第49卷第3期 中国科学技术大学学报 Vol. 49, No. 3
2019年3月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Mar. 2019

文章编号:0253-2778(2019)03-0173-09

An analytical approach to degree profile of a random tree

FENG Qunqiang, ZHENG Boze

(*Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei* 230026, *China*)

Abstract: The degree profile in a random planted plane tree was considered. For any $d \geqslant 1$, it was proven that under suitable normalization, the number of vertices of degree d in a random planted plane tree with n edges has asymptotic normality, as n goes to infinity. The asymptotic formulae for the expectation and variance of this random variable were also given. An analytical method was employed in the proof.

Key words: random tree; degree profile; saddle point method; Hurwitz theorem

CLC number: O211.4 **Document code**: A doi:10.3969/j.issn.0253-2778.2019.03.001

2010 Mathematics Subject Classification: 05C05;60C05;60F05

Citation: FENG Qunqiang, ZHENG Boze. An analytic approach to degree profile of a random tree[J]. Journal of University of Science and Technology of China, 2019,49(3):173-181.
冯群强,郑伯泽. 一种随机树度分布的解析方法[J]. 中国科学技术大学学报,2019,49(3):173-181.

一种随机树度分布的解析方法

冯群强,郑伯泽

(中国科学技术大学管理学院统计与金融系,安徽合肥 230026)

摘要:主要讨论了随机平面根树的度分布. 对任意 $d \geqslant 1$,证明了在含有 n 条边的随机平面根树中,当 $n \to \infty$ 时,度数为 d 的顶点数目在合适的正则化条件下具有渐近正态性,还给出了该数目期望和方差的渐近表达式. 在证明过程中主要使用了一种解析的方法.

关键词:随机树;度分布;鞍点法;Hurwitz 定理

0 Introduction

A planted plane tree is a plane tree whose root has only one child. In drawings, planted plane trees ascend from their roots. These trees have no labeling. It is well-known that the number of such trees with n edges is the Catalan number (see, for example, Ref. [1]):

$$C_n = \frac{1}{n}\binom{2n-2}{n-1} \tag{1}$$

The model of randomness in the growth of random planted plane trees induces a uniform distribution on the trees: All C_n planted plane trees with n edges are generated with equal probability.

Planted plane trees have been considered by many researchers. For example, The height problems of planted plane trees have been

Received: 2017-08-01; **Revised**: 2018-01-19

Foundation item: Supported by the National Natural Science Foundation of China (11771418).

Biography: FENG Qunqiang (corresponding author), male, born in 1979, PhD/ associate Prof. Research field: Probability and statistics. E-mail: fengqq@ustc.edu.cn

considered by Prodinger[2-4]; Gutjahr[5] studied the variance of level numbers in planted plane trees and some other families of trees; Chen et al.[6] gave a bijection between planted plane trees with elevated Dyck paths.

Vertices with degree 1 are referred to as leaves. But here, for convenience, we exclude the root of a planted plane tree as a leaf. For any integer $d \geqslant 1$, let $X_{n,d}$ be the number of (nonroot) vertices of degree d in a random planted plane tree with n edges. Note that a rooted tree with n edges has a root and n vertices, which implies that $\sum_{d=1}^{n-1} X_{n,d} = n$.

Degree profile (or degree distribution) is an important topic in the study of random trees and random networks. Via Pólya urn models, Mahmoud and Smythe[7], Janson[8] have discussed the asymptotic normality of the number of vertices with various degrees in random recursive trees. The degree profile of m-ary search trees has been studied in Ref. [9]. Zhang and Mahmoud[10] have also considered this problem both for Apollonian network which is a random network model, and for another random tree model, k-trees. Following their routes, in this paper we shall study the degree profile in random planted plane trees, i. e., the asymptotic distribution of $X_{n,d}$ for any fixed integer $d \geqslant 1$, as $n \to \infty$. Unlike their approaches, an analytical method is employed here.

To derive the asymptotic distribution of $X_{n,d}$, we shall define a bivariate generating function $B(x,y)$. In Section 1, we establish a functional equation for $B(x,y)$. Based on the functional equation, the asymptotic formula for the expectation of $X_{n,d}$ is given in Section 2. In the last section, we prove the asymptotic normality of a normalized version of $X_{n,d}$, from which one can easily get the asymptotic formula for the variance of $X_{n,d}$.

Throughout this paper, the limits are always to be taken as $n \to \infty$. For functions $f(n)$ and $g(n)$, we write $f(n) \sim g(n)$ if $\lim_{n\to\infty} f(n)/g(n) = 1$. We also use $\xrightarrow{\mathscr{D}}$ for convergence in distribution.

Our main result is as follows.

Theorem 0.1 For any integer $d \geqslant 1$, let $X_{n,d}$ be the number of vertices of degree d in a random planted plane tree of n edges, we have

$$\frac{X_{n,d} - E[X_{n,d}]}{\sqrt{\mathrm{Var}[X_{n,d}]}} \xrightarrow{\mathscr{D}} \mathscr{N}(0,1),$$

with

$$E[X_{n,d}] \sim \frac{1}{2^d} n, \mathrm{Var}[X_{n,d}] \sim \frac{2^{d+1} - (d-2)^2 - 2}{2^{2d+1}} n.$$

1 Functional equation

Let $\mathscr{P}$ be the set of all planted plane trees. For any tree $\omega \in \mathscr{P}$ and any integer $d \geqslant 1$, let $N(\omega)$ denote the number of edges and $M_d(w)$ be the number of vertices with degree d in w. We define a bivariate generating function as follows:

$$B(x,y) := \sum_{\omega \in \mathscr{P}} x^{N(\omega)} y^{M_d(w)}.$$

We can rewrite $B(x,y)$ as

$$B(x,y) = \sum_{n,k} C(n,k) x^n y^k \tag{2}$$

where $C(n,k)$ is the number of planted plane trees with n edges which contain exactly k vertices of degree d. Clearly, $B(x,y)$ and $C(n,k)$ both depend on d. For the explicit formula for $C(n,k)$, see Ref. [1].

Using the branch decomposition on the family $\mathscr{P}$, it is not hard to show that

$$\mathscr{P} - \{\epsilon\} = \{\epsilon\} \times (\mathscr{P} + \mathscr{P} \times \mathscr{P} + \mathscr{P} \times \mathscr{P} \times \mathscr{P} + \cdots),$$

with ϵ being a planted plane tree with no edges and $\times$ being the Cartesian product. For an example of branch decomposition of planted plane trees, see Fig. 1. Then, by the symbolic method (see, for example, Ref. [11]), one can show that $B(x,y)$ satisfies the following functional equation:

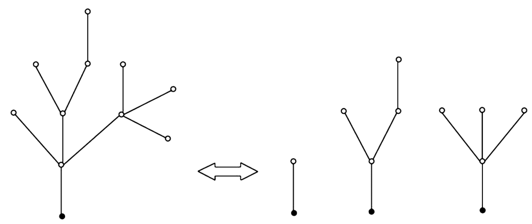

Fig. 1 Branch decomposition of a planted plane tree with three branches

$$B(x,y)=x(1+B(x,y)+B^2(x,y)+\cdots+ B^{d-2}(x,y)+yB^{d-1}(x,y)+B^d(x,y)+\cdots).$$

Simplifying this relation, we obtain the functional equation as follows:

$$B(x,y)=x\Big(\frac{1}{1-B(x,y)}+(y-1)B^{d-1}(x,y)\Big) \tag{3}$$

for $d\geqslant 1$. This functional equation will play an essential role in our study of the asymptotic distribution of $X_{n,d}$.

2 The mean

For $E[X_{n,d}]$, by the definition of expectation,

$$E[X_{n,d}]=\frac{1}{C_n}\sum_{k=1}^{n-1}kC(n,k).$$

For convenience, we define the following two functions:

$$P(x):=B(x,1),\ G(x):=\frac{\partial}{\partial y}B(x,y)\mid_{y=1}.$$

From Eq. (2), we have

$$P(x)=\sum_{n=1}^{\infty}C_nx^n;$$

$$G(x)=\sum_{n=1}^{\infty}\Big(\sum_{k=1}^{n-1}kC(n,k)\Big)x^n.$$

The generating function $P(x)$ does not depend on d, but $G(x)$ does. On the relation between $P(x)$ and $G(x)$, we have the following result.

Proposition 2.1 For any integer $d\geqslant 1$,

$$G(x)=\frac{xP^{d-1}(x)(1-P(x))}{1-2P(x)} \tag{4}$$

Proof Rewrite Eq. (3) as

$$B(1-B)=x(1+(y-1)B^{d-1}(1-B)).$$

Differentiating with respect to y for both sides of the above equation, we have

$$(1-2B)\frac{\partial B}{\partial y}=xB^{d-1}(1-B)+ x(y-1)((d-1)B^{d-2}-dB^{d-1})\frac{\partial B}{\partial y}.$$

Applying $y=1$ to the above equation and carrying out some simplifications, one can check easily that Eq. (4) is valid for any $d\geqslant 1$.

Since $n\geqslant 1$, we have $P(x)=x+O(x^2)$, as $x\to 0$. Thus, if we consider x as a complex variable, $P(\cdot)$ is a conformal mapping in a small neighborhood of $x=0$. As a result, $P(x)$ has an inverse function in this neighborhood. This fact is very useful for the following manipulation of integrals. Make use of Cauchy integral formula in Eq. (4) to get

$$\sum_{k=1}^{n-1}kC(n,k)=\frac{1}{2\pi i}\oint_C G(x)\frac{\mathrm{d}x}{x^{n+1}}= \frac{1}{2\pi i}\oint_C\frac{xP^{d-1}(x)(1-P(x))}{1-2P(x)}\frac{\mathrm{d}x}{x^{n+1}}.$$

Here C is a small circle centered at $x=0$ and lies in a neighborhood of $x=0$, in which $P(x)$ has an inverse function.

If let $y=1$ in Eq. (3), then

$$P(x)(1-P(x))=x.$$

Differentiating with respect to P, we have

$$\frac{\mathrm{d}x}{\mathrm{d}P}=1-2P.$$

Then

$$E[X_{n,d}]=\frac{1}{C_n}\frac{1}{2\pi i}\oint_C\frac{P^{d-1}(x)(1-P(x))}{1-2P(x)}\frac{\mathrm{d}x}{x^n}= \frac{1}{C_n}\frac{1}{2\pi i}\oint_{\tilde{C}}\frac{p^{d-1}(1-p)}{1-2p}\frac{1}{(p-p^2)^n}(1-2p)\mathrm{d}p= \frac{1}{C_n}\frac{1}{2\pi i}\oint_{\tilde{C}}\frac{p^{d-1}(1-p)}{(p-p^2)^n}\mathrm{d}p,$$

where $\tilde{C}$ is a small circle centered at $p=0$. If we denote

$$h(p):=-\ln(p-p^2),$$

then the expectation of $X_{n,d}$ can be expressed as

$$E[X_{n,d}]=\frac{1}{C_n}\frac{1}{2\pi i}\oint_{\tilde{C}}p^{d-1}(1-p)\mathrm{e}^{nh(p)}\mathrm{d}p \tag{5}$$

Eq. (5) can be treated with the classical saddle point method (see, for example, Ref. [12] or Ref. [13]). Here we shall briefly sketch the process and omit the details.

The saddle point is the root of equation $\frac{\mathrm{d}}{\mathrm{d}p}h(p)=0$. That is,

$$-\frac{1-2p}{p-p^2}=0.$$

Thus, one can easily get that the saddle point is $p=\frac{1}{2}$. Moreover, one can get $h''(\frac{1}{2})=8$. Then, by the saddle point formula, we have

$$\frac{1}{2\pi i}\oint_{\tilde{C}} p^{d-1}(1-p)\exp\{nh(p)\}\mathrm{d}p=$$

$$\frac{1}{2\pi i}\frac{1}{2^{d-1}}\left(1-\frac{1}{2}\right)\exp\left\{nh\left(\frac{1}{2}\right)\right\}\left(\frac{-2\pi}{nh''(\frac{1}{2})}\right)^{\frac{1}{2}}\cdot$$

$$\left(1+O\left(\frac{1}{n}\right)\right)=\frac{4^n}{2^{d+2}\sqrt{\pi n}}\left(1+O\left(\frac{1}{n}\right)\right) \quad (6)$$

By Stirling's formula

$$n!\ =n^{n+1/2}\mathrm{e}^{-n}\sqrt{2\pi}\left(1+O\left(\frac{1}{n}\right)\right),$$

and Eq. (1), C_n can be expressed as

$$C_n=\frac{4^{n-1}}{n\sqrt{\pi n}}\left(1+O\left(\frac{1}{n}\right)\right) \quad (7)$$

Substituting Eq. (6) and Eq. (5) into Eq. (5), we get the following result.

Proposition 2.2 Let $X_{n,d}$ be the number of vertices of degree d in a random planted plane tree with n edges. Then

$$E[X_{n,d}]=\frac{n}{2^d}\left(1+O\left(\frac{1}{n}\right)\right).$$

3 Asymptotic distribution

In principle, one can continue working out the variance of $X_{n,d}$ by following a similar method. However, it becomes more complicated. To avoid the difficulty, we introduce the random variable

$$Y_{n,d}:=\frac{X_{n,d}-\frac{n}{2^d}}{a\sqrt{n}},$$

with a being a positive constant, which only depends on d. If $Y_{n,d}$ converges to the standard normal random variable in distribution as $n\to\infty$, then the variance of $X_{n,d}$ is asymptotic to a^2n. It will be shown in the following that

$$a=\sqrt{\frac{2^{d+1}-(d-2)^2-2}{2^{2d+1}}} \quad (8)$$

In the sequel we shall prove the asymptotic normality of $Y_{n,d}$ under Eq. (8).

3.1 Expression of the characteristic function

The characteristic function of $Y_{n,d}$ can be expressed as

$$E[\exp\{itY_{n,d}\}]=E\left[\exp\left\{it\frac{X_{n,d}-n/2^d}{a\sqrt{n}}\right\}\right]=$$

$$\exp\left\{-it\frac{n/2^d}{a\sqrt{n}}\right\}E\left[\exp(it\frac{X_{n,d}}{a\sqrt{n}})\right].$$

By the definition of expectation, we have

$$E\left[\exp\left\{it\frac{X_{n,d}}{a\sqrt{n}}\right\}\right]=\frac{1}{C_n}\sum_{k=1}^{n-1}\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}\right)^kC(n,k).$$

Let

$$B_n(x):=B\left(x,\exp\left\{\frac{it}{a\sqrt{n}}\right\}\right).$$

Note that

$$\sum_{k=1}^{n-1}\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}\right)^kC(n,k)=[x^n]B_n(x) \quad (9)$$

Inserting $y=\exp\left\{\frac{it}{a\sqrt{n}}\right\}$ to Eq. (3), we see that $B_n(x)$ satisfies the following functional equation:

$$B_n(x)=x\left(\frac{1}{1-B_n(x)}+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)B_n^{d-1}(x)\right).$$

That is,

$$x=\frac{B_n(x)}{\frac{1}{1-B_n(x)}+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)B_n^{d-1}(x)}.$$

By Lagrange inversion formula (see, for example, Ref. [11, Appendix A.6]), we have

$$[x^n]B_n(x)=\frac{1}{n}[\lambda^{n-1}]\phi^n(\lambda) \quad (10)$$

where

$$\phi(\lambda):=\frac{1}{1-\lambda}+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)\lambda^{d-1}.$$

Using Cauchy integral formula in Eq. (9), we get

$$\sum_{k=1}^{n-1}\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}\right)^kC(n,k)=$$

$$\frac{1}{n}\frac{1}{2\pi i}\oint_C\frac{\left(\frac{1}{1-z}+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}\right)^n}{z^n}\mathrm{d}z=$$

$$\frac{1}{n}\frac{1}{2\pi i}\oint_C\left(\frac{1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)}{z(1-z)}\right)^n\mathrm{d}z,$$

where C can be taken as a small circle centered at $z=0$. Divide both sides of the above equation by C_n to get

$$E[\exp\{itY_{n,d}\}]=\frac{\exp\left\{-\frac{it\sqrt{n}}{2^da}\right\}}{nC_n}\cdot$$

$$\frac{1}{2\pi i}\oint_C\left(\frac{1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)}{z(1-z)}\right)^n dz \tag{11}$$

Now one can see that the key to our problem is to find the asymptotic expression of the integral in Eq. (11).

3.2 Asymptotic formula for the saddle point

If we denote

$$h_n(z):=\ln\left(\frac{1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)}{z(1-z)}\right),$$

then the integral in Eq. (11) can be simplified as

$$\frac{1}{2\pi i}\oint_C\left(\frac{1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)}{z(1-z)}\right)^n dz = \frac{1}{2\pi i}\oint_C \exp\{nh_n(z)\}dz.$$

Solve the equation $\frac{\partial}{\partial z}h_n(z)=0$ to find the saddle point.

$$\frac{\partial}{\partial z}h_n(z)=\frac{\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)((d-1)z^{d-2}-dz^{d-1})}{1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)}-\frac{1-2z}{z(1-z)}=0,$$

which implies that

$$(d-2)\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)^2-(1-2z)=0 \tag{12}$$

For any complex number α and real number $r>0$, let

$$D(\alpha,r):=\{z: |z-\alpha|<r\}.$$

For some basic properties of the saddle point, we shall apply Hurwitz theorem[14], which is stated as follows.

Let $\{f_n\}$ be a sequence of holomorphic functions on a connected open set G that converges uniformly on compact subsets of G to a holomorphic function f which is not constantly zero on G. If f has a zero of order m at z_0 then for every small enough $\rho>0$ and for sufficiently large integer n (depending on ρ), f_n has precisely m zeroes in the disk defined by $|z-z_0|<\rho$, including multiplicity. Furthermore, these zeroes converge to z_0 as $n\to\infty$.

Lemma 3.1 For all sufficiently large integer n, the polynomial equation

$$(d-2)\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)^2-(1-2z)=0$$

has a unique root r_n in the disc $D(0,\frac{2}{3})$. Furthermore,

$$r_n=\frac{1}{2}-\frac{d_1 it}{a\sqrt{n}}+O\left(\frac{1}{n}\right),$$

as $n\to\infty$, with

$$d_1=\frac{d-2}{2^{d+2}}.$$

Proof Consider the function

$$f_n(z):=(d-2)\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)^2-(1-2z),$$

and the function

$$f(z):=-1+2z.$$

Since $\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\to 0$, $f_n(z)$ converges uniformly to $f(z)$ in the disc $D(0,\frac{2}{3})$ as $n\to\infty$. It is obvious that $\frac{1}{2}$ is the unique root of $f(z)$. By Hurwitz theorem, for all sufficiently large positive integer n, $f_n(z)$ has a unique root r_n in $D(0,\frac{2}{3})$ and r_n converges to $\frac{1}{2}$. That is, $r_n=\frac{1}{2}+o(1)$.

For the asymptotic formula for r_n, here we use the bootstrap method. Let

$$r_n=\frac{1}{2}+\varepsilon_n \tag{13}$$

where $\varepsilon_n=o(1)$. Plug Eq. (13) into Eq. (12) to get

$$(d-2)\left(\frac{it}{a\sqrt{n}}+O\left(\frac{1}{n}\right)\right)\frac{1+O(\varepsilon_n)}{2^{d+1}}=-2\varepsilon_n.$$

Thus,

$$\frac{(d-2)it}{2^{d+1}a\sqrt{n}}+O\left(\frac{1}{n}\right)+O\left(\frac{\varepsilon_n}{\sqrt{n}}\right)+O\left(\frac{\varepsilon_n}{n}\right)=-2\varepsilon_n.$$

Comparing the order of both sides of the above equation, we have

$$\varepsilon_n=-\frac{(d-2)it}{2^{d+2}a\sqrt{n}}+O\left(\frac{1}{n}\right).$$

This finishes the proof for Lemma 3.1.

If r_n is expressed as

$$r_n=|r_n|\exp\{i\phi_n\},$$

then, from Lemma 3.1, we have

$$|r_n|=\frac{1}{2}+O\left(\frac{1}{n}\right),\ \phi_n=O\left(\frac{1}{\sqrt{n}}\right) \tag{14}$$

We remark that these two estimations occur frequently in the subsequent subsections.

3.3 Estimations and convergence

Choose a constant ε in the interval $(\frac{1}{3},\frac{1}{2})$. Now we can expand contour C of the integral in Eq. (11) to the circle $|z|=|r_n|$. That is, for all sufficiently large integer n,

$$\frac{1}{2\pi i}\oint_C \exp\{nh_n(z)\}\mathrm{d}z=$$

$$\frac{1}{2\pi i}\oint_{|z|=|r_n|}\exp\{nh_n(z)\}\mathrm{d}z:=J_1+J_2+J_3,$$

where

$$J_k=\frac{1}{2\pi i}\int_{C_k}\exp\{nh_n(z)\}\mathrm{d}z,\ k=1,2,3,$$

and where

$$C_1=\{z:|z|=|r_n|,\ |\arg z-\phi_n|\leqslant n^{-\varepsilon}\},$$
$$C_2=\{z:|z|=|r_n|,\ n^{-\varepsilon}<|\arg z-\phi_n|\leqslant\delta\},$$
$$C_3=\{z:|z|=|r_n|,\ \delta<|\arg z-\phi_n|\leqslant\pi\}.$$

Here $\delta>0$ being a sufficiently small constant depending only on d. Note that C_1 is a circular arc symmetrical about the saddle point r_n and it subtends an angle of size $2n^{-\varepsilon}$.

We estimate J_1, J_2 and J_3 in the following. It will be shown below that J_1 gives the main contribution. We start with J_3.

Lemma 3.2 There exists a constant $c_\delta\in(0,1)$, such that

$$J_3=O((4-c_\delta)^n).$$

Proof Similar to the convergence of $f_n(z)$ to $f(z)$, $|\exp\{h_n(z)\}|$ converges uniformly to $\left|\frac{1}{z(1-z)}\right|$ in $D(0,\frac{2}{3})\backslash D(0,\frac{1}{3})$ as $n\to\infty$. Let $z=\frac{1}{2}\exp\{i\theta\}$, $\theta\in(\delta,\pi)$. It is not hard to see that the function

$$g(\theta):=\left|\frac{1}{\frac{1}{2}\exp\{i\theta\}(1-\frac{1}{2}\exp\{i\theta\})}\right|=\frac{2}{\sqrt{\frac{5}{4}-\cos\theta}}$$

is strictly decreasing and $\max\limits_{0\leqslant\theta\leqslant\pi}g(\theta)=g(0)=4$. Thus, by Hurwitz theorem, the proof of Lemma 3.2 is complete.

Lemma 3.3 There exists a constant $c>0$, such that

$$J_2=O\left(\frac{4^n}{\sqrt{n}}\exp\{-cn^{1-2\varepsilon}\}\right).$$

Proof Through the transform $z=r_n\exp\{i\theta\}$, one can get

$$J_2=\frac{r_n}{2\pi}\int_{n^{-\varepsilon}\leqslant|\theta|\leqslant\delta}\left(\frac{1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)r_n^{d-1}\exp\{(d-1)i\theta\}(1-r_n\exp\{i\theta\})}{r_n\exp\{i\theta\}(1-r_n\exp\{i\theta\})}\right)^n\exp\{i\theta\}\mathrm{d}\theta.$$

Thus,

$$|J_2|\leqslant\frac{|r_n|^{-n+1}}{2\pi}\int_{n^{-\varepsilon}\leqslant|\theta|\leqslant\delta}\left|\frac{1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)r_n^{d-1}\exp\{(d-1)i\theta\}(1-r_n\exp\{i\theta\})}{1-r_n\exp\{i\theta\}}\right|^n\mathrm{d}\theta \tag{15}$$

We consider the case $n^{-\varepsilon}\leqslant\theta\leqslant\delta$ in the following. One can treat the other case $-\delta\leqslant\theta\leqslant-n^{-\varepsilon}$ in a similar way. By Eq. (14),

$$|1-r_n\exp\{i\theta\}|^n=|1-|r_n|\exp\{i(\phi_n+\theta)\}|^n=$$
$$(1+|r_n|^2-2|r_n|\cos(\phi_n+\theta))^{\frac{n}{2}}=$$

$$\left(\frac{5}{4}-\cos(\phi_n+\theta)+O\left(\frac{1}{n}\right)\right)^{\frac{n}{2}}\geqslant$$
$$\left(\frac{1}{4}+\frac{1}{3}(\phi_n+\theta)^2+O\left(\frac{1}{n}\right)\right)^{\frac{n}{2}},$$

since δ is sufficiently small. If n is sufficiently large, then $|\phi_n|<\theta/2$. It follows that

$$\left|\frac{1}{1-r_n\exp\{i\theta\}}\right|^n\leqslant$$
$$\left(\frac{1}{4}+\frac{1}{3}\left(-\frac{\theta}{2}+\theta\right)^2+O\left(\frac{1}{n}\right)\right)^{-\frac{n}{2}}=$$
$$2^n\left(1+\frac{\theta^2}{3}+O\left(\frac{1}{n}\right)\right)^{-\frac{n}{2}}=$$
$$2^n\exp\left\{-\frac{n}{2}\ln\left(1+\frac{\theta^2}{3}+O\left(\frac{1}{n}\right)\right)\right\}.$$

The relation $\ln(1+x)\geqslant\frac{x}{2}$ holds, if $x>0$ is sufficiently small. Thus, if δ is sufficiently small,

$$\left|\frac{1}{1-r_n\exp\{i\theta\}}\right|^n\leqslant 2^n\exp\left\{-\frac{n}{12}\theta^2+O(1)\right\}=$$
$$O\left(2^n\exp\left\{-\frac{n}{12}\theta^2\right\}\right)\tag{16}$$

Let

$$R_n(\theta):=r_n^{d-1}\exp\{(d-1)i\theta\}(1-r_n\exp\{i\theta\}).$$

Since $|R_n(\theta)|$ is bounded, one can get

$$\left|1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)r_n^{d-1}\cdot\right.$$
$$\left.\exp\{(d-1)i\theta\}(1-r_n\exp\{i\theta\})\right|^n=$$
$$\left|\exp\left\{n\ln\left(1+\left(\frac{it}{a\sqrt{n}}+O\left(\frac{1}{n}\right)\right)R_n(\theta)\right)\right\}\right|=$$
$$O\left(\exp\left\{-\frac{t}{a}\sqrt{n}\,\mathrm{Im}(R_n(\theta))\right\}\right),$$

where $\mathrm{Im}(R_n(\theta))$ is the imaginary part of $R_n(\theta)$. For $R_n(\theta)$, we have

$$R_n(\theta)=|r_n|^{d-1}\exp\{(d-1)i(\phi_n+\theta)\}\cdot$$
$$(1-|r_n|\exp\{i(\phi_n+\theta)\})=$$
$$\left(\frac{1}{2^{d-1}}+O\left(\frac{1}{n}\right)\right)(\cos[(d-1)(\phi_n+\theta)]+$$
$$i\sin[(d-1)(\phi_n+\theta)])\cdot$$
$$\left(1-\frac{1}{2}\cos(\phi_n+\theta)-\frac{1}{2}i\sin(\phi_n+\theta)+O\left(\frac{1}{n}\right)\right).$$

Then,

$$\mathrm{Im}(R_n(\theta))=$$
$$\frac{1}{2^{d-1}}\left[\left(1-\frac{1}{2}\cos(\phi_n+\theta)\right)\sin[(d-1)(\phi_n+\theta)]-\right.$$
$$\left.\frac{1}{2}\cos[(d-1)(\phi_n+\theta)]\sin(\phi_n+\theta)\right]+O\left(\frac{1}{n}\right)=$$
$$O(\theta)+O\left(\frac{1}{\sqrt{n}}\right).$$

Thus, there exists a constant $c_1>0$ not depending on n, such that

$$\left|1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)r_n^{d-1}\cdot\right.$$
$$\left.\exp\{(d-1)i\theta\}(1-r_n\exp\{i\theta\})\right|^n=$$
$$O(\exp\{c_1\theta\sqrt{n}\}).$$

Hence, by Eqs. (15) and (16),

$$|J_2|=O\left(2^n\int_{n^{-\varepsilon}}^{\delta}2^n\exp\left\{-\frac{n}{12}\theta^2+c_1\theta\sqrt{n}\right\}\mathrm{d}\theta\right)=$$
$$O\left(\frac{4^n}{\sqrt{n}}\int_{n^{1/2-\varepsilon}}^{\delta\sqrt{n}}\exp\left\{-\frac{\mu^2}{12}+c_1\mu\right\}\mathrm{d}\mu\right)=$$
$$O\left(\frac{4^n}{\sqrt{n}}\int_{n^{1/2-\varepsilon}}^{\delta\sqrt{n}}\exp\left\{-\frac{1}{12}(\mu-6c_1)^2\right\}\mathrm{d}\mu\right).$$

Recall that ε is chosen between $\frac{1}{3}$ and $\frac{1}{2}$. So the integral is of $O(\exp\{-cn^{1-2\varepsilon}\})$, which implies Lemma 3.3.

Lemma 3.4 For J_1, we have

$$J_1=\frac{4^{n-1}}{\sqrt{\pi n}}\exp\left\{\frac{it\sqrt{n}}{2^d a}-\frac{2^{d+1}-(d-2)^2-2}{2^{2(d+1)}}\cdot\frac{t^2}{a^2}\right\}\cdot$$
$$(1+O(n^{1-3\varepsilon}))\tag{17}$$

Proof Recall that

$$J_1=\frac{1}{2\pi i}\int_{C_1}\left(\left(1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)z^{d-1}(1-z)\right)/\right.$$
$$\left.(z(1-z))\right)^n\mathrm{d}z=\frac{1}{2\pi i}\int_{C_1}\exp\{nh_n(z)\}\mathrm{d}z.$$

The following argument acquires its inspiration from Ref. [12]. Note that the function $\ln\frac{1}{z(1-z)}$ is analytic at $z=\frac{1}{2}$ and the radius of convergence of its Taylor series is $\frac{1}{2}$. Moreover, $h_n(z)$ converges uniformly to $\ln\frac{1}{z(1-z)}$ in the disc $D(\frac{1}{2},\frac{1}{6})$, as $n\to\infty$. Since the saddle point $r_n\to\frac{1}{2}$, by Hurwitz theorem, the following claim holds: For any sufficiently large integer n, there exists a constant $\rho>0$ not dependent on n, such that the radius of convergence of Taylor series of $h_n(z)$ about the point $z=r_n$ is not smaller than ρ. Then, in $|z-r_n|\leqslant\rho$, we can expand $h_n(z)$ as

$$h_n(z)=h_n(r_n)+\frac{1}{2!}h''_n(r_n)(z-r_n)^2+F_n(z),$$

where

$$F_n(z)=a_{3,n}(z-r_n)^2+a_{4,n}(z-r_n)^4+\cdots,$$

with $a_{3,n}, a_{4,n},\cdots$ being Taylor coefficients.

Also note that $C_1\subset D(r_n,\rho)$ if n is sufficiently large. Let

$$M_n:=\max_{|z-r_n|=\rho}|h_n(z)|.$$

From Cauchy inequality (see, for example, Ref. [14]), we have

$$|a_{m,n}|\leqslant\frac{M_n}{\rho^m},\ m=3,4,\cdots.$$

Then, for any $z\in C_1$,

$$|F_n(z)|\leqslant\frac{M_n}{\rho^3}|z-r_n|^3+\frac{M_n}{\rho^4}|z-r_n|^4+\cdots=\frac{M_n|z-r_n|^3}{\rho^2(\rho-|z-r_n|)}\leqslant\frac{2M_n}{\rho^3}|z-r_n|^3.$$

Since $h_n(z)$ converges uniformly to $\ln\frac{1}{z(1-z)}$, there exists a constant $M>0$, such that $M_n\leqslant M$ for all n. Therefore, if $z\in C_1$,

$$|nF_n(z)|\leqslant\frac{2M}{\rho^3}n\,|r_n(\exp\{i\phi_n\}-1)|\leqslant\frac{2M}{\rho^3}n^{1-3\epsilon}\to 0.$$

Thus,

$$J_1=\frac{1}{2\pi i}\int_{C_1}\exp\left\{n\left(h_n(r_n)+\frac{1}{2}h''_n(r_n)(z-r_n)^2+F_n(z)\right)\right\}\mathrm{d}z=\frac{\exp\{nh_n(r_n)\}}{2\pi i}\int_{C_1}\exp\left\{\frac{n}{2}h''_n(r_n)(z-r_n)^2+nF_n(z)\right\}\mathrm{d}z=\left(\frac{\exp\{nh_n(r_n)\}}{2\pi i}\int_{C_1}\exp\left\{\frac{n}{2}h''_n(r_n)(z-r_n)^2\right\}\mathrm{d}z\right)(1+O(n^{1-3\epsilon})).$$

Replacing $z=r_n\exp\{i\theta\}$, we proceed,

$$J_1=\left(\frac{\exp\{nh_n(r_n)\}r_n}{2\pi}\int_{-n^{-\epsilon}}^{n^{-\epsilon}}\exp\left\{\frac{n}{2}h''_n(r_n)r_n^2(\exp\{i\theta\}-1)^2+i\theta\right\}\mathrm{d}\theta\right)(1+O(n^{1-3\epsilon})).$$

Since

$$(\exp\{i\theta\}-1)^2=-\theta^2+O(n^{-3\epsilon})\text{ and }\exp\{i\theta\}=1+O(n^{-\epsilon}),$$

we have

$$J_1=\left(\frac{\exp\{nh_n(r_n)\}r_n}{2\pi}\int_{-n^{-\epsilon}}^{n^{-\epsilon}}\exp\left\{-\frac{n}{2}h''_n(r_n)r_n^2\theta^2\right\}\mathrm{d}\theta\right)(1+O(n^{1-3\epsilon})).$$

Changing again the integral variable $\mu\to\theta\sqrt{n}$, to get that

$$J_1=\frac{\exp\{nh_n(r_n)\}r_n}{2\pi\sqrt{n}}\left(\int_{-n^{1/2-\epsilon}}^{n^{1/2-\epsilon}}\exp\left\{-\frac{h''_n(r_n)r_n^2}{2}\mu^2\right\}\mathrm{d}\mu\right)(1+O(n^{1-3\epsilon}))=$$

$$\frac{\exp\{nh_n(r_n)\}r_n}{2\pi\sqrt{n}}\left(\sqrt{\frac{2\pi}{h''_n(r_n)r_n^2}}+O(n^{1-3\epsilon})\right)(1+O(n^{1-3\epsilon}))=\frac{\exp\{nh_n(r_n)\}}{\sqrt{2\pi n h''_n(r_n)}}(1+O(n^{1-3\epsilon})).$$

After routine calculations, one can get that

$$h''_n(r_n)=8+O\left(\frac{1}{\sqrt{n}}\right)=8+O(n^{1-3\epsilon}).$$

Then,

$$J_1=\frac{\exp\{nh_n(r_n)\}}{4\sqrt{\pi n}}(1+O(n^{1-3\epsilon}))\tag{18}$$

Now $\exp\{nh_n(r_n)\}$ is the only term to be estimated. Recall that

$$r_n=\frac{1}{2}-\frac{d_1it}{a\sqrt{n}}+O\left(\frac{1}{n}\right),$$

with $d_1=\dfrac{d-2}{2^{d+2}}$. We have

$$\exp\{nh_n(r_n)\}=\left(\frac{1+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)r_n^{d-1}(1-r_n)}{r_n(1-r_n)}\right)^n=$$

$$\left(\frac{1}{r_n(1-r_n)}+\left(\exp\left\{\frac{it}{a\sqrt{n}}\right\}-1\right)r_n{}^{d-2}\right)^n=$$

$$\left(4-\frac{16d_1^2t^2}{a^2n}+\left(\frac{it}{a\sqrt{n}}-\frac{t^2}{2a^2n}\right)\left(\frac{1}{2}-\frac{d_1it}{a\sqrt{n}}\right)^{d-2}+O(n^{\frac{3}{2}})\right)^n=$$

$$4^n\left(1-\frac{4d_1^2t^2}{a^2n}+\frac{1}{2^d}\left(\frac{it}{a\sqrt{n}}-\frac{t^2}{2a^2n}\right)\left(1-\frac{2(d-2)d_1it}{a\sqrt{n}}\right)+O(n^{\frac{3}{2}})\right)^n=$$

$$4^n\left(1+\frac{it}{2^da\sqrt{n}}-\frac{2^{d+1}-(d-2)^2}{2^{2(d+1)}}\cdot\frac{t^2}{a^2n}+O(n^{\frac{3}{2}})\right)^n=$$

$$4^n\exp\left\{\frac{it\sqrt{n}}{2^da}-\frac{2^{d+1}-(d-2)^2-2}{2^{2(d+1)}}\cdot\frac{t^2}{a^2}\right\}(1+O(n^{-\frac{1}{2}})),$$

which implies with Eq. (18) that Eq. (17) holds.

The proof of Lemma 3.4 is complete.

It follows by Lemmas 3.2 and 3.3 that, as $n\to\infty$,

$$J_3=o(J_1),\ J_2=o(J_1).$$

If a takes value as in Eq. (8), then

$$J_1=\frac{4^{n-1}}{\sqrt{\pi n}}\exp\left\{\frac{it\sqrt{n}}{2^da}-\frac{t^2}{2}\right\}(1+O(n^{1-3\varepsilon})).$$

Hence, by Eqs. (7) and (11), we have

$$E[\exp\{itY_{n,d}\}]\sim\exp\left\{-\frac{t^2}{2}\right\},$$

which is the characteristic function for the standard normal distribution. The proof of Theorem 0.1 is complete.

Acknowledgements We are grateful to Prof. William Goh for helpful discussions on this topic.

References

[1] GOULDEN I, JACKSON D. Combinatorial Enumeration [M]. New York: John Wiley & Sons, 1983.

[2] PRODINGER H. A note on a result of R. Kemp on r-tuply rooted planted plane trees[J]. Computing, 1982, 28: 363-366.

[3] PRODINGER H. The height of planted plane trees revisited[J]. Ars Combinatoria, 1983, 16: 51-55.

[4] PRODINGER H. The average height of the d-th highest leaf of a planted plane tree[J]. Networks, 1985, 16: 67-75.

[5] GUTJAHR W. The variance of level numbers in certain families of trees[J]. Random Structures and Algorithms, 1992, 3: 361-374.

[6] CHEN W, LI N, SHAPIRO L. The buttery decomposition of plane trees[J]. Discrete Applied Mathematics, 2007, 155: 2187-2201.

[7] MAHMOUD H, SMYTHE R. Asymptoic joint normality of outdegrees of nodes in random recursive trees[J]. Random Structures and Algorithms, 1992, 3: 255-266.

[8] JANSON S. Asymptotic degree distribution in random recursive trees [J]. Random Structures and Algorithms, 2005, 26: 69-83.

[9] KALPATHY R, MAHMOUD H. Degree profile of m-ary search trees: A vehicle for data structure compression[J]. Probability in the Engineering and Informational Sciences, 2015, 30: 113-123.

[10] ZHANG P, MAHMOUD H. The degree profile and weight in Apollonian networks and k-trees [J]. Advances in Applied Probability, 2016, 48: 163-175.

[11] FLAJOLET P, SEDGEWICK R. Analytic Combinatorics [M]. Cambridge: Cambridge University Press, 2009.

[12] COPSON E. Asymptotic Expansions[M]. New York: Cambridge Universtiy Press, 1965.

[13] WONG R. Asymptotic Approximations of Integrals [M]. Philadelphia, PA: SIAM, 2001.

[14] TITCHMARSH E C. The Theory of Functions [M]. Reprinted. New York: Oxford University Press, 1991.