

# HORTON LAW IN SELF-SIMILAR TREES

YEVGENIY KOVCHEGOV<sup>\*,‡</sup> and ILYA ZALIAPIN<sup>†,§</sup>

*\*Department of Mathematics  
Oregon State University  
Corvallis, OR, 97331, USA*

*†Department of Mathematics and Statistics  
University of Nevada  
Reno, NV 89557, USA*

*‡kovchegy@math.oregonstate.edu  
§zal@unr.edu*

Received October 26, 2015

Accepted January 7, 2016

Published April 8, 2016

## Abstract

Self-similarity of random trees is related to the operation of pruning. Pruning  $\mathcal{R}$  cuts the leaves and their parental edges and removes the resulting chains of degree-two nodes from a finite tree. A Horton–Strahler order of a vertex  $v$  and its parental edge is defined as the minimal number of prunings necessary to eliminate the subtree rooted at  $v$ . A branch is a group of neighboring vertices and edges of the same order. The Horton numbers  $\mathcal{N}_k[K]$  and  $\mathcal{N}_{ij}[K]$  are defined as the expected number of branches of order  $k$ , and the expected number of order- $i$  branches that merged order- $j$  branches,  $j > i$ , respectively, in a finite tree of order  $K$ . The Tokunaga coefficients are defined as  $T_{ij}[K] = \mathcal{N}_{ij}[K]/\mathcal{N}_j[K]$ . The pruning decreases the orders of tree vertices by unity. A rooted full binary tree is said to be mean-self-similar if its Tokunaga coefficients are invariant with respect to pruning:  $T_k := T_{i,i+k}[K]$ . We show that for self-similar trees, the condition  $\limsup_{k \rightarrow \infty} T_k^{1/k} < \infty$  is necessary and sufficient for the existence of the strong Horton law:  $\mathcal{N}_k[K]/\mathcal{N}_1[K] \rightarrow R^{1-k}$ , as  $K \rightarrow \infty$  for some  $R > 0$  and every  $k \geq 1$ .

---

<sup>‡</sup>Corresponding author.

This work is a step toward providing rigorous foundations for the Horton law that, being omnipresent in natural branching systems, has escaped so far a formal explanation.

*Keywords:* Random Trees; Horton Law; Tokunaga Self-Similarity.

## 1. INTRODUCTION

Horton laws, which are akin to a power-law distribution of the element sizes in a branching system, epitomize the scale invariance of natural dendritic structures. It is very intuitive that the existence of Horton laws should be related to the self-similar organization of branching, defined in suitable terms. Such relation, however, has escaped a rigorous explanation, remaining for long time a part of science literature folklore.<sup>1,2</sup> This paper shows that a weak (mean) invariance under the operation of tree pruning is sufficient for the Horton law of branch numbers to hold in the strongest sense, hence explaining and unifying many earlier empirical observations and partial results in this direction.

We work with binary trees, although our results can be easily extended to the case of trees of a higher degree. Recall that *pruning* of a finite rooted full binary tree  $T$  cuts its leaves (vertices of degree one) and their parental edges, and removes the resulting chains of degree-two vertices and their parental edges (so-called series reduction). The *Horton–Strahler order*  $k$  of vertex  $v$  is the minimal number of prunings necessary to eliminate the subtree rooted at  $v$ . A *branch* is a sequence of neighboring vertices and their parental edges of the same order. We write  $N_k$  for the total number of branches of order  $k$  in a tree. A common empirical observation in the natural dendritic structures is that  $N_{k+1}/N_k \approx R$ ,  $3 \leq R \leq 5$ . This regularity was first described by Horton<sup>3,4</sup> in a study of river streams; it has been strongly corroborated in hydrology<sup>1,5–9</sup> and expanded to biology and other areas<sup>10</sup> since then. Similar relations, referred to as *Horton laws*, are reported for selected metric quantities, for example the average lengths of river streams ( $l_{k+1}/l_k \approx R_l$ ), average contributing areas of order- $k$  drainage basin ( $A_{k+1}/A_k \approx R_A$ ), etc.

Informally, Horton laws suggest that the branch order  $k$  is proportional to the logarithm of a suitably defined “size”  $S_k$  of the branch:  $S_k \propto B^k$ . A geometric distribution of the branch counts  $N_k$ ,

$$P(\text{a random branch has order } \geq k) = R^{-k}$$

is equivalent to a power-law distribution of branch sizes:

$$P(S_k \geq x) \propto x^{-\alpha}, \quad \alpha = \ln(R)/\ln(B).$$

Hence, the empirical Horton laws can be interpreted as a power-law distribution of system element sizes. This might hint at a scale-invariant organization of the respective branching structures, as power laws often accompany fractality.

For long time, the only rigorous result on validity of Horton laws was that of Shreve,<sup>11</sup> who demonstrated that in a uniform distribution of rooted binary trees with  $n$  leaves (that he called *topologically random networks*), the ratio  $N_{k+1}/N_k$  converges to 4 as  $n$  goes to infinity. This model is equivalent to the critical binary Galton–Watson tree conditioned to have  $n$  leaves.<sup>12,13</sup> Shreve<sup>26</sup> also showed that in a topologically random network the average number  $T_{ij}$  of side-branches of order  $i$  per branch of order  $j$  only depends on the relative ordering of the branches:  $T_{ij} = 2^{j-i-1}$ , as the tree size increases. We notice that pruning decreases the order of every branch by unity. Accordingly, Shreve’s result implies, in particular, that the average numbers  $T_{ij}$  are invariant under the pruning operation:  $T_{ij} = T_{i-1,j-1}$ . The topologically random network was hence the first example of a model that obeys both the Horton law of branch numbers and structural invariance with respect to pruning. The invariance with respect to pruning is called *self-similarity*, and may refer to the invariance of distributions, or the means of selected statistics (like is the case with Shreve’s result).

Tokunaga<sup>14</sup> introduced a broader class of mean-invariant models defined by the constraint  $T_{ij} = T_{j-i} = a c^{j-i-1}$  for positive  $a, c$ . The validity of the Tokunaga constraint has been empirically confirmed in numerous observed and modeled systems,<sup>7,10,15–17</sup> notably including diffusion limited aggregation,<sup>15,18</sup> and two-dimensional site percolation.<sup>19–21</sup> Furthermore, Burd *et al.*<sup>13</sup> demonstrated that the Tokunaga constraint with  $(a, c) = (1, 2)$  is the characteristic property of critical binary offspring distribution within the class of Galton–Watson (non-necessarily binary) trees, and that the

critical binary Galton–Watson trees are also distributionally invariant with respect to pruning. Zaliapin and Kovchegov<sup>22</sup> have shown that both Horton law with  $R = 4$  and the Tokunaga constraint with  $(a, c) = (1, 2)$  hold in a level-set tree representation of a symmetric random walk, and that in general such a tree is not equivalent to the critical binary Galton–Watson model.

McConnell and Gupta<sup>23</sup> have shown that the Tokunaga constraint is sufficient for a Horton law. Specifically, they proved that if the sequence of branch counts  $N_k$  is related to the Tokunaga coefficients  $T_k = ac^{k-1}$  via the recursive counting equation

$$N_k = 2N_{k+1} + \sum_{j=1}^{K-j} T_j N_{k+j},$$

$$1 \leq k \leq K-1, \quad K \geq 2, \quad (1)$$

then  $N_{k+1}/N_k \rightarrow R$  for any  $k \geq 1$  in the limit of large-order trees. In this case

$$R = \frac{2 + a + c + \sqrt{(2 + a + c)^2 - 8c}}{2}, \quad (2)$$

which was reported earlier (under the explicit assumption that Horton law holds) by Tokunaga,<sup>14</sup> Peckham,<sup>6</sup> and others.

Equation (2) suggests that different Horton exponents  $R$  can be easily attained by using the Tokunaga side-branching with different pairs  $(a, c)$ . At the same time, most of the existing rigorous results on the Horton laws in “natural” models (not formulated explicitly in terms of Horton branch counting) refer to the models equivalent to the Galton–Watson critical binary tree or its slight ramifications, with  $R = 4$ , or to trees with no side-branching and  $R = 2$ . Recently, the authors established<sup>24</sup> a weak version of the Horton law for the tree that describes the celebrated Kingman’s coalescent process; this system has  $R = 3.043827\dots$

This study expands the sufficient conditions for the Horton law (in its strong version defined in Sec. 2.4) to all sequences of side-branch coefficients such that  $\limsup_k T_k^{1/k} < \infty$ . We also show that this condition is necessary in the class of mean self-similar trees. The Horton exponent in this case is given by  $R = 1/w_0$ , where  $w_0$  is the only real root of

$$\hat{t}(z) = -1 + 2z + \sum_{k=1}^{\infty} z^k T_k$$

within the interval  $[0, 1/2]$ , which was conjectured by Peckham.<sup>6</sup> The results are obtained in a probabilistic setting and refer to the expectations of branch counts with respect to a probability measure on the space of finite rooted full binary trees of Horton–Strahler order  $K$ , as  $K$  increases. This setup allows us to relax the assumption of similar statistical structure of side-branching within each branch, which is a typical assumption in the studies of Horton laws and tree self-similarity.<sup>6,23</sup>

We start by reviewing the essential definitions in Sec. 2. Section 3 introduces self-similar trees in a probabilistic setting, and establishes the equivalence of prune-invariance and the constraint  $T_{ij} = T_{j-i}$ . The main results are proven in Sec. 4.

## 2. PRELIMINARIES

### 2.1. Rooted Trees

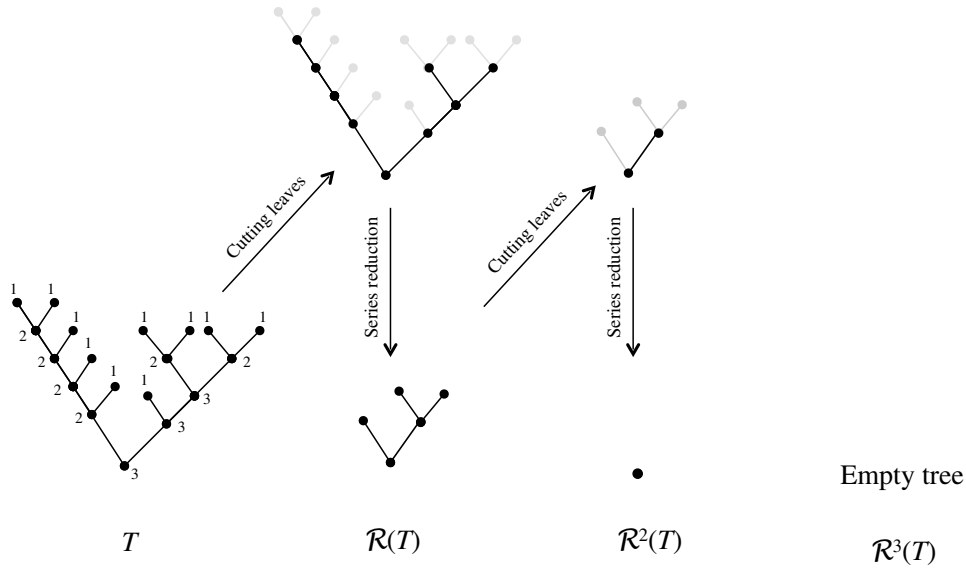
Recall that a simple graph is a collection of vertices connected by edges in such a way that each pair of vertices may have at most one connecting edge and there is no self-loops. A *tree* is a connected simple graph without cycles. In a *rooted tree*, one node is designated as a root; this imposes the parent–child relationship between the neighbor vertices. Specifically, of the two neighbor vertices the one closest to the root is called *parent*, and the other — *child*. In a rooted tree, each non-root vertex has the unique *parental* edge that connects this vertex to its parent. A *leaf* is a vertex with no children. The space of finite unlabeled rooted full binary trees, including the empty tree  $\phi$ , is denoted by  $\mathcal{T}$ . All internal vertices in a tree from  $\mathcal{T}$  have degree 3, leaves have degree 1, and the root has degree 2.

### 2.2. Tree Pruning

*Pruning* of a tree is an onto function  $\mathcal{R} : \mathcal{T} \rightarrow \mathcal{T}$ , whose value  $\mathcal{R}(T)$  for a tree  $T \neq \phi$  is obtained by removing the leaves and their parental edges from  $T$ , and then compressing the resulting tree from  $\tilde{T}$  by removing all degree-two chains (this operation is known as *series reduction*). We also set  $\mathcal{R}(\phi) = \phi$ .

### 2.3. Horton–Strahler Orders

The Horton–Strahler ordering of the vertices and edges of a finite rooted binary tree  $T \in \mathcal{T}$  is related to the iterations  $\mathcal{R}^k$  of the pruning operation.<sup>4,6,25</sup> Specifically, a vertex  $v \in T$  and its parental edge have order  $k = 1, 2, \dots$  if the subtree  $\tau_v \in T$



**Fig. 1** Example of pruning and Horton–Strahler ordering. The Horton–Strahler orders are shown next to each vertex of the initial tree  $T$ . The figure shows the two stages of each pruning – cutting the leaves (top row), and consecutive series reduction (bottom row). The order of the tree is  $k(T) = 3$  with  $N_1 = 10$ ,  $N_2 = 3$ ,  $N_3 = 1$ , and  $N_{1,2} = 3$ ,  $N_{1,3} = 1$ ,  $N_{2,3} = 1$ .

rooted at  $v$  is eliminated during the  $k$ th iteration of pruning:

$$k(v) = \min_{k \geq 1} (\mathcal{R}^k(\tau_v) = \phi).$$

The order  $k(T)$  of a non-empty tree coincides with the maximal order of its vertices. We also set  $k(\phi) = 0$ . A *branch* is defined as a union of neighboring vertices and edges of the same order. Figure 1 illustrates the operation of pruning and the definition of Horton–Strahler orders.

Equivalently, the Horton–Strahler ordering can be done by hierarchical counting.<sup>6,10,13</sup> In this approach, each leaf is assigned order  $k(\text{leaf}) = 1$ . An internal vertex  $p$  whose children have orders  $i$  and  $j$  is assigned the order

$$k(p) = \max(i, j) + \delta_{ij},$$

where  $\delta_{ij}$  is the Kronecker’s delta. The parental edge of a vertex has the same order as the vertex.

## 2.4. Horton Law

Let  $\mathcal{T}_K$ ,  $K \geq 1$ , be the subspace of finite binary trees of Horton–Strahler order  $K$ . Consider a set of probability measures  $\{\mu_K\}_{K \geq 1}$ , each of which is defined on  $\mathcal{T}_K$ , and write  $\mathbb{E}_K(\cdot)$  for the mathematical expectation with respect to  $\mu_K$ . Let  $N_k = N_k[T]$  be the number of branches of order  $k$  in a tree  $T \in \mathcal{T}$ . We define the average Horton numbers, which are the

main object for our analysis:

$$\mathcal{N}_k[K] = \mathbb{E}_K(N_k), \quad 1 \leq k \leq K, \quad K \geq 1.$$

**Definition 1.** We say that a sequence of measures  $\{\mu_K\}$  satisfies a *strong Horton law* if

$$\lim_{K \rightarrow \infty} \frac{\mathcal{N}_k[K]}{\mathcal{N}_1[K]} = R^{1-k} < \infty \quad \text{for any } k \geq 1.$$

## 2.5. Tokunaga Coefficients

Let  $N_{ij} = N_{ij}[T]$  denote the number of instances when an order- $i$  branch merges with an order- $j$  branch,  $1 \leq i < j$ , in a tree  $T$ . Such branches are referred to as *side-branches* of order  $ij$ . Define the respective expectation  $\mathcal{N}_{ij}[K] = \mathbb{E}_K(N_{ij})$ . The *Tokunaga coefficients*  $T_{ij}[K]$  for subspace  $\mathcal{T}_K$  are defined as

$$T_{ij}[K] = \frac{\mathcal{N}_{ij}[K]}{\mathcal{N}_j[K]}, \quad 1 \leq i < j \leq K. \quad (3)$$

**Remark 1.** Consider a situation when every branch of order  $j$  has the same expected number  $S_{ij}$  of side-branches of order  $i < j$ . Then

$$\begin{aligned} \mathcal{N}_{ij}[K] &= \mathbb{E}_K(N_{ij}) = \mathbb{E}_K(\mathbb{E}_K(N_{ij}|N_j)) \\ &= \mathbb{E}_K(N_j S_{ij}) = S_{ij} \mathbb{E}_K(N_j) = S_{ij} \mathcal{N}_j[K], \end{aligned}$$

and hence

$$T_{ij}[K] = \frac{\mathcal{N}_{ij}[K]}{\mathcal{N}_j[K]} = S_{ij}.$$

Such framework was considered by Shreve,<sup>26</sup> Tokunaga,<sup>14</sup> Burd *et al.*<sup>13</sup> and others. Our definition (3)

includes this situation as a special case, although in general it is free of the assumption of similar statistical structure of individual branches.

### 3. SELF-SIMILAR TREES

**Definition 2.** A set of measures  $\{\mu_K\}$  on  $\{\mathcal{T}_K\}$  is called *coordinated* if  $T_{ij} := T_{ij}[K]$  for all  $K \geq 2$  and  $1 \leq i < j \leq K$ .

For a set of coordinated measures  $\{\mu_K\}$ , the Tokunaga matrix  $\mathbb{T}_K$  for any  $K$  forms a  $K \times K$  matrix

$$\mathbb{T}_K = \begin{bmatrix} 0 & T_{1,2} & T_{1,3} & \dots & T_{1,K} \\ 0 & 0 & T_{2,3} & \dots & T_{2,K} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & T_{K-1,K} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which coincides with the restriction of any larger-order Tokunaga matrix  $\mathbb{T}_M$ ,  $M > K$ , to the first  $K \times K$  entries.

**Definition 3.** A collection of coordinated probability measures  $\{\mu_K\}$  on  $\{\mathcal{T}_K\}$  is called (mean) *self-similar* if  $T_{ij} = T_{j-i}$  for some sequence  $T_k \geq 0$ ,  $k = 1, 2, \dots$ , and any  $K \geq 2$ . The elements of the sequence  $T_k$  are also referred to as Tokunaga coefficients, which does not create confusion with  $T_{ij}$ .

For a self-similar collection of measures, the matrix of Tokunaga coefficients becomes Toeplitz:

$$\mathbb{T}_K = \begin{bmatrix} 0 & T_1 & T_2 & \dots & T_{K-1} \\ 0 & 0 & T_1 & \dots & T_{K-2} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & T_1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A variety of self-similar measures can be constructed for an arbitrary sequence of Tokunaga coefficients  $T_k > 0$ ,  $k \geq 1$ . Next, we give one natural example.

**Example 1. Independent Random Attachment.** The subspace  $\mathcal{T}_1$ , which consists of a single-vertex tree, possess a trivial unity mass measure. To construct a random tree from  $\mathcal{T}_2$ , we select a discrete probability distribution  $P_{1,2}(n)$ ,  $n = 0, 1, \dots$ , with the mean value  $T_1$ . A random tree  $T \in \mathcal{T}_2$  is obtained from the single-vertex tree  $\tau_1$  of order 1 via the following two operations. First, we attach two child vertices to the only vertex of  $\tau_1$ . This creates a

tree of order 2 with no side-branches — two leaves attached to the root. Second, we draw the number  $N_{1,2}$  from the distribution  $P_{1,2}$ , and attach  $N_{1,2}$  vertices to this tree so that they form side-branches of order  $\{1, 2\}$ .

In general, to construct a random tree from  $\mathcal{T}_K$  for  $K \geq 2$  we select a set of discrete probability distributions  $P_{k,K}(n)$ ,  $k = 1, \dots, K - 1$ , with the respective mean values  $T_k$ . A random tree  $T \in \mathcal{T}_K$  is constructed in iterative fashion, starting from the single-vertex tree  $\tau_1$  and increasing its order by adding new vertices. Specifically, to construct a random tree  $\tau_k$  of order  $k \geq 2$  from a random tree  $\tau_{k-1}$  of order  $k - 1$ , we perform the following operations. First, add two new child vertices to every leaf of  $\tau_{k-1}$  hence producing a tree  $\tilde{\tau}_k$  of order  $k$  with no side-branches of order 1. Second, for each branch of order  $2 \leq j \leq k$  in  $\tilde{\tau}_k$  draw a random number  $N_{1j}$  from the distribution  $P_{j-1,K}$  and attach  $N_{1j}$  new child vertices to this branch so that they form side-branches of order 1. Each new vertex is attached in random order with respect to the existing side-branches. Specifically, we notice that  $s \geq 0$  side-branches attached to a branch of order  $j$  are uniquely associated with  $s + 1$  edges within this branch. (When discussing the single branch of the maximal order  $k$ , we count one “imaginary” edge parental to the tree root.) The attachment of the new  $N_{1j}$  vertices among the  $s + 1$  edges is given by the equiprobable multinomial distribution with  $s + 1$  categories and  $N_{1j}$  trials.

According to Remark 1, the self-similarity condition  $T_{i,i+k}[K] = T_k$  holds within each subspace  $\mathcal{T}_K$ ,  $K \geq 2$ .

Notice that pruning defines a down-shift of the order subspaces, that is for  $K \geq 1$

$$\mathcal{R}(\mathcal{T}_K) = \mathcal{T}_{K-1}.$$

Moreover, pruning decreases the Horton–Strahler order of each vertex (and hence of each branch) by unity; in particular

$$N_k[T] = N_{k-1}[\mathcal{R}(T)], \quad k \geq 2, \tag{4}$$

$$N_{ij}[T] = N_{i-1,j-1}[\mathcal{R}(T)], \quad 2 \leq i < j. \tag{5}$$

This shift property allows us to establish connection between the values of Tokunaga coefficients for different orders  $K$ . Specifically, consider measure  $\mu_K^{\mathcal{R}}$  induced on  $\mathcal{T}_K$  by the pruning operator:

$$\mu_K^{\mathcal{R}}(A) = \mu_{K+1}(\mathcal{R}^{-1}(A)) \quad \forall A \subset \mathcal{T}_K.$$

The Tokunaga coefficients computed on  $\mathcal{T}_K$  using the induced measure  $\mu_K^{\mathcal{R}}$  are denoted by  $T_{ij}^{\mathcal{R}}[K]$ .

**Definition 4.** A collection of coordinated probability measures  $\{\mu_K\}$  on  $\{\mathcal{T}_K\}$  is called *self-similar* if  $T_{ij}[K] = T_{ij}^{\mathcal{R}}[K]$  for any  $K \geq 2$  and all  $1 \leq i < j \leq K$ .

**Lemma 1.** *The Definitions 3 and 4 are equivalent.*

**Proof.** The pruning-related index shift (4),(5) implies

$$T_{i+1,j+1}[K+1] = T_{ij}^{\mathcal{R}}[K]. \tag{6}$$

- [3  $\Rightarrow$  4]: If a coordinated set of measures  $\{\mu_K\}$  satisfy Definition 3, then

$$T_{i+1,j+1}[K+1] \stackrel{\text{Def3}}{=} T_{j-i} \stackrel{\text{Def3}}{=} T_{ij}[K+1] \stackrel{\text{coordination}}{=} T_{ij}[K].$$

Together with (6), this implies

$$T_{ij}[K] = T_{ij}^{\mathcal{R}}[K],$$

which means that Definition 4 is also satisfied.

- [4  $\Rightarrow$  3]: If a coordinated set of measures  $\{\mu_K\}$  satisfy Definition 4, then

$$T_{ij}[K] \stackrel{\text{Def4}}{=} T_{ij}^{\mathcal{R}}[K] \stackrel{\text{by(6)}}{=} T_{i+1,j+1}[K+1] \stackrel{\text{coordination}}{=} T_{i+1,j+1}[K].$$

Hence

$$T_{i+1,j+1}[K] = T_{ij}[K] = \dots = T_{1,j-i+1}[K] =: T_{j-i},$$

which means that Definition 3 is also satisfied.  $\square$

### 4. RESULTS

Consider a set of self-similar measures  $\{\mu_K\}_{K \geq 1}$  with Tokunaga coefficients  $\mathbb{T}_K$ . We define the vector of Horton indices as

$$\zeta_K = \begin{pmatrix} \mathcal{N}_1[K] \\ \mathcal{N}_2[K] \\ \vdots \\ \mathcal{N}_K[K] \end{pmatrix}.$$

We also define the vector of normalized Horton indices in  $\mathbb{R}^\infty$ ,

$$\xi_K := \frac{1}{\zeta_K(1)} \begin{pmatrix} \zeta_K \\ 0 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 \\ \mathcal{N}_2[K]/\mathcal{N}_1[K] \\ \vdots \\ \mathcal{N}_K[K]/\mathcal{N}_1[K] \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

The average number of side-branches of order  $1 \leq i < K$  within  $\mathcal{T}_K$  is  $\mathcal{N}_i[K] - 2\mathcal{N}_{i+1}[K]$ . At the same time, the number of side-branches of order  $i$  can be computed by counting the side-branches of order  $i$  for all larger-order branches:

$$\sum_{j=i+1}^K T_{ij} \mathcal{N}_j[K] = \sum_{m=1}^{K-i} T_m \mathcal{N}_{i+m}[K],$$

and therefore the vector of side-branches is  $\mathbb{T}_K \zeta_K$ . Thus

$$\mathbb{T}_K \zeta_K = \begin{bmatrix} 1 & -2 & 0 & \dots & 0 \\ 0 & 1 & -2 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & 1 & -2 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} \zeta_K. \tag{7}$$

This can also be written as

$$\mathcal{N}_k[K] = 2\mathcal{N}_{k+1}[K] + \sum_{j=1}^{K-j} T_j \mathcal{N}_{k+j}[K], \tag{8}$$

$$1 \leq k \leq K-1, \quad K \geq 2,$$

which is a probabilistic (mean) version of a deterministic counting equation (1).

Next, define

$$\mathbb{G}_K := \begin{bmatrix} -1 & T_1 + 2 & T_2 & \dots & T_{K-1} \\ 0 & -1 & T_1 + 2 & \dots & T_{K-2} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & -1 & T_1 + 2 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

which is a  $K \times K$  restriction of the following infinite dimensional linear operator to the first  $K$  dimensions:

$$\mathbb{G} := \begin{bmatrix} -1 & T_1 + 2 & T_2 & T_3 & \dots \\ 0 & -1 & T_1 + 2 & T_2 & \dots \\ 0 & 0 & -1 & T_1 + 2 & \ddots \\ 0 & 0 & 0 & -1 & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{9}$$

Equation (7) implies  $\mathbb{G}_K \zeta_K = -e_K$ , the  $K$ th coordinate basis vector, and therefore

$$\begin{pmatrix} \zeta_{K+1}(2) \\ \zeta_{K+1}(3) \\ \vdots \\ \zeta_{K+1}(K+1) \end{pmatrix} = -\mathbb{G}_K^{-1} e_K = \begin{pmatrix} \zeta_K(1) \\ \zeta_K(2) \\ \vdots \\ \zeta_K(K) \end{pmatrix}. \quad (10)$$

Thus we proved the following.

**Proposition 1.** *Let  $\{\mu_K\}$  be a set of self-similar measures on  $\{\mathcal{T}_K\}$ . Then for any  $K \geq 1$  and  $1 \leq j \leq K$ ,*

$$\mathcal{N}_{j+1}[K+1] = \zeta_{K+1}(j+1) = \zeta_K(j) = \mathcal{N}_j[K].$$

Accordingly, we also have

$$\begin{aligned} \mathcal{N}_{i+1,j+1}[K+1] &= \mathcal{N}_{ij}[K], \\ 1 \leq i < j \leq K, \quad K \geq 2. \end{aligned}$$

Observe that  $\mathbb{G}\xi_K = \frac{-1}{\zeta_K(1)}e_K$ , where, by construction,  $\zeta_K(1) = \mathcal{N}_1[K] \geq (T_1 + 2)^{K-1}$ . The following proposition formalizes the condition required for  $\lim_{K \rightarrow \infty} \xi_K = \xi$ , where  $\xi$  satisfies  $\mathbb{G}\xi = 0$  with coordinates  $\xi(j) = R^{1-j}$ . Finally, Theorem 1 at the end of this section provides a complete analysis of  $\lim_{K \rightarrow \infty} \xi_K$  in terms of the sequence  $T_j$  of Tokunaga coefficients.

**Proposition 2.** *Let  $\{\mu_K\}$  be a set of self-similar measures on  $\{\mathcal{T}_K\}$ . Suppose that the limit*

$$R = \lim_{K \rightarrow \infty} \frac{\zeta_{K+1}(1)}{\zeta_K(1)} = \lim_{K \rightarrow \infty} \frac{\mathcal{N}_1[K+1]}{\mathcal{N}_1[K]} \quad (11)$$

*exists and is finite. Then, the strong Horton law holds; that is, for each positive integer  $j$*

$$\xi(j) = \lim_{K \rightarrow \infty} \xi_K(j) = \lim_{K \rightarrow \infty} \frac{\mathcal{N}_j[K]}{\mathcal{N}_1[K]} = R^{1-j}.$$

*Conversely, if the limit (11) does not exist, then neither will  $\lim_{K \rightarrow \infty} \xi_K(j)$ . That is, the limit  $\lim_{K \rightarrow \infty} \xi_K(j)$  does not exist at least for some  $j$ .*

**Proof.** Suppose that the limit  $R = \lim_{K \rightarrow \infty} \frac{\zeta_{K+1}(1)}{\zeta_K(1)}$  exists and is finite. Proposition 1 implies for any

fixed integer  $m \geq 1$ ,

$$\frac{\zeta_K(m+1)}{\zeta_K(m)} = \frac{\zeta_{K-m}(1)}{\zeta_{K-m+1}(1)} \rightarrow R^{-1}.$$

Thus, for any fixed integer  $j \geq 2$ ,

$$\xi_K(j) = \frac{\zeta_K(j)}{\zeta_K(1)} = \prod_{m=1}^{j-1} \frac{\zeta_K(m+1)}{\zeta_K(m)} \rightarrow R^{1-j}.$$

Conversely, suppose the limit  $\lim_{K \rightarrow \infty} \frac{\zeta_{K+1}(1)}{\zeta_K(1)}$  does not exist. Then, taking  $j = 2$ , we obtain

$$\xi_K(2) = \frac{\zeta_K(2)}{\zeta_K(1)} = \frac{\zeta_{K-1}(1)}{\zeta_K(1)}$$

by Proposition 1. Thus,  $\lim_{K \rightarrow \infty} \xi_K(2)$  diverges.  $\square$

**Remark 2.** The conditions of Proposition 2 can be somewhat relaxed. Specifically, we only use the self-similarity requirement to prove the shift equality (10). Hence, this equality together with the existence of the limit (11) suffices to obtain the strong Horton law.

#### 4.1. Expressing $\zeta_K(1)$ from $\{T_j\}$

In this section, we express  $\zeta_K(1)$  in terms of the elements of the Tokunaga sequence  $\{T_j\}_{j=1,2,\dots}$ , under the assumption of a ‘‘tamed’’ Tokunaga sequence:  $\limsup_{j \rightarrow \infty} T_j^{1/j} < \infty$ . We define

$$t(i) = \begin{cases} -1 & i = 0, \\ T_1 + 2 & i = 1, \\ T_i & i \geq 2 \end{cases}$$

and let  $\hat{t}(z) = \sum_{j=0}^{\infty} z^j t(j) = -1 + 2z + \sum_{j=1}^{\infty} z^j T_j$ . The quantity  $\zeta_K(1)$  can be computed by counting, and expressed via convolution products as follows:

$$\begin{aligned} \zeta_{K+1}(1) &= \sum_{r=1}^K \sum_{\substack{j_1, j_2, \dots, j_r \geq 1 \\ j_1 + j_2 + \dots + j_r = K}} t(j_1) t(j_2) \cdots t(j_r) \\ &= \sum_{r=1}^K \underbrace{(t + \delta_0) * (t + \delta_0) * \cdots * (t + \delta_0)}_{r \text{ times}}(K) \\ &= \sum_{r=1}^{\infty} \underbrace{(t + \delta_0) * (t + \delta_0) * \cdots * (t + \delta_0)}_{r \text{ times}}(K), \end{aligned}$$

where  $\delta_0(j)$  is the Kronecker delta, and therefore,  $(t + \delta_0)(0) = 0$ . Hence, taking the  $z$ -transform of

$\zeta_{K+1}(1)$ , we obtain

$$\begin{aligned} \sum_{K=1}^{\infty} z^{K-1} \zeta_K(1) &= 1 + \sum_{r=1}^{\infty} [\widehat{(t + \delta_0)}(z)]^r \\ &= 1 + \sum_{r=1}^{\infty} [\hat{t}(z) + 1]^r = -\frac{1}{\hat{t}(z)} \end{aligned} \tag{12}$$

for  $|z|$  small enough.

For a holomorphic function expanding in a power series  $f(z) = \sum_{j=0}^{\infty} a_j z^j$  in a non-empty neighborhood of zero containing  $|z| \leq \rho$ , define  $\check{f}(j) = \frac{1}{2\pi i} \oint_{|z|=\rho} \frac{f(z)}{z^{j+1}} dz = a_j$ . Then we arrive with the following formula, expressing  $\zeta_K(1)$  from  $\{T_j\}$ ,

$$\zeta_{K+1}(1) = -\left(\frac{1}{\hat{t}}\right)(K). \tag{13}$$

**Lemma 2.** *Let  $w_0$  be the only real root of  $\hat{t}(z) = -1 + 2z + \sum_{j=1}^{\infty} z^j T_j$  in the interval  $(0, \frac{1}{2}]$ . Then, for any other root  $w$  of  $\hat{t}(z)$ , we have  $|w| > w_0$ .*

**Proof.** Observe that since  $\{T_j\}$  are all non-negative reals,  $\overline{\hat{t}(z)} = \hat{t}(\bar{z})$ , and that the radius of convergence of  $\sum_{j=1}^{\infty} z^j T_j$  must be greater than  $w_0$ . Suppose  $w = re^{i\theta}$  ( $0 \leq \theta < 2\pi$ ) is a root of magnitude at most  $w_0$ . That is  $\hat{t}(w) = 0$  and

$$r := |w| \leq w_0.$$

Then  $\hat{t}(\bar{w}) = 0$  and

$$\begin{aligned} 0 &= \frac{1}{2} [\hat{t}(w) + \hat{t}(\bar{w})] = -1 + 2r \cos(\theta) \\ &+ \sum_{j=1}^{\infty} r^j T_j \cos(j\theta). \end{aligned}$$

If  $r < w_0$ , then

$$\begin{aligned} 0 &= -1 + 2r \cos(\theta) + \sum_{j=1}^{\infty} r^j T_j \cos(j\theta) \\ &\leq -1 + 2r + \sum_{j=1}^{\infty} r^j T_j < -1 \\ &+ 2w_0 + \sum_{j=1}^{\infty} w_0^j T_j = 0 \end{aligned}$$

arriving to a contradiction. Thus  $r = w_0$ .

Next we show that  $\theta = 0$ . Suppose not. Then

$$\begin{aligned} 0 &= -1 + 2r \cos(\theta) + \sum_{j=1}^{\infty} r^j T_j \cos(j\theta) \\ &< -1 + 2r + \sum_{j=1}^{\infty} r^j T_j = -1 + 2w_0 \\ &+ \sum_{j=1}^{\infty} w_0^j T_j = 0 \end{aligned}$$

arriving to another contradiction. Hence  $r = w_0$ ,  $\theta = 0$ , and  $w = w_0$ .  $\square$

Let  $w_0$  denote the only root of  $\hat{t}(z)$  in the real line subinterval  $(0, \frac{1}{2}]$  as in Lemma 2. Recall that the radius of convergence  $L = (\limsup_{j \rightarrow \infty} T_j^{1/j})^{-1}$  of  $\hat{t}(z) = -1 + 2z + \sum_{j=1}^{\infty} z^j T_j$  is greater than  $w_0$ . Then, following Lemma 2, there is a positive real  $\gamma \in (w_0, L)$  such that

$$\gamma < w \quad \text{for all } w \neq w_0 \text{ such that } \hat{t}(w) = 0. \tag{14}$$

Now, (14) implies for  $0 < \rho < w_0$ ,

$$\begin{aligned} \zeta_K(1) &= \frac{-1}{2\pi i} \oint_{|z|=\rho} \frac{dz}{\hat{t}(z)z^K} \\ &= -\text{Res} \left( \frac{1}{\hat{t}(z)z^K}; w_0 \right) - \frac{1}{2\pi i} \oint_{|z|=\gamma} \frac{dz}{\hat{t}(z)z^K}. \end{aligned}$$

Observe that  $\text{Res}(\frac{1}{\hat{t}(z)z^K}; w_0)$  is a constant multiple of  $\frac{1}{w_0^K}$  since  $w_0$  is a root of  $\hat{t}(z)$  of algebraic multiplicity one. Thus, since  $w_0 < \gamma$  and  $|\frac{1}{2\pi i} \oint_{|z|=\gamma} \frac{dz}{\hat{t}(z)z^K}| \leq \frac{1}{\gamma^K \min_{|z|=\gamma} |\hat{t}(z)|}$ ,

$$\frac{\zeta_{K+1}(1)}{\zeta_K(1)} = \left| \frac{\zeta_{K+1}(1)}{\zeta_K(1)} \right| \rightarrow \frac{1}{w_0} \quad \text{as } K \rightarrow \infty.$$

Hence Proposition 2 will imply the following lemma.

**Lemma 3.** *Suppose  $\limsup_{j \rightarrow \infty} T_j^{1/j} < \infty$ . Then, for each positive integer  $j$ , the limit*

$$\xi(j) = \lim_{K \rightarrow \infty} \xi_K(j)$$

*exists, and  $\xi(j) = w_0^{j-1}$ .*



The converse is also true. Specifically, suppose the limit

$$R = \lim_{k \rightarrow \infty} \frac{\zeta_{k+1}(1)}{\zeta_k(1)}$$

exists and is finite. Then, since  $\zeta_k(1) \geq T_j^{k/j}$  for all  $j \in \mathbb{N}$  and  $k \in j\mathbb{N}$ ,

$$\limsup_{j \rightarrow \infty} T_j^{1/j} \leq \lim_{k \rightarrow \infty} [\zeta_k(1)]^{1/k} = R < \infty.$$

Hence, we have proven another lemma.

**Lemma 4.** *Suppose  $\limsup_{j \rightarrow \infty} T_j^{1/j} = \infty$ . Then, the limit  $\lim_{K \rightarrow \infty} \xi_K(j)$  does not exist at least for some  $j$ .*

We now combine the results in Lemmas 3 and 4 into the following theorem.

**Theorem 1.** *Suppose  $\limsup_{j \rightarrow \infty} T_j^{1/j} < \infty$ . Then, for each positive integer  $j$*

$$\xi(j) = \lim_{K \rightarrow \infty} \xi_K(j) = \lim_{K \rightarrow \infty} \frac{\mathcal{N}_j[K]}{\mathcal{N}_1[K]} = R^{1-j},$$

where  $1/R = w_0$  is the only real root of the function  $\hat{t}(z) = -1 + 2z + \sum_{j=1}^{\infty} z^j T_j$  in the interval  $(0, \frac{1}{2}]$ .

Conversely, if  $\limsup_{j \rightarrow \infty} T_j^{1/j} = \infty$ , then the limit  $\lim_{K \rightarrow \infty} \xi_K(j) = \lim_{K \rightarrow \infty} \frac{\mathcal{N}_j[K]}{\mathcal{N}_1[K]}$  does not exist at least for some  $j$ .

We notice that the fact that  $R$  is reciprocal to the solution of  $\hat{t}(z) = 0$  was noticed by Peckham,<sup>6</sup> under the assumption  $N_k \sim cR^{K-k}$ , as  $K \rightarrow \infty$ . Below we give several examples of Theorem 1.

**Example 1. Shallow side-branching.** Suppose  $T_k = 0$  for  $k \geq 3$ , that is we only have “shallow” side-branches of orders  $\{j-2, j\}$  and  $\{j-1, j\}$ . Then

$$\hat{t}(z) = -1 + (T_1 + 2)z + T_2 z^2.$$

The only root of this equation within  $[0, 1/2]$  is

$$w_0 = \frac{\sqrt{(T_1 + 2)^2 + 4T_2} - (T_1 + 2)}{2T_2},$$

which leads to

$$R = \frac{1}{w_0} = \frac{\sqrt{(T_1 + 2)^2 + 4T_2} + (T_1 + 2)}{2}.$$

In particular, if  $T_k = 0$  for  $k \geq 2$ , then  $R = T_1 + 2$ ; such trees are called “cyclic”.<sup>6</sup> This shows that the entire range of Horton exponents  $2 \leq R < \infty$  can be achieved by trees with only very shallow side-branching. This also shows that  $T_1 \geq 1$  leads to

$R \geq 3$ , which seems to be the case for most of the observed branching systems.

**Example 2. Tokunaga self-similarity.** suppose  $T_j = a c^{j-1}$ , where  $a, c > 0$ . This model received considerable attention in the literature,<sup>6,14,23</sup> in part because of its ability to closely describe river networks.<sup>17</sup> Here we have

$$\begin{aligned} \hat{t}(z) &= -1 + 2z + az \sum_{j=1}^{\infty} (cz)^{j-1} \\ &= -1 + 2z + \frac{az}{1 - cz} \\ &= \frac{-1 + (a + c + 2)z - 2cz^2}{1 - cz} \end{aligned}$$

which leads to

$$-\frac{1}{\hat{t}(z)} = \frac{1 - cz}{1 - (a + c + 2)z + 2cz^2},$$

and the discriminant is positive,  $(a + c + 2)^2 - 8c > (c + 2)^2 - 8c = (c - 2)^2 \geq 0$ . Therefore, there will be two positive roots,  $p_1 > p_2$  of the denominator  $1 - (a + c + 2)z + 2cz^2$ , and

$$-\frac{1}{\hat{t}(z)} = \frac{1 - cz}{2c(p_1 - p_2)(z - p_1)} - \frac{1 - cz}{2c(p_1 - p_2)(z - p_2)}.$$

Thus, since  $\frac{1}{z-p} = -\sum_{k=0}^{\infty} \frac{1}{p^{k+1}} z^k$  for  $|z| < |p|$ , formula (13) implies

$$\zeta_{K+1}(1) = \frac{1}{2c(p_1 - p_2)} \left( \frac{1 - cp_2}{p_2^{K+1}} - \frac{1 - cp_1}{p_1^{K+1}} \right) \tag{15}$$

for  $|z|$  small enough, where one can easily check that  $1 > cp_2$ . Therefore, the conditions of Proposition 2 are satisfied with

$$\frac{1}{R} = p_2 = \frac{a + c + 2 - \sqrt{(a + c + 2)^2 - 8c}}{4c}.$$

Hence,

$$R = \frac{a + c + 2 + \sqrt{(a + c + 2)^2 - 8c}}{2} \tag{16}$$

as in earlier works.<sup>6,14,23</sup> Also, in agreement with the Lemma 3,  $w_0 = p_2 = \frac{1}{R}$ .

**Example 3. “Differentiated Tokunaga” self-similarity.** Suppose  $T_j = a \cdot j c^{j-1}$ , where  $a, c > 0$ .

Then

$$\begin{aligned}\hat{t}(z) &= -1 + 2z + az \sum_{j=1}^{\infty} j(cz)^{j-1} \\ &= -1 + 2z + \frac{az}{(1-cz)^2} \\ &= \frac{2c^2z^3 - c(c+4)z^2 + (a+2c+2)z - 1}{(1-cz)^2}.\end{aligned}$$

Here  $\frac{1}{R} = w_0$  is the smallest positive real root of polynomial

$$p(z) = 2c^2z^3 - c(c+4)z^2 + (a+2c+2)z - 1.$$

Now, since  $\hat{t}(1/2) > 0$ ,  $w_0 \in (0, \frac{1}{2})$ . In this example, we cannot derive an explicit formula for  $R = R(a, c)$ . However, we solve  $p(w_0) = 0$  for  $c > 0$ , obtaining the following relation among  $a, c$  and  $R$ :

$$c = \frac{1}{w_0} + \sqrt{\frac{a}{(1-2w_0)w_0}} = \left(1 + \sqrt{\frac{a}{R-2}}\right) R.$$

## ACKNOWLEDGMENTS

This research is supported by the NSF DMS-1412557 award.

## REFERENCES

1. J. W. Kirchner, Statistical inevitability of Horton's laws and the apparent randomness of stream channel networks, *Geology* **21**(7) (1993) 591–594.
2. P. S. Dodds and D. H. Rothman, Unified view of scaling laws for river networks, *Phys. Rev. E* **59**(5) (1999) 4865.
3. R. E. Horton, Drainage-basin characteristics, *Eos, Trans. Am. Geophys. Union* **13**(1) (1932) 350–361.
4. R. E. Horton, Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology, *Geol. Soc. Am. Bull.* **56** (1945) 275–370.
5. R. L. Shreve, Statistical law of stream numbers, *J. Geol.* **74** (1966) 17–37.
6. S. D. Peckham, New results for self-similar trees with applications to river networks, *Water Resources Res.* **31** (1995) 1023–1029.
7. D. G. Tarboton, Fractal river networks, Horton's laws and Tokunaga cyclicity, *J. Hydrol.* **187**(1) (1996) 105–117.
8. V. K. Gupta and E. Waymire, Some mathematical aspects of rainfall, landforms and floods, in *Rainfall, Landforms and Floods*, O. Barndorff-Nielsen, V. K. Gupta and V. Perez-Abreu and E. Waymire (eds.) (World Scientific, 1998).
9. I. Rodriguez-Iturbe and A. Rinaldo, *Fractal River Networks: Chance and Self-Organization* (Cambridge University Press, New York, 1997).
10. W. I. Newman, D. L. Turcotte and A. M. Gabrielov, Fractal trees with side branching, *Fractals* **5** (1997) 603–614.
11. R. L. Shreve, Infinite topologically random channel networks, *J. Geol.* **75** (1967) 178–186.
12. J. Pitman, *Combinatorial Stochastic Processes*, Lecture Notes in Mathematics, Vol. 1875 (Springer-Verlag, 2006).
13. G. A. Burd, E. C. Waymire and R. D. Winn, A self-similar invariance of critical binary Galton–Watson trees, *Bernoulli* **6** (2000) 1–21.
14. E. Tokunaga, Consideration on the composition of drainage networks and their evolution, *Geographical Rep. Tokyo Metro. Univ.* **13** (1978) 1–27.
15. J. G. Masek and D. L. Turcotte, A diffusion limited aggregation model for the evolution of drainage networks, *Earth Planet. Sci. Lett.* **119** (1993) 379.
16. D. L. Turcotte, J. D. Pelletier and W. I. Newman, Networks with side branching in biology, *J. Theor. Biol.* **193** (1998) 577–592.
17. S. Zanardo, I. Zaliapin and E. Foufoula-Georgiou, Are American rivers Tokunaga self-similar? New results on fluvial network topology and its climatic dependence, *J. Geophys. Res.* **118** (2013) 166–183.
18. P. Ossadnik, Branch order and ramification analysis, of large diffusion limited aggregation clusters, *Phys. Rev. A* **45** (1992) 1058.
19. D. L. Turcotte, B. D. Malamud, G. Morein and W. I. Newman, An inverse cascade model for self-organized critical behavior, *Physica A* **268** (1999) 629–643.
20. G. Yakovlev, W. I. Newman, D. L. Turcotte and A. Gabrielov, An inverse cascade model for self-organized complexity and natural hazards, *Geophys. J. Int.* **163** (2005) 433–442.
21. I. Zaliapin, H. Wong and A. Gabrielov, Inverse cascade in percolation model: Hierarchical description of time-dependent scaling, *Phys. Rev. E* **71** (2006) 066118.
22. I. Zaliapin and Y. Kovchegov, Tokunaga and Horton self-similarity for level set trees of Markov chains, *Chaos Solitons Fractals* **45**(3) (2012) 358–372.
23. M. McConnell and V. Gupta, A proof of the Horton law of stream numbers for the Tokunaga model of river networks, *Fractals* **16** (2008) 227–233.
24. Y. Kovchegov and I. Zaliapin, Horton self-similarity of Kingman's coalescent tree, to appear in *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* (2016).
25. A. N. Strahler, Quantitative analysis of watershed geomorphology, *Trans. Am. Geophys. Union* **38** (1957) 913–920.
26. R. L. Shreve, Stream lengths and basin areas in topologically random channel networks, *J. Geology* **77** (1969) 397–414.