

The Generalized Zagreb Index for Non-Plane and Plane Recursive Trees

Qunqiang Feng

Department of Statistics and Finance
University of Science and Technology of China
Hefei 230026
P. R. China

Michael Fuchs

Department of Mathematical Sciences
National Chengchi University
Taipei 116
Taiwan

Tsan-Cheng Yu*

Department of Mathematics
Fu Jen Catholic University
New Taipei City, 242062
Taiwan

October 14, 2025

Abstract

The Zagreb index, which is defined as the sum of squares of degrees of the nodes of a tree, was studied in previous works by martingale techniques for random non-plane recursive trees and classes of random trees which are close to random plane recursive trees. These techniques are not easily amended to the generalized Zagreb index, which is defined similar but with squares replaced by higher powers. In this paper, we use the moment transfer approach to (i) obtain the first-order asymptotics of moments and to (ii) prove limit laws for the (suitable normalized) generalized Zagreb index for random non-plane and plane recursive trees; for the former, we show that for all higher powers the limit law is normal, for the latter, we show for cubes and fourth powers that its a non-normal law.

1 Introduction

The Zagreb indices are among the most extensively studied topological indices in chemical graph theory. Initially introduced by Gutman and Trinajstić in the 1970s [15], the first Zagreb index has been widely utilized to predict the physicochemical properties of chemical compounds [22]. It has also been extensively applied in quantitative structure-property relationship (QSPR) and quantitative structure-activity relationship (QSAR) studies (see, e.g., [6]). Recently, researchers have generalized this index to accommodate more complex molecular

*Corresponding Author, Email: tsancheng@mail.fju.edu.tw

structures, resulting in the development of the generalized Zagreb index (see [2,20,25] and references therein). These generalizations provide enhanced flexibility in modeling molecular structures and their associated properties [13].

The generalized Zagreb index is formally defined as follows. Let $G = (V, E)$ be a graph with vertex set V and edge set E . For any positive integer k , the k -th order generalized Zagreb index of G is defined as

$$Z_G^{(k)} = \sum_{v \in V} D_v^k = \sum_{uv \in E} (D_u^{k-1} + D_v^{k-1}),$$

where D_v denotes the degree of vertex v in G . Specially, $Z_G^{(k)}$ corresponds to the first Zagreb index when $k = 2$, and it is referred to as the forgotten topological index when $k = 3$ [13]. The mathematical properties of the generalized Zagreb index have been extensively studied. For example, extremal values for this index in trees and unicyclic graphs are determined in [17,19], and upper or lower bounds of $Z_G^{(k)}$ are derived for some specific classes of graphs, such as planar graphs [16], bipartite graphs [23], and line graphs [4]. For a comprehensive review of the first Zagreb index and its applications in chemistry, we refer to [14], and for relations between the generalized Zagreb index and other topological indices, see [1].

Random non-plane recursive trees (also just called recursive trees) and plane recursive trees (also called plane-oriented recursive trees or PORTs) are two fundamental structures in the study of random graphs and combinatorial probability (see, e.g., [8]). Using martingale techniques, the limit laws for the first Zagreb index of non-plane recursive trees and Barabási-Albert trees (a slight variant of plane recursive trees) are established in [9,10]; see also [24] for explicit expressions of the first two moments and the limit law of the first Zagreb index for another (slight) variant of plane recursive trees. Furthermore, with an application of Stein’s method, the asymptotic normality of the first Zagreb index for classical Erdős-Rényi (ER) random graphs is established as the graph size approaches infinity [11]. For the generalized Zagreb index of ER random graphs, its expectation is obtained in [7], and several limits laws are further demonstrated in [12]. Motivated by these studies, our work focuses on the asymptotic behavior of this index for non-plane and plane recursive trees. While extending prior work [9,10,24] might seem natural, the martingale techniques and Stein’s method are not applicable for general k . Instead, we employ the moment transfer approach in this work.

The rest of this paper is organized as follows. In the next section, we formally define random non-plane and plane recursive trees, obtain distributional recurrences for the general Zagreb index, and recall the moment transfer approach. In Section 3, we apply this method to derive the first-order asymptotics of moments and a central limit theorem for non-plane recursive trees (including the case $k = 2$ which was already obtained in earlier works). In Section 4, we show corresponding results for plane recursive trees (for $k \geq 3$ as the case $k = 2$ requires a different treatment). We end the paper with a conclusion in Section 5.

2 Recurrences and the Moment-Transfer Approach

Random non-plane and plane recursive trees. We first give precise definitions of the two random tree models considered in this paper.

First, a *random non-plane recursive tree* of size n is recursively built as follows. Start with a root node labeled by 1. In the i -th step, a node labeled by i joins to already built tree by

becoming the child of one of $i - 1$ existing nodes, where the parent node is picked uniformly at random from these $i - 1$ nodes. We stop when exactly n nodes have joined.

Note that the children of every non-leaf node in a random non-plane recursive tree are not ordered. On the other hand, in a *random plane recursive tree*, in the i -th step, we assume that a node with k children has $k + 1$ free places (one in front of the first child, the next between the first and the second child, etc.). The i -th node then chooses uniformly at random from one of those free places (among all the free places) and is attached there. Thus, in contrast to random non-plane recursive trees, the children of every non-leaf node in a random plane recursive tree have a left-to-right order.

The above definitions easily imply that there are $(n - 1)!$ different non-plane recursive trees of size n and

$$(2n - 3)!! = \frac{(2n - 2)!}{2^{n-1}(n - 1)!} = \frac{n!C_n}{2^{n-1}}$$

different plane recursive trees of size n , where C_n denotes the (shifted) Catalan numbers. Moreover, from this, we can re-derive the distribution of the size I_n of the left-most subtree of the root of a random non-plane or plane recursive tree (where for non-plane recursive trees, we order the subtrees of the root according to increasing labels of their roots); see [8] and [18].

Lemma 2.1. (i) *For non-plane recursive trees, I_n has a uniform distribution on the set $\{1, \dots, n - 1\}$.*

(ii) *For plane recursive trees,*

$$\mathbb{P}(I_n = j) = \frac{2(n - j)C_j C_{n-j}}{nC_n}, \quad (1 \leq j \leq n - 1).$$

Proof. Both results follow from (simple) counting arguments using the above numbers of non-plane and plane recursive trees and decomposing the tree into the left-most subtree of the root and the remaining tree (both of which are again non-plane and plane recursive trees, respectively). First, for part (i):

$$\mathbb{P}(I_n = j) = \binom{n-2}{j-1} \frac{(j-1)!(n-j-1)!}{(n-1)!} = \frac{1}{n-1},$$

where the binomial coefficient takes care of the relabeling of the two trees. Likewise for part (ii):

$$\mathbb{P}(I_n = j) = \binom{n-1}{j} \frac{(2j-3)!!(2n-2j-3)!!}{(2n-3)!!}$$

which simplifies to the claimed form. $\square$

Generalized Zagreb index and root degree. The decomposition of random non-plane and plane recursive trees of size n used in the proof of Lemma 2.1 can also be used to recursively compute the generalized Zagreb index, which we denote by $Z_n^{(k)}$ throughout this work.

Proposition 2.2. *The generalized Zagreb index of a random (non-plane or plane) recursive tree of size n satisfies, for $n \geq 2$,*

$$Z_n^{(k)} \stackrel{d}{=} Z_{I_n}^{(k)} + \tilde{Z}_{n-I_n}^{(k)} - R_{I_n}^k + (R_{I_n} + 1)^k - \tilde{R}_{n-I_n}^k + (\tilde{R}_{n-I_n} + 1)^k \quad (1)$$

with initial condition $Z_1^{(k)} = 0$. Here, R_n denotes the (random) degree of the root, $\tilde{Z}_n$ and $\tilde{R}_n$ are independent copies of Z_n and R_n , and (Z_n, R_n) and (I_n) are independent.

Proof. The decomposition of a random (non-plane or plane) recursive tree into the left-most subtree of the root and the remaining tree gives two independent (non-plane or plane) recursive trees of size I_n and $n - I_n$. Also, the Zagreb index is then the sum of Zagreb indices of these two trees where we have to correct the contributions at the two roots. For the root of the left-most subtree this correction gives the term $-R_{I_n}^k + (R_{I_n} + 1)^k$ since we are missing one edge of the root. Likewise, for the remaining tree, we have the correction $-\tilde{R}_{n-I_n}^k + (\tilde{R}_{n-I_n} + 1)^k$. $\square$

The same argument can be used to obtain a recurrence for the (random) root degree, too.

Lemma 2.3. *The degree of the root of a random (non-plane or plane) recursive tree of size n satisfies, for $n \geq 2$,*

$$R_n \stackrel{d}{=} \tilde{R}_{n-I_n} + 1 \quad (2)$$

with initial condition $R_1 = 0$.

From this result, we see that all (non-centered or centered) moments of R_n satisfy a recurrence of the type:

$$a_n = \sum_{j=1}^{n-1} \pi_{n,j} a_{n-j} + b_n, \quad (n \geq 2) \quad (3)$$

with initial condition $a_1 = 0$, where $\pi_{n,j} = \mathbb{P}(I_n = j)$ and b_n is a function which involves moments of lower order. By solving this recurrence (either exactly or asymptotically), a great deal of properties can be proved for the root degree; see Section 6.1.1 in [8] for random non-plane recursive trees and [18] for random plane recursive trees. We gather some of these properties (which are used below) in the next proposition.

Proposition 2.4. (i) *For random non-plane recursive trees, in distribution and with convergence of all moments,*

$$\frac{R_n - \log n}{\sqrt{\log n}} \xrightarrow{d} N(0, 1), \quad (n \rightarrow \infty),$$

where $N(0, 1)$ denotes a standard normal distribution.

(ii) *For random plane recursive trees, in distribution and with convergence of all moments,*

$$\frac{R_n}{\sqrt{n}} \xrightarrow{d} \text{Rayleigh}(\sqrt{2}),$$

where $\text{Rayleigh}(\sigma)$ denotes a Rayleigh distribution with parameter σ .

Similarly, all (non-centered and centered) moments of $Z_n^{(k)}$ satisfy the *two-sided* version of (3):

$$a_n = \sum_{j=1}^{n-1} \pi_{n,j} (a_j + a_{n-j}) + b_n, \quad (n \geq 2) \quad (4)$$

with initial condition $a_1 = 0$, where b_n is again a function of lower-order moments and moments of R_n .

This recurrence will be the starting point of our analysis of the generalized Zagreb index. To derive the first-order asymptotics of moments and limit laws from it, we are going to apply the *moment-transfer approach* which we explain next. (Similarly, one could re-prove Proposition 2.4 with this method from (3); see [5].)

Moment-transfer approach. Our goal is to prove limit laws for the generalized Zagreb index of non-plane and plane recursive trees. One way of proving a limit law in probability theory is to use the *method of moments*; see Section 30 in [3]. This method proceeds by (i) finding the first-order asymptotics of all moments (of the suitably normalized sequence of random variables) and (ii) identifying the limit law from the asymptotic moment sequence. The method is in particular well-suited for sequences of random variables which satisfy distributional recurrences such as (1) or (2) since all moments satisfy the same type of recurrence (see (4) for the former and (3) for the latter) which depends on the sequence b_n that involves lower-order moments. Thus, induction can be applied to find the asymptotics of all moments as follows: first the induction hypothesis is used to obtain the asymptotics of b_n . Then, one needs (general) results which bridge the asymptotics of b_n with that of a_n ; such results are called *asymptotic transfer results*. These are then used to complete the induction step. In summary, this method is called the *moment-transfer approach*. It has been used in numerous papers analyzing parameters of random trees; see for instance [5,18,21] and references therein.

Note that the distributional recurrence (1) of the Zagreb index depends also on R_n . Thus, in the induction step, we have to work with mixed moments of $Z_n^{(k)}$ and R_n . Consequently, we need an asymptotic transfer result not only for (4) but also for (3). Such results for both non-plane and plane recursive trees have been established in earlier work. We will state next the results which we need in Section 3 and Section 4.

First, for non-plane recursive trees, we need a further notation. Let the (modified) $\hat{\mathcal{O}}$ -notation be defined as the $\mathcal{O}$ -notation but with logarithmic terms suppressed, e.g., $\mathcal{O}(\log^k n) = \hat{\mathcal{O}}(1)$ for any integer $k \geq 0$. Then, we have the following asymptotic transfer results for non-plane recursive trees; see Lemma 2 and Lemma 6 in [21].

Lemma 2.5. *Consider (3) for non-plane recursive trees. If $b_n = \hat{\mathcal{O}}(n^\alpha)$ for $\alpha \geq 0$, then $a_n = \hat{\mathcal{O}}(n^\alpha)$.*

Lemma 2.6. *Consider (4) for non-plane recursive trees.*

- (i) *If $b_n = \hat{\mathcal{O}}(n^\alpha)$ with $0 \leq \alpha < 1$, then $a_n = \mu n + \hat{\mathcal{O}}(n^\alpha)$ where $\mu \in \mathbb{R}$;*
- (ii) *If $b_n \sim cn^\alpha$ with $\alpha > 1$, then $a_n \sim c(\alpha + 1)n^\alpha/(\alpha - 1)$.*

Remark 2.7. The constant in part (i) of the above lemma is given by:

$$\mu = 2 \sum_{j=2}^{\infty} \frac{b_j}{j(j+1)}. \quad (5)$$

Note that, under the assumption of part (i), the series converges.

Likewise, we have similar results for plane recursive trees; see [18].

Lemma 2.8. *Consider (3) for plane recursive trees. If $b_n \sim cn^\alpha$ with $\alpha > -1/2$, then $a_n \sim c\Gamma(\alpha + 1/2)n^{\alpha+1/2}/\Gamma(\alpha + 1)$, where $\Gamma(z)$ denotes the gamma function.*

Lemma 2.9. *Consider (4) for plane recursive trees.*

- (i) *If $b_n \sim c\sqrt{n}$ then $a_n \sim cn \log n/\sqrt{\pi}$;*
- (ii) *If $b_n \sim cn^\alpha$ with $\alpha > 1/2$, then $a_n \sim c\Gamma(\alpha - 1/2)n^{\alpha+1/2}/\Gamma(\alpha)$.*

3 Non-plane Recursive Trees

In this section, we prove a central limit theorem for the generalized Zagreb index of random non-plane recursive trees for all powers. We start by stating the main result of this section.

Theorem 3.1. *For random non-plane recursive trees, we have the following limit distribution result for the generalized Zagreb index: for $k \geq 2$,*

$$\frac{Z_n^{(k)} - \mu_k n}{\sigma_k \sqrt{n}} \xrightarrow{d} N(0, 1), \quad (n \rightarrow \infty),$$

where μ_k and σ_k are positive constants which are given in (9) and the proof of Lemma 3.5.

Remark 3.2. As mentioned in the introduction, the limit law for $k = 2$ (first Zagreb index) was already obtained in [9]. Our method works for this case, too, and gives an alternative proof.

We are going to prove the above result by computing the asymptotics of all moments. We first derive the mean and then consider higher moments. The starting point is (1) which we bring into the form:

$$Z_n^{(k)} \stackrel{d}{=} Z_{I_n}^{(k)} + \tilde{Z}_{n-I_n}^{(k)} + \sum_{\ell=0}^{k-1} \binom{k}{\ell} (R_{I_n}^\ell + \tilde{R}_{n-I_n}^\ell). \quad (6)$$

Mean value. By taking expectation on both sides of (6), we see that the mean $a_n := \mathbb{E}(Z_n^{(k)})$ satisfies (4) with

$$b_n = \sum_{\ell=0}^{k-1} \binom{k}{\ell} \frac{2}{n-1} \sum_{j=1}^{n-1} \mathbb{E}(R_j^\ell). \quad (7)$$

Next, by moment convergence in Proposition 2.4-(i):

$$\mathbb{E}(R_n - \log n)^m = \hat{\mathcal{O}}(1), \quad (m \geq 0), \quad (8)$$

which in turn implies that $\mathbb{E}(R_n^m) = \hat{\mathcal{O}}(1)$ for $m \geq 0$. Plugging this into (7) gives $b_n = \hat{\mathcal{O}}(1)$. Thus, from Lemma 2.6-(i):

$$\mu(n) := \mathbb{E}(Z_n^{(k)}) = \mu_k n + \hat{\mathcal{O}}(1), \quad (9)$$

where μ_k is a suitable constant given by (5). (This also shows that $\mu_k > 0$ as $b_n > 0$ for $n \geq 3$.)

For small values of k , the constant μ_k can be computed. For example, for $k = 2$, the above expression (7) for b_n becomes

$$b_n = 2 + \frac{4}{n-1} \sum_{j=1}^{n-1} \mathbb{E}(R_j).$$

Note that $\mathbb{E}(R_n) = H_{n-1}$, where $H_{n-1} := \sum_{j=1}^{n-1} (1/j)$ is the $(n-1)$ -st harmonic number. This can easily be derived from (2); see also Section 6.1.1 in [8]. Thus,

$$b_n = 2 + \frac{4}{n-1} \sum_{j=1}^{n-1} H_{j-1} = 4\Psi(n) + 4\gamma - 2,$$

where γ is Euler's constant and $\Psi(n)$ denotes the digamma function. Plugging this into (5) gives:

$$\mu_2 = 8 \sum_{j=2}^{\infty} \frac{\Psi(j)}{j(j+1)} + (2\gamma - 4) \sum_{j=2}^{\infty} \frac{1}{j(j+1)} = 6,$$

where the last step follows from

$$\sum_{j=2}^{\infty} \frac{1}{j(j+1)} = \frac{1}{2} \quad \text{and} \quad \sum_{j=2}^{\infty} \frac{\Psi(j)}{j(j+1)} = 1 - \frac{\gamma}{2}.$$

This value is the same as the one obtained in [9].

Higher moments. We next consider the asymptotics of higher moments, where we consider central moments. Therefore, set

$$\bar{Z}_n^{(k)} := Z_n^{(k)} - \mu(n).$$

Then, (6) becomes

$$\bar{Z}_n^{(k)} \stackrel{d}{=} \bar{Z}_{I_n}^{(k)} + \tilde{\bar{Z}}_{n-I_n}^{(k)} + \sum_{\ell=0}^{k-1} \binom{k}{\ell} (R_{I_n}^{\ell} + \tilde{R}_{n-I_n}^{\ell}) + \Delta(n), \quad (10)$$

where

$$\Delta(n) := \mu(I_n) + \mu(n - I_n) - \mu(n).$$

Note that from (9), for $1 \leq j \leq n-1$,

$$\begin{aligned} (\Delta(n)|I_n = j) &= \mu(j) + \mu(n-j) - \mu(n) \\ &= \mu_k j + \mu_k(n-j) - \mu_k n + \hat{\mathcal{O}}(1) = \hat{\mathcal{O}}(1). \end{aligned} \quad (11)$$

Since the recurrence (10) depends on R_n , we need to consider mixed moments. We first make the following important observation. (From now on, we suppress the dependence on k .)

Lemma 3.3. *The mixed moments $\mathbb{E}(\bar{Z}_n^r R_n^s)$ for $s > 0$ satisfy (3) and for $s = 0$ satisfy (4), where b_n is in both cases a function of $(\Delta(n)|I_n = j)$ and mixed moments of smaller order with respect to the lexicographic order.*

Proof. Raising (10) to the r -th power gives:

$$\begin{aligned} \mathbb{E}(\bar{Z}_n^r) &= \frac{1}{n-1} \sum_{j=1}^{n-1} \mathbb{E} \left(\bar{Z}_j + \tilde{\bar{Z}}_{n-j} + \sum_{\ell=0}^{k-1} \binom{k}{\ell} (R_j^{\ell} + \tilde{R}_{n-j}^{\ell}) + (\Delta(n)|I_n = j) \right)^r \\ &= \frac{1}{n-1} \sum_{j=1}^{n-1} \sum_{i_1+i_2+i_3=r} \binom{r}{i_1, i_2, i_3} \mathbb{E}(\bar{Z}_j^{i_1} \tilde{\bar{Z}}_{n-j}^{i_2} T_{n,j}^{i_3}), \end{aligned}$$

where

$$T_{n,j} := \sum_{\ell=0}^{k-1} \binom{k}{\ell} (R_j^{\ell} + \tilde{R}_{n-j}^{\ell}) + (\Delta(n)|I_n = j).$$

Likewise, for the mixed moments, by (2):

$$\mathbb{E}(\bar{Z}_n^r R_n^s) = \frac{1}{n-1} \sum_{j=1}^{n-1} \sum_{i_1+i_2+i_3=r} \binom{r}{i_1, i_2, i_3} \mathbb{E}(\bar{Z}_j^{i_1} \bar{Z}_{n-j}^{i_2} T_{n,j}^{i_3} (\bar{R}_{n-j} + 1)^s). \quad (12)$$

Expanding $T_{n,j}^{i_3}$ and $(\bar{R}_{n-j} + 1)^s$ in the above mean gives terms of the form

$$\mathbb{E}(\bar{Z}_j^{i_1} R_j^{i_1}) \mathbb{E}(\bar{Z}_{n-j}^{i_2} R_{n-j}^{\ell}), \quad (13)$$

where $i \leq i_3(k-1)$, $\ell \leq i_3(k-1) + s$, and $i_1 + i_2 + i_3 = r$. With respect to the lexicographic order, a highest-order term is obtained if we set $i_2 = r$ and $\ell = s$ which gives for (13) $\mathbb{E}(\bar{Z}_{n-j}^r R_{n-j}^s)$ since $i_1 = i_3 = 0$ and thus $i = 0$. Also note that the coefficient of this term is 1. If $s > 0$, the orders of all other terms are strictly smaller with respect to the lexicographic order which gives the first claim. On the other hand, for $s = 0$, we have another highest-order term, namely, by setting $i_1 = r$ and thus $i_2 = i_3 = 0$ which in turn gives $i = \ell = 0$, we obtain $\mathbb{E}(\bar{Z}_j^r)$ and thus the recurrence in this case is indeed two-sided. $\square$

We now use this lemma and induction to obtain the following proposition.

Proposition 3.4. *For $r, s \geq 0$,*

$$\mathbb{E}(\bar{Z}_n^r R_n^s) = \hat{\mathcal{O}}(n^{r/2}).$$

Moreover, for $s = 0$, we have the refinement:

$$\mathbb{E}(\bar{Z}_n^r) \sim g_r \sigma_k^r n^{r/2}, \quad (14)$$

where σ_k is a suitable constant and

$$g_r = \begin{cases} \frac{(2m)!}{2^m m!}, & \text{if } r = 2m; \\ 0, & \text{if } r = 2m + 1. \end{cases}$$

Proof. We use induction on (r, s) , which we again equip with the lexicographic order.

First, the claims hold for $r = 0$ and all $s \geq 0$ due to (8). Moreover, from the definition of $\bar{Z}_n$, the claims also hold for $(r, s) = (1, 0)$.

Next, we assume that the claims hold for all (r', s') which are smaller than (r, s) with respect to the lexicographic order, i.e., either $r' < r$ or $r = r'$ and $s' < s$. In order to prove the claims for (r, s) , we distinguish the two cases $s > 0$ and $s = 0$.

If $s > 0$, by Lemma 3.3, we have that $\mathbb{E}(\bar{Z}_n^r R_n^s)$ satisfies (3) with b_n being a function of $(\Delta(n)|I_n = j)$ and mixed moments of smaller order. Thus, (11) and the induction hypothesis yield $b_n = \hat{\mathcal{O}}(n^{r/2})$. From this, by Lemma 2.5, we obtain that $\mathbb{E}(\bar{Z}_n^r R_n^s) = \hat{\mathcal{O}}(n^{r/2})$. This proves the claim in this case.

If $s = 0$, then again by Lemma 3.3, we have that $\mathbb{E}(\bar{Z}_n^r)$ satisfies (4) with b_n given by

$$b_n = \sum_{i=1}^{r-1} \binom{r}{i} \frac{1}{n-1} \sum_{j=1}^{n-1} \mathbb{E}(\bar{Z}_j^i) \mathbb{E}(\bar{Z}_{n-j}^{r-i}) + \hat{\mathcal{O}}(n^{(r-1)/2}), \quad (15)$$

where this expression arises from (12), where we have separated the terms with $i_3 = 0$ (without the two main terms) and the remaining terms which satisfy the claimed bound by (11) and the induction hypothesis.

Now, for $r = 2$, the first term on the right-hand side of (15) equals 0 and thus $b_n = \hat{O}(n^{1/2})$. Applying Lemma 2.6-(i) gives $\mathbb{E}(\bar{Z}_n^2) \sim \sigma_k n$, for a suitable sequence σ_k , which proves the claim in this case.

Next, for $r > 2$, we plug the induction hypothesis into the first term on the right-hand side of (15) which yields

$$\begin{aligned} \sum_{i=1}^{r-1} \binom{r}{i} \frac{1}{n-1} \sum_{j=1}^{n-1} \mathbb{E}(\bar{Z}_j^i) \mathbb{E}(\bar{Z}_{n-j}^{r-i}) &\sim \sigma_k^r \sum_{i=1}^{r-1} \binom{r}{i} g_i g_{r-i} \frac{1}{n-1} \sum_{j=1}^{n-1} j^{i/2} (n-j)^{(r-i)/2} \\ &\sim \sigma_k^r n^{r/2} \left(\sum_{i=1}^{r-1} \binom{r}{i} g_i g_{r-i} \int_0^1 x^{i/2} (1-x)^{(r-i)/2} dx \right), \end{aligned}$$

where the constant inside the brackets for r odd equals 0 (since either $g_i = 0$ or $g_{r-i} = 0$ for $1 \leq i \leq r-1$) and for $r = 2m$ equals:

$$\begin{aligned} \sum_{i=1}^{2m-1} \binom{2m}{i} g_i g_{2m-i} \frac{\Gamma(i/2 + 1) \Gamma((2m-i)/2 + 1)}{\Gamma(m+2)} &= \sum_{i=1}^{m-1} \binom{2m}{2i} g_{2i} g_{2m-2i} \frac{i!(m-i)!}{(m+1)!} \\ &= \frac{(2m)!(m-1)!}{2^m(m+1)!}. \end{aligned}$$

Thus, in both cases, $b_n \sim \sigma_k^r n^{r/2} g_r (r/2 - 1)/(r/2 + 1)$. Applying Lemma 2.6-(ii) gives the claimed result also in this case. $\square$

The last step is to show that σ_k from the last proposition is strictly positive.

Lemma 3.5. *For all $k \geq 2$, we have $\sigma_k > 0$.*

Proof. In order to show the claim, we have to revisit the case $(r, s) = (2, 0)$ in the proof of the last proposition.

First, from (12), we see that $\mathbb{E}(\bar{Z}_n^2)$ satisfies (4) with

$$b_n = \mathbb{E}(T_{n, I_n}^2) + 4\mathbb{E}(\bar{Z}_{I_n} T_{n, I_n}) = \mathbb{E}(T_{n, I_n}^2) + 4 \sum_{\ell=0}^{k-1} \binom{k}{\ell} \mathbb{E}(\bar{Z}_{I_n} R_{I_n}^\ell).$$

Since σ_k is given by (5), the claimed result follows if (i) b_n is not identical to zero and (ii) $b_n \geq 0$. Item (i) is easy to check and thus, we only need to verify (ii). For this, we prove by induction on $\ell \geq 0$ that $\mathbb{E}(\bar{Z}_n R_n^\ell) \geq 0$.

The claim holds for $\ell = 0$. Next, from (12), we see that $\mathbb{E}(\bar{Z}_n R_n^\ell)$ satisfies (3) with

$$b_n = \sum_{i=0}^{\ell-1} \binom{\ell}{i} \mathbb{E}(\tilde{Z}_{n-I_n} \tilde{R}_{n-I_n}^i) + \mathbb{E}(T_{n, I_n} (\tilde{R}_{n-I_n} + 1)^\ell). \quad (16)$$

Note that the solution of (3) with $\pi_{n,j} = 1/(n-1)$ is given by

$$a_n = b_n + \sum_{j=2}^{n-1} \frac{b_j}{j}.$$

Thus, the claim follows if again $b_n \geq 0$. The first term on the right-hand side of (16) is non-negative by induction hypothesis. The second term, we rewrite as:

$$\mathbb{E}(T_{n,I_n}(\tilde{R}_{n-I_n} + 1)^\ell) = \sum_{i=0}^{k-1} \binom{k}{i} \mathbb{E}(R_{I_n}^i + \tilde{R}_{n-I_n}^i)(\tilde{R}_{n-I_n} + 1)^\ell + \Delta(n)\mathbb{E}(\tilde{R}_{n-I_n} + 1)^\ell.$$

Taking expectations on both sides of (10) gives

$$\Delta(n) = -2 \sum_{i=0}^{k-1} \binom{k}{i} \mathbb{E}(R_{I_n}^i).$$

Plugging this into the expression above and simplifying yields

$$\begin{aligned} \mathbb{E}(T_{n,I_n}(\tilde{R}_{n-I_n} + 1)^\ell) &= \sum_{i=0}^{k-1} \binom{k}{i} (\mathbb{E}(\tilde{R}_{n-I_n}^i(\tilde{R}_{n-I_n} + 1)^\ell) - \mathbb{E}(R_{I_n}^i)\mathbb{E}(\tilde{R}_{n-I_n} + 1)^\ell) \\ &= \sum_{i=0}^{k-1} \binom{k}{i} (\mathbb{E}(R_{I_n}^i(R_{I_n} + 1)^\ell) - \mathbb{E}(R_{I_n}^i)\mathbb{E}(R_{I_n} + 1)^\ell), \end{aligned}$$

where the last step follows by symmetry. Finally, by the binomial theorem,

$$\mathbb{E}(R_{I_n}^i(R_{I_n} + 1)^\ell) - \mathbb{E}(R_{I_n}^i)\mathbb{E}(R_{I_n} + 1)^\ell = \sum_{j=0}^{\ell} \binom{\ell}{j} (\mathbb{E}(R_{I_n}^{i+j}) - \mathbb{E}(R_{I_n}^i)\mathbb{E}(R_{I_n}^j))$$

and the expression inside the sum is non-negative because of Hölder's inequality. Thus,

$$\mathbb{E}(T_{n,I_n}(\tilde{R}_{n-I_n} + 1)^\ell) \geq 0,$$

which in turn implies that the b_n from (16) is non-negative. This concludes the induction and the proof of the lemma. $\square$

We can now complete the proof of Theorem 3.1.

Proof of Theorem 3.1. By Lemma 3.5 and (14), we obtain that $X_n = \bar{Z}_n/(\sigma_k n^{1/2})$ satisfies

$$\mathbb{E}(X_n^r) \longrightarrow g_r, \quad (n \rightarrow \infty).$$

Since $\mathbb{E}(N(0, 1))^r = g_r$, the method of moments gives $X_n \xrightarrow{d} N(0, 1)$ as claimed. $\square$

4 Plane Recursive Trees

This section contains our main result for random plane recursive trees which is the following limit distribution result.

Theorem 4.1. *For random plane recursive trees, we have the following limit distribution result for the generalized Zagreb index: for $k = 3$ or $k = 4$,*

$$\frac{Z_n^{(k)}}{n^{k/2}} \xrightarrow{d} Z^{(k)},$$

where $Z^{(k)}$ is a random variable that is uniquely characterized by the moment sequence $\{g_{r,0}\}_{r=0}^\infty$ which satisfies (27) and (28).

Remark 4.2. (i) For $k = 2$, a different normalization has to be used; see Remark 4.9.

(ii) For $k \geq 5$, the moment-transfer approach cannot be applied as the sequence $\{g_{r,0}\}_{r=0}^{\infty}$ grows too fast to imply that it uniquely characterizes $Z^{(k)}$; see Remark 4.8.

We again compute the asymptotics of all moments by first considering the mean and then higher moments, where for the mean we bring (1) into the following form:

$$Z_n^{(k)} \stackrel{d}{=} Z_{I_n}^{(k)} + \tilde{Z}_{n-I_n}^{(k)} + \sum_{\ell=0}^{k-1} \binom{k}{\ell} (R_{I_n}^{\ell} - (-1)^{k-\ell} R_n^{\ell}). \quad (17)$$

Note that this differs from (6) since we have used (2) to replace the last two terms of (1) by $-(R_n - 1)^k + R_n^k$ before expanding by the binomial theorem; this change will make the computation of the (asymptotic) mean easier. For higher moments we, however, again use (6).

In addition, we state a technical lemma which is needed below.

Lemma 4.3. For $u, v \geq 0$, we have

$$\sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} j^{u/2} (n-j)^{v/2} = \begin{cases} \mathcal{O}(n^{v/2}), & \text{if } u = 0; \\ \mathcal{O}(n^{v/2} \log n), & \text{if } u = 1; \\ \mathcal{O}(n^{(u+v-1)/2}), & \text{if } u \geq 2. \end{cases} \quad (18)$$

Proof. First recall the following asymptotic result for the Catalan numbers:

$$C_n \sim \frac{4^{n-1}}{\sqrt{\pi n^3}}, \quad (n \rightarrow \infty). \quad (19)$$

By plugging this into (18), we obtain for $u \geq 2$,

$$\begin{aligned} \sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} j^{u/2} (n-j)^{v/2} &= \mathcal{O} \left(n^{1/2} \sum_{j=1}^{n-1} j^{(u-3)/2} (n-j)^{(v-1)/2} \right) \\ &= \mathcal{O} \left(n^{(u+v-1)/2} \int_0^1 x^{(u-3)/2} (1-x)^{(v-1)/2} dx \right) \end{aligned}$$

and the result follows from this since the last integral is finite. On the other hand, for $u = 0$ and $u = 1$, the integral does not exist and one has to be more careful. More precisely, in these two cases, we replace the sum after the first step by:

$$\sum_{j=1}^{n-1} j^{(u-3)/2} ((n-j)^{(v-1)/2} - n^{(v-1)/2}) + n^{(v-1)/2} \sum_{j=1}^{n-1} j^{(u-3)/2}.$$

Then, the divergence issue is resolved and we can argue as above:

$$\begin{aligned} n^{1/2} \sum_{j=1}^{n-1} j^{(u-3)/2} (n-j)^{(v-1)/2} &= \mathcal{O} \left(n^{(u+v-1)/2} \int_0^1 x^{(u-3)/2} ((1-x)^{(v-1)/2} - 1) dx + n^{v/2} \sum_{j=1}^{n-1} j^{(u-3)/2} \right) \\ &= \begin{cases} \mathcal{O}(n^{v/2}), & \text{if } u = 0; \\ \mathcal{O}(n^{v/2} \log n), & \text{if } u = 1. \end{cases} \end{aligned}$$

This proves the claimed result. $\square$

Mean Value. Taking the mean value on both sides of (17) shows that $a_n = \mathbb{E}(Z_n^{(k)})$ satisfies (4) with

$$b_n = \sum_{\ell=0}^{k-1} \binom{k}{\ell} \left(\sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} \mathbb{E}(R_j^\ell) - (-1)^{k-\ell} \mathbb{E}(R_n^\ell) \right).$$

Next, from the moment convergence in Proposition 2.4-(ii), we obtain that

$$\mathbb{E}(R_n^m) \sim \frac{m! \sqrt{\pi}}{\Gamma((m+1)/2)} n^{m/2}, \quad (m \geq 0), \quad (20)$$

and in particular, $\mathbb{E}(R_n^m) = \mathcal{O}(n^{m/2})$ for $m \geq 0$. Thus, by (18) (with $u = \ell$ and $v = 0$):

$$b_n \sim k \mathbb{E}(R_n^{k-1}) \sim \frac{k! \sqrt{\pi}}{\Gamma(k/2)} n^{(k-1)/2}$$

and consequently, by Lemma 2.9,

$$\mathbb{E}(Z_n^{(k)}) \sim \begin{cases} 2n \log n, & \text{if } k = 2; \\ \frac{2k! \sqrt{\pi}}{(k-2)\Gamma((k-1)/2)} n^{k/2}, & \text{if } k \geq 3. \end{cases} \quad (21)$$

Remark 4.4. This difference between $k = 2$ and $k \geq 3$ is the reason why Theorem 4.1 does not include the case $k = 2$.

For small k , one can also compute closed-form expressions. We demonstrate this for $k = 2$; the computation for higher values of k is similar but gets more and more involved.

First, from (2) one obtains that

$$\mathbb{E}(R_n) = \sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} \mathbb{E}(R_{n-j}) + 1, \quad (n \geq 2),$$

where $\mathbb{E}(R_1) = 0$. This is equivalent to

$$nC_n \mathbb{E}(R_n) = 2 \sum_{j=1}^{n-1} C_j (n-j) C_{n-j} \mathbb{E}(R_{n-j}) + nC_n.$$

Set

$$A(z) := \sum_{n \geq 1} nC_n \mathbb{E}(R_n) z^n.$$

Then,

$$A(z) = 2 \left(\sum_{n \geq 1} C_n z^n \right) A(z) + \sum_{n \geq 2} nC_n z^n.$$

Note that

$$\sum_{n \geq 1} C_n z^n = \frac{1 - \sqrt{1 - 4z}}{2} \quad \text{and} \quad \sum_{n \geq 2} nC_n z^n = \frac{z}{\sqrt{1 - 4z}} - z.$$

Consequently,

$$A(z) = \frac{z}{1-4z} - \frac{z}{\sqrt{1-4z}}$$

and thus,

$$\mathbb{E}(R_n) = \frac{[z^n]A(z)}{nC_n} = \frac{4^{n-1}}{nC_n} - 1, \quad (22)$$

where the last step follows from standard computations. (Here, $[z^n]f(z)$ denotes the n -th coefficient in the Maclaurin series of $f(z)$.)

Now, let us turn to $\mathbb{E}(Z_n^{(2)})$ which due to (6) satisfies

$$\mathbb{E}(Z_n^{(2)}) = 2 \sum_{j=1}^{n-1} \frac{C_j C_{n-j}}{C_n} \mathbb{E}(Z_j^{(2)}) + 2 + 4 \sum_{j=1}^{n-1} \frac{C_j C_{n-j}}{C_n} \mathbb{E}(R_j), \quad (n \geq 2)$$

with $\mathbb{E}(Z_1^{(2)}) = 0$. This can be rewritten as

$$C_n \mathbb{E}(Z_n^{(2)}) = 2 \sum_{j=1}^{n-1} C_{n-j} C_j \mathbb{E}(Z_j^{(2)}) + 2C_n + 4 \sum_{j=1}^{n-1} C_{n-j} C_j \mathbb{E}(R_j)$$

which by setting

$$B(z) := \sum_{n \geq 1} C_n \mathbb{E}(Z_n^{(2)}) z^n$$

translates into

$$B(z) = 2 \left(\sum_{n \geq 1} C_n z^n \right) B(z) + 2 \sum_{n \geq 2} C_n z^n + 4 \left(\sum_{n \geq 1} C_n z^n \right) \sum_{n \geq 1} C_n \mathbb{E}(R_n) z^n.$$

From (22), we have

$$\sum_{n \geq 1} C_n \mathbb{E}(R_n) z^n = \frac{1}{4} \log \frac{1}{1-4z} - \frac{1 - \sqrt{1-4z}}{2}.$$

Collecting everything and simplifying gives:

$$B(z) = \frac{1}{2\sqrt{1-4z}} \log \frac{1}{1-4z} - \frac{2z}{\sqrt{1-4z}} - \frac{1}{2} \log \frac{1}{1-4z} + 1 - \sqrt{1-4z}.$$

Recall that

$$[z^n] \frac{1}{\sqrt{1-z}} \log \frac{1}{1-z} = \frac{1}{4^n} \binom{2n}{n} (2H_{2n} - H_n),$$

where $H_n = \sum_{j=1}^n (1/j)$ denotes the n -th Harmonic number. Consequently,

$$\mathbb{E}(Z_n^{(2)}) = (2n-1)(2H_{2n} - H_n) - 2n - \frac{4^n}{2nC_n} + 2$$

From this, we obtain the refined asymptotics:

$$\mathbb{E}(Z_n^{(2)}) = 2n \log n + (4 \log 2 + 2\gamma - 2)n + \mathcal{O}(\sqrt{n}), \quad (n \rightarrow \infty), \quad (23)$$

where γ denotes Euler's constant.

We next consider higher moments.

Higher Moments. Here, in contrast to the non-plane case, we do not need to shift by the mean value, which simplifies the proof. We drop from now on the superscript k .

We start with an analogue of Lemma 3.3.

Lemma 4.5. *The mixed moments $\mathbb{E}(Z_n^r R_n^s)$ for $s > 0$ satisfy (3) and for $s = 0$ satisfy (4), where b_n is in both cases a function of mixed moment of smaller order with respect to the lexicographic order.*

Proof. Similar as in the proof of Lemma 3.3, we obtain that

$$\mathbb{E}(Z_n^r R_n^s) = \sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} \sum_{i_1+i_2+i_3=r} \binom{r}{i_1, i_2, i_3} \mathbb{E}(Z_j^{i_1} \tilde{Z}_{n-j}^{i_2} T_{n,j}^{i_3} (\tilde{R}_{n-j} + 1)^s), \quad (24)$$

where

$$T_{n,j} := \sum_{\ell=0}^{k-1} \binom{k}{\ell} (R_j^\ell + \tilde{R}_{n-j}^\ell).$$

The rest of the arguments are identical to those in the proof of Lemma 3.3. $\square$

From this result and induction, we obtain the following.

Proposition 4.6. *For $r, s \geq 0$, we have*

$$\mathbb{E}(Z_n^r R_n^s) \sim g_{r,s} n^{(kr+s)/2} \quad (25)$$

where $g_{r,s}$ satisfies the recurrence in the proof below.

Proof. We use induction on (r, s) with respect to the lexicographic order, where the claim holds for $r = 0$ and arbitrary s because of (20) and for $(r, s) = (1, 0)$ because of (21).

We next assume that the claim holds for all (r', s') which are lexicographically smaller than (r, s) . We want to prove it for (r, s) .

First consider $s > 0$. Then, by Lemma 4.5, we know that $\mathbb{E}(Z_n^r R_n^s)$ satisfies (3), where b_n consists of the terms on the right-hand side of (24) except the term with $i_2 = r$ and $\tilde{R}_{n-j}^s$ in the expansion of $(\tilde{R}_{n-j} + 1)^s$. By expanding, we see that these terms are a linear combination of terms

$$\mathbb{E}(Z_j^{i_1} R_j^{\ell_1}) \mathbb{E}(Z_{n-j}^{i_2} R_{n-j}^{\ell_2+\ell_3}), \quad (26)$$

where $i_1 + i_2 + i_3 = r$, ℓ_1 and ℓ_2 arise from expanding $T_{n,j}^{i_3}$ and thus $\ell_1 + \ell_2 \leq i_3(k-1)$, and ℓ_3 arises from expanding $(\tilde{R}_{n-j} + 1)^s$ and thus $\ell_3 \leq s$. We plug the induction hypothesis into this and then apply Lemma 4.3. Note that whenever $i_1 + i_2$ is reduced by 1, we lose a factor of k in front of r in the exponent of (25). On the other hand, from $\ell_1 + \ell_2 \leq i_3(k-1)$, we see that we can gain at most a term $k-1$. Thus, by a careful analysis of the terms (26), we see that the following terms dominate:

- (i) $i_1 + i_2 = r$ with $i_2 < r$. Then, $i_3 = 0$ and thus $\ell_1 = \ell_2 = 0$. Moreover, $\ell_3 = s$.
- (ii) $i_2 = r$ and thus $i_1 = i_3 = 0$ which in turn implies that $\ell_1 = \ell_2 = 0$. Moreover, $\ell_3 = s-1$. (Recall that $\ell_3 = s$ is not allowed as this term was removed from b_n .)
- (iii) $i_2 = r-1$ and $i_1 = 0$, $\ell_1 = 0$ which implies that $\ell_2 = k-1$ and $\ell_3 = s$.

In all remaining cases, the term (26) contributes $\mathcal{O}(n^{(kr+s)/2-1})$. Thus, up to this error, b_n is given by:

$$\sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} \left(\sum_{\ell=1}^r \binom{r}{\ell} \mathbb{E}(Z_j^\ell) \mathbb{E}(Z_{n-j}^{r-\ell} R_{n-j}^s) + kr \mathbb{E}(Z_{n-j}^{r-1} R_{n-j}^{k+s-1}) + s \mathbb{E}(Z_{n-j}^r R_{n-j}^{s-1}) \right).$$

We consider the three terms in this expression separately. For the first one, by (19) and the induction hypothesis:

$$\begin{aligned} & \sum_{\ell=1}^r \binom{r}{\ell} \sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} \mathbb{E}(Z_j^\ell) \mathbb{E}(Z_{n-j}^{r-\ell} R_{n-j}^s) \\ & \sim \sum_{\ell=1}^r \binom{r}{\ell} g_{\ell,0} g_{r-\ell,s} \frac{\sqrt{n}}{2\sqrt{\pi}} \sum_{j=1}^{n-1} j^{(k\ell-3)/2} (n-j)^{(k(r-\ell)+s-1)/2} \\ & \sim \frac{n^{(kr+s-1)/2}}{2\sqrt{\pi}} \sum_{\ell=1}^r \binom{r}{\ell} g_{\ell,0} g_{r-\ell,s} \int_0^1 x^{(k\ell-3)/2} (1-x)^{(k(r-\ell)+s-1)/2} dx \\ & = \frac{n^{(kr+s-1)/2}}{2\sqrt{\pi} \Gamma((kr+s)/2)} \sum_{\ell=1}^r \binom{r}{\ell} g_{\ell,0} g_{r-\ell,s} \Gamma\left(\frac{k\ell-1}{2}\right) \Gamma\left(\frac{k(r-\ell)+s+1}{2}\right). \end{aligned}$$

Next, consider the second term

$$\begin{aligned} kr \sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} \mathbb{E}(Z_{n-j}^{r-1} R_{n-j}^{k+s-1}) & \sim 2kr g_{r-1,k+s-1} \sqrt{n} \sum_{j=1}^{n-1} C_j 4^{-j} (n-j)^{(kr+s)/2-1} \\ & \sim 2kr g_{r-1,k+s-1} n^{(kr+s-1)/2} \sum_{j \geq 1} C_j 4^{-j} \\ & \sim kr g_{r-1,k+s-1} n^{(kr+s-1)/2}, \end{aligned}$$

where the last step follows from

$$\sum_{j \geq 1} C_j 4^{-j} = \frac{1}{2}.$$

Likewise,

$$\begin{aligned} s \sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} \mathbb{E}(Z_{n-j}^r R_{n-j}^{s-1}) & \sim 2s g_{r,s-1} n^{(kr+s-1)/2} \sum_{j \geq 1} C_j 4^{-j} \\ & \sim s g_{r,s-1} n^{(kr+s-1)/2}. \end{aligned}$$

Using Lemma 2.6-(ii) gives now the claimed result with

$$\begin{aligned} g_{r,s} &= \frac{1}{2\sqrt{\pi} \Gamma((kr+s+1)/2)} \sum_{\ell=1}^r \binom{r}{\ell} g_{\ell,0} g_{r-\ell,s} \Gamma\left(\frac{k\ell-1}{2}\right) \Gamma\left(\frac{k(r-\ell)+s+1}{2}\right) \\ & \quad + \frac{\Gamma((kr+s)/2)}{\Gamma((kr+s+1)/2)} (kr g_{r-1,k+s-1} + s g_{r,s-1}), \quad (s > 0). \end{aligned} \quad (27)$$

Finally, for $s = 0$, the arguments are similar. First, by Lemma 4.5, $\mathbb{E}(Z_n^r R_n^s)$ satisfies (4) with

$$b_n = \sum_{j=1}^{n-1} \frac{2(n-j)C_j C_{n-j}}{nC_n} \left(\sum_{\ell=1}^{r-1} \binom{r}{\ell} \mathbb{E}(Z_j^\ell) \mathbb{E}(Z_{n-j}^{r-\ell}) + kr \mathbb{E}(Z_{n-j}^{r-1} R_{n-j}^{k-1}) \right) + \mathcal{O}(n^{(kr+s)/2-1}).$$

Thus, arguing as above, we obtain the claim with

$$\begin{aligned} g_{r,0} &= \frac{1}{\sqrt{\pi}(kr-2)\Gamma((kr-1)/2)} \sum_{\ell=1}^{r-1} \binom{r}{\ell} g_{\ell,0} g_{r-\ell,0} \Gamma\left(\frac{k\ell-1}{2}\right) \Gamma\left(\frac{k(r-\ell)+1}{2}\right) \\ &\quad + \frac{kr\Gamma(kr/2-1)}{\Gamma((kr-1)/2)} g_{r-1,k-1}. \end{aligned} \quad (28)$$

This concludes the proof. $\square$

Before we prove the main result of this section, we still need another technical lemma.

Lemma 4.7. *For $k = 3$ or $k = 4$, there is a unique random vector (Z, R) whose (r, s) -th mixed moment equals $g_{r,s}$.*

Proof. In the appendix, we prove that there is an absolute constant A such that

$$g_{r,s} \leq A^{kr+s} \sqrt{(kr+s)!}, \quad (r, s \geq 0). \quad (29)$$

Now, in order to establish the result, we only need to show that the two marginal sequences $\{g_{r,0}\}_{r=0}^\infty$ and $\{g_{0,s}\}_{s=0}^\infty$ uniquely characterize distributions. This is clear for the latter, as it is the (unique) moments sequence of a Rayleigh distribution.

As for the former, we use Carleman's condition for the Stieltjes moment problem which means that we have to check the divergence of $\sum_{r=0}^\infty g_{r,0}^{-1/(2r)}$:

$$\sum_{r=0}^\infty g_{r,0}^{-1/(2r)} \geq A^{-k/2} \sum_{r=0}^\infty (kr)!^{-1/(4r)}.$$

The series on the right hand side diverges if and only if $k = 3$ or $k = 4$. $\square$

Remark 4.8. It is not hard to see that the estimate $g_{r,0} \leq A^{kr} \sqrt{(kr)!}$ is sharp up to the base of the exponential growth term and thus the series $\sum_{r=0}^\infty g_{r,0}^{-1/(2r)}$ is indeed convergent for all $k \geq 5$. As a consequence, it is not clear how to apply the method of moments for $k \geq 5$.

Proof of Theorem 4.1. From Proposition 4.6 and Lemma 4.7, we can conclude that for $k = 3$ or $k = 4$:

$$\left(\frac{Z_n}{n^{k/2}}, \frac{R_n}{n^{1/2}} \right) \xrightarrow{d} (Z, R).$$

From this the result follows. $\square$

Remark 4.9. For $k = 2$, a similar result holds but with a different normalization. More precisely, in this case, we have to shift the mean (see (23)) as in Section 3 and the result becomes:

$$\frac{Z_n^{(2)}}{n} - 2 \log n - (4 \log 2 + 2\gamma - 2) \xrightarrow{d} Z^{(2)}, \quad (n \rightarrow \infty),$$

where $Z^{(2)}$ is a random variable which is again uniquely characterized by its moment sequence. This can either be proved with the method from this section, or more easily with martingale arguments; see [24] where these arguments are used for a closely related random model.

5 Conclusion

Due to its importance in chemistry, the first Zagreb index has been investigated for several random tree models and limit laws have been proved with martingale techniques and Stein's method; see, e.g., [9] and [24]. These techniques do not seem to work for the generalized Zagreb index. In this paper, we used another approach, namely, the *moment-transfer approach* to investigate the generalization.

We concentrated on random non-plane and plane recursive trees for which *transfer results*, on which the moment-transfer approach rests, have been established in previous work [18,21]. For random non-plane trees, we proved that the the generalized Zagreb index has linear mean and variance, and (when properly normalized) satisfies a central limit theorem regardless of the value of k (the main parameter of the generalized Zagreb index; $k = 2$ is the classical case). On the other hand, for random plane recursive trees, the mean grows with k and we proved that the limit law is non-normal for $2 \leq k \leq 4$. More precisely, we derived the asymptotic moment sequence for all $k \geq 3$ and showed that it characterizes the limit law for $k = 3$ and $k = 4$. For the mean, our result on the moments also includes the case $k = 2$ for which we have a closed-form expression; it corrects a previous result from [24] which was stated for random plane recursive trees even though the author investigated a slightly different random model. In addition, we commented on the limit law of the case $k = 2$ which requires a different normalization; see Remark 4.9. In this case, our method of proof works, too, but the result is more easily proved by using martingale techniques (that is why we did not provided details).

Acknowledgments

This research was initiated at the 2024 Institute for Mathematical Statistics – Asia-Pacific Rim Meeting (IMS–APRM 2024) which was held in Melbourne, Australia. A large part of the technical work was done while the second author was visiting the first author at the University of Science and Technology of China (USTC) in Hefei, China. He thanks the first author and USTC for hospitality. The second and third author were partially supported by the National Science and Technology Council (NSTC), Taiwan under grants NSTC-113-2115-M-004-004-MY3 (MF) and NSCT-114-2115-M-030-006 (TCY).

References

- [1] A. Ali, I. Gutman, E. Milovanović, and I. Milovanović (2018). Sum of powers of the degrees of graphs: extremal results and bounds, *MATCH Commun. Math. Comput. Chem.*, **80**, 5–84.
- [2] L. Bedratyuk and O. Savenko (2018). The star sequence and the general first Zagreb index, *MATCH Commun. Math. Comput. Chem.*, **79**, 407–414.
- [3] P. Billingsley (1995). *Probability and Measure*, 3rd edition, John Wiley & Sons, New York.
- [4] B. Chaluvvaraju, H. S. Boregowda, and I. N. Cangul (2021). Some inequalities for the first general Zagreb index of graphs and line graphs, *Natl. Acad. Sci.*, **91**, 79–88.

- [5] C.-Y. Chen and M. Fuchs (2011). On the moment-transfer approach for random variables satisfying a one-sided distributional recurrence, *Electron. J. Probab.*, **16**, 903–928.
- [6] J. Devillers and A. T. Balaban (Eds.). (1999). *Topological Indices and Related Descriptors in QSAR and QSPR*, CRC Press, London.
- [7] T. Došlić, M. A. Hosseinzadeh, S. Hossein-Zadeh, A. Iranmanesh, and F. Rezakhanlou (2020). On generalized Zagreb indices of random graphs. *MATCH Commun. Math. Comput. Chem.*, **84**, 499–511.
- [8] M. Drmota (2019). *Random Trees: An Interplay between Combinatorics and Probability*, 1st edition, Springer.
- [9] Q. Feng and Z. Hu (2011). On the Zagreb index of random recursive trees, *J. Appl. Prob.*, **48**, 1189–1196.
- [10] Q. Feng and Z. Hu (2013). Phase changes in the topological indices of scale-free trees, *J. Appl. Prob.*, **50**, 516–532.
- [11] Q. Feng, Z. Hu, and C. Su (2013). The Zagreb indices of random graphs. *Probab. Eng. Inf. Sci.*, **27**, 247–260.
- [12] Q. Feng, H. Ren, and Y. Tian (2025). Limit laws for the generalized Zagreb indices of random graphs, *Adv. Appl. Probab.*, accepted for publication.
- [13] B. Furtul and I. Gutman (2015). A forgotten topological index, *J. Math. Chem.*, **53**, 1184–1190.
- [14] I. Gutman and K. C. Das (2004). The first Zagreb index 30 years after, *MATCH Commun. Math. Comput. Chem.*, **50**, 83–92.
- [15] I. Gutman, and N. Trinajstić (1972). Graph theory and molecular orbitals. Total φ -electron energy of alternant hydrocarbons, *Chem. Phys. Lett.*, **17**, 535–538.
- [16] J. Harant, S. Jendrol and T. Madaras (2011). Upper bounds on the sum of powers of the degrees of a simple planar graph, *J. Graph Theory*, **67**, 112–123.
- [17] H. Hua and H. Deng (2007). On unicycle graphs with maximum and minimum zeroth-order general Randić indices, *J. Math. Chem.*, **41**, 173–181.
- [18] H.-K. Hwang (2006). Profiles of random trees: Plane-oriented recursive trees, *Random Struc. Algor.*, **30:3**, 380–413.
- [19] X. Li and H. Zhao (2004). Trees with the first three smallest and largest generalized topological indices, *MATCH Commun. Math. Comput. Chem.*, **50**, 57–62.
- [20] X. Li and J. Zheng (2005). A unified approach to the extremal trees for different indices, *MATCH Commun. Math. Comput. Chem.*, **54**, 195–208.
- [21] R. Neininger and H.-K. Hwang (2002). Phase change of limit laws in the quicksort recurrence under varying toll functions, *SIAM J. Comput.*, **31:6**, 1687–1722.

- [22] R. Todeschini, and V. Consonni (2009). *Molecular Descriptors for Chemoinformatics*, Wiley, New York.
- [23] K. Xu, K. Tang, H. Liu, and J. Wang (2015). The Zagreb indices of bipartite graphs with more edges, *J. Appl. Math. Inform.*, **33**, 365–377.
- [24] P. Zhang (2020). On several properties of a class of preferential attachment trees—plane-oriented recursive trees, *Probab. Eng. Inf. Sci.*, **35**, 839–857.
- [25] S. Zhang and H. Zhang (2006). Unicyclic graphs with the first three smallest and largest first general Zagreb index, *MATCH Commun. Math. Comput. Chem.*, **55**, 427–438.

Appendix

In this appendix we prove (29).

Proposition 5.1. *For the sequence $\{g_{r,s}\}_{r,s=0}^{\infty}$ from Proposition 4.6, we have the bound*

$$g_{r,s} \leq A^{kr+s} \sqrt{(kr+s)!},$$

where A is a suitable large constant.

We first need two lemmas.

Lemma 5.2. *For all integers $a, b \geq 1$, we have*

$$\frac{\Gamma(a/2)\Gamma(b/2)}{\sqrt{\pi}\Gamma((a+b-1)/2)} \leq 1.$$

Proof. We prove this by induction on $a+b = k$. Note that the claim is true for $k = 2$ and $k = 3$. Assume that it holds for $k' < k$. We will show it for $a+b = k$. By the functional equation for the Gamma function, we have

$$\frac{(a/2-1)\Gamma((a-2)/2)\Gamma(b/2)}{\sqrt{\pi}((a+b-1)/2-1)\Gamma((a+b-3)/2)} = \frac{a-2}{a-2+b-1} \cdot \frac{\Gamma((a-2)/2)\Gamma(b/2)}{\sqrt{\pi}\Gamma((a+b-3)/2)} \leq 1,$$

where the first factor is trivially bounded by 1 and the second factor is bounded by 1 because of the induction hypothesis. $\square$

Lemma 5.3. *For $k \geq 2$, we have*

$$\binom{r}{\ell}^2 \leq \binom{kr+s}{k\ell}, \quad (0 \leq \ell \leq r),$$

where $r, s \geq 0$ are integers.

Proof. Partition a set of kr objects into k blocks which all have size r . Then, clearly

$$\binom{r}{\ell}^k \leq \binom{kr}{k\ell}$$

as selecting ℓ elements from each block gives a selection of $k\ell$ elements from the original set. Thus,

$$\binom{r}{\ell}^2 \leq \binom{r}{\ell}^k \leq \binom{kr}{k\ell} \leq \binom{kr+s}{k\ell}$$

as claimed. $\square$

Proof of Proposition 5.1. We show the claimed bound by using induction on (r, s) which we equip with the lexicographic order.

First, for $r = 0$, we have

$$g_{0,s} = \frac{s! \sqrt{\pi}}{\Gamma((s+1)/2)}$$

as these are the moments of the Rayleigh distribution with parameter $\sqrt{2}$. Thus, the claim for the induction base follows by Stirling's formula for the Gamma function.

Next, we assume that the claim holds for (r', s') which are (lexicographically) smaller than (r, s) . We show that it holds for (r, s) by using the recurrences (27) and (28) where we concentrate on the former as details for the later are similar (and even easier as one term of (27) is missing in (28)).

We can break the right-hand side of (27) into three parts according to its three terms. First, for the first part:

$$\begin{aligned} & \frac{1}{2\sqrt{\pi}\Gamma((kr+s+1)/2)} \sum_{\ell=1}^r \binom{r}{\ell} g_{\ell,0} g_{r-\ell,s} \Gamma\left(\frac{k\ell-1}{2}\right) \Gamma\left(\frac{k(r-\ell)+s+1}{2}\right) \\ & \leq \frac{1}{kr+s-1} \sum_{\ell=1}^r \binom{r}{\ell} g_{\ell,0} g_{r-\ell,s} \\ & \leq \frac{A^{kr+s}}{kr+s-1} \sum_{\ell=1}^r \sqrt{\binom{r}{\ell}^2 (k\ell)!(k(r-\ell)+s)!} \\ & \leq \frac{r}{kr+s-1} A^{kr+s} \sqrt{(kr+s)!} \leq \frac{1}{3} A^{kr+s} \sqrt{(kr+s)!}, \end{aligned}$$

where we used Lemma 5.2 for the first estimate, the induction hypothesis for the second estimate, and Lemma 5.3 for the second last estimate. For the second and third part, recall that for $x \geq 1$,

$$\frac{\Gamma(x/2)}{\Gamma((x+1)/2)} \leq 2x^{-1/2}.$$

Thus,

$$\begin{aligned} & \frac{\Gamma((kr+s)/2)}{\Gamma((kr+s+1)/2)} (kr g_{r-1,k+s-1} + s g_{r,s-1}) \\ & \leq 2(kr+s)^{-1/2} \left(kr A^{kr+s-1} \sqrt{(kr+s-1)!} + s A^{kr+s-1} \sqrt{(kr+s-1)!} \right) \\ & \leq \frac{2}{A} A^{kr+s} \sqrt{(kr+s)!} \leq \frac{2}{3} A^{kr+s} \sqrt{(kr+s)!}, \end{aligned}$$

where the last step follows since we can assume that $A \geq 3$.

Combining the above two estimates for the first part and the second and third part completes the induction step and hence the proof. $\square$