On the length of a random minimum spanning tree.

Colin Cooper* Alan Frieze† Nate Ince‡ Svante Janson§ Joel Spencer¶

March 9, 2013

Abstract

We study the expected value of the length L_n of the minimum spanning tree of the complete graph K_n when each edge e is given an independent uniform $[0,1]$ edge weight. We sharpen the result of Frieze [6] that $\lim_{n\to\infty} \mathbb{E}(L_n) = \zeta(3)$ and show that $\mathbb{E}(L_n) = \zeta(3) + \frac{c_1}{n} + \frac{c_2+o(1)}{n^{4/3}}$ where c_1, c_2 are explicitly defined constants.

1 Introduction

We study the expected value of the length L_n of the minimum spanning tree of the complete graph K_n when each edge e is given an independent uniform $[0,1]$ edge weight X_e. It was shown in Frieze [6] that

$$\lim_{n\to\infty} \mathbb{E}(L_n) = \zeta(3) = \sum_{k=1}^{\infty} \frac{1}{k^3} = 1.202\dots \tag{1.1}$$

Since then there have been several generalisations and improvements. Steele [26] extended the applicability of (1.1) distribution-wise. Janson [11] proved a central limit theorem for L_n. Penrose [22], Frieze and McDiarmid [7], Beveridge, Frieze and McDiarmid [2], Frieze, Ruszinkó and Thoma [8] analysed L_n for graphs other than the complete graph. Fill and Steele [4] used the Tutte polynomial to compute $\mathbb{E}(L_n)$ exactly for small values and Gamarnik [9] computed $\mathbb{E}_{exp}(L_n)$ exactly up to $n \le 45$ using a more efficient algorithm, where $\mathbb{E}_{exp}(L_n)$ is the expectation when the distribution of the X_e is exponential with mean one. Li and Zhang [18] consider more general distributions and prove in particular that

$$\mathbb{E}_{exp}(L_n) - \mathbb{E}(L_n) = \frac{\zeta(3)}{n} + O\left(\frac{\log^2 n}{n^2}\right). \tag{1.2}$$

Flaxman [5] gives an upper bound on the lower tail of L_n.

*Department of Computer Science, King's College, University of London, London WC2R 2LS, UK

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15217, USA. Research supported in part by NSF Grant CCF1013110

‡Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15217, USA

§Department of Mathematics, Uppsala University, SE-75310 Uppsala, Sweden. Research supported in part by the Knut and Alice Wallenberg Foundation

¶Courant Institute, New York, NY 10012 (USA)

Equation (1.1) says that $\mathbb{E}(L_n) = \zeta(3) + o(1)$ as $n \to \infty$. Ideally, one would like to have an exact expansion for $\mathbb{E}(L_n)$ as there is for the assignment problem, see Wästlund [27] and the references therein. Such an expansion has proven elusive. In this work we improve the asymptotics of $E[L_n]$ by giving the secondary and tertiary terms.

Theorem 1.

$$\mathbb{E}(L_n) = \zeta(3) + \frac{c_1}{n} + \frac{c_2 + o(1)}{n^{4/3}}$$

where

$$c_1 = -1 - \zeta(3) - \frac{1}{2}\int_{x=0}^{\infty} \log\bigl(1-(1+x)e^{-x}\bigr)\,dx$$

and

$$\begin{aligned} c_2 &= \int_{x=0}^{\infty} \Bigl(x^{-3}\psi(x^{3/2})e^{-x^3/24} - x^{-3} - \sqrt{\frac{\pi}{8}}x^{-3/2} - \frac{1}{2}\Bigr)\,dx \\ &= \frac{2}{3}\int_{y=0}^{\infty} \Bigl(y^{-2}\psi(y)e^{-y^2/24} - y^{-2} - \sqrt{\frac{\pi}{8}}y^{-1} - \frac{1}{2}\Bigr)y^{-1/3}\,dy \end{aligned}$$

with ψ defined in (1.3) *below.*

The two integral expressions defining c_2 are equal by the change of variable $x = y^{2/3}$.

A numerical integration (with `Maple`) yields $c_1 = 0.0384956\ldots$. This shows that the rate of convergence to $\zeta(3)$ is order $1/n$ and is from above. Further numerical computations show that $c_2 \approx -1.7295$, and these are explained in an appendix.

To define ψ, we let the random variable $\mathcal{B}_{\mathrm{ex}} = \int_{s=0}^{1} B_{\mathrm{ex}}(s)\,ds$ be the area under a normalized Brownian excursion; we then let

$$\psi(t) = \mathbb{E}\, e^{t\mathcal{B}_{\mathrm{ex}}}, \tag{1.3}$$

the moment generating function ψ of $\mathcal{B}_{\mathrm{ex}}$. The Brownian excursion area $\mathcal{B}_{\mathrm{ex}}$ and its moments $\mathbb{E}\,\mathcal{B}_{\mathrm{ex}}^{\ell}$ and moment generating function ψ have been studied by several authors, see e.g. Louchard [19, 20] and the survey by Janson [12], where further references are given. From these results, we derive an expression, see (1.7), that will show c_2 is well-defined. Note that $\psi(t)$ is finite for all $t > 0$ (and thus (1.3) holds for all complex t); indeed, see [12, (53)] and the references there, it is well-known that

$$\mathbb{E}\,\mathcal{B}_{\mathrm{ex}}^{\ell} \sim \sqrt{18}\,\ell\,(12e)^{-\ell/2}\ell^{\ell/2} \qquad \text{as } \ell \to \infty, \tag{1.4}$$

and thus [13, Lemma 4.1(ii)] implies, cf. [13, Remarks 3.1 and 4.9] (where $\xi = 2\mathcal{B}_{\mathrm{ex}}$),

$$\psi(t) \sim \tfrac{1}{2}t^2 e^{t^2/24} \qquad \text{as } t \to +\infty. \tag{1.5}$$

More precisely, Janson and Louchard [15] show that the density f_{ex} of $\mathcal{B}_{\mathrm{ex}}$ satisfies

$$f_{\mathrm{ex}}(x) = \frac{72\sqrt{6}}{\sqrt{\pi}}x^2 e^{-6x^2}\bigl(1 + O(x^{-2})\bigr), \qquad x > 0, \tag{1.6}$$

from which routine calculations show that

$$\psi(t) = \int_{x=0}^{\infty} e^{tx} f_{\mathrm{ex}}(x)\,dx = \frac{t^2}{2}e^{t^2/24}\bigl(1 + O(t^{-2})\bigr), \qquad t > 0. \tag{1.7}$$

Hence the integrand in the second integral defining c_2 in Theorem 1 is $O(y^{-4/3})$ as $y \to \infty$. Moreover, $\psi(0) = 1$ and $\psi'(0) = \mathbb{E}\,\mathcal{B}_{\mathrm{ex}} = \sqrt{\pi/8}$, and thus a Taylor expansion shows that the integrand is $O(y^{-1/3})$ as $y \to 0$. (Similarly, the integrand in the first integral is $O(x^{-3/2})$ and $O(1)$.) Consequently, the integrals defining c_2 converge absolutely.

2 Proof of Theorem 1

We prove the theorem by using the expression (see Janson [11]),

$$\mathbb{E}(L_n) = \int_{p=0}^{1} \mathbb{E}(\kappa(G_{n,p}))dp - 1. \tag{2.1}$$

Here $\kappa(G_{n,p})$ is the (random) number of components in the random graph $G_{n,p}$.

To evaluate (2.1) we let $\kappa(k,j,p) = \kappa_n(k,j,p)$ denote the number of components of $G_{n,p}$ with k vertices and $k+j$ edges in $G_{n,p}$. The components neatly split into three categories: trees ($j=-1$), unicyclic ($j=0$) and complex ($j \geq 1$) components. These are evaluated separately.

Lemma 2.1.

(a)

$$\int_{p=0}^{1} \sum_{k\geq 1} \mathbb{E}(\kappa(k,-1,p))dp = \zeta(3) + \frac{3(\zeta(2)-\zeta(3))}{2n} - \frac{1}{n^{4/3}} \int_{x=0}^{\infty} x^{-3}(1-e^{-x^3/24})\,dx + o(n^{-4/3}).$$

(b)

$$\int_{p=0}^{1} \sum_{k\geq 3} \mathbb{E}(\kappa(k,0,p))dp = \frac{1}{2n}\left(\zeta(3) - 3\zeta(2) - \int_{x=0}^{\infty} \log\bigl(1-(1+x)e^{-x}\bigr)\,dx.\right) - \frac{\sqrt{\pi/8}}{n^{4/3}} \int_{x=0}^{\infty} x^{-3/2}(1-e^{-x^3/24})\,dx + o(n^{-4/3}).$$

(c) With $\psi_2(x) = \psi(x) - 1 - \sqrt{\pi/2}\,x$,

$$\int_{p=0}^{1} \sum_{k\geq 1}\sum_{j\geq 1} \mathbb{E}(\kappa(k,j,p))dp = 1 - \frac{1}{n} + \frac{1}{n^{4/3}} \int_{x=0}^{\infty} \left(x^{-3}\psi_2(x^{3/2})e^{-x^3/24} - \frac{1}{2}\right)dx + o(n^{-4/3}).$$

Remark 1. Tree components contribute the main $\zeta(3)$ addend. Unicyclic components contribute a secondary $O(\frac{1}{n})$ addend. Roughly speaking there are no complex components for $p \leq \frac{1}{n}$ and precisely one complex component (the famous "giant component") for $p \geq \frac{1}{n}$. Were this to be precisely the case the contribution of complex components would be $1-\frac{1}{n}$. The additional $\Theta(n^{-4/3})$ term in Lemma 2.1 (c) comes from the behavior of complex components in the critical window $p = \frac{1}{n} + \lambda n^{-4/3}$.

Remark 2. The coefficients of $n^{-4/3}$ in Lemma 2.1(a) and (b) are easily evaluated as $-\frac{1}{8}3^{-2/3}\Gamma(1/3)$ and $-\frac{1}{2}3^{-1/6}\sqrt{\pi}\,\Gamma(5/6)$, respectively, see the appendix. The coefficient in (c) is expressed as an infinite sum and evaluated numerically in the appendix.

Proof We assume in the proof tacitly that n is large enough when necessary. We let $C_1, \dots$ denote some unimportant universal constants.

Let $C(k,\ell)$ be the number of connected graphs on vertex set $[k]$ with ℓ edges. We begin by noting the standard formula

$$\mathbb{E}\,\kappa(k,j,p) = \binom{n}{k} C(k,k+j) p^{k+j} (1-p)^{k(n-k)+\binom{k}{2}-k-j}. \tag{2.2}$$

By Cayley's formula, $C(k,k-1) = k^{k-2}$. Moreover, Wright [28] proved that for every fixed $j \ge -1$,

$$C(k,k+j) \sim w_{j+1} k^{k+3j/2-1/2} \qquad \text{as } k \to \infty, \tag{2.3}$$

for some constants $w_\ell > 0$. (See also [14, §8] and the references there. In the notation of [28], $w_{j+1} = \rho_j$.) We have $w_0 = 1$ and $w_1 = \sqrt{\pi/8}$. It was shown in Spencer [25] that

$$w_\ell = \frac{\mathbb{E}\,\mathcal{B}_{\mathrm{ex}}^\ell}{\ell!}, \qquad \ell \ge 0, \tag{2.4}$$

where $\mathcal{B}_{\mathrm{ex}}$ is the Brownian excursion area defined above. See further Janson [12]. Hence,

$$\psi(t) = \mathbb{E}\, e^{t\mathcal{B}_{\mathrm{ex}}} = \sum_{\ell=0}^{\infty} w_\ell t^\ell. \tag{2.5}$$

Let

$$\begin{aligned} A(k,k+j) &= \int_{p=0}^{1} \mathbb{E}(\kappa(k,j,p))\, dp \\ &= \binom{n}{k} C(k,k+j) \int_{p=0}^{1} p^{k+j}(1-p)^{k(n-k)+\binom{k}{2}-k-j}\, dp \\ &= \binom{n}{k} C(k,k+j) \frac{(k+j)!\,(k(n-k)+\binom{k}{2}-k-j)!}{(k(n-k)+\binom{k}{2}+1)!} \\ &= \frac{C(k,k+j)\,(k+j)!}{k!} \times B(k,k+j) \end{aligned} \tag{2.6}$$

where, provided $k \le n$ and $k+j \le \binom{k}{2}$ (as in our case),

$$\begin{aligned} B(k,k+j) &= \frac{n!}{(n-k)!} \cdot \frac{(k(n-k)+\binom{k}{2}-k-j)!}{(k(n-k)+\binom{k}{2}+1)!} \\ &= \frac{1}{n^{j+1}k^{k+j+1}} \frac{\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right)}{\prod_{i=0}^{k+j}\left(1-\frac{k+1}{2n}-\frac{i-1}{kn}\right)} \\ &= \frac{1}{n^{j+1}k^{k+j+1}} \exp\left\{ \sum_{m=1}^{\infty} \frac{1}{mn^m} \left(\sum_{i=0}^{k+j} \left(\frac{k+1}{2} + \frac{i-1}{k} \right)^m - \sum_{i=0}^{k-1} i^m \right) \right\} \\ &= \frac{1}{n^{j+1}k^{k+j+1}} \exp\left\{ \sum_{m=1}^{\infty} \frac{t_m(k,j)}{mn^m} \right\}. \end{aligned} \tag{2.7}$$

Observe that as $\sum_{i=1}^{a} i^m \geq \int_0^a x^m\,dx$, for $\ell = k+j$ we have

$$t_m(k,j) = \sum_{i=0}^{\ell}\left(\frac{k+1}{2}+\frac{i-1}{k}\right)^m - \sum_{i=0}^{k-1} i^m \leq (\ell+1)\left(\frac{k+1}{2}+\frac{\ell-1}{k}\right)^m - \frac{(k-1)^{m+1}}{m+1}. \qquad (2.8)$$

This implies that, as is easily verified,

$$t_m(k,j) \leq 0 \text{ if } m \geq 2 \text{ and } j \in \{0,-1\} \text{ and } k \geq 100. \qquad (2.9)$$

Case (a): $1 \leq k \leq n$, $j=-1$ (Tree components).

Now we have by (2.7)

$$B(k,k-1) = \frac{1}{k^k}\exp\left\{\frac{1}{n}\sum_{i=0}^{k-1}\left(\frac{k+1}{2}+\frac{i-1}{k}\right) - \frac{1}{n}\sum_{i=0}^{k-1} i + \frac{1}{2n^2}\sum_{i=0}^{k-1}\left(\frac{k+1}{2}+\frac{i-1}{k}\right)^2 - \frac{1}{2n^2}\sum_{i=0}^{k-1} i^2 + \xi\right\}$$

where, using (2.9),

$$|\xi| \leq \sum_{m=3}^{\infty}\frac{10^{2m+1}}{mn^m} = O(n^{-3}) \qquad 1 \leq k \leq 100, \qquad (2.10)$$

$$0 \geq \xi \geq -\sum_{m=3}^{\infty}\frac{k^{m+1}}{m(m+1)n^m} \geq -\frac{k^4}{n^3} \qquad k > 100, \qquad (2.11)$$

and hence for all $k \leq n$,

$$\xi = O(k^4/n^3). \qquad (2.12)$$

This implies, after some calculation, that, for $1 \leq k \leq n$,

$$B(k,k-1) = \frac{1}{k^k}\exp\left\{\frac{3(k-1)}{2n} - \frac{k^3}{24n^2} + O\left(\frac{k^2}{n^2}+\frac{k^4}{n^3}\right)\right\}$$

and then, by (2.6),

$$\begin{aligned}\sum_{k=1}^{n^{0.7}} A(k,k-1) &= \sum_{k=1}^{n^{0.7}}\frac{k^{k-2}}{k}\cdot B(k,k-1)\\ &= \sum_{k=1}^{n^{0.7}}\frac{1}{k^3}\exp\left\{\frac{3(k-1)}{2n} - \frac{k^3}{24n^2} + O\left(\frac{k^2}{n^2}+\frac{k^4}{n^3}\right)\right\}\\ &= \sum_{k=1}^{n^{0.7}}\frac{e^{-k^3/24n^2}}{k^3}\left(1+\frac{3(k-1)}{2n}+O\left(\frac{k^2}{n^2}+\frac{k^4}{n^3}\right)\right).\end{aligned}$$

Now, by simple estimates,

$$\sum_{k=1}^{n^{0.7}}\frac{e^{-k^3/24n^2}}{k^3}\times O\left(\frac{k^2}{n^2}+\frac{k^4}{n^3}\right) = O(n^{-5/3}) \qquad (2.13)$$

and

$$\sum_{k=1}^{n^{0.7}} \frac{(1-e^{-k^3/24n^2})}{k^3}\left(1+\frac{3(k-1)}{2n}\right) = o(n^{-4/3}) + \sum_{k=n^{2/3}/\ln n}^{n^{2/3}\ln n} \frac{(1-e^{-k^3/24n^2})}{k^3}$$
$$= o(n^{-4/3}) + \frac{1}{n^{4/3}}\int_{x=0}^{\infty} x^{-3}(1-e^{-x^3/24})\,dx. \tag{2.14}$$

Thus

$$\sum_{k=1}^{n^{0.7}} A(k,k-1)$$
$$= \sum_{k=1}^{n^{0.7}} \frac{1}{k^3} + \frac{1}{n}\sum_{k=1}^{n^{0.7}} \frac{3(k-1)}{2k^3} - \frac{1}{n^{4/3}}\int_{x=0}^{\infty} x^{-3}(1-e^{-x^3/24})\,dx + o(n^{-4/3})$$
$$= \zeta(3) + O(n^{-1.4}) + \frac{3(\zeta(2)-\zeta(3))}{2n} + O(n^{-1.7}) - \frac{1}{n^{4/3}}\int_{x=0}^{\infty} x^{-3}(1-e^{-x^3/24})\,dx + o(n^{-4/3})$$
$$= \zeta(3) + \frac{3(\zeta(2)-\zeta(3))}{2n} - \frac{1}{n^{4/3}}\int_{x=0}^{\infty} x^{-3}(1-e^{-x^3/24})\,dx + o(n^{-4/3}). \tag{2.15}$$

When $k \geq n^{0.7}$ we have from (2.7) and (2.9) that

$$B(k,k-1) \leq \frac{1}{k^k}\exp\left(\frac{1}{n}\sum_{i=0}^{k-1}\left(\frac{k+1}{2}+\frac{i-1}{k}\right) - \frac{1}{n}\sum_{i=0}^{k-1} i\right) = \frac{1}{k^k}\exp\left\{\frac{3(k-1)}{2n}\right\} \leq \frac{e^{3/2}}{k^k}.$$

This implies that $A(k,k-1) \leq k^{-3}e^{3/2}$. This gives

$$\sum_{k>n^{0.7}} A(k,k-1) \leq \sum_{k>n^{0.7}} \frac{e^{3/2}}{k^3} = O(n^{-1.4}) = o(n^{-4/3}).$$

Together with (2.15), this verifies (a).

Case (b): $1 \leq k \leq n$, $j = 0$ (Unicyclic components).

Rényi [24] proved (see e.g. Bollobás [3, Theorem 5.18]) that, cf. the more general (2.3) above,

$$C(k,k) = \frac{(k-1)!}{2}\sum_{l=0}^{k-3}\frac{k^l}{l!} \sim \sqrt{\frac{\pi}{8}}k^{k-1/2}. \tag{2.16}$$

Now for $1 \leq k \leq n$ we have by (2.7)

$$B(k,k) = \frac{1}{nk^{k+1}}\exp\left\{\frac{1}{n}\sum_{i=0}^{k}\left(\frac{k+1}{2}+\frac{i-1}{k}\right) - \frac{1}{n}\sum_{i=0}^{k-1} i \right.$$
$$\left. + \frac{1}{2n^2}\sum_{i=0}^{k}\left(\frac{k+1}{2}+\frac{i-1}{k}\right)^2 - \frac{1}{2n^2}\sum_{i=0}^{k-1} i^2 + \xi\right\}$$

where ξ satisfies (2.10)–(2.12). Thus, after some calculation,

$$B(k,k) = \frac{1}{k^{k+1}n}\exp\left\{\frac{2k}{n} - \frac{1}{kn} - \frac{k^3}{24n^2} + O\left(\frac{k^2}{n^2}+\frac{k^4}{n^3}\right)\right\}$$

and then

$$\sum_{k=3}^{n^{0.7}} A(k,k) = \frac{1}{n}\sum_{k=3}^{n^{0.7}} \frac{C(k,k)}{k^{k+1}} \exp\left\{-\frac{k^3}{24n^2} + O\left(\frac{k}{n}+\frac{k^4}{n^3}\right)\right\}$$

$$= \frac{1}{n}\sum_{k=3}^{n^{0.7}} \frac{C(k,k)e^{-k^3/24n^2}}{k^{k+1}} \left\{1 + O\left(\frac{k}{n}+\frac{k^4}{n^3}\right)\right\}. \tag{2.17}$$

Now (2.16) implies

$$\frac{1}{n}\sum_{k=3}^{n^{0.7}} \frac{C(k,k)e^{-k^3/24n^2}}{k^{k+1}} \times O\left(\frac{k}{n}+\frac{k^4}{n^3}\right) = O(n^{-5/3}) \tag{2.18}$$

and

$$\frac{1}{n}\sum_{k=3}^{n^{0.7}} \frac{C(k,k)(1-e^{-k^3/24n^2})}{k^{k+1}} = o(n^{-4/3}) + \frac{1}{n}\sum_{k=n^{2/3}/\ln n}^{n^{2/3}\ln n} \frac{C(k,k)(1-e^{-k^3/24n^2})}{k^{k+1}}$$

$$= o(n^{-4/3}) + \frac{\sqrt{\pi/8}}{n^{4/3}}\int_{x=0}^{\infty} x^{-3/2}(1-e^{-x^3/24})\,dx. \tag{2.19}$$

It follows from (2.17), (2.18) and (2.19) that

$$\sum_{k=3}^{n^{0.7}} A(k,k) = \frac{1}{n}\sum_{k=3}^{\infty} \frac{C(k,k)}{k^{k+1}} - \frac{\sqrt{\pi/8}}{n^{4/3}}\int_{x=0}^{\infty} x^{-3/2}(1-e^{-x^3/24})\,dx + o(n^{-4/3}). \tag{2.20}$$

For $k > n^{0.7}$ we observe that $t_1(k,0) \le 2k$ in (2.8) and $t_m(k,0) \le 0$ for $m \ge 2$ and so

$$B(k,k) \le \frac{e^2}{k^{k+1}n}$$

and so

$$A(k,k) \le e^2\frac{C(k,k)}{k^{k+1}n} = O\left(\frac{1}{k^{3/2}n}\right).$$

It follows from this that

$$\sum_{k=n^{0.7}}^{n} A(k,k) = O(n^{-1.35}) = o(n^{-4/3}). \tag{2.21}$$

We are almost done, we need to simplify the sum $\sum_{k=3}^{\infty}\frac{C(k,k)}{k^{k+1}}$.

Now, by (2.16),

$$\sum_{k=3}^{\infty}\frac{2C(k,k)}{k^{k+1}} = \sum_{k=3}^{\infty}\frac{(k-1)!}{k^{k+1}}\sum_{i=0}^{k-3}\frac{k^i}{i!} = \sum_{i=0}^{\infty}\sum_{k=i+3}^{\infty}\frac{k^i}{k^{k+1}}\frac{(k-1)!}{i!}. \tag{2.22}$$

In the last double sum, let us also add the terms with $k = i+2$, $k = i+1$ and $k = i \ge 1$. The terms with $k = i+2$ add up to

$$\sum_{k=2}^{\infty}\frac{k^{k-2}}{k^{k+1}}\frac{(k-1)!}{(k-2)!} = \sum_{k=2}^{\infty}\frac{k-1}{k^3} = \sum_{k=1}^{\infty}\frac{k-1}{k^3} = \zeta(2) - \zeta(3).$$

The terms with $k = i + 1$ add up to

$$\sum_{k=1}^{\infty} \frac{k^{k-1}}{k^{k+1}} \frac{(k-1)!}{(k-1)!} = \sum_{k=1}^{\infty} \frac{1}{k^2} = \zeta(2).$$

The terms with $k = i \geq 1$ add up to

$$\sum_{k=1}^{\infty} \frac{k^{k}}{k^{k+1}} \frac{(k-1)!}{k!} = \sum_{k=1}^{\infty} \frac{1}{k^2} = \zeta(2).$$

Consequently, (2.22) yields

$$\sum_{k=3}^{\infty} \frac{2C(k,k)}{k^{k+1}} = \zeta(3) - 3\zeta(2) + \sum_{k=1}^{\infty}\sum_{i=0}^{k} \frac{k^i}{k^{k+1}} \frac{(k-1)!}{i!} = \zeta(3) - 3\zeta(2) + \sum_{k=1}^{\infty}\sum_{i=0}^{k} \frac{k!}{i!} k^{i-k-2}. \tag{2.23}$$

We transform the sum further:

$$\begin{aligned}
\sum_{k=1}^{\infty}\sum_{i=0}^{k} \frac{k!}{i!} k^{i-k-2} &= \sum_{k=1}^{\infty}\sum_{i=0}^{k} \binom{k}{i} (k-i)!\, k^{i-k-2} \\
&= \sum_{k=1}^{\infty}\sum_{i=0}^{k} \binom{k}{i} k^{-1} \int_{x=0}^{\infty} x^{k-i} e^{-kx}\, dx \\
&= \int_{x=0}^{\infty} \sum_{k=1}^{\infty}\sum_{i=0}^{k} k^{-1} \binom{k}{i} x^{k-i} e^{-kx}\, dx \\
&= \int_{x=0}^{\infty} \sum_{k=1}^{\infty} k^{-1} (1+x)^k e^{-kx}\, dx \\
&= \int_{x=0}^{\infty} -\log\bigl(1 - (1+x)e^{-x}\bigr)\, dx
\end{aligned}$$

Consequently, (2.23) yields

$$2\sum_{k=3}^{\infty} \frac{C(k,k)}{k^{k+1}} = \zeta(3) - 3\zeta(2) - \int_{x=0}^{\infty} \log\bigl(1 - (1+x)e^{-x}\bigr)\, dx. \tag{2.24}$$

Together with (2.20) and (2.21), this verifies (b).

Case (c): $1 \leq k \leq n$, $j \geq 1$ (Complex components).

Let

$$\kappa_{\mathsf{c}}(p) = \kappa_{\mathsf{c},n}(p) := \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \kappa(k,j,p), \tag{2.25}$$

i.e., the number of complex components in $G_{n,p}$, and

$$f_n(p) = \mathbb{E}\, \kappa_{\mathsf{c}}(p) = \sum_{k\geq 1}\sum_{j\geq 1} \mathbb{E}\, \kappa(k,j,p), \tag{2.26}$$

the expected number of complex components in $G_{n,p}$. The contribution to (2.1) from the complex components is thus $\int_{p=0}^{1} f_n(p)\, dp$. We make a change of variables and let

$$p = n^{-1} + \lambda n^{-4/3}, \tag{2.27}$$

which means that we focus on the critical window. We will assume this relation between p and λ in the rest of the proof. We thus define $\bar{f}_n(\lambda) = f_n(p) = f_n(n^{-1} + \lambda n^{-4/3})$, and obtain the contribution, letting $\mathbf{1}\{\dots\}$ denote the indicator of an event,

$$\begin{aligned}\int_{p=0}^{1} f_n(p)\,dp &= 1 - \frac{1}{n} + \int_{p=0}^{1} \bigl(f_n(p) - \mathbf{1}\{p > 1/n\}\bigr)\,dp \\ &= 1 - \frac{1}{n} + n^{-4/3}\int_{\lambda=-n^{1/3}}^{n^{4/3}-n^{1/3}} \bigl(\bar{f}_n(\lambda) - \mathbf{1}\{\lambda > 0\}\bigr)\,d\lambda. \qquad (2.28)\end{aligned}$$

We begin by showing that the integrand in the final integral converges pointwise. We define, cf. (2.5),

$$\psi_2(t) = \sum_{\ell=2}^{\infty} w_\ell t^l = \psi(t) - 1 - \sqrt{\pi/8}\,t, \qquad (2.29)$$

and

$$F(x,\lambda) = \frac{1}{6}x^3 - \frac{1}{2}x^2\lambda + \frac{1}{2}x\lambda^2 = \frac{x}{2}\Bigl(\lambda - \frac{x}{2}\Bigr)^2 + \frac{1}{24}x^3. \qquad (2.30)$$

Sublemma 2.2. *For any fixed $\lambda \in (-\infty,\infty)$, as $n \to \infty$,*

$$\bar{f}_n(\lambda) \to f(\lambda) = \frac{1}{\sqrt{2\pi}}\int_{x=0}^{\infty} \psi_2(x^{3/2})e^{-F(x,\lambda)}x^{-5/2}\,dx. \qquad (2.31)$$

Proof We note first that the integral in (2.31) is convergent; for small x we have $\psi_2(x) = O(x^2)$ and for large x we have $\psi_2(x) = O\bigl(x^2 e^{x^2/24}\bigr)$ by (1.5) while $e^{-F(x,\lambda)} \le e^{-x^3/6+\lambda x^2/2} = O(e^{-x^3/7})$ by (2.30), remember that λ is fixed in the integral.

We convert the sum over k in (2.26) to an integral by setting $k = \lceil xn^{2/3}\rceil$. Thus

$$\bar{f}_n(\lambda) = f_n(p) = \int_{x=0}^{\infty} \sum_{j\ge 1} \mathbb{E}\,\kappa\bigl(\lceil xn^{2/3}\rceil, j, p\bigr)n^{2/3}\,dx. \qquad (2.32)$$

For any fixed λ and fixed $x > 0$, $j \ge 1$, and $p = n^{-1} + \lambda n^{-4/3}$ and $k = \lceil xn^{2/3}\rceil$ as above, we have as $n \to \infty$ by (2.2) and (2.3) and standard calculations, see e.g. [17, Section 4] or [1, Section 11.10] for further details,

$$\begin{aligned}\mathbb{E}\,\kappa(k,j,p) &\sim \frac{n^k}{k!}\exp\Bigl(-\frac{k^2}{2n} - \frac{k^3}{6n^2}\Bigr)C(k,k+j)n^{-k-j}\bigl(1+\lambda n^{-1/3}\bigr)^k \exp\bigl(-p(nk - k^2/2)\bigr) \\ &\sim n^{-j}\frac{C(k,k+j)}{k!}\exp\bigl(-k - F(kn^{-2/3},\lambda)\bigr) \\ &\sim (2\pi)^{-1/2}w_{j+1}k^{-1}\Bigl(\frac{k^{3/2}}{n}\Bigr)^j e^{-F(kn^{-2/3},\lambda)} \\ &\sim n^{-2/3}(2\pi)^{-1/2}w_{j+1}x^{3j/2-1}e^{-F(x,\lambda)}.\end{aligned}$$

Thus, as $n \to \infty$,

$$n^{2/3}\,\mathbb{E}\,\kappa(\lceil xn^{2/3}\rceil, j, p) \to (2\pi)^{-1/2}w_{j+1}x^{3j/2-1}e^{-F(x,\lambda)}. \qquad (2.33)$$

Moreover, Bollobás [3, Theorem 5.20] has shown the uniform bound

$$C(k,k+j) \le \Bigl(\frac{C_1}{j}\Bigr)^{j/2} k^{k+(3j-1)/2} \qquad (2.34)$$

for some constant C_1 and all $k, j \geq 1$. Let $A \geq 1$ be a constant, and consider first only components of size $k \leq An^{2/3}$. For such k, all $j \geq 1$ and $p = n^{-1} + \lambda n^{-4/3}$, (2.2) and (2.34) yield by calculations similar to those above,

$$\begin{aligned}
\mathbb{E}\,\kappa(k,j,p) &\leq C_2 \frac{n^k}{k!} \exp\Bigl(-\frac{k^2}{2n}\Bigr) C(k,k+j) n^{-k-j} \bigl(1+\lambda n^{-1/3}\bigr)^{k+j} \exp\bigl(-p(nk - k^2/2 - j)\bigr) \\
&\leq C_3 n^{-j} \frac{C(k,k+j)}{k!} e^{-k+j\times o(1)} \\
&\leq C_3 n^{-j} \Bigl(\frac{2C_1}{j}\Bigr)^{j/2} k^{3j/2-1}
\end{aligned}$$

(with C_3 possibly depending on A) and thus

$$n^{2/3}\,\mathbb{E}\,\kappa(k,j,p) \leq C_3 \left(\frac{C_4 A^{3/2}}{j}\right)^{j/2}.$$

The sum over j of the right-hand side converges, and thus (2.33) and dominated convergence yield, recalling (2.29),

$$\int_{x=0}^{A} \sum_{j\geq 1} \mathbb{E}\,\kappa\bigl(\lceil xn^{2/3}\rceil, j, p\bigr) n^{2/3}\,dx \to \frac{1}{\sqrt{2\pi}} \int_{x=0}^{A} \psi_2(x^{3/2}) e^{-F(x,\lambda)} x^{-5/2}\,dx. \tag{2.35}$$

For $k > An^{2/3}$ we use the fact shown in [17, (6.6)] that the expected number of vertices in tree components of size at most $n^{2/3}$ is $n - O(n^{2/3})$; consequently, the expected number of vertices in all components (complex or not) of size larger than $n^{2/3}$ is $O(n^{2/3})$, and the expected number of components larger than $An^{2/3}$ is $\leq C_5/A$. The left-hand side of (2.35) thus converges uniformly to the right-hand side of (2.32) as $n \to \infty$, and the result (2.31) follows from (2.35) by letting $A \to \infty$. □

The next step is to use dominated convergence in (2.28). For this we use the following estimates. For convenience, we let $\kappa_{\mathsf{c}}(n^{-1} + \lambda n^{-4/3})$ and its expectation $\bar{f}_n(\lambda)$ be defined for all real λ, by trivially defining $\kappa_{\mathsf{c}}(p) = \kappa_{\mathsf{c}}(0) = 0$ for $p < 0$ and $\kappa_{\mathsf{c}}(p) = \kappa_{\mathsf{c}}(1) = 1$ for $p > 1$.

Sublemma 2.3. *There exist integrable functions $g_1(\lambda), g_2(\lambda), g_3(\lambda)$, not depending on n, such that*

(i)
$$\bar{f}_n(\lambda) = \mathbb{E}\,\kappa_{\mathsf{c}}(n^{-1} + \lambda n^{-4/3}) \leq g_1(\lambda), \qquad \lambda \leq 0,$$

(ii)
$$\mathbb{P}\bigl(\kappa_{\mathsf{c}}(n^{-1} + \lambda n^{-4/3}) = 0\bigr) \leq g_2(\lambda), \qquad \lambda \geq 0,$$

(iii)
$$\bar{f}_n(\lambda) - 1 = \mathbb{E}\,\kappa_{\mathsf{c}}(n^{-1} + \lambda n^{-4/3}) - 1 \leq g_3(\lambda), \qquad \lambda \geq 0.$$

Proof We use the method in Janson [10]. We consider $G(n,p)$, $p \in [0,1]$, as a random graph process in the usual way: we regard p as time, edges are added as p grows from 0 to 1, and an edge e is added at a time T_e with a uniform distribution on $[0,1]$, with all T_e independent.

As $G(n,p)$ evolves, there are at first only tree components, but later unicyclic components and complex components appear as edges are added to the graph. If we consider the complex components only, a new complex component is created if a new edge is added to a unicyclic component, or if it joins two unicyclic components. (Note that these are the only possibilities; we do not regard the growth of an already existing complex component as creating a new complex component. Creation of a new complex component may happen one or several times. It is shown in [14] that it happens only once with probability converging to $5\pi/18$, but we will not need this.) As evolution continues, the complex components may grow by merging with trees or unicyclic components, and they may merge with each other, until at the end only one complex component remains, containing all vertices.

Let $\varphi_n(k,p)$ be the intensity of creation on new complex components of size k, i.e., the probability of creating a new complex component of size k in the interval $[p, p+dp]$ is $\varphi_n(k,p)\,dp$. (For $p<0$, $p>1$ or $k>n$, we set $\varphi_n(k,p)=0$.) Further, let $\Phi_n(p)=\sum_{k\ge1}\varphi_n(k,p)$, the intensity of creation of complex components regardless of size. We change variables as above and define also

$$\begin{aligned}\psi_n(x,\lambda)&=n^{-2/3}\varphi_n(\lceil xn^{2/3}\rceil, n^{-1}+\lambda n^{-4/3}),\\ \Psi_n(\lambda)&=n^{-4/3}\Phi_n(n^{-1}+\lambda n^{-4/3})=\int_{x=0}^{\infty}\psi_n(x,\lambda)\,dx.\end{aligned}$$

(The notation is not exactly as in [10], where the two ways of creating a complex component are treated separately, but the estimates are the same.)

We have

$$\varphi_n(k,p)=\binom{n}{k}\hat{C}(k)p^k(1-p)^{(n-k)k+\binom{k}{2}-k-1}$$

where $\hat{C}(k)$ is the number of ways to create a multicyclic component by either adding an edge to a unicyclic component on $[k]$ or adding an edge joining two unicyclic components whose vertex sets are complementary subsets of $[k]$. The first case contributes

$$C(k,k)\left(\binom{k}{2}-k\right)=O(k^{k+3/2})$$

to $\hat{C}(k)$ and the second

$$\frac12\sum_{i=3}^{k-3}\binom{k}{i}C(i,i)C(k-i,k-i)i(k-i)\le C_6\sum_{i=3}^{k-3}\binom{k}{i}e^i i!\,e^{k-i}(k-i)!\le C_6ke^kk!=O(k^{k+3/2});$$

hence

$$\hat{C}(k)=O(k^{k+3/2})=O\big(ke^kk!\big).$$

(Cf. the more precise [10, (2.30)].) The intensity $\psi_n(x,\lambda)$ is bounded in [10, (2.12)–(2.19)] by calculations similar to those in the proof of Sublemma 2.2. (In these bounds, and our versions below, $\delta,\delta_1,\dots$ are some positive constants.)

We use the results of [10] with some small modifications: Equation (2.12) of [10] shows (together with the comments after it) that

$$\psi_n(x,\lambda)\le C_7xe^{-\delta x^3-\delta x\lambda^2}\qquad\text{for } k\le\delta_1 n \text{ and } -n^{1/3}\le\lambda\le\delta_2 n^{1/3}.$$

Then one line before (2.15) of [10] proves that

$$\psi_n(x,\lambda) \le C_8 x e^{-\delta x^3 - \delta_3 x\lambda n^{1/3}/3} \qquad \text{for } k \le \delta_3 n \text{ and } \lambda \ge \delta_2 n^{1/3}.$$

Because $\lambda \le n^{4/3}$ always, it is legitimate to replace $-\delta_3 x\lambda n^{1/3}/3$ by $-\delta_3 x\lambda^{5/4}$ to give

$$\psi_n(x,\lambda) \le C_8 x e^{-\delta x^3 - \delta_3 x\lambda^{5/4}/3} \qquad \text{for } k \le \delta_3 n \text{ and } \lambda \ge \delta_2 n^{1/3}.$$

Then (2.17) of [10] proves that

$$\psi_n(x,\lambda) \le C_9 n e^{-2\delta_5 n} \qquad \text{for } \min\{\delta_1,\delta_3\}\, n \le k \le n.$$

We replace this by, using $\min\{\delta_1,\delta_3\}\, n^{1/3} \le x \le n^{1/3}$ and $\lambda \le n^{4/3}$,

$$\psi_n(x,\lambda) \le C_{10} x e^{-\delta_5 x^3}(1+\lambda^4)^{-1}.$$

We therefore have, for all x and λ (recalling that $\psi_n(x,\lambda)=0$ if $x > n^{1/3}$, $\lambda < -n^{1/3}$ or $\lambda > n^{4/3}$),

$$0 \le \psi_n(x,\lambda) \le g(x,\lambda) = C_7 x e^{-\delta x^3 - \delta x\lambda^2} + C_8 x e^{-\delta x^3 - \delta_3 x|\lambda|^{5/4}/3} + C_{10} x e^{-\delta_5 x^3}(1+\lambda^4)^{-1}. \tag{2.36}$$

Integrating we find

$$\Psi(\lambda) \le \int_{x=0}^{\infty} g(x,\lambda)\,dx \le \frac{C_{11}}{1+|\lambda|^{5/2}}. \tag{2.37}$$

The number of complex components at any time is at most the number of complex components that have been created so far. Taking expectations we thus obtain, using (2.37),

$$\bar{f}_n(\lambda) = \mathbb{E}\,\kappa_{\mathsf{c}}(n^{-1}+\lambda n^{-4/3}) \le \int_{\mu=-\infty}^{\lambda} \Psi(\mu)\,d\mu \le \int_{\mu=-\infty}^{\lambda} \frac{C_{11}}{1+|\mu|^{5/2}}\,d\mu. \tag{2.38}$$

This verifies (i), with $g_1(\lambda) = C_{12}(1+|\lambda|^{3/2})^{-1}$ for $\lambda \le 0$.

Similarly, if there is no complex component at some time, at least one complex component has to be created later. Thus,

$$\mathbb{P}\big(\kappa_{\mathsf{c}}(n^{-1}+\lambda n^{-4/3})=0\big) \le \int_{\mu=\lambda}^{\infty} \Psi(\mu)\,d\mu \le \int_{\mu=\lambda}^{\infty} \frac{C_{11}}{1+|\mu|^{5/2}}\,d\mu, \tag{2.39}$$

which verifies (ii) with $g_2(\lambda) = C_{13}(1+\lambda^{3/2})^{-1}$ for $\lambda \ge 0$.

For (iii), let $Y(p) = \binom{\kappa_{\mathsf{c}}(p)}{2}$ be the number of pairs of complex components in $G_{n,p}$. Since $\kappa_{\mathsf{c}}(p)-1 \le Y(p)$, it suffices to estimate $\mathbb{E}\,Y(p)$.

If there is a pair of complex components in $G_{n,p}$, then these components have been created at some times p_1 and p_2 with $p_1 \le p_2 \le p$. The intensity of this happening, with sizes $k_1 = \lceil x_1 n^{2/3}\rceil$ and $k_2 = \lceil x_2 n^{2/3}\rceil$ of the components at the moments of their creations, is bounded in [10, (2.24)–(2.26)] by (using modifications as above, and g is defined in (2.36)),

$$C_{14}\, g(x_1,\lambda_1) g(x_2,\lambda_2)\,d\lambda_1\,d\lambda_2\,dx_1\,dx_2.$$

Moreover, if the two components still are distinct components in $G_{n,p}$, then, at least (ignoring further conditions from the growth of the components), the original vertex sets of sizes k_1 and k_2 are not connected by any edge in the time interval $[p_2, p]$; the (conditional) probability of this is

$$\Bigl(1-\frac{p-p_2}{1-p_2}\Bigr)^{k_1k_2} \le (1-(p-p_2))^{k_1k_2} \le e^{-k_1k_2(p-p_2)} \le e^{-x_1x_2(\lambda-\lambda_2)}.$$

Consequently,

$$\begin{aligned}
\bar f_n(\lambda) - 1 &\le \mathbb{E}\, Y(n^{-1}+\lambda n^{-4/3}) \\
&\le g_3(\lambda) = \int_{\lambda_1=-\infty}^{\lambda}\int_{\lambda_2=\lambda_1}^{\lambda}\int_{x_1=0}^{\infty}\int_{x_2=0}^{\infty} C_{14} g(x_1,\lambda_1)g(x_2,\lambda_2)e^{-x_1x_2(\lambda-\lambda_2)}\,d\lambda_1\,d\lambda_2\,dx_1\,dx_2.
\end{aligned}$$

This yields (iii), but it remains to verify that $\int_{\lambda=0}^{\infty} g_3(\lambda)\,d\lambda<\infty$. Indeed, by Fubini and (2.36),

$$\begin{aligned}
&\int_{\lambda=-\infty}^{\infty} g_3(\lambda)\,d\lambda \\
&\quad= \int_{\lambda_1=-\infty}^{\infty}\int_{\lambda_2=\lambda_1}^{\infty}\int_{x_1=0}^{\infty}\int_{x_2=0}^{\infty} C_{14} g(x_1,\lambda_1)g(x_2,\lambda_2)\int_{\lambda=\lambda_2}^{\infty} e^{-x_1x_2(\lambda-\lambda_2)}\,d\lambda\,d\lambda_1\,d\lambda_2\,dx_1\,dx_2 \\
&\quad= \int_{\lambda_1=-\infty}^{\infty}\int_{\lambda_2=\lambda_1}^{\infty}\int_{x_1=0}^{\infty}\int_{x_2=0}^{\infty} C_{14}\frac{g(x_1,\lambda_1)g(x_2,\lambda_2)}{x_1x_2}\,d\lambda_1\,d\lambda_2\,dx_1\,dx_2 \\
&\quad\le C_{14}\Bigl(\int_{\lambda=-\infty}^{\infty}\int_{x=0}^{\infty}\frac{g(x,\lambda)}{x}\,d\lambda\,dx\Bigr)^2 < \infty.
\end{aligned}$$

□

Sublemma 2.3(ii) implies that $1-\bar f_n(\lambda)\le g_2(\lambda)$ for $\lambda\ge 0$, and thus Sublemma 2.3 yields

$$\bigl|\bar f_n(\lambda) - \mathbf{1}\{\lambda>0\}\bigr| \le \begin{cases} g_1(\lambda), & \lambda\le 0,\\ g_2(\lambda)+g_3(\lambda), & \lambda>0.\end{cases}$$

This justifies using dominated convergence in the integral in (2.28), and Sublemma 2.2 implies

$$\int_{\lambda=-n^{1/3}}^{n^{4/3}-n^{1/3}} \bigl(\bar f_n(\lambda)-\mathbf{1}\{\lambda>0\}\bigr)\,d\lambda \to c_{2c} = \int_{\lambda=-\infty}^{\infty}\bigl(f(\lambda)-\mathbf{1}\{\lambda>0\}\bigr)\,d\lambda. \tag{2.40}$$

Hence (2.28) yields

$$\int_{p=0}^{1} f_n(p)\,dp = 1-\frac{1}{n}+c_{2c}n^{-4/3}+o(n^{-4/3}), \tag{2.41}$$

which is the sought result except for the expression for c_{2c}.

We transform the expression for c_{2c} in (2.40) by first writing it as

$$\begin{aligned}
c_{2c} &= \lim_{A\to\infty}\Bigl(-A+\int_{\lambda=-\infty}^{A} f(\lambda)\,d\lambda\Bigr) \\
&= \lim_{A\to\infty}\Bigl(-A+\frac{1}{\sqrt{2\pi}}\int_{\lambda=-\infty}^{A}\int_{x=0}^{\infty}\psi_2(x^{3/2})e^{-F(x,\lambda)}x^{-5/2}\,dx\,d\lambda\Bigr).
\end{aligned} \tag{2.42}$$

By (2.29) we have $\psi_2(t) = O(t^2)$ for small t, which together with (1.7) shows that

$$\psi_2(t) = O\big(t^2 e^{t^2/24}\big), \qquad t \geq 0$$

and thus by (2.30), for all $x > 0$ and $\lambda \in (-\infty, \infty)$,

$$\psi_2(x^{3/2})e^{-F(x,\lambda)} \leq C_{15} x^3 e^{-x(\lambda - x/2)^2/2}.$$

Hence, for $A > 0$, with the substitutions $x = 2A + s$ and $\lambda = A - t$,

$$\begin{aligned}
\int_{x>2A}\int_{\lambda<A} \psi_2(x^{3/2})e^{-F(x,\lambda)}x^{-5/2}\,dx\,d\lambda &\leq C_{15}\int_{x>2A}\int_{\lambda<A} e^{-x(\lambda-x/2)^2/2}x^{1/2}\,dx\,d\lambda \\
&= C_{15}\int_{s>0}\int_{t>0} e^{-(2A+s)(t+s/2)^2/2}(2A+s)^{1/2}\,dt\,ds \\
&\leq C_{15}\int_{s>0}\int_{t>0} e^{-(2A+s)(t^2/2+s^2/8)}(2A+s)^{1/2}\,dt\,ds \\
&= C_{16}\int_{s>0} e^{-(2A+s)s^2/8}\,ds \leq C_{17}A^{-1/2}.
\end{aligned}$$

Similar estimates show also

$$\int_{x<2A}\int_{\lambda>A} \psi_2(x^{3/2})e^{-F(x,\lambda)}x^{-5/2}\,dx\,d\lambda \leq C_{18}\int_{s=0}^{2A} e^{-(2A-s)s^2/8}\,ds \leq C_{19}A^{-1/2}.$$

Consequently, we can subtract and add these integrals to (2.42), yielding

$$c_{2c} = \lim_{A\to\infty}\left(-A + \frac{1}{\sqrt{2\pi}}\int_{\lambda=-\infty}^{\infty}\int_{x=0}^{2A} \psi_2(x^{3/2})e^{-F(x,\lambda)}x^{-5/2}\,dx\,d\lambda\right). \tag{2.43}$$

It follows from (2.30) that

$$\int_{\lambda=-\infty}^{\infty} e^{-F(x,\lambda)}\,d\lambda = e^{-x^3/24}\int_{\lambda=-\infty}^{\infty} e^{-x(\lambda-x/2)^2/2}\,d\lambda = e^{-x^3/24}\sqrt{2\pi/x}. \tag{2.44}$$

Hence (2.43) yields by Fubini

$$c_{2c} = \lim_{A\to\infty}\left(-A + \int_{x=0}^{2A}\psi_2(x^{3/2})e^{-x^3/24}x^{-3}\,dx\right) = \int_{x=0}^{\infty}\Big(x^{-3}\psi_2(x^{3/2})e^{-x^3/24} - \frac{1}{2}\Big)\,dx. \tag{2.45}$$

This completes the proof of Lemma 2.1 and the proof of Theorem 1. ☐

3 Final remarks

Remark 3. We have shown that when the X_e are uniform $[0,1]$ then $\mathbb{E}(L_n)$ converges to $\zeta(3)$ with an error term of order $1/n$. The constant c_1 is positive and so for large n we have $\mathbb{E}(L_n) > \zeta(3)$. Fill and Steele [4] computed $\mathbb{E}(L_n)$ for $n \leq 8$. $\mathbb{E}(L_n)$ increased monotonically and it was natural to conjecture from this that $\mathbb{E}(L_n)$ increases monotonically for all n. However, since $\mathbb{E}(L_n)$ converges to $\zeta(3)$ from above, we now see that this turns out not to be true. Note, however, that $c_2 < 0$, and that $|c_2|$ is much larger than c_1. Thus we expect that $\mathbb{E}\,L_n > \zeta(3)$ only for very large n.

We have, if our numerical estimates are correct, $|c_2|/c_1 \approx 45$, so a naive guess, ignoring higher order terms, would be that $\mathbb{E}\,L_n > \zeta(3)$ for $n > 45^3 \approx 10^5$. We don't want to conjecture this, as we have no idea about the next term.

Remark 4. By (1.2), we obtain for $\mathbb{E}_{exp}(L_n)$ the same result as in Theorem 1 except that c_1 is increased by $\zeta(3)$ (while c_2 remains the same). This gives a somewhat simpler c_1, which suggests that this version might be slightly simpler to analyze. Note that the formula (2.1) holds for $\mathbb{E}_{exp}(L_n)$ if we replace $G_{n,p}$ by the multigraph where each pair of vertices is connected by a $\mathrm{Po}(t)$ number of edges, and integrate for $t \in (0, \infty)$. This suggests that it might be profitable to make a version of the argument below using these multigraphs, but we have not pursued this. (Cf. the use of multigraphs in [14].)

Acknowledgement: In an earlier version of this paper, we showed that $\mathbb{E}(L_n) = \zeta(3) + \frac{c_1 + o(1)}{n}$. Nick Read pointed us to his article [23] which suggested that the $o(1/n)$ term could be replaced by $(c + o(1))/n^{4/3}$. This encouraged us to go the extra mile and find the next term and prove Nick's conjecture.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, 3rd ed, Wiley, New York (2008).

[2] A. Beveridge, A. M. Frieze and C. J. H. McDiarmid, *Minimum length spanning trees in regular graphs*, Combinatorica 18 (1998) 311–333.

[3] B. Bollobás, *Random Graphs*, Second Edition, Cambridge University Press (2001).

[4] J. A. Fill and M. J. Steele, *Exact expectations of minimal spanning trees for graphs with random edge weights*, Stein's Method and Applications, Singapore Univ. Press, Singapore (2005) 169–180.

[5] A. Flaxman, *The lower tail of the random minimum spanning tree*, Electronic Journal of Combinatorics 14 (2007).

[6] A. M. Frieze, *On the value of a random minimum spanning tree problem*, Discrete Applied Mathematics 10 (1985) 47–56.

[7] A. M. Frieze and C. J. H. McDiarmid, *On random minimum length spanning trees*, Combinatorica 9 (1989) 363–374.

[8] A. M. Frieze, M. Ruszinkó and L. Thoma, *A note on random minimum length spanning trees*, Electronic Journal of Combinatorics 7 (2000) R41.

[9] D. Gamarnik, *The expected value of random minimal length spanning tree of a complete graph.* Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2005), ACM, New York (2005) 700–704.

[10] S. Janson, *Multicyclic components in a random graph process*, Random Structures and Algorithms 4 (1993, 71–84.

[11] S. Janson, *The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph*, Random Structures and Algorithms 7 (1995) 337–355.

[12] S. Janson, *Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas*, Probability Surveys, 3 (2007) 80–145.

[13] S. Janson and P. Chassaing, *The center of mass of the ISE and the Wiener index of trees.* Electronic Comm. Probability 9 (2004) paper 20, 178–187.

[14] S. Janson, D.E. Knuth, T. Łuczak and B. Pittel, *The birth of the giant component*, Random Structures and Algorithms 3 (1993) 233–358.

[15] S. Janson and G. Louchard, *Tail estimates for the Brownian excursion area and other Brownian areas*, Electronic Journal Probability 12 (2007) 1600–1632.

[16] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley, New York (2000).

[17] S. Janson and J. Spencer, *A point process describing the component sizes in the critical window of the random graph evolution*, Combinatorics Probability and Computing 16 (2007) 631–658.

[18] W. Li and X. Zhang, *On the difference of expected lengths of minimum spanning trees*, Combinatorics, Probability and Computing 18 (2009) 423–434.

[19] G. Louchard, *Kac's formula, Lévy's local time and Brownian excursion.* Journal of Applied Probability 21 (1984) 479–499.

[20] G. Louchard, *The Brownian excursion area: a numerical analysis.* Computational Mathematics and Applications 10 (1984) 413–417. Erratum: Computational Mathematics and Applications Part A 12 (1986) 375.

[21] *NIST Digital Library of Mathematical Functions.* `http://dlmf.nist.gov/`

[22] M. Penrose, *Random minimum spanning tree and percolation on the n-cube*, Random Structures and Algorithms 12 (1998) 63–82.

[23] N. Read, *Minimum spanning trees and random resistor networks in d dimensions*, Physics Revue. E 72, 036114 (2005).

[24] A. Rényi, *Some remarks on the theory of trees*, Publ. Math. Inst. Hungar. Acad. Sci. 4 (1959) 73–85.

[25] J. Spencer, *Enumerating graphs and Brownian motion.* Communications in Pure and Applied Mathematics 50 (1997) 291–294.

[26] M. J. Steele, *On Frieze's $\zeta(3)$ limit for lengths of minimal spanning trees*, Discrete Applied Mathematics 18 (1987) 99–103.

[27] J. Wästlund, *An easy proof of the $\zeta(2)$ limit in the random assignment problem*, Electronic Communications in Probability 14 (2009) 261–269.

[28] E. M. Wright, *The number of connected sparsely edged graphs.* Journal of Graph Theory 1 (1977) 317–330.

A Appendix: Estimation of c_2

The constant c_2 in Theorem 1 is the sum of the three coefficients for $n^{-4/3}$ in Lemma 2.1(a)–(c), which we denote by c_{2a}, c_{2b} and c_{2c}. By the change of variable $t = x^3/24$, and integration by parts (cf. [21, §5.9.5]), we obtain, as said in Remark 2,

$$c_{2a} = \frac{24^{-2/3}}{3} \int_{t=0}^{\infty} t^{-5/3}(e^{-t} - 1)\,dt = -\frac{1}{8}3^{-2/3}\Gamma\Bigl(\frac{1}{3}\Bigr) = -0.16098\ldots, \tag{A.1}$$

$$c_{2b} = \sqrt{\frac{\pi}{8}}\frac{24^{-1/6}}{3} \int_{t=0}^{\infty} t^{-7/6}(e^{-t} - 1)\,dt = -\frac{1}{2}3^{-1/6}\sqrt{\pi}\,\Gamma\Bigl(\frac{5}{6}\Bigr) = -0.83298\ldots. \tag{A.2}$$

The coefficient c_{2c} is given by an integral in Lemma 2.1, see also (2.45). To evaluate c_{2c}, we change variables by $x = y^{1/3}$ and use the definition (2.29) of ψ_2 to obtain

$$\begin{aligned} c_{2c} &= \frac{1}{3}\int_{y=0}^{\infty} \Bigl(y^{-1}\psi_2(y^{1/2}) - \frac{1}{2}e^{y/24}\Bigr)e^{-y/24}y^{-2/3}\,dy \\ &= \frac{1}{3}\int_{y=0}^{\infty}\sum_{k=1}^{\infty}\Bigl(w_{2k}y^{k-1} + w_{2k+1}y^{k-1/2} - \frac{y^{k-1}}{2\cdot 24^{k-1}(k-1)!}\Bigr)e^{-y/24}y^{-2/3}\,dy. \end{aligned} \tag{A.3}$$

We interchange the order of integration and summation, which is justified below, and obtain

$$\begin{aligned} c_{2c} &= \frac{1}{3}\sum_{k=1}^{\infty}\int_{y=0}^{\infty}\Bigl(w_{2k}y^{k-1} + w_{2k+1}y^{k-1/2} - \frac{y^{k-1}}{2\cdot 24^{k-1}(k-1)!}\Bigr)e^{-y/24}y^{-2/3}\,dy \\ &= \frac{24^{1/3}}{3}\sum_{k=1}^{\infty}\Bigl(w_{2k}24^{k-1}\Gamma(k-2/3) + w_{2k+1}24^{k-1/2}\Gamma(k-1/6) - \frac{\Gamma(k-2/3)}{2\,\Gamma(k)}\Bigr). \end{aligned} \tag{A.4}$$

We note that (2.4) and (1.4) yield, together with Stirling's formula, $w_\ell \sim 6\cdot 24^{-\ell/2}/\Gamma(\ell/2)$, which implies that

$$w_{2k}24^{k-1}\Gamma(k-2/3) \sim w_{2k+1}24^{k-1/2}\Gamma(k-1/6) \sim \tfrac{1}{4}k^{-2/3} \qquad \text{as } k\to\infty$$

so the three terms in the sum in (A.4) are all of order $k^{-2/3}$, showing that we cannot sum them separately. However, their leading terms cancel. A more precise calculation using (1.6) yields

$$\mathbb{E}\,\mathcal{B}_{\mathrm{ex}}^r = \sqrt{18}\,r\Bigl(\frac{r}{12e}\Bigr)^{r/2}\bigl(1 + O(r^{-1})\bigr), \qquad r > 0, \tag{A.5}$$

and thus by (2.4) and Stirling's formula,

$$w_\ell = \frac{3\sqrt{\ell}}{\sqrt{\pi}}\Bigl(\frac{e}{12\ell}\Bigr)^{\ell/2}\bigl(1 + O(\ell^{-1})\bigr) = \frac{6\cdot 24^{-\ell/2}}{\Gamma(\ell/2)}\bigl(1 + O(\ell^{-1})\bigr), \qquad \ell \ge 1. \tag{A.6}$$

Hence,

$$w_{2k}24^{k-1}\Gamma(k-2/3) = \tfrac{1}{4}k^{-2/3}\bigl(1 + O(k^{-1})\bigr), \qquad \text{as } k\to\infty, \tag{A.7}$$

and the same estimate holds for $w_{2k+1}24^{k-1/2}\Gamma(k-1/6)$, while $\Gamma(k-2/3)/\Gamma(k) = k^{-2/3}\bigl(1+O(k^{-1})\bigr)$. Consequently, the summand in (A.4) is $O(k^{-5/3})$.

The constants w_k can be computed by a recursion formula, see [28] and [12], and a numerical summation of the first 1000 terms in (A.4) yields -0.7331. It can be shown, using (1.6) with the further second order term given in [15] (which replaces $O(x^{-2})$ by $-\frac{1}{9}x^{-2}+O(x^{-4})$), that the terms in the sum in (A.4) are $\sim -\frac{1}{6}k^{-5/3}$, and using this to estimate the sum of the terms with $k > 1000$ yields the estimate $c_{2c} \approx -0.7355$ which together with (A.1)–(A.2) yields

$$c_2 \approx -1.7295. \tag{A.8}$$

The tail estimate is not rigorous. Replacing $O(x^{-4})$ by $\le Cx^{-4}$ for some estimate C is what is needed to make the tail estimate rigorous. Nevertheless, it seems unlikely that the estimate in (A.8) *is very far off.*

To justify the interchange of summation and integration above, it is by Fubini's theorem sufficient to verify that

$$\sum_{k=1}^{\infty} \int_{y=0}^{\infty} \left| w_{2k} y^{k-1} + w_{2k+1} y^{k-1/2} - \frac{y^{k-1}}{2 \cdot 24^{k-1}(k-1)!} \right| e^{-y/24} y^{-2/3} \, dy < \infty. \tag{A.9}$$

Indeed, we claim that the integral in (A.9) is $O(k^{-7/6})$. Using (A.7), its analogue for $2k+1$, and $\Gamma(k-2/3)/\Gamma(k) = k^{-2/3}(1+O(k^{-1}))$, it follows easily that the integral is, after another change of variable $t = y/24$,

$$\frac{24^{1/3}}{4} k^{-2/3} \int_{t=0}^{\infty} \left| \frac{t^{k-7/6}}{\Gamma(k-1/6)} - \frac{t^{k-5/3}}{\Gamma(k-2/3)} \right| e^{-t} \, dt + O\big(k^{-5/3}\big). \tag{A.10}$$

Let I_k denote the integral in (A.10). By the Cauchy–Schwarz inequality,

$$\begin{aligned}
I_k^2 &\le \int_{t=0}^{\infty} t^{k-1} e^{-t} \, dt \cdot \int_{t=0}^{\infty} \left(\frac{t^{k-7/6}}{\Gamma(k-1/6)} - \frac{t^{k-5/3}}{\Gamma(k-2/3)} \right)^2 t^{1-k} e^{-t} \, dt \\
&= \Gamma(k) \left(\frac{\Gamma(k-2/6)}{\Gamma(k-1/6)^2} - 2\frac{\Gamma(k-5/6)}{\Gamma(k-1/6)\Gamma(k-4/6)} + \frac{\Gamma(k-8/6)}{\Gamma(k-4/6)^2} \right) \\
&= O\big(k^{-1}\big).
\end{aligned}$$

Consequently, $I_k = O(k^{-1/2})$, which shows that (A.10) is $O(k^{-7/6})$, and thus (A.9) holds as claimed above.