|T| \le qn^2, |N_n^+| \ge \eta n \big| N_n^+ \neq \emptyset\big] \le P_x\big[|T| \le qn^2 \big| |N_n^+| \ge \eta n\big]. \tag{7.19}$$

Given $\{|N_n^+| \ge \eta n\}$, let $w_1, \dots, w_{\lfloor \eta n \rfloor}$ be the first $\lfloor \eta n \rfloor$ vertices in N_n^+, and let T_w be the subtree of T rooted at w. Obviously, $|T| \le \sum_{i=1}^{\lfloor \eta n \rfloor} |T_{w_i}|$. Under $P[\,\cdot\,| |N_n^+| \ge \eta n]$, the random variables $|T_{w_i}|$ are independent. Moreover, since $w \in N_n^+$ implies $\varphi_w \ge h^*$, by stochastic domination (3.5), for any $u > 0$,

$$P_x\big[|T_{w_i}| \ge u \big| |N_n^+| \ge \eta n\big] \ge P_{h^*}\big[|\mathcal{C}_o \cap \mathbb{T}^+| \ge u\big]. \tag{7.20}$$

Denoting T_i', $i \ge 1$, i.i.d. random variables distributed as $|\mathcal{C}_o \cap \mathbb{T}^+|$ under P_{h^*}, it follows that

$$P_x\big[|T| \le qn^2 \big| |N_n^+| \ge \eta n\big] \le P\Big[\sum_{i=1}^{\lfloor \eta n \rfloor} T_i' \le qn^2\Big]. \tag{7.21}$$

By Theorem 3.3, T_i' are in the domain of attraction of the 1/2-stable random distribution. Therefore, $n^{-2}\sum_{i \le \lfloor \eta n \rfloor} T_i'$ converges in distribution to a non-negative 1/2-stable random variable (see e.g. [Fel71, Theorem XIII.6.2]). As a consequence, since the distribution of this random variable has no atom at 0, for any $\delta, \eta > 0$ there exists q small so that for all n large enough the right-hand side of (7.21) is bounded by $\delta/2$. This shows (7.18) and completes the proof. ☐

The second lemma needed to show Proposition 7.2 studies the probability that a randomly chosen vertex v_{τ_1} is (η, R)-bad.

Lemma 7.4. *Let $V := v_{\tau_1}$ be a uniformly chosen vertex of T. Then, for every $x \ge h^*$ and $\eta > 0$,*

$$\lim_{R \to \infty} \sup_{n \ge R} P_x\big[V \text{ is } (\eta, R)\text{-bad } \big| N_n^+ \neq \emptyset\big] = 0. \tag{7.22}$$

Proof. For arbitrary positive constants $q, h_1 < h_2$, it holds

$$\begin{aligned} &P_x\big[V \text{ is } (\eta, R)\text{-bad } \big| N_n^+ \neq \emptyset\big] \\ &\le P_x\big[|T| \le qn^2 \big| N_n^+ \neq \emptyset\big] + P_x\big[|T| > qn^2, H(V) < h_1 n \big| N_n^+ \neq \emptyset\big] \\ &\quad + P_x\big[|T| > qn^2, H(V) > h_2 n \big| N_n^+ \neq \emptyset\big] \\ &\quad + P_x\big[|T| > qn^2, H(V) \in [h_1 n, h_2 n], V \text{ is } (\eta, R)\text{-bad } \big| N_n^+ \neq \emptyset\big]. \end{aligned} \tag{7.23}$$

We will show that for every $\delta > 0$ there are q, h_1, h_2 such that for every $n \ge 0$ the first three summands on the right-hand side are smaller than $\delta/3$, and that for every fixed q, h_1, h_2 the fourth one satisfies

$$\lim_{R \to \infty} \sup_{n \ge R} P_x\big[|T| > qn^2, H(V) \in [h_1 n, h_2 n], V \text{ is } (\eta, R)\text{-bad } \big| N_n^+ \neq \emptyset\big] = 0. \tag{7.24}$$

This will imply the statement of the lemma.

Concerning the first summand, the fact that it is possible to choose $q > 0$ small so that $P_x[|T| \le qn^2 \big| N_n^+ \ne \emptyset] \le \delta/3$ for all $n \ge 0$ follows immediately from Lemma 7.3. From now on we keep q fixed in this way.

For the second summand, it holds

$$\begin{aligned} P_x\big[|T| > qn^2, H(V) < h_1 n \big| N_n^+ \ne \emptyset\big] &\le \frac{1}{P_x[N_n^+ \ne \emptyset]} P_x\big[|T| > qn^2, H(V) < h_1 n\big] \\ &\le \frac{1}{P_x[N_n^+ \ne \emptyset]} \frac{1}{qn^2} E_x\Big[\sum_{k=0}^{h_1 n} |N_k^+|\Big], \end{aligned} \tag{7.25}$$

where the last inequality follows from the fact that V is a uniformly chosen vertex of T, and thus on $\{|T| \ge qn^2\}$,

$$P_x[H_V \le h_1 n | \sigma(T)] = |T|^{-1} \sum_{k=0}^{h_1 n} |N_k^+| \le \frac{1}{qn^2} \sum_{k=0}^{h_1 n} |N_k^+|. \tag{7.26}$$

By Theorem 2.1, $P_x[N_n^+ \ne \emptyset]^{-1} \le c(x) n$, and, by Proposition 3.8, $E_x[|N_k^+|] \le c'(x)$ for all $k \ge 0$. Therefore, the right-hand side of (7.25) can be made smaller than $\delta/3$ by choosing h_1 small.

For the third summand in (7.23), note that $\{H(V) \ge h_2 n\} \subset \{N_{h_2 n}^+ \ne \emptyset\}$, and thus

$$P_x\big[|T| > qn^2, H(V) > h_2 n \big| N_n^+ \ne \emptyset\big] \le \frac{1}{P_x[N_n^+ \ne \emptyset]} P_x[N_{h_2 n}^+ \ne \emptyset], \tag{7.27}$$

which can be made arbitrarily small by using Theorem 2.1 and choosing h_2 large.

Finally, we show (7.24). Recall the notation from (7.2). Since V is a uniformly chosen vertex of T, for arbitrary $q, h_1, h_2, n, R > 0$ and $\varepsilon \in (0,1)$, on the event $\{|T| \ge qn^2\}$ it holds that

$$\begin{aligned} P_x\big[H(V) \in [h_1 n, h_2 n], V \text{ is } (\eta, R)\text{-bad } \big| \sigma(T, \varphi)\big] &= \frac{1}{|T|} \sum_{k=h_1 n}^{h_2 n} |N_k^{+,(\eta,R)}| \\ = \frac{1}{|T|} \sum_{k=h_1 n}^{h_2 n} \frac{|N_k^{+,(\eta,R)}|}{|N_k^+|} |N_k^+| &\le \varepsilon + \frac{1}{qn^2} \sum_{k=h_1 n}^{h_2 n} 1_{\{|N_k^{+,(\eta,R)}|/|N_k^+| \ge \varepsilon\}} |N_k^+|. \end{aligned} \tag{7.28}$$

As a consequence, the probability in (7.24) satisfies

$$\begin{aligned} &P_x\big[|T| > qn^2, H(V) \in [h_1 n, h_2 n], V \text{ is } (\eta, R)\text{-bad } \big| N_n^+ \ne \emptyset\big] \\ &\qquad \le \varepsilon + \frac{1}{qn^2} \sum_{k=h_1 n}^{h_2 n} E_x\big[1_{\{|N_k^{+,(\eta,R)}|/|N_k^+| \ge \varepsilon\}} |N_k^+| \big| N_n^+ \ne \emptyset\big]. \end{aligned} \tag{7.29}$$

By the Cauchy–Schwarz inequality, every summand in (7.29) is bounded by

$$E_x\big[|N_k^+|^2 \big| N_n^+ \ne \emptyset\big]^{1/2} P_x\big[|N_k^{+,(\eta,R)}|/|N_k^+| > \varepsilon, N_k^+ \ne \emptyset \big| N_n^+ \ne \emptyset\big]^{1/2}. \tag{7.30}$$

By Theorem 2.1 and Proposition 3.8, $E_x[|N_k^+|^2 \,|\, N_n^+ \ne \emptyset]^{1/2} \le c(x) n$. The second term satisfies,

$$P_x\Big[\frac{|N_k^{+,(\eta,R)}|}{|N_k^+|} > \varepsilon, N_k^+ \ne \emptyset \Big| N_n^+ \ne \emptyset\Big] \le P_x\Big[\frac{|N_k^{+,(\eta,R)}|}{|N_k^+|} > \varepsilon \Big| N_k^+ \ne \emptyset\Big] \frac{P_x[N_k^+ \ne \emptyset]}{P_x[N_n^+ \ne \emptyset]}. \tag{7.31}$$

By Theorem 2.1, for $k \geq h_1 n$, the fraction on the right-hand side can be bounded by a constant $c(x)$. Noting also that $|N_k^{+,(\eta,R)}| = 0$ for $R > k$, we obtain that (7.29) is bounded by

$$\varepsilon + \frac{1}{qn^2} c(x) n^2 (h_2 - h_1) \sup_{k \geq R} P_x\Big[\frac{|N_k^{+,(\eta,R)}|}{|N_k^+|} > \varepsilon \Big| N_k^+ \neq \emptyset\Big]^{1/2}. \tag{7.32}$$

By Proposition 7.1, the supremum here converges to 0 as $R \to \infty$. Since $\varepsilon > 0$ is arbitrary, this shows (7.24) and completes the proof. □

We can now prove Proposition 7.2. The proof resembles the one of Proposition 6.21 in [Pow19], with modifications required to account for the unbounded domain in our setting.

Proof of Proposition 7.2. It is enough to show the statement for $k = 2$. The general case where $k > 2$ is then obtained by a union bound. In the case $k = 2$, since the matrices are symmetric and the diagonal entries are zero, it is sufficient to control one off-diagonal entry. To this end, for $\varepsilon > 0$ let G_n be the event

$$G_n := \big\{|C_1^{-1}(D_n^H)_{1,2} - (D_n^{\bar{S}})_{1,2}| > \varepsilon\big\}. \tag{7.33}$$

In order to bound the probability of G_n, we will define events $A_n(\delta)$ (see (7.38)) satisfying $P[A_n(\delta)|N_n^+ \neq \emptyset] \leq \delta$ and show that for n large enough and the right choice of $\delta_n \to 0$, G_n is a subset of $A_n(\delta_n)$. This will imply the statement of the lemma.

In order to define the events $A_n(\delta)$ for $\delta > 0$, we first fix $M = M(\delta) \geq 1$ so that

$$\lim_{n\to\infty} P_x[N_{Mn}^+ \neq \emptyset | N_n^+ \neq \emptyset] \leq \delta/3, \tag{7.34}$$

which is possible by Theorem 2.1. We then set

$$\eta = \eta(\delta) := \varepsilon/(4M), \tag{7.35}$$

and, using Lemma 7.4 and the independence of τ^1 and τ^2, we fix $R = R(\delta)$ so that

$$\lim_{n\to\infty} P_x\big[v_{\tau^1} \text{ or } v_{\tau^2} \text{ is } (\eta, R)\text{-bad}\big| N_n^+ \neq \emptyset\big] \leq \delta/3. \tag{7.36}$$

Next, we fix $K = K(\delta)$ such that

$$\lim_{n\to\infty} P_x[\sup\{\chi(\varphi_u) : u \in \cup_{k=0}^R N_k^+\} \geq K | N_n^+ \neq \emptyset] \leq \delta/3. \tag{7.37}$$

This is possible since the event in (7.37) is $\mathcal{F}_R$-measurable, and thus, by Proposition 5.2, the left-hand side of (7.37) equals $Q_x[\sup\{\chi(\varphi_u) : u \in \cup_{k=0}^R N_k^+\} \geq K]$ (with Q_x as in Section 3.2). It is then straightforward to choose K large enough so that (7.37) is satisfied. Finally, we define the events A_n by

$$\begin{aligned} A_n = A_n(\delta) = \{N_{Mn}^+ \neq \emptyset\} &\cup \{v_{\tau^1} \text{ or } v_{\tau^2} \text{ is } (\eta, R)\text{-bad}\} \\ &\cup \{\sup\{\chi(\varphi_u) : u \in \cup_{k=0}^R N_k^+\} \geq K\}. \end{aligned} \tag{7.38}$$

By the choice of M, η, R, and K, it holds that $\lim_{n\to\infty} P_x[A_n(\delta)|N_n^+ \neq \emptyset] \leq \delta$. Therefore, for sequences δ_n converging to zero sufficiently slowly, it also holds that

$$\lim_{n\to\infty} P_x[A_n(\delta_n)|N_n^+ \neq \emptyset] = 0. \tag{7.39}$$

Next, we show that $G_n \subseteq A_n(\delta_n)$ for a well-chosen δ_n and large n. Let $B(\delta) = \{H(v_{\tau^1} \wedge v_{\tau^2}) > R\}$. We will show that for large enough n and a suitable choice of δ_n,

$$G_n \cap A_n^c(\delta_n) \cap B(\delta_n) = \emptyset \quad \text{and} \quad G_n \cap A_n^c(\delta_n) \cap B^c(\delta_n) = \emptyset. \tag{7.40}$$

From this, it easily follows that $G_n \cap A_n^c(\delta_n) = \emptyset$, and thus, $G_n \subseteq A_n(\delta_n)$. Then, if $\delta_n \to 0$ sufficiently slowly, by (7.39),

$$\lim_{n\to\infty} P_x[G_n | N_n^+ \neq \emptyset] \le \lim_{n\to\infty} P_x[A_n(\delta_n) | N_n^+ \neq \emptyset] = 0, \tag{7.41}$$

which is sufficient to prove the statement of the proposition.

It remains to show the existence of δ_n such that (7.39) and (7.40) hold. By definition, on A_n^c it holds

$$H(v_{\tau^i}) \le Mn \quad \text{and} \quad \left| \frac{\bar{S}(v_{\tau^i})}{H(v_{\tau^i})} - C_1^{-1} \right| \le \eta \qquad \text{for } i = 1, 2. \tag{7.42}$$

Thus, on $A_n^c \cap B$ it holds that $R < H(v_{\tau^1} \wedge v_{\tau^2}) \le Mn$. Since v_{τ^1} is not (η, R)-bad, $|\bar{S}(v_{\tau^1} \wedge v_{\tau^2})/H(v_{\tau^1} \wedge v_{\tau^2}) - C_1^{-1}| \le \eta$. Together with (7.42), the definition of the matrices in (7.13), and the definition of η, this implies

$$|(D_n^{\bar{S}})_{1,2} - C_1^{-1}(D_n^H)_{1,2}| \le \frac{1}{n} 4Mn\eta \le \varepsilon, \tag{7.43}$$

and thus $G_n \cap B \cap A_n^c = \emptyset$, proving the first half of (7.40). Next, on event $B^c \cap A_n^c$ it holds $H(v_{\tau^1} \wedge v_{\tau^2}) \le R$ and $\sup\{\chi(\varphi_u) : u \in \cup_{k\le R} N_k^+\} < K$. Therefore, $\bar{S}(v_{\tau^1} \wedge v_{\tau^2}) \le Kd^R$, and as a consequence $|\bar{S}(v_{\tau^1} \wedge v_{\tau^2}) - C_1^{-1} H(v_{\tau^1} \wedge v_{\tau^2})| \le Kd^R + RC_1^{-1}$. We choose a sequence δ_n converging to 0 sufficiently slowly so that (7.39) is satisfied and $K(\delta_n)d^{R(\delta_n)+1} + R(\delta_n)C_1^{-1} \le \varepsilon n$, and thus

$$|(D_n^{\bar{S}})_{1,2} - C_1^{-1}(D_n^H)_{1,2}| \le \frac{\varepsilon}{2} + \frac{K(\delta_n)d^{R(\delta_n)} + R(\delta_n)C_1^{-1}}{n} \le \varepsilon \tag{7.44}$$

for n large. Hence, $G_n \cap B^c(\delta_n) \cap A_n^c(\delta_n) = \emptyset$ for such n, completing the proof of (7.40). □

8. Proof of Theorem 2.3

In this section, we prove Theorem 2.3, relying on Lemmas 8.1 and 8.2. The argument follows the framework of *random metric measure spaces*, with key results originating from [GPW09] and [ADH13]. The proof closely mirrors that of Theorem 1.1 in [Pow19, Section 6.3.2], which itself builds on the aforementioned results.

Recall from Section 2 that $\mathbf{e}$ denotes a Brownian excursion conditioned to reach height 1. Also define $\hat{\mathbf{e}}$ to be (σC_1) times a Brownian excursion conditioned to reach height $(C_1\sigma)^{-1}$, where the constants σ and C_1 are given in (6.5) and (2.11), and write $(T_{\hat{\mathbf{e}}}, d_{\hat{\mathbf{e}}})$ for the real tree with contour function $\hat{\mathbf{e}}$. By Brownian scaling, $(T_{\hat{\mathbf{e}}}, d_{\hat{\mathbf{e}}})$ is isometrically equivalent to $(T_{\mathbf{e}}, d_{\mathbf{e}})$. Therefore, to show Theorem 2.3, it is enough to prove that if $(T_{x,n}, d_{x,n})$ is a random metric space whose law coincides with that of $(\mathcal{C}_o^{h^*} \cap \mathbb{T}^+, d/n)$ under $P_x[\cdot | N_n^+ \neq \emptyset]$ then

$$(T_{x,n}, d_{x,n}) \to (T_{\hat{\mathbf{e}}}, d_{\hat{\mathbf{e}}}) \quad \text{as } n \to \infty \tag{8.1}$$

in distribution with respect to the Gromov-Hausdorff topology.

It will prove useful to additionally equip $(T_{x,n}, d_{x,n})$ and $(T_{\hat{\mathbf{e}}}, d_{\hat{\mathbf{e}}})$ with measures and work in the space of metric measure spaces which we will call $\mathbb{X}$. For this purpose, we equip $(T_{x,n}, d_{x,n})$ with the measure $\mu_{x,n}$ which is the uniform measure among the vertices in $\mathcal{C}_o^{h^*} \cap \mathbb{T}^+$. We equip $(T_{\hat{\mathbf{e}}}, d_{\hat{\mathbf{e}}})$ with the measure $\mu_{\hat{\mathbf{e}}}$ which is the uniform measure on $[0, \hat{\tau})$ and $\hat{\tau}$ is the length of $\hat{\mathbf{e}}$. We write P_x^n for the law of $(T_{x,n}, d_{x,n}, \mu_{x,n})$ and $P_{\hat{\mathbf{e}}}$ for the law of $(T_{\hat{\mathbf{e}}}, d_{\hat{\mathbf{e}}}, \mu_{\hat{\mathbf{e}}})$, and use E_x^n and $E_{\hat{\mathbf{e}}}$ for the corresponding expectations.

Lemma 8.1. *For fixed $x \ge h^*$, the family $(T_{x,n}, d_{x,n}, \mu_{x,n})_{n\ge 0}$ is tight with respect to the Gromov-Hausdorff-Prokhorov metric.*

Proof. Fix $x \geq h^*$. Our aim is to show that for every $\varepsilon > 0$ there exists a relatively compact subset $\mathbb{K} \subset \mathbb{X}$ (with respect to the Gromov-Hausdorff-Prokhorov metric) such that

$$\inf_{n\geq 0} P_x^n[(T_{x,n}, d_{x,n}, \mu_{x,n}) \in \mathbb{K}] \geq 1 - \varepsilon. \tag{8.2}$$

To do this, we define the relatively compact set (where the relative compactness follows from Theorem 2.6 in [ADH13])

$$\mathbb{K}_{R,M} = \left\{ (X, r, \mu) \in \mathbb{X} \;\middle|\; \begin{array}{l} \mu(X) = 1,\ \mathrm{diam}(X) \leq 2R,\ \text{for all } k \geq 1,\ X \text{ can be} \\ \text{covered by fewer than } 2^{4k}M \text{ balls of radius } 2^{-k} \end{array} \right\} \tag{8.3}$$

and show that for $\varepsilon > 0$ there are $R(\varepsilon)$ and $M(\varepsilon)$ such that for $\mathbb{K} = \mathbb{K}_{R,M}$, (8.2) is satisfied. Define the events

$$A_{x,n}(R) = \{\mathrm{diam}(T_{x,n}) \leq 2R\} \quad \text{and} \tag{8.4}$$

$$B_{x,n}^{\delta}(M) = \{T_{x,n} \text{ can be covered with } < \delta^{-4}M \text{ balls of radius } \delta\}. \tag{8.5}$$

Then, since $\mu_{x,n}(T_{x,n}) = 1$, $\{(T_{x,n}, d_{x,n}, \mu_{x,n}) \in \mathbb{K}_{R,M}\} = A_{x,n}(R) \cap (\cap_{k\geq 1} B_{x,n}^{2^{-k}}(M))$, and thus $P_x^n[(T_{x,n}, d_{x,n}, \mu_{x,n}) \notin \mathbb{K}_{R,M}] \leq P_x^n[A_{x,n}(R)^c] + \sum_{k\geq 1} P_x^n[B_{x,n}^{2^{-k}}(M)^c \cap A_{x,n}(R)]$. To see (8.2), it is therefore enough to show that for $M = M(\varepsilon)$ and $R = R(\varepsilon)$ large enough

$$P_x^n[A_{x,n}(R)^c] \leq \varepsilon/2 \quad \text{and} \quad P_x^n[B_{x,n}^{\delta}(M)^c \cap A_{x,n}(R)] \leq \delta\varepsilon/2, \tag{8.6}$$

for all $\delta > 0$ and $n \geq 0$.

We will first find a $R(\varepsilon)$ such that the first equation of (8.6) is satisfied. Recall that the law of $(T_{x,n}, d_{x,n})$ under P_x^n is that of $(T, d/n)$ under $P_x[\,\cdot\,|N_n^+ \neq \emptyset]$. Thus, $P_x^n[\mathrm{diam}(T_{x,n}) \geq 2R] \leq P_x[|N_{Rn}^+| > 0 | N_n^+ \neq \emptyset]$, which by Theorem 2.1, is smaller than $\varepsilon/2$ for $R(\varepsilon)$ large enough. For such $R(\varepsilon)$ the first part of (8.6) is valid.

Next, we will find a $M(\varepsilon)$ such that also the second equation of (8.6) is met. We will distinguish thereby between the cases where $n\delta < 1$ and $n\delta \geq 1$. Assume first that $n\delta < 1$. Then, every δ-ball around a vertex $v \in T_{x,n}$ only contains the vertex v itself. This means that in this case $T_{x,n}$ can be covered with $< \delta^{-4}M$ balls of radius δ if and only if the cardinality of $T_{x,n}$ is $< \delta^{-4}M$. Therefore

$$P_x^n[B_{x,n}^{\delta}(M)^c \cap A_{x,n}(R)] \leq P_x[|T| \geq \delta^{-4}M | N_n^+ \neq \emptyset] \leq \frac{P_x[|T| \geq \delta^{-4}M]}{P_x[N_n^+ \neq \emptyset]}. \tag{8.7}$$

Using Theorem 2.1 and Proposition 3.3 this can be further bounded from above by

$$C\frac{\delta^2 n}{\sqrt{M}} \leq C\frac{\delta}{\sqrt{M}}, \tag{8.8}$$

for some constant C, where we also used that by assumption $\delta n < 1$. Choosing some $M = M(\varepsilon) \geq (2C/\varepsilon)^2$, the second equation of (8.6) is true for all n, δ with $n\delta < 1$.

Next we consider the case $n\delta \geq 1$. Set $b_{n,\delta} = \lfloor n\delta \rfloor$ and define the sets of vertices $V_j = \{v \in N_{jb_{n,\delta}}^+ : \exists w \in N_{(j+1)b_{n,\delta}}^+ \text{ with } v \prec w\}$ for $j \geq 0$. We will show that on

$$\{\mathrm{diam}(T_{x,n}) \leq 2R\} \cap \Big(\bigcap_{j=0}^{2Rn/b_{n,\delta}-1} \{|V_j| \leq \frac{Mb_{n,\delta}}{2Rn\delta^4}\}\Big) \tag{8.9}$$

$(T_{x,n}, d_{x,n})$ can be covered with fewer than $\delta^{-4}M$ balls of radius δ. Specifically, $T_{x,n}$ is covered by $\{B_\delta(v)\}_{v\in V}$, where $V = \cup_{j=0}^{2Rn/b_{n,\delta}-1} V_j$. To see this, note that since on $\{\mathrm{diam}(T_{x,n}) \leq 2R\}$ the height of $T_{x,n}$ is trivially bounded by $2Rn$ and thus, for every

$v \in T_{x,n}$ there is a $w \in V_j$ for some $0 \le j \le 2Rn/b_{n,\delta} - 1$ so that $d_{x,n}(v,w) \le \delta$. Further, it is clear that on (8.9) the cardinality of V is bounded by $(2Rn/b_{n,\delta})(Mb_{n,\delta}/2Rn\delta^4) = \delta^{-4}M$. This shows the claimed covering property. To prove the second part of (8.6), it thus suffices to show that for $M = M(R,\varepsilon)$ large enough,

$$P_x\Big[\bigcup_{j=0}^{2Rn/b_{n,\delta}-1} \{|V_j| > \frac{Mb_{n,\delta}}{2Rn\delta^4}\}\Big| N_n^+ \neq \emptyset\Big] \le \delta\varepsilon/2. \tag{8.10}$$

By conditioning on $\mathcal{F}_{jb_{n,\delta}}$ (recalling the definition of $\mathcal{F}_n$ from (3.20)) and using Theorem 2.1 and Proposition 3.8,

$$E_x\big[|V_j|\big] = E_x\big[E_x\big[|V_j|\big|\mathcal{F}_{jb_{n,\delta}}\big]\big] = E_x\Big[\sum_{v \in N^+_{jb_{n,\delta}}} P_{\varphi_v}[N^+_{b_{n,\delta}} \neq \emptyset]\Big] \tag{8.11}$$

$$= E_x\Big[\sum_{v \in N^+_{jb_{n,\delta}}} C\chi(\varphi_v)b_{n,\delta}^{-1}(1 + \varepsilon_{b_{n,\delta}}(\varphi_v))\Big] \le q(x)b_{n,\delta}^{-1} \tag{8.12}$$

for some function $q(x)$ independent of n and δ. Therefore, by conditioning on $\{N_n^+ \neq \emptyset\}$ and using the Markov inequality and Theorem 2.1,

$$P_x\Big[|V_j| > \frac{Mb_{n,\delta}}{2Rn\delta^4}\Big| N_n^+ \neq \emptyset\Big] \le \frac{2Rn\delta^4}{Mb_{n,\delta}} \frac{E_x[|V_j|]}{P_x[N_n^+ \neq \emptyset]} \le q'(x)\frac{2Rn^2\delta^4}{Mb_{n,\delta}^2}, \tag{8.13}$$

where q' is some other function independent of n and δ. This means that by a union bound, the left-hand side of (8.10) is bounded from above by

$$q'(x)\frac{2Rn^2\delta^4}{Mb_{n,\delta}^2}(2Rn/b_{n,\delta}) = q'(x)\frac{4R^2n^3\delta^4}{Mb_{n,\delta}^3} = \frac{q'(x)4R^2\delta}{M}\Big(\frac{n\delta}{\lfloor n\delta\rfloor}\Big)^3. \tag{8.14}$$

Using that for $x \ge 1$, $1 \le x/\lfloor x\rfloor \le 2$, we see that by choosing $M \ge 64R^2q'(x)/\varepsilon$, the right-hand side of (8.14) is bounded by $\delta\varepsilon/2$. For such choices of M, (8.10) is satisfied.

Finally, first choosing R as described above and then M as the maximum of the two M-values obtained for the cases $n\delta < 1$ and $n\delta \ge 1$, gives (R,M) such that (8.6) is satisfied. For such (R,M), (8.2) holds true for $\mathbb{K} = \mathbb{K}_{R,M}$ and the proof is finished. □

Lemma 8.2. *For fixed $x \ge h^*$,*

$$(T_{x,n}, d_{x,n}, \mu_{x,n}) \to (T_{\hat{\mathbf{e}}}, d_{\hat{\mathbf{e}}}, \mu_{\hat{\mathbf{e}}}) \quad \text{as } n \to \infty \tag{8.15}$$

in distribution with respect to the Gromov-Prokhorov metric.

Proof. By Corollary 3.1 in [GPW09], the convergence (8.15) holds true if and only if

(a) The family $\{P_x^n\}_{n\ge0}$ is relatively compact in the space of probability measures on $\mathbb{X}$ (with respect to the Gromov-weak topology).

(b) For every function $\Psi : \mathbb{X} \to \mathbb{R}$ of the form,

$$\Psi((X,r,\mu)) = \int \psi((r(x_i,x_j))_{1\le i<j\le k})\mu^{\otimes k}(\mathrm{d}(x_1,\dots,x_k)), \tag{8.16}$$

where $\psi : [0,\infty)^{\binom{k}{2}} \to \mathbb{R}$ is a continuous and bounded function, $E_x^n[\Psi] \to E_{\hat{\mathbf{e}}}[\Psi]$ as $n \to \infty$.

Since convergence in the Gromov-Hausdorff-Prokhorov sense implies convergence in the Gromov-weak sense, part (a) is implied by Lemma 8.1 and only part (b) is left to be shown.

To see (b), fix a function Ψ as described in (8.16) and recall the definition of D_n^H from (7.13). Note that $d_{x,n}(v,w) = n^{-1}(H(v)+H(w)-2H(v\wedge w))$, and thus

$$E_x^n[\Psi] = E_x\big[\psi(D_n^H)\big|N_n^+\neq\emptyset\big]. \tag{8.17}$$

By Proposition 7.2,

$$\lim_{n\to\infty}\big(E_x\big[\psi(D_n^H)\big|N_n^+\neq\emptyset\big]-E_x\big[\psi(D_n^{\bar S}/C_1^{-1})\big|N_n^+\neq\emptyset\big]\big)=0. \tag{8.18}$$

Assume for now that also

$$\lim_{n\to\infty}\big(E_x\big[\psi(D_n^{\bar S}/C_1^{-1})\big|N_n^+\neq\emptyset\big]-E_x\big[\psi(D_n^{\bar S}/C_1^{-1})\big|\sup_m \bar S_m\geq C_1^{-1}n\big]\big)=0. \tag{8.19}$$

By Proposition 6.10, $\lim_{n\to\infty}E_x[\psi(D_n^{\bar S}/C_1^{-1})|\sup_m\bar S_m\geq C_1^{-1}n]=E_{\hat{\mathbf e}}[\Psi]$, and thus in combination with (8.17), (8.18) and (8.19), it follows $\lim_{n\to\infty}E_x^n[\Psi]=E_{\hat{\mathbf e}}[\Psi]$, concluding the proof.

It remains to prove (8.19). Since ψ is bounded, it is enough to show that

$$\lim_{n\to\infty}P_x\big[N_n^+\neq\emptyset\big|\sup_m\bar S_m\geq C_1^{-1}n\big]=1\quad\text{and}\quad\lim_{n\to\infty}P_x\big[\sup_m\bar S_m\geq C_1^{-1}n\big|N_n^+\neq\emptyset\big]=1. \tag{8.20}$$

Using that $P[A|B]=P[B|A]P[A]/P[B]$, it is enough to show that

$$\lim_{n\to\infty}P_x[\sup_m\bar S_m\geq C_1^{-1}n|N_n^+\neq\emptyset]=1\quad\text{and}\quad\lim_{n\to\infty}\frac{P_x[\sup_m\bar S_m\geq C_1^{-1}n]}{P_x[N_n^+\neq\emptyset]}=1. \tag{8.21}$$

Observe that for every $\delta>0$,

$$\begin{aligned}&P_x\big[\sup_m\bar S_m\geq C_1^{-1}n\big|N_n^+\neq\emptyset\big]\\&\quad\geq P_x\big[\sup_m\bar S_m\geq C_1^{-1}n\big||N^+_{\lfloor n(1+\delta)\rfloor}|>0\big]\frac{P_x[|N^+_{\lfloor n(1+\delta)\rfloor}|>0]}{P_x[N_n^+\neq\emptyset]}.\end{aligned} \tag{8.22}$$

By Proposition 7.1, the first term in the product on the right-hand side converges to 1 as $n\to\infty$, and the second term converges to $(1+\delta)^{-1}$ by Theorem 2.1. This shows the first convergence of (8.21). By Theorem 2.1, $P_x[N_n^+\neq\emptyset]\sim\chi(x)C_1n^{-1}$ as $n\to\infty$. Therefore, we are left to prove that

$$\lim_{n\to\infty}\frac{P_x[\sup_m\bar S_m\geq C_1^{-1}n]}{\chi(x)C_1n^{-1}}=1 \tag{8.23}$$

to conclude the second limit of (8.21). We will do this in a similar way as, e.g., in [LG05, Section 1.4, p. 263].

Consider the depth-first traversal $(v_1,v_2,\dots)$ of the sequence of trees $(T^1,T^2,\dots)$. We set $S'_n=\sum_{w\in Y(v_n)}\chi(\varphi_w)$ and $S^i=\sup_n S'_n 1_{v_n\in T^i}$ for $i\geq1$ (so that $\sup_n\bar S_n=S^1$). Recall that by Proposition 6.8,

$$\lim_{n\to\infty}\Big(\frac{S'_{\lfloor nt\rfloor}}{\sqrt n},\frac{\Lambda_{\lfloor nt\rfloor}}{\sqrt n}\Big)_{t\geq0}=\Big(\sigma|B_t|,\frac{\sigma}{\chi(x)}L_t^0\Big)_{t\geq0}, \tag{8.24}$$

where the limit is in P_x-distribution with respect to the Skorokhod topology. Defining $\gamma_r=\inf\{t\geq0:L_t^0>r\}$ and $\tau_n=\inf\{t\geq0:n^{-1}\Lambda_{\lfloor n^2t\rfloor}>1\}$ and using Brownian scaling, this gives the joint convergence

$$\lim_{n\to\infty}\Big(\Big(\frac1n S'_{\lfloor n^2t\rfloor}\Big)_{t\geq0},\tau_n\Big)=\Big((\sigma|B_t|)_{t\geq0},\gamma_{\chi(x)/\sigma}\Big), \tag{8.25}$$

and therefore also

$$\lim_{n\to\infty}\Big(\frac{1}{n}S'_{\lfloor n^2(t\wedge\tau_n)\rfloor}\Big)_{t\geq 0}=\big(\sigma|B_{t\wedge\gamma_{\chi(x)/\sigma}}|\big)_{t\geq 0}. \tag{8.26}$$

From this we deduce that for every $y>0$, as $n\to\infty$,

$$\lim_{n\to\infty}P_x\Big[\sup_{1\leq i\leq n}S^i>ny\Big]=P_x\Big[\sup_{t\leq\gamma_{\chi(x)/\sigma}}\sigma|B_t|>y\Big]=1-\exp\Big(-\frac{\chi(x)}{y}\Big), \tag{8.27}$$

where the last equality is a consequence of excursion theory for Brownian motion (see e.g. Chapter XII in [RY99]). By the independence of the trees T^i (and thus the variables S^i), we also have that

$$P_x\Big[\sup_{1\leq i\leq n}S^i>ny\Big]=1-\big(1-P_x\big[S^1>ny\big]\big)^n=1-\Big(1-P_x\Big[\sup_m\bar{S}_m>ny\Big]\Big)^n. \tag{8.28}$$

Combining (8.27) and (8.28) then gives

$$P_x\Big[\sup_m\bar{S}_m>ny\Big]\sim\frac{\chi(x)}{y}n^{-1}\quad\text{as } n\to\infty, \tag{8.29}$$

which by setting $y=C_1^{-1}$ concludes (8.23) and finishes the proof. □

We now present the proof of Theorem 2.3, which is identical to that of Theorem 1.1 in [Pow19].

Proof of Theorem 2.3. Since convergence in the Gromov-Hausdorff-Prokhorov metric implies convergence in the Gromov-Prokhorov metric, Lemma 8.2 characterizes subsequential limits with respect to the Gromov-Hausdorff-Prokhorov topology. Hence, we obtain the convergence in distribution

$$(T_{x,n},d_{x,n},\mu_{x,n})\to(T_{\hat{\mathbf{e}}},d_{\hat{\mathbf{e}}},\mu_{\hat{\mathbf{e}}})\quad\text{as } n\to\infty \tag{8.30}$$

with respect to the Gromov-Hausdorff-Prokhorov metric. Now, since Gromov-Hausdorff-Prokhorov convergence further implies convergence in the Gromov-Hausdorff sense, claim (8.1) follows, completing the proof. □

Appendix A. Properties of χ

We prove here that χ is Lipschitz continuous on $[h^*,\infty)$, as stated in (3.19). The proof uses similar ideas as the proof of Proposition 3.1 of [AČ20].

Proof of (3.19). We will show that the derivative χ' is bounded on $[h^*,\infty)$. Recall that for $x\geq h^*$ it holds $\chi(x)=d\int_{[h^*,\infty)}\chi(z)\rho_Y(z-x/d)\,\mathrm{d}z$, where ρ_Y is the density of the Gaussian random variable Y. Differentiating this expression, using an integration by parts and the fact that $\chi'(z)=0$ for $z<h^*$, we obtain that for $x\geq h^*$

$$\begin{aligned}\chi'(x)&=-\int_{h^*}^{\infty}\chi(z)\rho_Y'\Big(z-\frac{x}{d}\Big)\,\mathrm{d}z\\&=\int_{[h^*,\infty)}\chi'(z)\rho_Y\Big(z-\frac{x}{d}\Big)\,\mathrm{d}z+\chi(h^*)\rho_Y\Big(h^*-\frac{x}{d}\Big)\\&=:E_Y\Big[\chi'\Big(Y+\frac{x}{d}\Big)\Big]+e(x).\end{aligned} \tag{A.1}$$

Defining $e(x) = 0$ for $x < h^*$, using that χ is increasing and thus $\chi' \geq 0$, this equality can be extended to inequality

$$\chi'(x) \leq E_Y\Big[\chi'\Big(Y + \frac{x}{d}\Big)\Big] + e(x) \qquad \text{for all } x \in \mathbb{R}. \tag{A.2}$$

Similarly as in the proof of Proposition 3.1 in [AČ20], we obtain an upper bound on χ' by iterating (A.2) an appropriate amount of times. For this, let $Y_1, \dots, Y_k$ be independent Gaussian random variables having the same distribution as Y, and define $Z_i = Y_1/d^{i-1} + Y_2/d^{i-2} + \cdots + Y_i$. Note that Z_i is a centred Gaussian random variable whose variance, denoted σ_i^2, is bounded uniformly in i. Then, by applying (A.2) k-times,

$$\begin{aligned}\chi'(x) &\leq E_{Y_1}\Big[E_{Y_2}\Big[\chi'\Big(Y_2 + \frac{Y_1 + x/d}{d}\Big) + e(Y_1 + x/d) + e(x)\Big]\Big]\\ &\leq \cdots \leq E\Big[\chi'\Big(\frac{x}{d^k} + \frac{Y_1}{d^{k-1}} + \frac{Y_2}{d^{k-2}} + \cdots + Y_k\Big)\Big] + E\Big[\sum_{i=0}^{k-1} e\Big(\frac{x}{d^i} + Z_i\Big)\Big]\\ &\leq E\Big[\chi'\Big(\frac{x}{d^k} + Z_k\Big)\Big] + \sum_{i=0}^{k-1} E\Big[e\Big(\frac{x}{d^i} + Z_i\Big)\Big].\end{aligned} \tag{A.3}$$

We now choose $k = k(x) = \lfloor \log_d(x) \rfloor$. Then the boundedness of the first summand can be proved exactly in the same way as the boundedness of $E[\chi(x/d^k + Z_k)]$ in the proof of Proposition 3.1 in [AČ20] (note that $(d-1)$ in [AČ20] corresponds to d in our setting), one only needs to verify that $\chi' \in L^2(\nu)$. This can be easily proved using (A.1) and the facts $\rho_Y'(x) = -cx\rho_Y(x)$, $L[\chi] = \chi$ and (3.18) which imply $|\chi'(x)| \leq cL[\chi^2](x) + cx^2$ for $x \geq h^*$. From this $\chi' \in L^2(\nu)$ follows from Proposition 3.2(b).

To bound the second summand on the right hand side of (A.3), we first observe that by the choice of $k(x)$ it holds that $x_0 := x/d^{k(x)} \in [1, d]$. With this notation, by rearranging the sum,

$$\sum_{i=0}^{k(x)-1} E[e(x/d^i + Z_i)] = \sum_{i=0}^{k(x)-1} E[e(x_0 d^i + Z_{k(x)-i})]. \tag{A.4}$$

Moreover, it is easy to see that for some constants c, c'

$$0 \leq e(x) := 1_{x \geq h^*} \chi(h^*) \rho_Y\Big(h^* - \frac{x}{d}\Big) \leq c' 1_{x \geq h^*} e^{-cx}. \tag{A.5}$$

Therefore,

$$\sum_{i=0}^{k(x)-1} E[e(x_0 d^i + Z_{k(x)-i})] \leq c' \sum_{i=0}^{\infty} E\big[e^{-c(x_0 d^i + Z_{k(x)-i})}\big] \leq c' \sum_{i=0}^{\infty} e^{-cx_0 d^i} e^{c^2 \sigma_{k(x)-i}^2/2}, \tag{A.6}$$

which, since σ_i^2 are bounded and $x_0 \in [1, d]$, is clearly bounded uniformly in $x \geq h^*$. □

Appendix B. Proof of the many-to-few formulas

We give here a proof of Proposition 3.5. Throughout the proof we use the notation from Section 3.2. We start by introducing a useful standard martingale related to φ when viewed as branching process.

Lemma B.1. *For $x \in \mathbb{R}$, $n \in \mathbb{N}$, let*

$$\zeta(x, n) := d^n \chi(x) \qquad \text{and} \qquad \zeta_n := \zeta(\xi_n^1, n). \tag{B.1}$$

Then, for every $k \in \mathbb{N}$, $(\zeta_n)_{n \in \mathbb{N}}$ is a P_x^k-martingale with respect to $\mathcal{F}_n^k$.

Proof. By definition of ζ_n, for $n \geq 1$,

$$d^{-(n-1)} E_x^k[\zeta_n \mid \mathcal{F}_{n-1}^k] = dE_x^k[\chi(\varphi_{\sigma_n^1}) \mid \mathcal{F}_{n-1}^k], \tag{B.2}$$

where σ_n^1 is the vertex carrying the first spine mark at level n. Conditionally on $\mathcal{F}_{n-1}^k$, σ_n^1 is uniformly distributed on $\mathrm{desc}(\sigma_{n-1}^1)$ and independent of φ. Therefore, this equals

$$\sum_{v \in \mathrm{desc}(\sigma_{n-1}^1)} E_x^k[\chi(\varphi_v) \mid \mathcal{F}_{n-1}^k] = L[\chi](\varphi_{\sigma_{n-1}^1}) = \chi(\varphi_{\sigma_{n-1}^1}) = \chi(\xi_{n-1}^1). \tag{B.3}$$

where in for the first equality we used (3.7) and the fact that $\chi(x) = 0$ on $(-\infty, h^*)$, and where the second equality follows from (3.10). The martingale property of ζ_n then follows directly from this computation. □

Due to Lemma B.1, Lemma 8 of [HR17] can directly be applied to our process. We restate it here for reader's convenience, with very minor adaptations coming from the fact that in our process every node has always d descendants (some of them might be in the cemetery state).

Lemma B.2 (Lemma 8 in [HR17])**.** *For any $k \geq 1$, let Y be a $\mathcal{F}_n^k$-measurable random variable which can be written as*

$$Y = \sum_{v_1,\dots,v_k \in S_n^+} Y(v_1,\dots,v_k) 1_{\{\sigma_n^1 = v_1,\dots,\sigma_n^k = v_k\}}, \tag{B.4}$$

where, for every $v_1,\dots,v_k \in S_n^+$, $Y(v_1,\dots,v_k)$ is a $\mathcal{F}_n$-measurable random variable. Then

$$E_x\Big[\sum_{v_1,\dots,v_k \in N_n^+} Y(v_1,\dots,v_k)\Big] = Q_x^k\Big[Y \prod_{v \in \mathrm{skel}(n)\setminus\{o\}} \frac{\zeta(\varphi_{p(v)}, |v|-1)}{\zeta(\varphi_v, |v|)} d^{l_{p(v)}}\Big]. \tag{B.5}$$

We are now ready to give the proof of Proposition 3.5.

Proof of Proposition 3.5. Statement (3.26) follows directly from Lemma B.2 with $k = 1$ and $Y(v_1) = f(\varphi_{v_1})$. Indeed, since $k = 1$, there is only one mark on every level and thus $Y = f(\xi_n^1)$, $\mathrm{skel}(n) \setminus \{o\} = \{\sigma_1^1, \dots, \sigma_n^1\}$, and $l_{p(v)} = 1$ for all $v \in \mathrm{skel}(n) \setminus \{o\}$. Therefore, by (B.5)

$$\begin{aligned} E_x\Big[\sum_{v \in N_n^+} f(\varphi_v)\Big] &= Q_x\Big[f(\xi_n) \prod_{v \in \{\sigma_1,\dots,\sigma_n\}} \Big(\frac{\zeta(\varphi_{p(v)}, |v|-1)}{\zeta(\varphi_v, |v|)} d\Big)\Big] \\ &= Q_x\Big[f(\xi_n) \prod_{i=1}^n \Big(\frac{\chi(\xi_{i-1}) d^{i-1}}{\chi(\xi_i) d^i} d\Big)\Big] = Q_x\Big[f(\xi_n^1) \frac{\chi(x)}{\chi(\xi_n)}\Big], \end{aligned} \tag{B.6}$$

as claimed in (3.26).

Similarly, statement (3.27) is a consequence of Lemma B.2 with $k = 2$. We set $Y(v_1, v_2) = f(\varphi_{v_1}) g(\varphi_{v_2})$. Then $Y = f(\xi_n^1) g(\xi_n^2)$ and by (B.5),

$$E_x\Big[\sum_{v,w \in N_n^+} f(\varphi_v) g(\varphi_w)\Big] = Q_x^2\Big[f(\xi_n^1) g(\xi_n^2) \prod_{v \in \mathrm{skel}(n)\setminus\{o\}} \frac{\zeta(\varphi_{p(v)}, |v|-1)}{\zeta(\varphi_v, |v|)} d^{l_{p(v)}}\Big]. \tag{B.7}$$

To consider the possible structures of $\mathrm{skel}(n) \setminus \{o\}$, set $s = \max\{k : \sigma_k^1 = \sigma_k^2\}$ to be the last time where the two spines agree. Note that due to the dynamics of marks under Q_x^2,

$$Q_x^2[s = k] = \begin{cases} (d-1)d^{-(k+1)}, & \text{if } k \in \{0,\dots,n-1\}, \\ d^{-n}, & \text{if } k = n, \end{cases} \tag{B.8}$$

and, by a simple computation,

$$\prod_{v\in\mathrm{skel}(n)\setminus\{o\}} \frac{\zeta(\varphi_{p(v)},|v|-1)}{\zeta(\varphi_v,|v|)} d^{l_{p(v)}} = \prod_{v\in\mathrm{skel}(n)\setminus\{o\}} \left(\frac{\chi(\varphi_{p(v)})}{\chi(\varphi_v)}\frac{1}{d}\right) d^{l_{p(v)}} = \frac{\chi(\xi_s^1)\chi(x)}{\chi(\xi_n^1)\chi(\xi_n^2)} d^s. \tag{B.9}$$

Therefore, (B.7) can be written as

$$\begin{aligned} E_x\Bigg[\sum_{v\in N_n^+} f(\varphi_v)\sum_{v\in N_n^+} g(\varphi_v)\Bigg] &= \sum_{k=0}^{n} d^k Q_x^2[s=k] Q_x^2\left[f(\xi_n^1)g(\xi_n^2)\frac{\chi(\xi_k^1)\chi(x)}{\chi(\xi_n^1)\chi(\xi_n^2)}\middle| s=k\right] \\ &= \frac{d-1}{d}\sum_{k=0}^{n-1} Q_x^2\left[\frac{f(\xi_n^1)}{\chi(\xi_n^1)}\frac{g(\xi_n^2)}{\chi(\xi_n^2)}\chi(\xi_s^1)\chi(x)\middle| s=k\right] \\ &\quad + Q_x^2\left[f(\xi_n^1)g(\xi_n^1)\frac{\chi(x)}{\chi(\xi_n^1)}\middle| s=n\right]. \end{aligned} \tag{B.10}$$

By construction, under Q_x^2, conditional on $s=k$, for times $i=1,\ldots,k$ the processes ξ_i^1 and ξ_i^2 are Markov chains which follow the same trajectory and have the same dynamics as ξ_i under Q_x^1. Further, for later times $i=k+1,\ldots,n$, they are independent Markov chains distributed according to $Q_{\xi_k^1}$. Therefore,

$$Q_x^2\left[\frac{f(\xi_n^1)}{\chi(\xi_n^1)}\frac{g(\xi_n^2)}{\chi(\xi_n^2)}\chi(\xi_s^1)\chi(x)\middle| s=k\right] = \chi(x)Q_x\left[\chi(\xi_k)Q_{\xi_k^1}^2\left[\frac{f(\xi_{n-k}^1)}{\chi(\xi_{n-k}^1)}\right]Q_{\xi_k^1}^2\left[\frac{g(\xi_{n-k}^2)}{\chi(\xi_{n-k}^2)}\right]\right], \tag{B.11}$$

and similarly

$$Q_x^2\left[f(\xi_n^1)g(\xi_n^1)\frac{\chi(x)}{\chi(\xi_n^1)}\middle| s=n\right] = \chi(x)Q_x\left[\frac{f(\xi_n)g(\xi_n)}{\chi(\xi_n)}\right]. \tag{B.12}$$

Combining (B.10)–(B.12) directly implies (3.27). □

References

[AČ20] Angelo Abächerli and Jiří Černý, *Level-set percolation of the Gaussian free field on regular graphs I: regular trees*, Electron. J. Probab. **25** (2020), Paper No. 65, 24. MR4115734

[ADH13] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit, *A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces*, Electron. J. Probab. **18** (2013), no. 14, 21. MR3035742

[Ald91] David Aldous, *The continuum random tree. I*, Ann. Probab. **19** (1991), no. 1, 1–28. MR1085326

[BDIM23] Dariusz Buraczewski, Congzao Dong, Alexander Iksanov, and Alexander Marynych, *Critical branching processes in a sparse random environment*, Mod. Stoch. Theory Appl. **10** (2023), no. 4, 397–411. MR4655407

[BFRS24] Florin Boenkost, Félix Foutel-Rodier, and Emmanuel Schertzer, *The genealogy of nearly critical branching processes in varying environment*, Preprint, available at arXiv:2207.11612, 2024.

[BLM87] Jean Bricmont, Joel L. Lebowitz, and Christian Maes, *Percolation in strongly correlated systems: the massless Gaussian field*, J. Statist. Phys. **48** (1987), no. 5-6, 1249–1268. MR914444

[CD24] Zhenhao Cai and Jian Ding, *One-arm probabilities for metric graph gaussian free fields below and at the critical dimension*, Preprint, available at arXiv:2406.02397, 2024.

[CD25] Zhenhao Cai and Jian Ding, *One-arm exponent of critical level-set for metric graph Gaussian free field in high dimensions*, Probability Theory and Related Fields **191** (2025), no. 3, 1035–1120. MR4898098

[CKKM24] Guillaume Conchon-Kerjan, Daniel Kious, and Cécile Mailler, *Scaling limit of critical random trees in random environment*, Electron. J. Probab. **29** (2024), Paper No. 112, 53. MR4779872

[ČL25] Jiří Černý and Ramon Locher, *Critical and near-critical level-set percolation of the Gaussian free field on regular trees*, 2025, pp. 746–767. MR4863059

[CN20] Alberto Chiarini and Maximilian Nitzschner, *Entropic repulsion for the Gaussian free field conditioned on disconnection by level-sets*, Probab. Theory Related Fields **177** (2020), no. 1-2, 525–575. MR4095021

[CR90] B. Chauvin and A. Rouault, *Supercritical branching Brownian motion and K-P-P equation in the critical speed-area*, Math. Nachr. **149** (1990), 41–59. MR1124793

[CTJP24] Natalia Cardona-Tobón, Arturo Jaramillo, and Sandra Palau, *Rates on Yaglom's limit for Galton-Watson processes in a varying environment*, ALEA Lat. Am. J. Probab. Math. Stat. **21** (2024), no. 1, 1–23. MR4703767

[DCGRS23] Hugo Duminil-Copin, Subhajit Goswami, Pierre-François Rodriguez, and Franco Severo, *Equality of critical parameters for percolation of Gaussian free field level sets*, Duke Math. J. **172** (2023), no. 5, 839–913. MR4568695

[DLG02] Thomas Duquesne and Jean-François Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque, no. 281, Société mathématique de France, 2002 (en). MR1954248

[DPR18] Alexander Drewitz, Alexis Prévost, and Pierre-Françcois Rodriguez, *The sign clusters of the massless Gaussian free field percolate on $\mathbb{Z}^d, d \geqslant 3$ (and more)*, Comm. Math. Phys. **362** (2018), no. 2, 513–546. MR3843421

[DPR25] Alexander Drewitz, Alexis Prévost, and Pierre-François Rodriguez, *Critical one-arm probability for the metric Gaussian free field in low dimensions*, Probability Theory and Related Fields (2025).

[dR17] Loïc de Raphélis, *Scaling limit of multitype Galton-Watson trees with infinitely many types*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 200–225. MR3606739

[DRS14] Alexander Drewitz, Balázs Ráth, and Artëm Sapozhnikov, *On chemical distances and shape theorems in percolation models with long-range correlations*, J. Math. Phys. **55** (2014), no. 8, 083307, 30. MR3390739

[Dur19] Rick Durrett, *Probability—theory and examples*, fifth ed., Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49, Cambridge University Press, Cambridge, 2019. MR3930614

[Fel71] William Feller, *An introduction to probability theory and its applications. Vol. II.*, Second edition, John Wiley & Sons Inc., New York, 1971. MR0270403

[GLLP22] Ion Grama, Ronan Lauvergnat, and Émile Le Page, *Limit theorems for critical branching processes in a finite-state-space Markovian environment*, Adv. in Appl. Probab. **54** (2022), no. 1, 111–140. MR4397862

[GPW09] Andreas Greven, Peter Pfaffelhuber, and Anita Winter, *Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees)*, Probab. Theory Related Fields **145** (2009), no. 1-2, 285–322. MR2520129

[Gri99] Geoffrey Grimmett, *Percolation*, Grundlehren der Mathematischen Wissenschaften, vol. 321, Springer-Verlag, Berlin, 1999. MR1707339

[GRS22] Subhajit Goswami, Pierre-François Rodriguez, and Franco Severo, *On the radius of Gaussian free field excursion clusters*, Ann. Probab. **50** (2022), no. 5, 1675–1724. MR4474499

[HH07] J. W. Harris and S. C. Harris, *Survival probabilities for branching Brownian motion with absorption*, Electron. Comm. Probab. **12** (2007), 81–92. MR2300218

[HHKW22] Simon C. Harris, Emma Horton, Andreas E. Kyprianou, and Minmin Wang, *Yaglom limit for critical nonlocal branching Markov processes*, Ann. Probab. **50** (2022), no. 6, 2373–2408. MR4499840

[HR17] Simon C. Harris and Matthew I. Roberts, *The many-to-few lemma and multiple spines*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 226–242. MR3606740

[Kol38] Andrey Nikolaevich Kolmogorov, *Zur Lösung einer biologischen Aufgabe*, Comm. Math. Mech. Chebyshev Univ. Tomsk **2** (1938), no. 1, 1–12.

[LG05] Jean-François Le Gall, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311. MR2203728

[LS86] Joel L. Lebowitz and H. Saleur, *Percolation in strongly correlated systems*, Phys. A **138** (1986), no. 1-2, 194–205. MR865243

[Mie08] Grégory Miermont, *Invariance principles for spatial multitype Galton-Watson trees*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 6, 1128–1161. MR2469338

[Mod71] Charles J. Mode, *Multitype branching processes. Theory and applications*, Modern Analytic and Computational Methods in Science and Mathematics, No. 34, American Elsevier Publishing Co., Inc., New York, 1971. MR0279901

[MS83] S. A. Molchanov and A. K. Stepanov, *Percolation in random fields. I*, Teoret. Mat. Fiz. **55** (1983), no. 2, 246–256. MR734878

[MS22] Pascal Maillard and Jason Schweinsberg, *Yaglom-type limit theorems for branching Brownian motion with absorption*, Ann. H. Lebesgue **5** (2022), 921–985. MR4526243

[Pow19] Ellen Powell, *An invariance principle for branching diffusions in bounded domains*, Probab. Theory Related Fields **173** (2019), no. 3-4, 999–1062. MR3936150

[PR15] Serguei Popov and Balázs Ráth, *On decoupling inequalities and percolation of excursion sets of the Gaussian free field*, J. Stat. Phys. **159** (2015), no. 2, 312–320. MR3325312

[PS22] Christoforos Panagiotis and Franco Severo, *Analyticity of Gaussian free field percolation observables*, Comm. Math. Phys. **396** (2022), no. 1, 187–223. MR4499015

[RS13] Pierre-François Rodriguez and Alain-Sol Sznitman, *Phase transition and level-set percolation for the Gaussian free field*, Comm. Math. Phys. **320** (2013), no. 2, 571–601. MR3053773

[RY99] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin, 1999. MR1725357

[Szn16] Alain-Sol Sznitman, *Coupling and an application to level-set percolation of the Gaussian free field*, Electron. J. Probab. **21** (2016), 1–26. MR3492939

[Szn19] Alain-Sol Sznitman, *On macroscopic holes in some supercritical strongly dependent percolation models*, Ann. Probab. **47** (2019), no. 4, 2459–2493. MR3980925

[Yag47] A. M. Yaglom, *Certain limit theorems of the theory of branching random processes*, Doklady Akad. Nauk SSSR (N.S.) **56** (1947), 795–798. MR22045