

On the limiting distribution of the number of improper edges for random trees

Wennie W.J. Ma ¹ and Kathy Q. Ji ²

Center for Applied Mathematics, KL-AAGDM

Tianjin University

Tianjin 300072, P.R. China

Emails: ¹winniema@tju.edu.cn, ²kathyji@tju.edu.cn

Abstract: Improper edges were introduced by Shor to refine Cayley’s formula for rooted labeled trees. Zeng established a connection between Shor’s refinement and the Ramanujan polynomials. Let $\mathcal{T}_n$ denote the set of rooted labeled trees on $[n] = \{1, \dots, n\}$. We prove that the number of improper edges in a uniformly random tree in $\mathcal{T}_n$ is asymptotically normal as $n \rightarrow \infty$, with mean and variance asymptotic to μn and $\sigma^2 n$, respectively, where $\mu = e - 2$ and $\sigma^2 = e^2 - 3e + 1$. This phenomenon was observed by Chen, and the proof presented here was developed through human–AI collaboration.

Keywords: Rooted labeled trees; improper edges; Ramanujan polynomials; asymptotic normality.

AMS Classification: 05C05, 05A16, 60F05.

1 Introduction

Motivated by an observation of Chen [3], we establish the asymptotic normality of the number of improper edges in uniformly random rooted labeled trees. Let $\mathcal{T}_n$ denote the set of rooted labeled trees on the vertex set $[n] = \{1, 2, \dots, n\}$, equipped with the uniform probability measure. Throughout this paper, rooted trees are understood to be unordered, so the children of each vertex are not equipped with a linear order. We identify each vertex with its label. For an edge (i, j) of a tree $T \in \mathcal{T}_n$, we always take i to be the parent of j . The edge (i, j) is called *proper* if i is smaller than the label of every vertex in the subtree rooted at j , including j itself; otherwise, it is called *improper*.

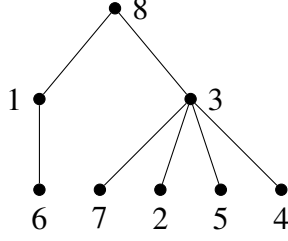


Figure 1: A rooted labeled tree on $[8]$ with three improper edges.

For example, the tree $T \in \mathcal{T}_8$ shown in Fig. 1 has exactly three improper edges: $(8, 1)$, $(8, 3)$, and $(3, 2)$. Its remaining four edges are proper. Our main result is the following.

Theorem 1.1. *Let X_n denote the number of improper edges in a uniformly random rooted labeled tree on $[n]$. As $n \rightarrow \infty$,*

$$\mathbb{E}X_n = \mu n - \frac{e}{2} + O(n^{-1}), \quad (1.1)$$

$$\text{Var}(X_n) = \sigma^2 n + \frac{e(3-e)}{2} + O(n^{-1}), \quad (1.2)$$

where

$$\mu = e - 2, \quad \sigma^2 = e^2 - 3e + 1.$$

Moreover,

$$\frac{X_n - \mu n}{\sigma \sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1). \quad (1.3)$$

Improper edges of rooted labeled trees were introduced by Shor [17] to refine Cayley's formula [2]. Let $r(n, k)$ denote the number of rooted labeled trees on $[n]$ with exactly k improper edges. Shor [17] proved that, for $n \geq 2$ and $0 \leq k \leq n - 1$,

$$r(n, k) = (n - 1)r(n - 1, k) + (n + k - 2)r(n - 1, k - 1), \quad (1.4)$$

with the initial condition $r(1, 0) = 1$ and the convention that $r(n, k) = 0$ for $k < 0$ or $k \geq n$. The numbers $r(n, k)$ is listed as sequences A054589 in OEIS. By Cayley's formula,

$$\sum_{k=0}^{n-1} r(n, k) = n^{n-1}.$$

For $n \geq 1$, define

$$R_n(q) = \sum_{k=0}^{n-1} r(n, k) q^k.$$

Below are the first few values of $R_n(q)$:

$$\begin{aligned}
R_1(q) &= 1, \\
R_2(q) &= 1 + q, \\
R_3(q) &= 2 + 4q + 3q^2, \\
R_4(q) &= 6 + 18q + 25q^2 + 15q^3, \\
R_5(q) &= 24 + 96q + 190q^2 + 210q^3 + 105q^4, \\
R_6(q) &= 120 + 600q + 1526q^2 + 2380q^3 + 2205q^4 + 945q^5.
\end{aligned}$$

Zeng [18] identified the connection between Shor's refinement of Cayley's formula and the Ramanujan polynomials [1, 15]. Dumont and Ramamonjisoa [10] independently obtained the interpretation of their coefficients in terms of rooted labeled trees with a prescribed number of improper edges. For further details, see Chen, Fu, and Wang [4] and Chen and Yang [5]. Chen, Wang, and Yang [6] proved that the sequence $(R_n(q))_{n \geq 1}$ is strongly q -log-convex; that is,

$$R_{m-1}(q)R_{n+1}(q) \geq_q R_m(q)R_n(q) \quad (n \geq m \geq 2),$$

where $f(q) \geq_q g(q)$ means that $f(q) - g(q)$ has nonnegative coefficients. Sokal [16] established the stronger property that the Hankel matrix

$$(R_{i+j+1}(q))_{i,j \geq 0}$$

is coefficientwise totally positive, meaning that each of its minors is a polynomial in q with nonnegative coefficients.

Although the recurrence (1.4) is relatively simple, it does not immediately yield the limiting distribution of the number of improper edges. In Section 2, we show that $R_n(q)$ is not real-rooted for $n \geq 3$, so standard methods for deducing asymptotic normality from real-rootedness do not apply directly. Instead, in Section 3, we prove the main result using an analytic approach based on singularity analysis and Hwang's quasi-powers theorem.

2 Non-real-rootedness of $R_n(q)$

Lin and Zeng [14, proof of Corollary 5.1] showed that the coefficient sequence of $R_n(q)$ is log-concave and hence unimodal. The following proposition shows, however, that $R_n(q)$ is not real-rooted for $n \geq 3$. Thus, these polynomials provide a family of examples of log-concavity without real-rootedness.

Proposition 2.1. *For $n \geq 3$, the polynomials $R_n(q)$ are not real-rooted.*

Proof. It follows by induction from (1.4) that, for $n \geq 2$,

$$r(n, 0) = (n-1)!, \quad r(n, 1) = (n-1)(n-1)!. \quad (2.1)$$

For $n \geq 2$, set

$$b_n = \frac{r(n, 2)}{(n-1)!}.$$

Using (2.1), the recurrence (1.4) gives

$$b_{n+1} = b_n + \frac{n+1}{n}(n-1) = b_n + n - \frac{1}{n}.$$

Since $b_2 = 0$, we obtain

$$b_n = \sum_{j=1}^{n-1} \left(j - \frac{1}{j} \right) = \binom{n}{2} - H_{n-1}, \quad (2.2)$$

where $H_m = \sum_{j=1}^m 1/j$ denotes the m th harmonic number.

Fix $n \geq 3$ and suppose, for a contradiction, that $R_n(q)$ is real-rooted. Since $R_n(q)$ has degree $n-1$ and strictly positive coefficients, all its zeros must be negative. Thus, by (2.1), there exist $\lambda_1, \dots, \lambda_{n-1} > 0$ such that

$$\frac{R_n(q)}{(n-1)!} = \prod_{j=1}^{n-1} (1 + \lambda_j q).$$

Comparing the coefficients of q and q^2 and using (2.1) and (2.2), we obtain

$$\sum_{j=1}^{n-1} \lambda_j = n-1, \quad \sum_{1 \leq i < j \leq n-1} \lambda_i \lambda_j = \binom{n}{2} - H_{n-1}.$$

By the Cauchy–Schwarz inequality,

$$\sum_{j=1}^{n-1} \lambda_j^2 \geq \frac{1}{n-1} \left(\sum_{j=1}^{n-1} \lambda_j \right)^2 = n-1.$$

Consequently,

$$\begin{aligned} \sum_{1 \leq i < j \leq n-1} \lambda_i \lambda_j &= \frac{1}{2} \left[\left(\sum_{j=1}^{n-1} \lambda_j \right)^2 - \sum_{j=1}^{n-1} \lambda_j^2 \right] \\ &\leq \frac{(n-1)^2 - (n-1)}{2} = \binom{n-1}{2}. \end{aligned}$$

On the other hand, since $H_{n-1} < n-1$ for $n \geq 3$,

$$\sum_{1 \leq i < j \leq n-1} \lambda_i \lambda_j = \binom{n}{2} - H_{n-1} > \binom{n}{2} - (n-1) = \binom{n-1}{2},$$

a contradiction. Therefore, $R_n(q)$ is not real-rooted for $n \geq 3$. ■

3 The limiting distribution

Throughout this paper, a labeled tree is understood to be rooted and unordered, with the vertices of a tree of size n labeled by the distinct elements of $[n] = \{1, \dots, n\}$. Recall that $\mathcal{T}_n$ denotes the set of such trees, and write $T_n = |\mathcal{T}_n|$. Cayley's formula [2] gives $T_n = n^{n-1}$. The exponential generating function

$$\mathcal{T}(z) := \sum_{n \geq 1} T_n \frac{z^n}{n!},$$

known as the *Cayley tree function*, is characterized via the functional equation

$$\mathcal{T}(z) = ze^{\mathcal{T}(z)}, \quad (3.1)$$

and is thus closely related to the Lambert W -function [7, 11].

Our starting point is the bivariate exponential generating function

$$F(z, q) = \sum_{n \geq 1} R_n(q) \frac{z^n}{n!}. \quad (3.2)$$

We express $F(z, q)$ in terms of the Cayley tree function $\mathcal{T}(z)$. By tracking the dominant singularity of $F(z, q)$ as q varies near 1, we obtain a singular expansion that is uniform for q in a complex neighborhood of 1. Singularity analysis then yields a quasi-powers expansion for the probability generating function

$$\mathbb{E}[q^{X_n}] = \sum_{k=0}^{n-1} \mathbb{P}(X_n = k) q^k = \frac{R_n(q)}{R_n(1)}.$$

Hwang's quasi-powers theorem consequently yields the asymptotic normality of X_n .

We begin by recalling some relevant background.

3.1 Preliminaries

We first recall the notions of a Δ -domain and Δ -analyticity, which are used to describe the singular behavior of $\mathcal{T}(z)$.

Definition 3.1. [11, Definition VI.1, p. 389] Given two numbers ϕ, R with $R > 1$ and $0 < \phi < \frac{\pi}{2}$, the open domain $\Delta(\phi, R)$ is defined as

$$\Delta(\phi, R) = \{z \mid |z| < R, z \neq 1, |\arg(z-1)| > \phi\}.$$

A domain is a Δ -domain at 1 if it is a $\Delta(\phi, R)$ for some R and ϕ . For a complex number $\zeta \neq 0$, a Δ -domain at ζ is the image by the mapping $z \mapsto \zeta z$ of a Δ -domain at 1. A function is Δ -analytic if it is analytic in some Δ -domain.

The following standard properties of the Cayley tree function will be used in our analysis; see Flajolet and Sedgewick [11, Example VI.8, pp. 403–404] and Corless et al. [7, Secs. 3–4, pp. 339–351].

Proposition 3.2. *The Cayley tree function $\mathcal{T}(z)$ has the following properties:*

- (i) *The radius of convergence of $\mathcal{T}(z)$ is e^{-1} . Moreover, the Cayley tree function $\mathcal{T}(z)$ is analytic in the slit plane*

$$\mathbb{C} \setminus [e^{-1}, \infty).$$

- (ii) *As $z \rightarrow e^{-1}$ in a Δ -domain at e^{-1} , the following Puiseux expansion holds:*

$$\mathcal{T}(z) = 1 - \sqrt{2(1 - ez)} + \frac{2}{3}(1 - ez) + O((1 - ez)^{3/2}), \quad (3.3)$$

where the branch of $(1 - ez)^{1/2}$ is chosen by analytic continuation from the interval $(0, e^{-1})$, so that

$$(1 - ez)^{1/2} > 0$$

for real $z \in (0, e^{-1})$.

- (iii) *If $z \neq 0$ and $\mathcal{T}(z) \neq 1$, then*

$$\mathcal{T}'(z) = \frac{\mathcal{T}(z)}{z(1 - \mathcal{T}(z))}. \quad (3.4)$$

These analytic properties also play a role in classical limit theorems for statistics on rooted labeled trees, see, for example, the root degree and the number of leaves [11, Examples IX.6 and IX.25], fixed-degree vertex counts [9, Theorem 2.3 and the subsequent remark], and occurrences of fixed tree patterns with matching internal vertex degrees [8, Theorem 1.1].

The two theorems below enable us to derive the asymptotic expansion of the normalized probability generating function $\mathbb{E}(q^{X_n})$ based on the asymptotic behavior of the bivariate exponential generating function $F(z, q)$.

Theorem 3.3. [11, Theorem VI.1 (Standard function scale), p. 381] *Let α be an arbitrary complex number in $\mathbb{C} \setminus \mathbb{Z}_{\leq 0}$. The coefficient of z^n in*

$$f(z) = (1 - z)^{-\alpha}$$

admits for large n a complete asymptotic expansion in descending powers of n ,

$$[z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)} \left(1 + \sum_{k=1}^{\infty} \frac{e_k}{n^k} \right),$$

where e_k is a polynomial in α of degree $2k$. In particular:

$$[z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)} \left(1 + \frac{\alpha(\alpha-1)}{2n} + \frac{\alpha(\alpha-1)(\alpha-2)(3\alpha-1)}{24n^2} + \frac{\alpha^2(\alpha-1)^2(\alpha-2)(\alpha-3)}{48n^3} + O\left(\frac{1}{n^4}\right) \right). \quad (3.5)$$

Theorem 3.4. [11, Theorem VI.3 (Transfer, Big-Oh and little-oh), p. 390] Let α, β be arbitrary real numbers, and let $f(z)$ be a function that is Δ -analytic.

1. Assume that $f(z)$ satisfies, in the intersection of a neighbourhood of 1 with its Δ -domain, the condition

$$f(z) = O \left((1-z)^{-\alpha} \left(\log \frac{1}{1-z} \right)^\beta \right).$$

Then

$$[z^n]f(z) = O \left(n^{\alpha-1} (\log n)^\beta \right).$$

2. Assume that $f(z)$ satisfies, in the intersection of a neighbourhood of 1 with its Δ -domain, the condition

$$f(z) = o \left((1-z)^{-\alpha} \left(\log \frac{1}{1-z} \right)^\beta \right).$$

Then

$$[z^n]f(z) = o \left(n^{\alpha-1} (\log n)^\beta \right).$$

Finally, we recall Hwang's quasi-powers theorem, which is the main probabilistic tool in our proof. We use the formulation in [11, Lemma IX.1, p. 646]; see also Hwang [12] for related results.

Theorem 3.5 (Quasi-powers, general distributions). Assume that the Laplace transforms $\lambda_n(s) = \mathbb{E}\{e^{sX_n}\}$ of a sequence of random variables X_n are analytic in a disc $|s| < \rho$, for some $\rho > 0$, and satisfy there is an expansion of the form

$$\lambda_n(s) = e^{\beta_n U(s) + V(s)} \left(1 + O \left(\frac{1}{\kappa_n} \right) \right), \quad (3.6)$$

with $\beta_n, \kappa_n \rightarrow +\infty$ as $n \rightarrow +\infty$, and $U(s), V(s)$ analytic in $|s| \leq \rho$. Assume also the variability condition, $U''(0) \neq 0$. Under these assumptions, the mean and variance of X_n satisfy

$$\mathbb{E}(X_n) = \beta_n U'(0) + V'(0) + O(\kappa_n^{-1}), \quad (3.7)$$

$$\mathbb{V}(X_n) = \beta_n U''(0) + V''(0) + O(\kappa_n^{-1}). \quad (3.8)$$

The distribution of $X_n^* := \frac{X_n - \beta_n U'(0)}{\sqrt{\beta_n U''(0)}}$ is asymptotically Gaussian, the speed of convergence to the Gaussian limit being $O(\kappa_n^{-1} + \beta_n^{-1/2})$.

3.2 Proof of Theorem 1.1

To apply Hwang's quasi-powers theorem and prove Theorem 1.1, the key step is to obtain an asymptotic expansion for the normalized probability generating function $\mathbb{E}(q^{X_n})$ of X_n . To this end, we first consider the bivariate exponential generating function $F(z, q)$ given by (3.2). Our goal is to express $F(z, q)$ explicitly in terms of the Cayley tree function and then analyze its dominant singularity uniformly in q .

Josuat-Vergès [13, Theorem 4.2] showed that

$$\sum_{n \geq 0} G_n(x) \frac{z^n}{n!} = \mathcal{T} \left(\frac{z+x}{1+x} \exp \left(-\frac{x}{1+x} \right) \right), \quad (3.9)$$

where $G_0(x-1) = \frac{x-1}{x} = 1 - \frac{1}{x}$ and $G_n(x-1) = R_n(x)$ for $n \geq 1$.

Setting $x = q - 1$ in (3.9) and shifting the constant term to the right-hand side yields

$$F(z, q) = \sum_{n \geq 1} R_n(q) \frac{z^n}{n!} = \frac{1}{q} - 1 + \mathcal{T} \left(e^{1/q-1} \frac{z+q-1}{q} \right). \quad (3.10)$$

The following uniform Δ -domains of $F(z, q)$ are needed in order to apply the transfer theorem uniformly with respect to the parameter q .

Lemma 3.6. *There exists $\varepsilon > 0$ such that, whenever $|q - 1| \leq \varepsilon$, the function $F(z, q)$, as a function of z , has the unique dominant singularity $z = \rho(q)$. Moreover, one can choose a family of Δ -domains, uniformly in q ,*

$$\Delta_q = \left\{ z \mid |z| < |\rho(q)| + \eta, z \neq \rho(q), \left| \arg \left(\frac{z}{\rho(q)} - 1 \right) \right| > \theta \right\}, \quad (3.11)$$

where $\eta > 0$ and $0 < \theta < \pi/2$ are independent of q , such that $F(z, q)$ is analytic in Δ_q .

Proof. Set

$$w = \omega(z, q) := A(q)z + B(q), \quad (3.12)$$

where

$$A(q) = \frac{e^{1/q-1}}{q}, \quad B(q) = \frac{e^{1/q-1}(q-1)}{q}.$$

Both $A(q)$ and $B(q)$ are analytic in a neighborhood of $q = 1$. Since $A(1) = 1$, continuity of $A(q)$ at $q = 1$ implies that $A(q) \neq 0$ for all q in a sufficiently small complex neighbourhood of 1. By Proposition 3.2 (i) and (ii), the principal branch of the Cayley tree function $\mathcal{T}(z)$ is analytic in $\mathbb{C} \setminus [e^{-1}, \infty)$ and its dominant singularity is the square-root branch point $z = e^{-1}$. From the explicit representation (3.10), the possible obstructions to analytic continuation of the chosen branch of $F(z, q)$ are contained in the preimage of the slit $[e^{-1}, \infty)$ under the affine map $z \mapsto \omega(z, q)$.

The preimage of the endpoint $w = e^{-1}$ is determined by

$$A(q)z + B(q) = e^{-1}.$$

Hence

$$z = \rho(q) := \frac{e^{-1} - B(q)}{A(q)} = 1 - q + qe^{-1/q}.$$

Evidently, $\rho(q)$ is analytic in a neighbourhood of $q = 1$ and $\rho(1) = e^{-1} > 0$.

Next we describe the whole preimage of the slit. Set $w(z, q) = e^{-1} + \lambda$ with $\lambda \geq 0$. By (3.12), we have

$$A(q)z + B(q) = e^{-1} + \lambda.$$

Hence, for $\lambda \geq 0$, we have

$$z = \rho(q) + \lambda d(q), \quad d(q) = \frac{1}{A(q)}.$$

Thus, the slit in the w -plane pulls back to the ray

$$L_q = \{\rho(q) + \lambda d(q) : \lambda \geq 0\}.$$

At $q = 1$, we have $\rho(1) = e^{-1}$, $d(1) = 1$. Therefore

$$\operatorname{Re}(\overline{\rho(1)}d(1)) = e^{-1} > 0.$$

By continuity of the function $\operatorname{Re}(\overline{\rho(q)}d(q))$ at $q = 1$, there exists $\varepsilon > 0$ such that $\operatorname{Re}(\overline{\rho(q)}d(q)) > 0$ for all $|q - 1| \leq \varepsilon$. Consequently, at this moment, for every $\lambda > 0$,

$$\begin{aligned} |\rho(q) + \lambda d(q)|^2 &= |\rho(q)|^2 + 2\lambda \operatorname{Re}(\overline{\rho(q)}d(q)) + \lambda^2 |d(q)|^2 \\ &> |\rho(q)|^2. \end{aligned}$$

Thus, the endpoint $z = \rho(q)$ is the unique point of minimal modulus on L_q . Since $A(q) \neq 0$, the representation (3.10) and Proposition 3.2(ii) imply that $z = \rho(q)$ is a square-root branch point of $F(z, q)$. Hence $z = \rho(q)$ is the unique dominant singularity of $F(z, q)$.

It remains to justify that $F(z, q)$ is analytic in Δ_q given by (3.11). Since

$$\frac{d(1)}{\rho(1)} = e > 0,$$

by continuity of the function $d(q)/\rho(q)$ at $q = 1$, there exists $\varepsilon > 0$ and $0 < \theta < \pi/2$ such that $\left| \arg \frac{d(q)}{\rho(q)} \right| \leq \theta$ for all $|q - 1| \leq \varepsilon$. Therefore $\forall z \in L_q \setminus \{\rho(q)\}$,

$$\left| \arg \left(\frac{z}{\rho(q)} - 1 \right) \right| = \left| \arg \lambda \frac{d(q)}{\rho(q)} \right| = \left| \arg \frac{d(q)}{\rho(q)} \right| \leq \theta.$$

Since q ranges over a sufficiently small compact neighbourhood of 1, the constants η and θ can be chosen uniformly. Since $\rho(q) \notin \Delta_q$, it follows that $\Delta_q \cap L_q = \emptyset$. Hence $F(z, q)$ is analytic in Δ_q . This completes the proof. $\blacksquare$

Define

$$\rho(q) = 1 - q + qe^{-1/q}, \quad \kappa(q) = \frac{e^{1/q}\rho(q)}{q}. \quad (3.13)$$

Then $\rho(1) = e^{-1}$ and $\kappa(1) = 1$.

Lemma 3.7. *There exists $\varepsilon > 0$ such that, as $n \rightarrow \infty$, uniformly for $|q - 1| \leq \varepsilon$,*

$$R_n(q) = \frac{n!}{\sqrt{2\pi}} \sqrt{\kappa(q)} \rho(q)^{-n} n^{-3/2} (1 + O(n^{-1})), \quad (3.14)$$

where $\rho(q)$ and $\kappa(q)$ are given in (3.13).

Proof. Under the affine substitution adopted in (3.12), we have

$$1 - ew(z, q) = 1 - e(e^{1/q-1} \frac{z + q - 1}{q}) = \kappa(q) \left(1 - \frac{z}{\rho(q)}\right). \quad (3.15)$$

Substituting this relation into (3.3) from Proposition 3.2, together with the identity (3.10), we further deduce that

$$F(z, q) = H_1(z, q) - \sqrt{2\kappa(q)} \left(1 - \frac{z}{\rho(q)}\right)^{1/2} + O\left(\left(1 - \frac{z}{\rho(q)}\right)^{3/2}\right),$$

where

$$H_1(z, q) = \frac{1}{q} + \frac{2}{3}\kappa(q) \left(1 - \frac{z}{\rho(q)}\right).$$

In particular, $H_1(z, q)$ is analytic in z at $z = \rho(q)$. By Lemma 3.6, this expansion holds in a family of Δ -domains uniformly in q .

For $n \geq 2$, the analytic term $H_1(z, q)$ has zero z^n -coefficient. Hence only the singular term and the remainder have to be considered.

By Theorem 3.3, we have

$$\begin{aligned} [z^n] \left(1 - \frac{z}{\rho(q)}\right)^{1/2} &= \frac{\rho(q)^{-n} n^{-3/2}}{\Gamma(-1/2)} \left(1 + \frac{3}{8n} + O(n^{-2})\right) \\ &= \frac{\rho(q)^{-n} n^{-3/2}}{\Gamma(-1/2)} (1 + O(n^{-1})). \end{aligned}$$

On the other hand, by Theorem 3.4,

$$[z^n] O\left(\left(1 - \frac{z}{\rho(q)}\right)^{3/2}\right) = O(\rho(q)^{-n} n^{-5/2}).$$

Since

$$\frac{\rho(q)^{-n} n^{-5/2}}{\rho(q)^{-n} n^{-3/2}} = n^{-1},$$

the remainder is of relative order $O(n^{-1})$ with respect to the leading term.

By Lemma 3.6, the corresponding Δ -domains and the constants in the local expansions can be chosen uniformly for q near 1. Consequently, the above coefficient estimates are uniform in q . As $n \rightarrow \infty$, we thus obtain

$$[z^n]F(z, q) = -\frac{\sqrt{2\kappa(q)}}{\Gamma(-1/2)}\rho(q)^{-n}n^{-3/2}(1 + O(n^{-1})), \quad (3.16)$$

uniformly for q in a neighbourhood of 1.

Substituting

$$\Gamma(-1/2) = -2\sqrt{\pi}, \quad R_n(q) = n![z^n]F(z, q),$$

into (3.16) yields the asymptotic estimate (3.14). This completes the proof. $\blacksquare$

We are now in a position to derive the asymptotic expansion of the normalized probability generating function $\mathbb{E}(q^{X_n})$ by virtue of Lemma 3.7.

Theorem 3.8. *Let X_n denote the number of improper edges in a uniformly random labeled rooted tree on $[n]$. There exists $\delta > 0$ such that, uniformly for $|s| \leq \delta$, as $n \rightarrow \infty$,*

$$\mathbb{E}(e^{sX_n}) = \sqrt{\kappa(e^s)} \left(\frac{\rho(1)}{\rho(e^s)} \right)^n (1 + O(n^{-1})), \quad (3.17)$$

where $\rho(q)$ and $\kappa(q)$ are given in (3.13).

Proof. The probability generating function of X_n is

$$\mathbb{E}(q^{X_n}) = \sum_{k=0}^{n-1} q^k \mathbb{P}(X_n = k) = \frac{R_n(q)}{R_n(1)}.$$

Setting $q = e^s$, we obtain the moment generating function

$$M_n(s) := \mathbb{E}(e^{sX_n}) = \frac{R_n(e^s)}{R_n(1)}.$$

By Lemma 3.7, there exist constants $\varepsilon, C > 0$ such that, for all sufficiently large n ,

$$R_n(q) = \frac{n!}{\sqrt{2\pi}} \sqrt{\kappa(q)} \rho(q)^{-n} n^{-3/2} (1 + \xi_n(q)), \quad |q - 1| \leq \varepsilon,$$

where the remainder satisfies

$$\sup_{|q-1| \leq \varepsilon} |\xi_n(q)| \leq \frac{C}{n}.$$

Choose $\delta > 0$ sufficiently small that $|e^s - 1| \leq \varepsilon$ whenever $|s| \leq \delta$. For all sufficiently large n , we have $C/n \leq 1/2$, and therefore

$$|1 + \xi_n(1)| \geq 1 - |\xi_n(1)| \geq 1 - \frac{C}{n} \geq \frac{1}{2}.$$

Since $\kappa(1) = 1$, we obtain

$$\begin{aligned} M_n(s) &= \frac{R_n(e^s)}{R_n(1)} \\ &= \sqrt{\kappa(e^s)} \left(\frac{\rho(1)}{\rho(e^s)} \right)^n \frac{1 + \xi_n(e^s)}{1 + \xi_n(1)} \\ &= \sqrt{\kappa(e^s)} \left(\frac{\rho(1)}{\rho(e^s)} \right)^n \left(1 + \frac{\xi_n(e^s) - \xi_n(1)}{1 + \xi_n(1)} \right). \end{aligned}$$

Moreover, uniformly for $|s| \leq \delta$,

$$\left| \frac{\xi_n(e^s) - \xi_n(1)}{1 + \xi_n(1)} \right| \leq \frac{|\xi_n(e^s)| + |\xi_n(1)|}{|1 + \xi_n(1)|} \leq \frac{2C/n}{1 - C/n} \leq \frac{4C}{n}.$$

Consequently, as $n \rightarrow \infty$,

$$M_n(s) = \sqrt{\kappa(e^s)} \left(\frac{\rho(1)}{\rho(e^s)} \right)^n (1 + O(n^{-1})),$$

uniformly for $|s| \leq \delta$. ■

We conclude the paper by proving our main result via Theorem 3.8 and Hwang's quasi-powers theorem 3.5.

Proof of Theorem 1.1. By Theorem 3.8, the moment generating function of X_n satisfies, uniformly for s in a complex neighbourhood of 0,

$$M_n(s) := \mathbb{E}(e^{sX_n}) = \exp(nU(s) + V(s)) (1 + O(n^{-1})), \quad (3.18)$$

where

$$U(s) = \log \frac{\rho(1)}{\rho(e^s)}, \quad V(s) = \frac{1}{2} \log \kappa(e^s).$$

Moreover,

$$U(0) = V(0) = 0.$$

Thus the hypotheses of Theorem 3.5 will be satisfied once we verify that $U''(0) > 0$.

We now compute the relevant derivatives. Since $\rho(q) = 1 - q + qe^{-1/q}$, we have

$$\rho'(q) = -1 + e^{-1/q} \left(1 + \frac{1}{q} \right), \quad \rho''(q) = \frac{e^{-1/q}}{q^3}.$$

Therefore

$$\rho(1) = e^{-1}, \quad \frac{\rho'(1)}{\rho(1)} = 2 - e, \quad \frac{\rho''(1)}{\rho(1)} = 1.$$

Let $D = q \frac{d}{dq}$. Under the substitution $q = e^s$, differentiation with respect to s corresponds to applying the operator D . Hence

$$U'(0) = -D \log \rho(q) \Big|_{q=1} = -\frac{\rho'(1)}{\rho(1)} = e - 2 = \mu.$$

Furthermore,

$$\begin{aligned} U''(0) &= -D^2 \log \rho(q)|_{q=1} \\ &= -\left(\frac{\rho'(1)}{\rho(1)} + \frac{\rho''(1)}{\rho(1)} - \left(\frac{\rho'(1)}{\rho(1)} \right)^2 \right) \\ &= e^2 - 3e + 1 = \sigma^2. \end{aligned}$$

Also,

$$\sigma^2 = e^2 - 3e + 1 > 0.$$

Next, since $\kappa(q) = \frac{e^{1/q} \rho(q)}{q}$, we have

$$\log \kappa(q) = \frac{1}{q} + \log \rho(q) - \log q.$$

It follows that

$$V(s) = \frac{1}{2} \log \kappa(e^s) = \frac{1}{2} (e^{-s} + \log \rho(e^s) - s).$$

Using the values computed above, we obtain

$$V'(0) = \frac{1}{2} \left(-1 + \frac{\rho'(1)}{\rho(1)} - 1 \right) = -\frac{e}{2}.$$

Moreover,

$$V''(0) = \frac{1}{2} \left(1 + D^2 \log \rho(q)|_{q=1} \right).$$

Since

$$D^2 \log \rho(q)|_{q=1} = -U''(0) = -(e^2 - 3e + 1),$$

we get

$$V''(0) = \frac{1}{2} (1 - e^2 + 3e - 1) = \frac{e(3 - e)}{2}.$$

From Theorem 3.5, we have

$$\mathbb{E}X_n = \mu n - \frac{e}{2} + O(n^{-1}),$$

and

$$\text{Var}(X_n) = \sigma^2 n + \frac{e(3 - e)}{2} + O(n^{-1}).$$

Moreover,

$$\frac{X_n - \mu n}{\sigma \sqrt{n}} \xrightarrow{d} N(0, 1).$$

This completes the proof. ■

Acknowledgment. This work was supported by the National Science Foundation of China.

References

- [1] B. C. Berndt, *Ramanujan's notebooks. Part I*, Springer, New York, 1985; MR0781125. [3](#)
- [2] A. Cayley, A theorem on trees, *Quart. J. Math.* 23 (1889), 376–378. [2](#), [5](#)
- [3] W. Y. C. Chen, Some observations and questions via Maple, talk at AlCoVE: an Algebraic Combinatorics Virtual Expedition, virtual conference, June 8–9, 2026. [1](#)
- [4] W. Y. C. Chen, A. M. Fu and E. L. Wang, A grammatical calculus for the Ramanujan polynomials, *Ramanujan J.*, 70 (2026) 35. [3](#)
- [5] W. Y. C. Chen, H. R. L. Yang, A context-free grammar for the Ramanujan-Shor polynomials, *Adv. in Appl. Math.*, 126 (2021) 101908. [3](#)
- [6] W. Y. C. Chen, L. X. W. Wang and A. L. B. Yang, Recurrence relations for strongly q -log-convex polynomials, *Canad. Math. Bull.*, 54 (2011) 217–229. [3](#)
- [7] R. M. Corless et al., On the Lambert W function, *Adv. Comput. Math.* **5** (1996), no. 4, 329–359; MR1414285. [5](#), [6](#)
- [8] F. Chyzak et al., The distribution of patterns in random trees, *Combin. Probab. Comput.* **17** (2008), no. 1, 21–59; MR2376422. [6](#)
- [9] M. Drmota and B. Gittenberger, The distribution of nodes of given degree in random trees, *J. Graph Theory* **31** (1999), no. 3, 227–253; MR1688949. [6](#)
- [10] D. Dumont and A. Ramamonjisoa, *Grammaire de Ramanujan et arbres de Cayley*, *Electron. J. Combin.* **3** (1996), no. 2, Research Paper 17, approx. 18 pp.; MR1392502. [3](#)
- [11] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge Univ. Press, Cambridge, 2009; MR2483235. [5](#), [6](#), [7](#)
- [12] H. K. Hwang, Large deviations for combinatorial distributions. I. Central limit theorems, *Annals of Applied Probability* **6** (1996), 297–319. [7](#)
- [13] M. Josuat-Vergès, Derivatives of the tree function, *Ramanujan J.* **38** (2015), no. 1, 1–15; MR3396483. [8](#)
- [14] Z. Lin and J. Zeng, Positivity properties of Jacobi–Stirling numbers and generalized Ramanujan polynomials, *Adv. in Appl. Math.* **53** (2014), 12–27. [3](#)
- [15] S. Ramanujan, *Notebooks*, Vol. 1, Tata Institute of Fundamental Research, Bombay, 1957. [3](#)
- [16] A. D. Sokal, Total positivity of some polynomial matrices that enumerate labeled trees and forests I: forests of rooted labeled trees, *Monatsh. Math.* **200** (2023), no. 2, 389–452; MR4544303. [3](#)

- [17] P. W. Shor, A new proof of Cayley's formula for counting labeled trees, *Journal of Combinatorial Theory, Series A* 71 (1995), 154–158. [2](#)
- [18] J. Zeng, A Ramanujan sequence that refines the Cayley formula for trees, *Ramanujan J.* 3 (1999), 45–54. [3](#)