On exponentially height-penalized random trees

Louigi Addario-Berry[1], Benoît Corsini[2], Neeladri Maitra[3], and Meltem Ünel[4]

[1]McGill University. Email: louigi.addario@mcgill.ca
[2]National University of Singapore. Email: bcorsini@nus.edu.sg
[3]University of Illinois, Urbana-Champaign. Email: nmaitra@illinois.edu
[4]Université Sorbonne Paris-Nord. Email: unel@lipn.univ-paris13.fr

December 19, 2025

Abstract

Given $n \in \mathbb{N}$ and $\mu \in \mathbb{R}$, a μ-*height-biased tree of size* n is a random plane tree $\mathbf{T}_n$ with n vertices with law given by $\mathbb{P}(\mathbf{T}_n = t) \propto e^{-\mu h(t)}$, where t ranges over fixed plane trees with n vertices, and $h(t)$ is the height of t.

Fix a sequence $(\mu_n)_{n \geq 1}$ of real numbers, and for $n \geq 1$ let $\mathbf{T}_n$ be a μ-height-biased tree of size n. In [19], the authors describe the asymptotic behaviour of $h(\mathbf{T}_n)$ when $\mu_n \equiv \mu \in \mathbb{R}$ is fixed. In this work, we extend their results to arbitrary sequences of positive parameters depending on n. Most notably, we show that such a tree behaves like a height-biased Continuum Random Tree (CRT) when μ_n is of order $1/\sqrt{n}$; that its height is asymptotically $(2\pi^2 n/\mu_n)^{1/3}$ when μ_n is of larger order than $1/\sqrt{n}$ and of smaller order than n; and that its height converges to a fixed constant when μ_n is of order at least n, with some random jumps under specific conditions on μ_n.

We additionally prove various results on second order behaviours, and large deviation principles for the height, for different regimes of μ_n. Finally, we describe new statistics of these trees, covering their widths, their root degrees, and the local structure around their roots.

Contents

1 Introduction

Trees have been at the heart of active research for many decades in combinatorics, probability, and theoretical computer science. Driven in part by the mathematical analysis of algorithms, asymptotic methods from combinatorics have been employed for the enumeration of different classes of trees. These results, being important in their own right, have also been essential in defining various probability measures on the different tree ensembles, hence for the probabilistic study of random tree models. A large class of random tree models is given by the so-called *simply generated trees*, introduced by Meir and Moon in [25]. These models offer a *local* picture by construction, in the sense that the weight of a tree is proportional to the weight factor associated to each vertex depending on the degree of the latter. A well studied sub-class of such trees is given by the genealogical trees of the Bienaymé processes (also known as Galton–Watson processes). The local and scaling limits coming from different variants of this model have been studied extensively and correspond to large universality classes [1, 20]. The uniform random plane tree model, which undergirds the models studied in this work, turns out to be an instance of the Bienaymé model with the offspring distribution being geometric of parameter $\frac{1}{2}$, and conditioned on the number of vertices.

In [19], the authors studied a random tree model with *non-local* weight factors where the height of the tree – a global statistic – is the argument of the weight function: for $n \in \mathbb{N}$ and $\mu \in \mathbb{R}$, the random plane tree $\mathbf{T}_n$ with n vertices with law given by $\mathbb{P}(\mathbf{T}_n = T_n) \propto e^{-\mu h(T_n)}$, where T_n ranges over fixed plane trees with n vertices, and $h(T_n)$ is the height of T_n. This model exhibits a phase transition at $\mu = 0$ where the local limit passes from a single-spine phase to a multi-spine one. The multi-spine tree, called the Poisson tree, obtained when $\mu > 0$, does not fall into a well-known universality class where we find the local limits of large Bienaymé trees. We note that the Poisson tree can also be obtained by conditioning the uniform plane trees on their width growing as a prescribed sequence with a certain growth-rate as a function of the size [2]. Hence, along with the interesting properties of this limiting object, biasing by height turns out to be a rather "natural" way to obtain trees with rather unusual properties.

In this work, we extend the results of [19] to the case where $\mu = \mu_n$ is a positive sequence that can depend on n. Before presenting a wide range of results on different properties of these *μ-height-biased trees*, let us first set some basic terminology regarding trees and introduce the model formally.

Notations and conventions for trees.   All trees arising in this paper are finite. We consider *rooted plane trees* which we also simply refer to as *trees* throughout the paper. Thus, the trees that we consider have a distinguished vertex which we call the *root* of the tree, and for any vertex, there is a left-to-right ordering of its children. We write $\mathbb{t}_n$ for the set of rooted plane trees with n nodes. It is well-known (see e.g., [14, Chapter 1]) that $|\mathbb{t}_n| = C_{n-1}$ where C_n is the n-th Catalan number:

$$C_n = \frac{1}{n+1}\binom{2n}{n}. \tag{1.1}$$

For a tree T, we slightly abuse notation and also use T to denote its vertex set. For any tree T, the *height* or *depth* of a vertex $v \in T$ is its distance from the root, i.e., the number of edges on the unique path connecting v with the root; we denote this quantity by $h(v)$. Note that by definition the root has height zero. We also define its *height* $h(T)$ of the tree T to be the largest depth of any of its nodes, i.e.,

$$h(T) = \max\left\{ h(v) : v \in T \right\}.$$

For a tree T, and non-negative integer k, the *k'th generation* of T is the set of vertices at distance exactly k from the root of T. We write $\mathcal{Z}_k = \mathcal{Z}_k(T)$ for the size of the k'th generation, and $w(T)$ for the *width* of T, which is defined to be the size of the largest generation, i.e.,

$$w(T) = \max\left\{ \mathcal{Z}_k : k \geq 0 \right\}. \tag{1.2}$$

Model definition. The model of random trees introduced in [19] is defined as follows. For any $n \in \mathbb{N}$ and $\mu \in \mathbb{R}$, let $\mathbf{T}_n^\mu$ be the random element of $\mathbb{t}_n$ with distribution given by

$$\mathbb{P}\left(\mathbf{T}_n^\mu = T\right) = \frac{e^{-\mu h(T)}}{Z_n^\mu}, \tag{1.3}$$

where

$$Z_n^\mu = \sum_{T \in \mathbb{t}_n} e^{-\mu h(T)} \tag{1.4}$$

is a normalizing constant, referred to as the *partition function*. We refer to $\mathbf{T}_n^\mu$ as a *μ-height-biased tree of size n*. Throughout the paper, by $\varnothing$ we denote the root of the random tree $\mathbf{T}_n$. When $\mu \equiv 0$ in (1.3), we obtain the uniform measure, and we denote the corresponding uniform random tree of size n by $\mathbb{T}_n$ for the rest of this paper.

In [19], the authors studied local properties of this model when μ is a fixed real, not depending on n, and observed three different asymptotic behaviours as $n \to \infty$, depending on whether μ was equal to 0, positive, or negative. The case $\mu = 0$ is the usual uniform plane tree whose asymptotic local limit is Kesten's tree associated to the geometric offspring distribution of parameter $1/2$, see [22], which we refer to as simply Kesten's tree (KT) for the rest of the paper. In the present work, our goal is to extend these results to consider the case where $\mu = \mu_n > 0$ can depend on n. Due to the large range of new behaviours arising from μ_n to depend on n, in this work we focus on the case that $\mu_n > 0$, and leave the case $\mu_n < 0$ for future work.) For the rest of the paper, we usually remove the subscript n from μ and the superscript μ from $\mathbf{T}_n^\mu$ and Z_n^μ, simply writing $\mathbf{T}_n$ and Z_n, except when we need to make the distinction between two distributions with different μ-parameters. After introducing some standard notation, in the sections that follow, we state our main results.

Notation. We write $x_n \asymp y_n$, $x_n \ll y_n$, $x_n \gg y_n$, $x_n = O(y_n)$, and $x_n = \Omega(y_n)$, respectively, to indicate that x_n/y_n is bounded away from 0 and ∞, that x_n/y_n converges to 0, that x_n/y_n diverges to ∞, that x_n/y_n is bounded, and that x_n/y_n is bounded away from 0. Sometimes it is convenient for us to write $x_n = o(y_n)$ for $x_n \ll y_n$. We also write $x_n \sim y_n$ to say that x_n/y_n converges to 1. For x_n and y_n possibly random sequences, we write $x_n \asymp_{\mathbb{P}} y_n$ to indicate that the ratio x_n/y_n is *bounded in probability* from above and below, i.e., for any $\epsilon > 0$, one can obtain a $K = K(\epsilon)$ so that $\mathbb{P}\left(K^{-1} \leq |x_n/y_n| \leq K\right) \geq 1 - \epsilon$. For any function $f : S \to \mathbb{R}$ for some set S, we denote $\|f\| := \sup_{s \in S} |f(s)|$ the supremum value of f, possibly infinite. By $\mathrm{Geo}(p)$ we denote the law of a discrete random variable having the geometric distribution with success parameter p.

1.1 Global properties in the Brownian regime

In this section, we state results on *global* properties of the tree $\mathbf{T}_n$ when $\mu = O(1/\sqrt{n})$. Note that in this regime, we are allowing $\mu < 0$. This is the only section of the paper where we allow μ to take negative values. We term this the *Brownian regime* due to the strong relation between our model in this regime and Brownian motions, in particular through the *Continuum Random Tree (CRT)*.

Our first result, on the scaling limit in the Brownian regime, establishes the convergence of the rescaled tree, seen as a metric space, toward a height-biased CRT. For more precise definitions about the objects appearing below, and the topology appearing in the statement of Theorem 1.1, we refer the reader to Section 2.

Let $\mathbf{e} = (\mathbf{e}(t) : t \in [0,1])$ be the standard Brownian excursion on $[0,1]$. Consider the random function $\mathbf{e}_\alpha = (\mathbf{e}_\alpha(t) : t \in [0,1])$ whose law ρ_α has Radon-Nikodým derivative with respect to the law ρ_0 of $\mathbf{e}$ given by

$$\frac{d\rho_\alpha}{d\rho_0}(f) := \frac{e^{-\alpha\|f\|}}{\int e^{-\alpha\|g\|}\,d\rho_0(g)}\,,$$

and where we write $\|g\| := \max_{t\in[0,1]}|g(t)|$ for any $g \in C([0,1],\mathbb{R})$, the set of all continuous functions from $[0,1]$ to $\mathbb{R}$. Define $(\mathcal{T}_\alpha, d_\alpha, \sigma_\alpha)$ be the measured random continuum tree encoded by the function $\mathbf{e}_\alpha$.

Theorem 1.1 (Scaling limit in the Brownian regime). *Let $\mu = \mu_n$ satisfy $\lim_{n\to\infty} \mu\sqrt{n} = \alpha \in \mathbb{R}$. Let $\mathbf{T}_n = \mathbf{T}^\mu$ be a μ-height-biased tree of size n, let d_n denote the graph distance on $\mathbf{T}_n$ with each edge having length 1, and let σ_n denote the uniform measure on the vertices of $\mathbf{T}_n$. Then*

$$\left(\mathbf{T}_n, \frac{d_n}{\sqrt{2n}}, \sigma_n\right) \xrightarrow{d} (\mathcal{T}_\alpha, d_\alpha, \sigma_\alpha),$$

where the convergence in distribution above is with respect to the Gromov-Hausdorff-Prokhorov (GHP) topology.

The proof of this theorem can be found in Section 3. Thus, the global geometry of the random tree $\mathbf{T}_n$ is completely described by appropriately tilting the measured continuum random tree encoded by $\mathbf{e}$. We next complement the above scaling limit with a characterization of the *width* of the tree, as defined in (1.2), whose proof can again be found in Section 3.

For the next result, let us introduce the random function $B_\alpha = (B_\alpha(t) : t \in [0,1]) \in C([0,1],\mathbb{R})$, whose law ξ_α has Radon-Nikodým derivative with respect to the law ρ_0 of $\mathbf{e}$ given by

$$\frac{d\xi_\alpha}{d\rho_0}(f) = \frac{e^{-\alpha\bar{f}}}{\int e^{-\alpha\bar{g}}\,d\rho_0(g)}\,, \tag{1.5}$$

where for $f \in C([0,1],\mathbb{R})$ we use the notation

$$\bar{f} := \frac{1}{2}\int_0^1 \frac{dt}{f(t)}.$$

Remark 1.2 (On the process B_α being well-defined). In general, for a continuous function f, the integral $\int_0^1 \frac{ds}{f(s)}$ need not be well-defined, since there can be a positive measure subset of $[0,1]$ where f vanishes. However, it is classical that this does not happen a.s. for the Brownian excursion $\mathbf{e}$ (see, e.g., [11] and the references therein), so that $\int_0^1 \frac{ds}{\mathbf{e}(s)}$ is well defined as a random variable. In particular, the law ρ_0 of $\mathbf{e}$ is supported by the set of functions f for which the integral $\int_0^1 \frac{ds}{f(s)}$ makes sense. Thus, the change of measure that defines ξ_α makes sense and therefore the process B_α is well-defined.

Theorem 1.3 (Width in the Brownian regime). *Let $\mu = \mu_n$ satisfy $\lim_{n\to\infty} \mu\sqrt{n} = \alpha \in \mathbb{R}$. Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. Then the sequence $(w(\mathbf{T}_n))_{n\geq 1}$ satisfies*

$$\frac{w(\mathbf{T}_n)}{\sqrt{n}} \xrightarrow{d} \|B_\alpha\|.$$

1.2 Global properties in the non-Brownian regime

The previous section identifies the global behavior of the tree $\mathbf{T}_n$ in the Brownian regime. We next complement the previous results by providing results on global properties of the tree $\mathbf{T}_n$ in the *non-Brownian regime*, when $\mu \gg 1/\sqrt{n}$. To state the results concisely, it is useful to further subdivide the non-Brownian regime into two parts.

- When $1/\sqrt{n} \ll \mu \ll n$, we are in the *intermediate regime*.

- When $\mu = \Omega(n)$, we are in the *extreme regime*.

We start by stating the results for the asymptotic behaviour of the height of $\mathbf{T}_n$ in the intermediate regime.

Theorem 1.4 (Asymptotic height in the intermediate regime). *Let $\mu = \mu_n$ be such that $1/\sqrt{n} \ll \mu \ll n$. Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. Then*

$$\frac{h(\mathbf{T}_n)}{\left(\frac{2\pi^2 n}{\mu}\right)^{1/3}} \xrightarrow{\mathbb{P}} 1 .$$

The proof of Theorem 1.4 can be found in Section 4.1. Note that $(2\pi^2 n/\mu)^{1/3}$ ranges from being of order $n^{1/2}$ when $\mu \asymp n^{-1/2}$, to being of order 1 when $\mu \asymp n$. We now extend this result to include the fluctuations around the height. As we present below, the fluctuations actually have a wide range of behaviour: from a standard central limit theorem when $\mu \ll n^{1/4}$, to a discrete central limit theorem when $\mu \asymp n^{1/4}$, and finally to *Bernoulli fluctuations* when $\mu \gg n^{1/4}$, the latter behavior also extending into the extreme regime.

Theorem 1.5 (Gaussian fluctuations for the height). *Let $\mu = \mu_n$ be such that $1/\sqrt{n} \ll \mu \ll n^{1/4}$. Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. Then*

$$\frac{h(\mathbf{T}_n) - \left(\frac{2\pi^2 n}{\mu}\right)^{1/3}}{\sqrt{\frac{1}{3}\left(\frac{2\pi^2 n}{\mu^4}\right)^{1/3}}} \xrightarrow{d} \mathcal{N}(0,1) ,$$

where $\mathcal{N}(0,1)$ is a standard normal random variable.

The proof of Theorem 1.5 can be found in Section 4.2. Observe that when μ is of order $n^{1/4}$, the denominator, which corresponds to the standard deviation of the height, becomes of constant order. In that regime, the previous result almost directly holds, except that the convergence cannot be towards a normal random variable, since $h(\mathbf{T}_n)$ is an integer, but towards the integer equivalent of a normal random variable. To state this result, for any $t \in (2,\infty)$ and $x > 0$ define the function

$$\lambda_x(t) := xt + \log\left(1 + \tan^2\left(\frac{\pi}{t}\right)\right) \tag{1.6}$$

and let $t_x := \operatorname{argmin} \lambda_x(t)$ be its minimizing argument. It is straightforward to check that λ_x is convex and thus admits a unique minimizer. We refer the reader to Proposition C.1 in Appendix C for further properties of λ_x and t_x.

Theorem 1.6 (Discrete central limit theorem for the height). *Let $\mu = \mu_n$ satisfy $\lim_{n\to\infty} \mu/n^{1/4} = \gamma > 0$. Assume that $t_{\mu/n} - \lfloor t_{\mu/n} \rfloor$ admits a limit $\delta \in [0,1]$. Then the asymptotic of the height $h(\mathbf{T}_n)$ is given by*

$$h(\mathbf{T}_n) - \lfloor t_{\mu/n} \rfloor + 2 \xrightarrow{d} X_{\gamma,\delta},$$

where $X_{\gamma,\delta}$ is an integer-valued random variable distributed as

$$\mathbb{P}\left(X_{\gamma,\delta} = k\right) = \frac{1}{C_{\gamma,\delta}} \exp\left(-3\left(\frac{\gamma^2}{4\pi}\right)^{2/3} (k-\delta)^2\right), \quad k \in \mathbb{Z}$$

with $C_{\gamma,\delta}$ being a normalizing constant corresponding to the sum over $k \in \mathbb{Z}$ of the exponential term.

Remark 1.7. We remark that if $\mu/n^{1/4}$ is of constant order but does not admit a limit, or if $t_{\mu/n} - \lfloor t_{\mu/n} \rfloor$ does not admit a limit, the asymptotic distribution of $h(\mathbf{T}_n) - \lfloor t_{\mu/n} \rfloor + 2$ is still close to that of $X_{\mu/n^{1/4}, t_{\mu/n} - \lfloor t_{\mu/n} \rfloor}$, although we do not obtain a distributional limit as stated in the last result.

The proof of Theorem 1.6 can be found in Section 4.2. We observe that, while the distribution of $X_{\gamma,\delta}$ strongly resembles that of a normal distribution, since it replaces the integral by a sum, it is different from the distribution of the integer part of a normal random variable.

After addressing the first order behavior of the height of $\mathbf{T}_n$ in Theorem 1.4 and its fluctuations in Theorem 1.5 and Theorem 1.6, the next result covers the probabilities of large deviations from the mean.

Theorem 1.8 (Large deviations of the height)**.** *Let $\mu = \mu_n$ be such that $1/\sqrt{n} \ll \mu \ll n$. Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. Then the sequence of random variables $(h(\mathbf{T}_n)/(2\pi^2 n/\mu)^{1/3})_{n\geq 1}$ satisfies a large deviation principle at speed $(\mu^2 n)^{1/3}$ and with good rate function given by*

$$
\Lambda^*(x) = \begin{cases} \left(\frac{\pi}{2}\right)^{2/3} \frac{(x-1)^2(2x+1)}{x^2} & \text{if } x > 0 \\ \infty & \text{otherwise.} \end{cases}
$$

The proof of Theorem 1.8 can be found in Section 4.2.2. With this last result on the height in the intermediate regime, we next move on to other functionals of $\mathbf{T}_n$. Our first result in this direction concerns the width of $\mathbf{T}_n$, as defined in (1.2). This result covers *almost* the entire intermediate regime, and extends beyond the intermediate regime well into the extreme regime.

Theorem 1.9 (Width in the non-Brownian regime)**.** *Let $\mu = \mu_n$ be such that $\mu \gg (\log n)^{3/2}/\sqrt{n}$. Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. Then for any $\epsilon > 0$, we can find $K = K(\epsilon) > 0$ large enough, such that*

$$
\liminf_{n\to\infty} \mathbb{P}\left(K^{-1} \leq \frac{w(\mathbf{T}_n)}{n \wedge \left((\mu n^2)^{1/3}\right)} \leq K \right) \geq 1 - \epsilon.
$$

The proof of Theorem 1.9 can be found in Section 5. Note that in the current version we do not characterize the width behavior for the entire intermediate regime: namely, our result has a gap for the regime $1/\sqrt{n} \ll \mu = O\left((\log n)^{3/2}/\sqrt{n}\right)$. In the last theorem, we only provide a tightness result on the scaled width. We conjecture a *law of large numbers* for the width in the entire non-Brownian regime:

Conjecture 1.10 (LLN for the width in the non-Brownian regime)**.** *Let $\mu = \mu_n$ be such that $\mu \gg 1/\sqrt{n}$. Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. There exists a constant $W > 0$ such that as $n \to \infty$,*

$$
\frac{w(\mathbf{T}_n)}{n \wedge \left((\mu n^2)^{1/3}\right)} \xrightarrow{\mathbb{P}} W.
$$

To prove Theorem 1.9, we employ a probabilistic technique involving a multiscale decomposition of the associated *random excursion* of the tree $\mathbf{T}_n$. More discussion on this technique and the proof of this result can be found in Section 5.

Finally, we state our main result on the height of $\mathbf{T}_n$ in the extreme regime, also covering fluctuations when $\mu \gg n^{1/4}$. Recall λ_x and t_x from (1.6). For any $x > 0$ define $m_x = \max\{\lfloor t_x \rfloor, 3\}$ and

$$
\delta_n = \mu + n \log\left(\frac{1 + \tan^2\left(\frac{\pi}{\lceil t_{\mu/n}\rceil}\right)}{1 + \tan^2\left(\frac{\pi}{m_{\mu/n}}\right)} \right). \tag{1.7}
$$

Theorem 1.11 (Height fluctuations beyond the Gaussian sub-regime)**.** *Let $\mu = \mu_n$ be such that $\mu \gg n^{1/4}$ and such that μ/n converges to some $c \in [0, +\infty]$. Further, assume that δ_n admits a limit $\delta \in [-\infty, \infty]$. Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. Then, $h(\mathbf{T}_n) - m_{\mu/n} + 2$ converges in distribution to a Bernoulli random variable with parameter*

$$
p_{c,\delta} = \begin{cases} \dfrac{m_c \tan^2\left(\frac{\pi}{m_c+1}\right) e^{-\delta}}{(m_c+1)\tan^2\left(\frac{\pi}{m_c}\right) + m_c \tan^2\left(\frac{\pi}{m_c+1}\right) e^{-\delta}} & \text{if } c > 0 \\ \dfrac{e^{-\delta}}{1 + e^{-\delta}} & \text{otherwise.} \end{cases}
$$

Observe that, since m_c diverges to ∞ when c converges to 0, the two formulas in the definition of $p_{c,\delta}$ are coherent with each other, although need to be split since the formula is not well-defined for $m_c = \infty$. The proof of Theorem 1.11 can be found in Section 4.3. This theorem states that $h(\mathbf{T}_n)$ is asymptotically deterministic, except in specific regimes for μ where it transitions from one value to another. This is to be expected, as the formula for the variance of $h(\mathbf{T}_n)$ as found in Theorem 1.5 converges to 0 whenever $\mu \gg n^{1/4}$, and since $h(\mathbf{T}_n)$ takes integer values and thus cannot change from one value to another without transition.

Furthermore, we observe that $p_{c,\delta}$ is equal to 0 when $\delta = \infty$ and to 1 when $\delta = -\infty$. This can be understood by fixing the limiting value $m_{\mu/n} = \tilde{m}$ and letting μ approach the threshold

$$n \log \left(\frac{1 + \tan^2\left(\frac{\pi}{\tilde{m}+1}\right)}{1 + \tan^2\left(\frac{\pi}{\tilde{m}}\right)} \right).$$

Indeed, when $\delta = -\infty$, meaning that μ is far below the threshold, then the height is less biased and thus higher, leading to $h(T_n) = \tilde{m}-1$. Then, as μ approaches the threshold and δ becomes of constant order, it becomes possible for $h(T_n)$ to decrease by one level, from $\tilde{m}-1$ to $\tilde{m}-2$. Finally, when μ has long passed the threshold and $\delta = +\infty$, the height-bias is too large and we have reduced the height by 1, leading to $h(T_n) = \tilde{m}-2$.

Remark 1.12 (Dichotomy of the concentration of the height about its mean). We observe the following dichotomy on the concentration behavior of the height about its mean. When $\mu = O(1/\sqrt{n})$, the height has random fluctuations at its highest order (Theorem 1.1), but also when $\mu = \Omega(n)$ (Theorem 1.11); while in the middle regime of $1/\sqrt{n} \ll \mu \ll n$, the height is more concentrated and satisfies a LLN (Theorem 1.4), where precise large-deviation estimates can be given (Theorem 1.8).

Using that the height of a tree with more than 2 nodes cannot be less than 1, it is interesting to extract from the previous result the regime in which $h(T_n)$ is minimal, and so the tree is simply the star S_n with $n-1$ leaves, as stated in the following corollary.

Corollary 1.13 (Observing a star). *Let $\mu = \mu_n$ be such that $\mu - n\log 2$ admits a limit $\delta \in \mathbb{R} \cup \{\pm\infty\}$. Then as $n \to \infty$*

$$\mathbb{P}\left(h(\mathbf{T}_n) = 1\right) = \mathbb{P}\left(T_n = S_n\right) \longrightarrow \frac{4}{4 + e^{-\delta}}.$$

Remark 1.14. Observe that this result does not require $\mu \gg n^{1/4}$, or even $\mu \gg 1/\sqrt{n}$. Indeed, when μ is sublinear, i.e., $\mu \ll n$, then $\delta = -\infty$, and in this case the corollary simply tells us that the probability of obtaining the star graph is asymptotically 0.

Let us quickly give the proof of the last corollary.

Proof of Corollary 1.13. This is a direct consequence of Theorem 1.11 by observing that

$$\log \left(\frac{1 + \tan^2\left(\frac{\pi}{3}\right)}{1 + \tan^2\left(\frac{\pi}{4}\right)} \right) = \log\left(\frac{1+3}{1+1} \right) = \log 2.$$

Indeed, if $m_{\mu/n} \neq 3$ and $\lceil t_{\mu/n} \rceil \neq 4$, then this means that μ is lower than the possible threshold considered here, leading to $\delta = -\infty$. Using the different results regarding the height of $h(T_n)$ (Theorem 1.1, 1.4, or 1.11), we see that $T_n = S_n$ with probability 0 here, thus satisfying the claimed result. On the other hand, if $m_{\mu/n} = 3$ and $\lceil t_{\mu/n} \rceil = 4$, then we can directly apply Theorem 1.11. Finally, if $m_{\mu/n} = \lceil t_{\mu/n} \rceil = 3$, then this means that $t_{\mu/n} \leq 3$, in which case

$$\frac{\mu}{n} \geq \frac{2\pi}{9} \tan \frac{\pi}{3} = \frac{2\pi}{3^{3/2}} > 1 > \log 2.$$

It follows that $\delta = +\infty$ and then Theorem 1.11 confirms that $T_n = S_n$ with high probability. $\square$

Let us note that in the very special case when $\mu \gg n$, we have a positive answer to Conjecture 1.10 due to Corollary 1.13. As $\delta = \infty$, we see a star with high probability, and thus $w(\mathbf{T}_n)/n \to 1$ in probability. This finishes all the statements about global functionals in the non-Brownian regime. In the next section, we state results on *local* properties of $\mathbf{T}_n$.

1.3 Results on local properties

In this section, we state a dichotomy between two qualitatively different local behaviors around the root of the tree $\mathbf{T}_n$, depending on whether or not μ diverges to infinity. We name these regimes as,

- **Local regime**, when $\mu = O(1)$.

- **Condensation regime**, where $\mu \gg 1$.

For the next result, the relevant definitions for the topology can be found in Section 2.

Theorem 1.15 (Local limit in the local regime). *Let $\mu = \mu_n$ be a sequence admitting a limit $\alpha \in [0, \infty)$. Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. If $\alpha > 0$, then the random rooted tree $(\mathbf{T}_n, \varnothing)$ converges locally to the infinite* Poisson tree *of parameter α as defined in [2]. Otherwise, if $\alpha = 0$, equivalently $\mu = o(1)$, then the random rooted tree $(\mathbf{T}_n, \varnothing)$ converges locally to Kesten's tree as defined in [22].*

The proof of Theorem 1.15 can be found in Section 6. We remark that the statement on the local limit when $\mu = \mu_n \equiv \alpha > 0$ is a constant sequence was already identified in [19, Theorem 4.3]. The above result additionally extends to the case where $\mu \ll 1$.

For any tree T, and any vertex $v \in T$, we denote by $\deg_T(v)$ the *number of children* of the vertex v in T, i.e., the number of neighbors of v in T not lying on the ancestral path of v to the root. When $\mu \gg 1$, we observe a condensation phenomenon at the root, i.e., its degree diverges.

Theorem 1.16 (Condensation at the root). *Let $\mu = \mu_n$ be such that $1 \ll \mu \ll n^{1/4}$. Let $\mathbf{T}_n$ be a tree distributed according to (1.3). Then the asymptotic behavior of the degree of the root $\varnothing$ in $\mathbf{T}_n$ is given by*

$$\frac{\deg_{\mathbf{T}_n}(\varnothing) - 2\mu}{\sqrt{6\mu}} \xrightarrow{d} \mathcal{N}(0, 1).$$

The proof of this theorem can be found in Section 6.2. This result shows that, in the non-local regime and up to $\mu \ll n^{1/4}$, we observe a condensation phenomenon as the root degree starts diverging at scale μ, as opposed to the local regime, where it is almost surely finite. We conjecture that the root degree satisfies a law of large numbers when $\mu = \Omega(n^{1/4})$.

Conjecture 1.17. *Assume $n \gg \mu = \Omega(n^{1/4})$. Then as $n \to \infty$, there is a constant $C > 0$ such that*

$$\frac{\deg_{\mathbf{T}_n}(\varnothing)}{\mu} \xrightarrow{\mathbb{P}} C.$$

The second order fluctuations of the root degree in the regime $n \gg \mu = \Omega(n^{1/4})$ appear to be more subtle, and we do not have a conjecture at this point.

Remark 1.18 (Root degree when $\mu = \Omega(n)$). Observe that Corollary 1.13 indirectly states that, whenever $\mu - n \log 2 \to \delta \in \mathbb{R} \cup \{\pm\infty\}$, then $\mathbb{P}\left(\deg_{\mathbf{T}_n}(\varnothing) = n - 1\right) = 4/(4 + e^{-\delta})$. This suggests that there exists a regime for μ where the limit $(n-1) - \deg_{\mathbf{T}_n}(\varnothing)$ is non-trivial. Moreover, and in order to connect the regime $\mu = \Omega(n)$ with that of $\mu \ll n$ from Conjecture 1.17, we also expect the existence of a regime for μ where $(n-1) - \deg_{\mathbf{T}_n}(\varnothing)$ diverges but $\deg_{\mathbf{T}_n}(\varnothing)/n$ is non-trivial. At the moment, it is unclear to us where the exact change of behavior occurs and how they impact the two types of limit (subtracting n and dividing by n).

Remark 1.19 (Comparison with the local results of [2]). As a result of Theorems 1.15 and 1.16, we observe the following qualitative picture in the transition of the local behavior around the root. As μ increases, locally around the root we begin with observing Kesten's tree, until it reaches a threshold when $\mu \to \alpha \in (0, \infty)$, when the behavior changes to that of the Poisson tree; beyond this threshold, we observe condensation at the root, in that its degree diverges. A similar picture was also observed in a different context, in the paper [2]. There, the authors consider Bienaymé trees conditoned on explicit generation sizes, thus penalizing the height *indirectly*, unlike the direct exponential penalization of the measure (1.3) that we consider. This suggests that this qualitative transition landscape is in a sense *universal*, and may emerge in other contexts.

Regime	$\mu = O(1/\sqrt{n})$	$1/\sqrt{n} \ll \mu \ll n$				$\mu = \Omega(n)$
Sub-regime	$\mu = O(1/\sqrt{n})$	$1/\sqrt{n} \ll \mu \ll 1$	$\mu \asymp 1$	$1 \ll \mu = O(n^{1/4})$	$n^{1/4} \ll \mu \ll n$	$\mu = \Omega(n)$
Height	$\asymp_{\mathbb{P}} \sqrt{n}$	$\sim (2\pi^2 n/\mu)^{1/3}$				$\asymp_{\mathbb{P}} 1$
LDP speed for height	NA	$\asymp (\mu^2 n)^{1/3}$				NA
Width	$\asymp_{\mathbb{P}} \sqrt{n}$	$\asymp (\mu n^2)^{1/3}$ (LLN conjectured)				$\asymp n$
Local behavior	Kesten's tree		Poisson tree	Condensation		
Root degree	$\asymp_{\mathbb{P}} 1$		$\asymp \mu$ (conjectured for $\mu = \Omega(n^{1/4})$)			$\asymp n$
Fluctuation of height	$\asymp_{\mathbb{P}} \sqrt{n}$	$\asymp_{\mathbb{P}} (n/\mu^4)^{1/6}$			$O(1)$	

Table 1: A summary of our different results together with Conjectures 1.17 and 1.10 depending on the regime of μ.

1.4 Related work and discussion

As has been mentioned before, the Poisson tree which is the local limit of the model in hand for $\mu > 0$ constant has first been obtained by conditioning the geometric Bienaymé trees on their generation sizes in [1]. The decomposition of the local limit around a fixed finite tree found in [19, Theorem 4.4] shows that, conditioned on the degree of the root being say M, one observes at least one and possibly more infinite subtrees emanating from those M vertices at height 1, together with finite subtrees grafted on the rest. Each one of these infinite subtrees is a height-biased tree on its own and they have a *joint* law. Tracing these infinite branches back to the stage where n is large (and $n \to \infty$ limit has not yet been taken), we find the large branches of size $\asymp n^{1/3}$. We expect a similar picture when $1 \ll \mu \ll n^{1/4}$: the degree of the root having mean 2μ, about half of which carrying large branches of size $\asymp (n/\mu)^{1/3}$, each one of them being a height-biased tree with some positive *constant* bias. Although we have a number of heuristics that support this expectation, we have not rigorously investigated it in the present work.

In [9], a very similar problem in case of one-dimensional simple random walk has been treated where the walk is penalized by its range R i.e. the weight of a walk of size n is proportional to $\exp(-\mu R)$. One again observes the stretched exponential in the asymptotics of the partition function. The limit laws for the range as $n \to \infty$ resemble closely to our results for the height and the windows of μ where different phases of the two models are observed are the same.

It is worth mentioning that the results from [9] require $\mu \gg (\log n)^{3/2}/\sqrt{n}$. This threshold is quite natural when studying such models, as it arises when the leading term of the partition function, of order $e^{(\mu^2 n)^{1/3}}$, becomes of the same order as any power of n or μ. To circumvent this issue, we end up relying on very precise asymptotics for the number of trees of a given height. In particular, we refer the reader to Theorem B.1 where we provide the exact asymptotic behaviour of the number of trees of size n and height $< m$:

$$H_{n,m} = \left| \left\{ T \in \mathbb{t}_n : h(t) < m \right\} \right|. \tag{1.8}$$

While the size of this set has been well-studied, with a closed form provided in [12], the best bound we could extract from the literature only covered the first order logarithm limit of $H_{n,m}$ (for example by applying the Mogul'skiĭ estimate [27]), which only provided results when $\mu \gg (\log n)^{3/2}/\sqrt{n}$ and not all the way to $\mu \gg 1/\sqrt{n}$.

The height-biased model where the weight function is power-like, i.e., where the weight of a tree T of size n is proportional to $h(T)^\alpha$ for α real, has been studied in [18]. There, the bias turns out to be not strong enough to affect the local limit even though the asymptotics of the partition function differ from the uniform case. The extension of these results to the case where $\alpha = \alpha_n$ depends on the size of the tree, remains open.

Before concluding this section, we mention some existing models of biased/weighted random trees, even though they exhibit different characteristics from ours. A well-known class of trees are the recursive trees (RT) that are constructed by starting from a root of label 1 and joining new (labelled) vertices step by step. In [24], the authors introduced *depth-weighted random recursive trees* where the probability of joining a new leaf to an existing vertex is proportional to its depth in the tree. They also observe a change in the height of the tree when the weight function is exponential in depth, whereas the order of the height does not differ from the uniform case when the weight function is power-like, similar to phenomena we observe between [18] and [17]. More recently, [6] studies recursive tree models where there is a preference for the new leaf to be joined towards *more recently arrived* vertices, thus indirectly making the obtained trees *taller* than when there is no such preference (i.e. taller than the

9

uniform random recursive tree (URRT) [28]). Indeed, the authors exhibit both logarithmic and polynomial height regimes for such trees. In another recent work [29] the authors studied a model of random trees that interpolates (in terms of the asymptotic form of the average path length, which is the sum of all distances to the root) between uniformly random rooted labelled trees (a sub-class of the so-called *simply generated trees* [21]) and recursive trees. There, an edge with the vertex closer to the root having label larger than the label of the other vertex it is adjacent to is called a *descent* and the weight of a tree of size n is proportional to a given parameter $q \in (0, 1]$ to the power of the number of descents; when $q = 1$, they recover the model of uniformly random labeled rooted trees.

Organization of the rest of the paper. In Section 2 we introduce some standard and useful notions related to viewing trees as excursions. Then in Section 3 we provide the proofs of Theorems 1.8 and 1.3 for global properties in the Brownian regime. In Section 4, we provide all the results related to the height of the tree in the non-Brownian regimes, i.e., proofs of Theorems 1.4, 1.5, 1.6, 1.8, 1.9 and 1.11. In Section 6, we provide the proofs of results on local behavior around the root, i.e., the proofs of Theorems 1.15 and 1.16. In Appendix A we prove a technical result on functions of certain exponential form. We use this result first in Appendix B to derive properties of the numbers $H_{n,m}$ as defined in (1.8), and later in Appendices C, D and E, where we provide the explicit analysis that derives the asymptotic behavior of the partition function as defined in (1.4).

2   Encoding trees with walks

In this section, we discuss *encodings* of random trees, using random walk excursions. These universal concepts will aid us in proving the scaling limit result of Theorem 1.1, as well as the result on the width in the non-Brownian regime, Theorem 1.3.

For any integer $k > 0$, let

$$\mathcal{W}^{(k)} := \left\{ (w_t : 0 \le t \le k) : w_0 = 0 \text{ and } |w_t - w_{t+1}| = 1 \text{ for } 0 \le t \le k-1 \right\}$$

be the set of *walks* starting at 0 of length k and step size 1. Let us further define

$$\mathcal{B}_{x,y}^{(k)} := \left\{ (x + w_t : 0 \le t \le k) : w \in \mathcal{W}^{(k)}, x + w_k = y \text{ and } |w_t - w_{t+1}| = 1 \text{ for } 0 \le t \le k-1 \right\}$$

for the set of *bridges* of length k and going from x to y, and

$$\mathcal{E}^{(k)} := \left\{ (w_t : 0 \le t \le k) : w \in \mathcal{W}^{(k)}, w_k = -1, w_t \ge 0 \text{ for } 0 \le t \le k-1 \text{ and } |w_t - w_{t+1}| = 1 \text{ for } 0 \le t \le k-1 \right\}$$

for the set of *excursions* of length k. It is worth noting that the set of bridges is empty whenever k and $y - x$ do not have the same parity, which also means that the set of excursions is empty whenever k is not odd. In the case of bridges, when $x = 0$ we simply denote it with $\mathcal{B}_y^{(k)}$. Since we will eventually be interested in scaling these objects and taking their limit, we equivalently see walks, bridges and excursions as their sequence of integer values or as continuous functions on $[0, k]$ where we linearly interpolate between the values at integer arguments.

Definition 2.1 (Contours)**.** For a plane tree T with n vertices, the contour walk $(C_t^{(T)}; 0 \le t \le 2n - 1)$ of T is an element of $\mathcal{E}^{(2n-1)}$, that traces the contour of the tree in the depth-first manner from left to right, starting at 0 and finishing at $2n - 1$. We refer the reader to Figure 1 for a visual representation.

The set of all rooted plane trees of size n is in bijective correspondence with the set $\mathcal{E}^{(2n-1)}$, with a tree simply being mapped to its contour walk. For the random trees $\mathbf{T}_n$ or $\mathbb{T}_n$, (recall $\mathbb{T}_n$ is the uniform tree, i.e., the random tree with law (1.3) corresponding to $\mu \equiv 0$) the contours are random variables taking values in $\mathcal{E}^{(2n-1)}$. In particular, thanks to the bijection mentioned above, the contour $(C_t^{(\mathbb{T}_n)}; 0 \le t \le 2n - 1)$ of the random tree $\mathbb{T}_n$, is a uniformly random element of $\mathcal{E}^{(2n-1)}$.

Remark 2.2 (Heights and maximums)**.** It is worth mentioning that the height of a tree exactly corresponds to the maximal value of its corresponding contour walk seen as a continuous function.

Sometimes we will translate events on bridges to events on excursions, and thus events on trees, which are a bit easier to understand. For this, we will need the *cycle lemma*. The result is relatively standard, and we refer the reader to [21, Lemma 15.3] for a proof.
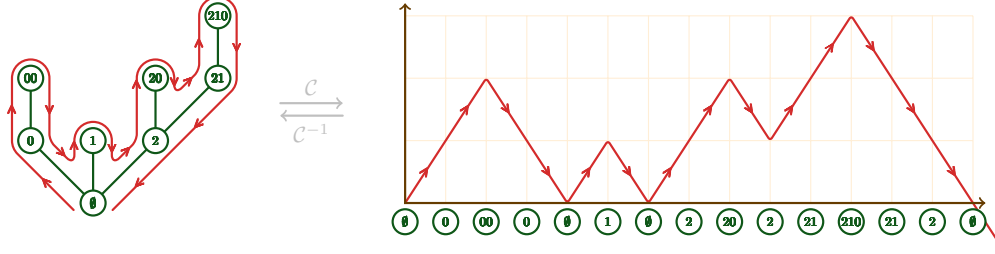
Figure 1: A representation of the bijection between a rooted plane tree with n nodes and its contour — an excursion of length $2n-1$. Each step of the excursion is represented by an arrow, both on the walk and on the tree, and corresponds to a single side of an edge (left or right). To highlight this relation, the nodes of the tree are shown at the bottom of the walk in the order in which they would be encountered when contouring around the tree.

Lemma 2.3 (Cycle lemma). *Fix $n \geq 1$. There is a map $\phi : \mathcal{B}_{-1}^{(2n-1)} \to \mathcal{E}^{(2n-1)}$ from the set of all bridges starting 0 and ending at -1, to the set of excursions, both of length $2n-1$, such that ϕ is a $(2n-1)$ to 1 map.*

Bridges that start at 0 and end at -1 are called *standard*. Sometimes it will be important to work with standard bridges. We will do this using the following result, that transforms general bridges starting at 0 to standard bridges with a constraint.

Lemma 2.4 (Cut bijection lemma). *Let $n \geq 1$ and $x \in \mathbb{Z}$. There is a function ϕ_x that bijectively sends the set $\mathcal{B}_x^{(n)}$ of bridges from 0 to x onto*

$$\tilde{\mathcal{B}}_x^{(n)} := \mathcal{B}_{-\mathbb{1}_{\{n \text{ is odd}\}}}^{(n)} \cap \left\{ \left(w_t : 0 \leq t \leq n \right) : \exists t, w_t = \lfloor x/2 \rfloor \right\}$$

of all bridges starting at 0, ending at -1 when n is odd or at 0 when n is even, and passing through $\lfloor x/2 \rfloor$.

Proof. We proceed by constructing such a function ϕ_x. Let $w \in \mathcal{W}^{(n)}$ be a walk such that there exists t satisfying $w_t = \lfloor x/2 \rfloor$. For such a w, let τ be its time of last passage from $\lfloor x/2 \rfloor$, i.e.,

$$\tau := \sup \left\{ t \geq 0 : w_t = \left\lfloor \frac{x}{2} \right\rfloor \right\}.$$

Then ϕ_x is the map such that $\phi_x(w)$ coincides with w on $0 \leq t \leq \tau$ and is the reflection of w around the horizontal axis $y = \lfloor x/2 \rfloor$ on $\tau \leq t \leq n$. In other words, seeing a walk as a sequence of ± 1 steps, w and $\phi_x(w)$ share the same first τ steps and then have opposite steps. We refer to Figure 2 for a representation of this function, where the original walk, in red and from the set $\mathcal{B}_x^{(n)}$, is transformed into its green reflection which belongs to $\tilde{\mathcal{B}}_x^{(n)}$.

As a first observation, for any w such that there exists t satisfying $w_t = \lfloor x/2 \rfloor$, we see that the last hitting time of $\lfloor x/2 \rfloor$ of w and $\phi_x(w)$ is the same. Indeed, $w_\tau = \lfloor x/2 \rfloor$ so we also have $\phi_x(w)_\tau = \lfloor x/2 \rfloor$ and since the two walks are the reflection of each other around $y = \lfloor x/2 \rfloor$ on (τ, n), each one intersects the axis if and only if the other one does. This remark implies that $\phi_x \cdot \phi_x(w) = w$, meaning that ϕ_x is a symmetry.

Assume now that w is such that $w_n = x$. In that case, we necessarily have some t such that $w_t = \lfloor x/2 \rfloor$, so $\phi_x(w)$ is well-defined. Since w and $\phi_x(w)$ coincide on $[0, \tau]$ and the steps of w from τ to n go from $\lfloor x/2 \rfloor$ to $x = \lfloor x/2 \rfloor + (x - \lfloor x/2 \rfloor)$, the steps of $\phi_x(w)$ from τ to n go from $\lfloor x/2 \rfloor$ to $\lfloor x/2 \rfloor - (x - \lfloor x/2 \rfloor) = 2\lfloor x/2 \rfloor - x$. Recalling that x has to have the the same parity as n, we see that $\phi_x(w)$ indeed ends at 0 when n is even and at -1 when n is odd. Conversely, if $w_n = -\mathbb{1}_{n \text{ is odd}}$ and there exists t such that $w_t = \lfloor x/2 \rfloor$, then $\phi_x(w)$ is a walk from 0 to $2\lfloor x/2 \rfloor + \mathbb{1}_{n \text{ is odd}} = x$. This implies that ϕ_x sends $\mathcal{B}_x^{(n)}$ into $\tilde{\mathcal{B}}_x^{(n)}$ and vice-versa. Using that $\phi_x \cdot \phi_x$ is the identity, this proves that ϕ_x bijectively maps to the sets. $\square$
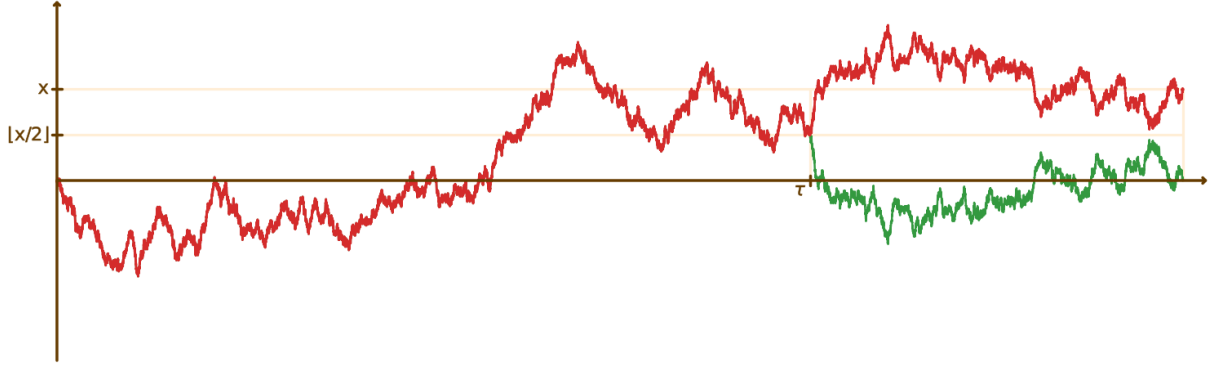
Figure 2: In the above picture for an element w of $\mathscr{B}_x^{(n)}$, displayed as the red path, we apply the map ϕ_x to reflect the segment of the path w from the last passage τ at $\lfloor x/2 \rfloor$ to n with respect to $y = \lfloor x/2 \rfloor$ to obtain the green segment. The image $\phi_x(w)$ is thus the concatenation of the red path up to time τ together with the green segment.

3 Proof of global results in the Brownian regime

In this section, we prove the scaling limit results in the Brownian regime, i.e., Theorems 1.1 and 1.3. We begin with Section 3.1 discussing some relevant concepts, and then move on to the actual proofs.

3.1 Gromov-Hausdorff-Prokhorov convergence and real trees

For more details on what follows, we refer the reader to [20, Section 3]. Informally, one can think of real trees as a connected collection of line segments in the plane with no loops. They can be constructed analogously to the discrete case as follows. Let $C([0,1], \mathbb{R}_+)$ be the space of continuous functions from $[0,1]$ into $\mathbb{R}_+$ equipped with the topology of uniform convergence. Given a continuous function g such that $g(0) = g(1) = 0$, the construction of the corresponding real tree is standard: for every $s, t \in [0,1]$, set

$$d_g(s,t) = g(s) + g(t) - 2m_g(s,t)$$

where

$$m_g(s,t) = \inf_{z \in [s \wedge t, s \vee t]} g(z).$$

Set $T_g = [0,1]/\sim$ where $s \sim t$ if and only if $d_g(s,t) = 0$. Note that d_g induces a distance on T_g, which we denote by d_g as well and is in fact a pseudo-metric. The canonical projection $p_g : [0,1] \to T_g$ is continuous hence T_g is compact. The compact metric space (T_g, d_g) is then a real tree, which we view as a rooted real tree with the root $\rho = p_g(0) = p_g(1)$.

In order to make sense of the convergence of rooted discrete trees towards real trees, we introduce the Gromov-Hausdorff-Prokhorov (GHP) distance on the set $\mathbb{K}$ of all isometry classes of rooted compact measured metric spaces, where a *rooted* (also commonly referred to as *pointed*) metric space is a metric space with a distinguished *root*.

Let (E, d) be a metric space with its Borel $\sigma-$algebra $\mathscr{B}(E)$. The Hausdorff distance d_H between two non-empty subsets X, Y of E is defined as

$$d_H(X,Y) = \inf\left\{\varepsilon > 0 : X \subset U^\varepsilon(Y) \text{ and } Y \subset U^\varepsilon(X)\right\},$$

where $U^\varepsilon(X) := \{x \in E : d(x, X) < \varepsilon\}$. Let (E_1, ρ_1, d_1) and (E_2, ρ_2, d_2) be two rooted compact metric spaces with metrics d_1 and d_2 and with root ρ_1 and ρ_2, respectively. The Gromov-Hausdorff distance d_{GH} between E_1 and E_2 is defined by

$$d_{GH}(E_1, E_2) = \inf_{\phi_1, \phi_2} \left\{d_H\big(\phi_1(E_1), \phi_2(E_2)\big) \vee d\big(\phi_1(\rho_1), \phi_2(\rho_2)\big)\right\}$$

where the infimum is over all possible choices of metric spaces (E, d) and the isometric embeddings $\phi_1 : E_1 \to E$ and $\phi_2 : E_2 \to E$ of E_1 and E_2 into E. Denoting by $P(E)$ the collection of all probability measures on $(E, \mathscr{B}(E))$, the Prokhorov distance $d_P : P(E) \times P(E) \to [0, \infty)$ between two Borel probability measures μ and ν on E is given by

$$d_P(\mu, \nu) = \inf \left\{ \varepsilon > 0 : \mu(A) \le \nu(U^\varepsilon(A)) + \varepsilon \text{ and } \nu(A) \le \mu(U^\varepsilon(A)) + \varepsilon, \ \forall A \in \mathscr{B}(E) \right\}.$$

Finally, the GHP distance between $(E_1, \rho_1, d_1, \mu_1)$ and $(E_2, \rho_2, d_2, \mu_2)$, where μ_i is a probability measure on E_i for $i = 1, 2$, is given by

$$d_{GHP}(E_1, E_2) = \inf \left\{ d_{GH}(\phi_1(E_1), \phi_2(E_2)) \vee d_P(\phi_{1*}\mu_1, \phi_{2*}\mu_2) \right\}$$

where the infimum is again over all possible choices of metric spaces (E, d) and the isometric embeddings $\phi_1 : E_1 \to E$ and $\phi_2 : E_2 \to E$ of E_1 and E_2 into E and ϕ_* stands for the pushforward measure. For more details on GH and GHP distances and the topologies they induce, we refer the reader to [10, 26, 20].

Definition 3.1 (Brownian Continuum Random Tree)**.** The Brownian Continuum Random Tree (BCRT in short) [4] is the random real tree $(T_{\mathbf{e}}, \rho_{\mathbf{e}}, d_{\mathbf{e}}, \sigma_{\mathbf{e}})$, where $(\mathbf{e}(t) : 0 \le t \le 1)$ is the standard Brownian excursion, and $\sigma_{\mathbf{e}}$ is the pushforward of the Lebesgue measure on $[0, 1]$ under the projection map of the quotient $\sim$.

For more discussion on this object and the Brownian excursion, we refer the reader to [15]. The conclusion of Theorem 1.1 says that when $\mu = O(1/\sqrt{n})$, the scaling limit of the tree $\mathbf{T}_n$ is obtained by appropriately tilting the above defined BCRT.

Finally, let us state without proof two useful results that we will need in the proof of Theorem 1.1. The first one is [20, Corollary 3.7]. Recall that, for any function f, $\|f\|$ denotes its maximum value.

Proposition 3.2. *Let $\mathscr{T}_g$ and $\mathscr{T}_{g'}$ be two real trees encoded by excursions $g, g' : [0, 1] \to \mathbb{R}_+$. Then*

$$d_{GH}(\mathscr{T}_g, \mathscr{T}_{g'}) \le 2\|g - g'\|.$$

The following proposition is [16, Proposition 4.14] adapted for our purposes.

Proposition 3.3. *Let $d_{\mathscr{T}}$ and $d_{\mathscr{T}'}$ be two pseudo-metrics on $[0, 1]$ with $\mathscr{T}$ and $\mathscr{T}'$ induced real trees equipped with the push-forward of the Lebesgue measure on $[0, 1]$. Then*

$$d_{GHP}(\mathscr{T}, \mathscr{T}') \le \frac{3}{2}\|d_{\mathscr{T}} - d_{\mathscr{T}'}\|.$$

3.2 Scaling limit in the Brownian regime

In this section, we provide the proof of theorem 1.1. Recall the discussion on real trees and GHP convergence from Section 3.1. In what follows, whenever the dominated convergence theorem is evoked, we implicitly use Skorokhod's embedding theorem (e.g., see [23]) to couple our random variables in a common probability space and pass to almost sure convergence.

Proof of Theorem 1.1. Recall the definition of the contour of a tree from Definition 2.1. Let $\mathbb{T}_n$ be a *uniformly* sampled tree, and let $\mathbb{C}^{(\mathbb{T}_n)}$ be the rescaled contour walk of the tree $\mathbb{T}_n$ obtained by applying the Brownian scaling on the original contour walk on $[0, 2n - 1]$, i.e.

$$\mathbb{C}^{(\mathbb{T}_n)}(t) = \frac{1}{\sqrt{n}} C^{(\mathbb{T}_n)}((2n - 1)t)$$

for $t \in [0, 1]$, where $C^{(\mathbb{T}_n)}$ is the contour of $\mathbb{T}_n$. Let us also denote the set of such rescaled contour walks of length $2n - 1$ by $\mathscr{C}^{(2n-1)}$. We note that $\mathscr{C}^{(2n-1)}$ is a finite subset of $C([0, 1], \mathbb{R}_+)$ and further, since the association of trees to their respective contours is bijective, we note that the law of $\mathbb{C}^{(\mathbb{T}_n)}$ is uniform on $\mathscr{C}^{(2n-1)}$.

Consider the contour walk $C^{(\mathbf{T}_n)}$ of the μ-height-biased tree $\mathbf{T}_n$. Note that by the definition of the distribution (1.3) and Remark 2.2, the law of $C^{(\mathbf{T}_n)}$, has a Radon-Nikodým derivative with respect to the law of $C^{(\mathbb{T}_n)}$, that is proportional to the function $\exp(-\mu\|f\|)$, on the space of all discrete contours of length $2n - 1$. As a consequence, the contour $\mathbb{C}^{(\mathbf{T}_n)}$ obtained after applying Brownian scaling on $C^{(\mathbf{T}_n)}$, i.e.,

$$\mathbb{C}^{(\mathbf{T}_n)}(t) = \frac{1}{\sqrt{n}} C^{(\mathbf{T}_n)}((2n - 1)t)$$

for $t \in [0,1]$, has a Radon-Nikodým derivative with respect to the rescaled uniform contour $\mathbb{C}^{(\mathbb{T}_n)}$, that is proportional to $\exp(-\mu\sqrt{n}\|f\|)$. Since the law of $\mathbb{C}^{(\mathbb{T}_n)}(t)$ is uniform on $\mathscr{C}^{(2n-1)}$, this is equivalent to simply saying that

$$\mathbb{P}\left(\mathbb{C}^{(\mathbf{T}_n)} = \mathbb{C}\right) \propto \exp\left(-\mu\sqrt{n}\|\mathbb{C}\|\right)$$

over all $\mathbb{C} \in \mathscr{C}^{(2n-1)}$. These observations at hand, let us first show that $\mathbb{C}^{(\mathbf{T}_n)}$ converges in distribution to $\mathbf{e}_\alpha$ in the supremum norm topology, $\mathbf{e}_\alpha$ is the tilted Brownian excursion as defined right before Theorem 1.1. Endowing $C([0,1],\mathbb{R}_+)$ with the supremum topology, by the Portmanteau theorem (see [23]), it is sufficient to show that for any continuous bounded function $\Lambda : C([0,1],\mathbb{R}_+) \to \mathbb{R}$,

$$\mathbb{E}\left[\Lambda\left(\mathbb{C}^{(\mathbf{T}_n)}\right)\right] = \frac{\mathbb{E}\left[\Lambda(\mathbb{C}^{(\mathbb{T}_n)})e^{-\mu\sqrt{n}\|\mathbb{C}^{(\mathbb{T}_n)}\|}\right]}{\mathbb{E}\left[e^{-\mu\sqrt{n}\|\mathbb{C}^{(\mathbb{T}_n)}\|}\right]} \longrightarrow \mathbb{E}[\Lambda(\mathbf{e}_\alpha)] .$$

It is standard that, in the uniform case, we have convergence of contours $\mathbb{C}^{(\mathbb{T}_n)}$ towards the standard Brownian excursion $(\mathbf{e}(t))_{t \in [0,1]}$ on $[0,1]$, see e.g., [4]. Since $\|\mathbb{C}^{(\mathbb{T}_n)}\|$ admits sub-Gaussian tails by [3, Theorem 1.2], we can deduce that $\exp(-\mu\sqrt{n}\|\mathbb{C}^{(\mathbb{T}_n)}\|)$ is uniformly integrable. Together with the boundedness of Λ this implies that

$$\frac{\mathbb{E}\left[\Lambda(\mathbb{C}^{(\mathbb{T}_n)})e^{-\mu\sqrt{n}\|\mathbb{C}^{(\mathbb{T}_n)}\|}\right]}{\mathbb{E}\left[e^{-\mu\sqrt{n}\|\mathbb{C}^{(\mathbb{T}_n)}\|}\right]} \longrightarrow \frac{\mathbb{E}\left[\Lambda(\mathbf{e})e^{-\alpha\|\mathbf{e}\|}\right]}{\mathbb{E}\left[e^{-\alpha\|\mathbf{e}\|}\right]}$$

which itself implies that

$$\mathbb{E}\left[\Lambda\left(\mathbb{C}^{(\mathbf{T}_n)}\right)\right] \longrightarrow \mathbb{E}\left[\Lambda(\mathbf{e}_\alpha)\right]$$

where the law ρ_α of $\mathbf{e}_\alpha$ satisfies

$$\frac{d\rho_\alpha}{d\rho_0}(f) = \frac{e^{-\alpha\|f\|}}{\int e^{-\alpha\|g\|}d\rho_0(g)} ,$$

with ρ_0 denoting the law of $\mathbf{e}$. In particular, $\mathbb{C}^{(\mathbf{T}_n)}$ converges in the topology of uniform convergence to $\mathbf{e}_\alpha$. Together with Proposition 3.2, the convergence of $\mathbb{C}^{(\mathbf{T}_n)}$ to $\mathbf{e}_\alpha$ implies the convergence in distribution of the corresponding real tree $\mathcal{T}_n^\mu$ to the continuum tree $\mathcal{T}_\alpha$ encoded by $\mathbf{e}_\alpha$, in the Gromov-Hausdorff sense. Let us now discuss how to get the improvement to the Gromov-Hausdorff-Prokhorov (GHP) sense of convergence.

For this improvement, let us first obtain some bounds in the deterministic setting. Fix some tree T_n of size n with the standard graph distance d_{gr} and consider the corresponding real tree $\mathcal{T}_n$ generated by its contour function with the distance function $d_{\mathcal{T}_n}$ assigning unit length to every edge. The discrete tree can be embedded in the real one isometrically: see for example [16, Example 4.9] for this standard construction. Further, denote by $a\mathcal{T}_n$ the real tree with distances scaled by a. Below, we will compute the Prokhorov distance between certain measures defined on $a\mathcal{T}_n$. When we say the push-forward of some measure originally defined on the discrete tree T_n, it is by the composition of these two maps to the real tree $a\mathcal{T}_n$.

Given T_n, recall that we write $\deg_{T_n}(v)$ for the degree of the vertex $v \in T_n$. The degree-biased measure $\sigma^*_{T_n}$ on T_n is the measure defined on the vertex set $V(T_n)$ of T_n satisfying

$$\sigma^*_{T_n}(v) = \frac{\deg_{T_n}(v)}{\sum\limits_{v \in T_n} \deg_{T_n}(v)} = \frac{\deg_{T_n}(v)}{2(n-1)} .$$

First, we bound the Prokhorov distance $d_{\mathrm{Prok}}(\mathrm{Leb}_{\mathcal{T}_n}, \sigma^*_{\mathcal{T}_n})$ between the push-forward of the degree-biased measure $\sigma^*_{\mathcal{T}_n}$ viewed as a measure on $a\mathcal{T}_n$ and the push-forward of Lebesgue measure $\mathrm{Leb}_{\mathcal{T}_n}$ from $[0,1]$ to $a\mathcal{T}_n$. For this purpose, first note that $2(n-1)$ is two times the number of edges of T_n. As each edge in the tree T_n corresponds to one up and one down step in the corresponding contour $\mathbb{C}^{(T_n)}$, we will call the set of points in the real tree $\mathcal{T}_n$ (equivalently in its rescaled version) induced by those two steps an *edge* as well. Hence the measure $\mathrm{Leb}^*_{\mathcal{T}_n}$ assigns mass $1/(n-1)$ to each edge of $a\mathcal{T}_n$.

For a metric space (E',d') and $U \subset E'$, recall the ε neighborhood of U is defined as $U^\varepsilon = \{x : d'(x,U) < \varepsilon\}$. It follows that for all $U \subset a\mathcal{T}_n$, $\mathrm{Leb}_{\mathcal{T}_n}(U) \leq \sigma^*_{T_n}(U^{\frac{a}{2}}) + \frac{a}{2}$ with the same inequality valid for $\mathrm{Leb}^*_{\mathcal{T}_n} \leftrightarrow \sigma^*_{\mathcal{T}_n}$. Hence

14

$$d_{\mathrm{Prok}}(\mathrm{Leb}^*_{\mathscr{T}_n}, \sigma^*_{\mathscr{T}_n}) \le \frac{a}{2}. \tag{3.1}$$

In a second step, we compare $\sigma^*_{\mathscr{T}_n}$ with the push-forward of the uniform measure on T_n to $a\mathscr{T}_n$ which we denote by $\sigma_{\mathscr{T}_n}$. Note that both the uniform and size-biased measures defined on T_n are supported on the vertices and denote the image of the vertex set $V(T_n)$ of T_n in $a\mathscr{T}_n$ under the aforementioned canonical embedding by $V(a\mathscr{T}_n)$. Fix any subset $S \subset V(a\mathscr{T}_n)$ and set $S_+ := \{v \in V(a\mathscr{T}_n) : d(v, S) \le a\}$. Writing $e(S)$ for the set of edges having at least one end-point in S, one has $e(S) \ge |S| - 1$, which can easily be seen by counting the edge that connects every $v \in S$ to its ancestor, except possibly the root. Defining the degree $\deg_{\mathscr{T}_n}(v)$ of a vertex $v \in V(a\mathscr{T})$ as the degree of its pre-image in the discrete tree T_n we have

$$\sigma^*_{\mathscr{T}_n}(S_+) = \sum_{v \in S_+} \sigma^*_{\mathscr{T}_n}(v) = \frac{\sum\limits_{v \in S_+} \deg_{\mathscr{T}_n}(v)}{2(n-1)} \ge \frac{2|e(S)|}{2(n-1)} \ge \frac{|S| - 1}{n} = \sigma_{\mathscr{T}_n}(S) - \frac{1}{n}$$

Note that S_+ contains the same vertices as the $(1 + \delta)a$-neighborhood of S for some $\delta < 1$ small. Hence the last inequality implies

$$\sigma_{\mathscr{T}_n}(S) \le \sigma^*_{\mathscr{T}_n}\big(S^{(1+\delta)a}\big) + \frac{1}{n}.$$

Then for $(1 + \delta)a \ge 1/n$, we have

$$\sigma_{\mathscr{T}_n}(S) \le \sigma^*_{\mathscr{T}_n}\big(S^{(1+\delta)a}\big) + (1 + \delta)a$$

and for $(1 + \delta)a < 1/n$, we have

$$\sigma^*_{\mathscr{T}_n}\big(S^{(1+\delta)a}\big) \le \sigma^*\big(S^{1/n}\big) \implies \sigma_{\mathscr{T}_n}(S) \le \sigma^*_{\mathscr{T}_n}\big(S^{1/n}\big) + \frac{1}{n}.$$

Furthermore, $\sigma^*_{\mathscr{T}_n}(S) \le \sigma_{\mathscr{T}_n}(S^{(1+\delta)a})$ trivially holds. Hence we conclude that

$$d_{\mathrm{Prok}}\big(\sigma^*_{\mathscr{T}_n}, \sigma_{\mathscr{T}_n}\big) \le \max\left((1 + \delta)a, \frac{1}{n}\right). \tag{3.2}$$

It now follows by an application of triangle inequality to (3.1) and (3.2) that

$$d_{\mathrm{Prok}}\big(\mathrm{Leb}^*_{\mathscr{T}_n}, \sigma_{\mathscr{T}_n}\big) \le \frac{a}{2} + \max\left((1 + \delta)a, \frac{1}{n}\right). \tag{3.3}$$

Finally, using Proposition 3.3, we have for any sequence $\mathscr{T}_n \to \mathscr{T}$,

$$d_{\mathrm{Prok}}\big(\mathrm{Leb}^*_{\mathscr{T}_n}, \mathrm{Leb}^*_{\mathscr{T}}\big) \le d_{GHP}\big(\mathscr{T}_n, \mathscr{T}\big) \le \frac{3}{2}\left\|d_{C^{(\mathscr{T}_n)}} - d_{C^{(\mathscr{T})}}\right\| \le 6\left\|C^{(\mathscr{T}_n)} - C^{(\mathscr{T})}\right\|. \tag{3.4}$$

Together with the GH convergence, (3.3) with a chosen to be $1/\sqrt{n}$ and (3.4) imply that

$$\left(\mathcal{T}^\mu_n, \frac{d_{\mathcal{T}^\mu_n}}{\sqrt{n}}, \sigma_{\mathcal{T}^\mu_n}\right) \xrightarrow{d} (\mathcal{T}_\alpha, d_\alpha, \sigma_\alpha)$$

as $n \to \infty$ in the GHP topology. Finally, using the canonical embedding of the discrete tree T_n into its corresponding real tree $\mathscr{T}_n$ one has

$$d_{GHP}\left(\left(T_n, \frac{d_{gr}}{\sqrt{n}}, \sigma_n\right), \left(\mathscr{T}_n, \frac{d_{\mathscr{T}_n}}{\sqrt{n}}, \sigma_{\mathscr{T}_n}\right)\right) \le \frac{1}{\sqrt{n}}.$$

Taking $T_n = \mathbf{T}_n$ and $\mathscr{T}_n = \mathcal{T}^\mu_n$ in the above display, the last two displays together imply the desired convergence in the statement of the theorem and conclude the proof. $\square$

15

3.3 Scaling limit of the width in the Brownian regime

In this section, we prove Theorem 1.3. Recall for any function f we denote its supremum value by $\|f\|$. The key result is that if we consider the uniform plane tree $\mathbb{T}_n$ on n vertices, and denote the corresponding height by H_n and width by W_n, then one has the joint convergence (see e.g., [15, Corollary 2.5.1])

$$\left(\frac{H_n}{\sqrt{n}}, \frac{W_n}{\sqrt{n}}\right) \xrightarrow{d} \left(\|\mathbf{e}\|, \frac{1}{2}\|\ell\|\right), \tag{3.5}$$

where $(\ell(x) : x \geq 0)$ is the *local-time process* of the standard Brownian excursion $(\mathbf{e}(t) : 0 \leq t \leq 1)$, defined by

$$\int_0^a \ell(x) dx = \int_0^1 \mathbb{1}_{[0,a]}(\mathbf{e}(s)) ds, \tag{3.6}$$

for any $a > 0$.

Proof of Theorem 1.3. Assume that $\mu\sqrt{n} \to \alpha \in \mathbb{R}$, and note that by the definition of our measure (1.3), for any bounded continuous function $F : \mathbb{R} \to \mathbb{R}$, one has

$$\mathbb{E}\left[F\left(\frac{w(\mathbf{T}_n)}{\sqrt{n}}\right)\right] = \frac{\mathbb{E}\left[F\left(\frac{W_n}{\sqrt{n}}\right)\exp\left(-\mu\sqrt{n}\frac{H_n}{\sqrt{n}}\right)\right]}{\mathbb{E}\left[\exp\left(-\mu\sqrt{n}\frac{H_n}{\sqrt{n}}\right)\right]}.$$

Observe that as $n \to \infty$, if $\mu \geq 0$, using (3.5) we immediately obtain as a consequence of dominated convergence theorem that

$$\mathbb{E}\left[F\left(\frac{w(\mathbf{T}_n)}{\sqrt{n}}\right)\right] \to \frac{\mathbb{E}\left[F\left(\frac{1}{2}\|\ell\|\right)\exp(-\alpha\|\mathbf{e}\|)\right]}{\mathbb{E}\left[\exp(-\alpha\|\mathbf{e}\|)\right]}. \tag{3.7}$$

A uniform integrability argument similar to the one used in the proof of Theorem 1.1 gives also the same conclusion for $\mu < 0$ satisfying $\mu\sqrt{n} \to \alpha \in [0, \infty)$.

Recall the well known identity of Jeulin, (see e.g., [11, Eq. (2.8)])

$$\left(2\|\mathbf{e}\|, \frac{1}{2}\|\ell\|\right) \stackrel{d}{=} \left(\int_0^1 \frac{ds}{\tilde{\mathbf{e}}(s)}, \|\tilde{\mathbf{e}}\|\right), \tag{3.8}$$

where $(\tilde{\mathbf{e}}(s) : 0 \leq 1 \leq 1)$ is distributed as another standard Brownian excursion.

Using the above, recalling the definition of B_α from the statement of Theorem 1.3, the RHS of (3.7) can be also written as,

$$\frac{\mathbb{E}\left[F(\|\tilde{\mathbf{e}}\|)\exp\left(-\frac{\alpha}{2}\int_0^1 \frac{ds}{\tilde{\mathbf{e}}(s)}\right)\right]}{\mathbb{E}\left[\exp\left(-\frac{\alpha}{2}\int_0^1 \frac{ds}{\tilde{\mathbf{e}}(s)}\right)\right]} = \mathbb{E}\left[F(\|B_\alpha\|)\right],$$

as claimed, where we recall the definition of the process B_α with law ξ_α from (1.5). $\qquad\square$

Remark 3.4 (Local times of max tilted excursions)**.** It is in fact possible to interpret the RHS of (3.7) as the maximum 'local time' of a random excursion on $[0,1]$. Namely, consider the random excursion $(\mathbf{e}_\alpha(t) : 0 \leq t \leq 1)$ with law ρ_α from Theorem 1.1, and recall that it has a Radon-Nikodým derivative with respect to the standard Brownian excursion given by

$$f \mapsto \frac{e^{-\alpha\|f\|}}{\int e^{-\alpha\|g\|} d\rho_0(g)},$$

for any $f \in C[0,1]$, where ρ_0 denotes the law of a standard Brownian excursion $(\mathbf{e}(t) : 0 \leq t \leq 1)$. Consider the local time process $(\ell(x); x \geq 0)$ of $\mathbf{e}(t)$ as defined in (3.6). We aim to define an analogous process $(\ell_\alpha(x) : x \geq 0)$ for the random excursion $\mathbf{e}_\alpha$. Observe that for any $0 < x_1 < x_2 < \cdots < x_m$, one can define the finite dimensional distributions as,

$$\mathbb{P}\left(\ell_\alpha(x_i) \in B_i, 1 \leq i \leq m\right) := \frac{\mathbb{E}\left[\mathbb{1}_{\{\ell(x_i) \in B_i, 1 \leq i \leq m\}} e^{-\alpha\|\mathbf{e}\|}\right]}{\mathbb{E}\left[e^{-\alpha\|\mathbf{e}\|}\right]},$$

16

where the B_i are arbitrary Borel subsets of $\mathbb{R}$. The Kolmogorov consistency theorem, and the existence of the process $(\ell(x) : x \geq 0)$ for the standard Brownian excursion $(\mathbf{e}(t) : 0 \leq t \leq 1)$ then guarantees the existence of a *unique* process $(\ell_\alpha(x) : x \geq 0)$ with finite dimensional distributions as given in the last display. Furthermore, the identity $\int_0^a \ell(x) dx = \int_0^1 \mathbb{1}_{[0,a]}(\mathbf{e}(s)) ds$ implies a similar identity,

$$\int_0^a \ell_\alpha(x) dx = \int_0^1 \mathbb{1}_{[0,a]}(\mathbf{e}_\alpha(s)) ds,$$

which means we can indeed interpret the RHS of (3.7) as the maximum local time of the random excursion $(\mathbf{e}_\alpha(t) : 0 \leq t \leq 1)$.

4   The partition function, and asymptotics of the height in the non-Brownian regime

In this section we prove all the results regarding the behaviour of the height in the non-Brownian regime, i.e., $\mu \gg 1/\sqrt{n}$. These comprise the Theorems 1.4, 1.5, 1.6, 1.8, and 1.11. The main ingredient for all these proofs is the following detailed result characterizing the asymptotic behaviour of the partition function as defined in (1.4). Recall the function $\lambda_x(t) = xt + \log(1 + \tan^2(\pi/t))$ and its unique minimizer t_x from (1.6).

Theorem 4.1. *Let $\mu = \mu_n$ be such that $\mu \gg 1/\sqrt{n}$. Then the partition function Z_n as defined in (1.4) behaves according to the following regimes.*

(1) *If $1/\sqrt{n} \ll \mu \ll n^{1/4}$, we have*

$$Z_n = (1 + o(1)) 4^n e^\mu (e^\mu - 1) \frac{\pi^{5/6}}{2^{1/3} 3^{1/2}} \frac{\mu^{1/3}}{n^{5/6}} \exp\left(-3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}\right).$$

(2) *If $(\log^{3/2} n)/\sqrt{n} \ll \mu \ll n$, we have*

$$Z_n = (1 + o(1)) 4^n e^\mu (e^\mu - 1) \frac{\mu}{2n} e^{-n\lambda_{\mu/n}(t_{\mu/n})} \sum_{m \in \mathbb{Z}} \exp\left(-\left(\frac{3}{2} + o(1)\right)\left(\frac{\mu^4}{2\pi^2 n}\right)^{1/3}(m - t_{\mu/n})^2\right).$$

(3) *If $n^{1/4} \ll \mu$, we have*

$$Z_n = (1 + o(1)) 4^n e^{2\mu} \sum_{m = \max\{\lfloor t_{\mu/n} \rfloor, 3\}}^{\lceil t_{\mu/n} \rceil} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} e^{-n\lambda_{\mu/n}(m)}.$$

Remark 4.2. Some of the regimes from Theorem 4.1 overlap, but it can be easily verified that the formulas are coherent with each other. More precisely, the first two agree thanks to properties of Riemann's sum and the last two are directly compared in the proof of Proposition E.2.

Proof. All of these results are covered in the different sections of the appendix. First of all, Lemma B.2 states that

$$Z_n = (1 + o(1)) 4^n e^\mu (e^\mu - 1) W_n$$

where W_n is defined in (B.3). Now, the three regimes directly follow from Theorem D.1, Proposition E.1, and Proposition E.2 respectively. $\square$

We now use this theorem to prove the asymptotic behaviour of the height of the trees as covered in Theorem 1.4, 1.5, 1.6, 1.8, and 1.11, respectively in Section 4.1, 4.2, 4.2, 4.2, and 4.3. Before doing, it is worth highlighting the results from Proposition C.1 regarding λ_x and t_x which are repeatedly used throughout this section. In particular, whenever $\mu \ll n$, the proposition states that

$$t_{\mu/n} = \left(\frac{2\pi^2 n}{\mu}\right)^{1/3} + o(1)$$

and that

$$n\lambda_{\mu/n}(t_{\mu/n}) = 3\left(\frac{\pi^2 \mu^2 n}{n}\right)^{1/3} + O\left(\left(\frac{\mu^4}{n}\right)^{1/3}\right).$$

4.1 First order behaviour

This section focuses on proving Theorem 1.4 and provides a first and straightforward application of Theorem 4.1.

Proof of Theorem 1.4. To prove this result, we simply apply a Chernoff bound (or, exponential Markov inequality) with the formula for the partition function from Theorem 4.1. In fact, we do not exactly prove the convergence stated in the theorem, but rather that

$$\frac{h(\mathbf{T}_n)}{t_{\mu/n}} \longrightarrow 1\,,$$

where $t_{\mu/n}$ is the minimum of $\lambda_{\mu/n}$ as defined in (1.6). Note that $t_{\mu/n}$ behaves as $(2\pi^2 n/\mu)^{1/3}$ thanks to Proposition C.1 (**1**). We start by fixing $\epsilon > 0$ and proving that

$$\mathbb{P}\left(\frac{h(\mathbf{T}_n)}{t_{\mu/n}} \geq 1 + \epsilon\right) \longrightarrow 0$$

for all considered regimes of μ. Observe that, for any (fixed) $\gamma \in (0, 1)$, by a Chernoff bound we have

$$\mathbb{P}\left(\frac{h(\mathbf{T}_n)}{t_{\mu/n}} \geq 1 + \epsilon\right) \leq \mathbb{E}\left[e^{\gamma\mu h(\mathbf{T}_n)}\right] e^{-\gamma(1+\epsilon)\mu t_{\mu/n}} = \frac{e^{-\gamma(1+\epsilon)\mu t_{\mu/n}} Z_n^{(1-\gamma)\mu}}{Z_n^\mu}\,.$$

Thus, if $1/\sqrt{n} \ll \mu \ll n^{1/4}$, Theorem 4.1 (**1**) tells us that

$$\mathbb{P}\left(\frac{h(\mathbf{T}_n)}{t_{\mu/n}} \geq 1 + \epsilon\right) \sim e^{-\gamma(1+\epsilon)\mu t_{\mu/n}} e^{-\gamma\mu} \frac{e^{(1-\gamma)\mu} - 1}{e^\mu - 1} (1-\gamma)^{1/3} \exp\left(3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} - 3\left(\frac{\pi^2(1-\gamma)^2\mu^2 n}{4}\right)^{1/3}\right).$$

Further use Proposition C.1 (**1**) to see that

$$\mu t_{\mu/n} = 2\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} + o(1)\,,$$

so that the previous bound can transform into

$$\mathbb{P}\left(\frac{h(\mathbf{T}_n)}{t_n} \geq 1 + \epsilon\right) \leq (1 + o(1)) \exp\left(-\left[3(1-\gamma)^{2/3} + 2\gamma(1+\epsilon) - 3\right]\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right). \tag{4.1}$$

Using that ϵ is fixed and that, for γ small enough, we have

$$\left[3(1-\gamma)^{2/3} + 2\gamma(1+\epsilon) - 3\right] \sim 2\epsilon\gamma > 0\,,$$

we can find some $\gamma > 0$ such that the right-hand side converges to 0, thus proving the upper bound when $1/\sqrt{n} \ll \mu \ll n^{1/4}$. Using the same method when $\mu \asymp n^{1/4}$ but with Theorem 4.1 (**2**) and observing that the sum appearing in the formula of Z_n is of order 1, we again obtain the bound (4.1). This leads to the same conclusion as for the case $\mu \ll n^{1/4}$. Finally, for the case $\mu \gg n^{1/4}$, observe again that the same method but applied with Theorem 4.1 (**3**) works, thanks to the fact that $(\mu^4/n)^{1/3} \ll (\mu^2 n)^{1/3}$.

To conclude the proof of the theorem, note that a Chernoff bound applied to the lower tail of the height also leads to

$$\mathbb{P}\left(\frac{h(\mathbf{T}_n)}{t_{\mu/n}} \leq 1 - \epsilon\right) \leq (A + o(1))e^{2\gamma\mu} \exp\left(-\left[3(1+\gamma)^{2/3} - 2\gamma(1-\epsilon) - 3\right]\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right),$$

where $A > 0$ is some positive constant and the convergence to 0 follows from the same argument as before and since $\mu \ll (\mu^2 n)^{1/3}$. □

4.2 Fluctuations and deviations from the mean

In this section, we extend the results from Section 4.1 to the second order behaviour and the large deviation of the height of T_n. We heavily rely on the asymptotic behaviour of the partition function from Theorem 4.1.

4.2.1 Second order fluctuations

In this section we provide the proofs of Theorems 1.5 and 1.6.

Proof of Theorem 1.5. Let us denote

$$\sigma = \sqrt{\frac{1}{3}\left(\frac{2\pi^2 n}{\mu^4}\right)^{1/3}}.$$

For any $t \in \mathbb{R}$, consider the moment-generating function of the shifted and scaled height

$$f(t) = f_n(t) = \mathbb{E}\left[e^{t\frac{h(\mathbf{T}_n)-(2\pi^2 n/\mu)^{1/3}}{\sigma}}\right] = e^{-t\frac{(2\pi^2 n/\mu)^{1/3}}{\sigma}}\frac{Z_n^{\mu-t/\sigma}}{Z_n^{\mu}}.$$

Now, applying Theorem 4.1 (**1**) on the numerator and the denominator of the ratio of the partition functions above, we obtain that

$$f(t) \sim e^{-t/\sigma}\frac{e^{\mu-t/\sigma}-1}{e^{\mu}-1}\left(1-\frac{t}{\sigma\mu}\right)^{1/3}\exp\left(-t\frac{(2\pi^2 n/\mu)^{1/3}}{\sigma}-3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\left[\left(1-\frac{t}{\sigma\mu}\right)^{2/3}-1\right]\right).$$

Observe first that $\sigma \asymp (n/\mu^4)^{1/6}$ so that $\sigma \gg 1$ but also $\sigma\mu \asymp (\mu^2 n)^{1/6} \gg 1$, implying that all the terms outside of the final exponential are of order 1. Combining this with a Taylor expansion of the term $(1-t/(\sigma\mu))^{2/3}$ leads to

$$f(t) \sim \exp\left(-t\frac{(2\pi^2 n/\mu)^{1/3}}{\sigma}-3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\left[-\frac{2}{3}\frac{t}{\sigma\mu}-(1+o(1))\frac{1}{9}\frac{t^2}{\sigma^2\mu^2}\right]\right)$$

$$= \exp\left((1+o(1))\frac{1}{2}\left[\frac{1}{3}\left(\frac{2\pi^2 n}{\mu^4}\right)^{1/3}\right]\frac{t^2}{\sigma^2}\right).$$

To conclude the proof, simply note that σ is exactly defined so that $f(t) \sim e^{t^2/2}$, as required for the convergence in distribution to the standard normal. $\square$

Proof of Theorem 1.6. Recall $\lambda_x(t) = xt + \log(1+\tan^2(\pi/t))$ and its unique minimizer t_x from (1.6). For this proof, we can directly use Theorem 4.1 (**2**) with the moment generating function of $h(\mathbf{T}_n)$ to obtain

$$\mathbb{E}\left[e^{sh(\mathbf{T}_n)}\right] = \frac{Z_n^{\mu-s}}{Z_n^{\mu}} \sim e^{-2s}\left(1-\frac{s}{\mu}\right)e^{n(\lambda_{\mu/n}(t_{\mu/n})-\lambda_{(\mu-s)/n}(t_{(\mu-s)/n}))}\frac{S_{(\mu-s)/n}}{S_{\mu/n}},$$

where for any $x > 0$,

$$S_x = \sum_{m\in\mathbb{Z}}\exp\left(-\left(\frac{3}{2}+o(1)\right)\left(\frac{n^3 x^4}{2\pi^2}\right)^{1/3}(m-t_x)^2\right),$$

corresponding to the sum in the asymptotics from Theorem 4.1 (**2**). It is worth observing that, since we consider $x \in \{\mu/n, (\mu-s)/n\}$, for both value it holds that $x \sim \gamma n^{-3/4}$. Moreover, Proposition C.1 (**1**) tells us that

$$t_{(\mu-s)/n} = \left(\frac{2\pi^2 n}{\mu-s}\right)^{1/3}+o(1) = \left(\frac{2\pi^2 n}{\mu}\right)^{1/3}\left(1+\frac{1}{3}\frac{s}{\mu}+o\left(\frac{1}{\mu}\right)\right)+o(1) = t_{\mu/n}+\left(\frac{2\pi^2}{\gamma^4}\right)^{1/3}\frac{s}{3}+o(1),$$

where we used that $\mu \sim \gamma n^{1/4}$, that these formulae work for arbitrary s, so also for $s = 0$, and that $t_{\mu/n} = (2\pi^2 n/\mu)^{1/3}+o(1)$, which is itself a direct consequence of Proposition C.1 (**1**). In a similar manner, using Proposition C.1 (**2**), we obtain that

$$n\lambda_{(\mu-s)/n}(t_{(\mu-s)/n}) = 3\left(\frac{\pi^2(\mu-s)^2 n}{4}\right)^{1/3}+\frac{1}{6}\left(\frac{(\mu-s)^4}{(2\pi)^4 n}\right)^{1/3}+o(1)$$

$$= 3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\left[1-\frac{2}{3}\frac{s}{\mu}-\frac{1}{9}\frac{s^2}{\mu^2}+o\left(\frac{1}{\mu^2}\right)\right]+\frac{1}{6}\left(\frac{\gamma}{2\pi}\right)^{4/3}+o(1)$$

$$= n\lambda_{\mu/n}(t_{\mu/n}) - s\left(\frac{2\pi^2 n}{\mu}\right)^{1/3} - \left(\frac{\pi}{2\gamma^2}\right)^{2/3}\frac{s^2}{3} + o(1),$$

where we used again that the formula holds for $s = 0$. Plugging the previous results in the original formula, along with the fact $s/\mu \ll 1$, we obtain that

$$\mathbb{E}\left[e^{sh(\mathbf{T}_n)}\right] \sim \exp\left(-2s + s\left(\frac{2\pi^2 n}{\mu}\right)^{1/3} + \left(\frac{\pi}{2\gamma^2}\right)^{2/3}\frac{s^2}{3}\right)\frac{S_{(\mu-s)/n}}{S_{\mu/n}},$$

where

$$S_{\mu/n} \sim \sum_{m\in\mathbb{Z}}\exp\left(-\frac{3}{2}\left(\frac{\gamma^4}{2\pi^2}\right)^{1/3}\left(m - \left(\frac{2\pi^2 n}{\mu}\right)^{1/3}\right)^2\right)$$

and

$$S_{(\mu-s)/n} \sim \sum_{m\in\mathbb{Z}}\exp\left(-\frac{3}{2}\left(\frac{\gamma^4}{2\pi^2}\right)^{1/3}\left(m - \left(\frac{2\pi^2 n}{\mu}\right)^{1/3} - \left(\frac{2\pi^2}{\gamma^4}\right)^{1/3}\frac{s}{3}\right)^2\right).$$

To conclude this proof, observe that the last term within the exponential simplifies as

$$\frac{3}{2}\left(\frac{\gamma^4}{2\pi^2}\right)^{1/3}\left(m - \left(\frac{2\pi^2 n}{\mu}\right)^{1/3} - \left(\frac{2\pi^2}{\gamma^4}\right)^{1/3}\frac{s}{3}\right)^2 = \frac{3}{2}\left(\frac{\gamma^4}{2\pi^2}\right)^{1/3}\left(m - \left(\frac{2\pi^2 n}{\mu}\right)^{1/3}\right)^2$$

$$- ms + s\left(\frac{2\pi^2 n}{\mu}\right)^{1/3} + \left(\frac{\pi}{2\gamma^2}\right)^{2/3}\frac{s^2}{3},$$

where we see that the first term exactly corresponds to the term in the exponential of the sum of $S_{\mu/m}$ and the last two terms are also the two terms in the exponential of the moment generating function. Doing all the subsequent simplifications lead to

$$\mathbb{E}\left[e^{sh(\mathbf{T}_n)}\right] \sim \frac{\displaystyle\sum_{m\in\mathbb{Z}}e^{(m-2)s}\exp\left(-\frac{3}{2}\left(\frac{\gamma^4}{2\pi^2}\right)^{1/3}\left(m - \left(\frac{2\pi^2 n}{\mu}\right)^{1/3}\right)^2\right)}{\displaystyle\sum_{m\in\mathbb{Z}}\exp\left(-\frac{3}{2}\left(\frac{\gamma^4}{2\pi^2}\right)^{1/3}\left(m - \left(\frac{2\pi^2 n}{\mu}\right)^{1/3}\right)^2\right)}.$$

The claimed convergence now follows by shifting the sum $m \in \mathbb{Z}$ by $m_{\mu/n} = \lfloor t_{\mu/n}\rfloor$ and recalling that Proposition C.1 (1) states that $t_{\mu/n} = (2\pi^2 n/\mu)^{1/3} + o(1)$. □

4.2.2 Large deviations

In this section we provide the proof of Theorem 1.8. Let us begin by stating the Gärtner-Ellis theorem [13, Theorem 2.3.6], as this is the key ingredient of our proof. We don't define the large deviations principle (LDP), or the notion of good rate functions here, and refer the reader for these standard concepts to [13, Chapter 1].

Theorem 4.3 (Gärtner-Ellis theorem). *Consider a sequence of random variables $(X_n)_{n\geq 1}$ and a sequence $(a_n)_{n\geq 1}$ with $a_n \gg 1$. Assume for every $\lambda \in \mathbb{R}$ the limit*

$$\lim_{n\to\infty}\frac{1}{a_n}\log\mathbb{E}\left[\exp\left(\lambda a_n X_n\right)\right] = \Lambda(\lambda), \tag{4.2}$$

exists point-wise as a convex function of λ, possibly assuming an extended real value for some λ, with the origin lying inside the interior of the domain $\mathcal{D}_\Lambda := \{\lambda \in \mathbb{R} : \Lambda(\lambda) < \infty\}$. Let Λ further satisfy the following:

- *Interior of $\mathcal{D}_\Lambda$ is non empty.*

- Λ *is differentiable at all points of the interior of* $\mathscr{D}_\Lambda$.

- $\Lambda(\cdot)$ *is steep, namely,* $\lim_{n\to\infty} |\Lambda'(\Lambda_n)| = \infty$ *whenever* λ_n *is a sequence of interior points of* $\mathscr{D}_\Lambda$ *converging to some boundary point of the interior of* $\mathscr{D}_\Lambda$.

Then the sequence $(X_n)_{n\geq 1}$ *satisfies a LDP with speed* a_n *and with good rate function* Λ^*, *the* Fenchel-Legendre *transform of* Λ, *defined as,*

$$\Lambda^*(x) = \sup_{\lambda \in \mathbb{R}} \left\{ \lambda x - \Lambda(\lambda) \right\}.$$

We are now ready to provide the proof of Theorem 1.8.

Proof of Theorem 1.8. Using Proposition C.1 (**1**), we see that $t_{\mu/n}$, the minimum of $\lambda_{\mu/n}$ as defined in (1.6), is of order $(2\pi^2 n/\mu)^{1/3}$. Thus, it is enough to prove that $h(\mathbf{T}_n)/t_{\mu/n}$ satisfies the claimed large deviation principle with speed $(\mu^2 n)^{1/3}$. Thus, the main technical part of the proof reduces to proving that

$$\Lambda(\gamma) = \lim_{n\to\infty} \frac{1}{(\mu^2 n)^{1/3}} \log \mathbb{E}\left[e^{(\mu^2 n)^{1/3}\gamma h(\mathbf{T}_n)/t_{\mu/n}} \right]$$

exists and has the desired properties of Theorem 4.3. To do so, we start by characterizing the above limit when replacing $(\mu^2 n)^{1/3}$ with $\mu t_{\mu/n}$, which has the same order, thanks to Proposition C.1 (**1**).

Assume first that $1/\sqrt{n} \ll \mu \ll n^{1/4}$. Then, using the asymptotic behaviour of the partition function as provided in Theorem 4.1 (**1**), we see that, for any $\gamma < 1$

$$\mathbb{E}\left[e^{\mu\gamma h(T_n)} \right] = \frac{Z_n^{\mu(1-\gamma)}}{Z_n^\mu} \sim \frac{e^{\mu(1-\gamma)}(e^{\mu(1-\gamma)}-1)}{e^\mu(e^\mu-1)}(1-\gamma)^{2/3}\exp\left(-3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\left((1-\gamma)^{1/3}-1\right)\right).$$

Observing that $\mu \ll (\mu^2 n)^{1/3}$ here, it follows that

$$\frac{1}{\mu t_{\mu/n}} \log \mathbb{E}\left[e^{\mu\gamma h(\mathbf{T}_n)} \right] = -\frac{3}{\mu t_{\mu/n}}\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\left((1-\gamma)^{2/3}-1\right) + o(1)$$

$$= \frac{3}{2}\left(1-(1-\gamma)^{2/3}\right) + o(1).$$

Note that this formula also works when $\gamma = 1$ by replacing the numerator with $C_{n-1} \asymp 4^n/n^{3/2}$.

Assume now that $n^{1/4} \ll \mu \ll n$. Then, using the asymptotic behaviour of the partition function as provided in Theorem 4.1 (**2**), we see that, for any $\gamma < 1$

$$\mathbb{E}\left[e^{\mu\gamma h(\mathbf{T}_n)} \right] = \frac{Z_n^{\mu(1-\gamma)}}{Z_n^\mu} \sim \frac{e^{2\mu(1-\gamma)}}{e^{2\mu}}(1-\gamma)e^{-n\lambda_{\mu(1-\gamma)/n}(t_{\mu(1-\gamma)/n})+n\lambda_{\mu/n}(t_{\mu/n})}R,$$

where R is a fraction between two sums, both of them of order $e^{O((\mu^4/n)^{1/3})} = e^{o(\mu^2 n)^{1/3}}$. Furthermore, thanks to Proposition C.1 (**2**), we know that

$$-n\lambda_{\mu(1-\gamma)/n}(t_{\mu(1-\gamma)/n}) + n\lambda_{\mu/n}(t_{\mu/n}) = -3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\left((1-\gamma)^{2/3}-1\right) + O\left(\left(\frac{\mu^4}{n}\right)^{1/3}\right).$$

Using again that $(\mu^4/n)^{1/3} \ll (\mu^2 n)^{1/3} \asymp \mu t_{\mu/n}$, we obtain that

$$\frac{1}{\mu t_{\mu/n}} \log \mathbb{E}\left[e^{\mu\gamma h(T_n)} \right] = \frac{3}{2}\left(1-(1-\gamma)^{2/3}\right) + o(1), \tag{4.3}$$

and can further check that this formula still holds when $\gamma = 1$.

Defining

$$\Gamma(\gamma) = \frac{3}{2}\left(1-(1-\gamma)^{2/3}\right),$$

we note thus that the limit $\Gamma(\gamma)$ exists for any $\gamma \leq 1$, is convex, differentiable on $(-\infty, 1)$, and its derivative diverges to ∞ at 1. It is worth observing that, since Γ is convex, we can also deduce that $\Gamma(\gamma) = \infty$ whenever $\gamma > 1$, without having to compute it (which would otherwise require an understanding of the partition function when μ is negative). All of these observations imply that Γ satisfies the properties in Theorem 4.3. Thus, by Theorem 4.3, it follows that $(h(\mathbf{T}_n)/t_{\mu/n})_{n\geq 1}$ satisfies an LDP with respect to the good rate function that is the Fenchel-Legendre transform of $\Gamma(\gamma) = (3/2)(1 - (1-\gamma)^{2/3})$ and with speed $\mu t_{\mu/n}$.

Observe that defining $\Lambda(\gamma) := \lim_{n\to\infty} \frac{1}{(\mu^2 n)^{1/3}} \log\mathbb{E}\left[e^{(\mu^2 n)\gamma h(\mathbf{T}_n)/t_{\mu/n}} \right]$, we have the relation

$$\Lambda(\gamma) \sim \frac{(2\pi^2)^{1/3}}{\mu t_{\mu/n}} \log\mathbb{E}\left[\exp\left(\mu \frac{\gamma + o(1)}{(2\pi^2)^{1/3}} h(\mathbf{T}_n) \right) \right] \sim (2\pi^2)^{1/3} \Gamma\left(\frac{\gamma}{(2\pi^2)^{1/3}} \right). \tag{4.4}$$

We now compute the rate function $\Lambda^*(\cdot)$. For any $x \in \mathbb{R}$, we have

$$x = \Lambda'(\gamma) \iff x = \Gamma'\left(\frac{\gamma}{(2\pi^2)^{1/3}} \right) = \left(1 - \frac{\gamma}{(2\pi^2)^{1/3}} \right)^{-1/3} \iff \gamma = (2\pi^2)^{1/3}\left(1 - \frac{1}{x^3} \right).$$

This value of γ is defined for any $x > 0$ and is less than $(2\pi^2)^{1/3}$, thus $\gamma \in \mathcal{D}_\Lambda$. In that case, we see that the supremum in the definition of the Fenchel-Legendre transform of Λ is reached when γ is the above value, leading to

$$\Lambda^*(x) = x(2\pi^2)^{1/3}\left(1 - \frac{1}{x^3} \right) - \Lambda\left((2\pi^2)^{1/3}\left(1 - \frac{1}{x^3} \right) \right)$$

$$= (2\pi^2)^{1/3}\left[x - \frac{1}{x^2} - \frac{3}{2}\left(1 - \left(\frac{1}{x^3} \right)^{2/3} \right) \right]$$

$$= (2\pi^2)^{1/3}\left[x + \frac{1}{2x^2} - \frac{3}{2} \right].$$

This agrees with the rate function appearing in the statement of Theorem 1.8 when $x > 0$, while for $x \leq 0$, the supremum in the definition of Fenchel-Legendre transform is infinite, as follows from (4.3) and the relation (4.4) between Λ and Γ, agreeing with the claim of Theorem 1.8. $\square$

4.3 Bernoulli fluctuations

In this section we provide the proof Theorem 1.11 by splitting the result into two propositions, covering the regimes $\mu \ll n$ and $\mu = \Omega(n)$. We then combine them together to prove Theorem 1.11. Recall the function λ_x, its minimizer t_x as defined in (1.6), and the definition of δ_n and $m_x = \max\{\lfloor t_x \rfloor, 3\}$ from (1.7).

Proposition 4.4. *Let $\mu = \mu_n$ be such that $n^{1/4} \ll \mu \ll n$, and assume that δ_n admits a limit δ (possibly infinite). Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. Then, $h(\mathbf{T}_n) - m_{\mu/n} + 2$ converges in distribution to a Bernoulli random variable with parameter*

$$p_\delta = \frac{e^{-\delta}}{1 + e^{-\delta}}.$$

Proof. Fix some $\gamma \in \mathbb{R}$. Recall that Theorem 4.1 (**3**) tells us that, in the regime considered here, we have

$$Z_n = \left(1 + o(1)\right)4^n e^{2\mu} \sum_{m=\max\{\lfloor t_{\mu/n} \rfloor, 3\}}^{\lceil t_{\mu/n} \rceil} \frac{\tan^2\left(\frac{\pi}{m} \right)}{m} e^{-n\lambda_{\mu/n}(m)},$$

Further using that $t_{\mu/n}$ and $t_{(\mu+\gamma)/n}$ diverge here thanks to Proposition C.1 (**1**), we can remove the max from the lower bound of the sum and identify $m_{\mu/n}$ with $\lfloor t_{\mu/n} \rfloor$. Finally, observing that $\lambda_{x+y}(m) = ym + \lambda_x(m)$, we see that

$$\mathbb{E}\left[e^{\gamma h(\mathbf{T}_n)} \right] = \frac{Z_n^{\mu-\gamma}}{Z_n^\mu} \sim e^{-2\gamma} \frac{\sum_{m=\lfloor t_{(\mu-\gamma)/n} \rfloor}^{\lceil t_{(\mu-\gamma)/n} \rceil} \frac{\tan^2\left(\frac{\pi}{m} \right)}{m} e^{\gamma m - n\lambda_{\mu/n}(m)}}{\sum_{\lfloor t_{\mu/n} \rfloor}^{\lceil t_{\mu/n} \rceil} \frac{\tan^2\left(\frac{\pi}{m} \right)}{m} e^{-n\lambda_{\mu/n}(m)}}. \tag{4.5}$$

To prove the Proposition, it is enough to show that

$$\mathbb{E}\left[e^{\gamma(h(\mathbf{T}_n)-m_{\mu/n}+2)}\right] \to 1 - p_\delta + p_\delta e^\gamma. \tag{4.6}$$

Note that the sequence $t_{\mu/n} - \lfloor t_{\mu/n} \rfloor$ is bounded in $[0,1]$, and thus admits subsequential limits. Below we separate our analysis into three possible cases of these limits, and exhibit that the result claimed by the Proposition is coherent across these cases. Further, to simplify notation, we hide the subsequence notation and assume the sequence itself has the assumed behavior.

Case: $t_{\mu/n} - \lfloor t_{\mu/n} \rfloor \to 0$. Using equation (4.5), if $t_{\mu/n} - \lfloor t_{\mu/n} \rfloor$ converges to 0, Proposition C.1 (**3**) tells us that

$$n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor \pm 1) - n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor) \geq Cn\left(\frac{\mu}{n}\right)^{2/3}\frac{1}{\left(\frac{n}{\mu}\right)^{2/3}} = C\left(\frac{\mu^4}{n}\right)^{1/3} \gg 1,$$

with a similar inequality when applied to $\lfloor t_{(\mu-\gamma)/n} \rfloor$, implying that we can ignore the term different from $\lfloor t_{\mu/n} \rfloor$ in both sums. In that case, we directly see that

$$\mathbb{E}\left[e^{\gamma h(\mathbf{T}_n)}\right] \sim e^{\gamma(\lfloor t_{\mu/n} \rfloor - 2)}.$$

To see that this implies (4.6), using $t_{\mu/n}$ diverges, it suffices to show that $\delta_n \to \infty$. But this is straightforward since, either $t_{\mu/n}$ is an integer, and then $\lfloor t_{\mu/n} \rfloor = \lceil t_{\mu/n} \rceil$ so that $\delta_n = \mu \gg 1$, or $t_{\mu/n}$ is not an integer, and then $\lceil t_{\mu/n} \rceil = \lfloor t_{\mu/n} \rfloor + 1$, leading to

$$\delta_n = n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor + 1) - n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor) \geq C\left(\frac{\mu^4}{n}\right)^{1/3} \gg 1.$$

Case: $t_{\mu/n} - \lfloor t_{\mu/n} \rfloor \to 1$. Note that $t_{\mu/n} - \lfloor t_{\mu/n} \rfloor$ converges to 1 is equivalent to saying that $t_{\mu/n} - \lceil t_{\mu/n} \rceil$ converges to 0, and similar arguments as in the previous case apply, leading to

$$\mathbb{E}\left[e^{\gamma h(\mathbf{T}_n)}\right] \sim e^{\gamma(\lceil t_{\mu/n} \rceil - 2)}.$$

To see that this is desired result, observe again that if $t_{\mu/n}$ is an integer the result follows, and otherwise $-\delta_n \gg 1$, leading to $\delta = -\infty$, and the last display again agrees with (4.6).

Case: $\lim_{n\to\infty} t_{\mu/n} - \lfloor t_{\mu/n} \rfloor \in (0,1)$. Assume finally that $t_{\mu/n} - \lfloor t_{\mu/n} \rfloor$ converges to some number in $(0,1)$. In that regime, Proposition C.1 (**1**) implies that

$$t_{(\mu-\gamma)/n} = \left(\frac{2\pi^2 n}{\mu-\gamma}\right)^{1/3} + o(1) = \left(\frac{2\pi^2 n}{\mu}\right)^{1/3} + O\left(\left(\frac{n}{\mu^4}\right)^{1/3}\right) + o(1) = t_{\mu/n} + o(1),$$

which means that $\lfloor t_{\mu/n} \rfloor = \lfloor t_{(\mu-\gamma)/n} \rfloor$ for n large enough. It follows that

$$\mathbb{E}\left[e^{\gamma h(\mathbf{T}_n)}\right] \sim e^{-2\gamma}\frac{\sum_{m=\lfloor t_{\mu/n} \rfloor}^{\lceil t_{\mu/n} \rceil}\frac{\tan^2\left(\frac{\pi}{m}\right)}{m}e^{\gamma m - n\lambda_{\mu/n}(m)}}{\sum_{m=\lfloor t_{\mu/n} \rfloor}^{\lceil t_{\mu/n} \rceil}\frac{\tan^2\left(\frac{\pi}{m}\right)}{m}e^{-n\lambda_{\mu/n}(m)}}$$

$$\sim e^{\gamma(\lfloor t_{\mu/n} \rfloor - 2)}\frac{1 + e^\gamma e^{n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor) - n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor + 1)}}{1 + e^{n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor) - n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor + 1)}},$$

where we used that $\lfloor t_n \rfloor \sim \lceil t_n \rceil \gg 1$ to take the fraction out of the sum. But now, we observe that the definition of δ_n implies that

$$n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor) - n\lambda_{\mu/n}(\lfloor t_{\mu/n} \rfloor + 1) = -\delta_n,$$

leading to

$$\mathbb{E}\left[e^{\gamma h(\mathbf{T}_n)}\right] \sim e^{\gamma(\lfloor t_{\mu/n} \rfloor - 2)}\frac{1 + e^\gamma e^{-\delta_n}}{1 + e^{-\delta_n}},$$

implying (4.6). □

23

Proposition 4.5. *Let $\mu = \mu_n$ be such that $c_n = \mu/n$ converges to some $c \in (0,\infty]$ and assume that δ_n admits a limit δ (possibly infinite). Let $\mathbf{T}_n$ be a μ-height-biased tree of size n. Then, $h(\mathbf{T}_n) - m_{\mu/n} + 2$ converges in distribution to a Bernoulli random variable with parameter*

$$p_{c,\delta} = \frac{m_c \tan^2\left(\frac{\pi}{m_c+1}\right) e^{-\delta}}{(m_c+1)\tan^2\left(\frac{\pi}{m_c}\right) + m_c \tan^2\left(\frac{\pi}{m_c+1}\right) e^{-\delta}}.$$

Proof. The proof is very similar to that of Proposition 4.4, except that Proposition C.1 does not apply here. We start by applying Theorem 4.1 to see that

$$Z_n \sim 4^n e^{2\mu} \sum_{m=m_{\mu/n}}^{\lceil t_{\mu/n}\rceil} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} e^{-n\lambda_{\mu/n}(m)},$$

so that, using again $\lambda_{x+y}(m) = ym + \lambda_x(m)$, we obtain

$$\mathbb{E}\left[e^{\gamma h(\mathbf{T}_n)}\right] = \frac{Z_n^{\mu-\gamma}}{Z_n^{\mu}} \sim e^{-2\gamma} \frac{\sum_{m=m_{(\mu-\gamma)/n}}^{\lceil t_{(\mu-\gamma)/n}\rceil} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} e^{\gamma m - n\lambda_{\mu/n}(m)}}{\sum_{m=m_{\mu/n}}^{\lceil t_{\mu/n}\rceil} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} e^{-n\lambda_{\mu/n}(m)}}.$$

Observe now that, whenever $t_{\mu/n} \le 3$, then the two sums can be reduced to $m = 3$ since the possible other term (corresponding to the case $\lceil t_{(\mu-\gamma)n}\rceil = 3$, implying $t_{(\mu-\gamma)/n} \sim 3$) is negligible. This leads to the desired result, since in that case $\delta_n = \mu$ and so $\delta = \infty$ with $p_{c,\delta} = 0$. We assume from now on that $t_{\mu/n} > 3$ so that $m_{\mu/n} = \lfloor t_{\mu/n}\rfloor$ and $m_{(\mu-\gamma)/n} = \lfloor t_{(\mu-\gamma)/n}\rfloor$ (since $x \mapsto t_x$ is decreasing).

Assume here that $t_{\mu/n}$ converges to some integer $\tilde{m}_n$, so that $t_{(\mu-\gamma)/n}$ does as well. Using that $\mu/n \sim c$ and the fact that λ_c is strictly convex, the previous two sums can be reduced to $m = \tilde{m}_n$ since the other terms decrease exponentially faster. In that case, we see that

$$\mathbb{E}\left[e^{\gamma h(\mathbf{T}_n)}\right] = \frac{Z_n^{\mu-\gamma}}{Z_n^{\mu}} \sim e^{\gamma(\tilde{m}_n - 2)}.$$

If $t_{\mu/n}$ is an integer, then $\delta_n = \mu$ and so $\delta = +\infty$. Otherwise, we see that

$$\delta_n = n\lambda_{\mu/n}(\lceil t_{\mu/n}\rceil) - n\lambda_{\mu/n}(\lfloor t_{\mu/n}\rfloor),$$

which is positive if $t_{\mu/n}$ converges to $\lfloor t_{\mu/n}\rfloor$ and negative if it converges to $\lceil t_{\mu/n}\rceil$. Moreover, using that $\mu/n \sim c$ and that λ_c is strictly convex, we see that $|\delta_n| \gg 1$. Thus, either $\delta = +\infty$, in which case $\tilde{m}_n = \lfloor t_{\mu/n}\rfloor$ and so $h(\mathbf{T}_n) \to \lfloor t_{\mu/n}\rfloor - 2$ and $p_{c,\delta} = 0$, or $\delta = -\infty$, in which case $\tilde{m}_n = \lceil t_{\mu/n}\rceil$ and so $h(\mathbf{T}_n) \to \lfloor t_{\mu/n}\rfloor - 1$ and $p_{c,\delta} = 1$. In all cases, the desired convergence is verified.

We assume now that $t_{\mu/n} - \lfloor t_{\mu/n}\rfloor$ converges to some $r \in (0,1)$. Observe that, for γ small enough and n large enough, this implies that $\lfloor t_{(\mu-\gamma)/n}\rfloor = \lfloor t_{\mu/n}\rfloor = m_c$. Recalling that we assume that $t_{\mu/n}$ and thus also $t_{(\mu-\gamma)/n}$ are both larger than 3, we see that

$$\mathbb{E}\left[e^{\gamma h(\mathbf{T}_n)}\right] \sim e^{-2\gamma} \frac{\sum_{m=m_c}^{m_c+1} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} e^{\gamma m - n\lambda_{\mu/n}(m)}}{\sum_{m=m_c}^{m_c+1} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} e^{-n\lambda_{\mu/n}(m)}}$$

$$= e^{\gamma(m_c - 2)} \frac{(m_c+1)\tan^2\left(\frac{\pi}{m_c}\right) + m\tan^2\left(\frac{\pi}{m_c+1}\right) e^{\gamma + n(\lambda_{\mu/n}(m_c) - \lambda_{\mu/n}(m_c+1))}}{(m_c+1)\tan^2\left(\frac{\pi}{m_c}\right) + m\tan^2\left(\frac{\pi}{m_c+1}\right) e^{n(\lambda_{\mu/n}(m_c) - \lambda_{\mu/n}(m_c+1))}}.$$

To conclude the proof, observe that

$$n\big(\lambda_{\mu/n}(m_c) - \lambda_{\mu/n}(m_c+1)\big) = -\mu + n\log\left(\frac{1 + \tan^2\left(\frac{\pi}{m_c}\right)}{1 + \tan^2\left(\frac{\pi}{m_c+1}\right)}\right) = -\delta_n,$$

leading to the desired limiting distribution. $\square$

With our two main results stated and proven, we can now straightforwardly combine them to obtain Theorem 1.11.

Proof of Theorem 1.11. The proof of Theorem 1.11 directly follows from Proposition 4.4 and 4.5 by observing that the case $\mu \ll n$ is equivalent to the case $c = 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

5 Width in the non-Brownian regime

In this section, we provide the proof of Theorem 1.9. We employ a multiscale decomposition technique using random excursions that encode our height biased trees.

Since we are interested in bounding the width of a tree by using its contour representation, it is worth pointing out the relation between the width and the corresponding walk. For a discrete contour $C = (C_t : 0 \le t \le 2n - 1)$ (seen as a continuous function on $[0, 2n - 1]$), we let

$$w(C) = \frac{1}{2} \max\left\{ \left| \{t : C_t = \ell\} \right| : \ell \in \mathbb{R} \right\}.$$

Note that for any discrete walk the maximum above is finite. By observing that every edge of the tree exactly corresponds to one up-step and one down-step of the contour, we see that $w(C)$ exactly counts the number of edges at a certain depth in the tree. By further using that each such edge leads to a unique node at the given depth, we see that $w(C^{(T)}) = w(T)$. It is worth noting that this equality does not apply in the case where the tree is a single node, but this definition will not be applied to this case. Finally, we observe that w is naturally extended to bridges and walks and that, for any bridge $B_{x,y}^{(n)}$ going from x to y and obtained as the concatenation of $B_{x,z}^{(k)}$ and $B_{z,y}^{(n-k)}$ respectively going from x to z and from z to y, we have

$$w\big(B_{x,y}^{(n)}\big) \le w\big(B_{x,z}^{(k)}\big) + w\big(B_{z,y}^{(n-k)}\big).$$

This inequality will prove to be useful in order to decompose our random excursion in a collection of bridges.

5.1 Relating height-biased and uniform trees

Proposition 5.1. *Let $\mu = \mu_n$ such that $1/\sqrt{n} \ll \mu \ll n$. Then for any $\epsilon > 0$, there exists $\delta > 0$ such that, for any sequence of events $(A_n)_{n \ge 1}$, we have*

$$\limsup_{n \to \infty} \frac{1}{(\mu^2 n)^{1/3}} \log\left(\frac{\mathbb{P}\left(\mathbf{T}_n \in A_n, \left| \frac{h(\mathbf{T}_n)}{(2\pi^2 n/\mu)^{1/3}} - 1 \right| \le \delta \right)}{\mathbb{P}\left(\mathbb{T}_n \in A_n \;\middle|\; \left| \frac{h(\mathbb{T}_n)}{(2\pi^2 n/\mu)^{1/3}} - 1 \right| \le \delta \right)} \right) \le \epsilon.$$

Proof. Start by considering a fix $\delta > 0$ and denote $h_n = (2\pi^2 n/\mu)^{1/3}$ for the asymptotic height in this regime, recall Theorem 1.4. By observing that $\mathbf{T}_n$ and $\mathbb{T}_n$ have the same distribution once we condition on their height, the numerator of the proposition can be transformed into

$$\mathbb{P}\left(\mathbf{T}_n \in A_n, \left| \frac{h(\mathbf{T}_n)}{h_n} - 1 \right| \le \delta \right) = \sum_{|k - h_n| \le \delta h_n} \mathbb{P}\left(\mathbf{T}_n \in A_n \;\middle|\; h(\mathbf{T}_n) = k \right) \mathbb{P}\left(h(\mathbf{T}_n) = k \right)$$

$$= \sum_{|k - h_n| \le \delta h_n} \mathbb{P}\left(\mathbb{T}_n \in A_n \;\middle|\; h(\mathbf{T}_n) = k \right) \mathbb{P}\left(h(\mathbf{T}_n) = k \right).$$

Moreover, using again the uniform distribution of $\mathbf{T}_n$ and $\mathbb{T}_n$ on trees of a given height, we have that

$$\mathbb{P}\left(h(\mathbf{T}_n) = k \right) = \frac{e^{-\mu k}}{Z_n} C_{n-1} \mathbb{P}\left(h(\mathbb{T}_n) = k \right).$$

Using that the sum is taken over all $|k - h_n| \le \delta h_n$, we see that $\mu k = \mu h_n + O(\delta (\mu^2 n)^{1/3})$, leading to

$$\mathbb{P}\left(\mathbf{T}_n \in A_n, \left| \frac{h(\mathbf{T}_n)}{h_n} - 1 \right| \le \delta \right) = \sum_{|k - h_n| \le \delta h_n} \mathbb{P}\left(\mathbb{T}_n \in A_n \;\middle|\; h(\mathbb{T}_n) = k \right) \frac{e^{-\mu k}}{Z_n} C_{n-1} \mathbb{P}\left(h(\mathbb{T}_n) = k \right)$$

$$= e^{-\mu h_n + O(\delta(\mu^2 n)^{1/3})} \frac{C_{n-1}}{Z_n} \mathbb{P}\left(\mathbb{T}_n \in A_n, \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \le \delta\right)$$

$$= e^{-\mu h_n + O(\delta(\mu^2 n)^{1/3})} \frac{C_{n-1}}{Z_n} \mathbb{P}\left(\mathbb{T}_n \in A_n \,\Big|\, \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \le \delta\right) \mathbb{P}\left(\left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \le \delta\right).$$

To conclude the proof, we first observe that

$$C_{n-1}\mathbb{P}\left(\left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \le \delta\right) = H_{n,\lfloor h_n(1+\delta)\rfloor + 1} - H_{n,\lceil h_n(1-\delta)\rceil} \sim \frac{4^n \pi^2}{(1+\delta)^3 h_n^3} \frac{1}{\left(1 + \tan^2 \frac{\pi}{\lfloor h_n(1+\delta)\rfloor + 1}\right)^n}$$

where $H_{n,m}$ is defined in (1.8) and its asymptotic behaviour is given in Theorem B.1. Denoting $\tilde{h}_n = \lfloor h_n(1+\delta)\rfloor + 1$ and observing that $\mu h_n = \mu \tilde{h}_n + O(\delta(\mu^2 n)^{1/3})$, it follows that

$$\frac{\mathbb{P}\left(\mathbf{T}_n \in A_n, \left|\frac{h(\mathbf{T}_n)}{h_n} - 1\right| \le \delta\right)}{\mathbb{P}\left(\mathbb{T}_n \in A_n \,\Big|\, \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \le \delta\right)} \sim \frac{e^{-\mu \tilde{h}_n + O(\delta(\mu^2 n)^{1/3})}}{Z_n} \frac{4^n \pi^2}{(1+\delta)^3 h_n^3} \frac{1}{\left(1 + \tan^2 \frac{\pi}{\tilde{h}_n}\right)^n}$$

$$= \frac{e^{-n\lambda_{\mu/n}(\tilde{h}_n) + O(\delta(\mu^2 n)^{1/3})}}{Z_n} \frac{4^n \pi^2}{(1+\delta)^3 h_n^3},$$

where λ_x is defined in (1.6). By considering $\delta > 0$ small enough, we observe that Proposition C.1 (**3**) tells us here that

$$n\lambda_{\mu/n}(\tilde{h}_n) = \lambda_{\mu/n}(t_{\mu/n}) + O\left(\delta^2 (\mu^2 n)^{1/3}\right)$$

Finally, the statement of the proposition follows from the different cases of the asymptotic of Z_n from Theorem 4.1. Indeed, in the case $1/\sqrt{n} \ll \mu \ll n^{1/4}$, we have that

$$\frac{\mathbb{P}\left(\mathbf{T}_n \in A_n, \left|\frac{h(\mathbf{T}_n)}{h_n} - 1\right| \le \delta\right)}{\mathbb{P}\left(\mathbb{T}_n \in A_n \,\Big|\, \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \le \delta\right)} \sim e^{-n\lambda_{\mu/n}(t_{\mu/n}) + O(\delta(\mu^2 n)^{1/3})} \frac{4^n \pi^2}{(1+\delta)^3 h_n^3} \frac{1}{4^n e^\mu (e^\mu - 1)} \frac{2^{1/3} 3^{1/2}}{\pi^{5/6}} \frac{n^{5/6}}{\mu^{1/3}} e^{3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}}$$

$$= e^{O(\delta(\mu^2 n)^{1/3})},$$

and the same result applies to $\mu \asymp n^{1/4}$ and to $n^{1/4} \ll \mu \ll n$. This concludes the proof as, for any $\epsilon > 0$, we are then able to find $\delta > 0$ small enough such that $O(\delta(\mu^2 n)^{1/3}) < \epsilon(\mu^2 n)^{1/3}$. $\qquad\square$

5.2 Bounding the width of a bridge

Proposition 5.2 (Subgaussian width tails for bridges). *Let* $B_x^{(n)} = (B_{x;t}^{(n)} : 0 \le t \le n)$ *be a uniform bridge of length* n *ending at* x, *that is a uniform element from* $\mathscr{B}_x^{(n)}$. *Then, for any* $A > 0$, *there exists* $R_1, R_2 > 0$ *such that, for any* $|x| \le A\sqrt{n}$ *and* $T \ge 0$, *we have*

$$\mathbb{P}\left(\frac{w\left(B_x^{(n)}\right)}{\sqrt{n}} > T\right) \le R_1 e^{-R_2 T^2}.$$

Proof. We first observe that the reflection ϕ_x defined in Lemma 2.4 can at most double the width of the walk, meaning that

$$\frac{1}{2} w\left(B_x^{(n)}\right) \le w\left(\phi_x\left(B_x^{(n)}\right)\right) \le 2 w\left(B_x^{(n)}\right).$$

Moreover, using that ϕ_x is a bijection and thus pushes forward uniform measures to uniform measures, and letting $e = -\mathbb{1}_{n \text{ is odd}}$ for the end-point of the new bridge, we have

$$\mathbb{P}\left(\frac{w\left(B_x^{(n)}\right)}{\sqrt{n}} > T\right) \le \mathbb{P}\left(\frac{w\left(B_e^{(n)}\right)}{\sqrt{n}} > \frac{T}{2} \,\Big|\, \exists t, B_{e;t}^{(n)} = \lfloor x/2\rfloor\right) \le \frac{\mathbb{P}\left(w\left(B_e^{(n)}\right) > \sqrt{n}T/2\right)}{\mathbb{P}\left(\exists t, B_{e;t}^{(n)} = \lfloor x/2\rfloor\right)}.$$

Using the distributional convergence of $B_e^{(n)}$ towards the Brownian bridge along with the fact that $|\lfloor x/2 \rfloor| \leq A\sqrt{n}$, we see that the denominator asymptotically satisfies

$$\mathbb{P}\left(\exists t, B_{e;t}^{(n)} = \lfloor x/2 \rfloor\right) \geq \mathbb{P}\left(\max_{0 \leq t \leq n} \frac{B_{e;t}^{(n)}}{\sqrt{n}} > A, \min_{0 \leq t \leq n} \frac{B_{e;t}^{(n)}}{\sqrt{n}} < -A\right)$$

$$= (1 + o(1))\mathbb{P}\left(\max_{0 \leq t \leq 1} B_t > A, \min_{0 \leq t \leq 1} B_t < -A\right),$$

where $(B_t : 0 \leq t \leq 1)$ is a standard Brownian bridge from 0 to 0. Furthermore, the right-hand side corresponds to a strictly positive function $K(A) > 0$, see for instance [8, Formula 1.15.8, Pg 180]. To conclude the proof, first observe that if n is even, we can extend $B_0^{(n)}$ to $B_{-1}^{(n+1)}$ and add at most +1 to the width. Then, by observing that the canonical map ϕ from Lemma 2.3 does not modify the width, the numerator

$$\mathbb{P}\left(w(B_e^{(n)}) > \sqrt{n}T/2\right) \leq \mathbb{P}\left(w(E^{(n)}) > \sqrt{n}T/2 - 1\right),$$

where $E^{(n)}$ is a uniform excursion of size either n or $n+1$ (depending on the parity of n). Finally, using that the width of an excursion is equal to the width of the corresponding tree, and that the distribution of the corresponding tree is a critical Bienamyé tree with $\text{Geo}(1/2)$ offspring distribution, conditioned to have size $\lceil n/2 \rceil$, we can use [3, Theorem 1.1] which states the existence of c_1, C_1 positive constants independent of T and n such that

$$\mathbb{P}\left(w(E^{(n)}) > \sqrt{n}T/2 - 1\right) \leq C_1 e^{-c_1(\sqrt{n}T/2-1)^2/n}.$$

Combining all previous inequalities lead to the existence of $R_1, R_2 > 0$ such that, for n large enough, we have

$$\mathbb{P}\left(\frac{w(B_x^{(n)})}{\sqrt{n}} > T\right) \leq (1 + o(1))\frac{C_1 e^{-c_1(\sqrt{n}T/2-1)^2/n}}{K(A)} \leq R_1 e^{-R_2 T^2}.$$

This can then be extended to the desired result by choosing R_1 and R_2 such that the inequality trivially holds when n is small. $\qquad\square$

Corollary 5.3 (Exponential moments of the width of uniform bridges)**.** *We denote by $B_{x,y}^{(n)}$ a uniform bridge of length n from x to y, that is a uniform element from $\mathscr{B}_{x,y}^{(n)}$. Then, for any $s \in \mathbb{R}$ and $A > 0$, there exists $M > 0$ such that, for all $n \geq 1$ and all $0 \leq x, y \leq A\sqrt{n}$, we have*

$$\mathbb{E}\left[\exp\left(s\frac{w(B_{x,y}^{(n)})}{\sqrt{n}}\right)\right] \leq M.$$

Proof. We start by observing that $B_{x,y}^{(n)}$ can be mapped to $B_{y-x}^{(n)}$ a bridge from 0 to $y - x$, by simply translating the walk. Moreover, shifting a walk does not affect its width, so we have

$$w(B_{x,y}^{(n)}) \stackrel{d}{=} w(B_{y-x}^{(n)}).$$

Further using that $0 < w(B_{y-x}^{(n)}) \leq n$, we see that the desired expected value re-writes as

$$\mathbb{E}\left[\exp\left(s\frac{w(B_{x,y}^{(n)})}{\sqrt{n}}\right)\right] = \mathbb{E}\left[\exp\left(s\frac{w(B_{y-x}^{(n)})}{\sqrt{n}}\right)\right] = \sum_{k=1}^{n} \mathbb{E}\left[\exp\left(s\frac{w(B_{y-x}^{(n)})}{\sqrt{n}}\right)\mathbb{1}_{\{\sqrt{k-1} < w(B_{y-x}^{(n)})/\sqrt{n} \leq \sqrt{k}\}}\right],$$

Now, thanks to Proposition 5.2 and since $|y - x| \leq A\sqrt{n}$ whenever $0 \leq x, y \leq A\sqrt{n}$, we have that

$$\mathbb{E}\left[\exp\left(s\frac{w(B_{x,y}^{(n)})}{\sqrt{n}}\right)\mathbb{1}_{\{\sqrt{k-1} < w(B_{y-x}^{(n)})/\sqrt{n} \leq \sqrt{k}\}}\right] \leq e^{s\sqrt{k}}\mathbb{P}\left(\frac{w(B_{y-x}^{(n)})}{\sqrt{n}} > \sqrt{k-1}\right) \leq R_1 e^{s\sqrt{k}-R_2(k-1)}.$$

Since the sum of the right-hand side above is finite even if we sum over all $k \in \mathbb{N}$, the corollary follows. $\qquad\square$

5.3 Width in the non-Brownian regime

To conclude this section, we now prove Theorem 1.9 stating that the width of a height-biased tree when $1/\sqrt{n} \ll \mu \ll n$ is of order $(\mu n^2)^{1/3}$, although without providing the exact constant. The proof relies on splitting the contour of a tree with bounded height into standard bridges whose height is the square of their length. A representation of this multiscale idea can be found in Figure 3.
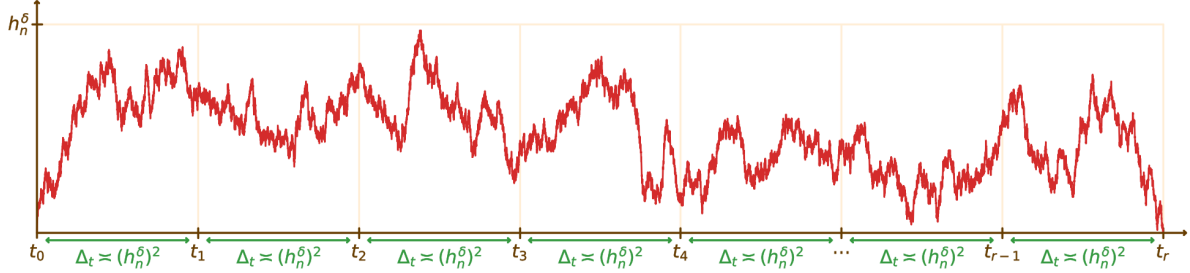

Figure 3: A representation of the multiscale decomposition used for the proof of Theorem 1.9. We bound the height with $h_n^\delta = (1+\delta)(2\pi^2 n/\mu)^{1/3}$ and split the contour of a uniform tree into $r \sim (\mu^2 n)^{1/3}$ bridges of length $\Delta_t \sim 2(n/\mu)^{2/3}$. This way the length of each bridge is of the same order as the square of its fluctuations, thus are objects from the *Brownian world*, admitting good enough width tails for our purposes, thanks to Proposition 5.2.

Proof of Theorem 1.9. To start the proof, first note that, since the width counts the number of nodes at a certain height in the tree, by summing over all possible heights we directly have the bound

$$w(\mathbf{T}_n) \cdot h(\mathbf{T}_n) \geq n.$$

Using the result of Theorem 1.4 which states that $h(\mathbf{T}_n)$ is asymptotically of order $(n/\mu)^{1/3}$, we directly see that $w(\mathbf{T}_n)$ is at least of order $(\mu n^2)^{1/3}$ as desired. This proves the lower bound of the theorem and we now focus on proving that $w(\mathbf{T}_n)$ is not of higher order than $(\mu n^2)^{1/3}$. Following the proof of Proposition 5.1, we let $h_n = (2\pi^2 n/\mu)^{1/3}$ be the asymptotic height of $\mathbf{T}_n$.

Fix for now some $K > 0$ and $\epsilon > 0$. We start by applying Proposition 5.1 to the event $A_n = \{w(T) > K(\mu n^2)\}$ to find $\delta > 0$ such that

$$\limsup_{n \to \infty} \frac{1}{(\mu^2 n)^{1/3}} \log \left(\frac{\mathbb{P}\left(w(\mathbf{T}_n) > K(\mu n^2)^{1/3}, \left| \frac{h(\mathbf{T}_n)}{h_n} - 1 \right| \leq \delta \right)}{\mathbb{P}\left(w(\mathbb{T}_n) > K(\mu n^2)^{1/3} \mid \left| \frac{h(\mathbb{T}_n)}{h_n} - 1 \right| \leq \delta \right)} \right) \leq \epsilon.$$

By further using Theorem 1.4, we know that $h(\mathbf{T}_n)$ converges in probability to h_n. Combining this with the previous result leads to

$$\mathbb{P}\left(w(\mathbf{T}_n) > K(\mu n^2)^{1/3} \right) \leq e^{O(\epsilon(\mu^2 n)^{1/3})} \mathbb{P}\left(w(\mathbb{T}_n) > K(\mu n^2)^{1/3} \mid \left| \frac{h(\mathbb{T}_n)}{h_n} - 1 \right| \leq \delta \right) + o(1). \tag{5.1}$$

Proving the theorem now boils down to proving an upper bound on the width of a uniform tree.

Given a uniform tree $\mathbb{T}_n$, recall we let $C^{(\mathbb{T}_n)} = (C_t^{(\mathbb{T}_n)} : 0 \leq t \leq 2n-1)$ to be its contour walk. We are going to split this walk into a sequence of bridges of length $2(n/\mu)^{2/3}$ so that, once conditioned to be within the range $[0, h_n)$, their length is exactly this range squared. We refer to Figure 3 for a representation of this idea.

Let $r = r_n = \lfloor (\mu^2 n)^{1/3} \rfloor$ and set $t_0, \ldots, t_r$ to evenly spaced time points from the interval $2n - 1$, meaning that

$$t_i = \left\lfloor \frac{i(2n-1)}{r} \right\rfloor.$$

We observe here that $r \sim (\mu^2 n)^{1/3}$ and that $\Delta_t = \max\{t_i - t_{i-1}\} \sim 2(n/\mu)^{2/3}$. Finally, we denote by $B_i = (C_{t+t_{i-1}}^{(\mathbb{T}_n)} : 0 \leq t \leq t_i - t_{i-1})$ the walk corresponding to the portion of $C^{(n)}$ between time t_{i-1} and time t_i (rescaled in time to start at 0). We now make two important observations.

1. Since $C^{(\mathbb{T}_n)}$ is obtained by concatenating $B_1, \ldots, B_r$, its width is bounded by the sum of their widths:

$$w\big(C^{(\mathbb{T}_n)}\big) \leq \sum_{i=1}^{r} w(B_i).$$

2. By the Markov property $B_1, B_2, \ldots, B_r$ are independent and uniform bridges conditionally given $C_{t_i}^{(\mathbb{T}_n)}$ for $0 \leq i \leq r$.

To use these observations, start by applying a standard exponential Markov's inequality to see that, for any $\rho > 0$, we have

$$\mathbb{P}\left(w(\mathbb{T}_n) > K(\mu n^2)^{1/3} \;\middle|\; \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \leq \delta\right) \leq e^{-\rho K(\mu n^2)^{1/3}} \mathbb{E}\left[e^{\rho w(\mathbb{T}_n)} \;\middle|\; \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \leq \delta\right].$$

We then combine the first observation and a nested expectation to obtain that

$$\mathbb{P}\left(w(\mathbb{T}_n) > K(\mu n^2)^{1/3} \;\middle|\; \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \leq \delta\right)$$

$$\leq e^{-\rho K(\mu n^2)^{1/3}} \mathbb{E}\left[\mathbb{E}\left[e^{\rho w(\mathbb{T}_n)} \;\middle|\; \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \leq \delta, C_{t_0}^{(\mathbb{T}_n)}, C_{t_1}^{(\mathbb{T}_n)}, \ldots, C_{t_r}^{(\mathbb{T}_n)}\right]\right]$$

$$\leq e^{-\rho K(\mu n^2)^{1/3}} \mathbb{E}\left[\mathbb{E}\left[e^{\rho[w(B_1) + \ldots + w(B_r)]} \;\middle|\; \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \leq \delta, C_{t_0}^{(\mathbb{T}_n)}, C_{t_1}^{(\mathbb{T}_n)}, \ldots, C_{t_r}^{(\mathbb{T}_n)}\right]\right].$$

The condition that $h(\mathbb{T}_n)$ is bounded can be reduced to simply having $0 \leq C_{t_i}^{\mathbb{T}_n} \leq (1+\delta)h_n$ by dividing with the probability that a walk of length $2n-1$ is a contour whose height is of order h_n, which itself is of order $e^{O(\log n) + O((\mu^2 n)^{1/3})}$. Finally, the second observation states that, under the above conditioning, all bridges are independent from each other and their endpoints both belong to the interval $[0, (1+\delta)h_n]$. Further using that the length of the bridges is bounded by $\Delta_t \sim 2(n/\mu)^{2/3}$ and that extending a walk can only increase its width, the previous bound can be transformed into

$$\mathbb{P}\left(w(\mathbb{T}_n) > K(\mu n^2)^{1/3} \;\middle|\; \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \leq \delta\right) \leq e^{-\rho K(\mu n^2)^{1/3} + O(\log n) + O((\mu^2 n)^{1/3})} \left(\max_{0 \leq x, y \leq A\sqrt{\Delta_t}} \mathbb{E}\left[e^{\rho w\big(B_{x,y}^{(\Delta_t)}\big)}\right]\right)^r,$$

where $A > 0$ is some constant such that $(1+\delta)h_n \leq A\sqrt{\Delta_t}$ and $B_{x,y}^{(\Delta_t)}$ is a uniform random bridge of length Δ_t going from x to y.

For the final part of this proof, start by replacing ρ with $1/\sqrt{\Delta_t}$ and apply Corollary 5.3 with the same A and $s = 1$ to find some $M > 0$ such that, for any $\Delta_t > 0$

$$\max_{0 \leq x, y \leq A\sqrt{\Delta_t}} \mathbb{E}\left[e^{w\big(B_{x,y}^{(\Delta_t)}\big)/\sqrt{\Delta_t}}\right] \leq M.$$

Finally, by recalling that $\mu \gg (\log n)^{3/2}/\sqrt{n}$, that $r \sim (\mu^2 n)^{1/3}$, and that $\Delta_t \sim 2(n/\mu)^{2/3}$, combining the previous two bounds lead to

$$\mathbb{P}\left(w(\mathbb{T}_n) > K(\mu n^2)^{1/3} \;\middle|\; \left|\frac{h(\mathbb{T}_n)}{h_n} - 1\right| \leq \delta\right) \leq \exp\left(-(1 + o(1))\left[\frac{K}{\sqrt{2}} - \log M + O(1)\right](\mu^2 n)^{1/3}\right).$$

Plugging this back into (5.1) finally leads to

$$\mathbb{P}\left(w(\mathbf{T}_n) > K(\mu n^2)^{1/3}\right) \leq \exp\left(-(1 + o(1))\left[\frac{K}{\sqrt{2}} - \log M + O(\epsilon) + O(1)\right](\mu^2 n)^{1/3}\right) + o(1)$$

Observe that the definitions of M, $O(\epsilon)$, and $O(1)$ only depend on ϵ, δ, and A and are thus independent of K. This means that we can take K large enough such that the exponential converges to 0, thus proving the desired upper bound and concluding the proof. $\qquad\square$

6   Local results around the root: local convergence and condensation

In this section we prove Theorems 1.15 and 1.16, completing the picture of local behavior around the root.

6.1   Local limits

Before moving on to the results, we briefly discuss the precise topology that the convergence in Theorem 1.15 takes place in. Even though the following definitions are standard, we include it for the sake of completeness. The notion of local convergence of sparse graphs goes back to works of Benjamini and Schramm [7], and Aldous and Steele [5]. In the context of rooted planar trees, we use the notion of local convergence as presented by Abraham and Delmas in [1], which respects the ordering of the children of every vertex.

Let T_1, T_2 be two rooted plane trees (possibly infinite). We equip the set of all plane trees $\mathfrak{t}$ with the metric

$$d_{\mathrm{loc}}(T_1, T_2) = \inf\left\{\frac{1}{1+R} : R \in \mathbb{N}, B_R(T_1) = B_R(T_2)\right\}$$

where $B_R(T)$ denotes the subtree of T spanned by the vertices at distance at most R from the root.

Definition 6.1 (Balls around a rooted tree)**.** For T a rooted tree and $a > 0$, define by $\mathscr{B}_a(T)$ the open ball of radius a around T in the local topology, i.e.,

$$\mathscr{B}_a(T) := \{T : d_{\mathrm{loc}}(T, T') < a\}.$$

In other words, for $r \in \mathbb{N}^*$, $\mathscr{B}_{1/r}(T)$ consists of those rooted trees that look identical with T up to generation r.

Remark 6.2. The set of finite trees $\mathfrak{t}_f$ is a countable dense subset of $\mathfrak{t} = \mathfrak{t}_f \cup \mathfrak{t}_\infty$, where $\mathfrak{t}_\infty$ stands for the set of infinite trees. It is also an easy exercise to check that the balls $\mathscr{B}_{1/r}$ are both closed and open, and two balls are either disjoint or one is contained in the other.

It is a well known result of Kesten [22] that the local limit of the uniform tree $\mathbb{T}_n$ is the the so-called Kesten's tree (KT) associated to the critical Geo(1/2) offspring distribution with mass function

$$p(n) = \left(\frac{1}{2}\right)^{n+1}, n = 0, 1, 2, \dots$$

which is a multi-type Biénaymé tree with two types of individuals: normal and special. The special individuals reproduce according to the size-biased distribution $p^*(n) = np(n), n = 0, 1, \dots$, whereas the normal ones reproduce according to $p(n)$. The root is special and reproduces according to $p^*(n) = np(n)$ where one of the children, chosen uniformly, is special and the others are normal. This is an infinite tree with a single spine, i.e. with a unique infinite path emanating from the root. When we write Kesten's tree or simply KT, we always refer to the one associated to the critical geometric offspring distribution as described above, since this is the only case we encounter in this work.

We now move on to prove Theorem 1.15, stating that the local limit of the tree $\mathbf{T}_n$ when $0 \leq \mu \ll 1$ is also KT, i.e., the height bias is not strong enough to change the local behavior. We know from [19] that when μ is a positive real constant, the local limit is the Poisson tree with parameter μ of [2]. This is an infinite tree which is not an instance of Kesten's tree, but it still has two types, with a backbone without leaves corresponding to individuals having an infinite progeny on which are grafted independent critical (finite) Bienaymé trees. Since $\mu = \alpha$ with $\alpha > 0$ is already covered in [19], and the $\mu \to \alpha$ case can be treated with the same techniques found therein up to minor modifications, we do not give further details on the explicit properties of this limiting tree and refer the reader to [2, 19].

For the $0 \leq \mu \ll 1$ case, we will need the following lemma for the proof of Theorem 1.15. Below, for notational convenience, we write for any event A of trees,

$$\mathbb{P}\left(\mathbf{T}_n \in A\right) = \nu_n^\mu(A).$$

Lemma 6.3. *Assume that $0 \leq \mu \ll 1$ and T_0 a finite tree of height r with K vertices at level r. For any $M \in \mathbb{N}$, it holds that*

$$\nu_n^\mu\big(\mathscr{B}_{1/r}(T_0)\big) \geq \big(1 + o(1)\big) K \frac{e^{-\mu r}}{4^{|T_0| - K}} \left(\sum_{S=1}^M C_{S-1} 4^{-S} \right)^{K-1}, \tag{6.1}$$

where $\mathscr{B}_{1/r}(T_0)$ is the open ball of radius $1/r$ around the tree T_0 in the local topology.

Proof. **Case $1/\sqrt{n} \ll \mu \ll 1$.** Since $\mu \to 0$ and $\mu^2 n \to \infty$ for the window in question, we can assume without loss of generality that $\mu = \mu_n$ and $\mu^2 n$ are respectively decreasing and increasing functions of n. Note that any tree $T \in \mathscr{B}_{1/r}(T_0)$ can be obtained by grafting K trees on the top K vertices of T_0, i.e., the *leaves* of T_0 at maximal height. The subtrees rooted at the top vertices of T_0 are referred to as *branches* in the rest of the proof. Let M be some fixed large constant and consider the subset of $\mathscr{B}_{1/r}(T_0)$ where all branches except one have size $< M$. Assuming that $n > KM + |T_0|$, exactly one branch has size $\geq M$ which we call the large one. We further restrict to the case where the height of the subtree that is grafted on the leaf corresponding to the one with the largest subtree, is at least M. In other words, the largest branch is also the highest with height at least M, and the total height of the tree T thus formed after the grafting procedure is at least $r + M$.

Let $\Omega(T_0, r)$ be the subset of $\mathscr{B}_{1/r}(T_0)$ consisting of trees of size n satisfying the properties above and whose large branch is precisely the first one, where we hide from notation the n-dependence of Ω. Then, by using the symmetry between all branches and only considering the case where the first branch is the largest one, we have

$$\nu_n^\mu\big(\mathscr{B}_{1/r}(T_0)\big) \geq K \cdot \nu_n^\mu\big(\Omega(T_0, r)\big).$$

We proceed to estimate the right hand side. Denoting the sizes of the subbranches grafted on the top K vertices by N_i, $i \in \{1, \ldots, K\}$, we have

$$\nu_n^\mu\big(\Omega(T_0, r)\big) \geq \frac{1}{Z_n^\mu} \sum_{\substack{n_2, \ldots, n_K < M \\ n_1 = n - n_2 - \cdots - n_K}} C_{n_2 - 1} \ldots C_{n_K - 1} \sum_{m=M}^n e^{-\mu(m+r)} (H_{n_1, m+1} - H_{n_1, m}),$$

where $H_{n,m}$ is the number of trees of height $< m$, as defined in (1.8).

Recall that $1/\sqrt{n} \ll \mu \ll 1$. For this window, we employ Lemma B.2 and Proposition D.3, where we choose ϵ_n to be $1/(\mu^2 n)^{1/6 - \delta}$ for any fixed $0 < \delta < 1/6$ to estimate the second sum on the right hand side. Using $n_1 < n$, $\mu_{n_1} > \mu_n$ and $\mu_{n_1}^2 n_1 < \mu_n^2 n$ we have, for n large,

$$\sum_{m=M}^n e^{-\mu_n m} \big(H_{n_1, m+1} - H_{n_1, m}\big) \geq \sum_{m=M}^n e^{-\mu_{n_1} m} \big(H_{n_1, m+1} - H_{n_1, m}\big)$$

$$\geq \big(1 + o(1)\big) \frac{4^{n_1} \pi^{5/6}}{2^{1/3} 3^{1/2}} \frac{\mu_n}{n} \exp\left(-3 \left(\frac{\pi^2 \mu_n^2 n}{4} \right)^{1/3} \right), \tag{6.2}$$

which together with Theorem 4.1 leads to

$$\nu_n^\mu\big(\Omega(T_0, r)\big) \geq \big(1 + o(1)\big) \frac{e^{-\mu r}}{4^{|T_0| - K}} \left(\sum_{s=1}^M C_{s-1} 4^{-s} \right)^{K-1}, \tag{6.3}$$

which concludes the proof for the regime in question once multiplied by K.

Case $\mu \sqrt{n} \to \alpha \in [0, \infty)$. Recall $\mathbb{T}_n$ is a uniform plane tree, i.e., sampled from (1.3) corresponding to $\mu = 0$. It follows from the fact that $h(\mathbb{T}_n)$ admits sub-Gaussian tails [3, Theorem 1.2],

$$Z_n^\mu = \mathbb{E}\left[e^{-\mu h(\mathbb{T}_n)} \right] \sim C_{n-1} \mathbb{E}\left[e^{-\alpha \|\mathbf{e}\|} \right]$$

where recall $\|\mathbf{e}\| = \sup_{t \in [0,1]} \mathbf{e}(t)$ with $\mathbf{e}$ being a standard Brownian excursion. We can then repeat the arguments above and analogously to (6.2) obtain a lower bound such that for any $\epsilon > 0$ and n sufficiently large

$$\sum_{m=M}^n e^{-\mu m} \big(H_{n_1, m+1} - H_{n_1, m}\big) \geq (1 - \epsilon) C_{n_1 - 1} \mathbb{E}\left[e^{-\alpha \|\mathbf{e}\|} \right]$$

which, using the well-known fact $C_n \sim \pi^{-1/2} 4^n n^{-3/2}$, simply reproduces the lower bound found in (6.3). This concludes the proof once multiplied by K. $\qquad\square$

Proof of Theorem 1.15. As mentioned before, when μ is constant, the local limit was already identified in [19]. When $\mu \to \alpha > 0$, the same analysis found in [19, Lemma 4.2] for the lower bound of the ball volumes can be repeated: for $\delta > 0$ and n sufficiently large, μ can be replaced by $\alpha - \delta$ or $\alpha + \delta$ wherever needed. Since δ is arbitrary, in the $n \to \infty$ limit one obtains the same ball volume which characterizes the local limit to be the same as $\mu = \alpha$ constant case.

In the $0 \le \mu \ll 1$ case, we apply Lemma 6.3. Note that taking $n \to \infty$ then $M \to \infty$, the right hand side of (6.1) is given by

$$2K \cdot 2^K \cdot 4^{-|T_0|}$$

which is equal to $\nu^{(0)}(\mathscr{B}_{1/r}(T_0))$ where $\nu^{(0)}$ is the law of the uniform tree $\mathbb{T}_n$. Remark 6.2 implies that any open subset of $\mathfrak{t}$ can be written as a disjoint union of balls. Then it follows from the Portemanteau theorem that the local limit when $0 \le \mu \ll 1$ is Kesten's tree.

$\square$

6.2 Condensation at the root away from the local regime

The goal of this section is to prove Theorem 1.16 regarding the asymptotic behaviour of the root degree in a random tree $\mathbf{T}_n$ distributed according to (1.3). To do so, we start by providing the following proposition, showing the asymptotic probability of the root degree to be a given value in the range of possibilities.

Proposition 6.4. *Let $\mu = \mu_n$ such that $1 \ll \mu \ll n^{1/4}$. Let $\mathbf{T}_n$ be a random tree sampled according to (1.3). Then, for any fixed $[-R, R]$, and $\rho \in [-R, R]$, we have that*

$$\mathbb{P}\left(\deg_{\mathbf{T}_n}(\varnothing) = \lfloor 2\mu + \rho\sqrt{\mu} \rfloor\right) \ge \left(1 + o(1)\right)\frac{e^{-\rho^2/12}}{\sqrt{12\pi\mu}},$$

where the $o(1)$ term is uniformly small over $\rho \in [-R, R]$.

Proof. Assume that $R > 0$ is fixed and $\rho \in [-R, R]$. We further write $r = \lfloor 2\mu + \sqrt{\mu}\rho \rfloor$ for the rest of the proof and assume without loss of generality that n is large enough so that $r < n - 1$. We recall that $H_{n,m}$ is the number of trees of size n and height $< m$, as defined in (1.8).

To start the proof, note that a tree of height $\le m$ and root degree r can be bijectively mapped with a forest of r trees of height $< m$ and whose total size equals $n - 1$, by simply looking at the subtrees or branches emanating from the offsprings of the root. By the last observation, the total number of trees of height m and root degree r is

$$\left|\left\{T \in \mathfrak{t}_n : h(T) = m, \deg_T(\varnothing) = r\right\}\right| = \sum_{n_1 + \cdots + n_r = n-1}\left(H_{n_1,m}\cdots H_{n_r,m} - H_{n_1,m-1}\cdots H_{n_r,m-1}\right),$$

where the sum is taken over $n_i \ge 1$. Using this formula, we obtain that

$$\mathbb{P}\left(\deg_{\mathbf{T}_n}(\varnothing) = r\right) = \frac{1}{Z_n}\sum_{m=2}^{n-1} e^{-\mu m}\sum_{n_1 + \dots + n_r = n-1}\left(H_{n_1,m}\cdots H_{n_r,m} - H_{n_1,m-1}\cdots H_{n_r,m-1}\right)$$

$$= \frac{1}{Z_n}\sum_{n_1 + \dots + n_r = n-1}\left[(e^\mu - 1)\sum_{m=3}^{n-1} e^{-\mu m}\prod_{i=1}^r H_{n_i,m-1} + e^{-\mu(n-1)}\prod_{i=1}^r C_{n_i - 1}\right].$$

Moreover, by considering an arbitrary $\epsilon = \epsilon_n$ such that $1 \gg \epsilon \gg 1/(\mu^2 n)^{1/6} \asymp (n/\mu^4)^{1/6}/t_n$, Theorem 1.5 tells us that we can reduce the previous sum over m by the sum over $|m - t_n| \le \epsilon t_n$. Further using that $\mu \gg 1$, we obtain the first approximate lower bound as

$$\mathbb{P}\left(\deg_{\mathbf{T}_n}(\varnothing) = r\right) \ge \left(1 + o(1)\right)\frac{e^\mu}{Z_n}\sum_{|m-t_n| \le \epsilon t_n} e^{-\mu m}\sum_{n_1 + \dots + n_r = n-1}\prod_{i=1}^r H_{n_i,m-1}.$$

In the case of this sum, since $m \ge (1 - \epsilon)t_n \gg 1$, we can apply Theorem B.1 stating that, for any $\gamma = \gamma_n \gg 1$ and uniformly over any $n_i > \gamma m^2$, we have that

$$H_{n_i,m-1} = \frac{4^n}{m}\frac{\tan^2\frac{\pi}{m}}{\left(1 + \tan^2\frac{\pi}{m}\right)^n} + O\left(\frac{4^n}{m^3}\frac{1}{\left(1 + \tan^2\frac{2\pi}{m}\right)^n}\right)$$

$$= \frac{4^n}{m} \frac{\tan^2 \frac{\pi}{m}}{\left(1 + \tan^2 \frac{\pi}{m}\right)^n} \left[1 + O\left(e^{-\frac{3n\pi^2}{m^2}}\right)\right].$$

Since this approximation only works when n_i is large enough, we split the sum over all n_i in the approximation of the probability according to whether $n_i \leq \gamma m^2$ or $n_i > \gamma m^2$. Without loss of generality, and to simplify the upcoming formulas, we further assume that γ is chosen so that γm^2 is an integer. We also observe that there is no constraint on the lower bound of the rate at which γ diverges. With that in mind, we can add finitely many diverging upper bounds on γ and the result would still apply. For any such upper bound that we add to γ, we mention that "γ can be chosen to diverge arbitrarily slowly", having in mind that this means that the related result holds as long as some undisclosed upper bound holds for γ.

To start the proof, observe that $m \sim t_n$ here, so that

$$\frac{rm^2}{n} \asymp \frac{\mu}{n} \left(\frac{n}{\mu}\right)^{2/3} = \left(\frac{\mu}{n}\right)^{1/3}.$$

Thus, by choosing γ diverging slowly enough (here meaning that $\gamma(\mu/n)^{1/3} \ll 1$), we can see that it is impossible for all n_i to be less than γm^2. By further re-ordering the n_i so that the first ones are larger than γm^2, we see that

$$\sum_{n_1+\ldots+n_r=n-1} \prod_{i=1}^r H_{n_i,m-1} = \sum_{\ell=1}^r \binom{r}{\ell} \sum_{\{1 \leq n_{\ell+1},\ldots,n_r \leq \gamma m^2\}} \sum_{\{n_1+\ldots+n_r=n-1:n_1,\ldots,n_\ell > \gamma m^2\}} \prod_{i=1}^r H_{n_i,m-1}.$$

Write for now $N = n_{\ell+1} + \ldots + n_r = O(\gamma(n^2\mu)^{1/3}) = o(n)$. Using the approximation for $H_{n_i,m-1}$ when $n_i > \gamma m^2$, we obtain that

$$\sum_{\{n_1+\ldots+n_r=n-1:n_1,\ldots,n_\ell > \gamma m^2\}} \prod_{i=1}^\ell H_{n_i,m-1} = \sum_{\{n_1+\ldots+n_\ell=n-1-N:n_i > \gamma m^2\}} \prod_{i=1}^\ell \frac{4^{n_i}}{m} \frac{\tan^2 \frac{\pi}{m}}{\left(1 + \tan^2 \frac{\pi}{m}\right)^{n_i}} \left[1 + O\left(e^{-\frac{3n\pi^2}{m^2}}\right)\right]$$

$$= \frac{4^{n-1-N}\pi^{2\ell}}{m^{3\ell}} \frac{1}{\left(1 + \tan^2 \frac{\pi}{m}\right)^{n-1-N}} \left[1 + O\left(e^{-\frac{3n\pi^2}{m^2}}\right)\right]^\ell \binom{n-2-N-\ell\gamma m^2}{\ell-1}.$$

It is now worth adding a constraint on ℓ to simplify the computations down the line.

The previous sum over ℓ covers the range $1, \ldots, r \sim 2\mu$. However, computations show that only the terms $\ell \sim \mu$ with a window of size $\sqrt{\mu}$ contribute significantly to the sum. For this reason, and since we are still considering a lower bound, thus allowing us to remove terms at our own discretion, we now focus on the terms $|\ell - \mu| \leq \gamma \sqrt{\mu}$ (assuming that $\gamma \ll \sqrt{\mu}$, so that this set of ℓ is contained within $[1, r]$). For any such ℓ, we now see that

$$2 + N + \ell\gamma m^2 = O(\gamma(n^2\mu)^{1/3}) = o(n),$$

from which we can deduce that

$$\binom{n-2-N-\ell\gamma m^2}{\ell-1} = \frac{n^{\ell-1}}{(\ell-1)!} \prod_{i=0}^{\ell-2} \left(1 - \frac{2+N+\ell\gamma m^2+i}{n}\right)$$

$$= \frac{n^{\ell-1}}{(\ell-1)!} \exp\left(O\left(\frac{\gamma\ell(n^2\mu)^{1/3}}{n}\right)\right).$$

Using that we consider $\ell \sim \mu$ here, the $O(\cdot)$ term becomes $O(\gamma(\mu^4/n)^{1/3})$ which, for γ diverging slowly enough, is $o(1)$. Further observing that $\ell e^{-3n\pi^2/m^2} \ll 1$, the above product on $H_{n_i,m-1}$ simplifies to

$$\sum_{n_1+\ldots+n_r=n-1} \prod_{i=1}^r H_{n_i,m-1} \geq (1+o(1)) \sum_{|\ell-\mu| \leq \gamma\sqrt{\mu}} \binom{r}{\ell} \sum_{1 \leq n_{\ell+1},\ldots,n_r \leq \gamma m^2} \prod_{i=\ell+1}^r H_{n_i,m-1} \frac{4^{n-1-N}\pi^{2\ell}}{m^{3\ell}} \frac{1}{\left(1 + \tan^2 \frac{\pi}{m}\right)^{n-1-N}} \frac{n^{\ell-1}}{(\ell-1)!}$$

$$= (1+o(1)) \left(\frac{4}{1 + \tan^2 \frac{\pi}{m}}\right)^{n-1} \sum_{|\ell-\mu| \leq \gamma\sqrt{\mu}} \binom{r}{\ell} \frac{\pi^{2\ell} n^{\ell-1}}{m^{3\ell}(\ell-1)!} \left(\sum_{s=1}^{\gamma m^2} H_{s,m-1}\left(\frac{1 + \tan^2 \frac{\pi}{m}}{4}\right)^s\right)^{r-\ell},$$

where we recall that $N = n_{\ell+1} + \ldots + n_r$.

33

Applying [12, Eq(9)] for $\theta \to 0$, we know that

$$\sum_{s=1}^{\infty} \frac{1}{4^s} H_{s,m-1} = \frac{m-1}{2m}.$$

Further recalling from Theorem B.1 that

$$H_{s,m-1} = O\left(\frac{4^s}{m^3}\right),$$

we can use the infinite sum to approximate the finite one and obtain

$$\sum_{s=1}^{\gamma m^2} H_{s,m-1} \left(\frac{1+\tan^2\frac{\pi}{m}}{4}\right)^s = \sum_{s=1}^{\gamma m^2} \frac{H_{s,m-1}}{4^s}\left[\left(1+\tan^2\frac{\pi}{m}\right)^s - 1\right] - \sum_{s>\gamma m^2} \frac{H_{s,m-1}}{4^s} + \frac{m-1}{2m}$$

$$= \sum_{s=1}^{\gamma m^2} \frac{H_{s,m-1}}{4^s}\left[\left(1+\tan^2\frac{\pi}{m}\right)^s - 1\right] + \frac{1}{2} + O\left(\frac{1}{m}\right).$$

Now, in order to bound the difference between the two finite sums, we can use that

$$s \mapsto \left[\left(1+\tan^2\frac{\pi}{m}\right)^s - 1\right]\frac{m^2 e^{-\gamma}}{s} = \left[e^{\frac{s\pi^2}{m^2}+O\left(\frac{s}{m^4}\right)} - 1\right]\frac{m^2 e^{-9.9\gamma}}{s}$$

is bounded by some constant C as long as $1 \le s \le \gamma m^2$ and γ diverges slowly enough (we used here that $\pi^2 < 9.9$). It follows that

$$\sum_{s=1}^{\gamma m^2} \frac{H_{s,m-1}}{4^s}\left[\left(1+\tan^2\frac{\pi}{m}\right)^s - 1\right] \le C e^{9.9\gamma} \sum_{s=1}^{\gamma m^2} \frac{H_{s,m-1}}{4^s}\frac{s}{m^2} = O\left(e^{9.9\gamma}\frac{\gamma^2 m^4}{m^5}\right) = O\left(\frac{e^{10\gamma}}{m}\right),$$

where we used again that $H_{s,m-1} = O(4^s/m^3)$ and further that $e^{9.9\gamma}\gamma^2 \ll e^{10\gamma}$. Pluggin this result back into the approximation for the finite sum, we eventually obtain that

$$\sum_{s=1}^{\gamma m^2} H_{s,m-1}\left(\frac{1+\tan^2\frac{\pi}{m}}{4}\right)^s = \frac{1}{2} + O\left(\frac{e^{10\gamma}}{m}\right).$$

Recall now that $r - \ell \asymp \mu$ along with $m \asymp (n/\mu)^{1/3}$ to see that $(r-\ell)/m \asymp (\mu^4/n)^{1/3} \ll 1$, allowing us to choose γ diverging slowly enough such that

$$\sum_{n_1+...+n_r=n-1} \prod_{i=1}^{r} H_{n_i,m-1} \ge (1+o(1))\left(\frac{4}{1+\tan^2\frac{\pi}{m}}\right)^{n-1} \sum_{|\ell-\mu|\le\gamma\sqrt{\mu}} \binom{r}{\ell} \frac{\pi^{2\ell} n^{\ell-1}}{m^{3\ell}(\ell-1)!}\frac{1}{2^{r-\ell}}.$$

Moreover, using that $|m - t_n| \le \epsilon t_n$ along with Proposition C.1 (**1**), we see that

$$m = \left(\frac{2\pi^2 n}{\mu}\right)^{1/3}\left[1+O\left(\left(\frac{\mu}{n}\right)^{2/3}\right)+O(\epsilon)\right].$$

Using that $\epsilon \gg 1/(\mu^2 n)^{1/6}$ which implies that $\mu\epsilon \gg (\mu^4/n)^{1/6}$, we can choose ϵ small enough so that it satisfies the previous assumptions and further $\mu\epsilon \ll 1$. In that scenario, and using that $\ell \sim \mu$, it follows that

$$m^{3\ell} \sim \left(\frac{2\pi^2 n}{\mu}\right)^{\ell},$$

which finally leads to

$$\sum_{n_1+...+n_r=n-1} \prod_{i=1}^{r} H_{n_i,m-1} \ge (1+o(1))\left(\frac{4}{1+\tan^2\frac{\pi}{m}}\right)^{n-1} \sum_{|\ell-\mu|\le\gamma\sqrt{\mu}} \binom{r}{\ell} \frac{\pi^{2\ell} n^{\ell-1}}{\left(\frac{2\pi^2 n}{\mu}\right)^{\ell}(\ell-1)!}\frac{1}{2^{r-\ell}}$$

$$= \left(1 + o(1)\right)\left(\frac{4}{1+\tan^2\frac{\pi}{m}}\right)^{n-1}\frac{1}{n2^r}\sum_{|\ell-\mu|\leq\gamma\sqrt{\mu}}\binom{r}{\ell}\frac{\mu^\ell}{(\ell-1)!}.$$

Relating this formula to the probability of the degree of the root, we eventually obtain that

$$\mathbb{P}\left(\deg_{\mathbf{T}_n}(\varnothing)=r\right)\geq\left(1+o(1)\right)\frac{e^\mu}{Z_n}\sum_{|m-t_n|\leq\epsilon t_n}e^{-\mu m}\left(\frac{4}{1+\tan^2\frac{\pi}{m}}\right)^{n-1}\frac{1}{n2^r}\sum_{|\ell-\mu|\leq\gamma\sqrt{\mu}}\binom{r}{\ell}\frac{\mu^\ell}{(\ell-1)!}$$

$$=\left(1+o(1)\right)\frac{e^\mu}{Z_n}\frac{4^{n-1}}{n2^r}\sum_{|\ell-\mu|\leq\gamma\sqrt{\mu}}\binom{r}{\ell}\frac{\mu^\ell}{(\ell-1)!}\sum_{|m-t_n|\leq\epsilon t_n}e^{-n\lambda_{\mu/n}(m)}.$$

We now use two comparisons to the integral to conclude this proof.

As a first step, we refer to the proof of Proposition D.3 to see that

$$\sum_{|m-t_n|\leq\epsilon t_n}e^{-n\lambda_{\mu/n}(m)}\sim e^{n\lambda_{\mu/n}(t_n)}\sqrt{2\pi}\frac{1}{\sqrt{3}}\left(\frac{2\pi^2 n}{\mu^4}\right)^{1/6},$$

which along with Theorem 4.1 and Proposition C.1 (**2**) leads to

$$\mathbb{P}\left(\deg_{\mathbf{T}_n}(\varnothing)=r\right)\geq\left(1+o(1)\right)\frac{e^\mu}{4^n e^{2\mu}\frac{\pi^{5/6}\mu^{1/3}}{2^{1/3}3^{1/2}n^{5/6}}}\frac{4^{n-1}}{n2^r}\sqrt{2\pi}\frac{1}{\sqrt{3}}\left(\frac{2\pi^2 n}{\mu^4}\right)^{1/6}\sum_{|\ell-\mu|\leq\gamma\sqrt{\mu}}\binom{r}{\ell}\frac{\mu^\ell}{(\ell-1)!}$$

$$=\left(1+o(1)\right)\frac{1}{2e^\mu 2^r}\sum_{|\ell-\mu|\leq\gamma\sqrt{\mu}}\binom{r}{\ell}\frac{\mu^{\ell-1}}{(\ell-1)!}.$$

For the final term, recall that $r=\lfloor 2\mu+\rho\sqrt{\mu}\rfloor$ with $\rho\in[-R,R]$ and let $\lambda=(\ell-\mu)/\sqrt{\mu}\in[-\gamma,\gamma]$. Further assume for now that $2\mu+\rho\sqrt{\mu}$ is an integer (this can be obtained by replacing ρ with $\tilde{\rho}=\rho+O(1/\sqrt{\mu})$, but we prefer avoiding the use of tilde for clarity of the formulas). Then, Stirling's formula tells us that

$$\binom{r}{\ell}\frac{\mu^{\ell-1}}{(\ell-1)!}=\left(1+o(1)\right)\sqrt{\frac{4\pi\mu}{(2\pi\mu)^3}}\exp\left((2\mu+\rho\sqrt{\mu})\log\left(2\mu+\rho\sqrt{\mu}\right)-(\mu+\lambda\sqrt{\mu}-1)\log\left(\frac{\mu+\lambda\sqrt{\mu}-1}{e\mu}\right)\right.$$

$$\left.-(\mu+\lambda\sqrt{\mu})\log\left(\mu+\lambda\sqrt{\mu}\right)-(\mu+(\rho-\lambda)\sqrt{\mu})\log\left(\mu+(\rho-\lambda)\sqrt{\mu}\right)\right).$$

Observe that, for any constant $c>0$ and $x>0$ possibly diverging, we have

$$(c\mu+x\sqrt{\mu})\log\left(c\mu+x\sqrt{\mu}\right)=(c\mu+x\sqrt{\mu})\left[\log(c\mu)+\frac{x\sqrt{\mu}}{c\mu}-\frac{x^2\mu}{2c^2\mu^2}+O\left(\frac{x^3}{\mu^{3/2}}\right)\right]$$

$$=(c\mu+x\sqrt{\mu})\log(c\mu)+x\sqrt{\mu}+\frac{x^2}{2c}+O\left(\frac{x^3}{\sqrt{\mu}}\right).$$

For x being equal to either ρ or λ, it is bounded by γ. Thus, we can choose γ diverging slowly enough so that the $O(\cdot)$ term in the RHS above is always $o(1)$. Plugging this formula in the previous equation, and simplifying the terms, we obtain that

$$\binom{r}{\ell}\frac{\mu^{\ell-1}}{(\ell-1)!}=\left(1+o(1)\right)\frac{1}{\sqrt{2\pi\mu}}\exp\left((2\mu+\rho\sqrt{\mu})\log 2+\frac{\rho^2}{4}-\frac{\lambda^2}{2}-\frac{(\rho-\lambda)^2}{2}+\mu-\frac{\lambda^2}{2}\right)$$

$$=\left(1+o(1)\right)\frac{e^\mu 2^r}{\sqrt{2\pi\mu}}\exp\left(-\frac{\rho^2}{4}-\frac{3\lambda^2}{2}+\rho\lambda\right)$$

$$=\left(1+o(1)\right)\frac{e^\mu 2^r}{\sqrt{2\pi\mu}}\exp\left(-\frac{\rho^2}{12}-\frac{3}{2}\left(\lambda-\frac{\rho}{3}\right)^2\right).$$

Observe that, while we assumed here that ρ was such that $2\mu+\rho\sqrt{\mu}$ was an integer, this assumption can now be removed without loss of generality, since any $O(1/\sqrt{\mu})$ error terms on ρ would not affect the asymptotic behaviour.

Plugging this back into the lower bound for the probability $\mathbb{P}\big(\deg_{\mathbf{T}_n}(\varnothing) = r\big)$, we now see that

$$\mathbb{P}\big(\deg_{\mathbf{T}_n}(\varnothing) = r\big) \geq \big(1 + o(1)\big) \frac{1}{2e^\mu 2^r} \sum_{|\ell - \mu| \leq \gamma \sqrt{\mu}} \frac{e^\mu 2^r}{\sqrt{2\pi\mu}} \exp\left(-\frac{\rho^2}{12} - \frac{3}{2}\left(\lambda - \frac{\rho}{3}\right)^2 \right)$$

$$= \big(1 + o(1)\big) \frac{e^{-\rho^2/12}}{\sqrt{8\pi\mu}} \sum_{|\ell - \mu| \leq \gamma \sqrt{\mu}} \exp\left(-\frac{3}{2}\left(\frac{\ell - \mu}{\sqrt{\mu}} - \frac{\rho}{3}\right)^2 \right).$$

To finally conclude this proof, observe that the Riemann approximation of the integral applies here, leading to

$$\sum_{|\ell - \mu| \leq \gamma \sqrt{\mu}} \exp\left(-\frac{3}{2}\left(\frac{\ell - \mu}{\sqrt{\mu}} - \frac{\rho}{3}\right)^2 \right) \sim \sqrt{\mu} \int_{\mathbb{R}} e^{-\frac{3(x-\rho)^2}{2}} \, dx = \sqrt{\frac{2\mu\pi}{3}}.$$

The desired result then follows from plugging this formula in the previous equation. $\qquad\square$

Proof of Theorem 1.16. With this result the proof of the asymptotic distribution of the root is quite straightforward. Indeed, for any $t \in \mathbb{R}$, we see that

$$\mathbb{P}\left(\frac{\deg_{\mathbf{T}_n}(\varnothing) - 2\mu}{\sqrt{6\mu}} \leq t \right) = \mathbb{P}\left(\deg_{\mathbf{T}_n}(\varnothing) \leq \lfloor 2\mu + t\sqrt{6\mu}\rfloor \right)$$

$$\geq \sqrt{6\mu} \int_{-R}^{t} \mathbb{P}\left(\deg_{\mathbf{T}_n}(\varnothing) = \lfloor 2\mu + u\sqrt{6\mu}\rfloor \right) du,$$

for any $R > t$. Thus, by applying Proposition 6.4, it follows that

$$\mathbb{P}\left(\frac{\deg_{\mathbf{T}_n}(\varnothing) - 2\mu}{\sqrt{6\mu}} \leq t \right) \geq \big(1 + o(1)\big) \sqrt{6\mu} \int_{-R}^{t} \frac{e^{-(\sqrt{6}u)^2/12}}{\sqrt{12\pi\mu}} du.$$

By first letting $n \to \infty$ using that the $o(1)$ is uniformly small in $[-R, R]$, and then letting $R \to \infty$, the right-hand side above equals $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-u^2/2} du$, which is the distribution function of the standard normal random variable. The Portmanteau theorem now finishes the proof. $\qquad\square$

Remark 6.5 (Local CLT)**.** In fact, the same argument above shows that the inequality in the statement of Proposition 6.4 is in fact an equality, yielding a *local CLT*.

A Bounding functions with polynomial factor and exponent

This section focuses on proving the single result below, which we later use for the proof of both Theorem B.1 and Theorem D.1. The proof uses a method where exponentials are iterated until reaching the desired behaviour.

Lemma A.1. *For any real number $\alpha \in \mathbb{R}$ and any two strictly positive numbers $\beta, \gamma > 0$, consider the function*

$$F_t(x, y) = \frac{t^\alpha}{x^\beta} \exp\left(-\frac{t}{y^\gamma}\right).$$

Let $a_+ = a_+(t)$ be strictly positive and bounded, and $\epsilon = \epsilon(t)$ converging to 0 as t diverges. Then there exists $k = k(t)$ an integer depending on t and a sequence of increasing values $(a_0 \leq \ldots \leq a_k)$ (implicitly depending on t) such that $a_0 = \epsilon$, $a_k = a_+$, and satisfying

$$\lim_{t \to \infty} \sum_{i=1}^{k} F_t(a_{i-1}, a_i) = 0.$$

Proof. In order to prove the theorem, we directly provide a proper sequence and show that the desired sum converges to 0. Before doing so, note that requiring $a_0 = \epsilon$ is equivalent to requiring $a_0 \leq \epsilon$ as the sum is a decreasing function of a_0. Similarly, by observing that $F(a, b)$ converges to 0 for any fix $a, b > 0$ and that the sum is increasing with a_k (at fixed k), we can assume without loss of generality that $a_+ = 1$. Consider now the sequence defined inductively by $b_0 = 1$ and

$$b_{i+1} = \exp\left(\frac{\gamma t}{2\beta} b_i\right).$$

Further define $k = \min\{i : b_i \geq 1/\epsilon^\gamma\}$ and let $a_i = b_{k-i}^{-1/\gamma}$ for $0 \leq i \leq k$. We prove that $(a_0, \ldots, a_k)$ satisfies all the assumptions of the lemma.

Before proving that the desired sum converges to 0, let us show that the initial assumptions are met. By definition, we see that $a_k = b_0^{-1/\gamma} = 1$ since $b_0 = 1$. Moreover, by the definition of k, we know that $b_k \geq 1/\epsilon^\gamma$ and so $a_0 = b_k^{-1/\gamma} \leq \epsilon$. Finally, since $e^x \geq x$ for any $x \in \mathbb{R}$ and using that t can be chosen arbitrarily large, for example such that $t \geq 2\beta/\gamma$, we see that $b_{i+1} \geq b_i$ and so $a_{i+1} \geq a_i$. This proves that all initial assumptions are met and we now focus on proving the convergence of the sum.

Start by observing that, for any $1 \leq i \leq k$, we have

$$F_t(a_{i-1}, a_i) = \frac{t^\alpha}{a_{i-1}^\beta} \exp\left(-\frac{t}{a_i^\gamma}\right) = t^\alpha b_{k-i+1}^{\beta/\gamma} \exp\left(-t b_{k-i}\right) = t^\alpha \exp\left(-\frac{t}{2} b_{k-i}\right).$$

It follows that the sum can be re-written as

$$\sum_{i=1}^{k} F_t(a_{i-1}, a_i) = \sum_{i=0}^{k-1} t^\alpha \exp\left(-\frac{t}{2} b_i\right).$$

But now, use $e^x \geq x + 1$ for any $x \in \mathbb{R}$ and a simple induction to see that $b_i \geq i$, which leads to

$$\sum_{i=1}^{k} F_t(a_{i-1}, a_i) \leq \sum_{i=0}^{k-1} t^\alpha \exp\left(-\frac{t}{2} i\right) \leq \frac{t^\alpha e^{-t/2}}{1 - e^{-t/2}},$$

and the desired convergence to 0 follows. $\qquad\square$

Corollary A.2. *Let $\alpha \in \mathbb{R}$ be a real number and B, Γ be two continuous functions. Define the function*

$$F_t(x, y) = \frac{t^\alpha}{B(x)} \exp\left(-\frac{t}{\Gamma(y)}\right).$$

Assume that B, Γ are strictly positive on $(0, a_+]$ and that there exists $C_\beta, C_\gamma > 0$ and $\beta, \gamma > 0$ such that $B(x) \sim C_\beta x^\beta$ and $\Gamma(x) \sim C_\gamma x^\gamma$ when $x \to 0$. Then, for any $\epsilon = \epsilon(t)$ converging to 0 as t diverges and $\tilde{a}_+ = \tilde{a}_+(t) \leq a_+$, there exists

$k = k(t)$ an integer depending on t and a sequence of increasing values $(a_0 \leq \ldots \leq a_k)$ (implicitly depending on t) such that $a_0 = \epsilon$, $a_k = \tilde{a}_+$, and satisfying

$$\lim_{t \to \infty} \sum_{i=1}^{k} F_t(a_{i-1}, a_i) = 0.$$

Proof. From the assumption on B and Γ, we can find $C > 0$ such that, for any $0 \leq x \leq \tilde{a}_+$, we have $B(x) \geq C^{-1} x^{\beta}$ and $\Gamma(x) \leq C x^{\gamma}$. Now, let $\epsilon = a_0 \leq \cdots \leq a_k = \tilde{a}_+$ be the sequence satisfying

$$\lim_{t \to \infty} \sum_{i=1}^{k} \frac{t^{\alpha}}{a_{i-1}^{\beta}} \exp\left(-\frac{t}{a_i^{\gamma}}\right) = 0,$$

which exists according to Lemma A.1. Then, we observe that this allows to bound the sum of F_t, as

$$\sum_{i=1}^{k+1} F_t(a_{i-1}, a_i) \leq \sum_{i=1}^{k} \frac{C t^{\alpha}}{a_{i-1}^{\beta}} \exp\left(-\frac{t}{C a_i^{\gamma}}\right) = C^{1+\alpha} \sum_{i=1}^{k} \frac{(t')^{\alpha}}{a_{i-1}^{\beta}} \exp\left(-\frac{t'}{a_i^{\gamma}}\right),$$

where we used the change of variable $t' = t/C$. Using that $t \to \infty$ if and only if $t' \to \infty$ along with the definition of the a_i proves the desired convergence to 0. $\qquad \square$

B   Trees of bounded height

In this section, we provide some useful properties regarding the number of trees of bounded height. Recall from (1.8) we let $H_{n,m}$ be the number of trees of size n and height less than m, that is

$$H_{n,m} = \left| \left\{ T \in \mathfrak{t}_n : h(t) < m \right\} \right|.$$

A closed formula for $H_{n,m}$ is already known [12, (14)] and is given by

$$H_{n,m} = \frac{4^n}{m+1} \sum_{k=1}^{\lfloor m/2 \rfloor} \sin^2\left(\frac{\pi k}{m+1}\right) \cos^{2n-2}\left(\frac{\pi k}{m+1}\right) = \frac{4^n}{m+1} \sum_{k=1}^{\lfloor m/2 \rfloor} \frac{\tan^2\left(\frac{\pi k}{m+1}\right)}{\left(1 + \tan^2\left(\frac{\pi k}{m+1}\right)\right)^n}. \tag{B.1}$$

Using this formula, we now prove the following theorem, bounding $H_{n,m}$ for some choices of m.

Theorem B.1. *Let h_n be such that $4 \leq h_n \ll \sqrt{n}$. Then, for any $4 \leq m \leq h_n$, we have*

$$H_{n,m} = \frac{4^n}{m+1} \frac{\tan^2\left(\frac{\pi}{m+1}\right)}{\left(1 + \tan^2\left(\frac{\pi}{m+1}\right)\right)^n} + O\left(\frac{4^n}{m^3} \frac{1}{\left(1 + \tan^2\left(\frac{2\pi}{m+1}\right)\right)^n}\right)$$

$$= \left(1 + o(1)\right) \frac{4^n}{m+1} \frac{\tan^2\left(\frac{\pi}{m+1}\right)}{\left(1 + \tan^2\left(\frac{\pi}{m+1}\right)\right)^n}$$

where the $o(\cdot)$ and $O(\cdot)$ terms only depend on h_n, and are uniformly bounded over all such m. Moreover, the second asymptotic equality trivially holds when $m \in \{2, 3\}$.

Proof. Let h_n satisfy the conditions and consider $4 \leq m \leq h_n$. Observe that the result is equivalent to saying that the first two terms of (B.1) dominate the sum, with the first term being the dominant one. The fact that the second term is negligible with respect to the first one directly follows from their ratio being an exponential with a diverging term (recalling that $m \leq h_n$ and $n/h_n^2 \gg 1$). Thus, we now prove that the term $k = 2$ dominates all the terms corresponding to $k \geq 2$ in (B.1).

Consider first the case m finite. In that case $H_{n,m}$ is a finite sum of exponential terms in n with negative exponents. Moreover all exponents are distinct, so only the exponent of minimal value contributes to the asymptotic behaviour. Further using that $m \geq 4$ so that the sum has at least two terms, we see that the given asymptotic behaviour holds. Since this is true for m fixed and $n \to \infty$, we can also directly extend it to $\tilde{m} = \tilde{m}_n$ diverging slowly

enough (for example $\tilde{m} = (1-\epsilon)\sqrt{3\pi^2 n/\log n}$ for any $\epsilon > 0$ works, but the exact rate is not relevant to the rest of the proof). Thus the first statement of the lemma is proven for all $4 \le m \le \tilde{m}$ and we now focus on $\tilde{m} \le m \le h_n$ (assuming that $\tilde{m} \le h_n$). For the rest of the proof we repeatedly use that, under the assumption that $\tilde{m} \le m \le h_n \ll \sqrt{n}$, any $o(\cdot)$ and $O(\cdot)$ terms related to the fact that m or n/m^2 both diverge to infinity is uniform over all such m.

To extend the result of the lemma to all $m \le h_n$, we first prove that the terms of the sum with $k > \log m$ are negligible. Let us observe that, for any $\alpha, \beta > 0$ such that $\alpha + \beta < \pi/2$, we have

$$1 + \tan^2(\alpha + \beta) = 1 + \left(\frac{\tan\alpha + \tan\beta}{1 - \tan\alpha\tan\beta}\right)^2 = \frac{(1 + \tan^2\alpha)(1 + \tan^2\beta)}{(1 - \tan\alpha\tan\beta)^2} \ge (1 + \tan^2\alpha)(1 + \tan^2\beta).$$

Following this inequality, a simple induction leads to

$$1 + \tan^2\left(\frac{\pi k}{m+1}\right) \ge \left(1 + \tan^2\left(\frac{\pi}{m+1}\right)\right)^k.$$

Using this bound along with the fact that $k \mapsto \tan^2(\pi k/(m+1))$ is increasing, we observe that

$$\sum_{\log m \le k \le m/2} \frac{\tan^2\left(\frac{\pi k}{m+1}\right)}{\left(1 + \tan^2\left(\frac{\pi k}{m+1}\right)\right)^n} \le \tan^2\left(\frac{\pi m}{2(m+1)}\right) \sum_{\log m \le k \le m/2} \frac{1}{\left(1 + \tan^2\left(\frac{\pi}{m+1}\right)\right)^{kn}}$$

$$\le \tan^2\left(\frac{\pi}{2} - \frac{\pi}{2(m+1)}\right) \frac{1}{\left(1 + \tan^2\left(\frac{\pi}{m+1}\right)\right)^{n\log m}} \frac{1}{1 - \left(1 + \tan^2\left(\frac{\pi}{m+1}\right)\right)^{-n}}$$

$$= (1 + o(1)) \frac{4m^2}{\pi^2} e^{-\frac{\pi n}{(m+1)^2}\log m + O\left(\frac{n\log m}{m^4}\right)}$$

Recall that the second term is of order

$$\frac{\tan^2\left(\frac{2\pi}{m+1}\right)}{\left(1 + \tan^2\left(\frac{2\pi}{m+1}\right)\right)^n} = (1 + o(1)) \frac{4\pi^2}{m^2} e^{-\frac{4n\pi^2}{(m+1)^2} + O\left(\frac{n}{m^4}\right)}$$

to see that

$$\frac{\left(1 + \tan^2\left(\frac{2\pi}{m+1}\right)\right)^n}{\tan^2\left(\frac{2\pi}{m+1}\right)} \sum_{\log m \le k \le m/2} \frac{\tan^2\frac{\pi k}{m+1}}{\left(1 + \tan^2\frac{\pi k}{m+1}\right)^n} \le (1 + o(1)) \frac{m^4}{\pi^4} e^{-(1+o(1))\frac{n\pi^2}{m^2}\log m}.$$

The fact that the sum for $k \ge \log m$ is of smaller order than the second term simply follows from the fact that n/m^2 diverges and that $m^4 e^{-x\log m}$ converges to 0 as $m, x \to \infty$. We now prove that the rest of the sum is negligible.

Observe that, uniformly over any $\tilde{m} \le m \le h_n$ and $2 \le k < \log m$, we have

$$\frac{\tan^2\left(\frac{\pi k}{m+1}\right)}{\left(1 + \tan^2\left(\frac{\pi k}{m+1}\right)\right)^n} = (1 + o(1)) \frac{\pi^2 k^2}{m^2} e^{-\frac{\pi^2 k^2 n}{(m+1)^2} + O\left(\frac{nk^4}{m^4}\right)} = e^{-(1+o(1))\frac{\pi^2 k^2 n}{m^2}}.$$

From this behaviour, we see that the terms are decreasing in k. Thus, for any $3 \le a \le b < \log m$, using the minimal term to bound the sum along with the fact that $a \le b$, it follows

$$\frac{\left(1 + \tan^2\left(\frac{2\pi}{m+1}\right)\right)^2}{\tan^2\left(\frac{2\pi}{m+1}\right)} \sum_{k=a}^{b} \frac{\tan^2\left(\frac{\pi k}{m+1}\right)}{\left(1 + \tan^2\left(\frac{\pi k}{m+1}\right)\right)^n} \le (1 + o(1)) b^3 e^{-(1+o(1))\frac{\pi^2(a^2-4)n}{(m+1)^2}},$$

where all $o(\cdot)$ terms are uniform over all the parameters considered here. Apply now Corollary A.2 with $\alpha = 0$, $B(x) = x^3$, $\Gamma(x) = x^2/(4 - x^2)$, $a_+ = 1/3$, mapping t with $\pi^2 n/(m+1)^2$, and $\epsilon = 1/\log n$ to find $a_0 \le \cdots \le a_K$ such that

$$\lim_{n \to \infty} \sum_{i=1}^{K} \frac{1}{a_{i-1}^2} \exp\left(-\left(\frac{1}{a_i^2} - 1\right) \frac{\pi^2 n}{(m+1)^2}\right) = 0.$$

Using this sequence and combining with the previous inequality, we obtain that

$$
\frac{\left(1+\tan^2\left(\frac{2\pi}{m+1}\right)\right)^2}{\tan^2\left(\frac{2\pi}{m+1}\right)} \sum_{3\le k<\log m} \frac{\tan^2\left(\frac{\pi k}{m+1}\right)}{\left(1+\tan^2\left(\frac{\pi k}{m+1}\right)\right)^n} = \sum_{i=1}^{K} \frac{\left(1+\tan^2\left(\frac{2\pi}{m+1}\right)\right)^2}{\tan^2\left(\frac{2\pi}{m+1}\right)} \sum_{1/a_i\le k<1/a_{i-1}} \frac{\tan^2\left(\frac{\pi k}{m+1}\right)}{\left(1+\tan^2\left(\frac{\pi k}{m+1}\right)\right)^n}
$$

$$
\le \left(1+o(1)\right)\sum_{i=1}^{K} \frac{1}{a_{i-1}^2}\exp\left(-\left(\frac{1}{a_i^2}-1\right)\frac{\pi^2 n}{(m+1)^2}+O\left(\frac{n\log^4 n}{m^4}\right)\right),
$$

where we used the monotonicity of the upper bound (increasing with a_i and decreasing with a_{i-1}) to relax the integer constraints. By the definition of the a_i terms, the upper bound converges to 0 and so the sum of the terms from 3 to $\log n$ are negligible with respect to the second term. Combined with the fact that the sum after $\log n$ is also negligible proves the desired result. □

The number of trees of bounded height is very useful in order to characterize the behaviour of the partition function Z_n introduced in (1.4), as its definition implies that

$$
Z_n^\mu = \sum_{T\in\mathfrak{t}_n} e^{-\mu h(T)} = \sum_{m=1}^{n} \left(H_{n,m+1}-H_{n,m}\right)e^{-\mu m}.
$$

Moreover, using that $H_{n,m}=C_{n-1}$ whenever $m\ge n$ and that $H_{n,1}=0$ (assuming $n>1$), we see that the previous formula re-writes as

$$
Z_n^\mu = C_{n-1}e^{-\mu(n-1)} + e^\mu(e^\mu-1)\sum_{m=2}^{n-1} H_{n,m}e^{-\mu(m+1)}. \tag{B.2}
$$

Thus, it is rather natural to introduce the sum

$$
W_n = W_n^\mu = \frac{1}{4^n}\sum_{m=3}^{n} H_{n,m-1}e^{-\mu m}, \tag{B.3}
$$

which itself contains all of the non-trivial asymptotic behaviour of Z_n, as stated by the following lemma.

Lemma B.2. *Let $\mu=\mu_n$ be such that $\mu\gg 1/\sqrt{n}$. Then, Z_n as defined in (1.4) is related to W_n introduced in (B.3) by*

$$
Z_n = \left(1+o(1)\right)4^n e^\mu(e^\mu-1)W_n
$$

Proof. The statement of this lemma simply says that the term $C_{n-1}e^{-\mu(n-1)}$ can be ignored in the formula (B.2). To see that it is the case, first note that, for any $m\le n$,

$$
W_n \ge \frac{1}{4^n}H_{n,m}e^{-\mu m} \ge \frac{\tan^2\frac{\pi}{m}}{m}\frac{e^{-\mu m}}{\left(1+\tan^2\frac{\pi}{m}\right)^n},
$$

from which it follows that, for any $m\gg 1$, we have

$$
\frac{C_{n-1}e^{-\mu(n-1)}}{4^n e^\mu(e^\mu-1)W_n} \le \left(C+o(1)\right)\frac{m^3}{n^{3/2}}\frac{1}{e^\mu-1}\exp\left(-\mu(n-m)+\frac{n\pi^2}{m^2}+O\left(\frac{n}{m^4}\right)\right).
$$

Use now $m=\lfloor\sqrt{n}\rfloor$ to obtain that

$$
\frac{C_{n-1}e^{-\mu(n-1)}}{4^n e^\mu(e^\mu-1)W_n} \le \left(C+o(1)\right)\frac{e^{-(1+o(1))\mu n}}{e^\mu-1}.
$$

Recalling that $\mu\gg 1/\sqrt{n}$, we see that the right-hand side necessarily converges to 0 and so the asymptotic behaviour of Z_n is characterized by the term W_n. Observe that the convergence occurs further away than $\mu\gg 1/\sqrt{n}$ and can be pushed as far as $\mu\gg(\log n)/n$ with the same method. However, we do not need this result to go further than $1/\sqrt{n}$ and so keep it as is here. □

40

C Exponent of the partition function

Thanks to Lemma B.2, we see that characterizing the behaviour of Z_n is similar to characterizing the behaviour of W_n. Furthermore, combining the formula of W_n from (B.3) along with Theorem B.1, we see that, approximately,

$$W_n \asymp \sum_{m=3}^{n} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} \frac{e^{-\mu m}}{\left(1 + \tan^2\left(\frac{\pi}{m}\right)\right)^n}.$$

Observe that this approximation only works as long as the trees contributing to the partition function all have their height of smaller order than $\sqrt{n}$. While this might seem like a circular argument, we in fact prove that only the trees of height $\ll \sqrt{n}$ contribute to W_n, and only intend to provide a high level view of the argument here.

In order to characterize this sum, we see that the first fraction in the sum behaves as a polynomial, whereas the second one has exponential terms. It is thus natural to focus on the behaviour of the exponential terms in order to characterize W_n, as we expect them to be the main contributor to the sum. We now use this observation to define and study the exponent function.

At this point, it is useful to redefine for convenience the function λ_x for any $x > 0$ and its minimizer t_x, already defined in (1.6),

$$\lambda_x(t) = xt + \log\left(1 + \tan^2\frac{\pi}{t}\right), \tag{C.1}$$

whose domain is $t \in (2, \infty)$. From this definition, we directly see that λ_x diverges to ∞ when t approaches any of its boundary, implying that it admits a minimum (non necessarily unique). Further noting that

$$\lambda_x'(t) = x - \frac{2\pi}{t^2} \tan\frac{\pi}{t},$$

we see that λ_x is strictly convex and thus the minimum t_x is unique. Observe that t_x satisfies

$$x = \frac{2\pi}{t_x^2} \tan\frac{\pi}{t_x}.$$

We now provide an important result regarding the asymptotic behaviour of t_x and λ_x when x converges to 0.

Proposition C.1. *Let $x > 0$ and λ_x be the function defined in (C.1). Let t_x be the minimizer of λ_x. Then, as x converges to 0, we have the following asymptotic behaviour.*

(**1**) *The minimizer t_x of the function satisfies*

$$t_x = \left(\frac{2\pi^2}{x}\right)^{1/3}\left[1 + \frac{1}{9}\left(\frac{\pi x}{2}\right)^{2/3} + O\left(x^{4/3}\right)\right] = \left(\frac{2\pi^2}{x}\right)^{1/3} + O\left(x^{1/3}\right)$$

(**2**) *The minimum of the function satisfies*

$$\lambda_x(t_x) = 3\left(\frac{\pi x}{2}\right)^{2/3} + \frac{1}{6}\left(\frac{x}{2\pi}\right)^{4/3} + O\left(x^2\right),$$

(**3**) *For any $\epsilon = \epsilon_x$ converging to 0 as x goes to 0 (allowed to be negative), the asymptotic behaviour of λ_x at $(1+\epsilon)t_x$ is given by*

$$\lambda_x\big((1+\epsilon)t_x\big) - \lambda_x(t_x) = 3\left(\frac{\pi x}{2}\right)^{2/3}\epsilon^2 + O\left(\epsilon^3 x^{2/3} + \epsilon^2 x^{4/3}\right),$$

and the $O(\cdot)$ term is uniform over any window of the form $[-\delta, \delta]$ for $\delta = \delta_x \ll 1$.

Proof. We start with the behaviour of t_x as given in (**1**). Since the second equality directly follows from the first one, we only show that the first asymptotic holds. While the method to obtain this result in the first place requires inductively deducting the terms of t_x, we now simply verify that the given formula satisfies our assumption. Indeed, by first observing that t_x necessarily diverges to ∞ when x converges to 0, and by using the Taylor expansion of the tangent function, we obtain

$$\frac{2\pi}{t_x^2}\tan\left(\frac{\pi}{t_x}\right) = \frac{2\pi}{t_x^2}\left(\frac{\pi}{t_x} + \frac{\pi^3}{3t_x^3} + O\left(\frac{1}{t_x^5}\right)\right) = \frac{2\pi^2}{t_x^3} + \frac{2\pi^4}{3t_x^5} + O\left(\frac{1}{t_x^7}\right).$$

Thus, replacing t_x by the given formula leads to

$$\frac{2\pi}{t_x^2}\tan^2\left(\frac{\pi}{t_x}\right) = \frac{2\pi^2}{\frac{2\pi^2}{x}}\left[1-\frac{1}{3}\left(\frac{\pi x}{2}\right)^{2/3}+O\left(x^{4/3}\right)\right]+\frac{2\pi^4}{3\left(\frac{2\pi^2}{x}\right)^{5/3}}\left[1+O\left(x^{2/3}\right)\right]+O\left(x^{7/3}\right)$$

$$= x+\left[\frac{2\pi^4}{3(2\pi^2)^{5/3}}-\frac{1}{3}\left(\frac{\pi}{2}\right)^{2/3}\right]x^{5/3}+O\left(x^{7/3}\right).$$

Observe that the $x^{5/3}$ order term in the previous formula is actually multiplied by 0. This implies that the approximation is correct and furthermore we see that the $O(\cdot)$ terms are exactly the right order as it matches the terms from the expansion of the tangent function. We now move on to the asymptotic behaviour of λ_x.

Starting with the definition of λ_x from (C.1), since t_x diverges, we have

$$\lambda_x(t_x) = xt_x+\log\left(1+\left(\frac{\pi}{t_x}+\frac{\pi^3}{3t_x^3}+O\left(\frac{1}{t_x^5}\right)\right)^2\right)$$

$$= xt_x+\log\left(1+\frac{\pi^2}{t_x^2}\left(1+\frac{2\pi^2}{3t_x^2}+O\left(\frac{1}{t_x^4}\right)\right)\right)$$

$$= xt_x+\left[\frac{\pi^2}{t_x^2}\left(1+\frac{2\pi^2}{3t_x^2}+O\left(\frac{1}{t_x^4}\right)\right)-\frac{1}{2}\frac{\pi^4}{t_x^4}\left(1+O\left(\frac{1}{t_x^2}\right)\right)+O\left(\frac{1}{t_x^6}\right)\right]$$

$$= xt_x+\frac{\pi^2}{t_x^2}+\frac{\pi^4}{6t_x^4}+O\left(\frac{1}{t_x^6}\right).$$

To prove (**2**), plug in this equation the previous asymptotic behaviour of t_x to obtain

$$\lambda_x(t_x) = x\left(\frac{2\pi^2}{x}\right)^{1/3}\left[1+\frac{1}{9}\left(\frac{\pi x}{2}\right)^{2/3}+O\left(x^{4/3}\right)\right]+\frac{\pi^2}{\left(\frac{2\pi^2}{x}\right)^{2/3}}\left[1-\frac{2}{9}\left(\frac{\pi x}{2}\right)^{2/3}+O\left(x^{4/3}\right)\right]$$

$$+\frac{\pi^4}{6\left(\frac{2\pi^2}{x}\right)^{4/3}}\left[1+O\left(x^{2/3}\right)\right]+O\left(x^2\right)$$

$$= 3\left(\frac{\pi x}{2}\right)^{2/3}+\frac{1}{6}\left(\frac{x}{2\pi}\right)^{4/3}+O\left(x^2\right),$$

as desired.

Finally, for (**3**), start with the definition of λ_x from (C.1) to see that

$$\lambda_x\left((1+\epsilon)t_x\right)-\lambda_x(t_x) = x\epsilon t_x+\log\left(\frac{1+\tan^2\left(\frac{\pi}{(1+\epsilon)t_x}\right)}{1+\tan^2\left(\frac{\pi}{t_x}\right)}\right).$$

Now, using the trigonometric identity

$$1+\tan^2(\alpha+\beta) = 1+\left(\frac{\tan\alpha+\tan\beta}{1-\tan\alpha\tan\beta}\right)^2 = \frac{\left(1+\tan^2\alpha\right)\left(1+\tan^2\beta\right)}{\left(1-\tan\alpha\tan\beta\right)^2}$$

with $\alpha = \pi/t_n$ and $\beta = -\pi\epsilon_n/((1+\epsilon_n)t_n)$ leads to

$$\lambda_x\left((1+\epsilon)t_x\right)-\lambda_x(t_x) = x\epsilon t_x+\log\left(\frac{1+\tan^2\left(\frac{\pi\epsilon}{(1+\epsilon)t_x}\right)}{\left(1+\tan\frac{\pi}{t_x}\tan\left(\frac{\pi\epsilon}{(1+\epsilon)t_x}\right)\right)^2}\right).$$

Using that $\epsilon \ll 1$ and $t_x \gg 1$, all the terms within the logarithm can be extended using Taylor approximation. Moreover, we recall that t_x being the minimum of λ_x, it satisfies

$$x = \frac{2\pi}{t_x^2}\tan\left(\frac{\pi}{t_x}\right),$$

42

which leads to the following approximation

$$\lambda_x\big((1+\epsilon)t_x\big) - \lambda_x(t_x) = x\epsilon t_x + \left[\frac{\pi^2\epsilon^2}{(1+\epsilon)^2 t_x^2} - 2\tan\left(\frac{\pi}{t_x}\right)\frac{\pi\epsilon}{(1+\epsilon)t_x} + O\left(\frac{\epsilon^2}{t_x^4}\right)\right]$$

$$= x\epsilon t_x + \frac{\pi^2\epsilon^2}{(1+\epsilon)^2 t_x^2} - \frac{x\epsilon t_x}{1+\epsilon} + O\left(\frac{\epsilon^2}{t_x^4}\right)$$

$$= \left[(1+\epsilon)xt_x + \frac{\pi^2}{t_x^2}\right]\frac{\epsilon^2}{(1+\epsilon)^2} + O\left(\frac{\epsilon^2}{t_x^4}\right).$$

To conclude the proof, simply use the asymptotic value of t_x obtained above to find that

$$\lambda_x\big((1+\epsilon)t_x\big) - \lambda_x(t_x) = \left(\frac{\pi x}{2}\right)^{2/3}\left[2(1+\epsilon)\left[1+O\big(x^{2/3}\big)\right] + \left[1+O\big(x^{2/3}\big)\right]\right]\frac{\epsilon^2}{(1+\epsilon)^2} + O\big(\epsilon^2 x^{4/3}\big)$$

$$= 3\left(\frac{\pi x}{2}\right)^{2/3}\epsilon^2 + O\big(\epsilon^3 x^{2/3} + \epsilon^2 x^{4/3}\big).$$

The uniformity of the $O(\cdot)$ terms directly comes from the fact that all terms depending on x are used within the $O(\cdot)$ and not as constants. □

D   Asymptotic of the partition function: part I

The goal of this section is to show the following theorem, stating the asymptotic behaviour of the term W_n introduced in (B.3).

Theorem D.1. *Let $\mu = \mu_n$ be such that $1/\sqrt{n} \ll \mu \ll n^{1/4}$. Then, as n diverges to infinity, the sum W_n defined in (B.3) satisfies*

$$W_n = \big(1 + o(1)\big)\frac{\pi^{5/6}}{2^{1/3}3^{1/2}}\frac{\mu^{1/3}}{n^{5/6}}\exp\left(-3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right).$$

Due to its proof being rather intricate, we split it into three sub-sections. We start by focusing on the terms of the sum which end up being the main contributors (in Section D.1), then prove that the remaining terms are negligible (in Section D.2), and finally combine all the different results together to prove that Theorem D.1 holds (in Section D.3).

D.1   Main contribution

In this section we focus on the part of the sum that contributes the most in the sum of W_n as defined in (B.3). This corresponds to the terms close to t_n where t_n is the minimizer of the function $\lambda_{\mu/n}$, defined in (C.1). Before proving the desired convergence, we provide a quick lemma regarding the integral approximation of a Gaussian sum.

Lemma D.2. *Let μ_n, σ_n be arbitrary (with $\sigma_n > 0$) and $\ell_n, r_n \gg 1$ as well as $\ell_n/\sigma_n, r_n/\sigma_n \gg 1$. Then, we can approximate the gaussian integral with the sum from $\mu_n - \ell_n$ to $\mu_n + r_n$ as such*

$$\left|\sum_{\mu_n - \ell_n < m < \mu_n + r_n}\exp\left(-\frac{(m-\mu_n)^2}{2\sigma_n^2}\right) - \sqrt{2\pi}\sigma_n\right| \le 1 + o(\sigma_n).$$

Proof. Using that $u \mapsto e^{-u^2/2\sigma_n^2}$ is decreasing on $[0,\infty)$, for any $m \ge \lceil\mu_n\rceil + 1 \ge \mu_n + 1$, the comparison to the integral gives us that

$$\exp\left(-\frac{(m-\mu_n)^2}{2\sigma_n^2}\right) \le \int_{m-1}^m \exp\left(-\frac{(u-\mu_n)^2}{2\sigma_n^2}\right)du \le \exp\left(-\frac{(m-\mu_n-1)^2}{2\sigma_n^2}\right).$$

Thus, by summing on m from $\lceil\mu_n\rceil + 1$ to $\lceil\mu_n + r_n\rceil - 1$, a telescopic sum appears and leads to

$$0 \le \int_{\lceil\mu_n\rceil}^{\lceil\mu_n+r_n\rceil-1}\exp\left(-\frac{(u-\mu_n)^2}{2\sigma_n^2}\right)du - \sum_{m=\lceil\mu_n\rceil+1}^{\lceil\mu_n+r_n\rceil-1}\exp\left(-\frac{(m-\mu_n)^2}{2\sigma_n^2}\right)$$

$$\leq \exp\left(-\frac{(\lceil \mu_n \rceil - \mu_n)^2}{2\sigma_n^2}\right) - \exp\left(-\frac{(\lceil \mu_n + r_n \rceil - \mu_n - 1)^2}{2\sigma_n^2}\right)$$

$$\leq \exp\left(-\frac{(\lceil \mu_n \rceil - \mu_n)^2}{2\sigma_n^2}\right).$$

Observe that by subtracting the last term everywhere and reversing all inequality by multiplying them with -1, this is the same as

$$0 \leq \sum_{m=\lceil \mu_n \rceil}^{\lceil \mu_n + r_n \rceil - 1} \exp\left(-\frac{(m - \mu_n)^2}{2\sigma_n^2}\right) - \int_{\lceil \mu_n \rceil}^{\lceil \mu_n + r_n \rceil - 1} \exp\left(-\frac{(u - \mu_n)^2}{2\sigma_n^2}\right) du$$

$$\leq \exp\left(-\frac{(\lceil \mu_n \rceil - \mu_n)^2}{2\sigma_n^2}\right)$$

In a similar manner, using that $u \mapsto e^{-u^2/2\sigma_n^2}$ is increasing on $(-\infty, 0]$ and by summing over m from $\lfloor \mu_n - \ell_n \rfloor + 1$ to $\lceil \mu_n \rceil - 1$, we obtain

$$0 \leq \sum_{m=\lfloor \mu_n - \ell_n \rfloor + 1}^{\lceil \mu_n \rceil - 1} \exp\left(-\frac{(m - \mu_n)^2}{2\sigma_n^2}\right) - \int_{\lfloor \mu_n - \ell_n \rfloor}^{\lceil \mu_n \rceil - 1} \exp\left(-\frac{(u - \mu_n)^2}{2\sigma_n^2}\right) du$$

$$\leq \exp\left(-\frac{(\lceil \mu_n \rceil - \mu_n - 1)^2}{2\sigma_n^2}\right) - \exp\left(-\frac{(\lfloor \mu_n - \ell_n \rfloor - \mu_n)^2}{2\sigma_n^2}\right)$$

$$\leq \exp\left(-\frac{(\lceil \mu_n \rceil - \mu_n - 1)^2}{2\sigma_n^2}\right).$$

Before concluding, observe that, as m is an integer, we have that $m < \mu_n + r_n$ is equivalent to $m \leq \lceil \mu_n + r_n \rceil - 1$ and $m > \mu_n - \ell_n$ is equivalent to $m \geq \lfloor \mu_n - \ell_n \rfloor + 1$. Thus, if we combine the previous two displayed inequalities we obtain on one hand that

$$\sum_{\mu_n - \ell_n < m < \mu_n + r_n} \exp\left(-\frac{(m - \mu_n)^2}{2\sigma_n^2}\right) - \int_{\lfloor \mu_n - \ell_n \rfloor}^{\lceil \mu_n + r_n \rceil - 1} \exp\left(-\frac{(u - \mu_n)^2}{2\sigma_n^2}\right) du \geq -\int_{\lceil \mu_n \rceil - 1}^{\lceil \mu_n \rceil} e^{-(u - \mu_n)^2/2\sigma_n^2} du.$$

Observing that the exponential term is less than 1, the lower bound is itself larger than -1. On the other hand, we have that

$$\sum_{\mu_n - \ell_n < m < \mu_n + r_n} \exp\left(-\frac{(m - \mu_n)^2}{2\sigma_n^2}\right) - \int_{\lfloor \mu_n - \ell_n \rfloor}^{\lceil \mu_n + r_n \rceil - 1} \exp\left(-\frac{(u - \mu_n)^2}{2\sigma_n^2}\right) du$$

$$\leq \exp\left(-\frac{(\lceil \mu_n \rceil - \mu_n)^2}{2\sigma_n^2}\right) + \exp\left(-\frac{(\lceil \mu_n \rceil - \mu_n - 1)^2}{2\sigma_n^2}\right) - \int_{\lceil \mu_n \rceil - 1}^{\lceil \mu_n \rceil} e^{-(u - \mu_n)^2/2\sigma_n^2} du,$$

and by observing that the integral is larger than the minimum of the two other values and using that the exponential terms are less than 1, we can replace the upper bound by 1. The two previous results thus imply that

$$\left| \sum_{\mu_n - \ell_n < m < \mu_n + r_n} \exp\left(-\frac{(m - \mu_n)^2}{2\sigma_n^2}\right) - \int_{\lfloor \mu_n - \ell_n \rfloor}^{\lceil \mu_n + r_n \rceil - 1} \exp\left(-\frac{(u - \mu_n)^2}{2\sigma_n^2}\right) du \right| \leq 1.$$

To conclude the proof, simply recall ℓ_n/σ_n and r_n/σ_n both diverge to see that

$$\int_{\lfloor \mu_n - \ell_n \rfloor}^{\lceil \mu_n + r_n \rceil - 1} \exp\left(-\frac{(u - \mu_n)^2}{2\sigma_n^2}\right) du = \sigma_n \int_{\frac{\lfloor \mu_n - \ell_n \rfloor - \mu_n}{\sigma_n}}^{\frac{\lceil \mu_n + r_n \rceil - 1 - \mu_n}{\sigma_n}} \exp\left(-\frac{u^2}{2}\right) du$$

$$= \sigma_n\left(\sqrt{2\pi} + o(1)\right),$$

leading to the desired result. □

With this lemma and the results from Section C regarding the exponent function λ_x, we now have all the tools to prove the asymptotic behaviour of the main terms of the sum in (B.3).

Proposition D.3. *Recall $t_{\mu/n}$ is the minimizer of $\lambda_{\mu/n}$. Let ϵ_n be such that $\epsilon_n \ll 1$ and $\epsilon_n t_n \gg 1$. Further assume that $\epsilon_n(\mu^2 n)^{1/6} \gg 1$. Then, under these assumptions, we have that*

$$\frac{1}{4^n} \sum_{|m - t_{\mu/n}| < \epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = \big(1 + o(1)\big) \frac{\pi^{5/6}}{2^{1/3} 3^{1/2}} \frac{\mu^{1/3}}{n^{5/6}} \exp\left(-3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}\right).$$

Proof. We start by recalling that we assume here that $1/\sqrt{n} \ll \mu \ll n^{1/4}$. Moreover, we observe that this implies that $\mu/n \ll 1$ and so we can apply the results of Proposition C.1 to $x = \mu/n$. In that scenario, and since $t_{\mu/n}$ is of order $(n/\mu)^{1/3}$ and that $\epsilon_n \ll 1$, we can see that $1 \ll (1 - \epsilon) t_n \le (1 + \epsilon_n) t_{\mu/n} \ll \sqrt{n}$. Thus we can uniformly apply the approximation of $H_{n,m-1}$ from Theorem B.1 and see that

$$\frac{1}{4^n} \sum_{|m - t_{\mu/n}| < \epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = \big(1 + o(1)\big) \frac{1}{4^n} \sum_m \frac{4^n}{m} \frac{\tan^2\left(\frac{\pi}{m}\right)}{\left(1 + \tan^2\left(\frac{\pi}{m}\right)\right)^n} e^{-\mu m}$$

$$= \big(1 + o(1)\big) \frac{\pi^2}{t_n^3} \sum_m e^{-n \lambda_{\mu/n}(m)},$$

where we used the fact that $\epsilon_n \ll 1$ and $t_n \gg 1$ to remove the polynomial terms out of the sum. Now, for any sequence $m = m_n$ between $(1 - \epsilon_n) t_{\mu/n}$ and $(1 + \epsilon_n) t_{\mu/n}$, the sequence $\gamma_n = m/t_{\mu/n} - 1 \ll 1$ and we can apply Proposition C.1 (**3**) to obtain that

$$e^{-n \lambda_{\mu/n}(m)} = e^{-n \lambda_{\mu/n}(t_{\mu/n})} \exp\left(-n\big(3 + o(1)\big)\left(\frac{\pi \mu}{2n}\right)^{2/3} \gamma_n^2\right),$$

where the $o(\cdot)$ term is uniform over such m. Thus, plugging this formula back into the previous sum and replacing γ_n by its value in relation to m, we obtain

$$\frac{1}{4^n} \sum_{|m - t_{\mu/n}| < \epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = \big(1 + o(1)\big) \frac{\pi^2}{t_{\mu/n}^3} e^{-n \lambda_{\mu/n}(t_{\mu/n})} \sum_m e^{-n \frac{3 + o(1)}{t_{\mu/n}^2} \left(\frac{\pi \mu}{2n}\right)^{2/3} (m - t_{\mu/n})^2}.$$

Consider now the statement of Lemma D.2 with $\mu = t_{\mu/n}$, $2\sigma_n^2 = \frac{1}{n} \frac{t_{\mu/n}^2}{3 + o(1)} \left(\frac{2n}{\pi \mu}\right)^{2/3}$, and $\ell_n, r_n = \epsilon_n t_{\mu/n}$. First of all, thanks to the asymptotic behaviour of $t_{\mu/n}$ following from Proposition C.1 (**1**), we establish that

$$\sigma_n^2 = \big(1 + o(1)\big) \frac{1}{n} \left(\frac{2\pi^2 n}{\mu}\right)^{2/3} \frac{1}{6} \left(\frac{2n}{\pi \mu}\right)^{2/3} = \big(1 + o(1)\big) \frac{1}{3} \left(\frac{2\pi^2 n}{\mu^4}\right)^{1/3}.$$

It is worth noting that the assumption that $\mu \ll n^{1/4}$ implies that $\sigma_n \to \infty$. Now, we already know that $\ell_n = r_n \gg 1$ since $\epsilon_n t_{\mu/n} \gg 1$. For the second assumption, observe that

$$\frac{\ell_n}{\sigma_n} = \big(1 + o(1)\big) \epsilon_n \left(\frac{2\pi^2 n}{\mu}\right)^{1/3} \sqrt{3} \left(\frac{\mu^4}{2\pi^2 n}\right)^{1/6} = \big(C + o(1)\big) \epsilon_n (\mu^2 n)^{1/6},$$

which diverges once again thanks to the assumption on ϵ_n. This means that we can apply Lemma D.2 to the sum above and, further using that $\sigma_n \to \infty$, obtain

$$\frac{1}{4^n} \sum_{|m - t_{\mu/n}| < \epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = \big(1 + o(1)\big) \frac{\pi^2}{t_{\mu/n}^3} e^{-n \lambda_{\mu/n}(t_{\mu/n})} \sqrt{2\pi} \sigma_n.$$

To conclude this proof recall the asymptotic behaviour of $t_{\mu/n}$ and σ_n as well as that of $\lambda_{\mu/n}(t_{\mu/n})$ from Proposition C.1 (**1**) and (**2**) to see that

$$\frac{1}{4^n} \sum_{|m - t_{\mu/n}| < \epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = \big(1 + o(1)\big) \frac{\pi^2}{\frac{2\pi^2 n}{\mu}} e^{-3n\left(\frac{\pi \mu}{2n}\right)^{2/3} + O\left(n\left(\frac{\mu}{n}\right)^{4/3}\right)} \sqrt{2\pi} \frac{1}{\sqrt{3}} \left(\frac{2\pi^2 n}{\mu^4}\right)^{1/6}$$

$$= \big(1 + o(1)\big) \frac{\pi^{5/6}}{2^{1/3} 3^{1/2}} \frac{\mu^{1/3}}{n^{5/6}} e^{-3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}},$$

where we used that $O(n(\mu/n)^{4/3}) = O((\mu^4/n)^{1/3}) = o(1)$. This result is exactly the claimed asymptotic behaviour. $\square$

D.2 Negligible terms

We now prove that Theorem D.1 holds by showing that the terms away from $t_{\mu/n}$ in the sum (B.3) are negligible, where $t_{\mu/n}$ is the location of the minimum of the function $\lambda_{\mu/n}$ from (C.1). Before doing so, we consider the sequence of terms defined by

$$X_{n,m} = \frac{1}{4^n} \frac{n^{5/6}}{\mu^{1/3}} H_{n,m-1} \exp\left(-\mu m + 3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}\right), \tag{D.1}$$

which exactly correspond the terms of the definition of W_n in (B.3) divided by the desired asymptotic behaviour of W_n from Theorem D.1 (up to constant terms). Thus, thanks to Proposition D.3 which shows that the main contribution of the terms in W_n has the desired asymptotic behaviour, it suffices to show that the sum of the $X_{n,m}$ when $|m - t_{\mu/n}| \geq \epsilon_n t_{\mu/n}$ converges to 0 for some appropriate ϵ_n. We divide this proof in two parts, depending on the absolute value of $m - t_{\mu/n}$. We start with the easiest one of the two, corresponding to $m - t_{\mu/n} > 0$.

Proposition D.4. *Recall* $t_{\mu/n} = \operatorname{argmin}_{t\in(0,2)} \lambda_{\mu/n}(t)$, *and* $X_{n,m}$ *be as defined as in* (D.1), *and* $\epsilon_n \ll 1$ *be such that* $\epsilon_n t_n \gg 1$ *and* $\epsilon_n^2 (\mu^2 n)^{1/3} \gg \log(\mu^2 n)$. *Then, the sum over* $m \geq (1 + \epsilon_n) t_{\mu/n}$ *of* $X_{n,m}$ *is negligible, as given by*

$$\lim_{n\to\infty} \sum_{(1+\epsilon_n)t_{\mu/n}\leq m\leq n} X_{n,m} = 0.$$

Proof. We prove this bound by splitting the sum in two parts. For the highest values of m, namely $m > c t_{\mu/n}$ for some fixed $c > 0$, we crudely use that $H_{n,m-1} \leq C_{n-1}$ where C_{n-1} is the $(n-1)$-th Catalan number. This implies that

$$\sum_{m\geq ct_{\mu/n}} X_{n,m} \leq \frac{C_{n-1}}{4^n} \frac{n^{5/6}}{\mu^{1/3}} \exp\left(3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right) \sum_{m\geq ct_{\mu/n}} e^{-\mu m}$$

$$\leq \frac{C_{n-1}}{4^n} \frac{n^{5/6}}{\mu^{1/3}} \frac{1}{1-e^{-\mu}} \exp\left(3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} - \mu c t_{\mu/n}\right).$$

We now recall that $C_{n-1}/4^n \sim 1/(4\sqrt{\pi} n^{3/2})$ and that Proposition C.1 (**1**) tells us that

$$t_{\mu/n} = \left(\frac{2\pi^2 n}{\mu}\right)^{1/3} + O\left(\left(\frac{\mu}{n}\right)^{1/3}\right).$$

Moreover, using that $e^\mu \geq 1 + \mu$, we can check that $1/(1 - e^{-\mu}) \leq (1+\mu)/\mu$. Finally, combining the previous remarks, it follows that

$$\sum_{m\geq ct_{\mu/n}} X_{n,m} \leq \frac{1}{4\sqrt{\pi}} \frac{(1+\mu)}{n^{2/3}\mu^{4/3}} \exp\left(3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} - c(2\pi^2\mu^2 n)^{1/3} + O\left(\left(\frac{\mu^4}{n}\right)^{1/3}\right)\right)$$

$$= \frac{1+o(1)}{4\sqrt{\pi}} \frac{(1+\mu)}{(\mu^2 n)^{2/3}} \exp\left(-(2c-3)\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right).$$

Thus, for any $2c - 3 > 0$, the right-hand side converges to 0 as desired. Without loss of generality, we now fix $c = 2$.

For the rest of the sum, we aim to bound $X_{n,m}$ when $(1 + \epsilon_n) t_{\mu/n} \leq m < 2 t_{\mu/n}$. Observe that $2 t_{\mu/n} \ll \sqrt{n}$ so that the result from Theorem B.1 applies and we have

$$\sum_{(1+\epsilon_n)t_{\mu/n}\leq m<2t_{\mu/n}} X_{n,m} = \frac{1+o(1)}{4^n} \frac{n^{5/6}}{\mu^{1/3}} \sum_m \frac{4^n}{m} \frac{\tan^2\left(\frac{\pi}{m}\right)}{\left(1+\tan^2\left(\frac{\pi}{m}\right)\right)^n} \exp\left(-\mu m + 3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right).$$

Now recalling the definition of λ_x from (C.1) and using that $\tan^2 \pi/m \sim \pi^2/m^2$ uniformly for all m in the limit of the sum, we see that the previous equation re-writes as

$$\sum_{(1+\epsilon_n)t_{\mu/n}\leq m<2t_{\mu/n}} X_{n,m} = \left(1+o(1)\right) \frac{n^{5/6}}{\mu^{1/3}} \sum_m \frac{\pi^2}{m^3} \exp\left(-n\lambda_{\mu/n}(m) + 3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right).$$

Using that $m \mapsto e^{-n\lambda_{\mu/n}(m)}$ and $m \mapsto 1/m^3$ are both decreasing when $m \geq t_{\mu/n}$, it follows that

$$\sum_{(1+\epsilon_n)t_{\mu/n} \leq m < 2t_{\mu/n}} X_{n,m}$$

$$\leq \left(1 + o(1)\right) \frac{n^{5/6}}{\mu^{1/3}} \left(2t_{\mu/n} - (1+\epsilon_n)t_{\mu/n}\right) \frac{\pi^2}{((1+\epsilon_n)t_{\mu/n})^3} \exp\left(-n\lambda_{\mu/n}\left((1+\epsilon_n)t_{\mu/n}\right) + 3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right).$$

To conclude this proof, first recall that $\epsilon_n \ll 1$ and $t_{\mu/n} \sim (2\pi^2 n/\mu)^{1/3}$, along with the asymptotic behaviour of $\lambda_{\mu/n}(t_{\mu/n})$ and $\lambda_{\mu/n}((1+\epsilon_n)t_{\mu/n})$ from Proposition C.1 (**2**) and (**3**), to see that

$$\sum_{(1+\epsilon_n)t_{\mu/n} \leq m < 2t_{\mu/n}} X_{n,m} \leq \left(1 + o(1)\right) \frac{n^{5/6}}{\mu^{1/3}} \frac{\pi^2}{\left(\frac{2\pi^2 n}{\mu}\right)^{2/3}} \exp\left(-(3 + o(1))\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} \epsilon_n^2\right).$$

To conclude the proof, simply note that

$$\frac{n^{5/6}}{\mu^{1/3}} \frac{\pi^2}{\left(\frac{2\pi^2 n}{\mu}\right)^{2/3}} = \left(\frac{\pi^4\mu^2 n}{16}\right)^{1/6}$$

and recall that we assume that $\epsilon_n^2 (\mu^2 n)^{1/3} \gg \log(\mu^2 n)$, which is exactly the desired assumption to prove the convergence to 0 of the above right-hand side. This proves that the second part of the sum over m also converges to 0, thus proving the statement of the lemma. $\qquad\square$

Proposition D.5. *Recall* $t_{\mu/n} = \operatorname{argmin}_{t \in (0,2)} \lambda_{\mu/n}(t)$, *and let* $X_{n,m}$ *be defined as in* (D.1), *and* $\epsilon_n \ll 1$ *be such that* $\epsilon_n t_{\mu/n} \gg 1$ *and* $\epsilon_n^2 (\mu^2 n)^{1/3} \gg \log(\mu^2 n)$. *Then, the sum over* $m \leq (1-\epsilon_n)t_{\mu/n}$ *of* $X_{n,m}$ *is negligible, i.e.,*

$$\lim_{n \to \infty} \sum_{3 \leq m \leq (1-\epsilon_n)t_{\mu/n}} X_{n,m} = 0.$$

Proof. Similar to the proof of Proposition D.4, we need to split the sum in two parts in order to prove the desired convergence. Fix for now $\delta \in (0,1)$ arbitrary. For the values of m closest to $t_{\mu/n}$, use the asymptotic behaviour of $H_{n,m}$ from Theorem B.1, along with the divergence of m and the definition of $\lambda_{\mu/n}$ from (C.1) to see that

$$\sum_{\delta t_{\mu/n} \leq m \leq (1-\epsilon_n)t_{\mu/n}} X_{n,m} = \left(1 + o(1)\right) \frac{n^{5/6}}{\mu^{1/3}} \exp\left(3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\right) \sum_m \frac{\pi^2}{m^3} e^{-n\lambda_{\mu/n}(m)}.$$

Further using the fact that $m \mapsto \lambda_{\mu/n}(m)$ and $m \mapsto 1/m^3$ are decreasing on the set of m considered here, we can bound the terms of the sum by taking the extreme values and obtain

$$\sum_{\delta t_{\mu/n} \leq m \leq (1-\epsilon_n)t_{\mu/n}} X_{n,m} \leq \left(1 + o(1)\right) \frac{n^{5/6}}{\mu^{1/3}} \frac{\pi^2}{\delta^3 t_n^2} \exp\left(3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} - n\lambda_{\mu/n}\left((1-\epsilon_n)t_{\mu/n}\right)\right).$$

To conclude this first part of the proof, simply recall the asymptotic behaviour of $t_{\mu/n}$ and of $\lambda_{\mu/n}$ from Proposition C.1 (**1**), (**2**) and (**3**) to obtain

$$\sum_{\delta t_{\mu/n} \leq m \leq (1-\epsilon_n)t_{\mu/n}} X_{n,m} \leq \left(1 + o(1)\right) \frac{n^{5/6}}{\mu^{1/3}} \frac{\pi^2}{\delta^3 \left(\frac{2\pi^2 n}{\mu}\right)^{2/3}} \exp\left(-(3 + o(1))\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} \epsilon_n^2\right)$$

$$\leq \left(1 + o(1)\right) \frac{(\pi^4\mu^2 n)^{1/6}}{\delta^3} \exp\left(-(3 + o(1))\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} \epsilon_n^2\right).$$

Thus, thanks to the assumption that $\epsilon_n^2 (\mu^2 n)^{1/3} \gg \log(\mu^2 n)$, the right-hand side converges to 0 as desired. We now prove that the second part of the sum converges to 0 as well.

Start by using the result from Theorem B.1 along with the definition of $X_{n,m}$ from (D.1) to see that, for any $3 \leq m \leq t_{\mu/n} \ll \sqrt{n}$ we have

$$X_{n,m} \leq \left(1 + o(1)\right) \frac{n^{5/6}}{\mu^{1/3}} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} \exp\left(3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3} - n\lambda_{\mu/n}(m)\right),$$

where the $o(\cdot)$ term does not depend on m. Moreover, we observe that $m^2 \tan^2(\pi/m)$ is bounded since $m \geq 3$. Finally, recalling the definition of $\lambda_{\mu/n}$ from (C.1), we see that for any $m \geq 3$, we have

$$m^2 \lambda_{\mu/n}(m) \geq m^2 \log\left(1 + \tan^2\left(\frac{\pi}{m}\right)\right),$$

and the right hand side is bounded away from 0. Thus, there exists $C > 0$ such that, for any n large enough and $3 \leq m \leq t_{\mu/n}$ we have

$$X_{n,m} \leq C \frac{(2\pi^2)^{2/3} n^{5/6}}{m^3 \mu^{1/3}} \exp\left(3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3} - \frac{1}{C} \frac{\pi^2 n}{m^2}\right),$$

It follows that, for any $x \leq y$ such that $3 \leq x t_{\mu/n} \leq y t_{\mu/n} \leq t_{\mu/n}$ we have

$$\sum_{x t_{\mu/n} \leq m < y t_{\mu/n}} X_{n,m} \leq \sum_{x t_{\mu/n} \leq m < y t_{\mu/n}} C \frac{(2\pi^2)^{2/3} n^{5/6}}{m^3 \mu^{1/3}} \exp\left(3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3} - \frac{1}{C} \frac{\pi^2 n}{m^2}\right)$$

$$\leq (y - x) t_{\mu/n} \cdot C \frac{(2\pi^2)^{2/3} n^{5/6}}{(x t_n)^3 \mu^{1/3}} \exp\left(3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3} - \frac{1}{C} \frac{\pi^2 n}{(y t_{\mu/n})^2}\right).$$

Note that $y \leq 1$ and use the asymptotic behaviour of $t_{\mu/n}$ from Proposition C.1 (**1**) to see that

$$\sum_{x t_{\mu/n} \leq m < y t_{\mu/n}} X_{n,m} \leq C \frac{(\mu^2 n)^{1/6}}{x^3} \exp\left(-\left(\frac{1}{Cy^2} - 3\right)\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}\right).$$

But now we observe that this upper bound is exactly of the same form as F_t from Corollary A.2 where we match t with $(\mu^2 n)^{1/3}$ and have $\alpha = 1/2$, $B(x) = x^3/C$, and $\Gamma(y) = (1/Cy^2 - 3)^{-1}$. Observe that Γ might diverge when Cy^2 approaches 1/3. However, by setting $a_+ < 1/\sqrt{3C}$, we see that all hypothesis from Corollary A.2 apply and it follows that we can find $3/t_{\mu/n} = a_0 \leq \cdots \leq a_k = a_+$ such that

$$\sum_{i=1}^{k} C \frac{(\mu^2 n)^{1/6}}{a_{i-1}^3} \exp\left(-\left(\frac{1}{Ca_i^2} - 3\right)\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}\right) \longrightarrow 0.$$

This then implies that

$$\sum_{3 \leq m < a_+ t_{\mu/n}} X_{n,m} \longrightarrow 0.$$

To conclude the proof, simply recall that the sum from $\delta t_{\mu/n}$ to $(1 - \epsilon_n) t_{\mu/n}$ converges to 0 as well, for any arbitrary $\delta > 0$. Thus, by setting $\delta = a_+$, the sum over m from 3 to $(1 - \epsilon_n) t_{\mu/n}$ converges to 0, as desired. □

D.3   Combining all results

In here, we combine all the previous results in order to prove Theorem D.1.

Proof of Theorem D.1. Before starting the proof, we recall that we assume that $1/\sqrt{n} \ll \mu \ll n^{1/4}$. Under these assumptions, observe that it is possible to find $\epsilon_n \ll 1$ such that $\epsilon_n t_{\mu/n} \gg 1$, $\epsilon_n (\mu^2 n)^{1/6} \gg 1$, and $\epsilon_n^2 (\mu^2 n)^{1/3} \gg \log(\mu^2 n)$. Indeed, in that case, since $t_{\mu/n}$ is of order $(n/\mu)^{1/3}$ it diverges to ∞ so it is possible to have $\epsilon_n t_{\mu/n} \gg 1$. Moreover, the second condition is actually a consequence of the third one which can also be satisfied since $\log(\mu^2 n)/(\mu^2 n)^{1/3} \ll 1$. We now choose such an ϵ_n which, by definition, satisfies the conditions from Proposition D.3, D.4 and D.5.

Recall now the definition of W_n from (B.3) and see that Proposition D.3 states that

$$\frac{1}{4^n} \sum_{|m-t_{\mu/n}|<\epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = \left(1 + o(1)\right) \frac{\pi^{5/6}}{2^{1/3} 3^{1/2}} \frac{\mu^{1/3}}{n^{5/6}} \exp\left(-3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}\right).$$

Moreover, using the definition of $X_{n,m}$ from (D.1) along with Proposition D.4 and D.5, we see that

$$\frac{1}{4^n} \sum_{|m-t_{\mu/n}|\geq \epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = o\left(\frac{\mu^{1/3}}{n^{5/6}} \exp\left(-3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}\right)\right).$$

Combining the previous two results directly implies that

$$W_n = \frac{1}{4^n} \sum_{m=3}^{n} H_{n,m-1} e^{-\mu m} = \left(1 + o(1)\right) \frac{\pi^{5/6}}{2^{1/3} 3^{1/2}} \frac{\mu^{1/3}}{n^{5/6}} \exp\left(-3\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3}\right),$$

exactly as desired. □

E   Asymptotic of the partition function: part II

The goal of this section is to extend the results from Theorem D.1 regarding the asymptotic of W_n as defined in (B.3). We actually start with a result overlapping with Theorem D.1, thus showing that the difficult part of the proof of the theorem is actually when μ is "close" to $1/\sqrt{n}$, or more precisely when μ is at most of order $(\log^{3/2} n)/\sqrt{n}$. We then provide a second result, overlapping with the first one and covering the case $\mu \gg n^{1/4}$, for which we show that the sum of W_n can be reduced to at most two terms. The two sections below cover these two cases.

E.1   Generic behaviour

In the case where μ is in the intermediate range but not too close to $1/\sqrt{n}$, the proof of Theorem D.1 can actually be simplified and does not require bounds as subtle as to prove the case close to $1/\sqrt{n}$. We provide here the complete proof of the asymptotic of W_n in that regime.

Proposition E.1. *Let $\mu = \mu_n$ be such that $(\log^{3/2} n)/\sqrt{n} \ll \mu \ll n$ and recall $t_{\mu/n} = \mathrm{argmin}_{t\in(0,2)} \lambda_{\mu/n}(t)$. Then, for any $\epsilon_n \ll 1$ such that $(\mu^2 n)^{1/3} \epsilon_n^2 \gg \log n$ and $\epsilon \gg (\mu/n)^{1/3}$ (which are both possible thanks to the assumption on μ) the sum W_n defined in (B.3) satisfies*

$$W_n = \left(1 + o(1)\right) \frac{\mu}{2n} e^{-n\lambda_{\mu/n}(t_{\mu/n})} \sum_{|m-t_{\mu/n}|<\epsilon_n t_{\mu/n}} \exp\left(-\left(\frac{3}{2} + o(1)\right)\left(\frac{\mu^4}{2\pi^2 n}\right)^{1/3} (m - t_{\mu/n})^2\right),$$

where the $o(\cdot)$ in the exponential are uniform over all terms of the sum.

Proof. To prove this result we split the sum into four parts. We start with the main contribution of the sum, which almost directly follows previous results. Indeed, for any $\epsilon_n \ll 1$, Theorem B.1 tells us that

$$\frac{1}{4^n} \sum_{|m-t_{\mu/n}|<\epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = \left(1 + o(1)\right) \frac{\pi^2}{t_{\mu/n}^3} \sum_{|m-t_{\mu/n}|\leq \epsilon_n t_{\mu/n}} e^{-n\lambda_{\mu/n}(m)},$$

where $\lambda_{\mu/n}$ is defined in (C.1). But then using the asymptotic behaviour of $t_{\mu/n}$ along with the behaviour of $\lambda_{\mu/n}$ close to $t_{\mu/n}$ with $\gamma = (m/t_{\mu/n} - 1)$ from Proposition C.1 (**1**) and (**3**), we see that

$$\frac{1}{4^n} \sum_{|m-t_{\mu/n}|<\epsilon_n t_{\mu/n}} H_{n,m-1} e^{-\mu m} = \left(1 + o(1)\right) \frac{n}{2\mu} e^{-n\lambda_{\mu/n}(t_{\mu/n})} \sum_{|m-t_{\mu/n}|<\epsilon_n t_{\mu/n}} \exp\left(-(3 + o(1))\left(\frac{\pi^2 \mu^2 n}{4}\right)^{1/3} \gamma^2\right)$$

$$= \left(1 + o(1)\right) \frac{n}{2\mu} e^{-n\lambda_{\mu/n}(t_{\mu/n})} \sum_{|m-t_{\mu/n}|<\epsilon_n t_{\mu/n}} \exp\left(-\left(\frac{3}{2} + o(1)\right)\left(\frac{\mu^4}{2\pi^2 n}\right)^{1/3} (m - t_{\mu/n})^2\right),$$

which is exactly the desired asymptotic. Before focusing on the other parts of the sum, first recall the approximation for $n\lambda_{\mu/n}(t_{\mu/n})$ from Proposition C.1:

$$n\lambda_{\mu/n}(t_{\mu/n}) = 3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} + O\left(\left(\frac{\mu^4}{n}\right)^{1/3}\right).$$

We now use this asymptotic behaviour to show that the other parts of the sum are negligible.

For the first terms of the sum, use Theorem B.1 to obtain

$$\frac{1}{4^n}\sum_{3\le m\le(1-\epsilon_n)t_{\mu/n}} H_{n,m-1}e^{-\mu m} = \left(1+o(1)\right)\sum_{3\le m\le(1-\epsilon_n)t_{\mu/n}}\frac{\tan^2\left(\frac{\pi}{m}\right)}{m}e^{-n\lambda_{\mu/n}(m)},$$

where $\lambda_{\mu/n}$ is defined in (C.1). Using that $m\mapsto\lambda_{\mu/n}(m)$ is decreasing on $(2, t_{\mu/n})$, we can bound the exponential term by its value at $m = (1-\epsilon_n)t_{\mu/n}$. Further bounding the other terms by their maximum, corresponding to $m = 3$, we obtain

$$\frac{1}{4^n}\sum_{3\le m\le(1-\epsilon_n)t_{\mu/n}} H_{n,m-1}e^{-\mu m} \le \left(1+o(1)\right)\frac{\tan^2\left(\frac{\pi}{3}\right)}{3}t_{\mu/n}e^{-n\lambda_{\mu/n}((1-\epsilon_n)t_{\mu/n})}.$$

Putting any constant term into a flexible C and dividing by the desired asymptotic leads to

$$\frac{n}{\mu}\exp\left(3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} + O\left(\left(\frac{\mu^4}{n}\right)^{1/3}\right)\right)\frac{1}{4^n}\sum_{3\le m\le(1-\epsilon_n)t_{\mu/n}} H_{n,m-1}e^{-\mu m}$$

$$\le \left(C+o(1)\right)\left(\frac{n}{\mu}\right)^{4/3}\exp\left(-\left(3+o(1)\right)\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\epsilon_n^2 + O\left(\left(\frac{\mu^4}{n}\right)^{1/3}\right)\right).$$

Recall now that we assumed that $(\mu^2 n)^{1/3}\epsilon_n^2 \gg \log n$ as well as $\epsilon \gg (\mu/n)^{1/3}$ which are exactly the two conditions needed for the first term in the exponential to be the dominant term, proving that this part of the sum is negligible with respect to the desired asymptotic behaviour.

Consider now $r_n \gg 1$ such that $r_n t_{\mu/n} \ll \sqrt{n}$ and focus on the sum from $(1+\epsilon_n)t_{\mu/n}$ to $r_n t_{\mu/n}$. Applying the same method as for the previous case, we see that

$$\frac{1}{4^n}\sum_{(1+\epsilon_n)t_{\mu/n}\le m<r_n t_{\mu/n}} H_{n,m-1}e^{-\mu m} \le \left(1+o(1)\right)\frac{\pi^2 r_n}{t_{\mu/n}^2}e^{-n\lambda_{\mu/n}((1+\epsilon_n)t_{\mu/n})},$$

leading again to

$$\frac{n}{\mu}\exp\left(3\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3} + O\left(\left(\frac{\mu^4}{n}\right)^{1/3}\right)\right)\frac{1}{4^n}\sum_{(1+\epsilon_n)t_{\mu/n}\le m<r_n t_{\mu/n}} H_{n,m-1}e^{-\mu m}$$

$$\le \left(C+o(1)\right)r_n\left(\frac{n}{\mu}\right)^{1/3}\exp\left(-\left(3+o(1)\right)\left(\frac{\pi^2\mu^2 n}{4}\right)^{1/3}\epsilon_n^2 + O\left(\left(\frac{\mu^4}{n}\right)^{1/3}\right)\right).$$

For the same reasons as the previous case, the right-hand side converges to 0 and this part of the sum is also negligible.

For the final part of the sum, consider the remaining values for m and upper bound $H_{n,m-1}$ with $C_{n-1}\sim 4^n/(4\sqrt{\pi}n^{3/2})$ to obtain

$$\frac{1}{4^n}\sum_{r_n t_{\mu/n}\le m\le n} H_{n,m-1}e^{-\mu m} \le \left(1+o(1)\right)\frac{1}{4\sqrt{\pi}n^{3/2}}\frac{e^{-\mu r_n t_{\mu/n}}}{1-e^{-\mu}}.$$

Using that $\mu t_{\mu/n}$ is of order $(\mu^2 n)^{1/3}$, we see that the same arguments as before proves that this part of the sum is also negligible. This concludes the proof of the proposition. $\qquad\square$

E.2 Finite sum behaviour

We conclude this section by covering the case where $\mu \gg n^{1/4}$, for which the sum of W_n can be reduced to only a couple of terms.

Proposition E.2. *Let $\mu = \mu_n$ be such that $\mu \gg n^{1/4}$. Recall $t_{\mu/n} = \operatorname{argmin}_{t \in (0,2)} \lambda_{\mu/n}(t)$. Then we have*

$$W_n = \left(1 + o(1)\right) \sum_{m=\max\{\lfloor t_{\mu/n} \rfloor, 3\}}^{\lceil t_{\mu/n} \rceil} \frac{\tan^2 \frac{\pi}{m}}{m} \exp\left(-\mu m - n \log\left(1 + \tan^2 \frac{\pi}{m}\right)\right).$$

Proof. We start by assuming that μ/n is bounded away from 0. In that case, we fix an arbitrarily large value M. Using that $H_{n,m-1} \leq C_{n-1}$, we see that

$$\frac{1}{4^n} \sum_{m > M} H_{n,m-1} e^{-\mu m} \leq \left(1 + o(1)\right) \frac{e^{-\mu M}}{4\sqrt{\pi}\sqrt{n}}.$$

Moreover, considering the term $m = 3$ and observing the definition from (B.1) implies that $H_{n,2} = 1$, we see that

$$\frac{1}{4^n} H_{n,2} e^{-3\mu} = \frac{e^{-3\mu}}{4^n},$$

from which we observe that the sum over $m > M$ is negligible with respect to the first term as long as $c(M-3) > \log 4$ where $c = \liminf \mu/n$. It follows that, for any such M, we have

$$W_n = \left(1 + o(1)\right) \frac{1}{4^n} \sum_{m=3}^{M} H_{n,m-1} e^{-\mu m},$$

and since this is a finite sum, we can directly apply the result from Theorem B.1 and replace $H_{n,m-1}$ by their equivalent asymptotic behaviour:

$$W_n = \left(1 + o(1)\right) \sum_{m=3}^{M} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} \frac{e^{-\mu m}}{\left(1 + \tan^2\left(\frac{\pi}{m}\right)\right)^n} = \left(1 + o(1)\right) \sum_{m=3}^{M} \frac{\tan^2\left(\frac{\pi}{m}\right)}{m} e^{-n\lambda_{\mu/n}(m)},$$

where $\lambda_{\mu/n}$ is defined in (C.1). Observing that $\lambda_{\mu/n}(m)$ is bounded away from 0 in that scenario, we see that all terms are at least of exponential decay and so only the exponent with minimal value can be kept. Moreover, using that $\lambda_{\mu/n}$ is strictly convex, we see that the difference between two distinct exponents, except possibly for $\lfloor t_{\mu/n} \rfloor$ and $\lceil t_{\mu/n} \rceil$, is bounded away from 0 as well. Thus only these two values can be kept in the sum and further, since it is possible that $t_{\mu/n} < 3$, we need to restrict it to $m \geq 3$.

Assume now that $\mu \ll n$. We recall that we also assume that $\mu \gg n^{1/4}$. Under these assumptions, we can apply Proposition E.1 and obtain that

$$W_n = \left(1 + o(1)\right) \frac{\mu}{2n} e^{-\lambda_n(t_{\mu/n})} \sum_{|m - t_{\mu/n}| < \epsilon_n t_{\mu/n}} \exp\left(-\left(\frac{3}{2} + o(1)\right)\left(\frac{\mu^4}{2\pi^2 n}\right)^{1/3} \left(m - t_{\mu/n}\right)^2\right).$$

But now, since $\mu^4/n \gg 1$, a geometric bound on the sum tells us that the first term asymptotically characterizes the whole sum, leading to

$$\sum_{m \geq \lceil t_{\mu/n} \rceil + 1} \exp\left(-\left(\frac{3}{2} + o(1)\right)\left(\frac{\mu^4}{2\pi^2 n}\right)^{1/3} \left(m - t_{\mu/n}\right)^2\right) = o\left(\exp\left(-\left(\frac{3}{2} + o(1)\right)\left(\frac{\mu^4}{2\pi^2 n}\right)^{1/3} \left(\lceil t_{\mu/n} \rceil - t_{\mu/n}\right)^2\right)\right)$$

as well as

$$\sum_{m \leq \lfloor t_{\mu/n} \rfloor - 1}^{\infty} \exp\left(-\left(\frac{3}{2} + o(1)\right)\left(\frac{\mu^4}{2\pi^2 n}\right)^{1/3} \left(m - t_{\mu/n}\right)^2\right) = o\left(\exp\left(-\left(\frac{3}{2} + o(1)\right)\left(\frac{\mu^4}{2\pi^2 n}\right)^{1/3} \left(\lfloor t_{\mu/n} \rfloor - t_{\mu/n}\right)^2\right)\right).$$

It follows that all terms except for $m \in \{\lfloor t_{\mu/n} \rfloor, \lceil t_{\mu/n} \rceil\}$ are negligible, thus proving the desired result. $\square$

References

[1] R. Abraham and J.-F. Delmas. An introduction to galton-watson trees and their local limits. *arXiv preprint arXiv:1506.05571*, 2015.

[2] R. Abraham, A. Bouaziz, and J.-F. Delmas. Very fat geometric galton-watson trees. *ESAIM: Probability and Statistics*, 24:294–314, 2020.

[3] L. Addario-Berry, L. Devroye, and S. Janson. Sub-Gaussian tail bounds for the width and height of conditioned Galton–Watson trees. *The Annals of Probability*, 41(2):1072 – 1087, 2013. doi:10.1214/12-AOP758. URL https://doi.org/10.1214/12-AOP758.

[4] D. Aldous. The continuum random tree. i. *The annals of probability*, pages 1–28, 1991.

[5] D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004. doi:10.1007/978-3-662-09444-0_1.

[6] O. Angel, S. Bhamidi, S. Donderwinkel, M. Neeladri, and S. Akshay. Evolution of recursive trees with limited memory. *arXiv preprint arXiv:2510.18856*, 2025.

[7] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13, 2001. ISSN 1083-6489. doi:10.1214/EJP.v6-96.

[8] A. N. Borodin and P. Salminen. *Handbook of Brownian motion-facts and formulae*. Birkhäuser, 2012.

[9] N. Bouchot. Scaling limits for the random walk penalized by its range in dimension one. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 21:791–813, 2024.

[10] D. Burago, Y. Burago, S. Ivanov, et al. *A course in metric geometry*, volume 33. American Mathematical Society Providence, 2001.

[11] P. Chassaing, J.-F. Marckert, and M. Yor. The height and width of simple trees. In *Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities*, pages 17–30. Springer, 2000.

[12] N. G. de Bruijn, D. E. Knuth, and S. Rice. The average height of planted plane trees. In *Graph theory and computing*, pages 15–22. Elsevier, 1972.

[13] A. Dembo. *Large deviations techniques and applications*. Springer, 2009.

[14] M. Drmota. *Random trees: an interplay between combinatorics and probability*. Springer, 2009.

[15] T. Duquesne and J.-F. Le Gall. *Random trees, Lévy processes and spatial branching processes*, volume 281. Société mathématique de France Paris, France, 2002.

[16] T. Duquesne, R. Khanfir, S. Lin, and N. Torri. Scaling limits of tree-valued branching random walks. *Electronic Journal of Probability*, 27:1–54, 2022.

[17] B. Durhuus. Probabilistic aspects of infinite trees and surfaces. *Acta Physica Polonica Series B*, 34(10):4795–4812, 2003.

[18] B. Durhuus and M. Ünel. Trees with power-like height dependent weight. *Electronic Journal of Probability*, 27:1–24, 2022.

[19] B. Durhuus and M. Ünel. Trees with exponential height dependent weight. *Probability Theory and Related Fields*, 186(3):999–1043, 2023.

[20] J.-F. L. Gall and G. Miermont. Scaling limits of random trees and planar maps. *arXiv preprint arXiv:1101.4856*, 2011.

[21] S. Janson. Simply generated trees, conditioned galton—watson trees, random allocations and condensation. *Discrete Mathematics & Theoretical Computer Science*, 9:103–252, 2012.

[22] H. Kesten. Subdiffusive behavior of random walk on a random cluster. In *Annales de l'IHP Probabilités et statistiques*, volume 22-4, pages 425–487, 1986.

[23] A. Klenke. *Probability theory: a comprehensive course*. Springer, 2008.

[24] K. Leckey, D. Mitsche, and N. Wormald. The height of depth-weighted random recursive trees. *Random Structures & Algorithms*, 56(3):851–866, 2020.

[25] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian journal of Mathematics*, 30(5): 997–1015, 1978.

[26] G. Miermont. The brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Mathematica*, 210, 2013.

[27] A. Mogul'skii. Small deviations in a space of trajectories. *Theory of Probability & Its Applications*, 19(4):726–736, 1975.

[28] B. Pittel. Note on the heights of random recursive trees and random m-ary search trees. *Random Structures & Algorithms*, 5(2):337–347, 1994.

[29] P. Thévenin and S. Wagner. Local limits of descent-biased permutations and trees. *arXiv preprint arXiv:2312.11183*, 2023.