

# Tokunaga self-similarity arises naturally from time invariance

Yevgeniy Kovchegov\*

*Department of Mathematics, Oregon State University, Corvallis, OR, 97331-4605, USA*

Ilya Zaliapin†

*Department of Mathematics and Statistics, University of Nevada, Reno, NV, 89557-0084, USA*

(Dated: April 9, 2018)

The Tokunaga condition is an algebraic rule that provides a detailed description of the branching structure in a self-similar tree. Despite a solid empirical validation and practical convenience, the Tokunaga condition lacks a theoretical justification. Such a justification is suggested in this work. We define a geometric branching processes  $\mathcal{G}(s)$  that generates self-similar rooted trees. The main result establishes the equivalence between the invariance of  $\mathcal{G}(s)$  with respect to a time shift and a one-parametric version of the Tokunaga condition. In the parameter region where the process satisfies the Tokunaga condition (and hence is time invariant),  $\mathcal{G}(s)$  enjoys many of the symmetries observed in a critical binary Galton-Watson branching process and reproduce the latter for a particular parameter value.

**The Tokunaga condition, also called Tokunaga self-similarity, is a particularly simple parameterization of branching that gives a constructive description of a self-similar tree. It has originated in hydrology and has been empirically rediscovered in a variety of other fields. The lack of a theoretical justification behind the Tokunaga condition raises the question of whether this is an actual physical constraint or merely a convenient ad hoc approximation that cannot be rejected with available data. The answer to this question affects theoretical and modeling efforts across disciplines. We argue that the Tokunaga condition is an actual physical constraint. Specifically, we establish a measure-theoretic equivalence between a one-parametric version of the Tokunaga condition and time invariance of a branching process that generates random trees. The methodology and results of this work are readily applicable to a multitude of questions related to scaling laws in trees (Horton law, fractal dimension, etc.). We also solve an abstract non-linear problem that can be of general interest to the non-linear dynamics community.**

## INTRODUCTION

Tree-shaped fractal formations, from the namesake botanical trees to river tributary networks to the systems of canyons and mountain crests that define the Earth topography, have always fascinated the great minds (e.g., [1]) providing inspiration in science, art, and architecture [2, 3]. A quantitative understanding of the branching patterns is instrumental in hydrology [4–9], geomorphology [10, 11], statistical seismology [12–19], statistical physics of fracture [20–22], vascular analysis [23], brain studies [24], ecology [25], biology [26], and beyond, en-

couraging further rigorous treatment. This study establishes equivalence between the Tokunaga self-similarity – a widely recognized algebraic parameterization of a self-similar tree – and a measure-theoretic invariance of a tree distribution with respect to a depth shift.

Much of the self-similar tree analysis has originated in the studies of river networks; it refers to the rooted trees with no edge lengths. The celebrated Horton-Strahler ordering scheme [28, 29] assigns integer orders to tree vertices and edges, beginning with  $\text{ord} = 1$  at the leaves and incrementing by unity every time the two edges of the same order meet (see Def. 1, Fig. 2). A sequence of adjacent vertices/edges with the same order is called a *branch* (Fig. 2c). Let  $N_j$  be the number of branches of order  $j$  and  $N_{ij}$ ,  $i < j$ , be the number of *side branches* of order  $\{ij\}$  – instances when an order- $i$  branch merges with an order- $j$  branch in a finite tree  $T$ . The average number of branches of order  $i$  in a single branch of order  $j$  can be traced with the following ratio [30]

$$T_{ij} = \frac{\mathbb{E}[N_{ij}]}{\mathbb{E}[N_j]}. \quad (1)$$

Here  $\mathbb{E}[x]$  is the expected value with respect to a suitable probability measure on the examined collection of trees.

Ronald L. Shreve [31] pioneered the study of side branching in a *topologically random channel network*, which is a uniform distribution of rooted binary trees with a given number of leaves. This model is equivalent to a critical binary Galton-Watson tree, conditioned on the number of leaves [32, 33]. Recall that a Galton-Watson process describes growth of a population that begins with a single progenitor. At each discrete time step, each population member disappears giving birth to a random number  $k \geq 0$  of offspring according to a distribution  $\{p_k\}$ . A Galton-Watson tree describes a trajectory of this process. The members are represented by the tree vertices; the progenitor corresponds to the tree root; the edges connect parents to their offspring. The

case  $p_0 + p_2 = 1$  corresponds to a *binary tree*, the constant average progeny case  $p_0 = p_2 = 1/2$  is called *critical*.

Shreve's calculations imply  $T_{ij} = 2^{j-i-1}$ , indicating, in particular, that the side branching only depends on the difference between the branch orders, not on their absolute values. Eiji Tokunaga [34] generalized this idea by introducing conditions

$$T_{ij} = T_{j-i} \quad \text{for some sequence } \{T_k \geq 0\}_{k \geq 1}, \quad (2a)$$

$$T_k = a c^{k-1}, \quad k \geq 1, \quad a \geq 0, \quad c > 0. \quad (2b)$$

The Tokunaga model (2) completely specifies a combinatorial tree shape (up to a permutation of side branch attachment within a given branch) with only two parameters ( $a, c$ ), hence suggesting a conventional modeling paradigm. The empirical validity of the Tokunaga constraints (2) has been strongly confirmed for a variety of river networks at different geographic locations [10, 35–38], as well as in other types of data represented by trees, including botanical trees [39], the veins of botanical leaves [26, 27], clusters of dynamically limited aggregation [39, 40], percolation and forest-fire model clusters [41, 42], earthquake aftershock sequences [14, 15, 19], tree representation of symmetric random walks [43], and hierarchical clustering [44]. The conditions (2), however, lack a theoretical justification.

This work suggests such a justification, by exploring a space  $\mathcal{T}$  of finite binary rooted trees with an arbitrary probability measure. We introduce the *Horton pruning* – cutting the tree leaves (Fig. 1), which is closely connected to the concept of Horton-Strahler orders. We showed in [30] that prune invariant measures are abundant on  $\mathcal{T}$ , and every prune invariant measure necessarily satisfies the constraint (2a). This motivates a *geometric branching process* (Def. 3) that induces a prune invariant measure for an arbitrary sequence  $\{T_k\}$ . A trajectory of this process is a random *geometric tree* (Def. 4). The geometric tree that satisfies (2b) with  $(a, c) = (1, 2)$  is equivalent to the critical binary Galton-Watson tree. A *time invariant* distribution preserves the average counts of branches of each order under the operation of a unit time shift in a geometric branching process (Def. 5, Fig. 3). Our main result (Thm. 2) is that for the geometric trees, time invariance is equivalent to the Tokunaga constraint (2b) with  $a = c - 1$  and  $c \geq 1$ . Proofs are given in a separate section at the end of the paper.

## HORTON PRUNE INVARIANCE

Denote by  $\mathcal{T}$  the space of finite unlabeled rooted reduced binary trees with no planar embedding, including the *empty tree*  $\phi$  comprised of a single root vertex and no edges. The existence of the root induces a parent-offspring relation for each pair of adjacent vertices: the one closer to the root is called parent, the other – off-

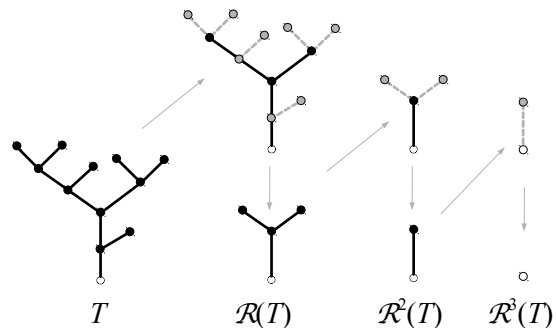


FIG. 1. Pruning: an illustration. The bottom row shows the trajectory  $T \rightarrow \mathcal{R}(T) \rightarrow \mathcal{R}^2(T) \rightarrow \mathcal{R}^3(T) = \phi$  of an initial tree  $T$ . The top row highlights the edges and vertices (dashed gray) to be pruned at each step. Notice two series reduction (elimination of a degree-2 vertex) that occur after the first pruning.

spring. The absence of labeling means that the tree vertices are not distinguishable (other than by their position within a tree). The absence of planar embedding refers to the lack of ordering (left/right) among the offspring of the same parent. The term *reduced* refers to the absence of vertices of degree 2.

The concept of tree self-similarity is related to the pruning operation [30, 32, 35]. *Pruning* of a tree is a map  $\mathcal{R} : \mathcal{T} \rightarrow \mathcal{T}$ , whose value  $\mathcal{R}(T)$  for  $T \neq \phi$  is obtained by removing the leaves and their parental edges from  $T$ , followed by series reduction (removing vertices of degree 2). We also set  $\mathcal{R}(\phi) = \phi$ . Pruning is an onto but not invertible operator, with an infinite preimage  $\mathcal{R}^{-1}(T)$  for any  $T \in \mathcal{T}$ . Pruning induces a contracting flow on  $\mathcal{T}$ . The trajectory of each tree  $T$  under  $\mathcal{R}(\cdot)$  is uniquely determined and finite (Fig. 1):

$$T \equiv \mathcal{R}^0(T) \rightarrow \mathcal{R}^1(T) \rightarrow \cdots \rightarrow \mathcal{R}^k(T) = \phi.$$

**Definition 1 (Horton-Strahler orders).** The Horton-Strahler order  $\text{ord}(T) \in \mathbb{Z}_+$  of a tree  $T \in \mathcal{T}$  is the minimal number of prunings necessary to erase the tree:

$$\text{ord}(T) = \min \{k \geq 0 : \mathcal{R}^k(T) = \phi\}. \quad (3)$$

The order  $\text{ord}(v)$  of a vertex  $v \in T$  is the order of the *descendant subtree*  $T_v \subset T$  that includes all vertices descendant to  $v$  (including  $v$ ) in  $T$ , together with their parental edges (Fig. 2a,b). The order of the parental edge of  $v$  is  $\text{ord}(v)$ . A *branch* is a collection of adjacent vertices/edges of the same order (Fig. 2c). A *side branch* of order  $\{ij\}$ ,  $i < j$ , is a branch of order  $i$  that merges a branch of a higher order  $j$ . The tree in Fig. 2c has two side branches of order  $\{12\}$ ; a situation when two branches of the same order merge (and form a branch of a higher order) is *not* a side branch.

**Definition 2 (Prune invariance).** Consider a probability measure  $\mu$  on  $\mathcal{T}$  such that  $\mu(\phi) = 0$ . Let

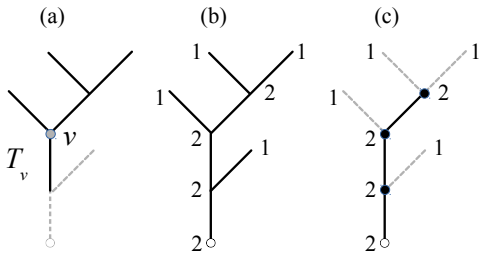


FIG. 2. Horton-Strahler ordering: an illustration. (a) Subtree  $T_v$  (solid) descendant to vertex  $v$  in  $T$  (dashed gray). (b) Horton-Strahler orders of tree vertices. The tree order  $\text{ord}(T) = 2$  is assigned to the tree root. (c) A branch of order  $\text{ord} = 2$  consists of three vertices and their parental edges (solid) in  $T$  (dashed gray).

$\nu(T) = \mu \circ \mathcal{R}^{-1}(T) = \mu(\mathcal{R}^{-1}(T))$ . (Note that  $\nu(\phi) > 0$ .) Measure  $\mu$  is called *invariant* with respect to the pruning operation if for any tree  $T \in \mathcal{T}$  we have

$$\nu(T | T \neq \phi) = \mu(T). \quad (4)$$

We restrict our analysis to *coordinated* measures [30]. Informally, the coordination property asserts that all descendant subtrees of a given order  $\kappa$  have the same probability structure, independent of how deep or shallow they occur in a given tree. Formally, a tree measure  $\mu(\tau)$  is coordinated if, conditioned on  $\text{ord}(\tau) = K \geq 1$ , for any  $\kappa < K$ , a uniformly selected descendant subtree of  $\tau$  of order  $\kappa$  is distributed as  $\mu(T | \text{ord}(T) = \kappa)$ . A tree measure  $\mu(T)$ , and a random tree drawn from  $\mu(T)$ , is called *self-similar* if it is coordinated and prune invariant.

The authors have shown [30] that for the coordinated measures, the prune invariance implies the order invariance constraint (2a). It was also shown that prune invariance implies a geometric distribution of tree orders. Accordingly, if one is interested in prune invariant trees, it is natural to consider measures specified by sequences  $\{T_k\}$ , with a geometric order distribution. The next section introduces a natural class of such measures.

## GEOMETRIC BRANCHING PROCESS

We define here a multi-type discrete time branching process that generates a prune invariant measure on  $\mathcal{T}$  that satisfy (2a) for an arbitrary sequence  $\{T_k\}$ . In general, this can be done in numerous ways [30]. We restrict ourselves to Markov processes that facilitate subsequent analysis.

We say that a random variable  $X$  has a geometric distribution with parameter  $r$ ,  $X \stackrel{d}{=} \text{Geom}(r)$ , if its probability mass function is given by

$$\text{Prob}(X = k) = r(1 - r)^k, \quad k = 0, 1, 2, \dots$$

**Definition 3 (Geometric branching process).** Consider a sequence of *Tokunaga coefficients*  $\{T_k \geq 0\}_{k \geq 1}$  and  $0 < p < 1$ . Define  $S_K := 1 + T_1 + \dots + T_K$  for  $K \geq 0$  by assuming  $T_0 = 0$ . The *geometric branching process*  $\mathcal{G}(s; T_k, p) \equiv \mathcal{G}(s)$  describes a discrete time population growth:

- (i) The process starts at  $s = 0$  with a progenitor of order  $\text{ord}(\mathcal{G})$  such that  $\text{ord}(\mathcal{G}) - 1 \stackrel{d}{=} \text{Geom}(p)$ .
- (ii) At every time instant  $s > 0$ , each population member of order  $K \in \{1, \dots, \text{ord}(\mathcal{G})\}$  terminates with probability  $q_K = S_{K-1}^{-1}$ , independently of other members. At termination, a member of order  $K > 1$  produces two offspring of order  $(K - 1)$ ; and a member of order  $K = 1$  terminates with leaving no offspring.
- (iii) At every time instant  $s > 0$ , each population member of order  $K \in \{1, \dots, \text{ord}(\mathcal{G})\}$  survives (does not terminate) with probability  $1 - q_K = 1 - S_{K-1}^{-1}$ , independently of other members. In this case, it produces a single offspring (side branch). The offspring order  $i$ ,  $1 \leq i < K$ , is drawn from the distribution

$$p_{K,i} = \frac{T_{K-i}}{T_1 + \dots + T_{K-1}}. \quad (5)$$

The geometric branching process is Markov in the space of ornamented trees – trees with vertex orders. The numbers of side branches are independent for distinct branches. Denote by  $m_K$  the total number of side branches within a randomly selected branch of order  $K$ , and by  $m_{K,i}$  the number of side branches of order  $i$ ,  $i = 1, \dots, K - 1$ , within that branch. Branches of order 1 have unit length and no side branches:  $m_1 = 0$ . The property (ii) of the definition implies for any  $K > 1$

$$m_K \stackrel{d}{=} \text{Geom}(S_{K-1}^{-1}),$$

with  $\mathbb{E}[m_K] = S_{K-1} - 1 = T_1 + \dots + T_{K-1}$ . Combining this with (5) we find

$$m_{K,i} \stackrel{d}{=} \text{Geom}\left([1 + T_{K-i}]^{-1}\right),$$

with  $\mathbb{E}[m_{K,i}] = T_{K-i}$ . The independence of branches implies  $\mathbb{E}[N_{ij}] = \mathbb{E}[N_j] \mathbb{E}[m_{j,i}]$  and hence

$$T_{ij} = \frac{\mathbb{E}[N_{ij}]}{\mathbb{E}[N_j]} = \frac{\mathbb{E}[N_j] \mathbb{E}[m_{j,i}]}{\mathbb{E}[N_j]} = T_{j-i}.$$

This means that a trajectory of the geometric branching process is a (random) tree  $\mathcal{G}(T_k, p) \in \mathcal{T}$  that satisfies the condition (2a) with the sequence  $\{T_k\}$ . We notice that essential elements of the resulting trees (tree order, branch lengths, numbers of side branches of a given order within a selected branch) are described by geometric laws – hence the model name.

A tree  $\mathcal{G}(T_k, p) \in \mathcal{T}$  generated by the geometric branching process can be equivalently defined via the following construction.

**Definition 4 (Geometric tree).** Consider a sequence  $\{T_k \geq 0\}_{k \geq 1}$  and  $0 < p < 1$ . We say that a (random) tree  $\mathcal{G}(T_k, p) \in \mathcal{T}$  is a *geometric tree* if and only if:

- (i) The tree order satisfies  $\text{ord}(\mathcal{G}) - 1 \stackrel{d}{=} \text{Geom}(p)$ .
- (ii) The total number  $m_K$  of side branches within a branch of order  $K \geq 2$  is  $\text{Geom}(S_{K-1}^{-1})$ .
- (iii) For a branch of order  $K \geq 2$ , conditioned on  $m_K = m$ , the assignment of orders for  $m$  side branches is done according to multinomial distribution with  $m$  trials and success probabilities  $p_{K,j}$  of (5) for any  $j = 1, \dots, K-1$ .
- (iv) The number of side branches and their orders are independent in distinct branches.

The Def. (4) is convenient for recursive tree generation. A tree of order  $K = 1$  consists of two vertices (root and leaf) connected by an edge. To generate a random tree of general order, one first generates a random geometric order  $K$  and starts with a complete binary tree of depth  $K$ , which we call *skeleton*. All leaves in the skeleton have the same depth  $K$ , and all vertices at depth  $\kappa \leq K$  have the same Horton-Strahler order  $\kappa$ . Accordingly, the skeleton has order  $K$  equal to the tree depth. The final tree is obtained by adding side branches of lower orders to every branch of the skeleton, according to the definition rules. Every side branch is generated according to the same procedure.

It follows from the properties of a critical binary Galton-Watson tree [32] that the geometric tree with  $p = 1/2$  and  $T_k = 2^{k-1}$  is the critical binary Galton-Watson tree. Notably, the next statement shows that the measure induced by a geometric tree (and hence by a geometric branching process) is prune invariant (Def. 2).

**Theorem 1 (Prune invariance).** *Given an arbitrary sequence  $\{T_k \geq 0\}_{k \geq 1}$  and  $0 < p < 1$ , the probability measure for the geometric random tree  $\mathcal{G}(T_k, p)$  is invariant with respect to Horton pruning.*

*Proof.* See the Proofs section. □

A geometric tree measure is coordinated by construction. The coordination together with prune invariance implies that the geometric measure is self-similar.

### TIME INVARIANCE

Let  $x_i(s)$ ,  $i \geq 1$ , denote the average number of vertices of order  $i$  at time  $s$  in the process  $\mathcal{G}(s)$ , and  $\mathbf{x}(s) =$

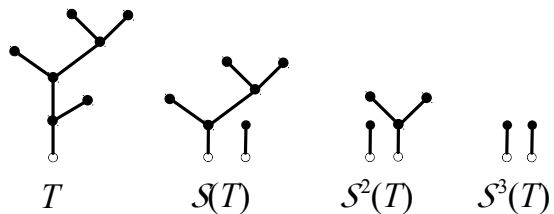


FIG. 3. Time shift: an illustration. The figure shows forests obtained by consecutive application of the time shift operator  $\mathcal{S}$  to a tree  $T$  shown on the left. At every step, we remove the root edge from each existing tree. This terminates the trees of order  $\text{ord} = 1$ , and splits any other tree in two new trees.

$(x_1(s), x_2(s), \dots)^T$  be the state vector. By definition we have

$$\mathbf{x}(0) = \pi := \sum_{K=1}^{\infty} p(1-p)^{K-1} \mathbf{e}_K,$$

where  $\mathbf{e}_K$  are standard basis vectors. Furthermore, if  $q_{a,b}$ ,  $a \geq b$ , denotes the probability that a vertex of order  $\text{ord} = a + \mathbb{1}_{\{a=b\}}$  that exists at time  $s$  splits into a pair of vertices of orders  $(a, b)$  at time  $(s+1)$ , then

$$x_K(s+1) = 2x_{K+1}(s)q_{K,K} + x_K(s)(1 - q_{K-1, K-1}) + \sum_{i=K+1}^{\infty} x_i(s)q_{i,K}. \quad (6)$$

The first term in the right-hand side of (6) corresponds to a split of an order- $(K+1)$  vertex into two vertices of order  $K$ , the second – to a split of an order- $K$  vertex into a vertex of order  $K$  and a vertex of a smaller order, and the third – to a split of a vertex of order  $i > K$  into a vertex of order  $K$  and a vertex of order  $i$ . The geometric branching implies

$$q_{a,b} = \begin{cases} S_a^{-1} & \text{if } a = b, \\ T_{a-b} S_{a-1}^{-1} & \text{if } b < a. \end{cases} \quad (7)$$

Summing up, the system (6) can be written as

$$\mathbf{x}(s+1) - \mathbf{x}(s) = \mathbb{G} \mathbf{x}(s), \quad (8)$$

where

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

and

$$\mathbb{S} = \text{diag}\{S_0, S_1, \dots\}.$$

In this setup, the *unit time shift* operator  $\mathcal{S}$ , which advances the process time by unity, can be applied to individual trees and forests (collection of trees) – it removes

the root edge from a tree  $T$ . A consecutive applications of  $d$  time shifts to a tree  $T$  is equivalent to removing the vertices/edges at depth less than  $d$  from the root (Fig. 3). Next we define an invariance with respect to the time shift  $\mathcal{S}$ .

**Definition 5 (Time invariance).** Process  $\mathcal{G}(s)$  is called *time invariant* if and only if the state vector  $\mathbf{x}(s)$  is invariant with respect to a unit time shift  $\mathcal{S}$ :

$$\mathbf{x}(s) = \mathbf{x}(0) \equiv \pi \quad \forall s \iff \mathbb{G}\mathcal{S}^{-1}\pi = \mathbf{0}. \quad (10)$$

Now we formulate the main result of this paper.

**Theorem 2 (Critical Tokunaga process).** *A geometric branching process  $\mathcal{G}(s; T_k, p)$  is time invariant if and only if*

$$p = 1/2 \text{ and } T_k = (c-1)c^{k-1} \text{ for any } c \geq 1. \quad (11)$$

*We call this family a critical Tokunaga process, and the respective trees – critical Tokunaga trees.*

Theorem 2 is proven in the Proofs section via solving a nonlinear system of equations that writes (10) in terms of ratios  $S_k/S_{k+1}$ .

Let  $K$  be the order of a random tree  $\mathcal{G}(T_k, p)$ , and, conditioned on  $K > 1$ , let  $K_a, K_b$  be the orders of its two subtrees,  $T_a$  and  $T_b$ , rooted at the internal vertex closest to the root, randomly and uniformly permuted. We call  $T_a$  and  $T_b$  the *principal* subtrees of  $T$ .

**Corollary 1.** *Let  $\mathcal{G}$  be a critical Tokunaga tree. Then the distribution of the principal subtree  $T_a$  (and hence  $T_b$ ) matches that of the initial tree  $\mathcal{G}$ . The distributions of  $T_a$  and  $T_b$  are independent if and only if  $c = 2$ .*

*Proof.* See the Proofs section. □

## DISCUSSION

We suggest a measure-theoretic justification for the algebraic Tokunaga constraint (2) with  $a = c - 1$  and  $c \geq 1$ , in terms of invariance of geometric branching process  $\mathcal{G}(s; T_k, p)$  with respect to a time shift. As a result, there appears a family of critical Tokunaga branching processes (11) characterized by invariance to both the Horton pruning (erasing a tree from leaves to the root) and time shift (erasing a tree from the root to leaves).

*Extending properties of critical Galton-Watson tree.* Our results extend some remarkable properties of the critical binary Galton-Watson process ( $T_k = 2^{k-1}$ ) to a one-parametric family of critical Tokunaga processes ( $T_k = (c-1)c^{k-1}$ ):

(i) Burd et al. [32] have shown that condition (2a) characterizes the invariance with respect to the Horton pruning in the family of binary Galton-Watson trees. Our work [30] established that (2a) follows from the prune

invariance for any coordinated measure (recall that any Galton-Watson measure is coordinated, but not vice-versa). Moreover, coordination is a necessary property – one can construct a multitude of uncoordinated prune invariant measures that do not satisfy the constraint (2a).

(ii) Among binary Galton-Watson trees, the condition (2a), and hence prune invariance, only holds in the *critical* case ( $p_0 = p_2 = 1/2$ ). In this case (2b) is also satisfied with  $a = c - 1 = 1$  [32]. We show here that the condition (2b) with  $a = c - 1$  is equivalent to time invariance (which implies criticality) in a broader family of geometric trees.

(iii) The forest obtained from descendant subtrees rooted at every vertex of a critical binary Galton-Watson tree  $T$  approximates the frequency structure of the entire space of trees, as the order of  $T$  increases. The same property holds for the trees generated by a critical Tokunaga branching process with any  $c \geq 1$ .

*Independence of Horton-Strahler ordering.* Our results relate the conditions (2a) and (2b), expressed in terms of the Horton-Strahler orders, to a discrete erasure of a tree from the leaves (Horton pruning) or from the root (time invariance), respectively. Such erasures, that do not use the notion of order, provide more natural process-generating constraints.

*Ease of simulation.* Generation of geometric trees for arbitrary parameters ( $p, \{T_k\}$ ) is easily implemented on a computer (see discussion after Def. 4). This model may hence facilitate analysis in a range of simulation-heavy problems, from structure and transport on river networks to phylogenetic tree analysis.

*Horton law.* The Horton law – a geometric decay of the branch counts  $N_i$  – is among the most fundamental regularities found in observed trees, both static and dynamic [4, 10, 35, 36, 39, 45, 46]. It was shown in [30] that a strong version of the Horton law is satisfied in a self-similar tree, provided  $\limsup_{k \rightarrow \infty} T_k^{1/k} < \infty$ . In particular, for a critical Tokunaga tree with  $T_k = (c-1)c^{k-1}$  and  $c \geq 1$  the Horton ratios  $N_i/N_{i+1}$  asymptotically converge to  $R_b = 2c$ , for any  $i \geq 1$ , as the tree order increases.

*Fractal dimension of trees.* Recall that a fractal dimension of a tree is defined as  $d = \ln R_b / \ln R_r$  [39], with the asymptotic Horton ratios  $R_b = \lim N_i/N_{i+1}$  and  $R_r = \lim r_{i+1}/r_i$ , where  $r_i$  is the average length of a branch of order  $i$ . If we assume that the lengths of edges in a geometric tree are independent equally distributed random variables with a unit mean, then  $\mathbf{E}(r_i) = S_{i-1}$ . For the critical Tokunaga model this gives  $R_r = c$  and the respective fractal dimension

$$d_c = \frac{\ln 2c}{\ln c} = 1 + \ln(2)/\ln(c). \quad (12)$$

Hence,  $1 \leq d_c < \infty$  depending on the value of  $c$ . In particular,  $d_{2^{1/k}} = 1 + k$ , which corresponds to a volume-filling tree in  $(1 + k)$  dimensions for any  $k \geq 1$ .

*Horton ratio.* The Horton ratio reported in the observed large trees is approximately within  $2.8 < R_b < 6$ .

The expression (12) for the fractal dimension of a critical Tokunaga tree, together with  $R_b = 2c$ , gives  $\log_2(R_b) = \frac{d_c}{d_c - 1}$ . Accordingly, the observed empirical range of the Horton ratios is reproduced by the critical Tokunaga trees with  $1.4 < c < 3$  and dimensions  $1.6 < d_c < 3$ . This covers all the tree dimensions that may exist in a 3-dimensional world, excluding the range  $d_c < 1.6$  that corresponds to almost “linear”, and probably less studied, trees.

On the condition  $a = c - 1$ . The constraint (2b) with  $a = c - 1$  that defines the critical Tokunaga process appears in the Random Self-similar Network model [37, 47], introduced in hydrological studies. At the same time, the multiple documented instances when the Tokunaga constraint (2b) is tested in observations and models typically refer to  $a \neq c - 1$  [10, 35–41, 44]. Although the observed deviations are not large in absolute value (e.g.,  $c \approx 2.6$ ,  $a \approx 1.1$  for river basins [35, 38]), they seem to be statistically significant [38]. It remains an open problem to either find a theoretical justification for the condition (2b) with  $a \neq c - 1$ , or otherwise explain the apparent deviation from the time invariance that leads to  $a \neq c - 1$  in observed trees.

## PROOFS

### Proof of Theorem 1

*Proof.* Consider a random geometric tree  $\mathcal{G} \equiv \mathcal{G}(T_k, p)$ . Let  $T := \mathcal{R}(\mathcal{G})$  be the respective pruned tree. We show below that tree  $T$  is geometric, i.e., it satisfies properties (i)-(iv) of Def. (4). Observe that the Horton pruning  $\mathcal{R}$  eliminates branches of order 1 in the initial tree  $\mathcal{G}$  and decreases the order of every non-leaf branch by unity. Hence, the event  $\{\text{ord}(\mathcal{G}) > 1\}$  is the same as  $\{T \neq \phi\}$ .

(i) We have

$$P(\text{ord}(T) = k | T \neq \phi) = \frac{P(\text{ord}(\mathcal{G}) = k + 1)}{P(\text{ord}(\mathcal{G}) > 1)},$$

which implies  $\{\text{ord}(T) | T \neq \phi\} - 1 \stackrel{d}{=} \text{Geom}(p)$ .

(ii) Pruning  $\mathcal{R}$  acts as a *Bernoulli thinning* on the number  $m_K$  of side branches within every branch of order  $K \geq 2$ , with thinning probability  $p_{K,1}$  of (5). Recall that application of a Bernoulli thinning with removal probability  $1 - q$  to a geometric random variable with parameter  $r$  results in a geometric random variable with parameter

$$r_{\text{thinned}} = \frac{r}{q(1 - r) + r}.$$

To obtain the distribution for the number  $m_K$  of side branches in a branch of order  $K$  of a pruned tree  $T$ , we apply a Bernoulli thinning to a branch of order  $(K + 1)$  in the original tree  $\mathcal{G}$ . Using  $r = S_K^{-1}$  and  $1 - q = p_{K+1,1}$  shows that  $m_K \stackrel{d}{=} \text{Geom}(S_{K-1}^{-1})$ , which establishes (ii).

(iii) Follows from the properties of multinomial distribution.

(iv) Follows from independence of the branch structure in  $\mathcal{G}$ .  $\square$

### Proof of Theorem 2

A proof of Theorem 2 is based on the following two lemmas proven in separate sections below.

**Lemma 1.** *A geometric branching process  $\mathcal{G}(s)$  is time invariant if and only if  $p = 1/2$  and the sequence  $\{T_k\}$  solves the following (nonlinear) system of equations:*

$$\frac{S_0}{S_k} = \sum_{i=1}^{\infty} 2^{-i} \frac{S_i}{S_{k+i}} \quad \text{for all } k \geq 1. \quad (13)$$

Let  $a_k = S_k/S_{k+1} \leq 1$  for all  $k \geq 0$ . Then, for any  $i \geq 0$  and any  $k > 0$  we have  $S_i/S_{k+i} = a_i a_{i+1} \dots a_{i+k-1}$ . The system (13) rewrites in terms of  $a_i$  as

$$\begin{aligned} \frac{1}{2}a_1 + \frac{1}{4}a_2 + \frac{1}{8}a_3 + \dots &= a_0, \\ \frac{1}{2}a_1a_2 + \frac{1}{4}a_2a_3 + \frac{1}{8}a_3a_4 + \dots &= a_0a_1, \\ \frac{1}{2}a_1a_2a_3 + \frac{1}{4}a_2a_3a_4 + \frac{1}{8}a_3a_4a_5 + \dots &= a_0a_1a_2, \end{aligned}$$

and so on, which can be summarized as

$$\sum_{j=1}^{\infty} \frac{1}{2^j} \prod_{k=j}^{n+j-1} a_k = \prod_{k=0}^{n-1} a_k, \quad \text{for all } n \in \mathbb{N}. \quad (14)$$

Assume that  $\{a_0, a_1, a_2, \dots\}$  is a solution to system (14). Then  $\{1, a_1/a_0, a_2/a_0, \dots\}$  is also a solution, since each equation only includes multinomial terms of the same degree.

**Lemma 2.** *The system (14) with  $a_0 = 1$  has a unique solution  $a_0 = a_1 = a_2 = \dots = 1$ .*

Lemma 2 implies  $a_k = S_k/S_{k+1} = 1/c$  for some  $c \geq 1$ . Hence  $S_1 = 1 + T_1 = c$  and  $T_1 = c - 1$ . Furthermore,

$$S_{k+1} = c S_k = c^k$$

and, accordingly,

$$T_{k+1} = S_{k+1} - S_k = (c - 1)c^{k-1},$$

which completes the proof.

### Proof of Lemma 1

*Proof.* Assume that the process is time invariant. Then the process progeny is constant in time and equals unity:

$$\|\pi\|_1 = \sum_{k=1}^{\infty} p(1 - p)^{k-1} = 1.$$

Observe that in one time step, every vertex of order  $\text{ord} = 1$  terminates, and any vertex of order  $\text{ord} > 1$  splits in two. Hence, the process progeny at  $s = 1$  is

$$2 \sum_{k=2}^{\infty} p(1-p)^{k-1} = 2(1-p) = 1,$$

which implies  $p = 1/2$ . Accordingly,  $p(1-p)^{k-1} = 2^{-k}$  and the time invariance (10) takes the following coordinate form

$$-\frac{2^{-k}}{S_{k-1}} + 2^{-(k+1)} \frac{T_1 + 2}{S_k} + \sum_{i=k+2}^{\infty} 2^{-i} \frac{T_{i-k}}{S_{i-1}} = 0, \text{ for all } k \geq 1. \quad (15)$$

Multiplying (15) by  $2^k$  and observing that  $T_k = S_k - S_{k-1}$  we obtain:

$$\begin{aligned} & -\frac{1}{S_{k-1}} + \frac{1}{2} \frac{T_1 + 2}{S_k} + \sum_{i=2}^{\infty} 2^{-i} \frac{T_i}{S_{k+i-1}} = 0, \\ & \frac{1}{S_{k-1}} - \sum_{i=1}^{\infty} 2^{-i} \frac{S_i}{S_{k+i-1}} \\ & = \frac{1}{S_k} - \frac{1}{2S_k} - \sum_{i=2}^{\infty} 2^{-i} \frac{S_{i-1}}{S_{k+i-1}}, \\ & \frac{1}{S_{k-1}} - \sum_{i=1}^{\infty} 2^{-i} \frac{S_i}{S_{k+i-1}} \\ & = \frac{1}{2} \left( \frac{1}{S_k} - \sum_{i=1}^{\infty} 2^{-i} \frac{S_i}{S_{k+i}} \right), \end{aligned} \quad (16)$$

We prove (13) by induction. Base: For  $k = 1$  we have

$$\begin{aligned} \frac{1}{2} &= \frac{1}{2S_1} + \sum_{i=1}^{\infty} 2^{-(i+1)} \frac{S_i - S_{i-1}}{S_i}, \\ 1 &= \frac{1}{S_1} + \sum_{i=1}^{\infty} 2^{-i} - \sum_{i=1}^{\infty} 2^{-i} \frac{S_{i-1}}{S_i}, \end{aligned}$$

which gives

$$\frac{1}{S_1} = \sum_{i=1}^{\infty} 2^{-i} \frac{S_i}{S_{i+1}}.$$

Step: Assume that the statement is proven for  $(k-1)$ . Then the left-hand side of (16) vanishes, and the right-hand part rewrites as (13). This establishes necessity.

Conversely, we showed that the system (13) is equivalent to (10) in case  $p = 1/2$ . This establishes sufficiency.  $\square$

## Proof of Lemma 2

*Proof.* We consider two cases.

*Case I.* Suppose the sequence  $\{a_j\}$  has a maximum: there exists an index  $i \in \mathbb{N}$  such that  $a_i = \max_{j \in \mathbb{N}} a_j$ . Define

$$w_{j,\ell} := \frac{1}{2^j} \prod_{k=j}^{\ell+j-1} a_k \left[ \prod_{k=0}^{\ell-1} a_k \right]^{-1}.$$

Using  $n = \ell$  in (14) we obtain that for any  $\ell \in \mathbb{N}$ ,

$$\sum_{j=1}^{\infty} w_{j,\ell} = 1, \quad (17)$$

and using  $n = \ell + 1$  we find that an arbitrary  $a_\ell$  is the weighted average of  $\{a_{\ell+j}\}_{j=1,2,\dots}$ :

$$\sum_{j=1}^{\infty} w_{j,\ell} a_{\ell+j} = a_\ell. \quad (18)$$

Hence, since  $a_i = \max_{j \in \mathbb{N}} a_j$ ,

$$a_i = a_{i+1} = a_{i+2} = a_{i+3} = \dots = a.$$

Similarly, letting  $\ell = i - 1$  in (17) and (18), we obtain  $a_{i-1} = a$ . Recursively, by plugging in  $\ell = i - 2, i - 3, \dots$ , we show that

$$a_1 = a_2 = \dots = a_{i-1} = a_i = a_{i+1} = \dots = a.$$

Finally,  $\frac{1}{2}a_1 + \frac{1}{4}a_2 + \frac{1}{8}a_3 + \dots = 1$  implies  $a = 1$ .

*Case II.* Suppose there is no  $\max_{j \in \mathbb{N}} a_j$ . Let  $U := \limsup_{j \rightarrow \infty} a_j$ .

From (14) we know via cancellation that

$$\begin{aligned} \frac{1}{2}a_n &+ \frac{1}{4} \frac{a_n a_{n+1}}{a_1} + \frac{1}{8} \frac{a_n a_{n+1} a_{n+2}}{a_1 a_2} + \dots \\ &+ \frac{1}{2^{n-1}} \frac{\prod_{k=n}^{2n-2} a_k}{\prod_{k=0}^{n-2} a_k} + \sum_{j=n}^{\infty} \frac{1}{2^j} \frac{\prod_{k=j}^{n+j-1} a_k}{\prod_{k=0}^{n-1} a_k} = 1. \end{aligned} \quad (19)$$

Thus,  $2^{-1}a_n < 1$  and  $U \leq 2$ . The absence of maximum implies  $a_j < U \leq 2$  for all  $j \in \mathbb{N}$ .

Plugging  $n + 1$  in (14), we obtain

$$\begin{aligned} & \left( \frac{1}{2}a_n \right) a_{n+1} + \left( \frac{1}{4} \frac{a_n a_{n+1}}{a_1} \right) a_{n+2} + \dots \\ & + \frac{1}{2^{n-1}} \frac{\prod_{k=n}^{2n-2} a_k}{\prod_{k=0}^{n-2} a_k} a_{n+j-1} + \sum_{j=n}^{\infty} \frac{1}{2^j} \frac{\prod_{k=j}^{n+j-1} a_k}{\prod_{k=0}^{n-1} a_k} a_{n+j} = a_n. \end{aligned}$$

Thus, since  $a_j < U$  for all  $j \in \mathbb{N}$ ,

$$\begin{aligned} & \left(\frac{1}{2}a_n\right) a_{n+1} + \left(\frac{1}{4} \frac{a_n a_{n+1}}{a_1}\right) U + \dots \\ & + \frac{1}{2^{n-1}} \frac{\prod_{k=n}^{2n-2} a_k}{\prod_{k=0}^{n-2} a_k} U + \sum_{j=n}^{\infty} \frac{1}{2^j} \frac{\prod_{k=j}^{n+j-1} a_k}{\prod_{k=0}^{n-1} a_k} U > a_n \end{aligned}$$

which simplifies via (19) to

$$\left(\frac{a_n}{2}\right) a_{n+1} + \left(1 - \frac{a_n}{2}\right) U > a_n. \quad (20)$$

For all  $\varepsilon \in (0, 1)$ , there are infinitely many  $n \in \mathbb{N}$  such that  $a_n > (1 - \varepsilon)U$ . Then, for any such  $n$ , the above inequality (20) implies

$$a_{n+1} > 2 - \frac{2}{a_n} U + U > 2 - \frac{2\varepsilon}{1 - \varepsilon} + U = (1 - \varphi(\varepsilon))U,$$

where

$$\varphi(x) := \frac{2x}{(1-x)U}.$$

Let  $\varphi^{(k)} = \varphi \circ \dots \circ \varphi$ . Repeating the argument for any given number of iterations  $K \in \mathbb{N}$ , we obtain

$$\begin{aligned} a_{n+2} &> (1 - \varphi^{(2)}(\varepsilon))U, \quad a_{n+3} > (1 - \varphi^{(3)}(\varepsilon))U, \\ &\dots, a_{n+K} > (1 - \varphi^{(K)}(\varepsilon))U. \end{aligned}$$

Thus, given any  $K \in \mathbb{N}$ , fix  $\varepsilon \in (0, 1)$  small enough so that such that  $\varphi^{(k)}(\varepsilon) \in (0, 1)$  for all  $k = 1, 2, \dots, K$ . Then, taking  $n > K$  such that  $a_n > (1 - \varepsilon)U$ , we obtain from (19) that

$$\begin{aligned} 1 &> \frac{1}{2}a_n + \frac{1}{4} \frac{a_n a_{n+1}}{a_1} + \frac{1}{8} \frac{a_n a_{n+1} a_{n+2}}{a_1 a_2} + \dots \\ & + \frac{1}{2^{K+1}} \frac{\prod_{k=n}^{n+K} a_k}{\prod_{k=0}^K a_k} \\ & > \frac{1}{2}(1 - \varepsilon)U + \frac{1}{4} \frac{(1 - \varepsilon)(1 - \varphi(\varepsilon))U^2}{U} + \dots \\ & + \frac{1}{2^{K+1}} \frac{U^{K+1} \prod_{k=0}^K (1 - \varphi^{(k)}(\varepsilon))}{U^K}. \end{aligned}$$

Now, since  $\varepsilon$  can be chosen arbitrarily small,

$$1 \geq \left(1 - \frac{1}{2^{K+1}}\right) U.$$

Finally, since  $K$  can be selected arbitrarily large, we have proven that  $1 \geq U$ . However, this will contradict the assumption of *Case II*. Indeed, if  $a_j < U \leq 1$  for all  $j \in \mathbb{N}$ , then

$$\frac{1}{2}a_1 + \frac{1}{4}a_2 + \frac{1}{8}a_3 + \dots < 1,$$

contradicting the first equation in the statement of the theorem. Thus, the assumptions of *Case II* cannot be satisfied. We conclude that there exists a maximal element in the sequence  $\{a_j\}_{j=1,2,\dots}$  as assumed in *Case I*, implying the statement of the theorem.  $\square$

### Proof of Corollary 1

*Proof.* Let  $\text{ord}(\mathcal{G})$  denote the (random) order of a random geometric tree  $\mathcal{G}$ . We apply a unit time shift  $\mathcal{S}$  to the tree  $\mathcal{G}$ . Conditioned on  $\text{ord}(\mathcal{G}) > 1$ , at instant  $s = 1$  there exist exactly two vertices that are the roots of the principal subtrees  $T_a$  and  $T_b$ . Since the trees  $T_a$  and  $T_b$  have the same distribution, their roots have the same order distribution. Denote by  $y_k$  the probability that the tree  $T_a$  has order  $k \geq 1$  and let  $\mathbf{y} = (y_1, y_2, \dots)^T$ . By Thm. 2 criticality is equivalent to time invariance. In a critical tree we have  $p = 1/2$ , which, together with time invariance, implies

$$\mathbf{x}(0) = \mathbf{x}(1) = 2\mathbf{y}\pi_1 + \mathbf{0}(1 - \pi_1) = 2\mathbf{y}\frac{1}{2} + \mathbf{0}\frac{1}{2} = \mathbf{y}.$$

This establishes the first statement.

The second statement follows from examining the joint distribution  $q_{a,b}$  of (7).  $\square$

### ACKNOWLEDGEMENT

The authors are grateful to the Editor and two referees whose comments helped to improve the initial version of this work. The authors acknowledge financial support from the National Science Foundation, awards DMS 1412557 (Y.K.) and EAR 1723033 (I.Z.).

---

\* kovchegy@math.oregonstate.edu

† zal@unr.edu

- [1] Richter, J. P. *The notebooks of Leonardo da Vinci*. Dover: New York, NY, USA (1970) Vol. 1. Unabridged edition of the work first published in London in 1883.
- [2] Rian, I. M. and Sassone, M. *Tree-inspired dendriforms and fractal-like branching structures in architecture: A brief historical overview*. *Frontiers of Architectural Research*, (2014) 3(3) 298–323.
- [3] Viennot, X. G. *Trees everywhere*. In *Colloquium on Trees in Algebra and Programming* (1990) 18–41. Springer, Berlin, Heidelberg.
- [4] Rodriguez-Iturbe, I. and Rinaldo, A. *Fractal river basins: chance and self-organization*. Cambridge University Press, (2001).
- [5] Tarboton, D. G., Bras, R. L., and Rodriguez-Iturbe, I. *The fractal nature of river networks*. *Water Resources Research*, (1988) 24(8) 1317–1322.



- [6] Banavar, J. R., Maritan, A., and Rinaldo, A. *Size and form in efficient transportation networks*. *Nature*, (1999) 399(6732) 130.
- [7] Maritan, A., Rinaldo, A., Rigon, R., Giacometti, A., and Rodriguez-Iturbe, I. *Scaling laws for river networks*. *Physical Review E*, (1996) 53(2) 1510.
- [8] Biswal, B. and Marani, M. *Geomorphological origin of recession curves*. *Geophysical Research Letters*, (2010) 37(24).
- [9] Kirchner, J. W., Feng, X., and Neal, C. *Fractal stream chemistry and its implications for contaminant transport in catchments*. *Nature*, (2000) 403(6769) 524.
- [10] Dodds, P. S. and Rothman, D. H. *Scaling, universality, and geomorphology*. *Annual Review of Earth and Planetary Sciences*, (2000) 28(1) 571–610.
- [11] Heckmann, T., Schwanghart, W., and Phillips, J. D. *Graph theory — Recent developments of its application in geomorphology*. *Geomorphology*, (2015) 243 130–146.
- [12] Newman, W. I., Turcotte, D. L., and Gabrielov, A. M. *Log-periodic behavior of a hierarchical failure model with applications to precursory seismic activation*. *Physical Review E*, (1995) 52(5) 4827.
- [13] Baiesi, M. and Paczuski, M. *Scale-free networks of earthquakes and aftershocks*. *Physical Review E*, (2004) 69(6) 066106.
- [14] Turcotte, D. L., Holliday, J. R., and Rundle, J. B. *BASS, an alternative to ETAS*. *Geophysical Research Letters*, (2007) 34(12).
- [15] Holliday, J. R., Turcotte, D. L., and Rundle, J. B. *Self-similar branching of aftershock sequences*. *Physica A: Statistical Mechanics and its Applications*, (2008) 387(4) 933–943.
- [16] Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H. *Clustering analysis of seismicity and aftershock identification*. *Physical Review Letters*, (2008) 101(1) 018501.
- [17] Gu, C., Schumann, A. Y., Baiesi, M., and Davidsen, J. *Triggering cascades and statistical properties of aftershocks*. *Journal of Geophysical Research: Solid Earth*, (2013) 118(8) 4278–4295.
- [18] Zaliapin, I. and Ben-Zion, Y. *Earthquake clusters in southern California I: Identification and stability*. *Journal of Geophysical Research: Solid Earth*, (2013) 118(6) 2847–2864.
- [19] Yoder, M. R., Van Aalsburg, J., Turcotte, D. L., Abaimov, S. G., and Rundle, J. B. *Statistical variability and Tokunaga branching of aftershock sequences utilizing BASS model simulations*. *Pure and Applied Geophysics*, (2013) 170(1-2) 155–171.
- [20] Gabrielov, A., Zaliapin, I., Newman, W. I., and Keilis-Borok, V. I. *Colliding cascades model for earthquake prediction*. *Geophysical Journal International*, (2000) 143(2) 427–437.
- [21] Zaliapin, I., Keilis-Borok, V., and Ghil, M. *A Boolean delay equation model of colliding cascades. Part I: Multiple seismic regimes*. *Journal of Statistical Physics*, (2003) 111(3-4) 815–837.
- [22] Zaliapin, I., Keilis-Borok, V., and Ghil, M. *A Boolean delay equation model of colliding cascades. Part II: Prediction of critical transitions*. *Journal of Statistical Physics*, (2003) 111(3-4) 839–861.
- [23] Kassab, G. S. *The coronary vasculature and its reconstruction*. *Annals of Biomedical Engineering*, (2000) 28(8) 903–915.
- [24] Cassot, F., Lauwers, F., Fouard, C., Prohaska, S., and Lauwers-Cances, V. *A novel three-dimensional computer-assisted method for a quantitative study of microvascular networks of the human cerebral cortex*. *Microcirculation*, (2006) 13(1) 1–18.
- [25] Campbell Grant, E. H., Lowe, W. H., and Fagan, W. F. *Living in the branches: population dynamics and ecological processes in dendritic networks*. *Ecology Letters*, (2007) 10(2) 165–175.
- [26] Turcotte, D. L., Pelletier, J. D., and Newman, W. I. *Networks with side branching in biology*. *Journal of Theoretical Biology*, (1998) 193(4) 577–592.
- [27] Pelletier, J. D. and Turcotte, D. L. *Shapes of river networks and leaves: are they statistically similar?* *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, (2000) 355(1394) 307–311.
- [28] Horton, R. E. *Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology*. *Geological Society of America Bulletin*, (1945) 56(3) 275–370.
- [29] Strahler, A. N. *Quantitative analysis of watershed geomorphology*. *Transactions of American Geophysical Union*, (1957) 38 913–920.
- [30] Kovchegov Y., and Zaliapin, I. *Horton Law in Self-Similar Trees*. *Fractals*, (2016) 24 1650017
- [31] Shreve, R. L. *Stream lengths and basin areas in topologically random channel networks*. *The Journal of Geology*, (1969) 77(4) 397–414.
- [32] Burd, G. A., Waymire, E. C., and Winn, R. D. *A self-similar invariance of critical binary Galton-Watson trees*. *Bernoulli*, (2000) 6 1–21.
- [33] Pitman, J. *Combinatorial Stochastic Processes*. *Lecture Notes in Mathematics*, (2006) 1875 Springer-Verlag.
- [34] Tokunaga, E. *Consideration on the composition of drainage networks and their evolution*. *Geographical reports of Tokyo Metropolitan University*, (1978) 13 1–27.
- [35] Peckham, S. D. *New results for self-similar trees with applications to river networks*. *Water Resources Research*, (1995) 31 1023–1029.
- [36] Tarboton, D. G. *Fractal river networks, Horton’s laws and Tokunaga cyclicity*. *Journal of Hydrology*, (1996) 187(1-2) 105–117.
- [37] Mantilla, R., Troutman, B. M., and Gupta, V. K. *Testing statistical self-similarity in the topology of river networks*. *Journal of Geophysical Research: Earth Surface*, (2010) 115(F3).
- [38] Zanardo, S., Zaliapin, I., and Foufoula-Georgiou, E. *Are American rivers Tokunaga self-similar? New results on fluvial network topology and its climatic dependence*. *Journal of Geophysical Research: Earth Surface*, (2013) 118(1) 166–183.
- [39] Newman, W. I., Turcotte, D. L., and Gabrielov, A. M. *Fractal trees with side branching*. *Fractals*, (1997) 5 603–614.
- [40] Ossadnik, P. *Branch order and ramification analysis of large diffusion-limited-aggregation clusters*. *Physical Review A*, (1992) 45(2) 1058.
- [41] Zaliapin, I., Wong, H. H. L., and Gabrielov, A. *Inverse cascade in a percolation model: Hierarchical description of time-dependent scaling*. *Physical Review E*, (2005) 71(6) 066118.
- [42] Yakovlev, G., Newman, W. I., Turcotte, D. L., and Gabrielov, A. *An inverse cascade model for self-organized complexity and natural hazards*. *Geophysical Journal In-*

- ternational, (2005) 163(2) 433–442.
- [43] Zaliapin, I. and Kovchegov, Y. *Tokunaga and Horton self-similarity for level set trees of Markov chains*. *Chaos, Solitons & Fractals*, (2012) 45(3) 358–372.
- [44] Gabrielov, A., Newman, W. I., and Turcotte, D. L. *Exactly soluble hierarchical clustering model: Inverse cascades, self-similarity, and scaling*. *Physical Review E*, (1999) 60(5) 5293.
- [45] Kovchegov, Y. and Zaliapin, I. *Horton self-similarity of Kingman’s coalescent tree*. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, (2017) 53(3) 1069–1107. Institut Henri Poincaré.
- [46] Zaliapin, I., Foufoula-Georgiou, E., and Ghil, M. *Transport on river networks: A dynamic tree approach*. *Journal of Geophysical Research: Earth Surface*, (2010) 115(F2).
- [47] Veitzer, S. A. and Gupta, V. K. *Random self-similar river networks and derivations of generalized Horton Laws in terms of statistical simple scaling*. *Water Resources Research*, (2000) 36(4) 1033–1048, doi:10.1029/1999WR900327.