PRECISE LOGARITHMIC ASYMPTOTICS FOR THE RIGHT TAILS OF SOME LIMIT RANDOM VARIABLES FOR RANDOM TREES

JAMES ALLEN FILL AND SVANTE JANSON

ABSTRACT. For certain random variables that arise as limits of functionals of random finite trees, we obtain precise asymptotics for the logarithm of the right-hand tail. Our results are based on the facts (i) that the random variables we study can be represented as functionals of a Brownian excursion and (ii) that a large deviation principle with good rate function is known explicitly for Brownian excursion. Examples include limit distributions of the total path length and of the Wiener index in conditioned Galton–Watson trees (also known as simply generated trees). In the case of Wiener index (where we recover results proved by Svante Janson and Philippe Chassaing by a different method) and for some other examples, a key constant is expressed as the solution to a certain optimization problem, but the constant's precise value remains unknown.

1. INTRODUCTION

Many authors have proved convergence in distribution of various functionals of various kinds of random trees. Many have also considered large-deviation estimates and tail bounds.

In this paper, in an attempt to understand several random variables that arise as such limits, we will obtain precise logarithmic asymptotics for their (right-hand) tails. For example, we will treat the limit distributions of the total path length and of the Wiener index in conditioned Galton–Watson trees (a.k.a. simply generated trees), where we recover results proved by Csörgő, Shi and Yor [8] and Janson and Chassaing [20] by a different method.

The results will be of the "quasi-Gaussian" type

$$\mathbb{P}(X > x) = \exp\left[-(1 + o(1))x^2/(2\gamma^2)\right],$$

Date: January 9, 2007; slightly revised September 5, 2007.

2000 *Mathematics Subject Classification.* 60F10; 60C05, 60J65.

Key words and phrases. large deviations, tail asymptotics, Galton–Watson trees, simply generated families of trees, Brownian excursion, variational problems, total path length, Wiener index.

Research supported by NSF grants DMS-0104167 and DMS-0406104 and by The Johns Hopkins University's Acheson J. Duncan Fund for the Advancement of Research in Statistics.

as $x \to \infty$, for some positive number γ that differs from case to case; γ is given as the solution to a variational problem. In some cases, we can solve the variational problem and find γ explicitly, while in other cases we only find bounds for γ. (Note, however, that the asymptotic distributions are not exactly Gaussian. Indeed, the examples we study will all be positive random variables.)

Our results are based on the fact that the limit random variables we study here can be represented as functionals of a (normalized) Brownian excursion; these representations have been established previously by various authors, and go back to the theory by Aldous [1, 2] of the continuum random tree.

Remark 1.1. We consider only tail asymptotics for the limiting random variables and not for the actual functionals of random trees of a finite size. That is, for certain random variables X_n associated with trees of size n with distributional limit X, we find lead-order asymptotics of $\ln \mathbb{P}(X > x)$ as x becomes large, *i.e.*, of $\lim_n \ln \mathbb{P}(X_n > x)$. More interesting would be large deviations for the sequence (X_n) itself, that is, asymptotics of $\ln \mathbb{P}(X_n > x_n)$ for sequences $x_n \to \infty$. Such results, however, fall outside the scope of this paper. Moreover, among the applications we consider, only in the case of height (our warm-up Example 4.1) are such large-deviation results known by any method; see Flajolet et al. [15].

Remark 1.2. Not all limit variables for random trees have quasi-Gaussian tails. One well-known example is the total path length in a binary search tree (under the so-called random permutation model), where Knessl and Szpankowski [24] give very sharp tail estimates for the limit distribution; roughly put, they assert that the left tail decays doubly exponentially and that the right tail decays exponentially. But their results rely on several unproven regularity assumptions (as noted in their paper), and it is still an intriguing open problem to verify the assumptions and prove these results rigorously.

Of course, the family of trees just cited is not simply generated. A counterexample functional for simply generated families is the total left path length minus the total right path length in a uniformly random binary tree, which is a measure of the asymmetry of the tree. (Note that the sum is the total path length treated in Example 4.2 below.) This difference converges after suitable scaling to the center of mass of integrated super-Brownian excursion (ISE), or equivalently to the integral of the head of a Brownian snake, see [20] or [19]. For this limit variable S we have $-\ln\bigl(\mathbb{P}(S > x)\bigr) \sim \frac{3}{4}10^{1/3}\, x^{4/3}$, as shown in [20]. See also Marckert [25] and Bousquet-Mélou and Janson [4].

The main theorem (together with a technical extension) giving large deviations for functionals of Brownian excursion is stated in Section 2. Each application of the main theorem results in a variational problem; techniques for solving such problems are discussed in Section 3. We present applications in Section 4. Finally, the main theorem is proved in Section 5.

2. A general theorem

2.1. **Some notation.** We introduce the following notation. $C[0,1]$ is the usual space of continuous functions on $[0,1]$ equipped with the supremum metric $\|f-g\|_\infty := \sup_t |f(t)-g(t)|$. We let

$$\begin{aligned} C_{\mathrm{bm}}[0,1] &:= \{f \in C[0,1] : f(0) = 0\},\\ C_{\mathrm{br}}[0,1] &:= \{f \in C[0,1] : f(0) = f(1) = 0\},\\ C_{\mathrm{ex}}[0,1] &:= \{f \in C[0,1] : f(0) = f(1) = 0,\ f \geq 0\}; \end{aligned}$$

these are regarded as subsets of $C[0,1]$ equipped with the same metric. Note that these spaces are closed subspaces of $C[0,1]$ and thus complete, separable metric spaces.

We further let B be a standard Brownian motion on $[0,1]$, B_{br} a Brownian bridge, and B_{ex} a normalized Brownian excursion (with length 1); these are random elements of $C_{\mathrm{bm}}[0,1]$, $C_{\mathrm{br}}[0,1]$, and $C_{\mathrm{ex}}[0,1]$, respectively (which explains our notation).

Further, let H be the Sobolev space of all absolutely continuous functions $f \in C[0,1]$ such that $\|f'\|_2^2 = \int_0^1 |f'(t)|^2 \,\mathrm{d}t < \infty$. (The derivative f' exists a.e., and all statements below about f' for $f \in H$ should be interpreted a.e.) We define

$$H_{\mathrm{bm}} := H \cap C_{\mathrm{bm}}[0,1], \quad H_{\mathrm{br}} := H \cap C_{\mathrm{br}}[0,1], \quad H_{\mathrm{ex}} := H \cap C_{\mathrm{ex}}[0,1].$$

(The space H_{bm} is known as the Cameron–Martin space for Brownian motion, see [5] and [16, Example 8.19].) Similarly, let K be the closed unit ball in H, *i.e.*, the set of $f \in H$ such that $\|f'\|_2^2 = \int_0^1 |f'(t)|^2 \,\mathrm{d}t \leq 1$, and let

$$K_{\mathrm{ex}} := K \cap C_{\mathrm{ex}}[0,1]. \tag{1}$$

2.2. **The main result.** We can now state a general theorem for functionals of Brownian excursions. We give asymptotic results for the distribution itself as well as for the moment generating function (*i.e.*, for the Laplace transform) and for the moments. These three results are equivalent (see the proof) but often useful in different situations.

Theorem 2.1. *Let $X = \Phi(B_{\mathrm{ex}})$, where Φ is a continuous nonnegative functional on $C_{\mathrm{ex}}[0,1]$ that is positively homogeneous [i.e., $\Phi(tf) = t\Phi(f)$ when $f \in C_{\mathrm{ex}}[0,1]$ and $t \geq 0$] and not identically 0. Let*

$$\gamma := \max\{\Phi(f) : f \in K_{\mathrm{ex}}\}. \tag{2}$$

Then $0 < \gamma < \infty$ and

$$-\ln \mathbb{P}(X > x) \sim \frac{x^2}{2\gamma^2} \qquad \text{as } x \to \infty, \tag{3}$$

$$\ln \mathbb{E}\, e^{tX} \sim \frac{\gamma^2}{2} t^2 \qquad \text{as } t \to \infty, \tag{4}$$

$$\big(\mathbb{E}\, X^r\big)^{1/r} \sim \frac{\gamma}{\sqrt{e}} r^{1/2} \qquad \text{as } r \to \infty. \tag{5}$$

Remark 2.2. K_{ex} is a compact subset of $C_{\text{ex}}[0,1]$, see Lemma 5.1 below. Hence, the maximum in (2) exists and is finite.

Remark 2.3. It follows from the proof that if the maximum in (2) is attained at a unique f_0, then this f_0 is the typical shape of the B_{ex} giving exceptionally large $X = \Phi(B_{\text{ex}})$, in the sense that if B_{ex}^x has the conditional distribution of B_{ex} given $\Phi(B_{\text{ex}}) > x$, then $x^{-1}B_{\text{ex}}^x$ converges in probability, as $x \to \infty$, to $\Phi(f_0)^{-1}f_0$. In other words, B_{ex}^x is, for large x, with large probability close to a suitable multiple of f_0. (Of course B_{ex}^x is much less smooth than f_0, so it looks different on small scales; a Brownian excursion a.s. does not belong to H.)

Theorem 2.1 will be proved in Section 5. We give several applications in Section 4.

2.3. **An extension.** In one of the applications in Section 4, the functional Φ is not continuous on $C_{\text{ex}}[0,1]$ and we need an extension (Theorem 2.4 below, also proved in Section 5) to Hölder spaces. We define, for $0 < \beta \le 1$, the Hölder space $C^\beta[0,1]$ as the space of all functions $f \in C[0,1]$ such that $|f(x)-f(y)| \le C|x-y|^\beta$ for some C and all $x, y \in [0,1]$; $C^\beta[0,1]$ is equipped with the metric given by the norm $\sup_x |f(x)| + \sup_{x\neq y} |f(x) - f(y)|/|x - y|^\beta$. Recall that $B, B_{\text{br}}, B_{\text{ex}} \in C^\beta[0,1]$ a.s. for all $\beta < 1/2$. We define $C_{\text{ex}}^\beta[0,1] := C^\beta[0,1] \cap C_{\text{ex}}[0,1]$, regarded as a subset of $C^\beta[0,1]$. Note that, for all $\beta \le 1/2$, it follows from the Cauchy–Schwarz inequality, as in (11) below, that $H \subset C^\beta[0,1]$, and thus $K_{\text{ex}} \subset C_{\text{ex}}^\beta[0,1]$.

Theorem 2.4. *If $0 < \beta < 1/2$, then Theorem 2.1 remains valid if $C_{\text{ex}}[0,1]$ is replaced by $C_{\text{ex}}^\beta[0,1]$.*

3. Finding γ

To find γ explicitly for the examples in which we have interest, we begin with some simplifications. We assume that Φ is a continuous functional on K_{ex} as in Theorem 2.1. We begin by listing some properties Φ may have.

- (A1) Φ is *symmetric*: if $\check{f}(x) := f(1-x)$ then $\Phi(\check{f}) = \Phi(f)$.
- (A2) Φ is *concave*: $\Phi\big(\frac{1}{2}(f+g)\big) \ge \frac{1}{2}\Phi(f) + \frac{1}{2}\Phi(g)$. (For positively homogeneous Φ this is equivalent to superadditivity.)
- (A3) Φ is *monotone*: $f_1 \le f_2$ implies $\Phi(f_1) \le \Phi(f_2)$.

Our first of two lemmas shows that if Φ has certain of these properties, then the search space for f maximizing $\Phi(f)$ may be suitably narrowed from K_{ex}.

Lemma 3.1. *Let Φ be a continuous functional defined on K_{ex}.*

- (i) *If Φ is symmetric and concave, then $\max_{K_{\text{ex}}} \Phi(f)$ is attained by an f which is symmetric ($\check{f} = f$).*

(ii) *If Φ is monotone, then* $\max_{K_{\mathrm{ex}}} \Phi(f)$ *is attained by a unimodal f, i.e., an f such that $f' \geq 0$ on $(0,a)$ and $f' \leq 0$ on $(a,1)$ for some $a \in (0,1)$.*

(iii) *If Φ is symmetric, concave, and monotone, then* $\max_{K_{\mathrm{ex}}} \Phi(f)$ *is attained by a symmetric unimodal f, i.e., an f such that $f' \geq 0$ on $(0,1/2)$ and $f'(x) = -f'(1-x)$ on $(1/2,1)$.*

Proof. (i): Let $f \in K_{\mathrm{ex}}$ maximize $\Phi(f)$, and let $g = \frac{1}{2}(f + \check{f})$. Then $g \in K_{\mathrm{ex}}$ is symmetric and, by the assumptions, $\Phi(g) \geq \frac{1}{2}\Phi(f) + \frac{1}{2}\Phi(\check{f}) = \Phi(f)$. Hence g, too, maximizes Φ.

(ii): Let $f \in K_{\mathrm{ex}}$ maximize $\Phi(f)$, and define $g \in H$ by $g(0) = 0$ and

$$g'(x) = \begin{cases} |f'(x)| & \text{if } x < a \\ -|f'(x)| & \text{if } x > a \end{cases}$$

where a is such that $\int_0^a |f'(x)| \,\mathrm{d}x = \frac{1}{2}\int_0^1 |f'(x)| \,\mathrm{d}x$. Then $g(1) = g(0) = 0$, $g \geq 0$, and $\|g'\|_2 = \|f'\|_2$, so $g \in K_{\mathrm{ex}}$. For $x \leq a$ we have

$$f(x) \leq \int_0^x |f'(y)| \,\mathrm{d}y = \int_0^x g'(y) \,\mathrm{d}y = g(x)$$

and for $x \geq a$

$$f(x) \leq \int_x^1 |f'(y)| \,\mathrm{d}y = \int_x^1 -g'(y) \,\mathrm{d}y = g(x),$$

so $f \leq g$ and thus $\Phi(g) \geq \Phi(f)$. Hence g too maximizes Φ.

(iii): Argue first as for (i) and then as for (ii). ☐

Our second lemma concerns maximization of $\Phi(f)$ over a certain smaller class of functions f than K_{ex}.

Lemma 3.2. (i) *Let K_{su} be the subset of K_{ex} consisting of symmetric unimodal functions. Suppose that Φ is a continuous functional on K_{su} such that for some nonnegative function $h \in L^2[0,1/2]$*

$$\Phi(f) = \int_0^{1/2} f'(t)h(t) \,\mathrm{d}t, \qquad f \in K_{\mathrm{su}}. \tag{6}$$

Then

$$\max_{K_{\mathrm{su}}} \Phi = \frac{1}{\sqrt{2}}\|h\|_{L^2[0,1/2]} = \left(\frac{1}{2}\int_0^{1/2} h(t)^2 \,\mathrm{d}t\right)^{1/2}. \tag{7}$$

A maximizing $f \in K_{\mathrm{su}}$ is given by $f' = (2\max_{K_{\mathrm{su}}} \Phi)^{-1} h$ on $[0,1/2]$.

(ii) *Suppose that Φ is a continuous symmetric, concave, monotone functional on K_{ex} such that (6) holds for some nonnegative function $h \in L^2[0,1/2]$ and all $f \in K_{\mathrm{su}}$. Then $\max_{K_{\mathrm{ex}}} \Phi = \max_{K_{\mathrm{su}}} \Phi$ is given by (7).*

Proof. (i): This is immediate by Hilbert space theory, since

$$\left\{f'|_{[0,1/2]} : f \in K_{\mathrm{su}}\right\} = \left\{g \in L^2[0,1/2] : g \geq 0 \text{ and } \|g\|_2 = 1/\sqrt{2}\right\}.$$

(ii): This follows from (i) and Lemma 3.1(iii). ☐

4. Applications

We give several applications of the general theorem to functionals of interest for random trees. In all cases, the random trees that we consider are conditioned Galton–Watson trees, also known as simply generated trees. As is well-known, this includes several important types of random trees, for example random planar trees, random labelled trees, and random binary trees (in each case uniformly distributed over all trees of the given type with a given number n of vertices). As is shown in the references given below, the functionals we study have limit distributions as the size n of the random trees tends to infinity, after proper normalization. Moreover, these limit distributions do not depend on the particular class of random trees (within the class of conditioned Galton–Watson trees) except for a simple scale factor. In the results below, we therefore will not usually discuss the random trees.

Moreover, since the asymptotic results always are given by (3), (4), (5), we will only give the value of γ.

Example 4.1 (Height and width)**.** For both the height and the width of a conditioned Galton–Watson tree, the limit distribution (after suitable rescaling) is given by the same random variable, viz. $\max B_{\mathrm{ex}}$, see Aldous [1] and Chassaing, Marckert and Yor [6]; see also [18, Section 7]. The distribution of this random variable is well-known [7, 23], see [3] for much more information; in particular,

$$\mathbb{P}(\max_t B_{\mathrm{ex}}(t) \le x) = 1 + 2\sum_{k=1}^{\infty}(1 - 4k^2x^2)\exp(-2k^2x^2), \qquad x > 0. \tag{8}$$

Hence the asymptotics we obtain from Theorem 2.1 do not yield anything new, but they serve as a simple warm-up exemplifying our results.

Thus, let $\Phi(f) := \max f$. This functional is symmetric and monotone, but not concave. By Lemma 3.1(ii), the maximum is attained for a unimodal f, but we cannot use Lemma 3.2. We can in this case easily argue directly. Let $f \in K_{\mathrm{ex}}$. By the Cauchy–Schwarz inequality, for every $x \in [0, 1]$,

$$f(x) = \tfrac12 \int_0^x f'(t)\,\mathrm{d}t - \tfrac12 \int_x^1 f'(t)\,\mathrm{d}t = \int_0^1 f'(t)\tfrac12 \operatorname{sign}(x-t)\,\mathrm{d}t \le \tfrac12 \|f'\|_2 \le \tfrac12,$$

with equality if $x = 1/2$ and $f'(t) = \operatorname{sign}(\frac12 - t)$, *i.e.*, if $f(x) = x$ for $0 \le x \le 1/2$ and $f(x) = 1 - x$ for $1/2 \le x \le 1$. Thus $\gamma = \max_{K_{\mathrm{ex}}} \Phi = 1/2$, so $\ln \mathbb{P}(\max_t B_{\mathrm{ex}}(t) > x) \sim -2x^2$, in accordance with (8).

Remark. (a) The maximizing f in Example 4.1 is easily seen to be unique. We guess that the same is true in all examples below, but we have not checked this.

(b) Example 4.1 shows that the maximum γ may be attained on K_{su} even if Φ is not concave.

(c) It follows from Theorem 1.2 in Flajolet et al. [15] that the height H_n of a conditioned critical Galton–Watson tree with offspring distribution having variance σ^2, when normalized to $X_n := \sigma H_n/(2\sqrt{n})$, satisfies, for any $c > 0$, the "zone of convergence" result

$$\mathbb{P}(X_n > x) \sim \mathbb{P}(\max_t B_{\mathrm{ex}}(t) > x) \quad \text{uniformly for } x < c\sqrt{\log n};$$

and hence that

$$-\ln \mathbb{P}(X_n > x_n) \sim 2x_n^2 \quad \text{as } n \to \infty$$

provided $x_n \to \infty$ and $x_n = O(\sqrt{\log n})$. Presumably, similar such results hold for other functionals treated below, but our techniques cannot yield these more delicate results.

Example 4.2 (Total path length)**.** It is also well known that the asymptotic distribution of the total path length in a conditioned Galton–Watson tree is given by the Brownian excursion area $\int_0^1 B_{\mathrm{ex}}(t)\,\mathrm{d}t$ [1, 2]. Thus, we now let $\Phi(f) = \int_0^1 f$. Here Φ is symmetric, concave (in fact, linear), and monotone. If $f \in K_{\mathrm{su}}$, then by integration by parts,

$$\Phi(f) = 2\int_0^{1/2} f(t)\,\mathrm{d}t = 2\int_0^{1/2} (\tfrac12 - t) f'(t)\,\mathrm{d}t.$$

Hence Lemma 3.2 applies with $h(t) = 1 - 2t$, which gives

$$\gamma = \max_{K_{\mathrm{su}}} \Phi = \left(\frac12 \int_0^{1/2} (1-2t)^2\,\mathrm{d}t\right)^{1/2} = \frac{1}{\sqrt{12}}.$$

This agrees with the tail asymptotics given by Csörgő, Shi and Yor, [8, Proof of Theorem 3.1]. A maximizing f is given by $f' = h/(2\gamma) = \sqrt{3}(1-2t)$ on $[0, 1/2]$, and thus $f(t) = \sqrt{3}t(1-t)$, $t \in [0,1]$.

The variable $\xi = 2\int_0^1 B_{\mathrm{ex}}$ is studied in [20]. Theorem 2.1 applies with $\Phi(f) = 2\int_0^1 f$. This is simply twice the functional just studied, and hence $\gamma = 2/\sqrt{12} = 1/\sqrt{3}$, in agreement with the result in [20, Remark 4.9] (obtained there by a different method). A maximizing f is again $f(t) = \sqrt{3}t(1-t)$.

In the following examples we use the notation

$$m(f; s, t) := \inf\{f(u) : u \in [s, t]\}$$

for a function f on $[0, 1]$ and $0 \le s \le t \le 1$.

Example 4.3. Another random variable studied in [20] is $\eta := 4\iint_{s<t} m(B_{\mathrm{ex}}; s, t)\,\mathrm{d}s\,\mathrm{d}t$; this arises as the limit for the sum, over all pairs of vertices in the random tree, of the depth of the last common ancestor. Theorem 2.1 applies with $\Phi(f) = 4\iint_{s<t} m(f; s, t)\,\mathrm{d}s\,\mathrm{d}t$. This Φ is symmetric, concave, and monotone. For $f \in K_{\mathrm{su}}$, $m(f; s, t) = f(s)$ when $|\frac12 - s| > |\frac12 - t|$, and $m(f; s, t) = f(t)$

otherwise. Hence, using the symmetry of f,

$$\begin{aligned}\Phi(f) &= 8 \iint\limits_{\substack{s<t \\ |\frac12-s|>|\frac12-t|}} f(s)\,\mathrm{d}s\,\mathrm{d}t = 8\int_0^{1/2}\int_s^{1-s} f(s)\,\mathrm{d}t\,\mathrm{d}s \\ &= 8\int_0^{1/2}(1-2s)f(s)\,\mathrm{d}s = 2\int_0^{1/2}(1-2s)^2 f'(s)\,\mathrm{d}s.\end{aligned}$$

Thus Lemma 3.2 applies with $h(t) = 2(1-2t)^2$ and

$$\gamma = \max_{K_{\mathrm{su}}}\Phi = \left(\frac12\int_0^{1/2} 4(1-2t)^4\,\mathrm{d}t\right)^{1/2} = \frac{1}{\sqrt5}.$$

This gives a new proof of the result in [20, Theorem 4.6]. A maximizing function is given by $f'(t) = h/2\gamma = \sqrt5(1-2t)^2$, $t \le 1/2$, and thus $f(t) = \frac{\sqrt5}{6}\bigl(1-|1-2t|^3\bigr)$, $t\in[0,1]$.

Example 4.4 (Wiener index)**.** It is shown in [17] that for the Wiener index of the random tree, the limit random variable is $\zeta := \xi - \eta$, with ξ and η given in the preceding examples. Thus Theorem 2.1 applies to ζ with $\Phi(f) = 2\iint_{s<t}\bigl(f(s)+f(t)-2m(f;s,t)\bigr)\,\mathrm{d}s\,\mathrm{d}t$. This Φ is symmetric, but neither concave (on the contrary, it is convex) nor monotone. For $f\in K_{\mathrm{su}}$, Examples 4.2 and 4.3 show that (6) holds with $h(t) = 2(1-2t)-2(1-2t)^2 = 4t(1-2t)$. Hence Lemma 3.2 shows that

$$\max_{K_{\mathrm{su}}}\Phi = \left(\frac12\int_0^{1/2} 16t^2(1-2t)^2\,\mathrm{d}t\right)^{1/2} = \frac{1}{\sqrt{30}}.$$

However, we do not know whether this also is the maximum γ over K_{ex}, so we can only conclude $\gamma \ge 1/\sqrt{30}$.

An upper bound can be found as follows. If $f\in K_{\mathrm{ex}}$, let $0<s<t<1$ and let v be a minimum point for f in $[s,t]$, *i.e.*, a point $v\in[s,t]$ such that $f(v) = \min\{f(u) : u\in[s,t]\} = m(f;s,t)$. Then

$$f(s)+f(t)-2m(f;s,t) = -\int_s^v f'(u)\,\mathrm{d}u + \int_v^t f'(u)\,\mathrm{d}u \le \int_s^t |f'(u)|\,\mathrm{d}u.$$

Thus, by the Cauchy–Schwarz inequality and the assumption $f \in K$,

$$\begin{aligned}
\Phi(f) &= 2\iint_{s<t} [f(s) + f(t) - 2m(f;s,t)]\,\mathrm{d}s\,\mathrm{d}t \\
&\le 2\iint_{s<t} \left[\int_s^t |f'(u)|\,\mathrm{d}u\right]\mathrm{d}s\,\mathrm{d}t \\
&= 2\int_0^1 |f'(u)| \left[\iint_{0<s<u<t<1} \mathrm{d}s\,\mathrm{d}t\right]\mathrm{d}u \\
&= 2\int_0^1 |f'(u)|\,u(1-u)\,\mathrm{d}u \\
&\le 2\|f'\|_2 \left[\int_0^1 u^2(1-u)^2\,\mathrm{d}u\right]^{1/2} \\
&= 2\|f'\|_2/\sqrt{30} \le 2/\sqrt{30}.
\end{aligned}$$

Consequently, $\gamma \le 2/\sqrt{30}$, and combining this with the lower bound above we find $1/\sqrt{30} \le \gamma \le 2/\sqrt{30}$.

Problem. Find γ for the random variable ζ.

Fill and Kapur [12] [13] have studied the sum, over all vertices v in the random tree, of the αth power of the size of the subtree rooted at v; here $0 < \alpha < \infty$ is a parameter. For $\alpha > 1/2$, which is the only range we shall consider here, they show that, after suitable scaling, there is a limit distribution characterized by its moments. Let Y_α have this distribution. Fill and Janson [11] show that Y_α can be represented as $\Phi(B_{\mathrm{ex}})$ with

$$\begin{aligned}
\Phi(f) = \alpha \int_0^1 &\left[t^{\alpha-1} + (1-t)^{\alpha-1}\right] f(t)\,\mathrm{d}t \\
&- \alpha(\alpha-1) \iint\limits_{0<s<t<1} (t-s)^{\alpha-2} \left[f(s) + f(t) - 2m(f;s,t)\right]\mathrm{d}s\,\mathrm{d}t. \qquad (9)
\end{aligned}$$

Note that for $\alpha = 1$ this reduces to $2\int_0^1 f$, and thus $W_1 = \xi$ in Example 4.2. Moreover, if $\alpha > 1$, then (9) simplifies to

$$\Phi(f) = 2\alpha(\alpha-1) \iint\limits_{0<s<t<1} (t-s)^{\alpha-2} m(f;s,t)\,\mathrm{d}s\,\mathrm{d}t. \qquad (10)$$

In particular, $W_2 = \eta$ in Example 4.3.

Example 4.5. If $\alpha > 1$, Theorem 2.1 applies with Φ given by (10). This Φ is symmetric, concave, and monotone. For $f \in K_{\mathrm{su}}$, as in Example 4.3,

$$\begin{aligned}
\Phi(f) &= 4\alpha(\alpha-1) \iint\limits_{\substack{s<t \\ |1/2-s|>|1/2-t|}} (t-s)^{\alpha-2} f(s)\,\mathrm{d}s\,\mathrm{d}t \\
&= 4\alpha(\alpha-1) \int_0^{1/2} \int_s^{1-s} (t-s)^{\alpha-2} f(s)\,\mathrm{d}t\,\mathrm{d}s \\
&= 4\alpha \int_0^{1/2} (1-2s)^{\alpha-1} f(s)\,\mathrm{d}s = 2\int_0^{1/2} (1-2s)^{\alpha} f'(s)\,\mathrm{d}s.
\end{aligned}$$

Thus Lemma 3.2 applies with $h(t) = 2(1-2t)^{\alpha}$ and

$$\gamma := \max_{K_{\mathrm{ex}}} \Phi = \left(\frac{1}{2} \int_0^{1/2} 4(1-2t)^{2\alpha}\,\mathrm{d}t \right)^{1/2} = \frac{1}{\sqrt{2\alpha+1}}.$$

This has been found (in the form (5)) by Fill and Kapur [14]. A maximizing function is given by $f'(t) = h(t)/(2\gamma) = \sqrt{2\alpha+1}(1-2t)^{\alpha}$, $t \le 1/2$, and thus $f(t) = \frac{\sqrt{2\alpha+1}}{2(\alpha+1)}\bigl(1 - |1-2t|^{\alpha+1}\bigr)$, $t \in [0,1]$.

Example 4.6. Now let $1/2 < \alpha < 1$. In this case, the formula (10) cannot be used (the integral diverges unless f is constant; moreover, the factor in front is negative), so we have to use (9). When $\alpha < 1$, this functional Φ is *not* continuous on $C_{\mathrm{ex}}[0,1]$. It is, however, continuous on the Hölder space $C^{\beta}_{\mathrm{ex}}[0,1]$ when $\alpha+\beta > 1$, as is easily verified. We thus choose $\beta \in (1-\alpha, 1/2)$ and use Theorem 2.4.

Nevertheless, there are further problems. When $\alpha < 1$, the functional Φ is neither monotone nor concave (it is instead convex), so we cannot apply Lemma 3.2.

For $f \in K_{\mathrm{su}}$ we find, in similar fashion as for Example 4.5, omitting the details,

$$\Phi(f) = 2\int_0^{1/2} (1-2s)^{\alpha} f'(s)\,\mathrm{d}s,$$

and thus, also for $\alpha < 1$,

$$\max_{K_{\mathrm{su}}} \Phi = \frac{1}{\sqrt{2\alpha+1}}.$$

However, for $\alpha < 1$, we do not know whether this also is the maximum over K_{ex}, so we can only conclude $\gamma \ge 1/\sqrt{2\alpha+1}$.

To get an upper bound, assume $f \in K_{\mathrm{ex}}$ and denote the two integrals in (9) by $\Phi_1(f)$ and $\Phi_2(f)$. An integration by parts yields

$$\Phi_1(f) = \int_0^1 f'(u)\,[(1-u)^{\alpha} - u^{\alpha}]\; du \le \int_0^1 |f'(u)|\,|(1-u)^{\alpha} - u^{\alpha}|\; du,$$

while an argument as in Example 4.4 yields

$$\Phi_2(f) \le \int_0^1 |f'(u)| \left[u^\alpha + (1-u)^\alpha - 1\right] du.$$

Hence, if we define

$$h(u) := |(1-u)^\alpha - u^\alpha| + u^\alpha + (1-u)^\alpha - 1 = \begin{cases} 2(1-u)^\alpha - 1, & \text{if } 0 \le u \le 1/2, \\ 2u^\alpha - 1, & \text{if } 1/2 \le u \le 1, \end{cases}$$

we have, for $f \in K_{\mathrm{ex}}$,

$$\Phi(f) = \Phi_1(f) + \Phi_2(f) \le \int_0^1 |f'(u)| h(u)\, du \le \|f'\|_2 \left[\int_0^1 h(u)^2\, du\right]^{1/2},$$

and thus

$$\gamma \le \left(\int_0^1 h(u)^2\, du\right)^{1/2} = \left(\frac{8}{2\alpha+1}(1 - 2^{-2\alpha-1}) - \frac{8}{\alpha+1}(1 - 2^{-\alpha-1}) + 1\right)^{1/2}.$$

Denoting the right hand side by $\psi(\alpha)^{1/2}$, we have verified (first graphically using `Maple`, and then rigorously using calculus) that $(2\alpha+1)\psi(\alpha)$ is decreasing on $[1/2, 1]$, and thus the maximum is attained for $\alpha = 1/2$, which gives the value $8(\sqrt{2}-1)/3$. Hence, for $1/2 < \alpha < 1$,

$$\gamma \le \psi(\alpha)^{1/2} \le \left(\frac{8(\sqrt{2}-1)}{3(2\alpha+1)}\right)^{1/2} \le \frac{1.051}{\sqrt{2\alpha+1}}.$$

Hence our upper and lower bound differ by a factor less than 1.051 (and the ratio tends to 1 as $\alpha \to 1$).

Problem. Find γ for W_α when $\alpha < 1$.

5. Proof of Theorems 2.1 and 2.4

Proof of Theorem 2.1. We begin with a simple lemma, see *e.g.* [21, Lemma 27.7].

Lemma 5.1. *The set K_{ex} defined at* (1) *is a compact subset of $C_{\mathrm{ex}}[0,1]$.*

Proof. If $f \in K$ and $0 \le x \le y \le 1$, then the Cauchy–Schwarz inequality yields

$$|f(x) - f(y)|^2 = \left|\int_x^y f'(t)\,\mathrm{d}t\right|^2 \le \int_x^y \mathrm{d}t \int_x^y |f'(t)|^2\,\mathrm{d}t \le y - x. \tag{11}$$

Since further $f \in K_{\mathrm{ex}}$ implies $f(0)=0$, it follows from the Arzelà–Ascoli theorem that K_{ex} is relatively compact in $C[0,1]$, and thus in $C_{\mathrm{ex}}[0,1]$.

It remains to show that K_{ex} is a closed subset of $C_{\mathrm{ex}}[0,1]$. Thus, assume that $f_n \in K_{\mathrm{ex}}$ and that $f_n \to f$ in $C[0,1]$. The functions f'_n belong to the unit ball of $L^2[0,1]$, so by weak compactness there exists a subsequence f'_{n_k} that converges weakly in L^2, say to g. Define $F(x) := \int_0^x g(t)\,\mathrm{d}t$. Then

$F' = g$ a.e., and $F \in K_{\mathrm{ex}}$. Moreover, the weak convergence $f'_{n_k} \to g$ along the subsequence implies, for every $x \in [0,1]$,

$$f_{n_k}(x) = \int_0^x f'_{n_k}(t)\,\mathrm{d}t = \langle f_{n_k}, \mathbf{1}_{[0,x]}\rangle \to \langle g, \mathbf{1}_{[0,x]}\rangle = \int_0^x g(t)\,\mathrm{d}t = F(x).$$

Hence $f = F \in K_{\mathrm{ex}}$. $\square$

As noted at Remark 2.2, Lemma 5.1 shows that the maximum γ in (2) exists and is finite. Moreover, $\gamma > 0$, because otherwise $\Phi(f) = 0$ for every $f \in K_{\mathrm{ex}}$. By homogeneity, this would imply $\Phi(f) = 0$ for every $f \in H_{\mathrm{ex}}$. However, H_{ex} is dense in $C_{\mathrm{ex}}[0,1]$, as can be seen by approximating a continuous function by piecewise linear functions, and since Φ is assumed to be continuous, this would imply that Φ vanishes identically on $C_{\mathrm{ex}}[0,1]$, contrary to our assumption.

To prove Theorem 2.1, we use some notations and results from large deviation theory, see for example Kallenberg [21, Chapter 27] or Dembo and Zeitouni [10].

Definition ([21, pp. 545–546])**.** A family $(X_\varepsilon)_{\varepsilon>0}$ of random elements in some metric space S satisfies the *Large Deviation Principle (LDP) with good rate function I* if $I : S \to [0,\infty]$ is a function such that the level sets $\{x \in S : I(x) \le r\}$ are compact for all finite r and, for every Borel set $A \subseteq S$,

$$-\inf_{x\in A^\circ} I(x) \le \liminf_{\varepsilon\to 0}\bigl(\varepsilon \ln \mathbb{P}(X_\varepsilon \in A)\bigr) \le \limsup_{\varepsilon\to 0}\bigl(\varepsilon \ln \mathbb{P}(X_\varepsilon \in A)\bigr) \le -\inf_{x\in \overline{A}} I(x).$$

We begin with two central facts.

Fact 1 ([21, Theorem 27.6])**.** *If B is a Brownian motion, then $(\varepsilon^{1/2}B)$ satisfies the LDP in $C_{\mathrm{bm}}[0,1]$ with good rate function $I(f) = \frac{1}{2}\|f'\|_2^2$ for $f \in H_{\mathrm{bm}}$ and $I(f) = \infty$ otherwise.*

Fact 2 ([21, Theorem 27.11])**.** *If $F : S \to T$ is a continuous mapping of one metric space into another, and X_ε satisfies the LDP in S with good rate function I, then $F(X_\varepsilon)$ satisfies the LDP in T with good rate function $J(y) := \inf\{I(x) : F(x) = y\}$.*

For the first application of Fact 2, note that a Brownian bridge may be constructed by $B_{\mathrm{br}}(t) := B(t) - tB(1)$. Hence, let $F(f)(t) := f(t) - tf(1)$. This is a continuous map $C_{\mathrm{bm}}[0,1] \to C_{\mathrm{br}}[0,1]$ and $F(B) = B_{\mathrm{br}}$. It is easily seen that $J(f) = \inf_{a\in\mathbb{R}} \frac{1}{2}\|f' + a\|_2^2 = I(f)$ for $f \in H_{\mathrm{br}}$ and $J(f) = \infty$ otherwise. Hence Facts 1 and 2 yield the LDP for the Brownian bridge:

Fact 3 ([21, Exercise 27.10])**.** *If B_{br} is a Brownian bridge, then $(\varepsilon^{1/2}B_{\mathrm{br}})$ satisfies the LDP in $C_{\mathrm{br}}[0,1]$ with good rate function $I(f) = \frac{1}{2}\|f'\|_2^2$ for $f \in H_{\mathrm{br}}$ and $I(f) = \infty$ otherwise.*

Moreover [21], Facts 1 and 3 extend readily to d-dimensional Brownian motion and bridge, respectively, if we replace the spaces $C_{\mathrm{bm}}[0,1]$, H_{bm}, $C_{\mathrm{br}}[0,1]$, and H_{br} by the corresponding spaces $C_{\mathrm{bm}}([0,1],\mathbb{R}^d)$, and so on, of functions with values in $\mathbb{R}^d$, interpreting $\|f'\|_2^2 = \int_0^1 |f'(t)|^2 \, dt$ with $|f'(t)|$ the usual Euclidean length of the vector $f'(t)$ in $\mathbb{R}^d$.

Turning to the Brownian excursion, we use the result that B_{ex} has the same distribution as the process $|B_{\text{br}}^{(3)}|$, where $B_{\text{br}}^{(3)}$ is 3-dimensional Brownian bridge, see, *e.g.*, Revuz and Yor [26, Theorem XII.(4.2)]. We can thus apply Fact 2 with $S = C_{\text{br}}([0,1], \mathbb{R}^3)$, $T = C_{\text{ex}}[0,1]$, $F(g) = |g|$ and $X_\varepsilon = \varepsilon^{1/2} B_{\text{br}}^{(3)}$. Recalling Fact 3 and noting that $|F(g)'| \le |g'|$, it is easily seen that $J(f) := \inf\{I(g) : g \in C_{\text{br}}([0,1], \mathbb{R}^3) \text{ and } |g| = f\}$ equals $\frac{1}{2}\|f'\|_2^2$ for $f \in H_{\text{ex}}$ and equals ∞ otherwise, and we obtain the following result.

Fact 4 (Serlet [27]). *If B_{ex} is a normalized Brownian excursion, then $(\varepsilon^{1/2} B_{\text{ex}})$ satisfies the LDP in $C_{\text{ex}}[0,1]$ with good rate function $I(f) = \frac{1}{2}\|f'\|_2^2$ for $f \in H_{\text{ex}}$ and $I(f) = \infty$ otherwise.*

(It is also possible, but more complicated, to prove this from Fact 3 using the result by Vervaat [28] that the random process $B_{\text{br}}(t) - \min B_{\text{br}}$ has the same distribution as $B_{\text{ex}}(U+t)$, where U is uniform on $[0,1]$ and independent of B_{ex}, and addition is modulo 1.)

Finally, we apply Fact 2 once more, now to $\Phi : C_{\text{ex}}[0,1] \to \mathbb{R}$ and find that $\varepsilon^{1/2} X = \Phi(\varepsilon^{1/2} B_{\text{ex}})$ satisfies the LDP in $[0, \infty)$ with the good rate function, for $x > 0$,

$$\begin{aligned}\inf_{f \in H_{\text{ex}} \colon \Phi(f) = x} \tfrac{1}{2}\|f'\|_2^2 &= \inf_{f \in H_{\text{ex}} \colon \Phi(f) \neq 0} \tfrac{1}{2}\left\|\frac{xf'}{\Phi(f)}\right\|_2^2 \\ &= \inf_{f \in H_{\text{ex}} \colon \|f'\|_2 = 1,\, \Phi(f) \neq 0} \tfrac{1}{2}\left(\frac{x}{\Phi(f)}\right)^2 = \frac{1}{2\gamma^2} x^2.\end{aligned}$$

Taking $A = (1, \infty)$ and $\varepsilon = x^{-2}$ in the definition of LDP, this proves (3). Finally, (4) and (5) follow easily from (3) by integration; indeed, the (more difficult) converses hold too, see Davies [9] and Kasahara [22] or [20, Theorem 4.5]. This completes the proof of Theorem 2.1. $\square$

Proof of Theorem 2.4. We begin by observing that the following extension of Fact 1 holds, also in d dimensions.

Fact 5. *If $0 < \beta < 1/2$, then $(\varepsilon^{1/2} B)$ satisfies the LDP in $C_{\text{bm}}^\beta[0,1] := C_{\text{bm}}[0,1] \cap C^\beta[0,1]$ with good rate function $I(f) = \frac{1}{2}\|f'\|_2^2$ for $f \in H \cap C_{\text{bm}}^\beta[0,1]$ and $I(f) = \infty$ otherwise.*

Indeed, by [21, Theorem 27.11(ii)]], this follows from Fact 1 and the property that $(\varepsilon^{1/2} B)$ is *exponentially tight* in $C_{\text{bm}}^\beta[0,1]$, *i.e.*, that for every $M < \infty$ there exists a compact subset $K \subset C_{\text{bm}}^\beta[0,1]$ such that

$$\limsup_{\varepsilon \to 0} \Bigl(\varepsilon \ln \mathbb{P}(\varepsilon^{1/2} B \notin K)\Bigr) \le -M; \tag{12}$$

this exponential tightness is easily verified by choosing a γ with $\beta < \gamma < 1/2$ and taking $K = \{f \in C_{\text{bm}}^\beta[0,1] : \|f\|_{C^\gamma} \le L\}$ for a large L. We omit the verifications that K is compact and satisfies (12) if $L = L(M)$ is large enough.

The rest of proof of Theorem 2.4 is entirely the same as for Theorem 2.1. □

Acknowledgements. We thank Philippe Chassaing, Philippe Flajolet, and Nevin Kapur for helpful comments.

References

[1] D. Aldous, The continuum random tree II: an overview. *Stochastic Analysis (Durham, 1990)*, 23–70, London Math. Soc. Lecture Note Ser. 167, Cambridge Univ. Press, Cambridge, 1991.
[2] D. Aldous, The continuum random tree III. *Ann. Probab.* **21** (1993), no. 1, 248–289.
[3] P. Biane, J. Pitman, and M. Yor, Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. *Bull. Amer. Math. Soc. (N.S.)* **38** (2001), no. 4, 435–465.
[4] M. Bousquet-Mélou and S. Janson, The density of the ISE and local limit laws for embedded trees. *Ann. Appl. Probab.* **16** (2006), no. 3, 1597–1632.
[5] R. H. Cameron and W. T. Martin, Transformations of Wiener integrals under translations. *Ann. of Math.* (2) **45** (1944), 386–396.
[6] P. Chassaing, J.F. Marckert, and M. Yor, The height and width of simple trees. *Mathematics and Computer Science (Versailles, 2000)*, 17–30, Trends Math., Birkhäuser, Basel, 2000.
[7] K.L. Chung, Excursions in Brownian motion. *Ark. Mat.* **14** (1976), no. 2, 155–177.
[8] M. Csörgő, Z. Shi, and M. Yor, Some asymptotic properties of the local time of the uniform empirical process. *Bernoulli* **5** (1999), no. 6, 1035–1058.
[9] L. Davies. Tail probabilities for positive random variables with entire characteristic functions of very regular growth. *Z. Angew. Math. Mech.* **56** (1976), no. 3, T334–T336.
[10] A. Dembo and O. Zeitouni, *Large deviations techniques and applications.* Jones and Bartlett Publishers, Boston, MA, 1993.
[11] J. A. Fill and S. Janson, Brownian excursion representation for the sum of αth powers of subtree sizes for conditioned Galton–Watson trees. In preparation.
[12] J. A. Fill and N. Kapur, Limiting distributions for additive functionals on Catalan trees. *Theoretical Computer Science* **326** (2004), 69–102.
[13] J. A. Fill and N. Kapur, An invariance principle for simply generated families of trees. Unpublished manuscript.
[14] J. A. Fill and N. Kapur, Catalan trees with toll (n^α): asymptotics of moments for $\alpha \geq 1$. Unpublished manuscript.
[15] P. Flajolet, Z. Gao, A. Odlyzko, and B. Richmond, The distribution of heights of binary and other simple trees. *Combin. Probab. Comput.* **2** (1993), 145–156.
[16] S. Janson, *Gaussian Hilbert Spaces.* Cambridge University Press, Cambridge, UK, 1997.
[17] S. Janson, The Wiener index of simply generated random trees. *Random Struct. Alg.* **22** (2003), no. 4, 337–358.
[18] S. Janson, Random cutting and records in deterministic and random trees. *Random Struct. Alg.* **29** (2006), 139–179.
[19] S. Janson, Left and right pathlengths in random binary trees. *Algorithmica* **46** (2006), 419–429.
[20] S. Janson and P. Chassaing, The center of mass of the ISE and the Wiener index of trees. *Electronic Comm. Probab.* **9** (2004), paper 20, 178–187.
[21] O. Kallenberg, *Foundations of modern probability.* 2nd ed., Springer, New York, 2002.

[22] Y. Kasahara. Tauberian theorems of exponential type. *J. Math. Kyoto Univ.* **18** (1978), no. 2, 209–219.

[23] D.P. Kennedy, The distribution of the maximum Brownian excursion. *J. Appl. Probab.* **13** (1976), no. 2, 371–376.

[24] C. Knessl and W. Szpankowski, Quicksort algorithm again revisited. *Discrete Math. Theor. Comput. Sci.* **3** (1999), 43–64.

[25] J.-F. Marckert, The rotation correspondence is asymptotically a dilatation. *Random Struct. Alg.* **24** (2004), no. 2, 118–132.

[26] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion.* 3^{rd} edition, Springer, Berlin, 1999.

[27] L. Serlet, A large deviation principle for the Brownian snake. *Stochastic Process. Appl.* **67** (1997), no. 1, 101–115.

[28] W. Vervaat, A relation between Brownian bridge and Brownian excursion. *Ann. Probab.* **7** (1979), no. 1, 143–149.

Department of Applied Mathematics and Statistics, The Johns Hopkins University

E-mail address: `jimfill@jhu.edu`

URL: `http://www.ams.jhu.edu/~fill/`

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: `svante.janson@math.uu.se`

URL: `http://www.math.uu.se/~svante/`