CENTRAL LIMIT THEOREM FOR BIASED RANDOM WALK ON MULTI-TYPE GALTON-WATSON TREES

AMIR DEMBO* AND NIKE SUN†

Abstract. Let $\mathcal{T}$ be a rooted supercritical multi-type Galton-Watson (MGW) tree with types coming from a finite alphabet, conditioned to non-extinction. The λ-biased random walk $(X_t)_{t\geq 0}$ on $\mathcal{T}$ is the nearest-neighbor random walk which, when at a vertex v with d_v offspring, moves closer to the root with probability $\lambda/(\lambda+d_v)$, and to each of the offspring with probability $1/(\lambda+d_v)$. This walk is recurrent for $\lambda \geq \rho$ and transient for $0 \leq \lambda < \rho$, with ρ the Perron-Frobenius eigenvalue for the (assumed) irreducible matrix of expected offspring numbers. Subject to finite moments of order $p > 4$ for the offspring distributions, we prove the following quenched CLT for λ-biased random walk at the critical value $\lambda = \rho$: for almost every $\mathcal{T}$, the process $|X_{\lfloor nt \rfloor}|/\sqrt{n}$ converges in law as $n \to \infty$ to a reflected Brownian motion rescaled by an explicit constant. This result was proved under some stronger assumptions by Peres and Zeitouni (2008) for single-type Galton-Watson trees. Following their approach, our proof is based on a new explicit description of a reversing measure for the walk from the point of view of the particle (generalizing the measure constructed in the single-type setting by Peres and Zeitouni), and the construction of appropriate harmonic coordinates. In carrying out this program we prove a zero-one law and moment and conductance estimates for MGW trees, which may be of independent interest. In addition, we extend our construction of the reversing measure to a biased random walk with random environment (RWRE) on MGW trees, again at a critical value of the bias. We compare this result against a transience-recurrence criterion for the RWRE generalizing a result of Faraud (2008) for Galton-Watson trees.

1. Introduction

Let $\mathcal{T}$ denote an infinite tree with root o. The λ-biased random walk on $\mathcal{T}$, hereafter denoted $\mathsf{RW}_\lambda(\mathcal{T})$, is the Markov chain $(X_t)_{t\geq 0}$ with $X_0 = o$ such that given $X_t = v$ with offspring number d_v and $v \neq o$, X_{t+1} equals the parent of v with probability $\lambda/(\lambda + d_v)$, and is uniformly distributed among the offspring of v otherwise (and if $X_t = o$, then X_{t+1} is uniformly distributed among the offspring of o).

For supercritical Galton-Watson trees without leaves, if ρ denotes the mean offspring number, then RW_λ is a.s. recurrent if and only if $\lambda \geq \rho$ ([20, Thm. 4.3 and Propn. 6.4]), and ergodic if and only if $\lambda > \rho$ ([15, Propn. 9-131] and [20, p. 944 and p. 954]). With $|v|$ denoting the (graph) distance from vertex v to the root o, $|X_t|/t$ converges a.s. to a speed V, with $V = V(\lambda)$ deterministic, positive for $\lambda < \rho$ and zero otherwise (see [22, 23] for $\lambda < \rho$ and [27] for $\lambda = \rho$; the case $\lambda > \rho$ follows trivially from positive recurrence).

Date: June 12, 2011.

2010 *Mathematics Subject Classification.* Primary 60F05, 60K37; Secondary 60J80, 60G50.

Key words and phrases. Multi-type Galton-Watson tree, biased random walk, central limit theorem, random walk with random environment.

*Departments of Mathematics and Statistics, Stanford University. Research partially supported by NSF grant DMS-0806211.

†Department of Statistics, Stanford University. Research partially supported by NSF grants DMS-0806211 and DMS-0502385, and Department of Defense (AFRL/AFOSR) NDSEG Fellowship.

Further, subject to no leaves and finite exponential moments for the offspring distribution, a quenched CLT for RW_λ ($\lambda \le \rho$) on single-type Galton-Watson trees was shown by Peres and Zeitouni [27], and extended to the setting of random walk with random environment (RWRE) by Faraud [10]. In contrast, if leaves occur, there emerges a zero-speed transient regime $\lambda < \lambda_c$ (for $\lambda_c < \rho$) [23] where the leaves "trap" the random walk and create slow-down. It follows from the results of Ben Arous et al. [2] that in this setting, for sufficiently small λ there cannot be a (functional) CLT with diffusive scaling. In this paper we consider the critical case $\lambda = \rho$, where [27, Thm. 1] proves that on a.e. Galton-Watson tree, the processes $(|X_{\lfloor nt \rfloor}|/\sqrt{n})_{t\ge 0}$ converge in law to the absolute value of a (deterministically) scaled Brownian motion. Their proof is based on the construction of harmonic coordinates and an explicit description of a reversing probability measure IGWR for RW_ρ "from the point of view of the particle." Having such an explicit description is a very delicate property: even for Galton-Watson trees, no such description is known for $\lambda < \rho$ except at $\lambda = 1$ which is done by [22, Thm. 3.1]. One thus might be led to believe that [27, Thm. 1] is a particular property resulting from the independence inherent in the Galton-Watson law.

Here we show to the contrary that such a quenched CLT extends to the much larger family of supercritical multi-type Galton-Watson trees with finite type space. We allow for leaves (but condition on non-extinction), demonstrating that at $\lambda = \rho$ the "trapping" phenomenon of [2] does not arise. We also replace the assumption of exponential moments for the offspring distribution by an assumption of finite moments of order $p > 4$, so that our result restricted to the single-type case strengthens [27, Thm. 1]. However, the main interest of our result lies in moving from an i.i.d. to a Markovian structure for the random tree.

As in [27], the key ingredient in our proof is the construction of an explicit reversing (probability) measure IMGWR for RW_λ from the point of view of the particle, generalizing IGWR to the multi-type setting, for λ at the critical value on the boundary between transience and recurrence. See §2 for the details of the construction which may be of independent interest.

The model we consider is as follows: let Ω be the space of rooted trees with type, where each vertex v is given a type χ_v from a finite alphabet $\mathcal{Q}$. We let $\mathcal{B}_\Omega$ be the σ-algebra on Ω generated by the cylinder sets (determined by the restrictions of trees to finite neighborhoods of the root). We write $\mathcal{T}$ for a generic element of Ω and o for its root. A *multi-type Galton-Watson* tree is a random element $\mathcal{T} \in \Omega$, generated from a starting type $a_0 \in \mathcal{Q}$ and a collection of probability measures $\mathbf{q}^a$ ($a \in \mathcal{Q}$) on

$$\mathcal{Q}^* := \bigcup_{\ell \ge 0} \mathcal{Q}^\ell,$$

as follows: begin with a root vertex o of type $\chi_o = a_0$. Supposing inductively that the first n levels of $\mathcal{T}$ have been constructed, each vertex v at the n-th level generates random offspring according to law $\mathbf{q}^{\chi_v}$. For our purposes the ordering of the children does not matter, so each $\mathbf{q}^a$ may equivalently be regarded as a probability measure on configurations $\underline{x} = (x_b)_{b\in\mathcal{Q}} \in (\mathbb{Z}_{\ge 0})^{\mathcal{Q}}$, where x_b is the number of children of type b. Continuing to construct successive generations in this Markovian fashion, we denote the resulting law on $(\Omega, \mathcal{B}_\Omega)$ by MGW^{a_0}. We denote by MGW any mixture of the measures MGW^{a_0}, $a_0 \in \mathcal{Q}$, and let $\mathbb{X}$ denote the event of extinction.

For $a, b \in \mathcal{Q}$ let

$$A(a,b) = \sum_{\underline{x}} \mathbf{q}^a(\underline{x}) x_b,$$

the expected number of offspring of type b at a vertex of type a. (Unless otherwise specified, the implicit assumption hereafter is that $\mathbb{E}_{\mathbf{q}^a}[|\underline{x}|] < \infty$ for all $a \in \mathcal{Q}$ where $|\underline{x}| \equiv \sum_b x_b$.) Throughout the paper we will refer to the following assumptions:

(H1) The matrix $A \equiv (A(a,b))_{a,b\in\mathcal{Q}}$ is irreducible with Perron-Frobenius eigenvalue ρ.

(H2) A is positive regular (every entry of A^{n_0} is positive for some $n_0 \in \mathbb{N}$), $\rho > 1$, and $\mathbb{E}_{\mathbf{q}^a}[|\underline{x}|\log|\underline{x}|] < \infty$ for all $a \in \mathcal{Q}$.

(H3^p) $\mathbb{E}_{\mathbf{q}^a}[|\underline{x}|^p] < \infty$ for all $a \in \mathcal{Q}$.

Note that (H1) and $\rho > 1$ together imply $\mathsf{MGW}^a(\mathbb{X}) < 1$ for all $a \in \mathcal{Q}$. We take all real-valued processes to be in the space $D[0,\infty)$ equipped with the topology of uniform convergence on compact intervals. Our main theorem is the following:

Theorem 1.1. *Under* (H1), (H2), *and* (H3^p) *with $p > 4$, for* MGW*-a.e.* $\mathcal{T} \notin \mathbb{X}$*, if* $X \sim \mathsf{RW}_\rho(\mathcal{T})$ *then the processes* $(|X_{\lfloor nt\rfloor}|/(\sigma\sqrt{n}))_{t\geq 0}$ *converge in law in* $D[0,\infty)$ *to the absolute value of a standard Brownian motion for* σ *a deterministic positive constant (see* (2.5)*).*

Let $\mathsf{RW}^{\text{cts}}_\lambda(\mathcal{T})$ denote the continuous-time version of $\mathsf{RW}_\lambda(\mathcal{T})$, which when at $v \in \mathcal{T}$ moves to the parent of v (if $v \neq o$) at rate λ and to each offspring of v at rate 1.

Corollary 1.2. *Under the assumptions of Thm. 1.1, for* MGW*-a.e.* $\mathcal{T} \notin \mathbb{X}$*, if* $X^{\text{cts}} \sim \mathsf{RW}^{\text{cts}}_\rho(\mathcal{T})$ *then the processes* $(|X^{\text{cts}}_{nt}|/(\sigma\sqrt{2\rho n}))_{t\geq 0}$ *converge in law in* $D[0,\infty)$ *to the absolute value of a standard Brownian motion.*

By moving the root of the tree to the current position of the random walk, RW_λ on the tree induces a random walk on the space Ω, the "walk from the point of view of the particle." As in [27, §3], to make the latter process Markovian we amend the state space so as to keep track of the ancestry of the vertices. Specifically, we consider the space $\Omega_\downarrow$ of pairs $(\mathcal{T},\xi)$, where $\mathcal{T}$ is an infinite tree and $\xi = (o = v_0, v_1, v_2, \ldots)$ is a ray emanating from the root o; this ray indicates the ancestry of each vertex in the tree. Let $\mathcal{B}_{\Omega_\downarrow}$ denote the σ-algebra generated by the cylinder sets. We define a height function h on $\mathcal{T}$ as follows: set $h(v_n) = -n$, and for $v \notin \xi$ set

$$h(v) = h(R_v) + d(v,\xi) \tag{1.1}$$

where d denotes graph distance and R_v is the nearest vertex to v on ξ (see Fig. 1). We denote by $\mathsf{RW}_\lambda(\mathcal{T},\xi)$ the λ-biased random walk $(Y_t)_{t\geq 0}$ on $(\mathcal{T},\xi)$, where the bias goes in the direction of decreasing height. With $\mathcal{T}^v$ the tree $\mathcal{T}$ rooted at v instead of o, and ξ^v the unique ray emanating from v such that $\xi \cap \xi^v$ is an infinite ray, let

$$(\mathcal{T},\xi)^{Y_t} := (\mathcal{T}^{Y_t}, \xi^{Y_t}), \quad t \geq 0. \tag{1.2}$$

This is a Markov process with state space $\Omega_\downarrow$, and we hereafter refer to it as $\mathcal{T}\mathsf{RW}_\lambda$. Let $\mathsf{RW}^{\text{cts}}_\lambda$ denote the continuous-time version of $\mathsf{RW}_\lambda(\mathcal{T},\xi)$ (moving in the direction of increasing height at rate 1 and in the direction of decreasing height at rate λ), and let $\mathcal{T}\mathsf{RW}^{\text{cts}}_\lambda$ denote the induced continuous-time process on the space $\Omega_\downarrow$.

As in the single-type Galton-Watson case considered in [27], the key to our proof lies in finding an explicit reversing measure IMGW for $\mathcal{T}\mathsf{RW}^{\text{cts}}_\rho$, which is then easily translated to a reversing measure IMGWR for $\mathcal{T}\mathsf{RW}_\rho$. For a tree $\mathcal{T}$ (with or without marked ray) and for any vertex $v \in \mathcal{T}$, we denote by $\mathcal{T}^{(v)}$ the subtree induced by v and its descendants, where descent is in direction of increasing distance from the root for a rooted tree, and in the direction of increasing height for a tree with marked ray. If μ is a law on trees we use $\mu \otimes \mathsf{RW}_\lambda$ to denote the joint law of the tree and the realization of RW_λ on the tree.

Theorem 1.3. *Assume* (H1).

(a) There exists a reversing probability measure IMGW *for* $\mathcal{T}\mathsf{RW}^{\text{cts}}_\rho$*, and if we define*

$$\frac{d\mathsf{IMGWR}}{d\mathsf{IMGW}} = \frac{d_o + \rho}{2\rho},$$

then IMGWR *is a reversing probability measure for* $\mathcal{T}\mathsf{RW}_\rho$.

(b) If $((\mathcal{T},\xi),(Y_t)_{t\geq 0}) \sim \mathsf{IMGWR}\otimes\mathsf{RW}_\rho$ *then the stationary sequence* $(\mathcal{T}^{(Y_t)})_{t\geq 0}$ *is ergodic.*

The IMGW trees always have an infinite ray ξ, though the trees coming off the ray may be finite. The measures IMGW, IMGWR are the multi-type analogues of the measures IGW, IGWR of [27]. Thm. 1.3 and the construction of harmonic coordinates allow us to prove the following quenched CLT for RW_ρ on IMGWR trees, which will be used to deduce Thm. 1.1.

Theorem 1.4. *Under* (H1), (H2), *and* (H3^p) *with* $p>2$, *for* IMGWR*-a.e.* $(\mathcal{T},\xi)$, *if* $Y\sim\mathsf{RW}_\rho(\mathcal{T},\xi)$ *then the processes* $(h(Y_{\lfloor nt\rfloor})/(\sigma\sqrt{n}))_{t\geq 0}$ *converge in law in* $D[0,\infty)$ *to a standard Brownian motion.*

In the setting of RW_λ on MGW trees, $\lambda=\rho$ represents the onset of recurrence. Indeed, MGW-a.e. tree $\mathcal{T}$ has branching number $\operatorname{br}\mathcal{T}=\rho$ [20, Propn. 6.5], therefore $\mathsf{RW}_\lambda(\mathcal{T})$ is transient for $\lambda<\rho$ and recurrent for $\lambda>\rho$ [20, Thm. 4.3]. In fact, recurrence for all $\lambda\geq\rho$ follows from a simple conductance calculation (for the general theory see [24, Ch. 2]), therefore ρ is the boundary between transience and recurrence for RW_λ on MGW trees. Further ρ is the boundary between non-ergodicity and ergodicity, with RW_ρ null recurrent (see [15, Propn. 9-131] and [20, p. 944 and p. 954]) and of zero speed (e.g. from the bound of Lem. 3.5).

We believe that the existence of a reversing measure and CLT is a feature of the onset of recurrence in a more general setting. Indeed, we will consider random walk with random environment (RWRE) on MGW trees, again with an adjustable bias parameter, defined formally as follows: we continue to suppose a finite type space $\mathcal{Q}$, but for each $a\in\mathcal{Q}$ we now let $\mathbf{q}^a$ be a probability measure on

$$\bar{\mathcal{Q}}^* := \bigcup_{\ell\geq 0}\bar{\mathcal{Q}}^\ell, \quad \bar{\mathcal{Q}}=\mathcal{Q}\times\mathbb{R}_{>0};$$

we write an element of $\bar{\mathcal{Q}}^\ell$ as

$$(\underline{y},\underline{\alpha})\equiv((y_1,\alpha_1),\ldots,(y_\ell,\alpha_\ell)).$$

The collection of distributions $\mathbf{q}^a$ $(a\in\mathcal{Q})$, together with a starting type $a_0\in\mathcal{Q}$, specifies a law $\overline{\mathsf{MGW}}^{a_0}$ on the space Ω of rooted trees where each vertex v has a type $\chi_v\in\mathcal{Q}$, and every $v\neq o$ has a weight $\alpha_v\in\mathbb{R}_{>0}$. As before we let $\overline{\mathsf{MGW}}$ denote any mixture of the $\overline{\mathsf{MGW}}^{a_0}$. Fixing such a tree $\mathcal{T}$ (the environment), the λ-biased random walk on $\mathcal{T}$, which we denote RWRE_λ, is the Markov chain $(X_t)_{t\geq 0}$ with $X_0=o$ which, at $v\neq o$ with parent u and offspring $(\underline{y},\underline{\alpha})\in\bar{\mathcal{Q}}^\ell$, moves to the j-th offspring with probability $\alpha_j/(\lambda+\sum_{i=1}^\ell\alpha_i)$ and to the parent with probability $\lambda/(\lambda+\sum_{i=1}^\ell\alpha_i)$. If $X_t=o$ then X_{t+1} is chosen from the offspring w of o according to weights α_w. (Note that RW_ρ corresponds to the case $\alpha_v=1$ for all v.) We let $\mathsf{RWRE}_\lambda^{\mathrm{cts}}$ denote the continuous-time version of RWRE_λ. This model, studied for Galton-Watson trees in [10], allows for quite general distributions on each neighborhood (a vertex v, its type, and its offspring vector $(\underline{y}^v,\underline{\alpha}^v)\in\bar{\mathcal{Q}}^*$), but conditioned on types the weights in different neighborhoods must be independent. For $\gamma\in\mathbb{R}$ and $a,b\in\mathcal{Q}$, let

$$\bar{A}^{(\gamma)}(a,b)\equiv\int\sum_j \mathbf{1}_{\{y_j=b\}}\alpha_j^\gamma\, d\mathbf{q}^a(\underline{y},\underline{\alpha}) \tag{1.3}$$

(not necessarily finite for all γ). Let $\bar{\rho}(\gamma)$ be the Perron-Frobenius eigenvalue of $\bar{A}^{(\gamma)}$ where well-defined (i.e. where $\bar{A}^{(\gamma)}$ has finite entries and is irreducible), and ∞ otherwise. We will prove the following characterization of the transience-recurrence boundary for RWRE_λ, extending part of [10, Thm. 1.1]:

Theorem 1.5. *Suppose $\bar{A}^{(0)}$ is positive regular, and $\bar{\rho}(\gamma) < \infty$ for γ in an open neighborhood of 0. For $\lambda > 0$ let*

$$p_\lambda := \inf_{0 \le \gamma \le 1} \frac{\bar{\rho}(\gamma)}{\lambda^\gamma}.$$

(a) If $p_\lambda < 1$, then RWRE_λ is positive recurrent.
(b) If $p_\lambda > 1$, then RWRE_λ is transient.

Thus the recurrence-transience boundary occurs at the unique value ρ° for which $p_{\rho^\circ} = 1$. On the other hand, let $\Omega_\downarrow$ denote the space of trees with ray (where each vertex v has type χ_v and weight α_v), and let $\mathcal{T}\mathsf{RWRE}_\lambda$ and $\mathcal{T}\mathsf{RWRE}^{\text{cts}}_\lambda$ denote the Markov chains in $\Omega_\downarrow$ induced by RWRE_λ and $\mathsf{RWRE}^{\text{cts}}_\lambda$ respectively. We then have the following generalization of Thm. 1.3 (a):

Theorem 1.6. *Suppose $\overline{\mathsf{MGW}}$ is such that $\bar{A} \equiv \bar{A}^{(1)}$ is irreducible with Perron-Frobenius eigenvalue $\bar{\rho} \equiv \bar{\rho}(1)$. Then there exists a reversing probability measure $\overline{\mathsf{IMGW}}$ on $\Omega_\downarrow$ for $\mathcal{T}\mathsf{RWRE}^{\text{cts}}_{\bar{\rho}}$. If we let α_{0j} denote the weight for the j-th child of the root v_0, and set*

$$\frac{d\overline{\mathsf{IMGWR}}}{d\overline{\mathsf{IMGW}}} = \frac{\bar{\rho} + \sum_{j=1}^{d_o} \alpha_{0j}}{2\bar{\rho}},$$

then $\overline{\mathsf{IMGWR}}$ is a reversing probability measure for $\mathcal{T}\mathsf{RWRE}_{\bar{\rho}}$.

We can see that ρ° matches $\bar{\rho}$ if and only if the function $\gamma \mapsto \bar{\rho}(\gamma)/(\rho^\circ)^\gamma$ attains its infimum at $\gamma = 1$. If this fails, Thm. 1.6 still gives a reversing measure at $\bar{\rho}$, but $\bar{\rho} > \rho^\circ$ and the walk is already positive recurrent above ρ°. However, at least in the single-type case, we have $\rho^\circ = \bar{\rho}$ in all cases in which a CLT is possible: indeed, if

$$\kappa := \inf \left\{ \gamma \ge 0 : \frac{\bar{\rho}(\gamma)}{(\rho^\circ)^\gamma} = 1 \right\},$$

by results of [13] a CLT cannot hold unless $\kappa \ge 2$ (see [10, p. 3]). We expect $\kappa \ge 2$ also to be a necessary condition in the multi-type case, and thus Thm. 1.5 and Thm. 1.6 support the claim that reversing measures occur at the boundary between transience and recurrence *in cases in which a CLT is possible*. However, even in the single-type case the random environment creates technical difficulties, and the RWRE CLT of [10] requires some restriction on κ. While we expect that the methods of this paper and [10] can also be adapted to extend the RWRE CLT to the multi-type setting under the same restrictions on κ, new ideas are required to achieve a CLT for the entire regime $\kappa \ge 2$.

Outline of the paper.

- In §2.1 we construct the reversing measure IMGWR for $\mathcal{T}\mathsf{RW}_\rho$ (and, in §2.2, its generalization $\overline{\mathsf{IMGWR}}$ for $\mathcal{T}\mathsf{RWRE}_{\bar{\rho}}$). We show that if $((\mathcal{T}, \xi), (Y_t)_{t \ge 0}) \sim \mathsf{IMGWR} \otimes \mathsf{RW}_\rho$ then the stationary sequence $(\mathcal{T}^{(Y_t)})_{t \ge 0}$ is ergodic.

  Conditioning on $(\mathcal{T}, \xi) \sim \mathsf{IMGWR}$, we construct in §2.4 a function $v \mapsto S_v$ ($v \in \mathcal{T}$) which is harmonic with respect to the transition probabilities of $\mathsf{RW}_\rho(\mathcal{T})$. Then, if $Y \sim \mathsf{RW}_\rho(\mathcal{T}, \xi)$, $M_t := S_{Y_t}$ is a martingale. Using the stationarity and ergodicity of the process $(\mathcal{T}^{(Y_t)})_{t \ge 0}$ under IMGWR, we are able to control the quadratic variation of M_t and thus to show that it satisfies a CLT.
- In §3, adapting the methods of [27, 10], we prove Thm. 1.4 by showing that $h(Y_t)$ is uniformly well approximated by M_t/η, for η an explicit constant.

- In §4 we prove Thm. 1.1. First, using the "shifted coupling" introduced in [27] (see also [10]), given $(\mathcal{T}, (X_t)_{t\geq 0})$ we construct $((\widehat{\mathcal{T}}, \xi), (Y_t)_{t\geq 0})$ with marginal law $\mathsf{IMGW}_0 \otimes \mathsf{RW}_\rho$ such that excursions of X into undiscovered territory are coupled with excursions of Y into undiscovered territory away from ξ. From this we obtain an annealed CLT for X by controlling the amount of time spent outside the coupled excursions as well as the drift of Y along ξ. Since (X, Y) depend on the realization of $(\mathcal{T}, (\widehat{\mathcal{T}}, \xi))$, we do not see how to use the shifted coupling *directly* to prove a quenched MGW CLT. Instead we adapt the method of [5] to deduce it from the annealed CLT by controlling the correlation between two realizations of RW_ρ on a single MGW tree $\mathcal{T}$ (as was done in [27, §7] in the case $\lambda < \rho$).
- In §5 we prove Thm. 1.5. The main result needed is a large deviations estimate (Lem. 5.2) on the conductances at the n-th level of the tree, which extends the estimates of [21, p. 129] and [10, p. 7] to our setting of Markovian dependency. The result then follows by adapting the proof of [21, Thm. 1] (see also [10, Propn. 1.1] for the single-type case).
- In §6 are collected some basic properties of MGW which are needed in the course of our proof and which may be of independent interest:
  - In §6.1 we prove a zero-one law which is used for the proof of ergodicity in Thm. 1.3.
  - In §6.2 we adapt the methods of [4] to prove that moments for the offspring distributions translate directly to moments for the normalized population size defined in §2.3 (see Propn. 2.2).
  - In §6.3 we prove the existence of harmonic moments for the normalized population size (extending part of [25, Thm. 1.1] to the multi-type setting) and use this result to prove conductance estimates (Propn. 4.4) used in the proof of Thm. 1.1.

Open problems. We conclude this section by mentioning some open problems in this area. These problems are open even for single-type Galton-Watson trees.

(1) Does a CLT with diffusive scaling hold for RW_ρ in the entire regime $p \geq 2$?
(2) Does a CLT with diffusive scaling hold for $\mathsf{RWRE}_{\bar{\rho}}$ in the entire regime $\kappa \geq 2$?
(3) What happens on the critical tree ($\rho = 1$) conditioned to non-extinction?
(4) Does a CLT with any scaling (or other limit laws) hold for RW_ρ when $p < 2$?

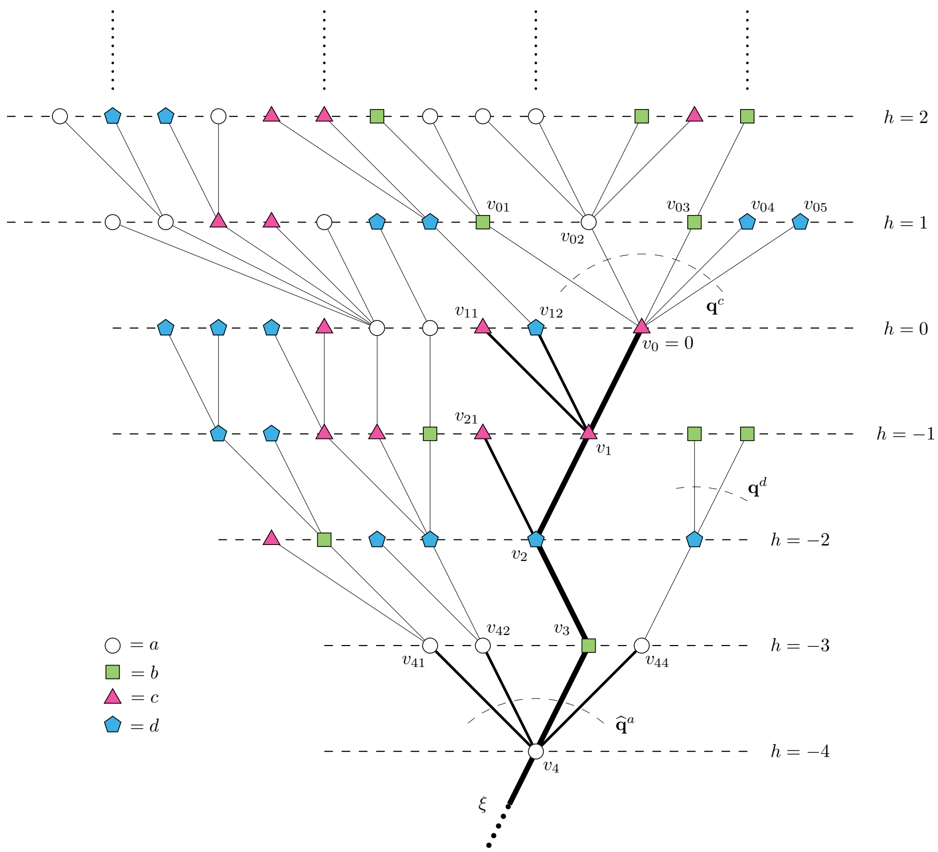
A common feature of these problems is that while the reversing measure for the process from the perspective of the particle is given by Thm. 1.3, the method of martingale approximation used in [27, 10] and in this paper seem not to be directly applicable.

Acknowledgements. We are very grateful to Ofer Zeitouni for suggesting the proof of ergodicity in Thm. 1.3 and the method of going from the annealed to quenched CLT in the proof of Thm. 1.1. A.D. thanks Alexander Fribergh for discussions on the results of [2] which motivated us to extend our results to trees with leaves. N.S. thanks Yuval Peres for several helpful conversations about the papers [27, 17] which led to the construction of the reversing measure. We thank the anonymous referee for valuable comments on an earlier draft which helped to improve the presentation of the paper.

2. Reversing probability measures for $\mathcal{T}\mathsf{RW}_\rho$ and $\mathcal{T}\mathsf{RWRE}_{\bar{\rho}}$

Consider a multi-type Galton-Watson measure MGW with offspring distributions $(\mathbf{q}^a)_{a\in\mathcal{Q}}$ and mean matrix A. Hereafter we let $\underline{e}$ and $\underline{g}$ denote the right and left eigenvectors respectively associated to the Perron-Frobenius eigenvalue of A, normalized so that $\sum_a g_a = \sum_a e_a = 1$. Since our results are stated for MGW-a.e. tree, with no loss of generality we set $\mathbf{g}(a) \equiv \mathsf{MGW}(\chi_o = a) = g_a$. Unless otherwise specified, X and Y denote RW_ρ on trees without and with marked ray respectively.

FIGURE 1. IMGW_0 tree



2.1. **Construction of IMGWR.** Assuming only (H1), in this section we use ideas from [17] to construct the reversing measure IMGWR on the space $\Omega_\downarrow$. We begin by constructing two auxiliary measures. Let the infinite ray ξ (without types) be given. For some $n > 0$, we let vertex v_n be given a type χ_n according to a distribution $\boldsymbol{\pi}$, to be determined shortly. It is then given offspring $\underline{x}^{v_n}$ according to the *inflated* offspring distribution $\widehat{\mathbf{q}}^{\chi_n}$, where

$$\widehat{\mathbf{q}}^a(\underline{x}) := \mathbf{q}^a(\underline{x})\frac{\langle \underline{x}, \underline{e}\rangle}{\rho e_a} \quad \forall a \in \mathcal{Q};$$

note that $\widehat{\mathbf{q}}^a(|\underline{x}| \geq 1) = 1$. One offspring w of v_n is then identified with the next vertex v_{n-1} along ξ, where each w is chosen with probability $e_{\chi_w}/\langle \underline{x}^{v_n}, \underline{e}\rangle$. We proceed in this manner along the ray ending with the identification of $v_0 = o$. The sequence of types $\chi_n, \chi_{n-1}, \ldots$ seen along the ray is then (by (H1)) an irreducible Markov chain with transition probabilities

$$K(a,b) = \sum_{\underline{x}} \widehat{\mathbf{q}}^a(\underline{x})\frac{e_b x_b}{\langle \underline{x}, \underline{e}\rangle} = \sum_{\underline{x}} \mathbf{q}^a(\underline{x})\frac{e_b x_b}{\rho e_a} = \frac{e_b}{\rho e_a}A(a,b). \tag{2.1}$$

This chain has stationary distribution $\boldsymbol{\pi}(a) = e_a g_a/\langle \underline{e}, \underline{g}\rangle$, so starting with $\chi_n \sim \boldsymbol{\pi}$ yields a consistent family of distributions for $(v_n, \ldots, v_1)$ and their (immediate) offspring, with types. By Kolmogorov's existence theorem, this uniquely specifies the distribution of the *backbone* of the tree, that is, of the ray ξ together with all (immediate) offspring of the vertices v_i, $i > 0$.

To each of these offspring (off the ray) and to o, we attach an independently chosen MGW tree conditioned on the given type, and denote by IMGW_0 the resulting measure on the pair $(\mathcal{T},\xi)$. The *inflated multi-type Galton-Watson* measure IMGW is obtained from IMGW_0 by an additional biasing according to the root type χ_o. Specifically, we set

$$\frac{d\mathsf{IMGW}}{d\mathsf{IMGW}_0} = \frac{1/e_{\chi_o}}{\mathbb{E}_{\boldsymbol{\pi}}[1/e_\chi]} = \frac{\mathbb{E}_{\mathbf{g}}[e_\chi]}{e_{\chi_o}},$$

where χ denotes a random variable on $\mathcal{Q}$ with the specified distribution. We note that under IMGW, $\chi_o \sim \mathbf{g}$ and so $\mathcal{T}^{(o)}$ has marginal law MGW, which implies

$$\mathbb{E}_{\mathsf{IMGW}}[d_o] = \mathbb{E}_{\mathsf{MGW}}[d_o] = \sum_a g_a \sum_b A(a,b) = \rho.$$

With this in mind, we define the probability measure IMGWR such that

$$\frac{d\mathsf{IMGWR}}{d\mathsf{IMGW}} = \frac{d_o+\rho}{\mathbb{E}_{\mathsf{IMGW}}[d_o+\rho]} = \frac{d_o+\rho}{2\rho}, \tag{2.2}$$

and proceed to show that it is a reversing measure for $\mathcal{T}\mathsf{RW}_\rho$. From now on we adopt the notation that if μ is a law on trees $\mathcal{T}$ (with or without marked ray) and $a \in \mathcal{Q}$, μ^a refers to the law conditioned on $\chi_o = a$.

Proof of Thm. 1.3. We first prove (a). If MGW produces a.s. a d-regular tree then the result holds trivially, so we omit this case in what follows. For notational convenience we view IMGW_0, etc. as measures on planar trees with marked ray (by letting the offspring of each vertex be ordered in an exchangeable manner). Given $(\mathcal{T},\xi)$ we denote by $(v_{i1},\ldots,v_{id_{v_i}})$ the offspring of $v_i\in\xi$, and we use the shorthand i for v_i and ij for v_{ij}.

Let $\mathcal{S}$ denote the map $(\mathcal{T},\xi)\mapsto(\mathcal{T}^1,\xi^1)$. We will show that for $A,B\in\mathcal{B}_{\Omega_\downarrow}$,

$$\int_A \mathbf{p}((\mathcal{T},\xi),B)\ d\mathsf{IMGWR}(\mathcal{T},\xi) = \int_B \mathbf{p}((\mathcal{T}',\xi'),A)\ d\mathsf{IMGWR}(\mathcal{T}',\xi') \tag{2.3}$$

where $\mathbf{p}((\mathcal{T},\xi),B)$ denotes the transition kernel of the process $\mathcal{T}\mathsf{RW}_\rho$. For $(\mathcal{T},\xi)\sim\mathsf{IMGW}_0$, let $\underleftarrow{\mathsf{IMGW}}{}_0^a$ denote the law of the subtree with marked ray $(\mathcal{T}\backslash\mathcal{T}^{(i-1)},\xi^i)$ conditioned on $\chi_{i-1}=a$, for any $i\ge 1$. Then

$$d\mathsf{IMGW}_0(\mathcal{T},\xi) = \boldsymbol{\pi}(\chi_1)\widehat{\mathbf{q}}^{\chi_1}(\underline{x}^1)\frac{e_{\chi_0}}{\langle \underline{x}^1,\underline{e}\rangle} d\underleftarrow{\mathsf{IMGW}}{}_0^{\chi_1}(\mathcal{T}\backslash\mathcal{T}^{(1)},\xi^2)\prod_{j=1}^{d_1} d\mathsf{MGW}^{\chi_{1j}}(\mathcal{T}^{(1j)}).$$

Let $\mathcal{P}$ denote the collection of $\mathcal{B}_{\Omega_\downarrow}$-measurable sets on which $\mathcal{S}$ is injective, and suppose $B\in\mathcal{P}$. If μ is a measure on $\Omega_\downarrow$, $\mathcal{S}_B^*\mu := (\mu\circ\mathcal{S})\mathbf{1}_B$ is a well-defined measure on $\Omega_\downarrow$. Then

$$d\mathcal{S}_B^*\mathsf{IMGW}_0(\mathcal{T},\xi) = \mathbf{1}_{\{(\mathcal{T},\xi)\in B\}}\boldsymbol{\pi}(\chi_1)\mathbf{q}^{\chi_1}(\underline{x}^1) d\underleftarrow{\mathsf{IMGW}}{}_0^{\chi_1}(\mathcal{T}\backslash\mathcal{T}^{(1)},\xi^2)\prod_{j=1}^{d_1} d\mathsf{MGW}^{\chi_{1j}}(\mathcal{T}^{(1j)}).$$

so

$$\frac{d\mathcal{S}_B^*\mathsf{IMGW}_0}{d\mathsf{IMGW}_0} = \mathbf{1}_B\frac{\rho e_{\chi_1}}{e_{\chi_0}}. \tag{2.4}$$

We then verify that

$$\frac{d\mathcal{S}_B^*\mathsf{IMGW}}{d\mathsf{IMGW}} = \mathbf{1}_B\frac{\left(\frac{d\mathcal{S}_B^*\mathsf{IMGW}}{d\mathcal{S}_B^*\mathsf{IMGW}_0}\right)}{\left(\frac{d\mathsf{IMGW}}{d\mathsf{IMGW}_0}\right)}\frac{d\mathcal{S}_B^*\mathsf{IMGW}_0}{d\mathsf{IMGW}_0} = \mathbf{1}_B\frac{\left(\frac{d\mathsf{IMGW}}{d\mathsf{IMGW}_0}\circ\mathcal{S}\right)}{\left(\frac{d\mathsf{IMGW}}{d\mathsf{IMGW}_0}\right)}\frac{d\mathcal{S}_B^*\mathsf{IMGW}_0}{d\mathsf{IMGW}_0} = \mathbf{1}_B\frac{1/e_{\chi_1}}{1/e_{\chi_o}}\frac{\rho e_{\chi_1}}{e_{\chi_o}} = \mathbf{1}_B\rho,$$

and similarly

$$\frac{d\mathcal{S}_B^*\mathsf{IMGWR}}{d\mathsf{IMGWR}} = \mathbf{1}_B \frac{\rho(d_1+\rho)}{d_o+\rho}.$$

The left-hand side of (2.3) can be written as

$$\int_{A\cap\mathcal{S}^{-1}B} \frac{\rho}{d_o+\rho}\, d\mathsf{IMGWR}(\mathcal{T},\xi) + \int_A \frac{1}{d_o+\rho}\sum_{i=1}^{d_o} \mathbf{1}_{\{(\mathcal{T}^{0i},\xi^{0i})\in B\}}\, d\mathsf{IMGWR}(\mathcal{T},\xi).$$

Using the injectivity of $\mathcal{S}$ on B, the second integral can be written as

$$\begin{aligned}\int_{A\cap\mathcal{S}B} \frac{1}{d_o+\rho}\, d\mathsf{IMGWR}(\mathcal{T},\xi) &= \int_{\mathcal{S}^{-1}A\cap B} \frac{1}{d_1+\rho}\, d\mathcal{S}_B^*\mathsf{IMGWR}(\mathcal{T},\xi)\\ &= \int_{\mathcal{S}^{-1}A\cap B} \frac{\rho}{d_o+\rho}\, d\mathsf{IMGWR}(\mathcal{T},\xi),\end{aligned}$$

from which it is clear that the two sides of (2.3) must agree.

Next we claim that $\mathcal{P}$ generates $\mathcal{B}_{\Omega_\downarrow}$ up to null sets. Since we are in the non-degenerate setting, IMGW_0-a.e. pair $(\mathcal{T},\xi)$ is such that the infinite typed subtrees $\mathcal{T}^{(1i)}$ and $\mathcal{T}^{(1j)}$ are non-isomorphic for $i\neq j$, so $\mathcal{B}_{\Omega_\downarrow}$ is generated up to null sets by cylinder sets determined by finite trees $\mathcal{T}_n$ with no two subtrees $\mathcal{T}_n^{(1j)}$ isomorphic. Given such a set F, decompose F into the disjoint union of the sets $F_j = \{(\mathcal{T},\xi)\in F : o = 1j\}$ (i.e., o is the j-th child of 1). Then $F_j\in\mathcal{P}$, and the claim follows.

To conclude, for fixed A let $\widetilde{\mathcal{B}}_{\Omega_\downarrow}$ denote the collection of sets $B\in\mathcal{B}_{\Omega_\downarrow}$ for which (2.3) holds. From the above $\widetilde{\mathcal{B}}_{\Omega_\downarrow}$ contains the $\boldsymbol{\pi}$-system $\mathcal{P}$. Further $\widetilde{\mathcal{B}}_{\Omega_\downarrow}$ is closed under monotone limits and countable disjoint unions, and in particular it contains $\Omega_\downarrow$ since $\Omega_\downarrow$ can be decomposed as a countable disjoint union of sets in $\mathcal{P}$ by a similar argument as above. Thus by the $\boldsymbol{\pi}$-λ theorem (2.3) holds for all $B\in\sigma(\mathcal{P})$, and extends to all $B\in\mathcal{B}_{\Omega_\downarrow}$ again using the claim above.

We prove (b) by adapting the method of [30, Cor. 2.1.25]. Letting ν denote the law of the sequence $(\mathcal{T}^{(Y_0)},\mathcal{T}^{(Y_1)},\dots)$ on the space Ω^∞ of rooted trees, and $\mathfrak{S}$ the shift $(\mathcal{T}_0,\mathcal{T}_1,\dots)\mapsto(\mathcal{T}_1,\mathcal{T}_2,\dots)$ on Ω^∞, the content of (b) is that the measure-preserving system $(\Omega^\infty,\mathcal{F}^\infty,\nu,\mathfrak{S})$ is ergodic.

To this end, let $\mathcal{I}$ denote the $\mathfrak{S}$-invariant σ-field and fix $B\in\mathcal{I}$, i.e. $\nu(\mathfrak{S}^{-1}B)=\nu(B)$. We define $\phi:\Omega\to[0,1]$ by

$$\phi(\mathcal{T}) = \nu(\,(\mathcal{T}_t)_{t\geq 0}\in B|\,\mathcal{T}_0=\mathcal{T}).$$

Let $\mathcal{H}_t=\sigma(\mathcal{T}_1,\dots,\mathcal{T}_t)$: by the shift-invariance of B,

$$\nu[(\mathcal{T}_s)_{s\geq 0}\in B|\mathcal{H}_t] = \nu[(\mathcal{T}_s)_{s\geq t}\in B|\mathcal{H}_t] = \phi(\mathcal{T}_t),$$

so $(\phi(\mathcal{T}_t))_{t\geq 0}$ is a martingale with respect to the filtration $(\mathcal{H}_t)_{t\geq 0}$. By Lévy's upward theorem, ν-a.s. $\phi(\mathcal{T}_t)\to\mathbf{1}_B$ as $t\to\infty$.

It follows that for $0<a\leq b<1$,

$$\frac{1}{t}\sum_{i=0}^{t-1}\mathbf{1}_{\{\phi(\mathcal{T}_i)\in[a,b]\}}$$

converges ν-a.s. to 0. On the other hand, by the Birkhoff ergodic theorem, it converges ν-a.s. to $\nu(\phi(\mathcal{T})\in[a,b]|\mathcal{I})$. Taking expectations on both sides we find $\phi(\mathcal{T})\in\{0,1\}$ ν-a.s., that is, $\phi=\mathbf{1}_{C_0}$ for some $C_0\subseteq\Omega$. Further, by the martingale property

$$\phi(\mathcal{T}) = \mathbb{E}_\nu[\phi(\mathcal{T}_1)|\mathcal{T}_0=\mathcal{T}] = \frac{\rho}{\rho+d_o}\mathbb{E}[\phi(\mathcal{T}^{(1)})|\mathcal{T}] + \frac{1}{\rho+d_o}\sum_{j=1}^{d_o}\phi(\mathcal{T}^{(0j)}),$$

which implies $\mathcal{T} \in C_0$ if and only if $\mathcal{T}^{(v)} \in C_0$ for all $v \in \partial o$. As we will show in Propn. 6.2, this implies $\mathsf{MGW}(C_0) \in \{0,1\}$. Then, if ν_0 denotes the marginal of ν on $\mathcal{T}_0$, we have $\nu_0 \ll \mathsf{MGW}$ (since $\mathsf{IMGWR} \ll \mathsf{IMGW}$ and $\mathcal{T}^{(o)} \sim \mathsf{MGW}$ under IMGW) and so

$$\nu(B) = \mathbb{E}_\nu[\phi] = \nu_0(C_0) \in \{0,1\}\,,$$

which completes the proof of ergodicity. □

2.2. **Biased random walk with random environment.** We now extend the methods of the previous section to prove Thm. 1.6. Let $\underline{\bar{e}}, \underline{\bar{g}}$ denote the right and left Perron-Frobenius eigenvectors of $\bar{A} \equiv \bar{A}^{(1)}$, normalized to have sum 1; as before we set $\mathbf{g}(a) \equiv \overline{\mathsf{MGW}}(\chi_o = a)$ to be $\bar{g}_a$.

We proceed much as in the deterministic environment setting, although the notation becomes more complicated. For $\underline{y} \in \mathcal{Q}^\ell$ write $\underline{\bar{e}}(\underline{y}) = (\bar{e}_{y_j})_{j=1}^{\ell}$. For $a \in \mathcal{Q}$

$$\mathbb{E}_{\mathbf{q}^a}[\langle \underline{\bar{e}}(\underline{y}), \underline{\alpha}\rangle] = \sum_b \bar{A}(a,b)\bar{e}_b = \bar{\rho}\bar{e}_a,$$

so we define the inflated offspring measure $\widehat{\mathbf{q}}^a$ by

$$\frac{d\widehat{\mathbf{q}}^a}{d\mathbf{q}^a} = \frac{\langle \underline{\bar{e}}(\underline{y}), \underline{\alpha}\rangle}{\bar{\rho}\bar{e}_a}.$$

We then construct the measure on $\Omega_\downarrow$: let the infinite ray ξ (without types or weights) be given, and for some $n > 0$ let v_n have type χ_n. It is given offspring $(\underline{y}^{v_n}, \underline{\alpha}^{v_n}) \sim \widehat{\mathbf{q}}^{\chi_n}$. One offspring w of v_n is identified with the next vertex v_{n-1} along ξ, where each w is chosen with probability

$$\frac{\bar{e}_{\chi_w}\alpha_w}{\langle \underline{\bar{e}}(\underline{y}), \underline{\alpha}^{v_n}\rangle}.$$

Continuing the procedure along the ray up to $v_0 = o$, the sequence of types $\chi_n, \chi_{n-1}, \ldots$ seen along ξ is an irreducible Markov chain with transition probabilities

$$\bar{K}(a,b) = \mathbb{E}_{\widehat{\mathbf{q}}^a}\left[\frac{\bar{e}_b \sum_j \alpha_j \mathbf{1}_{\{y_j = b\}}}{\langle \underline{\bar{e}}(\underline{y}), \underline{\alpha}\rangle}\right] = \frac{\bar{e}_b}{\bar{\rho}\bar{e}_a}\bar{A}(a,b)$$

and stationary distribution $\bar{\boldsymbol{\pi}}(a) = \bar{e}_a\bar{g}_a/\langle \underline{\bar{e}}, \underline{\bar{g}}\rangle$. Thus, starting with $\chi_n \sim \bar{\boldsymbol{\pi}}$ and applying Kolmogorov's existence theorem, we obtain a measure $\overline{\mathsf{IMGW}}_0$ on $\Omega_\downarrow$ which is a generalization of IMGW_0.

Proof of Thm. 1.6. We argue as in the proof of Thm. 1.3 (a), the key difference being that the trees now have edge weights as well as types. If $\overline{\mathsf{MGW}}$ produces a.s. a d-regular tree with constant weights then the result holds trivially, so we omit this case in what follows. For $(\mathcal{T},\xi) \sim \overline{\mathsf{IMGW}}_0$, let $\underleftarrow{\overline{\mathsf{IMGW}}}{}_0^a$ denote the law of the subtree with marked ray $(\mathcal{T}\backslash\mathcal{T}^{(i-1)}, \xi^i)$ conditioned on $\chi_{i-1} = a$, for any $i \geq 1$. Let $\mathcal{S} : (\mathcal{T},\xi) \mapsto (\mathcal{T}^1, \xi^1)$; we emphasize that $\mathcal{S}$ is a mapping on trees with types and edge weights. Let $\mathcal{P}$ denote the collection of $\mathcal{B}_{\Omega_\downarrow}$-measurable sets on which $\mathcal{S}$ is injective. For $A \in \mathcal{B}_{\Omega_\downarrow}$ and $B \in \mathcal{P}$, we compute

$$d\overline{\mathsf{IMGW}}_0(\mathcal{T},\xi) = \bar{\boldsymbol{\pi}}(\chi_1)\widehat{\mathbf{q}}^{\chi_1}(\underline{y}^1, \underline{\alpha}^1)\frac{\bar{e}_{\chi_0}\alpha_0}{\langle \underline{\bar{e}}(\underline{y}^1), \underline{\alpha}^1\rangle}d\underleftarrow{\overline{\mathsf{IMGW}}}{}_0^{\chi_1}(\mathcal{T}\backslash\mathcal{T}^{(1)}, \xi^2)\prod_{j=1}^{d_1} d\overline{\mathsf{MGW}}^{\chi_{1j}}(\mathcal{T}^{(1j)}),$$

$$d\mathcal{S}_B^*\overline{\mathsf{IMGW}}_0(\mathcal{T},\xi) = \mathbf{1}_{\{(\mathcal{T},\xi)\in B\}}\bar{\boldsymbol{\pi}}(\chi_1)\mathbf{q}^{\chi_1}(\underline{y}^1, \underline{\alpha}^1)d\underleftarrow{\overline{\mathsf{IMGW}}}{}_0^{\chi_1}(\mathcal{T}\backslash\mathcal{T}^{(1)}, \xi^2)\prod_{j=1}^{d_1} d\overline{\mathsf{MGW}}^{\chi_{1j}}(\mathcal{T}^{(1j)}),$$

so $\bar{e}_{\chi_0}\alpha_0 \; d\mathcal{S}_B^*\overline{\mathsf{IMGW}}_0 = \mathbf{1}_B\bar{\rho}\bar{e}_{\chi_1} \; d\overline{\mathsf{IMGW}}_0$. Letting

$$\frac{d\overline{\mathsf{IMGW}}}{d\overline{\mathsf{IMGW}}_0} = \frac{1/\bar{e}_{\chi_o}}{\mathbb{E}_{\bar{\boldsymbol{\pi}}}[1/\bar{e}_\chi]} = \frac{\mathbb{E}_{\bar{\mathbf{g}}}[\chi]}{\bar{e}_{\chi_0}}, \quad \frac{d\overline{\mathsf{IMGWR}}}{d\overline{\mathsf{IMGW}}} = \frac{\bar{\rho}+\sum_{j=1}^{d_o}\alpha_{0j}}{\mathbb{E}_{\overline{\mathsf{IMGW}}}[\bar{\rho}+\sum_{j=1}^{d_o}\alpha_{0j}]} = \frac{\bar{\rho}+\sum_{j=1}^{d_o}\alpha_{0j}}{2\bar{\rho}},$$

we obtain

$$\alpha_0 \; d\mathcal{S}_B^*\overline{\mathsf{IMGW}} = \mathbf{1}_B\bar{\rho} \; d\overline{\mathsf{IMGW}}, \quad \frac{\alpha_0}{\bar{\rho}+\sum_{j=1}^{d_1}\alpha_{1j}} \; d\mathcal{S}_B^*\overline{\mathsf{IMGWR}} = \mathbf{1}_B \frac{\bar{\rho}}{\bar{\rho}+\sum_{j=1}^{d_0}\alpha_{0j}} \; d\overline{\mathsf{IMGWR}}.$$

The analogue of (2.3) thus holds for all $B \in \mathcal{P}$, and we extend to all $B \in \mathcal{B}_{\Omega_\downarrow}$ by essentially the same argument used in the proof of Thm. 1.3 (a): by the non-degeneracy assumption, $\overline{\mathsf{IMGW}}_0$-a.e. pair $(\mathcal{T},\xi)$ is such that the infinite subtrees $\mathcal{T}^{(1i)}$ and $\mathcal{T}^{(1j)}$ for $i \neq j$ are non-isomorphic as weighted typed trees. □

2.3. **IMGW as a weak limit.** We next provide an alternative characterization of the inflated Galton-Watson measure IMGW in a multi-type analogue of [27, Lem. 1].

To this end, we will define the notion of "normalized population size" for rooted trees $\mathcal{T}$ with type. Let $\mathcal{T}_n$ denote the subtree induced by $\{v \in \mathcal{T} : |v| \le n\}$, and D_n the set $\{v \in \mathcal{T} : |v| = n\}$. Let $(\mathcal{F}_n)_{n\ge 0}$ denote the natural filtration of the tree, i.e., $\mathcal{F}_n$ is the σ-algebra generated by $\mathcal{T}_n$ (a finite tree with vertex types). Let $\underline{Z}_n = (Z_n(b))_{b\in\mathcal{Q}} \in (\mathbb{Z}_{\ge 0})^{\mathcal{Q}}$ count the number of vertices of each type at level n, so $\underline{Z}_n$ is $\mathcal{F}_n$-measurable. Then

$$\widehat{Z}_n := \frac{\langle \underline{Z}_n, \underline{e}\rangle}{\rho^n} = \frac{1}{\rho^n}\sum_{v\in D_n} e_{\chi_v}$$

is a non-negative $(\mathcal{F}_n)$-martingale under MGW^a for every a, with $\mathbb{E}_{\mathsf{MGW}^a}[\widehat{Z}_0] = e_a$ (see e.g. [11, p. 49]). By the normalized population size of the tree we mean the a.s. limit of $\widehat{Z}_n$, denoted W_o. For $v \in \mathcal{T}$ we use W_v to denote the normalized population size of $\mathcal{T}^{(v)}$. Under (H1) and (H2), it follows from the multi-type Kesten-Stigum theorem (see [16], or the conceptual proof of [17]) that $W_o > 0$ a.s. on the event of non-extinction, and $\mathbb{E}_{\mathsf{MGW}^a}[W_o] = e_a$.

For $a \in \mathcal{Q}$ let Q_n^a be a probability measure on (infinite) rooted trees defined by

$$\frac{d\mathsf{Q}_n^a}{d\mathsf{MGW}^a} = \frac{\widehat{Z}_n}{e_a}.$$

For $\mathcal{T} \sim \mathsf{Q}_n^a$ choose $v_n \in D_n$ at random with probabilities proportional to weights $e_{\chi_{v_n}}$, and let $\mathsf{Q}_{n\star}^a$ denote the law of the resulting pair $(\mathcal{T}, v_n)$. Let $\mathsf{Q}_{n\star} := \sum_{a\in\mathcal{Q}} \boldsymbol{\pi}_a \mathsf{Q}_{n\star}^a$ and $\mathsf{Q}_n := \sum_{a\in\mathcal{Q}} \boldsymbol{\pi}_a \mathsf{Q}_n^a$, so that $d\mathsf{Q}_n/d\mathsf{MGW} = \widehat{Z}_n/\mathbb{E}_{\mathbf{g}}[e_\chi]$. Finally let $\mathsf{IMGW}_0(n)$ denote the law of $(\mathcal{T},\xi_0)^{v_n}$ (see (1.2) for this notation), where $(\mathcal{T}, v_n) \sim \mathsf{Q}_{n\star}$ and ξ_0 is any infinite ray emanating from o not sharing an edge with the geodesic from o to v_n.

Proposition 2.1. *Under* (H1), $\mathsf{IMGW}_0(n)$ *converges weakly to* IMGW_0.

The proposition can be seen from the following explicit construction of $\mathsf{Q}_{n\star}^a$: begin with $v_0 \equiv o$ of type a, and suppose inductively that we have constructed $(\mathcal{T}_i, v_i)$ $(i < n)$ where $\mathcal{T}_i$ is the tree up to level i and v_i is the i-th vertex on the geodesic from o to v_n. Then v_i is given offspring $\underline{x}^{v_i}$ according to $\widehat{\mathbf{q}}^{\chi_{v_i}}$, and one of these offspring w is randomly chosen (according to weights e_w) to be distinguished as v_{i+1}. Meanwhile all other vertices $v \in D_i\backslash\{v_i\}$ are given offspring $\underline{x}^v$ according to

$\mathbf{q}^{\chi_v}$. Once $(\mathcal{T}_n, v_n)$ has been constructed, attach to each $v \in D_n$ an independent MGW^{χ_v} tree. For $N \ge n$,

$$\frac{\mathsf{Q}^a_{n\star}(\mathcal{T}_N, v_n)}{\mathsf{MGW}^a(\mathcal{T}_N)} = \prod_{i=0}^{n-1} \frac{\langle \underline{x}^{v_i}, \underline{e}\rangle}{\rho e_{\chi_{v_i}}} \frac{e_{\chi_{v_{i+1}}}}{\langle \underline{x}^{v_i}, \underline{e}\rangle} = \frac{e_{\chi_{v_n}}}{\rho^n e_a},$$

from which it follows that

$$\frac{d\mathsf{Q}^a_n}{d\mathsf{MGW}^a} = \frac{\widehat{Z}_n}{e_a}, \quad \frac{d\mathsf{Q}_n}{d\mathsf{MGW}} = \frac{\widehat{Z}_n}{\mathbb{E}_{\mathbf{g}}[e_\chi]}.$$

Letting $n \to \infty$ in $\mathsf{Q}^a_{n\star}, \mathsf{Q}_{n\star}$ we obtain the measures $\widehat{\mathsf{MGW}}^a_\star, \widehat{\mathsf{MGW}}_\star$ of [17] on rooted trees with infinite marked ray, as well as the corresponding marginals $\widehat{\mathsf{MGW}}^a, \widehat{\mathsf{MGW}}$ on trees without marked ray which satisfy

$$\frac{d\widehat{\mathsf{MGW}}^a_n}{d\mathsf{MGW}^a}\bigg|_{\mathcal{F}_n} = \frac{\widehat{Z}_n}{e_a}, \quad \frac{d\widehat{\mathsf{MGW}}^a_n}{d\mathsf{MGW}^a} = \frac{\widehat{Z}_n}{\mathbb{E}_{\mathbf{g}}[e_\chi]}.$$

By the Kesten-Stigum theorem and Scheffé's lemma (see e.g. [29, §5.10]), $\widehat{Z}_n \xrightarrow{L^1} W_o$, hence

$$\frac{d\widehat{\mathsf{MGW}}^a}{d\mathsf{MGW}^a} = \frac{W_o}{e_a}, \quad \frac{d\widehat{\mathsf{MGW}}}{d\mathsf{MGW}} = \frac{W_o}{\mathbb{E}_{\mathsf{MGW}}[W_o]} = \frac{W_o}{\mathbb{E}_{\mathbf{g}}[e_\chi]}.$$

We remark that although $\widehat{\mathsf{MGW}}_\star$ and IMGWR are both measures on trees with rays, they are not in general equivalent unless K is reversible.

Proof of Propn. 2.1. Take $(\mathcal{T}'_n, v'_n) \sim \widehat{\mathsf{MGW}}_\star$, delete everything below the first n levels, and attach to each $v \in D_n$ a standard MGW^{χ_v} tree. If we denote the resulting rooted tree by $\mathcal{T}'$, it follows from the above discussion that $(\mathcal{T}', v'_n)$ has law $\mathsf{Q}_{n\star}$. From this description it is clear that if $(\mathcal{T}, \xi) \sim \mathsf{IMGW}_0$ then $(\mathcal{T}^{(v_n)}, o) \sim \mathsf{Q}_{n\star}$ (noting that the root type has law $\boldsymbol{\pi}$ under $\mathsf{Q}_{n\star}$, which agrees with the law of $e_{\chi_{v_n}}$ under IMGW_0). In other words the portion of $(\mathcal{T}, \xi)$ descended from v_n has the same distribution under $\mathsf{IMGW}_0(n)$ as under IMGW_0, proving the result. □

The following result will be shown in §6.2:

Proposition 2.2. *If* (H1), (H2), *and* (H3^p) *hold with* $p > 1$, *then* $\mathbb{E}_{\mathsf{MGW}}[W_o^p] < \infty$.

We define for future reference

$$\eta := \mathbb{E}_{\widehat{\mathsf{MGW}}}[W_o] = \frac{\mathbb{E}_{\mathsf{MGW}}[W_o^2]}{\mathbb{E}_{\mathbf{g}}[e_\chi]}, \quad \sigma^2 := \frac{\mathbb{E}_{\mathbf{g}}[e_\chi]^2}{\mathbb{E}_{\mathsf{MGW}}[W_o^2]}; \tag{2.5}$$

given the proposition these quantities are finite under (H3^2).

2.4. **Harmonic coordinates for RW_ρ and CLT.** From now on, if μ is a probability measure on trees (with or without marked ray), we use μ as shorthand also for $\mu \otimes \mathsf{RW}_\rho$. We write $\mathbb{P}_\mathcal{T}$ for the law of the quenched random walk $\mathsf{RW}_\rho(\mathcal{T})$ and $\mathbb{E}_\mathcal{T}$ for expectation with respect to $\mathbb{P}_\mathcal{T}$, and let $(\mathcal{G}_t^\mathcal{T})_{t\ge 0}$ denote the corresponding filtration of the walk. Given $\mathcal{T}$, for a vertex $v \in \mathcal{T}$ we let ∂v denote the neighbors of v, and $\partial^+ v$ the offspring of v, i.e., $\partial^+ v = \partial v \cap \mathcal{T}^{(v)}$. We write $v \le w$ if $w \in \mathcal{T}^{(v)}$, with $v < w$ if $w \ne v$.

For $v \in \mathcal{T}$ recall that W_v denotes the normalized population size of the subtree $\mathcal{T}^{(v)}$.[1] For vertices $v \in \mathcal{T}$ we define S_v as in [27, §3]: if $\mathcal{T}$ is a rooted tree, let

$$S_v := \sum_{o<u\le v} W_u. \tag{2.6}$$

If $\mathcal{T}$ has marked ray ξ, recalling (1.1) we set

$$S_v := S_{R_v} + S_v^\xi \quad \text{where} \quad S_{R_v} := -\sum_{u\in\xi, o\ge u>R_v} W_u, \quad S_v^\xi := \sum_{R_v<u\le v} W_u. \tag{2.7}$$

While on MGW-a.e. $\mathcal{T}$ the map $v \mapsto S_v$ is harmonic except at o with respect to the transition probabilities of $\mathsf{RW}_\rho(\mathcal{T})$, on IMGW-a.e. $(\mathcal{T},\xi)$ the map $v \mapsto S_v$ is harmonic at every vertex with respect to the transition probabilities of $\mathsf{RW}_\rho(\mathcal{T},\xi)$. Thus, if $(Y_t)_{t\ge0} \sim \mathsf{RW}_\rho(\mathcal{T},\xi)$, $M_t := S_{Y_t}$ will be a martingale given a fixed realization of the tree; we regard it as providing "harmonic coordinates" for the random walk. Using the reversing measure IMGWR it is easy to prove a quenched CLT for M (extending [27, Cor. 1]):

Proposition 2.3. *Under* (H1), (H2), *and* (H3^2), *on* IMGW*-a.e.* $(\mathcal{T},\xi)$ *the process* $M_{\lfloor nt\rfloor}/(\eta\sigma\sqrt{n})$ *converges in distribution to a standard Brownian motion as* $n\to\infty$.

Proof. We check the conditions of the Lindeberg-Feller martingale CLT (see e.g. [9, Thm. 7.7.4]): letting

$$V_n = \frac{1}{n}\sum_{t=0}^{n-1} \mathbb{E}_{\mathcal{T}}[(M_{t+1}-M_t)^2|\mathcal{G}_t^{\mathcal{T}}],$$

we verify that for IMGW-a.e. $(\mathcal{T},\xi)$,

(i) $V_n \to \eta^2\sigma^2$ in probability and
(ii) for all $\epsilon>0$, $\frac{1}{n}\sum_{t=0}^{n-1}\mathbb{E}_{\mathcal{T}}[(M_{t+1}-M_t)^2\mathbf{1}_{\{|M_{t+1}-M_t|>\epsilon\sqrt{n}\}}] \to 0$.

Let Y_n denote the random walk on $(\mathcal{T},\xi)$: we rewrite V_n in terms of the induced random walk on $\Omega_\downarrow$ as

$$V_n = \frac{1}{n}\sum_{t=1}^{n-1}\varphi[(\mathcal{T},\xi)^{Y_n}], \quad \varphi[(\mathcal{T},\xi)] := \frac{\rho}{\rho+d_o}W_o^2 + \frac{1}{\rho+d_o}\sum_{j=1}^{d_o} W_{0j}^2.$$

The key observation is that φ is a function only of the subtree $\mathcal{T}^{(o)}$ descended from o. We saw in Thm. 1.3 (b) that for $(\mathcal{T},\xi)\sim\mathsf{IMGWR}$ the sequence $(\mathcal{T}^{(Y_0)},\mathcal{T}^{(Y_1)},\ldots)$ is stationary and ergodic, and hence $V_n \to \mathbb{E}_{\mathsf{IMGWR}}[\varphi]$, IMGWR-a.s. by the Birkhoff ergodic theorem ([9, Thm. 6.2.1]) as soon as it is verified that $\varphi \in L^1(\mathsf{IMGWR})$. We compute

$$\mathbb{E}_{\mathsf{IMGWR}}[\varphi] = \frac{1}{2\rho}\mathbb{E}_{\mathsf{MGW}}\left[\rho W_o^2 + \sum_{v\in\partial o} W_v^2\right] = \mathbb{E}_{\mathsf{MGW}}[W_o^2] = \eta^2\sigma^2,$$

so condition (i) is proved. Condition (ii) is checked similarly using dominated convergence. ☐

Remark 2.4. To give some indication of how our results might be extended to $\mathsf{RWRE}_{\bar\rho}$, we note that the main ingredient needed is the appropriate generalization of the normalized population

[1] Note that if $\mathcal{T}$ has a marked ray ξ, then for $v\in\xi$, $\widehat{Z}_n^v = \langle \underline{Z}_n^v, \underline{e}\rangle/\rho^n$ is not necessarily a martingale for the first $|h(v)|$ steps. Nevertheless it is eventually a martingale so we can still define W_v to be the a.s. limit of $\widehat{Z}_n^v$.

size: we define it to be the random variable $\overline{W}_o$ which is the a.s. limit of the martingale $\widehat{Z}_n \equiv \widehat{Z}_n^{(1)}$ defined by (5.1). If $\overline{W}_v$ denotes the normalized population size of $\mathcal{T}^{(v)}$, then

$$\bar{\rho}\overline{W}_v = \sum_{w \in \partial^+ v} \alpha_w \overline{W}_w,$$

so the $\overline{W}_v$ can be used to define harmonic coordinates for the RWRE. In the single-type case, $\overline{W}_o$ has finite second moment if and only if $\kappa \geq 2$ [19, Thm. 2.1], so clearly Propn. 2.3 cannot apply outside this regime. We emphasize again that due to the same technical barriers which arise in [10], simple adaptations of our proof will not cover the full regime $\kappa \geq 2$.

3. Proof of Thm. 1.4

In this section we prove the quenched CLT for IMGWR trees by controlling the corrector

$$\varepsilon_t := \frac{M_t}{\eta} - h(Y_t)$$

on the interval $0 \leq t \leq n$. For $1/2 < \delta < 1$ and $n \geq 0$ fixed, let $\tau^n(j)$, for $j\lfloor n^\delta \rfloor \leq n$ denote integer times chosen uniformly at random (independently of one another and of the random walk Y) from the interval $[j\lfloor n^\delta \rfloor, (j+1)\lfloor n^\delta \rfloor)$.

Proposition 3.1. *Assume* (H1), (H2), *and* (H3^p) *with* $p > 2$. *There exists* $\delta_0 \equiv \delta_0(p) \in (1/2, 1)$ *such that for* $\delta_0 \leq \delta < 1$ *and* $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_{(\mathcal{T},\xi)}\left(\max_{j\lfloor n^\delta \rfloor \leq n} \left|\varepsilon_{\tau^n(j)}\right| \geq \epsilon\sqrt{n} \right) = 0, \quad \mathsf{IMGWR}\text{-}a.s. \tag{3.1}$$

Further, for any ϵ' *with* $2\epsilon' + \delta < 1$,

$$\lim_{n\to\infty} \mathbb{P}_{(\mathcal{T},\xi)}\left(\max_{r,s \leq n, |r-s| \leq n^\delta} |h(Y_r) - h(Y_s)| \geq n^{1/2-\epsilon'} \right) = 0, \quad \mathsf{IMGWR}\text{-}a.s. \tag{3.2}$$

Given this proposition, we can prove the quenched CLT for RW_ρ on IMGWR trees:

Proof of Thm. 1.4. If $t \leq n$ then $|t - \tau^n(j)| \leq \lfloor n^\delta \rfloor$ for some j, so

$$\max_{t \leq n} |\varepsilon_t| \leq \max_{r,s \leq n, |r-s| \leq \lfloor n^\delta \rfloor} \left| \frac{M_r}{\eta} - \frac{M_s}{\eta} \right| + \max_{j\lfloor n^\delta \rfloor \leq n} |\varepsilon_{\tau^n(j)}| + \max_{r,s \leq n, |r-s| \leq \lfloor n^\delta \rfloor} |h(Y_r) - h(Y_s)|.$$

M satisfies a CLT by Propn. 2.3, and it follows from (3.1) and (3.2) that

$$\lim_{n\to\infty} \mathbb{P}_{(\mathcal{T},\xi)}\left(\max_{t \leq n} |\varepsilon_t| \geq \epsilon\sqrt{n} \right) = 0, \quad \mathsf{IMGWR}\text{-a.s.},$$

which gives the result. □

The remainder of this section is devoted to the proof of Propn. 3.1. By (2.7),

$$\frac{1}{\sqrt{n}} \max_{j\lfloor n^\delta \rfloor \leq n} \left|\varepsilon_{\tau^n(j)}\right| \leq E_1 + E_2 \tag{3.3}$$

where, with $R_t \equiv R_{Y_t}$ denoting the nearest ancestor of Y_t on ξ,

$$E_1 := \frac{1}{\sqrt{n}} \max_{t \leq 2n} \left| \frac{S_{R_t}}{\eta} - h(R_t) \right|, \quad E_2 := \frac{1}{\sqrt{n}} \max_{j\lfloor n^\delta \rfloor \leq n} \left| \frac{S^\xi_{Y_{\tau^n(j)}}}{\eta} - d(Y_{\tau^n(j)}, \xi) \right|.$$

Our strategy will be to control the two terms individually. We will make use of the following classical result:

Lemma 3.2 ([28, p. 60])**.** *If $z_1, \dots, z_n$ are independent random variables with $\mathbb{E}z_i = 0$ and $\mathbb{E}|z_i|^p < \infty$, then*

$$\mathbb{E}\left[\left|\sum_{i=1}^{n} z_i\right|^p\right] \le \begin{cases} 2\sum_{i=1}^n \mathbb{E}[|z_i|^p] & \text{if } 1 \le p \le 2, \\ C(p) n^{p/2-1} \sum_{i=1}^n \mathbb{E}[|z_i|^p] & \text{if } p \ge 2. \end{cases} \tag{3.4}$$

Let $\mathcal{T}$ be a rooted tree. Recalling the definition of S_v for rooted trees ((2.6)), for $\epsilon > 0$ we write

$$A_n^\epsilon = \left\{ v \in D_n : \left| \frac{S_v}{n} - \eta \right| > \epsilon \right\}, \quad n \ge 1. \tag{3.5}$$

The following lemma says that the harmonic coordinates $(S_v)_{v\in\mathcal{T}}$, rescaled by η, are a good approximation to the actual coordinates $|v|$ on the MGW rooted trees. Let $H_o := \min\{t > 0 : X_t = o\}$ denote the first return time to the root by the walk X.

Lemma 3.3. *Assume* (H1), (H2), *and* (H3^p) *with $p \ge 2$. For any $\epsilon > 0$, the expected number of visits to A_k^ϵ during a single excursion away from the root is*

$$\mathbb{E}_{\mathsf{MGW}}\left[\sum_{t=0}^{H_o} \mathbf{1}_{\{X_t \in A_k^\epsilon\}}\right] \le \frac{C}{\rho^k} \mathbb{E}_{\mathsf{MGW}}\left[\sum_{v\in A_k^\epsilon} (1 + d_v)\right] \le C(p, \epsilon) k^{-p/2}.$$

Proof. If $v \in \mathcal{T}$ with $|v| = k \ge 1$, a simple conductance calculation (see [24, Ch. 2]) gives

$$\mathbb{E}_{\mathcal{T}}\left[\sum_{t=0}^{H_o} \mathbf{1}_{\{X_t = v\}}\right] = \frac{\mathbb{P}_{\mathcal{T}}(o \to v)}{\mathbb{P}_{\mathcal{T}}(v \to o)} = \frac{\rho + d_v}{d_o \rho^k}, \tag{3.6}$$

so the first inequality follows. For the second we follow the proof of [27, Lem. 3] (in particular the estimate [27, (20)]) and of [10, Lem. 4.2]. Recall from §2.3 the probability Q_k^a on rooted trees $\mathcal{T}$ obtained from a size-biasing of MGW^a, and further the probability $\mathsf{Q}_{k\star}^a$ on rooted trees $\mathcal{T}$ with a marked path $(o = v_0, \dots, v_k)$ to level k.

$$\begin{aligned} \mathbb{E}_{\mathsf{MGW}^a}\left[\sum_{v\in A_k^\epsilon} (1 + d_v)\right] &\le C \mathbb{E}_{\mathsf{MGW}^a}\left[\sum_{v\in A_k^\epsilon} e_{\chi_v}(1 + d_v)\right] \\ = \mathbb{E}_{\mathsf{Q}^a}&\left[\frac{\sum_{v\in A_k^\epsilon} e_{\chi_v}(1+d_v)}{\langle \underline{Z}_k, \underline{e}\rangle}\right] \mathbb{E}_{\mathsf{MGW}^a}[\langle \underline{Z}_k, \underline{e}\rangle] = e_a \rho^k \mathbb{E}_{\mathsf{Q}_{k\star}^a}\left[(1 + d_{v_k})\mathbf{1}_{\{v_k \in A_k^\epsilon\}}\right], \end{aligned}$$

so it suffices to show

$$\mathbb{E}_{\mathsf{Q}_{k\star}^a}\left[(1 + d_{v_k})\mathbf{1}_{\{v_k\in A_k^\epsilon\}}\right] \le C(p, \epsilon) k^{-p/2}.$$

To this end, writing $W_i \equiv W_{v_i}$, for $i < k$ we decompose $W_i \equiv W_{i+1}/\rho + \widetilde{W}_i$ where $\widetilde{W}_i$ is the normalized population size of $\mathcal{T}^{(v_i)} \setminus \mathcal{T}^{(v_{i+1})}$. Then

$$W_i = \sum_{j=i}^{k-1} \frac{\widetilde{W}_j}{\rho^{j-i}} + \frac{W_k}{\rho^{k-i}},$$

so

$$\frac{S_{v_k}}{k} - \eta = \frac{1}{k} C_k W_k + \frac{1}{k}\sum_{i=1}^{k-1} C_i \widetilde{W}_i - \eta, \quad C_i := \sum_{j=0}^{i-1} \rho^{-j} \le C_\infty := \frac{\rho}{\rho - 1}.$$

Conditional on the types $(\chi_i \equiv \chi_{v_i})_{i=1}^k$, the random variables $\widetilde{W}_1, \ldots, \widetilde{W}_{k-1}$ are independent of one another and of the pair (W_k, d_{v_k}), and all these random variables have finite moments of order p by Propn. 2.2. Therefore

$$\mathbb{E}_{\mathsf{Q}^a_{k\star}}\left[(1+d_{v_k})\mathbf{1}_{\{v_k\in A^\epsilon_k\}}\right] \le C\mathsf{Q}^a_{k\star}\left(\left|\frac{1}{k}\sum_{i=1}^{k-1} C_i\widetilde{W}_i - \eta\right| \ge \epsilon/2\right) + \mathbb{E}_{\mathsf{Q}^a_{k\star}}\left[d_{v_k}\mathbf{1}_{\{C_kW_k/k\ge\epsilon/2\}}\right]$$

By (H3^p), Markov's inequality, and Hölder's inequality, the second term is

$$\le \left(\frac{2C_k}{k\epsilon}\right)^{p-1}\mathbb{E}_{\mathsf{Q}^a_{k\star}}\left[d_{v_k}W_k^{p-1}\right] \le \frac{C(p,\epsilon)}{k^{p-1}} \le \frac{C(p,\epsilon)}{k^{p/2}}$$

since $p \ge 2$. As for the first term, by (3.4) and Markov's inequality,

$$\begin{aligned}
&\mathsf{Q}^a_{k\star}\left(\left|\frac{1}{k}\sum_{i=1}^{k-1} C_i\widetilde{W}_i - \mathbb{E}_{\mathsf{Q}^a_{k\star}}\left[\frac{1}{k}\sum_{i=1}^{k-1} C_i\widetilde{W}_i\middle|(\chi_i)_{i=1}^k\right]\right| > \epsilon/4\right)\\
&\le C(p)\frac{k^{p/2-1}}{(k\epsilon)^p}\left\{\sum_{i=1}^{k-1}\mathbb{E}[|C_i(\widetilde{W}_i - \mathbb{E}[\widetilde{W}_i|\chi_i])|^p]\right\} \le C(p,\epsilon)k^{-p/2}.
\end{aligned}$$

On the other hand,

$$\eta_k := \mathbb{E}_{\mathsf{Q}^a_{k\star}}\left[\frac{1}{k}\sum_{i=1}^{k-1} C_i\widetilde{W}_i\right] \to C_\infty\mathbb{E}_{\widehat{\mathsf{MGW}}}[\widetilde{W}_o] = \mathbb{E}_{\widehat{\mathsf{MGW}}}[W_o] = \eta,$$

and so

$$\mathsf{Q}^a_{k\star}\left(\left|\mathbb{E}_{\mathsf{Q}^a_{k\star}}\left[\frac{1}{k}\sum_{i=1}^{k-1} C_i\widetilde{W}_i\middle|(\chi_i)_{i=1}^k\right] - \eta\right| > \epsilon/4\right)$$

decays exponentially in k by [8, Thm. 3.1.2]. Combining these estimates completes the proof. □

The next step is to use Lem. 3.3 to show that on the IMGWR trees, S^ξ_v/η is a good approximation to $d(v,\xi)$. Recalling (2.7), set

$$B^\epsilon_k(\mathcal{T},\xi) = \left\{w \in \mathcal{T} : d(w,\xi) = k, \left|\frac{S^\xi_w}{k} - \eta\right| > \epsilon\right\}, \quad B^\epsilon(\mathcal{T},\xi) = \bigcup_{k\ge1} B^\epsilon_k(\mathcal{T},\xi). \tag{3.7}$$

Lemma 3.4. *Assume* (H1), (H2), *and* (H3^p) *with* $p > 2$. *There exists* $\delta_0 \equiv \delta_0(p) \in (1/2,1)$ *such that for* $\delta_0 \le \delta < 1$ *and* $\epsilon > 0$,

$$\lim_{n\to\infty}\mathbb{P}_{(\mathcal{T},\xi)}\left(\exists j \in \mathbb{Z}_{\ge0}, j\lfloor n^\delta\rfloor \le n : Y_{\tau^n(j)} \in B^\epsilon(\mathcal{T},\xi)\right) = 0, \quad \mathsf{IMGWR}\text{-}a.s.$$

Proof. We modify the proof of [10, (22)]. Let $\epsilon > 0$ be given; without loss $\epsilon < \eta/2$. Recall from the proof of Propn. 2.1 that if $(\mathcal{T},\xi) \sim \mathsf{IMGW}_0$ then $(\mathcal{T}^{(v_k)}, o) \sim \mathsf{Q}_{k\star}$. Thus a consequence of the proof of Lem. 3.3 is that

$$\mathsf{IMGW}_0(|S_{v_k}/k + \eta| \ge \epsilon) \le C(p,\epsilon)k^{-p/2}. \tag{3.8}$$

With $p > 2$, it follows from Borel-Cantelli (using that IMGW_0 and IMGWR are mutually absolutely continuous) that for IMGWR-a.e. $(\mathcal{T},\xi)$ there exists k_0 so that $|S_{v_k}/k+\eta| < \epsilon$ for all $k \ge k_0$. For any $1/2 < \gamma < \delta$, let $H_{\lfloor n^\gamma\rfloor}$ denote the hitting time of $v_{\lfloor n^\gamma\rfloor}$ and $\Gamma_n := \{H_{\lfloor n^\gamma\rfloor} \le 2n\}$: for n sufficiently large

$$\Gamma_n \subseteq \left\{\inf_{t\le 2n} M_t \le S_{v_{\lfloor n^\gamma\rfloor}} \le (-\eta+\epsilon)\lfloor n^\gamma\rfloor\right\},$$

so $\mathbb{P}_{(\mathcal{T},\xi)}(\Gamma_n) \to 0$ as $n \to \infty$ by the invariance principle for M proved in Propn. 2.3.

On the event Γ_n^c, we decompose the walk into (possibly empty) excursions away from ξ started at v_i, $0 \le i \le \lfloor n^\gamma \rfloor$, to find

$$\mathbb{P}_{(\mathcal{T},\xi)}\left(\left\{\exists j \in \mathbb{Z}_{\ge 0}, j\lfloor n^\delta \rfloor \le n : Y_{\tau^n(j)} \in B^\epsilon(\mathcal{T},\xi)\right\} \cap \Gamma_n^c\right) \le \frac{1}{\lfloor n^\delta \rfloor}\mathbb{E}_{(\mathcal{T},\xi)}\left[\sum_{t=0}^{2n} \mathbf{1}_{\{Y_t \in B^\epsilon(\mathcal{T},\xi)\}}\right]$$
$$\le \frac{1}{\lfloor n^\delta \rfloor}\sum_{i=0}^{\lfloor n^\gamma \rfloor - 1} \mathbb{E}_{(\mathcal{T},\xi)}\left[\sum_{t=0}^{2n} \mathbf{1}_{\{Y_t \in B^\epsilon \cap \mathcal{T}^{(v_i)}\}}\right] \equiv \frac{1}{\lfloor n^\delta \rfloor}\sum_{i=0}^{\lfloor n^\gamma \rfloor - 1} G_i. \tag{3.9}$$

If L_i denotes the number of visits to v_i before $H_{\lfloor n^\gamma \rfloor}$ and N_{ij} denotes the number of visits to $B^\epsilon \cap \mathcal{T}^{(v_i)}$ during the j-th excursion away from ξ started at v_i, then Wald's identity (see e.g. [3, Exercise 22.8]) implies

$$G_i \le \mathbb{E}_{(\mathcal{T},\xi)}[L_i]\mathbb{E}_{(\mathcal{T},\xi)}[N_{i1}].$$

By a conductance estimate $\mathbb{E}_{(\mathcal{T},\xi)}[L_i] \le \mathbb{P}(v_i \to v_{\lfloor n^\gamma \rfloor})^{-1} \le Cd_i$. During a single excursion away from ξ the walk can visit only one of the $\mathcal{T}^{(w)}$ for $w \in \partial^+ v \backslash v_{i-1}$, so to bound $\mathbb{E}_{(\mathcal{T},\xi)}[N_{i1}]$ it suffices to consider an MGW rooted tree $\mathcal{T}'$ (without ray): letting

$$\widetilde{A}_k^\epsilon \equiv \widetilde{A}_k^\epsilon(\mathcal{T}') := \left\{v \in D_k : \left|\frac{W_o + S_v}{k+1} - \eta\right| > \epsilon\right\}, \quad \widetilde{A}^\epsilon := \bigcup_{k \ge 0} \widetilde{A}_k^\epsilon,$$

it follows from a (very slight) modification of Lem. 3.3 that

$$\mathbb{E}_{\mathsf{IMGW}_0}[N_{i1}|i \cup \partial^+ i] \le C\mathbb{E}_{\mathsf{MGW}}\left[\sum_{t=0}^{H_o} \mathbf{1}_{\{X_t \in \widetilde{A}^\epsilon\}}\right] \le C(p,\epsilon)\sum_{k \ge 1} k^{-p/2} \le C(p,\epsilon)$$

(using $p > 2$). Therefore

$$\mathbb{E}_{\mathsf{IMGW}_0}[G_i] \le \mathbb{E}_{\mathsf{IMGW}_0}[Cd_i \cdot \mathbb{E}_{\mathsf{IMGW}_0}[N_{i1}|i \cup \partial^+ i]] \le C(p,\epsilon).$$

It follows that

$$\frac{1}{\lfloor n^\delta \rfloor}\sum_{i=0}^{\lfloor n^\gamma \rfloor - 1} G_i \to 0, \quad \mathsf{IMGWR}\text{-a.s.}$$

The proof is completed by recalling (3.9). □

Proof of Propn. 3.1, (3.1). Recall the decomposition (3.3). For any k_0,

$$E_1 \le \frac{1}{\sqrt{n}}\max_{i \le k_0}\left|\frac{S_{v_i}}{\eta} - h(v_i)\right| + \frac{1}{\sqrt{n}}\max_{t \le 2n, h(R_t) > k_0}\left|\frac{S_{R_t}}{\eta} - h(R_t)\right|.$$

The first term clearly tends to zero as $n \to \infty$ with k_0 fixed. The second term is bounded above by

$$\left(\frac{1}{\sqrt{n}}\max_{t \le 2n}|M_t|\right)\sup_{i > k_0}\left|\frac{1}{\eta} - \frac{h(v_i)}{S_{v_i}}\right|.$$

Recalling (3.8), the second factor can be made arbitrarily small by taking k_0 large. We also have

$$E_2 \le \left(\frac{1}{\sqrt{n}}\max_{t \le 2n}|M_t|\right)\max_{j\lfloor n^\delta \rfloor \le n}\left|\frac{1}{\eta} - \frac{d(Y_{\tau^n(j)},\xi)}{S^\xi_{Y_{\tau^n(j)}}}\right|,$$

and in view of Lem. 3.4 the second factor tends to zero in probability. By the invariance principle for M proved in Propn. 2.3, $\max_{t\le 2n}|M_t|/\sqrt{n}$ stays bounded in probability as $n\to\infty$, so the result follows. $\square$

The proof of (3.2) is based on some *a priori* annealed estimates for RW_ρ:

Lemma 3.5. *There exists a constant $C<\infty$ such that for $m,n\ge 1$,*

$$\mathsf{MGW}\left(\max_{t\le n}|X_t|\ge m\right)\le Cne^{-(m+1)^2/(2n)}.$$

Proof. We modify the proof of [27, Lem. 5]. Take the finite tree with vertices $\{w\in\mathcal{T}:|w|\le m\}$, and make this into a wired tree $\mathcal{T}^\star$ by adding a new vertex $o^\star$ which is joined by an edge to each vertex in D_m. Define the modified random walk $X^\star$ on $\mathcal{T}^\star$ which follows the law of RW_ρ except at $o^\star$ where it moves to a vertex chosen uniformly at random from D_m. Then

$$\mathbb{P}_{\mathcal{T}}(\max_{t\le n}|X_t|\ge m)\le 2\sum_{t=1}^{n+1}\mathbb{P}_{\mathcal{T}^\star}(X_t=o^\star).$$

By the Carne-Varopoulos inequality (see [24, Thm. 13.4]),

$$\mathbb{P}_{\mathcal{T}^\star}(X^\star_t=o^\star)\le 2\sqrt{\frac{|D_m|}{\rho^{m-1}}}e^{-(m+1)^2/(2t)}.$$

Taking expectations gives

$$\mathsf{MGW}(|X^\star_t|=o^\star)\le Ce^{-(m+1)^2/(2t)},$$

and summing over $1\le t\le n+1$ gives the result. $\square$

Corollary 3.6. *There exists a constant $C<\infty$ such that for any $m,n\ge 1$,*

$$\mathsf{IMGWR}\left(\max_{t\le n}|h(Y_t)|\ge m\right)\le Cn^2e^{-m^2/(2n)}.$$

Proof. We argue as in the proof of [27, Cor. 2]. By decomposing into at most n excursions away from height zero and using the stationarity of IMGWR, we find

$$\begin{aligned}&\mathsf{IMGWR}\left(\max_{t\le n}h(Y_t)\ge m\right)\\&\le n\cdot\mathsf{IMGWR}\left(\exists t\le n: h(Y_t)\ge m, h(Y_s)>0\ \forall 0\le s\le t\right)\\&\le Cn\cdot\mathsf{MGW}\left(\max_{t\le n}|X_t|\ge m-1\right)\le Cn^2e^{-m^2/(2n)},\end{aligned}$$

by Lem. 3.5. The same bound holds for $\mathsf{IMGWR}(\min_{t\le n}h(Y_t)\le -m)$ by the reversibility of IMGWR, and the result follows. $\square$

Proof of Propn. 3.1, (3.2). By stationarity of IMGWR and Cor. 3.6, for any fixed s

$$\mathsf{IMGWR}\left(\max_{0\le u\le n^\delta}|h(Y_{s+u})-h(Y_s)|\ge n^{1/2-\epsilon'}\right)\le Cn^{2\delta}e^{-n^{1-2\epsilon'-\delta}/2},$$

and summing over $s\le n$ gives

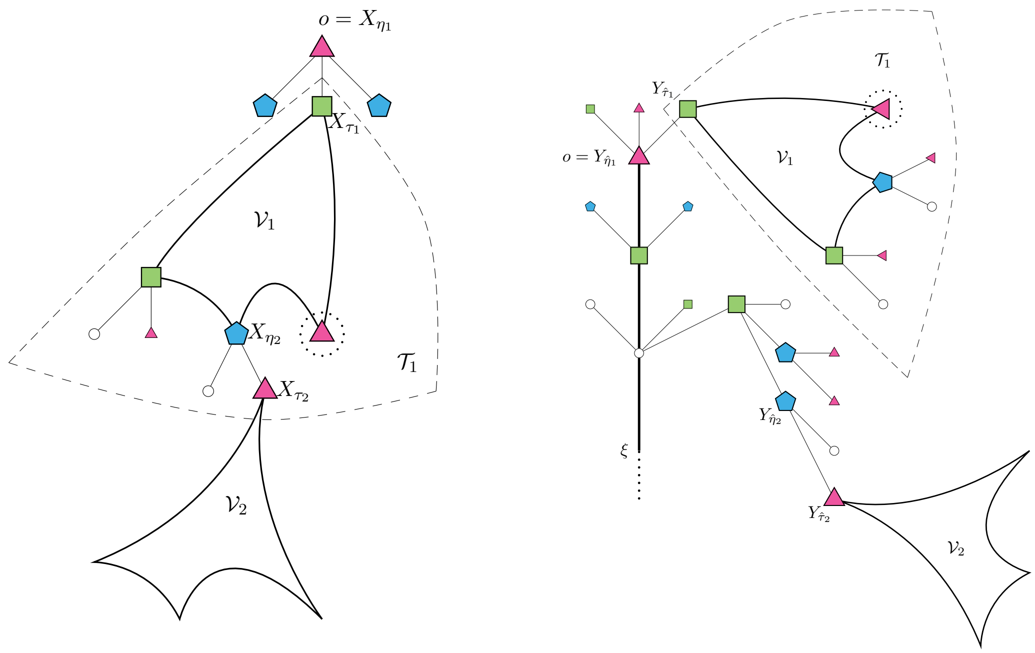
$$\mathsf{IMGWR}\left(\max_{r,s\le t,|r-s|\le n^\delta}|h(Y_r)-h(Y_s)|\ge n^{1/2-\epsilon'}\right)\le Cn^{2\delta+1}e^{-n^{1-2\epsilon'-\delta}/2},$$

which is summable in n provided $2\epsilon'+\delta<1$. The result then follows from Markov's inequality and Borel-Cantelli. $\square$

4. From IMGW to MGW by shifted coupling

4.1. **The shifted coupling construction.** To transfer the CLT on IMGWR trees to a CLT on MGW trees, we use a shifted coupling procedure which is essentially the same as the one described in [27, §6]. We begin by reviewing this construction, following the notation of [27] as much as possible.

For any tree $\mathcal{T}$ (with or without marked ray) let $\mathcal{L}\mathcal{T}$ denote the set of leaves and $\mathcal{T}^\circ := \mathcal{T} \backslash \mathcal{L}\mathcal{T}$. Given trees $\mathcal{T}_i$ rooted at o_i $(i = 1, 2)$ and $v \in \mathcal{L}\mathcal{T}_1$, let $\mathcal{T}_1 \circ^v \mathcal{T}_2$ denote the tree rooted at o_1 which is obtained by identifying the root of $\mathcal{T}_2$ with v.



(A) MGW excursions (B) IMGW_0 excursions

Figure 2. Shifted coupling

Now let $(\mathcal{T}, (X_t)_{t\geq 0}) \sim \mathsf{MGW} \otimes \mathsf{RW}_\rho$. Set $\tau_0 = 0, \eta_0 = 0$, and let $\mathcal{U}_0$ be the subtree consisting of vertices $\{o\} \cup \partial^+ o$. For $i \geq 1$, let

$$\tau_i := \min \{t > \eta_{i-1} : X_t \in \mathcal{L}\mathcal{U}_{i-1}\} \qquad \text{excursion start,}$$

$$\eta_i := \min \left\{t > \tau_i : X_t \in \mathcal{U}_{i-1}^\circ\right\} \qquad \text{excursion end,}$$

$$\mathcal{V}_i := \{v \in \mathcal{T} : X_t = v \text{ for some } t \in [\tau_i, \eta_i)\} \qquad \text{vertices visited during } i\text{-th excursion,}$$

$$\bar{\mathcal{V}}_i := \mathcal{V}_i \cup \left\{w \in \mathcal{T} : w \in \partial^+ v \text{ for some } v \in \mathcal{V}_i\right\},$$

$$\mathcal{T}_i := \text{subtree of } \mathcal{T} \text{ with vertices } \bar{\mathcal{V}}_i \text{ and root } X_{\tau_i},$$

$$\mathcal{U}_i := \mathcal{U}_{i-1} \circ^{X_{\tau_i}} \mathcal{T}_i \qquad \text{subtree explored by time } \eta_i.$$

See Fig. 2a; the vertices visited by X are enlarged and the dotted circle marks a vertex with no children. We take the convention $\min \varnothing \equiv \infty$, and let $H_{\mathbb{X}} := \max \{i : \eta_i < \infty\}$ be the total number of excursions (so $\{H_{\mathbb{X}} < \infty\} = \mathbb{X}$).

Next we construct a *coupled* realization $((\widehat{\mathcal{T}}, \xi), (Y_t)_{t\geq 0}) \sim \mathsf{IMGW}_0 \otimes \mathsf{RW}_\rho$ as follows: first construct the backbone $\widehat{\mathcal{U}}_0$ of the tree (ξ and $\partial^+ v_i$ for $i \geq 1$, together with types) in the manner described in §2.1. Set $\hat{\tau}_0 = \hat{\eta}_0 = 0$, and start a ρ-biased random walk Y_t on $\widehat{\mathcal{U}}_0$ with $Y_0 = o$. The goal is to match excursions of Y into unexplored territory with the excursions of X into unexplored territory. For the single-type case this is accomplished by waiting until Y reaches a leaf node of $\widehat{\mathcal{U}}_0$, attaching $\mathcal{T}_1$ to the leaf node, letting Y take the path of X through $\mathcal{T}_1$, and then repeating the procedure inductively.

In the multi-type case we cannot attach $\mathcal{T}_1$ at a leaf node unless the types match, so we make the following natural modification: let $\hat{\tau}_0 = \hat{\eta}_0 = 0$, and let o_i denote the root of $\mathcal{T}_i$. For $i \geq 1$, consider the process $(Y_t)_{t \geq \hat{\eta}_{i-1}}$, and suppose it has reached $v \in \mathcal{L}\widehat{\mathcal{U}}_{i-1}$. If $\chi_v \neq \chi_{o_i}$, we enlarge $\widehat{\mathcal{U}}_{i-1}$ by attaching offspring $\partial^+ v$ according to the law $\mathbf{q}^{\chi_v}$. In this manner we continue to let Y travel according to RW_ρ, attaching offspring as needed; by abuse of notation we continue to call the enlarged tree $\widehat{\mathcal{U}}_{i-1}$. This process terminates at time

$$\hat{\tau}_i := \min \left\{ t > \hat{\eta}_{i-1} : Y_t \in \mathcal{L}\widehat{\mathcal{U}}_{i-1}, \chi_{Y_t} = \chi_{o_i} \right\},$$

and we then set

$$\begin{aligned} \widehat{\mathcal{U}}_i &:= \widehat{\mathcal{U}}_{i-1} \circ^{Y_{\hat{\tau}_i}} \mathcal{T}_i && \text{extended tree,} \\ \hat{\eta}_i &:= \hat{\tau}_i + \eta_i - \tau_i && \text{excursion end,} \\ (Y_{t'})_{t' \in [\hat{\tau}_i, \hat{\eta}_i)} &:= (X_t)_{t \in [\tau_i, \eta_i)} && \text{coupled excursion,} \\ Y_{\hat{\eta}_i} &:= \text{ancestor of } Y_{\hat{\tau}_i} = Y_{\hat{\tau}_i - 1}. \end{aligned}$$

See Fig. 2b; the vertices visited by Y are enlarged. Finally, with $\widehat{\mathcal{U}} := \lim_i \widehat{\mathcal{U}}_i$, we define $\widehat{\mathcal{T}}$ by attaching to each vertex of $\mathcal{L}\widehat{\mathcal{U}}$ an MGW tree, depending on type. We thus obtain the following extension of [27, Lem. 8]:

Lemma 4.1. *If $(\mathcal{T}, (X_t)_{t\geq 0}) \sim \mathsf{MGW} \otimes \mathsf{RW}_\rho$ then the marginal law of $((\widehat{\mathcal{T}}, \xi), (Y_t)_{t \geq 0})$ arising from the above construction is $\mathsf{IMGW}_0 \otimes \mathsf{RW}_\rho$.*

4.2. **Annealed CLT for MGW.** In this section we will transfer the IMGWR quenched CLT to the following *annealed* CLT on MGW trees:

Proposition 4.2. *Assume* (H1), (H2) *and* (H3^p) *with $p > 4$. If X has law $\mathsf{MGW} \otimes \mathsf{RW}_\rho$ conditioned on $\mathbb{X}^c$, the processes $(|X_{\lfloor nt \rfloor}|/(\sigma\sqrt{n}))_{t\geq 0}$ converge in law to the absolute value of a standard Brownian motion.*

Let $\mathcal{R}_t = h(Y_t) - \min_{0 \leq s \leq t} h(Y_s) \geq 0$. By Thm. 1.4, for IMGW_0-a.e. $(\widehat{\mathcal{T}}, \xi)$, the process $\mathcal{R}_{\lfloor nt \rfloor}/(\sigma\sqrt{n})$ converges to a Brownian motion minus its running minimum, which is the same in law as the absolute value of a Brownian motion (see e.g. [14, Thm. 3.6.17]). Thus to deduce Propn. 4.2 we need to estimate the relation between the processes $|X_n|$ and $\mathcal{R}_n$. To this end, let $\mathbf{t}$, $\hat{\mathbf{t}}$ be monotone increasing bijections

$$\mathbf{t} : \mathbb{Z}_{\geq 0} \to \bigcup_{i \geq 1} [\eta_{i-1}, \tau_i), \quad \hat{\mathbf{t}} : \mathbb{Z}_{\geq 0} \to \bigcup_{i \geq 1} [\hat{\eta}_{i-1}, \hat{\tau}_i).$$

We make the following notations (the left column refers to the MGW tree, while the right column refers to the IMGW_0 tree):

$$\begin{aligned}
&\widetilde{X}_s = X_{\mathbf{t}(s)}, && \widetilde{Y}_s = Y_{\mathbf{t}(s)} \\
&\mathcal{H}_s = \sigma(X_t : t \le \mathbf{t}(s)) && \widehat{\mathcal{H}}_s = \sigma(Y_t : t \le \mathbf{t}(s)), \\
&J_i = \mathbf{t}^{-1}[\eta_{i-1}, \tau_i), && \widehat{J}_i = \hat{\mathbf{t}}^{-1}[\hat{\eta}_{i-1}, \hat{\tau}_i); \\
&I_n = \max\{i : \tau_i \le n\}, && \widehat{I}_n = \max\{i : \hat{\tau}_i \le n\}; \\
&\Delta_n = \textstyle\sum_{i=1}^{I_n} |J_i|, && \widehat{\Delta}_n = \textstyle\sum_{i=1}^{\widehat{I}_n} |\widehat{J}_i|; \\
&\Delta_n^\alpha = \textstyle\sum_{i=1}^{I_n} \left|\left\{s \in J_i : |\widetilde{X}_s| \le n^\alpha\right\}\right|, && \widehat{\Delta}_n^\alpha = \textstyle\sum_{i=1}^{\widehat{I}_n} \left|\left\{s \in \widehat{J}_i : d(\widetilde{Y}_s, \xi) \le n^\alpha\right\}\right|.
\end{aligned}$$

In words, given the walk X on the MGW tree, $\widetilde{X}_s$ is the "inter-excursion process" adapted to the filtration $\mathcal{H}_s$, J_i refers to the i-th inter-excursion interval, I_n is the number of such intervals contained in $[0, n)$, Δ_n is the total length of these intervals, and Δ_n^α is the length of these intervals except for times spent at distance more than n^α from the root. The right column defines the analogous objects for the walk on the IMGW_0 tree. Finally, let

$$\mathbf{B}_n := \max_{0 \le r \le s \le n} \{h(Y_s) - h(Y_r) : Y_r, Y_s \in \xi\}$$

denote the maximum displacement by time n against the backward drift on ξ.

Lemma 4.3. *Assume* (H1), (H2), *and* (H3^p) *with* $p > 4$. *Then*

(a) There exists $\alpha \equiv \alpha(p) < 1/2$ *so that* $\lim_{n\to\infty} \mathsf{MGW}(\Delta_n^\alpha \ne \Delta_n | \mathbb{X}^c) = \lim_{n\to\infty} \mathsf{IMGW}_0(\widehat{\Delta}_n^\alpha \ne \widehat{\Delta}_n | \mathbb{X}^c) = 0$*;*

(b) $\mathbf{1}_{\mathbb{X}^c} |\Delta_n|/n \xrightarrow{p} 0$ *under* MGW*;*

(c) $\mathbf{1}_{\mathbb{X}^c} |\widehat{\Delta}_n|/n \xrightarrow{p} 0$ *under* IMGW_0*;*

(d) $\mathbf{1}_{\mathbb{X}^c} \mathbf{B}_n/\sqrt{n} \xrightarrow{p} 0$ *under* IMGW_0.

Given this lemma, we can prove the annealed MGW CLT:

Proof of Propn. 4.2. Let $\mathbf{s} : \mathbb{Z}_{\ge 0} \to \mathbb{Z}_{\ge 0}$ be a nondecreasing map which maps each $[\hat{\tau}_i, \hat{\eta}_i)$ bijectively to $[\tau_i, \eta_i)$. Then for $t \in [\hat{\tau}_i, \hat{\eta}_i)$ we have $|X_{\mathbf{s}(t)}| - |X_{\tau_i}| = d(Y_t, \xi) - d(Y_{\hat{\tau}_i}, \xi)$, so, recalling that R_v denotes the nearest ancestor of v on ξ,

$$||X_{\mathbf{s}(t)}| - \mathcal{R}_t| = \left||X_{\mathbf{s}(t)}| - d(Y_t, \xi) - h(R_t) + \min_{i \le t} h(Y_i)\right| \le |X_{\tau_i}| + d(Y_{\hat{\tau}_i}, \xi) + |\mathbf{B}_n|.$$

It follows that on the event $\{\Delta_n^\alpha = \Delta_n\} \cap \{\widehat{\Delta}_n^\alpha = \widehat{\Delta}_n\}$,

$$\frac{1}{\sqrt{n}} \max_{0 \le t \le n} ||X_{\mathbf{s}(t)}| - \mathcal{R}_t| \le \frac{2n^\alpha + |\mathbf{B}_n|}{\sqrt{n}},$$

so by Thm. 1.4, Lem. 4.1, and Lem. 4.3 (a) and (e) the processes $(|X_{\mathbf{s}(\lfloor nt \rfloor)}|/(\sigma\sqrt{n}))_{t \ge 0}$ converge in law to a reflected Brownian motion. On the other hand, Lem. 4.3 (b) and (c) imply that $n^{-1} \max_{0 \le t \le 1} (\mathbf{s}(\lfloor nt \rfloor) - \lfloor nt \rfloor) \xrightarrow{p} 0$ so we obtain the CLT for the processes $(|X_{\lfloor nt \rfloor}|/(\sigma\sqrt{n}))_{t \ge 0}$ from the continuity of Brownian motion. □

The remainder of this subsection is devoted to the proof of Lem. 4.3. Let $\mathcal{C}_{o,\ell} \equiv \mathcal{C}(o, \ell)$ denote the conductance between o and D_ℓ in $\mathcal{T}$, with respect to the stationary measure ϖ for $\mathsf{RW}_\rho(\mathcal{T})$ with the normalization $\varpi(o) = d_o$. The following estimate, a version of [26, Lem. 2.2], will be proved in §6.3:

Proposition 4.4. *(a) Under* (H1), (H2), *and* (H3^2), *there exist* $0 < r, C < \infty$ *such that for all* $\epsilon > 0$, $\mathsf{MGW}(\mathcal{C}_{o,k}^{-1} \geq k^{1+\epsilon} | \mathbb{X}^c) \leq C k^{-r\epsilon}$.

(b) If further (H3^p) *holds with* $p > 2$, *then for* MGW*-a.e.* $\mathcal{T} \notin \mathbb{X}$ *there exists a random constant* $C_{\mathcal{T}} < \infty$ *such that* $\mathcal{C}_{o,k}^{-1} \leq C_{\mathcal{T}} k$ *for all* k.

Corollary 4.5. *Assume* (H1), (H2), *and* (H3^p) *with* $p > 2$, *and let* $N^\alpha(n) := \sum_{t=0}^n \mathbf{1}_{\{|X_t| \leq n^\alpha\}}$. *There exists* $0 < c, C < \infty$ *(depending on* α*) such that for all* $\epsilon > 0$,

$$\mathsf{MGW}(N^\alpha(n) \geq n^{1/2+\alpha+\epsilon} | \mathbb{X}^c) \leq C n^{-c\epsilon}.$$

Proof. By Markov's inequality it suffices to show

$$\mathsf{MGW}(\mathbb{E}_{\mathcal{T}}[N^\alpha(n)] \geq n^{1/2+\alpha+\epsilon} | \mathbb{X}^c) \leq C n^{-c\epsilon}.$$

If $H_n^{\epsilon'}$ denotes the first hitting time of level $\lfloor n^{1/2+\epsilon'} \rfloor$, then $\mathsf{MGW}(H_n^{\epsilon'} \leq n) \leq e^{-cn^{2\epsilon'}}$ by Lem. 3.5, so it suffices to show this result replacing $N^\alpha(n)$ with $N^\alpha(H_n^{\epsilon'})$ for some $\epsilon' > 0$.

Letting $L_o(n)$ denote the number of visits to o by time n, $\mathbb{E}_{\mathcal{T}}[L_o(H_n^{\epsilon'})] \leq d_o \mathcal{C}(o, \lfloor n^{1/2+\epsilon'} \rfloor)^{-1}$ so by Propn. 4.4

$$\mathsf{MGW}\left(\mathbb{E}_{\mathcal{T}}[L_o(H_n^{\epsilon'})] \geq n^{1/2+2\epsilon'} | \mathbb{X}^c\right) \leq C n^{-c\epsilon'}. \tag{4.1}$$

If V^α denotes the number of visits to levels $[0, \lfloor n^\alpha \rfloor]$ during a single excursion away from o, then (3.6) implies

$$\mathbb{E}_{\mathcal{T}}[V^\alpha] \leq C \sum_{k=0}^{\lfloor n^\alpha \rfloor + 1} |D_k|,$$

so by Markov's inequality

$$\mathsf{MGW}(\mathbb{E}_{\mathcal{T}}[V^\alpha] \geq n^{\alpha+\epsilon'}) \leq C n^{-\epsilon'}. \tag{4.2}$$

By Wald's identity

$$\mathbb{E}_{\mathcal{T}}[N^\alpha(H_n^{\epsilon'})] \leq \mathbb{E}_{\mathcal{T}}[L_o(H_n^{\epsilon'})] \mathbb{E}_{\mathcal{T}}[V^\alpha],$$

so setting $\epsilon' = \epsilon/3$ gives the desired result. □

Most of the technical estimates required for the proof of Lem. 4.3 are contained in the following auxiliary lemma (cf. [10, Lem. 7.3]). For fixed n let $\ell(n) \equiv \lfloor (\log n)^{3/2} \rfloor$ and define the sequence of $(\mathcal{H}_s)$-stopping times

$$\Theta_0 = 0, \quad \Theta_{j+1} = \min\left\{ s > \Theta_j : \left| |\widetilde{X}_s| - |\widetilde{X}_{\Theta_j}| = \ell(n) \right| \right\}$$

and similarly the sequence of $(\widehat{\mathcal{H}}_s)$-stopping times

$$\widehat{\Theta}_0 = 0, \quad \widehat{\Theta}_{j+1} = \min\left\{ s > \widehat{\Theta}_j : \left| d(\widetilde{Y}_s, \xi) - d(\widetilde{Y}_{\widehat{\Theta}_j}, \xi) \right| = \ell(n) \right\}.$$

Lemma 4.6. *Assume* (H1) *and* (H2).

(a) Assuming (H3^p),

$$\mathsf{MGW}\left(\max_{t \leq n} W_{X_t} > n^{1/4-\epsilon} \right) \leq C(\epsilon, p) n^{1-p(1/4-\epsilon)}$$

$$\mathsf{IMGW}_0\left(\max_{t \leq n} \mathbf{1}_{\{Y_t \notin \xi\}} W_{Y_t} > n^{1/4-\epsilon} \right) \leq C(\epsilon, p) n^{1-p(1/4-\epsilon)}.$$

In particular, if $p > 4$, *there exists* $\epsilon \equiv \epsilon(p) > 0$ *such that these probabilities tend to zero.*

(b) Assuming (H3^p) *with* $p > 2$, *for any* $\epsilon > 0$ *there exists* $C \equiv C(p, \epsilon)$ *such that*

$$\mathsf{MGW}\left(I_n \leq n^{1/2+\epsilon} \middle| \mathbb{X}^c\right) \leq e^{-Cn^{\epsilon/2}}, \quad \mathsf{IMGW}_0\left(\widehat{I}_n \leq n^{1/2+\epsilon}\right) \leq e^{-Cn^{\epsilon/2}}.$$

(c) With $\Theta_j, \widehat{\Theta}_j$ *defined as above,*

$$\mathsf{MGW}(\exists i \leq I_n, \Theta_{j-1}, \Theta_j \in J_i, |\widetilde{X}_{\Theta_j}| > |\widetilde{X}_{\Theta_{j-1}}|) \leq Cn^2 \rho^{-\ell(n)},$$
$$\mathsf{IMGW}_0(\exists i \leq \widehat{I}_n, \widehat{\Theta}_{j-1}, \widehat{\Theta}_j \in \widehat{J}_j, d(\widetilde{Y}_{\widehat{\Theta}_j}, \xi) > d(\widetilde{Y}_{\widehat{\Theta}_{j-1}}, \xi)) \leq e^{-c\ell(n)/2}.$$

(d) Assuming (H3^p) *with* $p > 2$, *for any* $\epsilon > 0$ *there exists* $\alpha \equiv \alpha(p) \in (0, 1/2)$ *such that if*

$$A^{(n)} := \bigcup_{k=\lfloor n^\alpha \rfloor - \lfloor (\log n)^2 \rfloor}^{\lfloor n^\alpha \rfloor} A_k^\epsilon, \quad B^{(n)} := \bigcup_{k=\lfloor n^\alpha \rfloor - \lfloor (\log n)^2 \rfloor}^{\lfloor n^\alpha \rfloor} B_k^\epsilon,$$

(where A_k^ϵ, B_k^ϵ *are as defined in* (3.5), (3.7) *respectively) then*

$$\lim_{n\to\infty} \mathsf{MGW}\left(\exists t \leq n : X_t \in A^{(n)}\right) = 0, \tag{4.3}$$
$$\lim_{n\to\infty} \mathsf{IMGW}_0\left(\exists t \leq n : Y_t \in B^{(n)}\right) = 0. \tag{4.4}$$

Further, if $p > 4$ *then* α *can be chosen such that for constants* $0 < c, C < \infty$ *depending on* ϵ, p,

$$\mathsf{MGW}\left(\exists t \leq n : X_t \in \bigcup_{k \geq n^\alpha} A_k^\epsilon\right) \leq Cn^{-c}. \tag{4.5}$$

Proof. (a) See proof of [27, (63)].

(b) We will show that with probability $\geq 1 - e^{-Cn^{\epsilon/2}}$ conditioned on $\mathbb{X}^c$, one of the first $\lfloor n^{1/2+\epsilon} \rfloor$ excursions has length $\eta_i - \tau_i > n$ which certainly implies the result. Conditioning on $\mathbb{X}^c$ is needed simply to ensure $H_{\mathbb{X}} = \infty$; for the purpose of proving the claim we may artifically define $\eta_i - \tau_i = \infty$ for $i > H_{\mathbb{X}}$.

Then, conditioned on $(\eta_j - \tau_j)_{j=1}^{i-1}$, the probability that $\eta_i - \tau_i > n$ dominates

$$C \cdot \mathsf{MGW}(H_o > n)$$

where H_o is the first return time to o. For a fixed rooted tree $\mathcal{T}$, let $H_n^{\epsilon/2}$ denote the first hitting time of level $\lfloor n^{1/2+\epsilon/2} \rfloor$. Then

$$\mathbb{P}_{\mathcal{T}}(H_o > n) \geq \mathbb{P}_{\mathcal{T}}(H_o > H_n^{\epsilon/2} > n) \geq \mathbb{P}_{\mathcal{T}}(H_o > H_n^{\epsilon/2}) - \mathbb{P}_{\mathcal{T}}(H_n^{\epsilon/2} \leq n),$$

so Propn. 4.4 and Lem. 3.5 imply $\mathsf{MGW}(H_o > n) \geq C/n^{1/2+\epsilon/2}$ for $C \equiv C(p, \epsilon)$. Thus the probability that none of the first $\lfloor n^{1/2+\epsilon} \rfloor$ excursions has length $> n$ is

$$\leq \left(1 - \frac{C}{n^{1/2+\epsilon/2}}\right)^{n^{1/2+\epsilon}} \leq e^{-Cn^{\epsilon/2}},$$

which proves the result.

(c) See proof of [27, Lem. 11] for the bound for MGW. For the IMGW_0 tree, the event of interest implies one of two possibilities:

1. there exist times $t_0 < t_1 < t_2 \leq n$ with $Y_{t_0} = Y_{t_2}$ and $d(Y_{t_0}, \xi) = d(Y_{t_1}, \xi) + \lfloor \ell(n)/2 \rfloor$, or
2. there exist times $t_1 < t_2 \leq n$ with $d(Y_{t_2}, \xi) = d(Y_{t_1}, \xi) + \lfloor \ell(n)/2 \rfloor$ such that some $a \in \mathcal{Q}$ does not appear on the geodesic between the Y_{t_i}.

By a simple random walk estimate and summing over at most n^2 possibilities for (Y_{t_0}, Y_{t_1}), the first event has probability $\leq Cn^2\rho^{-\lfloor \ell(n)/2 \rfloor}$. The second event has probability $\leq e^{-c\ell(n)}$ by the construction of IMGW_0 and the irreducibility of the Markov chain, and combining these estimates gives the result.

(d) For the proof of (4.3), arguing as in Cor. 4.5 it suffices to show

$$\lim_{n\to\infty} \mathsf{MGW}(\exists t \leq H_n^\epsilon : X_t \in A^{(n)}) = 0.$$

By Wald's identity

$$\mathbb{P}_{\mathcal{T}}(\exists t \leq H_n^\epsilon : X_t \in A^{(n)}) \leq \mathbb{E}_{\mathcal{T}}[L_o(H_n^\epsilon)] \cdot P_n$$

where P_n is the probability (conditioned on $\mathcal{T}$) to visit $A^{(n)}$ during a single excursion away from o. By (4.1), $\mathbb{E}_{\mathcal{T}}[L_o(H_n^\epsilon)] \leq n^{1/2+2\epsilon}$ except with probability Cn^{-c} for C, c depending on p, ϵ. By Lem. 3.3, $\mathbb{E}_{\mathsf{MGW}}[P_n] \leq C(\alpha, \epsilon, p)(\log n)^2(n^\alpha)^{-p/2}$, so, using $p > 2$ and choosing α sufficiently close to 1/2 and ϵ sufficiently small, it follows from Markov's inequality that $\mathsf{MGW}(P_n \geq n^{-(1/2+3\epsilon)}) \leq Cn^{-c}$, and (4.3) follows. On the other hand, if P_n' is the probability (conditioned on $\mathcal{T}$) to visit $\bigcup_{k\geq n^\alpha} A_k^\epsilon$ during a single excursion away from o, then $\mathbb{E}_{\mathsf{MGW}}[P_n] \leq C(\alpha, \epsilon, p)(n^\alpha)^{1-p/2}$, so (4.5) follows by a very similar argument.

For (4.4), arguing as in the proof of Lem. 3.4 it suffices to show

$$\lim_{n\to\infty} \sum_{i=0}^{\lfloor n^\gamma \rfloor - 1} \mathsf{IMGW}_0(\exists t \leq n : Y_t \in B^{(n)} \cap \mathcal{T}^{(v_i)}) = 0.$$

If $\widehat{P}_i$ denotes the probability conditioned on $(\mathcal{T}, \xi)$ to visit $B^{(n)}$ during a single excursion away from ξ started at v_i, then

$$\mathbb{P}_{(\mathcal{T},\xi)}(\exists t \leq n : Y_t \in B^{(n)} \cap \mathcal{T}^{(v_i)}) \leq \mathbb{E}_{(\mathcal{T},\xi)}[L_i]\widehat{P}_i \leq Cd_i\widehat{P}_i.$$

By a slight modification of Lem. 3.3,

$$\mathbb{E}_{\mathsf{IMGW}_0}[\widehat{P}_i | i \cup \partial^+ i] \leq C(\epsilon, p)(\log n)^2 n^{-\alpha p/2},$$

therefore

$$\sum_{i=0}^{\lfloor n^\gamma \rfloor - 1} \mathsf{IMGW}_0(\exists t \leq n : Y_t \in B^{(n)} \cap \mathcal{T}^{(v_i)}) \leq C(\epsilon, p)(\log n)^2 n^{\gamma - \alpha p/2}.$$

Since $p > 2$ this tends to zero by choosing γ, α sufficiently close to 1/2. □

Proof of Lem. 4.3. (a) See proof of [10, (28),(29)], making use of Lem. 4.6 to replace [10, Lem. 5.1].

(b) By (a) it suffices to show $\mathbf{1}_{\mathbb{X}^c} N^\alpha(n)/n \xrightarrow{p} 0$ under MGW, and a stronger result was proved in Cor. 4.5.

(c) By (a) it suffices to show $\mathbf{1}_{\mathbb{X}^c} \widehat{N}^\alpha(n)/n \xrightarrow{p} 0$ under IMGW_0 where $\widehat{N}^\alpha(n) := \sum_{t=0}^n \mathbf{1}_{\{d(Y_t,\xi)\leq n^\alpha\}}$. This can be deduced from the estimate (4.2) by the same arguments which were used in the proofs of Lem. 3.4 and Lem. 4.6 (d).

(d) The height process $(h_s)_{s\geq 0}$ for the walk Y restricted to ξ (i.e. deleting all excursions away from ξ) is a random walk on $\mathbb{Z}_{\leq 0}$ with a ρ-bias in the negative direction, so it clearly suffices to show that for $\mathbf{B}_n' := \max_{0\leq r\leq s\leq n}\{h_s - h_r\}$, we have $\mathbf{B}_n'/\sqrt{n} \xrightarrow{p} 0$. Set $\sigma_0 = 0$ and define for $j > 0$

$$\sigma_j := \inf\{i > \sigma_{j-1} : h_i = h_{\sigma_{j-1}} - 1\}.$$

Now the processes $(\widetilde{h}_s^{(j)} \equiv h_s - h_{\sigma_j})_{\sigma_j \le s \le \sigma_{j+1}}$ are i.i.d., and clearly $\sigma_n \ge n$, so

$$\mathbf{B}'_n \le \max_{0 \le j < n} \left(\max_{\sigma_j \le s \le \sigma_{j+1}} \widetilde{h}_s^{(j)} \right).$$

The probability that $\max_{\sigma_j \le s \le \sigma_{j+1}} \widetilde{h}_s^{(j)} \ge m$ is at most the probability that a random walk on $\mathbb{Z}$ started at 0 with a ρ-bias in the negative direction will reach m before -1, which is $(1 - \rho^{-1})/(\rho^m - \rho^{-1}) \le \rho^{-m}$. Therefore $\mathbb{P}_{\mathcal{T}}(\mathbf{B}'_n \ge m) \le n\rho^{-m}$, so in fact we have the stronger result that $\mathbf{B}'_n/\sqrt{n} \to 0$ a.s. on MGW-a.e. $\mathcal{T}$. □

4.3. **From annealed to quenched CLT.** We now describe how to move from the annealed to the quenched CLT; the proof is motivated by some ideas in [27, §6-7].

For $X \sim \mathsf{RW}_\rho(\mathcal{T})$ for any rooted tree $\mathcal{T}$, let $\mathbb{B}_t^n(X)$ denote the polygonal interpolation of $k/n \mapsto |X_k|/\sqrt{n}$. Let $\mathbf{t}_X^\epsilon$ be the unique increasing bijection

$$\mathbb{Z}_{\ge 0} \to \mathbb{Z}_{\ge 0} \backslash \{t \ge 0 : |X_t| \le n^\epsilon\};$$

and let $\widetilde{X}_s := X_{\mathbf{t}_X^\epsilon(s)}$, so that $\widetilde{X}$ is the time change of X obtained by killing time spent within n^ϵ of o (the total time killed during the first n steps of X is precisely $n - (\mathbf{t}_X^\epsilon)^{-1}(n) = N^\epsilon(n)$).

Lemma 4.7. *Assume* (H1), (H2), *and* (H3^p) *with* $p > 4$. *For* ϵ *sufficiently small, there exists* $\alpha < 1/2$ *such that*

$$\mathsf{MGW}\left(\max_{t \le n} ||X_t| - |\widetilde{X}_t|| \ge n^\alpha \right) \le n^{-c}.$$

Proof. We use the method in the proof of [27, (55)]. By Cor. 4.5, $\mathsf{MGW}(N^\epsilon(n) \ge n^{1/2+2\epsilon}) \le Cn^{-c\epsilon}$, so it suffices to consider $|X_r - X_s|$ for $0 \le r \le s \le n$ with $|r - s| \le n^{1/2+2\epsilon}$. By Lem. 4.6 (a), if $\psi \equiv \psi_n := \min\{t \ge 0 : W_{X_t} > n^{1/4-2\epsilon}\}$, then $\mathsf{MGW}(\psi_n \le n) \le Cn^{-c}$ for ϵ sufficiently small. If H_o denotes the first return time to the root, then $M'_s := S_{X_{s \wedge H_o}}$ is a martingale with differences bounded in absolute value by $n^{1/2+2\epsilon}$ until time ψ, when it has a *positive* jump of size $> n^{1/2+2\epsilon}$. The truncated process

$$\widetilde{M}_s = M'_{s \wedge \psi} - \left(M'_\psi - M'_{\psi-1} \right) \mathbf{1}_{\{\psi \le s\}}$$

is then a supermartingale with all differences $\le n^{1/2+2\epsilon}$ in absolute value. Decomposing into excursions away from the root and applying the Azuma-Hoeffding inequality gives

$$\begin{aligned} \mathsf{MGW}\left(\max_{\substack{0 \le r \le s \le n \\ |r-s| \le n^{1/2+2\epsilon}}} |S_{X_r} - S_{X_s}| \ge 2n^\alpha \right) &\le \mathsf{MGW}(\psi_n \le n) + n \cdot \mathsf{MGW}\left(\max_{\substack{0 \le r \le s \le n \\ |r-s| \le n^{1/2+2\epsilon}}} |\widetilde{M}_r - \widetilde{M}_s| \ge n^\alpha \right) \\ &\le Cn^{-c} + n^3 \exp\left\{ -\frac{n^{2\alpha}}{n^{1/2+2\epsilon} n^{2(1/4-2\epsilon)}} \right\}, \end{aligned}$$

which can be made $\le Cn^{-c}$ by choosing α sufficiently close to 1/2. But by (4.5) the same bound holds replacing S_{X_t} with $\eta|X_t|$, which proves the result. □

Proof (Thm. 1.1). We use similar ideas as in the proof of [27, Thm. 3]. Regard $\mathbb{B}_t^n(X)$ as an element of $C[0, T]$ with the norm

$$d_T(u, u') = \left(\sup_{0 \le t \le T} |u(t) - u'(t)| \right) \wedge 1.$$

By [5, Lem. 4.1], the theorem will follow once we show that for all F Lipschitz on $C[0,T]$ with Lipschitz constant ≤ 1 and for all $1 < b \le 2$,

$$\sum_{k\ge 1} \mathrm{Var}_{\mathsf{MGW}} \left(\mathbb{E}_{\mathcal{T}}[F(\mathbb{B}^{\lfloor b^k \rfloor})] \right) < \infty.$$

To this end, let $\mathcal{T} \sim \mathsf{MGW}$, and let X^1, X^2 be independent realizations of $\mathsf{RW}_\rho(\mathcal{T})$, and write $\mathbb{B}^{n,i} \equiv \mathbb{B}^n(X^i)$ and $\widetilde{\mathbb{B}}^{n,i} \equiv \mathbb{B}^n(\widetilde{X}^i)$. Then

$$\mathrm{Var}_{\mathsf{MGW}}\left(\mathbb{E}_{\mathcal{T}}[F(\mathbb{B}^n)]\right) = \mathbb{E}_{\mathsf{MGW}}\left[F(\mathbb{B}^{n,1})F(\mathbb{B}^{n,2})\right] - \mathbb{E}_{\mathsf{MGW}}\left[F(\mathbb{B}^{n,1})\right]\mathbb{E}_{\mathsf{MGW}}\left[F(\mathbb{B}^{n,2})\right].$$

By subtracting a constant we can take F to be uniformly bounded by 1, so by Lem. 4.7 the above is

$$\le \mathbb{E}_{\mathsf{MGW}}\left[F(\widetilde{\mathbb{B}}^{n,1})F(\widetilde{\mathbb{B}}^{n,2})\right] - \mathbb{E}_{\mathsf{MGW}}\left[F(\widetilde{\mathbb{B}}^{n,1})\right]\mathbb{E}_{\mathsf{MGW}}\left[F(\widetilde{\mathbb{B}}^{n,2})\right] + Cn^{-c}.$$

But if $\mathcal{A}_n$ denotes the event that the paths of $\widetilde{X}^1$ and $\widetilde{X}^2$ up to time n intersect, then $\mathsf{MGW}(\mathcal{A}_n) \le Cn^{-c}$: indeed, using Cor. 4.5 again, it suffices to note that the chance that X^2 hits a given vertex v by time $2n$ (say) with $|v| > n^\epsilon$ is $\le Cn\rho^{-n^\epsilon}$, so the chance that X^2 hits the path of X^1 above n^ϵ is $\le Cn^2\rho^{-n^\epsilon}$. On the event $\mathcal{A}_n^c$ the random variables $F(\widetilde{\mathbb{B}}^{n,1})$ and $F(\widetilde{\mathbb{B}}^{n,2})$ are independent, so finally we obtain

$$\mathrm{Var}_{\mathsf{MGW}}(\mathbb{E}_{\mathcal{T}}[F(\mathbb{B}^n)]) \le Cn^{-c},$$

and the theorem follows. □

Proof of Cor. 1.2. Given $(\mathcal{T}, X) \sim \mathsf{MGW} \otimes \mathsf{RW}_\rho$ we can obtain $(\mathcal{T}, X^{\mathrm{cts}}) \sim \mathsf{MGW} \otimes \mathsf{RW}_\rho^{\mathrm{cts}}$ by taking $(E_i)_{i\ge 1}$ i.i.d. exponential random variables with unit mean independent of X, and setting

$$X_t^{\mathrm{cts}} = X_{\boldsymbol{\theta}(t)}, \quad \boldsymbol{\theta}(t) = \max\left\{ i : \sum_{j=1}^{i} \frac{E_j}{\rho + d_{X_{j-1}}} \le t \right\};$$

similarly we can obtain $((\widehat{\mathcal{T}}, \xi), Y^{\mathrm{cts}}) \sim \mathsf{IMGW}_0 \otimes \mathsf{RW}_\rho^{\mathrm{cts}}$ from $((\widehat{\mathcal{T}}, \xi), Y) \sim \mathsf{IMGW}_0 \otimes \mathsf{RW}_\rho$. Thus a shifted coupling of $(\mathcal{T}, X)$ with $((\widehat{\mathcal{T}}, \xi), Y)$ (as constructed in §4) naturally gives rise to a shifted coupling of $(\mathcal{T}, X^{\mathrm{cts}}) \sim \mathsf{MGW} \otimes \mathsf{RW}_\rho^{\mathrm{cts}}$ with $((\widehat{\mathcal{T}}, \xi), Y^{\mathrm{cts}}) \sim \mathsf{IMGW}_0 \otimes \mathsf{RW}_\rho^{\mathrm{cts}}$ by using sequences $(E_i)_{i\ge 1}$ for X^{cts} and $(\widehat{E}_i)_{i\ge 1}$ for Y^{cts} which are marginally i.i.d. exponential but such that the jump times match during the coupled excursions.

By Thm. 1.3 (b) and the exponential decay of the $\widehat{E}_i$, it holds IMGWR-a.s. that

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\widehat{E}_i}{\rho + d_{Y_i}} \to \mathbb{E}_{\mathsf{IMGWR}}\left[\frac{1}{\rho + d_o}\right] = \frac{1}{2\rho}.$$

It follows from Lem. 4.3 that $\mathbf{1}_{\mathbb{X}^c} n^{-1} \sum_{i=1}^n E_i/(\rho + d_{X_i}) \to 1/(2\rho)$ in MGW-probability, hence MGW-a.s. along subsequences. Along such a subsequence it is easily seen that

$$\max_{0\le t\le 1}\left|\frac{\boldsymbol{\theta}(nt)}{n} - 2\rho t\right| \to 0$$

so it follows from Thm. 1.1 that $(|X_{nt}^{\mathrm{cts}}|/(\sigma\sqrt{n}))$ is uniformly well approximated on $0 \le t \le 1$ by $|B_{\boldsymbol{\theta}(nt)/n}|$, hence by $\sqrt{2\rho}|B_t|$, for B a standard Brownian motion. □

5. Transience-recurrence boundary for RWRE

In this section we prove Thm. 1.5. Our proof is a straightforward adaptation of that of [21, Thm. 1] or [10, Propn. 1.1] once we supply the needed large deviations estimate (Lem. 5.2) on the conductances at the n-th level of the tree, extending the estimates of [21, p. 129] and [10, p. 7] to our setting of Markovian dependency.

Let $\mathcal{D} \equiv \{\gamma : \max_{a,b} \bar{A}^{(\gamma)}(a,b) < \infty\}$, where $\bar{A}^{(\gamma)}$ is as defined in (1.3). Recall that $\bar{\rho}(\gamma)$ denotes the Perron-Frobenius eigenvalue of $\bar{A}^{(\gamma)}$ with $\bar{\rho}(\gamma) = \infty$ for $\gamma \notin \mathcal{D}$. The following lemma collects some basic properties of $\bar{\rho}$.

Lemma 5.1. *Under the hypotheses of Thm. 1.5, $\bar{\rho}$ is lower semicontinuous, log-convex, and differentiable in $\mathcal{D}$.*

Proof. Lower semicontinuity at $\gamma \in \mathcal{D}$ follows from Fatou's lemma, so it remains to consider γ on the boundary of $\mathcal{D}$: we must show that if $\gamma \to \gamma_\infty$ with $\max_{a,b} \bar{A}^{(\gamma)}(a,b) \to \infty$ then $\bar{\rho}(\gamma) \to \infty$. Recall the min-max characterization

$$\bar{\rho}(\gamma) = \max_{\underline{x} \geq 0, \underline{x} \neq 0} \min_{a : x_a \neq 0} \frac{(\bar{A}^{(\gamma)} \underline{x})_a}{x_a}$$

(see e.g. [12, Cor. 8.3.3]), which implies $\bar{\rho}(\gamma) \geq \max_a \bar{A}^{(\gamma)}(a,a)$. Since $\bar{A}^{(0)}$ is positive regular, so is $\bar{A}^{(\gamma)}$ for all $\gamma \in \mathcal{D}$, and by taking matrix powers we may suppose without loss that $\min_{a,b} \bar{A}^{(\gamma)}(a,b) \geq \epsilon$ for all γ in a neighborhood of γ_∞. The min-max characterization applied to the vectors $\underline{x} \geq 0$ with $(x_b/x_a)^2 = \epsilon/\bar{A}^{(\gamma)}(a,b)$ and all other entries equal to zero gives

$$\bar{\rho}(\gamma)^2 \geq \epsilon \max_{a \neq b} A^{(\gamma)}(a,b),$$

which proves lower semicontinuity. The entries of $\bar{A}^{(\gamma)}$ are log-convex in γ by Hölder's inequality, so $\bar{\rho}$ is log-convex by monotonicity and log-convexity of the Perron-Frobenius eigenvalue in the entries of the matrix (see e.g. [12, Cor. 8.1.19] and [6, Exercise 4.34]). For differentiability of $\bar{\rho}$ in $\mathcal{D}$ see [8, p. 75]. □

For $\gamma \in \mathcal{D}$ let $\underline{\bar{e}}^{(\gamma)}$ and $\underline{\bar{g}}^{(\gamma)}$ denote the associated left and right eigenvectors; we use the shorthand

$$\underline{e} \equiv \underline{\bar{e}}^{(0)}, \quad \underline{g} \equiv \underline{\bar{g}}^{(0)}.$$

For $\mathcal{T} \sim \overline{\mathsf{MGW}}$ and $v \in \mathcal{T}$, let

$$C_v \equiv \prod_{o < u \leq v} \alpha_u,$$

the conductance of the edge leading to v. The natural generalization of the martingale introduced in §2.3 is

$$\widehat{Z}_n^{(\gamma)} = \frac{1}{\bar{\rho}(\gamma)^n} \sum_{v \in D_n} \bar{e}_{\chi_v}^{(\gamma)} C_v^\gamma; \tag{5.1}$$

this is a multi-type Mandelbrot's martingale and has been studied in various contexts, for example as the Laplace transform of the branching random walk with increments $\log \alpha_v$ [7, 18]. Using this martingale we can make a change of measure (similar to the one of §2.3) and control the conductances at the n-th level by controlling the conductance of a *random* vertex. Specifically, for each $a \in \mathcal{Q}$ define a size-biased measure $\overline{\mathsf{Q}}_n^a$ on Ω by

$$\frac{d\overline{\mathsf{Q}}_n^a}{d\overline{\mathsf{MGW}}^a} = \frac{\widehat{Z}_n^{(0)}}{e_a}.$$

We then let $\overline{\mathsf{Q}}^a_{n\star}$ denote the measure on pairs $(\mathcal{T}, v_n)$ obtained by letting $\mathcal{T} \sim \overline{\mathsf{Q}}^a_n$ and choosing $v_n \in D_n$ according to weights e_{χ_v}.

Lemma 5.2. *Under the hypotheses of Thm. 1.5, for each $a \in \mathcal{Q}$, under $\overline{\mathsf{Q}}^a_n$ the random variables $n^{-1} \log C_{v_n}$ satisfy a large deviation principle with good rate function $\Lambda^*(x) \equiv \sup_\gamma(\gamma x - \Lambda(\gamma))$, where $\Lambda(\gamma) = \log \bar{\rho}(\gamma) - \log \bar{\rho}(0)$. In particular, for any $0 < z < y$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \overline{\mathsf{Q}}^a_{n\star}(C_{v_n} > z^n) \ge -\sup_{\gamma \ge 0}(\gamma \log y - \Lambda(\gamma)). \tag{5.2}$$

Proof. Let Λ_n denote the cumulant generating function of $n^{-1} \log C_{v_n}$ with respect to $\overline{\mathsf{Q}}^a_{n\star}$, that is,

$$\Lambda_n(\gamma) = \log \mathbb{E}_{\mathsf{Q}^a_{n\star}}[C_{v_n}^{\gamma/n}].$$

Then

$$\begin{aligned} e^{\Lambda_n(n\gamma)} &= \mathbb{E}_{\mathsf{Q}^a_{n\star}}[C^\gamma_{v_n}] = \mathbb{E}_{\mathsf{Q}^a_n}\left[\frac{\sum_{v\in D_n} e_{\chi_v} C^\gamma_v}{\sum_{v\in D_n} e_{\chi_v}}\right] = \frac{1}{e_a \bar{\rho}(0)^n} \mathbb{E}_{\mathsf{MGW}^a}\left[\sum_{v\in D_n} e_{\chi_v} C^\gamma_v\right] \\ &\asymp \frac{\bar{\rho}(\gamma)^n}{\bar{\rho}(0)^n} \mathbb{E}_{\mathsf{MGW}^a}[\widehat{Z}^{(\gamma)}_n] \asymp \frac{\bar{\rho}(\gamma)^n}{\bar{\rho}(0)^n}, \end{aligned}$$

where $\asymp$ indicates equivalence up to constant factors depending only on $\underline{e}$ and $\underline{e}^{(\gamma)}$. Thus

$$\lim_{n\to\infty} \frac{1}{n} \Lambda(n\gamma) = \log \bar{\rho}(\gamma) - \log \bar{\rho}(0) = \Lambda(\gamma).$$

By Lem. 5.1 this is an essentially smooth convex function in the sense of [8, Defn. 2.3.5], so the large deviation principle follows from the Gärtner-Ellis theorem (see [8, Thm. 2.3.6]). In particular, for any $0 < z < y$, [8, (2.3.8)] implies

$$\liminf_{n\to\infty} \frac{1}{n} \log \overline{\mathsf{Q}}^a_{n\star}(C_{v_n} > z^n) \ge -\inf_{x > \log z} \Lambda^*(x) \ge -\Lambda^*(\log y)$$

(making use of [8, Lem. 2.3.9]). The result (5.2) follows immediately if $\log y = \Lambda'(\gamma)$ for some $\gamma \ge 0$. Next, by the assumption that $\Lambda < \infty$ in a neighborhood of 0, $\Lambda^{*\prime}(x) > 0$ for sufficiently large x and so Λ^* attains its global infimum at x_0 with $\Lambda^*(x_0) = -\Lambda(0)$. If $\log y \le \inf_{\gamma\ge 0} \Lambda'(\gamma)$, then $x_0 \ge \log y$ so (5.2) again follows. ☐

Thm. 1.5 now follows by adapting the proof of [21, Thm. 1]:

Proof of Thm. 1.5. Since the bias λ can always be absorbed into the environment variables α_v ($v \in \mathcal{T}$), we may take $\lambda = 1$ from now on, and write $p \equiv p_1 = \min_{0\le\gamma\le 1} \bar{\rho}(\gamma)$.

(a) If $p < 1$ for some $\gamma \in [0, 1]$, making use of the martingale $\widehat{Z}^{(\gamma)}_n$ it is easy to see that the sum of the conductances C_v, $v \in \mathcal{T}$ is finite $\overline{\mathsf{MGW}}$-a.s., hence the random walk is a.s. positive recurrent.

(b) Suppose $p > 1$. We will show there exists $w < 1$ such that

$$\liminf_{n\to\infty} w^n \sum_{v\in D_n} C_v > 0; \tag{5.3}$$

transience then follows from [20, Cor. 4.2]. By the proof on [21, p. 129],

$$p = \max_{0<y\le 1} y \left(\inf_{\gamma\ge 0} y^{-\gamma} \bar{\rho}(\gamma)\right),$$

and we fix $y \in (0, 1]$ achieving this maximum. Then Lem. 5.2 implies that

$$\liminf_{n\to\infty} \frac{1}{n} \log \overline{\mathsf{Q}}^a_{n\star}(C_{v_n} > z^n) \ge -\log[y\bar{\rho}(0)/p] \quad \forall z < y, \forall a \in \mathcal{Q}.$$

Therefore we can choose $z < y$, $j \in \mathbb{N}$, $\epsilon > 0$, and $w < 1$ such that

$$\min_{a\in\mathcal{Q}} \overline{\mathsf{Q}}^a_{j\star}(\{C_{v_j} > z^j\} \cap \{\alpha_w \geq \epsilon \ \forall o < w \leq v_j\}) \geq q > (wz\bar{\rho}(0))^{-j}.$$

Now consider the percolation process described on [21, p. 130]: let $\mathcal{T}'$ be the tree with vertices $\{v \in \mathcal{T} : |v| \equiv 0 \bmod j\}$, with an edge $v \to w$ if and only if $|w| = |v| + k$ in $\mathcal{T}$. Form a random subgraph of $\mathcal{T}'$ by keeping the edge $v \to w$ if and only if

$$\prod_{v<u\leq w} \alpha_u > z^j \quad \text{and} \quad \min_{v<u\leq w} \alpha_u \geq \epsilon,$$

in which case we write $v \rightsquigarrow w$. Every open cluster of this percolation process is a multi-type Galton-Watson tree with mean offpsring numbers

$$A^*(a,b) = \mathbb{E}_{\overline{\mathsf{MGW}}^a}\left[\sum_{v\in D_j} \mathbf{1}_{\{\chi_v=b\}}\mathbf{1}_{\{o\rightsquigarrow v\}}\right] = \frac{e_a\bar{\rho}(0)^j}{e_b}\mathbb{E}_{\overline{\mathsf{Q}}^a_{j\star}}[\mathbf{1}_{\{\chi_{v_j}=b\}}\mathbf{1}_{\{o\rightsquigarrow v_j\}}].$$

Then

$$\sum_b A^*(a,b)e_b \geq e_a\bar{\rho}(0)^j q > e_a(wz)^{-j},$$

so A^* has Perron-Frobenius eigenvalue larger than $(wz)^{-j}$ and there exists a.s. an open cluster which is an infinite tree of branching number larger than $(wz)^{-j}$. It follows that for some random $X > 0$ (depending on ϵ and the location of the infinite cluster)

$$\liminf_{n\to\infty} w^{nk} \sum_{v\in D_{nk}} C_v \geq X \liminf_{n\to\infty} \sum_{v\in D_{nk}} (wz)^{nk} > 0,$$

which concludes the proof of (5.3). □

6. Appendix: general properties of MGW trees

We begin with some definitions. For $a \in \mathcal{Q}$ and $\underline{s} \in [0,1]^{\mathcal{Q}}$, let

$$F(\underline{s}) := (F^a(\underline{s}))_{a\in\mathcal{Q}}, \quad F^a(\underline{s}) := \mathbb{E}_{\mathbf{q}^a}\left[\prod_{b\in\mathcal{Q}} s_b^{x_b}\right];$$

we refer to F as the generating function of the MGW tree. If $F^{(n)}$ denotes the n-fold composition of F, then for all $a \in \mathcal{Q}$

$$\mathbb{E}_{\mathsf{MGW}^a}\left[\prod_{b\in\mathcal{Q}} s_b^{Z_n(b)}\right] = (F^{(n)}(\underline{s}))_a, \quad \mathsf{MGW}^a(|\underline{Z}_n| = 0) = (F^{(n)}(\underline{0}))_a.$$

Recall that $\mathbb{X}$ denotes the event of extinction, and let $q_a := \mathsf{MGW}^a(\mathbb{X})$. The case of $\mathsf{MGW}(\mathbb{X}) > 0$ can often be reduced to the simpler case of an a.s. infinite tree without leaves by the following transformation which is discussed in [1, §I.12] for the single-type case. For $\mathcal{T} \sim \mathsf{MGW}$, consider the subtree $\mathcal{T}^\infty$ consisting of those vertices v of *infinite descent*, i.e. with $|\mathcal{T}^{(v)}| = \infty$. Conditioned on $\mathbb{X}^c$, $\mathcal{T}^\infty$ is an a.s. infinite tree without leaves, following a transformation of the original MGW given by the generating function $\widetilde{F}(\underline{s}) = (\widetilde{F}^a(\underline{s}))_{a\in\mathcal{Q}}$, where

$$\widetilde{F}^a(\underline{s}) = \frac{1}{1-q_a}\sum_{\underline{x}} \mathbf{q}^a(\underline{x}) \prod_{b\in\mathcal{Q}} \sum_{y_b\leq x_b} \binom{x_b}{y_b} q_b^{x_b-y_b}(1-q_b)^{y_b} s_b^{y_b} = \frac{F_a((q_b + (1-q_b)s_b)_{b\in\mathcal{Q}})}{1-q_a}.$$

The transformed law has mean matrix

$$\widetilde{A} = D^{-1}AD, \quad D = \mathrm{diag}((1-q_a)_{a\in\mathcal{Q}}),$$

so in particular it has the same Perron-Frobenius eigenvalue as A. Finally, it is clear that if the original law satisfies (H3^p) then so does the transformed law.

6.1. **Zero-one law of perpetuating properties.** The following zero-one law of MGW trees was used in the proof of Thm. 1.3, and is similar in spirit to [24, Propn. 5.6]. Recall that Ω denotes the space of rooted trees with Borel σ-algebra $\mathcal{B}_\Omega$; we refer to sets $P \in \mathcal{B}_\Omega$ as "tree properties."

Definition 6.1. We say that a tree property $P \in \mathcal{B}_\Omega$ is *perpetuating* if $\mathcal{T} \in P$ if and only if $\mathcal{T}^{(v)} \in P$ for all $v \in \partial o$.

The following proposition does not require (H1) since no moment assumptions are needed; instead we will say the law MGW is *irreducible* if there is a positive MGW^a-probability of having a vertex of type b in the tree for all $a, b \in \mathcal{Q}$.

Proposition 6.2. *If* MGW *is irreducible and* $\mathsf{MGW}(|\underline{Z}_1| > 1) > 0$*, then all perpetuating properties are* MGW*-trivial.*

Proof. Let $\underline{Z}_n^\infty$ count the vertices at level n of infinite descent. Then

$$\mathsf{MGW}^a(P|\mathbb{X}^c) \le \sum_{\underline{x}} \mathsf{MGW}^a(\underline{Z}_n^\infty = \underline{x}|\mathbb{X}^c) \prod_{b\in\mathcal{Q}} \mathsf{MGW}^b(P|\mathbb{X}^c)^{x_b}.$$

Note that $\mathsf{MGW}^a(|\underline{Z}_n^\infty| = 0|\mathbb{X}^c) = 0$, so by considering a for which $\mathsf{MGW}^a(P|\mathbb{X}^c)$ is maximal, it follows from irreducbility that $\mathsf{MGW}^a(P|\mathbb{X}^c)$ is constant in a. Then, since $\mathsf{MGW}(|\underline{Z}_n| > 1|\mathbb{X}^c) > 0$, we have that $\mathsf{MGW}(P|\mathbb{X}^c) \in \{0,1\}$.

If $\mathsf{MGW}(\mathbb{X}) = 0$ we are done so assume $\mathsf{MGW}(\mathbb{X}) > 0$, and suppose without loss that the law of χ_o is the stationary distribution for the Markov chain with transition probabilities $p(a,b) = \sum_{\underline{x}} \mathbf{q}^a(\underline{x}) x_b/|x|$; this is chosen for the convenient property that if $\mathcal{T} \sim \mathsf{MGW}$ and $\mathcal{T}' = \mathcal{T}^{(v)}$ for v chosen uniformly at random from ∂o, then $\mathcal{T} \sim \mathsf{MGW}$. Then

$$\mathsf{MGW}(P|\mathbb{X}^c) \ge \mathsf{MGW}(\mathcal{T}' \in \mathbb{X}|\mathcal{T} \in \mathbb{X}^c)\mathsf{MGW}(P|\mathbb{X})$$

and $\mathsf{MGW}(\mathcal{T}' \in \mathbb{X}|\mathcal{T} \in \mathbb{X}^c) > 0$ by assumption, so if $\mathsf{MGW}(P|\mathbb{X}^c) = 0$ then $\mathsf{MGW}(P = 0)$. Finally notice that P is perpetuating if and only if P^c is perpetuating, so $\mathsf{MGW}(P|\mathbb{X}^c) = 1$ implies $\mathsf{MGW}(P = 1)$ which completes the proof. ☐

6.2. **Positive moments of the normalized population size.** In this section we prove that moment conditions on the MGW offspring distribution translate direction to moment conditions on the normalized population size of the entire tree.

Lemma 6.3. *If* (H1)*,* (H2) *and* (H3^n) *hold with* $n \in \mathbb{Z}_{\ge 2}$*, then* $\mathbb{E}_{\mathsf{MGW}}[W_o^n] < \infty$*.*

Proof. We follow the proof of [4, Thm. 0]. For each $a \in \mathcal{Q}$ let

$$\phi^a(s) := \mathbb{E}_{\mathsf{MGW}^a}[e^{-sW_o}], \quad \Phi(s) := (\phi^a(s))_{s\in\mathcal{Q}}, \quad \phi(s) := \mathbb{E}_{\mathsf{MGW}}[e^{-sW_o}] = \langle \underline{g}, \Phi(s)\rangle.$$

We have the functional equation $\Phi(s) = F(\Phi(\gamma s))$ where $\gamma \equiv 1/\rho$. By the assumption (H3^n), if $\underline{v} \in (\mathbb{R}_{\ge 0})^{\mathcal{Q}}$ with $|\underline{v}| = 1$,

$$\begin{aligned} f_n(t;\underline{v}) &:= (-1)^{n+1} \left\{ \mathbb{E}_{\mathsf{MGW}}[e^{-t\langle \underline{v}, \underline{Z}_1\rangle}] - \sum_{r=0}^n \mathbb{E}_{\mathsf{MGW}}[\langle \underline{v}, \underline{Z}_1\rangle^r] \frac{(-t)^r}{r!} \right\} \\ &= \frac{t^n}{n!} \mathbb{E}_{\mathsf{MGW}} \left[\langle \underline{v}, \underline{Z}_1\rangle^n (1 - e^{-\zeta(t;v)\langle \underline{v}, \underline{Z}_1\rangle}) \right] \end{aligned} \tag{6.1}$$

where $0 \le \zeta(t; v) \le t$. From this it is clear that $f_n(t; \underline{v}) = o(t^n)$ uniformly in $\underline{v}$ as $t \to 0$. We have $\phi_a(s) = 1 - e_a s + o(s)$, and the result will follow once we show the existence of constants $m_2, \dots, m_n$ such that

$$\phi_n(s) \equiv (-1)^{n+1} \left\{ \phi(s) - 1 + \langle \underline{g}, \underline{e} \rangle s - \sum_{r=2}^{n} m_r \frac{(-s)^r}{r!} \right\} = o(s^n) \quad (s \to 0).$$

Suppose inductively the existence of $m_1, \dots, m_k$ with $k < n$: then, writing $\Phi(\gamma s) \equiv e^{-\underline{t}} \equiv e^{-t\underline{v}}$ with $|\underline{v}| = 1$ (note $\underline{v}$ depends on t), we have

$$tv_a = -\log \phi^a(\gamma s) = e_a \gamma s + \sum_{r=2}^{k} c_r^a s^r + o(s^k) \quad (s \to 0). \tag{6.2}$$

Let $\psi(s) := [\phi(s) - 1 + \langle \underline{g}, \underline{e} \rangle s]/s$. By the functional equation and (6.1),

$$\begin{aligned} \psi(s) - \psi(\gamma s) &= \frac{1}{s} \left[\mathbb{E}_{\mathsf{MGW}}[e^{-t\langle \underline{v}, \underline{Z}_1 \rangle}] - 1 + \rho(1 - \langle \underline{g}, e^{-t\underline{v}} \rangle) \right] \\ &= \frac{1}{s} \left[\rho(1 - \langle \underline{g}, e^{-t\underline{v}} \rangle - t\langle \underline{g}, \underline{v} \rangle) + \sum_{r=2}^{k+1} \mathbb{E}_{\mathsf{MGW}}[\langle \underline{v}, \underline{Z}_1 \rangle^r] \frac{(-t)^r}{r!} + o(t^{k+1}) \right], \end{aligned}$$

and it then follows from (6.2) that

$$\psi(s) - \psi(\gamma s) = \sum_{r=1}^{k} b_r s^r + o(s^k).$$

Since $\lim_{s \downarrow 0} \psi(s) = 0$,

$$\psi(s) = \sum_{j \ge 0} [\psi(\gamma^j s) - \psi(\gamma^{j+1} s)] = \sum_{r=1}^{k} \frac{b_r}{1 - \gamma^r} s^r + o(s^k)$$

and by the definition of ψ this verifies the inductive hypothesis. □

Corollary 6.4. *Under the notation and hypotheses of Lem. 6.3,*

$$(-1)^{n+1} [\phi_n(s) - \rho \phi_n(\gamma s)] = f_n(t; \underline{v}_t) + O(t^{n+1})$$

where $\underline{v}_t$ is defined by $\Phi(\gamma s) \equiv e^{-t\underline{v}_t}$.

Proof. By the definition of ϕ_n and the functional equation,

$$(-1)^{n+1} [\phi_n(s) - \rho \phi_n(\gamma s)] = \mathbb{E}_{\mathsf{MGW}}[e^{-t\langle \underline{v}_t, \underline{Z}_1 \rangle}] - 1 + \rho - \rho \langle \underline{g}, e^{-t\underline{v}_t} \rangle - \sum_{r=2}^{n} m_r (1 - \gamma^{r-1}) \frac{(-s)^r}{r!}$$

Since s appears only in powers ≥ 2 on the right-hand side, it follows from (6.2) that

$$(-1)^{n+1} [\phi_n(s) - \rho \phi_n(\gamma s)] = \left\{ \mathbb{E}_{\mathsf{MGW}}[e^{-t\langle \underline{v}_t, \underline{Z}_1 \rangle}] - 1 - \sum_{r=2}^{m} b_r \frac{(-t)^r}{r!} \right\} + O(t^{n+1}).$$

Since the left-hand side is $o(s^n)$, the expression in braces is *a fortiori* equal to $f_n(t; \underline{v}_t)$. □

Proof of Propn. 2.2. By Lem. 6.3 we may take $\alpha = n+\beta$ for $\beta \in (0,1)$. By (6.1) and Fubini's theorem,

$$\int_0^1 f_n(t;\underline{v}_t)t^{-(1+\alpha)}\,dt \le \frac{1}{n!}\int_0^1 t^{-(1+\beta)}\mathbb{E}[|\underline{Z}_1|^n(1-e^{-t|\underline{Z}_1|})]\,dt$$
$$\le \frac{\mathbb{E}_{\mathsf{MGW}}[|\underline{Z}_1|^\alpha]}{n!}\int_0^\infty \frac{1-e^{-t}}{t^{1+\beta}}\,dt < \infty. \tag{6.3}$$

By [4, Thm. B] the result will follow if we can show

$$\int_0^1 \phi_n(s)s^{-(1+\alpha)}\,ds < \infty.$$

This follows from the proof of [4, Propn. 5] replacing [4, (3.13)] with (6.3) and [4, (3.9)] with Cor. 6.4. □

6.3. **Harmonic moments and conductance estimates.** In this section we prove the existence of harmonic moments for the normalized population size, extending part of [25, Thm. 1] to the multi-type setting (using a similar proof). Using this result we adapt the methods of [26, Lem. 2.2] to prove the conductance estimates used in the proofs of Cor. 4.5 and Lem. 4.6.

Lemma 6.5. *Assume* (H1) *and* (H2)*. There exists some $r>0$ for which*

$$\mathbb{E}_{\mathsf{MGW}}[W_o^{-r}|\mathbb{X}^c] \le \limsup_{n\to\infty}\mathbb{E}_{\mathsf{MGW}}[\widehat{Z}_n^{-r}|\mathbb{X}^c] < \infty.$$

Proof. By the above discussion we may reduce to the case $\mathbf{q}^a(|\underline{x}|=0)=0$ for all $a\in\mathcal{Q}$. Expanding $F^{(n)}(\underline{s})$ as a power series we find

$$(F^{(n)}(\underline{s}))_a \le \sum_{b\in\mathcal{Q}}\mathsf{MGW}^a(\underline{Z}_n=(b)) + \mathsf{MGW}^a(|\underline{Z}_n|>1)\|\underline{s}\|_\infty^2.$$

It follows from (H1) that for all $a\in\mathcal{Q}$ there exists n_0 such that $\mathsf{MGW}^a(|\underline{Z}_{n_0}|>1)>0$ for all $a\in\mathcal{Q}$. Then, given $s_0<1$, there exists $\gamma<1$ such that

$$(F^{(n)}(\underline{s}) \le C\gamma^n\|\underline{s}\|_\infty \quad \forall\|\underline{s}\|_\infty \le s_0, a\in\mathcal{Q}.$$

Write

$$f_n(u) = \mathbb{E}_{\mathsf{MGW}}[e^{-u\langle\underline{Z}_n,\underline{e}\rangle}] = \langle\underline{g}, F^{(n)}((e^{-ue_b})_{b\in\mathcal{Q}})\rangle.$$

By Fubini's theorem,

$$\mathbb{E}_{\mathsf{MGW}}[\widehat{Z}_n^{-r}] = \frac{\rho^{nr}}{\Gamma(r)}\int_0^\infty f_n(u)u^{r-1}\,du.$$

We break up the integral into three parts. First, by a change of variables,

$$\rho^{nr}\int_0^{\rho^{-n}} f_n(u)u^{r-1}\,du = \int_0^1 \mathbb{E}_{\mathsf{MGW}}[e^{-u\widehat{Z}_n}]u^{r-1}\,du \le \frac{1}{r}.$$

Next, for $u\ge 1$ we have $\|(e^{-ue_b})_{b\in\mathcal{Q}}\|_\infty \le e^{-e_{\min}}$, so for some $\gamma<1$

$$\rho^{nr}\int_1^\infty f_n(u)u^{r-1}\,du \le C(\gamma\rho^r)^n\int_1^\infty e^{-e_{\min}u}u^{r-1}\,du.$$

Choosing $r > 0$ small enough so that $\gamma\rho^r < 1$, this tends to zero as $n \to \infty$. It remains to consider the integral over $[\rho^{-n}, 1]$,

$$I := \rho^{nr} \int_{\rho^{-n}}^{1} f_n(u) u^{r-1} \, du = \rho^{nr} \sum_{i=1}^{n} \int_{1/\rho^i}^{1/\rho^{i-1}} f_n(u) u^{r-1} \, du$$
$$= \sum_{i=1}^{n} \rho^{r(n-i)} \int_1^{\rho} \mathbb{E}_{\mathsf{MGW}}[e^{-u\langle \underline{Z}_n, \underline{e}\rangle/\rho^i}] u^{r-1} \, du.$$

By conditioning on the first $n-i$ levels of the tree,

$$\mathbb{E}_{\mathsf{MGW}}[e^{-u\langle \underline{Z}_n, \underline{e}\rangle/\rho^i}] = \mathbb{E}_{\mathsf{MGW}}\left[\prod_{a\in\mathcal{Q}} \Phi_i^a(u)^{Z_{n-i}(a)}\right] = \langle \underline{g}, F^{(n-i)}[(\Phi_i^a(u))_{a\in\mathcal{Q}}]\rangle,$$

where $\Phi_i^a(u) := \mathbb{E}_{\mathsf{MGW}^a}[e^{-u\widehat{Z}_i}]$. But

$$\sup_{i\ge1} \sup_{u\ge1} \Phi_i^a(u) = \sup_{i\ge1} \Phi_i^a(1) < 1,$$

since $\Phi_i^a(1) < 1$ for all a, i and $\Phi_i^a(1) \to \mathbb{E}_{\mathsf{MGW}^a}[e^{-uW_o}]$ which is less than 1 by the Kesten-Stigum theorem as noted in §2.3 (using (H2)). Thus

$$I \le C \sum_{i=1}^{n} (\gamma\rho^r)^{n-i} \int_1^{\rho} u^{r-1} \, du$$

which is bounded in n for small enough r. Putting the estimates together the result follows. □

Proof of Propn. 4.4. (a) Recall that a *unit flow* is a non-negative function U on the vertices of $\mathcal{T}$ such that for all $v \in \mathcal{T}$, $U(v) = \sum_{w\in\partial^+ v} U(w)$. For $v \in D_\ell$ define

$$U(v) = \frac{W_v}{\sum_{u\in D_\ell} W_u} = \frac{W_v}{\rho^\ell W_o};$$

it is easily seen that U is a well-defined unit flow on $\mathbb{X}^c$. It gives positive flow only to vertices of infinite descent, so by the discussion at the beginning of this section we may reduce to the case $\mathbf{q}^a(|\underline{x}| = 0) = 0$ for all $a \in \mathcal{Q}$. By Thomson's principle [24, §2.4]

$$\mathcal{C}_{o,k}^{-1} \le \sum_{\ell=1}^{k} \rho^\ell \sum_{v\in D_\ell} U(v)^2 = \frac{1}{W_o^2} \sum_{\ell=1}^{k} \frac{1}{\rho^\ell} \sum_{v\in D_\ell} W_v^2.$$

By Hölder's inequality,

$$\mathbb{E}_{\mathsf{MGW}}\left[\mathcal{C}_{o,k}^{-r}\right] = \mathbb{E}_{\mathsf{MGW}}\left[\sum_{\ell=1}^{k} \frac{1}{\rho^\ell} \sum_{v\in D_\ell} W_v^2\right]^r \mathbb{E}[W_o^{-2r/(1-r)}]^{1-r} \le Ck^r$$

for r sufficiently small, using Lem. 6.5 and $p \ge 2$. It follows from Markov's inequality that $\mathsf{MGW}(\mathcal{C}_{o,k}^{-1} \ge k^{1+\epsilon}) \le Ck^{-r\epsilon}$.

(b) We claim there exist $0 < c, C < \infty$ deterministic such that

$$\mathsf{MGW}\left(\frac{1}{|D_k|} \sum_{v\in D_k} W_v^2 \ge C\right) \le \rho^{-ck} \tag{6.4}$$

Assuming the claim, we have

$$\rho^k \sum_{v\in D_k} U(v)^2 = \frac{1}{\rho^k}\frac{\sum_{v\in D_k} W_v^2}{W_o^2} \le \frac{1}{W_o^2}\frac{1}{e_{\min}}\frac{|D_k|_{\underline{e}}}{\rho^k}\frac{1}{|D_k|}\sum_{v\in D_k} W_v^2,$$

so

$$\limsup_{k\to\infty} \rho^k \sum_{v\in D_k} U(v)^2 \le \frac{C}{e_{\min} W_o} < \infty,$$

and $\mathcal{C}_{o,k}^{-1} \le C_{\mathcal{T}} k$ follows from Thomson's principle.

It remains to prove (6.4). Since $p > 2$, the W_v have finite moments of order $1+r$ for some $0 < r \le 1$, so by (3.4) and Lem. 6.5

$$\mathsf{MGW}\left(\left|\frac{1}{|D_k|}\sum_{v\in D_k}(W_v^2 - \mathbb{E}_{\mathsf{MGW}^{\chi_v}}[W_v^2])\right| \ge \epsilon\right) \le C(\epsilon,r)\mathbb{E}_{\mathsf{MGW}}\left[|D_k|^{-r}\right] \le C(\epsilon,r)\rho^{-rk},$$

for r sufficiently small. But the quantity

$$\frac{1}{|D_k|}\sum_{v\in D_k}\mathbb{E}_{\mathsf{MGW}^{\chi_v}}[W_v^2]$$

is bounded uniformly in k by a deterministic constant, so (6.4) follows by Borel-Cantelli. □

References

[1] K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].

[2] G. Ben Arous, A. Fribergh, N. Gantert, and A. Hammond. Biased random walks on Galton-Watson trees with leaves. To appear.

[3] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[4] N. H. Bingham and R. A. Doney. Asymptotic properties of supercritical branching processes. I. The Galton-Watson process. *Advances in Appl. Probability*, 6:711–731, 1974.

[5] E. Bolthausen and A.-S. Sznitman. On the static and dynamic points of view for certain random walks in random environment. *Methods Appl. Anal.*, 9(3):345–375, 2002. Special issue dedicated to Daniel W. Stroock and Srinivasa S. R. Varadhan on the occasion of their 60th birthday.

[6] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.

[7] M. Bramson, P. Ney, and J. Tao. The population composition of a multitype branching random walk. *Ann. Appl. Probab.*, 2(3):575–596, 1992.

[8] A. Dembo and O. Zeitouni. *Large deviations: techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.

[9] R. Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, 3rd edition, 2004.

[10] G. Faraud. A central limit theorem for random walk in a random environment on marked Galton-Watson trees. *Electron. J. Probab.*, 16:no. 6, 174–215, 2011.

[11] T. E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)].

[12] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[13] Y. Hu and Z. Shi. A subdiffusive behaviour of recurrent random walk in random environment on a regular tree. *Probab. Theory Related Fields*, 138(3-4):521–549, 2007.

[14] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[15] J. G. Kemeny, J. L. Snell, and A. W. Knapp. *Denumerable Markov chains*. Springer-Verlag, New York, second edition, 1976.

[16] H. Kesten and B. P. Stigum. A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.*, 37:1211–1223, 1966.

[17] T. Kurtz, R. Lyons, R. Pemantle, and Y. Peres. A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 181–185. Springer, New York, 1997.

[18] A. E. Kyprianou and A. Rahimzadeh Sani. Martingale convergence and the functional equation in the multi-type branching random walk. *Bernoulli*, 7(4):593–604, 2001.

[19] Q. Liu. On generalized multiplicative cascades. *Stochastic Process. Appl.*, 86(2):263–286, 2000.

[20] R. Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18(3):931–958, 1990.

[21] R. Lyons and R. Pemantle. Random walk in a random environment and first-passage percolation on trees. *Ann. Probab.*, 20(1):125–136, 1992.

[22] R. Lyons, R. Pemantle, and Y. Peres. Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems*, 15(3):593–619, 1995.

[23] R. Lyons, R. Pemantle, and Y. Peres. Biased random walks on Galton-Watson trees. *Probab. Theory Related Fields*, 106(2):249–264, 1996.

[24] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Preprint, 2009.

[25] P. E. Ney and A. N. Vidyashankar. Harmonic moments and large deviation rates for supercritical branching processes. *Ann. Appl. Probab.*, 13(2):475–489, 2003.

[26] R. Pemantle and Y. Peres. Galton-Watson trees with the same mean have the same polar sets. *Ann. Probab.*, 23(3):1102–1124, 1995.

[27] Y. Peres and O. Zeitouni. A central limit theorem for biased random walks on Galton-Watson trees. *Probab. Theory Related Fields*, 140(3-4):595–629, 2008.

[28] V. V. Petrov. *Sums of independent random variables*. Springer-Verlag, New York, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

[29] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[30] O. Zeitouni. Random walks in random environment. In *Lectures on probability theory and statistics*, volume 1837 of *Lecture Notes in Math.*, pages 189–312. Springer, Berlin, 2004.