PHYSICAL REVIEW E **105**, 014301 (2022)

Critical Tokunaga model for river networks

Yevgeniy Kovchegov [*]
Department of Mathematics, Oregon State University, Corvallis, Oregon 97331, USA

Ilya Zaliapin [†]
Department of Mathematics and Statistics, University of Nevada, Reno, Nevada 89557, USA

Efi Foufoula-Georgiou [‡]
Department of Civil and Environmental Engineering and Department of Earth System Science,
University of California, Irvine, California 92697, USA

(Received 4 June 2021; accepted 9 December 2021; published 3 January 2022)

The hierarchical organization and self-similarity in river basins have been topics of extensive research in hydrology and geomorphology starting with the pioneering work of Horton in 1945. Despite significant theoretical and applied advances, however, the mathematical origin of and relation among Horton's laws for different stream attributes remain unsettled. Here we capitalize on a recently developed theory of random self-similar trees to elucidate the origin of Horton's laws, Hack's laws, basin fractal dimensions, power-law distributions of link attributes, and power-law relations between distinct attributes. In particular, we introduce a one-parametric family of self-similar critical Tokunaga trees that includes the celebrated Shreve's random topology model and extends to trees that approximate the observed river networks with realistic exponents. The results offer tools to increase our understanding of landscape organization under different hydroclimatic forcings, and to extend scaling relationships useful for hydrologic prediction to resolutions higher than those observed.

DOI: 10.1103/PhysRevE.105.014301

I. INTRODUCTION

In a pioneering study "of streams and their drainage basins," Horton [1] introduced the concept of river stream order and formulated two fundamental laws of the composition of stream-drainage nets. The law of stream numbers postulates a geometric decay of the numbers N_K of streams of increasing order K, with the exponent R_B. The law of stream lengths postulates a geometric growth of the average length L_K of streams of increasing order K, with the exponent R_L. During the 20th century, geometric dependence on the stream order has been documented for multiple stream attributes, including upstream area, magnitude (number of upstream headwater channels, also called sources), the total channel length, the longest stream length, link slope, mean annual discharge, energy expenditure, etc. [2]; all such relations are also referred to as Horton's laws. Despite their elemental role in describing the key regularities in river stream networks (such as fractal dimension, Hack's law, etc.), Horton's laws remain an empirical finding and their origin and apparent ubiquity remain unsettled [3].

The first attempt at a rigorous explanation of Horton's laws was made by Shreve [4] in the 1960s, who examined a "topologically random population of channel networks," where all topologically distinct networks with given number of first-order streams are equally likely. This model is equivalent to the celebrated critical binary Galton-Watson branching process with a given population size [5,6]. Shreve's calculations imply that in this model the Horton's law of stream numbers holds with $R_B = 4$. Although not attempted by Shreve, it can be shown [5] that the law of stream lengths also holds here with $R_L = 2$ under the assumption of constant or identically distributed link lengths. Albeit insightful and mathematically tractable, the random topology model deviates from observations, which became apparent with the development of improved methods for extracting river networks [7,8]. This called for developing alternative modeling approaches for river networks.

It has proven challenging to find a model that would be mathematically tractable and flexible enough to reproduce the Horton exponents and other scaling laws observed in river basins. One end of the modeling spectrum is occupied by conceptual approaches, such as the Peano fractal basin ([2], their Sec. 2.4; [1], their Fig. 25) or Scheidegger's lattice model [9,10]. These models provide an invaluable insight into the origin of the observed scalings; they however lack realistic dendritic patterns and values of scaling exponents. On the other end are simulation approaches that can generate visually appealing networks that closely fit selected exponents, but can be analytically opaque. The optimal channel network (OCN) model [11–16] is a well-recognized simulation technique. Following the energy expenditure minimization principle, the model creates random drainage basins on a planar lattice (or

---

[*] kovchegy@math.oregonstate.edu

[†] zal@unr.edu

[‡] efi@uci.edu

more general graphs). We refer to [2] for a comprehensive discussion of these and other models.

Despite the progress achieved by the modeling efforts of the 20th century, the following essential questions remain unanswered: What are sufficient conditions for Horton's laws? What are the values of the Horton exponents for river basins? How are the Horton exponents for different stream attributes related to each other and to other basin parameters? There is a consensus that Horton's laws are connected to the self-similar structure of a basin [2,8,17,18], which is informally understood as invariance of the basin's statistical structure with respect to changing the scale of analysis (zooming in or out) [19]. However, a commonly accepted mathematical definition of tree self-similarity is still lacking. Three alternative definitions have been proposed: the Toeplitz property of the Tokunaga coefficients [8,20]; the invariance of a tree distribution with respect to the Horton pruning (cutting the source streams) [5]; and statistical self-similarity of basin attributes [17,21]. The unsettled questions are as follows: How are the alternative definitions related? Is self-similarity (any version) sufficient for selected Horton's laws? These questions are relevant in other areas beyond hydrogeomorphology where Horton's laws and related scalings have been reported, including computer science [22,23], statistical seismology [24–27], vascular analysis [28], brain studies [29], ecology [30], and biology [31].

We answer the above questions within a self-consistent mathematical theory of random self-similar trees recently developed by the authors [32]. In particular, we propose a concept of tree self-similarity that unifies the alternative existing definitions, rigorously explains the appearance and parametrization of Horton's laws, and offers an approach to modeling a variety of dendritic systems. The goal of this paper is to adapt and extend the theory to the studies of river networks. Most notably, we show that two fundamental and practically appealing properties—coordination and Horton prune invariance—result in trees that enjoy a wealth of scaling laws observed in landscape dissection. Furthermore, we propose a one-parametric family of critical Tokunaga trees, which reproduces multiple Horton's laws and related scalings reported in river network studies, with realistic values of the respective parameters. The critical Tokunaga family yields rigorous relations among scaling exponents that have been empirically documented in multiple earlier studies, and serves as a useful analytic and modeling tool for further analysis of river network structure and dynamics. Our results also offer a computationally efficient algorithm of generating self-similar trees with arbitrary parameters (Horton exponents, fractal dimensions, etc.), which facilitates ensemble simulations.

We represent a stream network that drains a single basin (watershed, catchment) as a rooted binary tree. The basin outlet (point furthest downstream) corresponds to the tree root, sources (points furthest upstream) correspond to leaves, junctions (points where two streams meet) correspond to internal vertices, and links (stream segments between two successive nodes) correspond to edges. This graph-theoretic nomenclature provides a link to the probability and combinatorics literature on the topic. We assume that all examined trees belong to the space $\mathcal{L}$ of finite binary rooted planted trees with positive edge lengths [32]. Recall that a tree is called *planted* if the tree root has degree 1 (the most downstream link goes to an ocean or another large water body instead of merging with another stream). The space $\mathcal{L}$ includes the empty tree ϕ composed of the root vertex. We also consider the space $\mathcal{T}$ of combinatorial projections of trees from $\mathcal{L}$, that is, trees with the same combinatorial structure but no edge lengths.

II. RESULTS

A. Review of Horton's laws and implied scaling relationships

The Horton-Strahler orders for river streams have been discussed extensively in the literature and nicely reviewed and summarized in [2]. In this section we introduce the orders through the viewpoint of Horton pruning—this streamlines our exposition and prepares the reader for material that follows. We also discuss Horton's laws and their implications.

Consider the map $\mathcal{R} : \mathcal{L} \to \mathcal{L}$ that removes the source links from a tree $T \in \mathcal{L}$. The Horton-Strahler order of a tree T [1,33,34] is the minimal number of Horton prunings that completely eliminates it:

$$\mathrm{ord}(T) = \min\{k \geqslant 0 \,:\, \mathcal{R}^k(T) = \phi\}. \tag{1}$$

The Horton pruning and Horton-Strahler orders are illustrated in Fig. 1 (see Appendix A for details and an alternative computational definition).

Horton's law of stream numbers [1] postulates a geometric decay of the stream counts N_K of increasing order K with Horton exponent $R_B \geqslant 2$:

$$N_K \propto R_B^{-K} \quad \text{or} \quad \frac{N_K}{N_{K+1}} = R_B. \tag{2}$$

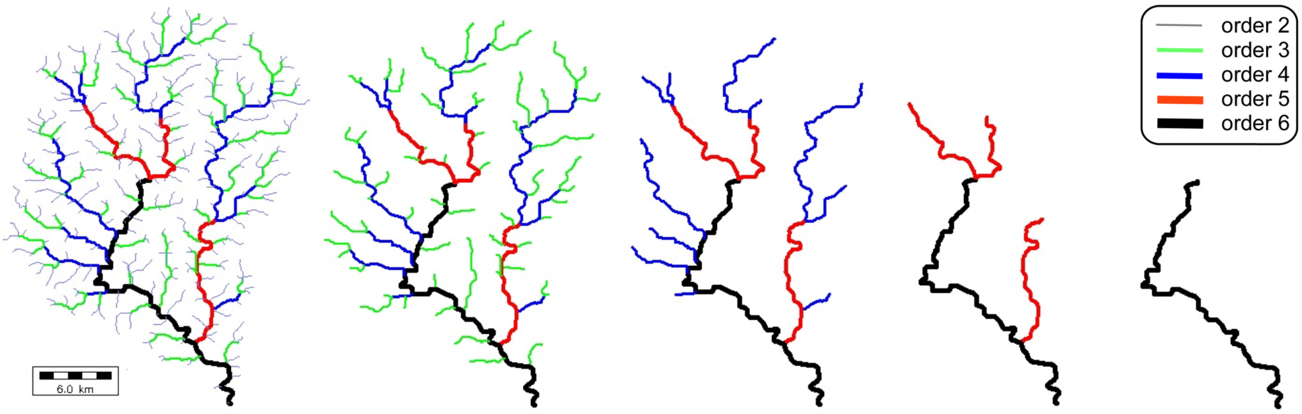
The notation $x \propto y$ stands for $x = \text{const} \times y$. The lower bound on R_B follows from the definition of Horton-Strahler orders, since it takes at least two streams of order K to create a single stream of order $K + 1$ (see Appendix A). The value of R_B reported in large river basins is close to 4.5 [2,4,17,19,21,33,35–39]. Figure 2(a) (cyan circles) shows the Horton's law for stream numbers in the stream network of Beaver Creek of Fig. 1; here $R_B \approx 4.6$.

Horton's law of stream lengths [1] postulates a geometric growth of the average length L_K of streams of increasing order K with exponent R_L:

$$L_K \propto R_L^{K} \quad \text{or} \quad \frac{L_{K+1}}{L_K} = R_L. \tag{3}$$

The value of R_L in river networks is around 2.5 [2]. Figure 2(d) (magenta squares) shows the Horton's law for stream lengths in the Beaver Creek network of Fig. 1; here $R_L \approx 2.3$.

Similarly to the laws (2) and (3) discussed above, a geometric scaling of any average stream attribute Z_K with order K is also called *Horton's law*, and the respective exponent is called *Horton exponent* [8,20]. Horton's laws are documented for multiple physical and combinatorial attributes, including upstream area, magnitude (number of upstream sources), total upstream channel length, length of the longest channel to the divide, etc. [2]. Figure 2 illustrates Horton's laws for seven stream attributes of the Beaver Creek network that is shown in Fig. 1. We use a convention that the Horton exponent is greater than unity, which always can be achieved by selecting



(a) 356 streams (b) 80 streams (c) 15 streams (d) 4 streams (e) 1 stream

FIG. 1. Horton pruning of the stream network of Beaver creek, Floyd County, KY. Streams of orders $K = 2, \ldots, 6$ are shown by different colors (see legend on the right). Streams of order 1 (source streams) are not shown for visual convenience. (a) The first Horton pruning, after eliminating streams of order $K = 1$. (b–e) Second to fifth consecutive Horton prunings. The sixth pruning completely eliminates the network. The channel extraction is done using RIVERTOOLS software [45].

the sign of the exponent K in the Horton's law [e.g., Eqs. (2) and (3)].

Horton's laws imply power-law frequencies of link attributes and power-law relations between the average values of distinct attributes. Suppose that stream attributes Z and Y satisfy Horton's law with Horton exponents R_Z and R_Y, respectively. Using each of the laws to express the channel order K and equating these expressions, we find

$$Z_K \propto Y_K^{\alpha}, \quad \text{with } \alpha = \frac{\log R_Z}{\log R_Y}. \tag{4}$$

Equation (4) is a punctured (by discrete orders) version of a power-law relation $Z \propto Y^{\alpha}$ that abounds among hydrologic quantities. A well-studied example is Hack's law that relates the length L of the longest stream in a basin to the basin area A via $L \propto A^{\mathbf{h}}$ with $\mathbf{h} \approx 0.6$ [40,41]. Equation (4) suggests that the parameter $\mathbf{h}$ is expressed via the exponents R_L and R_A of the Horton's laws for length L and area A as

$$\mathbf{h} = \frac{\log R_L}{\log R_A}. \tag{5}$$

Next, consider the value $Z_{(i)}$ of an attribute Z calculated at link i in a large basin. Assuming Horton's laws for N_K and S_K with exponents R_B and R_S, respectively, and considering a limit of an infinitely large basin we approximate the distribution of link attributes as (see Appendix B)

$$|i : Z_{(i)} \geqslant z| \propto z^{-\beta}, \quad \beta = \frac{\log R_B - \log R_S}{\log R_Z}, \tag{6}$$

where $|A|$ denotes the number of elements in a set A.

Such power laws are documented for the upstream contributing area, length of the longest channel to the divide, water discharge, or energy expenditure. For example, analyses of river basins (e.g., [14,42,43]) extracted from digital elevation models suggest

$$|i : A_{(i)} \geqslant a| \propto a^{-\beta_A}, \quad \beta_A \approx 0.43 \tag{7}$$

and

$$|i : \Lambda_{(i)} \geqslant l| \propto l^{-\beta_\Lambda}, \quad \beta_\Lambda \approx 0.8, \tag{8}$$

where $A_{(i)}$ is the area upstream of link i and $\Lambda_{(i)}$ is the distance from link i to the furthest source (or, equivalently, to the basin divide) measured along the channel network.

Horton's laws [e.g., Eqs. (2) and (3)] and the implied scaling relations [Eqs. (4) and (6)] provide key observational constraints in modeling river networks [2,12,39,42]. Our paper explains the appearance of Horton's laws in terms of tree self-similarity and offers a parametric toolbox for the analysis and modeling of river networks and other branching structures that exhibit such scaling relations.

B. Tree self-similarity and Tokunaga sequence

We introduce the concept of self-similarity for random trees that encompasses the existing definitions and satisfies practical intuitive expectations. The proposed definition applies to a distribution of trees from a suitable space such as $\mathcal{T}$ or $\mathcal{L}$ and combines two fundamental properties—coordination and Horton prune invariance.

Coordination means that the (random) structure of a river basin is determined by its order. For example, a basin with outlet of order 3 and a sub-basin of order 3 within a basin of order 9 have, statistically, the same structure. This assumption is at the heart of analyses based on the Horton-Strahler orders; it has been imposed, explicitly or implicitly, in the mainstream studies of river networks [1,2,4,7,8,21,36]. A distribution that satisfies the coordination property is called *coordinated*. We refer to [32] for a measure-theoretic definition of coordination.

Horton prune invariance formalizes the expectation that the scaling laws of hydrology are (by and large) independent of the data resolution. The Horton pruning $\mathcal{R}$ is a natural model for the change of resolution in a river network. Indeed, better observations lead to detecting smaller streams, which increases the basin order. Pruning a basin by order is roughly equivalent to decreasing the resolution of stream detection. The Horton prune invariance requires that the statistical structure of trees remains the same after zooming in or out.

Definition 1 (Self-similar tree). A coordinated distribution μ on the space $\mathcal{T}$ of combinatorial trees is called *self-similar*

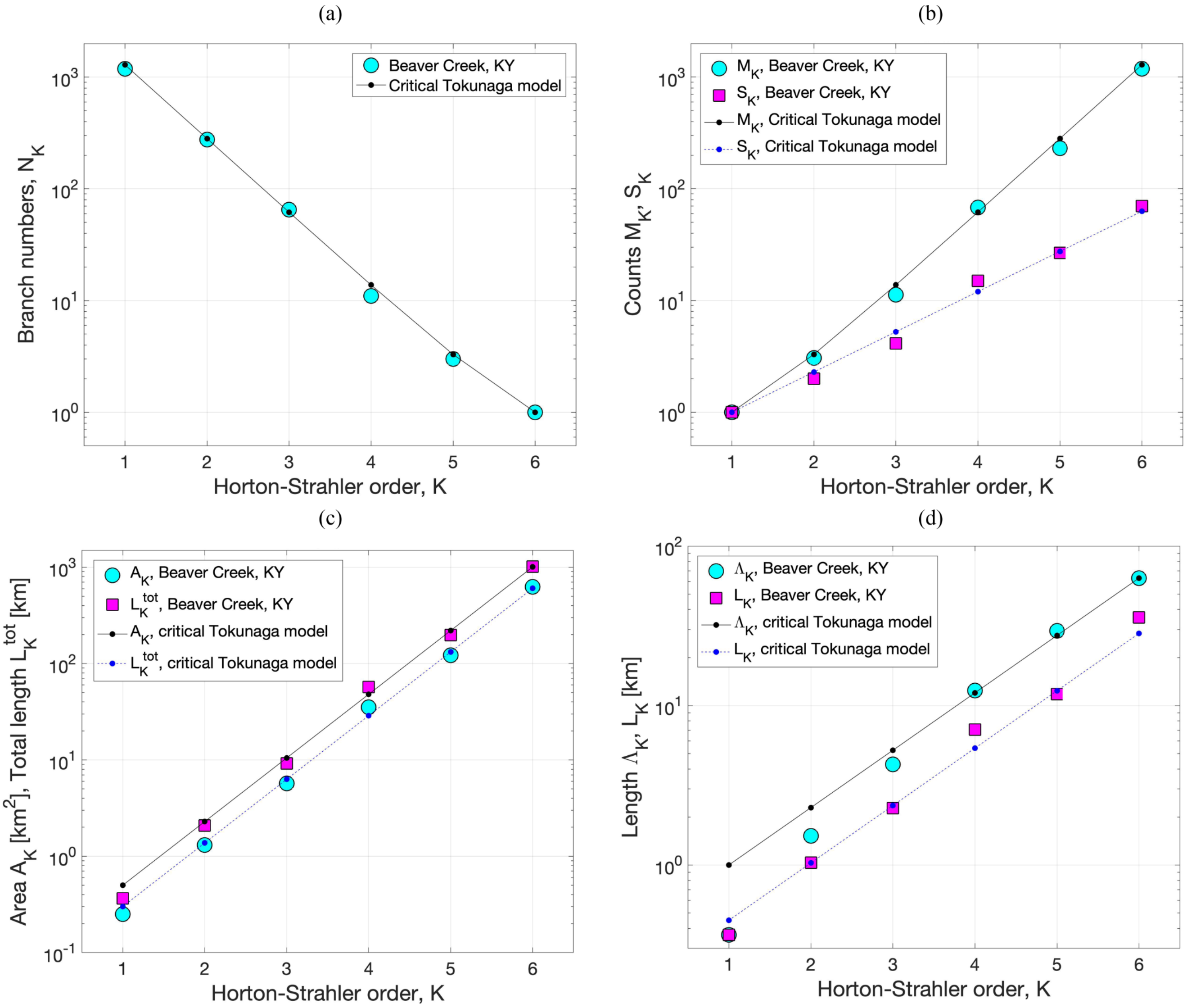


FIG. 2. Critical Tokunaga fit to the Horton's laws in the stream network of Beaver creek. The stream network is shown in Fig. 1. Symbols correspond to the values of the observed attributes. Lines and dots show the theoretical fit by the critical Tokunaga model with $c = 2.3$. (a) Stream numbers N_K. The model fit is given by (18); it has asymptotic slope $-\log_{10}(2c) \approx -0.66$. (b) Average stream magnitude M_K (cyan circles) and average number of links S_K (magenta squares). The fit for M_K is given by (19); it has asymptotic slope $\log_{10}(2c) \approx 0.66$. The fit for S_K is given by c^{K-1}; it has theoretical slope $\log_{10} c \approx 0.36$. (c) Average total contributing area A_K (cyan circles) and average total upstream channel length $L_K^{\rm tot}$ (magenta squares). The fitting lines, according to a combination of Eqs. (19) and (G3), have theoretical slope $\log_{10}(2c) \approx 0.66$. (d) Average stream length L_K (magenta squares) and average length Λ_K of the longest stream to the divide (cyan circles). The fitting lines, according to Eqs. (28) and (30), have theoretical slope $\log_{10}(c) \approx 0.36$.

if it is invariant with respect to Horton pruning [5,44]:

$$\mu[\mathcal{R}^{-1}(T)|T \neq \phi] = \mu(T) \quad \text{for any} \quad T \in \mathcal{T}. \tag{9}$$

Recall that ϕ denotes an empty tree. Equation (9) states that, for any nonempty tree T, the total probability assigned by μ to the collection of trees that result in T after pruning—these trees are denoted by $R^{-1}(T)$—is the same as the probability of T. Informally, consider a forest of trees generated by measure μ, where each tree T occurs multiple times according to its probability $\mu(T)$. The forest is self-similar if after pruning each tree by $\mathcal{R}$ we obtain the same forest. This definition can be extended to trees with edge lengths from space $\mathcal{L}$; see Def. 9 in [32]. In that case, we allow the edge lengths to scale after pruning by a multiplicative scaling constant $\zeta > 0$. We use a conventional abuse of terminology by saying that a tree T is self-similar if T is a random tree drawn from a self-similar distribution μ.

A measure-theoretic Definition 1 might be not appealing for practical analyses that oftentimes involve only a handful of finite basins. A bridge from this definition to easily computed stream statistics is provided by the Tokunaga sequence.

Definition 2 (Tokunaga coefficients [5,8,32,48,53]). Fix a coordinated measure μ on $\mathcal{T}$ and denote by $n_{i,j}$ the random number of streams of order i per a randomly selected stream of order j with respect to μ. For any pair $i < j$, the Tokunaga coefficient $T_{i,j} = T_{i,j}(\mu)$ is the expected value of $n_{i,j}$:

$$T_{i,j} = E_\mu[n_{i,j}], \tag{10}$$

where E_μ denotes the mathematical expectation with respect to μ.

We can arrange the Tokunaga coefficients for trees of a given order K in an upper triangular matrix:

$$\mathbb{T}_K = \begin{bmatrix} 0 & T_{1,2} & T_{1,3} & \dots & T_{1,K} \\ 0 & 0 & T_{2,3} & \dots & T_{2,K} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & T_{K-1,K} \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \tag{11}$$

Theorem 1 (Tokunaga sequence [32]). Suppose μ is a self-similar measure on $\mathcal{T}$. Then the Tokunaga coefficients satisfy the Toeplitz property: $T_{i,i+k} = T_k$ for any positive integer pair i, k. In this case the Tokunaga matrix becomes Toeplitz:

$$\mathbb{T}_K = \begin{bmatrix} 0 & T_1 & T_2 & \dots & T_{K-1} \\ 0 & 0 & T_1 & \dots & T_{K-2} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & T_1 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \tag{12}$$

Proof. Consider the pushforward measure $\mathcal{R}_*(\mu)$ induced on $\mathcal{T}$ by the Horton pruning operator:

$$\mathcal{R}_*(\mu)(A) = \mu[\mathcal{R}^{-1}(A)], \quad \forall A \subset \mathcal{T}. \tag{13}$$

Since Horton pruning decreases the order of every stream by 1, the Tokunaga coefficients $T^R_{i,j}$ computed on $\mathcal{T}$ with respect to $\mathcal{R}_*(\mu)$ satisfy $T^R_{i,j} = T_{i+1,j+1}$. The self-similarity of μ implies $T^R_{i,j} = T_{i,j}$. Combining these relations, we find $T_{i,j} = T_{i+1,j+1}$. This establishes the desired Toeplitz property of the Tokunaga coefficients. ■

The proof of Theorem 1 shows that in coordinated trees both prune invariance and Toeplitz property take the same algebraic form $T_{i,j} = T_{i+1,j+1}$. This leads to the following corollary.

Corollary 1 (Prune invariance vs Toeplitz). Suppose μ is a coordinated measure on $\mathcal{T}$. Then the Toeplitz property $T_{i,i+k} = T_k$ and Horton prune invariance of Eq. (9) are equivalent (i.e., both hold or do not hold at the same time).

We refer to the elements T_k of the Tokunaga sequence as Tokunaga coefficients, which creates no confusion with the original double-indexed coefficients $T_{i,j}$. According to Theorem 1, each self-similar measure μ corresponds to a unique non-negative sequence of Tokunaga coefficients T_k such that

$$T_k = T_{i,i+k} = E_\mu[n_{i,i+k}] \quad \text{for all} \quad i, k > 0. \tag{14}$$

The Tokunaga coefficients T_k provide a fundamental parametrization of a self-similar tree and constitute the main tool of respective applied analyses.

Our Definition 1 of tree self-similarity unifies the alternative definitions used in the literature. Burd *et al.* [5] define self-similarity in Galton-Watson trees as the Horton prune invariance; this is a special case of our definition since the Galton-Watson trees are coordinated [32]. Peckham [8] and Newman *et al.* [20] define self-similarity as the Toeplitz property for Tokunaga coefficients; this is equivalent to our definition in coordinated trees (Corollary 1). The coordination assumption is further justified in [32] by showing that the Toeplitz property alone, without coordination, allows for a multitude of obscure measures that are hardly useful in practice. Gupta and Waymire [17] and Peckham and Gupta [21] suggested a concept of statistical self-similarity that requires a random stream attribute Z to have distribution that scales with order. It can be shown (Sec. 7 in [32]) that (i) statistical self-similarity for some attributes (e.g., for any discrete attribute) may only hold asymptotically and (ii) multiple attributes, including stream length, magnitude, and total basin length, are statistically self-similar in a limit of an infinitely large basin that is self-similar according to our Definition 1.

C. Horton's laws for stream numbers and magnitudes in self-similar trees

We now capitalize on the concept of tree self-similarity introduced above to establish a key emergent property of self-similar trees—Horton's laws for stream numbers and magnitudes, conveniently parametrized by the Tokunaga sequence.

Consider the mean number

$$\mathcal{N}_i[K] = E_\mu[N_i(T) \,|\, \mathrm{ord}(T) = K]$$

of streams of order i in a basin of order K and the mean magnitude (number of upstream sources) M_i of a stream of order i. Observe that for a fixed K the stream counts $\mathcal{N}_i[K]$ form a decreasing sequence in i, and the sequence's first term $\mathcal{N}_1[K]$ increases with K. At the same time, the average magnitudes M_i form an increasing sequence in i; its first elements $M_1, \dots, M_j$ are independent of basin order K for any $K \geqslant j$. This explains the notational dependence on K in the average stream counts and absence of such in the average magnitudes. The definition implies $\mathcal{N}_K[K] = M_1 = 1$ and $\mathcal{N}_1[K] = M_K$ for any tree distribution. Moreover, in self-similar trees the two sequences are deterministically related as [8,44]

$$\mathcal{N}_{K-j+1}[K] = \mathcal{N}_1[j] = M_k \quad \text{for any} \quad 1 \leqslant j \leqslant K. \tag{15}$$

Consider the generating function $T(z) = \sum_{k=1}^{\infty} T_k z^k$ of the Tokunaga coefficients and define

$$\hat{t}(z) = -1 + 2z + T(z). \tag{16}$$

Theorem 2 (Horton's law for stream numbers, magnitudes in self-similar trees [32,44]). Consider a self-similar tree T with Tokunaga sequence T_k and suppose that

$$\limsup_{k\to\infty} T_k^{1/k} < \infty. \tag{17}$$

Then, the stream numbers $\mathcal{N}_i[K]$ and the magnitudes M_i obey Horton's laws:

$$\lim_{K\to\infty} \left(\mathcal{N}_1[K]\, R_B^{-K}\right) = M < \infty, \tag{18}$$

$$\lim_{i\to\infty} \left(M_i\, R_M^{-i}\right) = M < \infty. \tag{19}$$

The Horton exponents are given by

$$R_B = R_M = 1/w_0, \tag{20}$$

where w_0 is the only real root of the function $\hat{t}(z)$ in the interval $(0, 1/2]$ and M is a positive real constant given by

$$M = -\frac{1}{w_0}\lim_{z\to w_0}\frac{z(z-w_0)}{\hat{t}(z)}. \tag{21}$$

Proof is based on the analysis of generating functions for the sequences $\mathcal{N}_i[K]$ and M_i (Appendix C); it is given in Appendix D.

Theorem 2 implies that Horton's laws for mean stream numbers $\mathcal{N}_i[K]$ and magnitudes M_i hold in almost any self-similar tree. More specifically, the theorem shows that the validity of the laws is determined by the Tokunaga sequence T_k, and not the whole distribution μ. Furthermore, the only restriction on the admissible Tokunaga sequences is given by Eq. (17), which prohibits superexponential growth of T_k, such as $T_k = k!$ or $T_k = k^k$. The theorem establishes a strong form of Horton's law [Eqs. (18) and (19)], which implies a weaker version that is often reported in applied literature:

$$\lim_{K\to\infty}\frac{\mathcal{N}_i[K]}{\mathcal{N}_{i+1}[K]} = R_B \quad \text{for any } i, \text{ and}$$

$$\lim_{i\to\infty}\frac{M_{i+1}}{M_i} = R_M. \tag{22}$$

Theorem 2 emphasizes the existence of a multitude of self-similar measures with the same Horton exponent. Assume we fix R_B and hence the root w_0 of $\hat{t}(z)$ according to Eq. (20). Equation (16) readily asserts that there is an infinite number of Tokunaga sequences that correspond to an arbitrary w_0 within $(0, 1/2]$. For example, if $R_B = 4$, then $w_0 = 1/4$ and one needs $T(z) = 1/2$. This can be achieved by selecting any of $T_k = \{2, 0, \dots\}$, $\{1, 4, 0, \dots\}$, $\{0, 8, 0, \dots\}$, $\{2^{k-1}\}$, etc., where "…" denotes trailing zeros.

We observe that Horton's law of stream numbers in Theorem 2 [Eq. (18)] is an asymptotic statement, different from the ideal Horton's law for stream numbers (2) which is commonly used in the literature. This is not a mathematical peculiarity—the ideal Horton's law is merely a convenient approximation to the actual behavior of stream counts. Its approximate nature is not related to the finite size of the observed basins—the ideal Horton's law rarely holds in theoretical trees of arbitrarily large size. Formally, we show in Appendix E that the ideal Horton's law for stream numbers in a self-similar tree holds if and only if $T_k = 0$ for $k > 1$. Realistically, Horton's laws are asymptotic statements of different strength. The strongest form of Horton's law for stream numbers is that of Eq. (18), which implies a weaker version of Eq. (22). Accordingly, the power relations among different stream attributes (4) and power-law frequencies of link attributes (6) that we have derived from the ideal Horton's law of Eq. (2) remain heuristic. A formal analysis based on actual Horton's laws [like those in Eqs. (18) and (19)], which will be presented elsewhere, confirms the results of Eqs. (4) and (6) and reveals additional solutions with oscillatory tail behavior.

The asymptotic nature of Horton's laws triggers a natural question of whether one can observe them in finite trees. The most general approach to answering this question is given by the implicit form of the average branch counts [32]:

$$\mathcal{N}_1[K+1] = -\widetilde{\left(\frac{1}{\hat{t}}\right)}(K), \tag{23}$$

where we write

$$\widetilde{f}(k) = \frac{1}{2\pi i}\oint_{|z|=\Delta}\frac{f(z)}{z^{k+1}}dz = a_k \tag{24}$$

for an analytic function $f(z)$ represented by a power series $f(z) = \sum_{k=0}^{\infty} a_k z^k$ in a nonempty disk $|z| \leqslant \Delta$ (see Appendix C). Informally, the rate of convergence in Horton's law for stream numbers is determined by how well the only real zero of $\hat{t}(z)$ within $(0, 1/2]$ is separated from other zeros. A comprehensive treatment of this issue, including estimation of the convergence rates in the general case, can be found in [32, Sec. 4].

In many cases the sequence $\mathcal{N}_i[K]$ can be found explicitly. For example, consider the Tokunaga sequence $T_k = (c-1)c^{k-1}$ with $c > 1$; it corresponds to the critical Tokunaga tree discussed below in Sec. II E. Kovchegov and Zaliapin [[32], Sec. 7.6.3] have shown that in this case $R_B = 2c$ and

$$\mathcal{N}_{K-j+1}[K] = \frac{R_B^j + R_B - 2}{2(R_B - 1)}, \quad j = 1, \dots, K.$$

Accordingly, with $c = 2.3$ (which we suggest in this paper as a suitable value for the observed river basins) the sequence $\mathcal{N}_1[K]R_B^{-K}$ takes the form

$$0.217, 0.156, 0.143, 0.140, 0.139, \dots$$

with all later terms being equal to the theoretical limit value of 0.139, all values being rounded to the third significant digit. In other words, the Horton's law asymptotic (18) is closely attained starting from $K \gtrapprox 3$. This fast convergence is typical in Horton's laws.

Horton's laws for other stream attributes may or may not hold depending on additional assumptions about T_k and other details of basin organization. A comprehensive treatment is possible using the generating function approach outlined in Appendix C. Most importantly, further analysis often requires specifying a concrete self-similar distribution μ, not only its Tokunaga sequence T_k. Below we examine a particularly useful family of distributions.

D. Random attachment model of self-similar trees

According to Theorem 1, every self-similar measure corresponds to a unique Tokunaga sequence T_k. At the same time, a multitude of self-similar measures can be constructed for a given Tokunaga sequence. Here we introduce a particularly symmetric random tree (tree distribution) for a given Tokunaga sequence and establish its key properties. We use Poisson attachment construction within exponential segments; this ensures that the link lengths have exponential distribution and the attachment of streams of lower orders to a given stream of a larger order is done in uniform random fashion. We refer to this construction as the *random attachment model* (RAM).

The RAM specifies a tree distribution on $\mathcal{L}$ by a non-negative Tokunaga sequence T_k, the order distribution $\pi_K = P[\text{ord}(T) = K]$, and the distribution of stream lengths. The model assumes that the lengths of streams of order j are independent exponential random variables with rate λ_j. Hence, the model is specified by three non-negative sequences:

$$\{T_k\}, \ \{\lambda_j\}, \ \text{and } \{\pi_K\}.$$

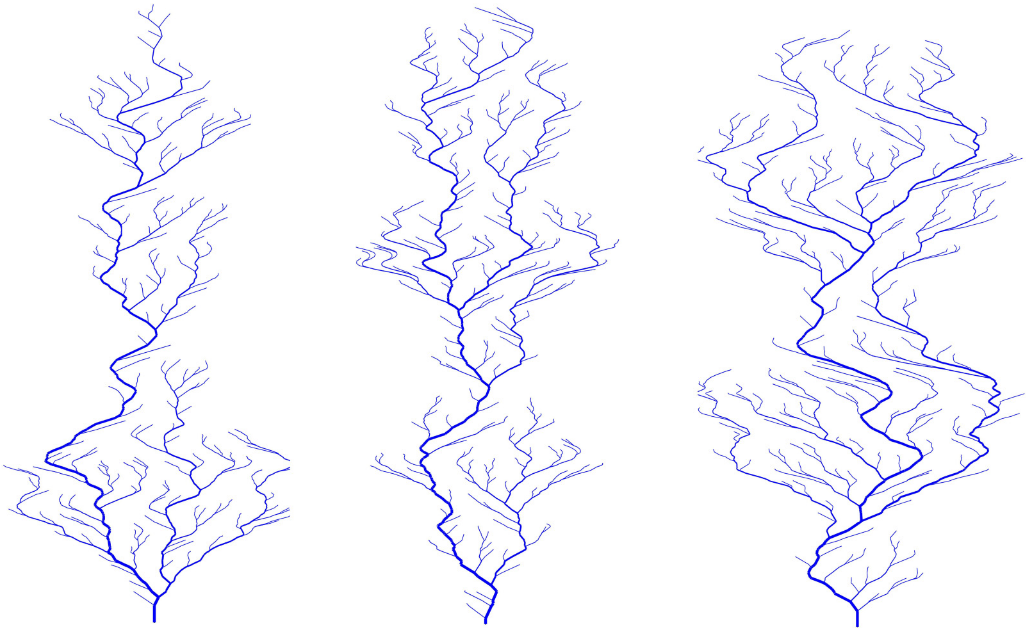


FIG. 3. Examples of critical Tokunaga trees (special case of RAM). The trees are generated by the critical Tokunaga process with $c = 2.3$ and Horton-Strahler order $K = 5$. The linewidth is proportional to the contributing area. The figure accurately represents the tree combinatorial structure; the edge lengths are scaled for a better planar embedding. We notice that the RAM generates trees with no planar embedding. The current figure uses an *ad hoc* embedding; accordingly, the related purely geometric properties, such as junction angles or spacing between channels, are not a part of the model.

Each sequence specifies a particular attribute of a random tree: the Tokunaga coefficients T_k specify the combinatorial structure of a tree of a given order; the probabilities π_K specify the frequencies of trees of different orders; and the rates λ_j specify the link lengths.

A random tree is constructed in a hierarchical fashion, starting from the stream of the highest order and adding side tributaries of consecutively smaller orders. The tree order K is selected according to the distribution π_K. At the first step we generate the main stream that will have order K in the final tree; its length is an exponential random variable with rate λ_K. At each of the remaining $K-1$ steps, we add streams of lower orders to the existing tree by a Poisson attachment procedure. The streams added at step m will have order $i(m) = K - m + 1$ in the final tree. The lengths of the newly added streams are independent exponential random variables with rate $\lambda_{i(m)}$. The new streams are added in two steps. First, we consider the existing tree as a one-dimensional metric space (union of link segments) and generate a collection of points on this space according to a homogeneous Poisson process. The process intensity depends on the order of a link within the final tree. Specifically, within every link added at step $K - j + 1$ (that will have order j in the final tree) the Poisson intensity is $\lambda_j T_{j-i(m)}$. A single stream is then attached to each Poisson point. Second, we add two new streams to each source stream of the current tree (except the sources just added during this step). The first part of this procedure (Poisson attachment) ensures that the tree has Tokunaga coefficients T_k, and the second part (adding stream pairs) increases the tree order by one at each step.

The trees generated by RAM can be equivalently represented as trajectories of a continuous-time multitype hierarchical branching process, with time evolving from the root upstream and member types corresponding to the stream orders. This approach, explored by the authors in [32], yields the joint distribution of the orders $K_1 < K_2$ of subtrees that share a common root of order K:

$$P(K_1 = j, K_2 = m|K) = \begin{cases} S_K^{-1} & \text{if } j = m = K-1, \\ T_{K-j}S_K^{-1} & \text{if } j < m = K. \end{cases} \tag{25}$$

We now use this result to propose a computationally efficient recursive construction of RAM trees. A tree of order 1 is a stream of exponential length with rate γ. To create a tree of order $K > 1$ we first generate a link of exponential length with rate $\lambda_K S_K$, where

$$S_K = 1 + \sum_{i=1}^{K-1} T_i. \tag{26}$$

To this link we attach two conditionally (conditioned on the order K) independent trees the orders of which are drawn from Eq. (25). Each of these trees is generated according to the same recursive procedure. This algorithm generates trees with up to 10^6 edges within seconds, providing a flexible computational framework for ensemble simulations based on independent statistical realizations of a tree with fixed parameters. Examples of RAM stream networks are shown in Fig. 3.

Another useful result of the branching process theory establishes the necessary and sufficient conditions for a RAM tree to be self-similar according to Definition 1. These conditions explicitly parametrize the probabilities π_K and stream length rates λ_j for an arbitrary Tokunaga sequence T_k. This emphasizes the richness of the self-similar family.

Theorem 3 (Self-similar RAM; [32, Theorem 11]). Suppose T is a random tree generated by the RAM with parameters

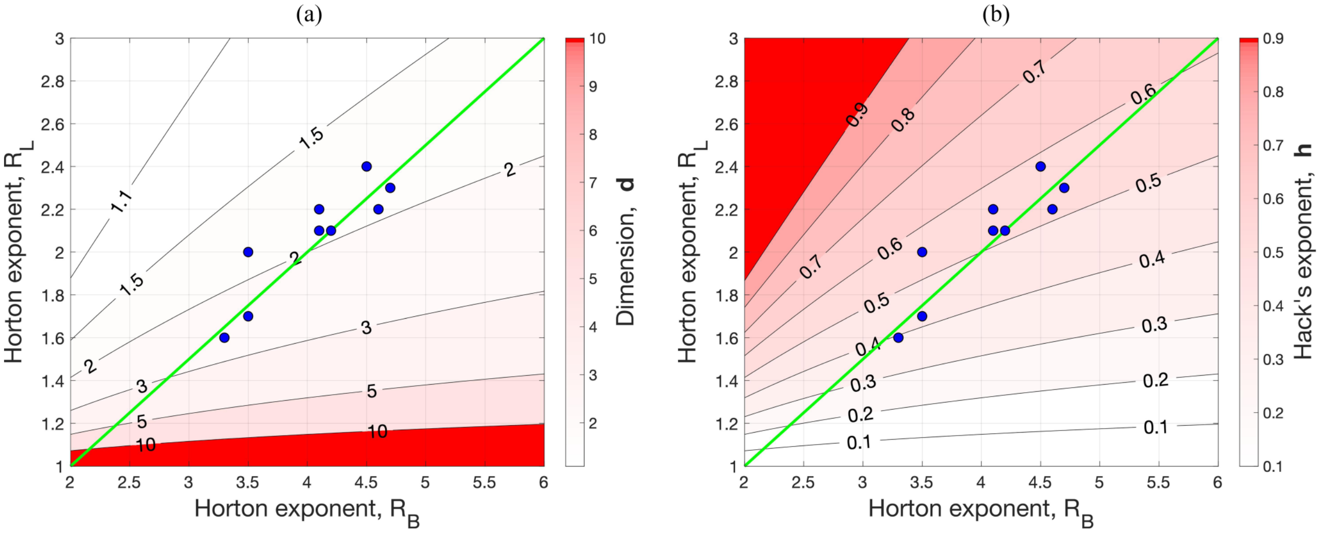


FIG. 4. Fractal dimension $\mathbf{d}$ (a) of Eq. (29) and Hack's exponent $\mathbf{h} = \mathbf{d}^{-1}$ (b) of a self-similar RAM tree in the limit of infinite size as a function of the Horton exponents R_B and R_L. Selected levels of $\mathbf{d}$ and $\mathbf{h}$ are shown by marked black lines. Green thick lines correspond to the critical Tokunaga process [Definition 3, Eq. (31)] for which $R_B = 2R_L$. Blue dots depict the pairs (R_B, R_L) estimated in nine real river basins by [42] (see also [2, Table 2.1]).

$\{T_k\}$, $\{\lambda_j\}$, and $\{\pi_K\}$. Then T is a coordinated tree. Tree T is self-similar with scaling constant $\zeta > 0$ (see Definition 1 and its discussion) if and only if

$$\pi_K = p(1-p)^{K-1} \quad (K \geqslant 1) \quad \text{and}$$
$$\lambda_j = \gamma\,\zeta^{1-j} \quad (j \geqslant 1) \tag{27}$$

for some parameters $p \in (0,1)$, $\gamma > 0$, and $\zeta > 0$ (and any Tokunaga sequence T_k).

Corollary 2 (Horton's law for the stream lengths). Consider a self-similar RAM tree with parameters given by Eq. (27). Then the average length L_j of a stream of order j satisfies

$$L_j R_L^{-j} = \frac{1}{\zeta\gamma} < \infty \quad \text{with } R_L = \zeta. \tag{28}$$

The proof is given in Appendix F. We notice that the Horton's law holds here in an exact form, without a limit in order j.

We show below that the two well-known empirical properties of self-similar river basins—fractal dimension and Horton's law for the longest stream length—formally hold in a self-similar RAM tree.

Theorem 4 (Fractal dimension of a self-similar tree). Consider a self-similar RAM tree with Tokunaga sequence T_k and other parameters given by Eq. (27). Let w_0 be the only real root of the function $\hat{t}(z)$ [Eq. (16)] in the interval $(0, 1/2]$. Then, the fractal dimension of the tree in the limit of infinite order and after a suitable length rescaling is given by

$$\mathbf{d} = \max\{1, \mathbf{d}_0\}, \quad \mathbf{d}_0 = -\frac{\log w_0}{\log \zeta} = \frac{\log R_B}{\log R_L}. \tag{29}$$

Proof is given in Appendix H.

Equation (29) coincides with the expression first obtained by La Barbera and Rosso [46] using a heuristic assumption of a basin with an ideal Horton's law of stream numbers. Figure 4(a) shows a map of $\mathbf{d}$ as a function of the Horton exponents R_B and R_L.

Theorem 5 (Horton's law for the length of the longest stream). Consider a self-similar RAM basin with parameters given by Eq. (27). Let Λ_k denote the average length of the longest stream in a basin of order k. Then

$$\lim_{k\to\infty} \Lambda_k R_\Lambda^{-k} = \text{const} < \infty \quad \text{with} \quad R_\Lambda = \zeta. \tag{30}$$

Proof is given in Appendix I.

E. Critical Tokunaga tree and emergent scaling relations

Observations on river networks have supported a basic constraint that the link length distribution is independent of the position of the link within a basin [2,47]. This motivates one to describe a family of trees that respect this property. Surprisingly, this leads to a one-parameter family of critical Tokunaga trees that satisfy multiple additional symmetries and include the celebrated Shreve model [4,5] as a special case.

The length of a link of order K in the RAM model is an exponential random variable with rate $\lambda_K S_K$, which is a direct consequence of using Poisson attachment along exponentially distributed streams. The order-independent link length implies $\lambda_K S_K = \text{const}$. Using the general form of λ_K in a self-similar RAM tree (27), one can select γ such that $S_K = \zeta^{K-1}$. This corresponds to the unique form of the Tokunaga sequence in a self-similar RAM tree with identically distributed link lengths:

$$T_k = (c-1)\,c^{k-1} \quad (k \geqslant 1), \quad c = \zeta \geqslant 1. \tag{31}$$

Definition 3 (Critical Tokunaga tree). A self-similar RAM with $p = 1/2$ and the Tokunaga sequence of Eq. (31) is called the *critical Tokunaga tree* or *critical Tokunaga model*. By Theorem 3, for this model we have $\pi_K = 2^{-K}$ and $\lambda_j = \gamma\,\zeta^{1-j}$.

Figure 3 shows three examples of critical Tokunaga trees of order $K = 5$ with parameter $c = 2.3$, which gives a close approximation to the observed river networks (see Table I). Equation (31) is a special case of the two-parameter sequence

TABLE I. Scaling exponents in the critical Tokunaga model. Selected scaling exponents (first column) in the critical Tokunaga model are expressed via the model parameter $c \geqslant 1$ (second column), fractal dimension $\mathbf{d}$ (third column), and Hack's exponent $\mathbf{h}$ (fourth column). Columns 5–7 show the values of the exponents in the critical Tokunaga model for $c = 2.0, 2.3, 2.5$. For comparison, column 8 shows the values estimated in the OCN model. Column 9 summarizes estimations in the observed river networks. The agreement of the exponents of the critical Tokunaga model with $c = 2.3$ (column 6) with those observed from real basins is noted.

	Expressed via			Critical Tokunaga model				
Exponent	c	$\mathbf{d}$	$\mathbf{h}$	$c = 2.0$[a]	$c = 2.3$	$c = 2.5$	OCN[b]	Real basins[c]
$R_B = R_M = R_A$	$2c$	$2^{\mathbf{d}/(\mathbf{d}-1)}$	$2^{1/(1-\mathbf{h})}$	4	4.6	5.0	4	4.1–4.8
$R_S = R_L$	c	$2^{1/(\mathbf{d}-1)}$	$2^{\mathbf{h}/(1-\mathbf{h})}$	2	2.3	2.5	2	2.1–2.7
$\mathbf{d} = \frac{\log R_B}{\log R_L}$	$\log_c(2c)$	$\mathbf{d}$	$\mathbf{h}^{-1}$	2	1.832	1.756	2	1.7–2.0
$\mathbf{h} = \frac{\log R_L}{\log R_B}$	$\log_{2c} c$	$\mathbf{d}^{-1}$	$\mathbf{h}$	0.5	0.546	0.569	0.57	0.5–0.6
β_A	$\log_{2c} 2$	$1 - \mathbf{d}^{-1}$	$1 - \mathbf{h}$	0.5	0.454	0.431	0.43	0.4–0.5
β_Λ	$\log_c 2$	$\mathbf{d} - 1$	$\mathbf{h}^{-1} - 1$	1	0.832	0.756	0.8	0.65–0.9

[a] Equivalent to the critical binary Galton-Watson branching process with independent and identically distributed exponential edge lengths.

[b] Average values estimated in simulated OCN basins, according to [2,55].

[c] According to [2,7,8,14,38,41,42,47,53].

$T_k = ac^{k-1}$ introduced by Tokunaga [48] to approximate river basin branching; this sequence has been examined in detail in [8,19,20,38,44]. The one-parameter sequence of Eq. (31) appears in the random self-similar network model of Veitzer and Gupta [49], which uses a purely combinatorial algorithm of recursive local replacement of the network generators to construct random trees. A theoretical underpinning of this constraint is revealed via the prism of branching process analysis. Kovchegov and Zaliapin [50] have shown that a random tree T generated by the critical Tokunaga model is critical and time invariant in both combinatorial and metric forms [32]. In particular, the condition $p = 1/2$ is necessary and sufficient for criticality. Moreover, the geometric branching process (that generates the combinatorial part of a RAM tree) is time invariant if and only if it corresponds to the critical Tokunaga model. Recall that *criticality* means that a branching process has unit average population size after an arbitrary but fixed time advancement (in both discrete and continuous versions). *Time invariance* means that the frequency of orders of subtrees that survive after a given time advancement coincides with the initial order distribution.

It is natural to assume that the local contributing area of a link (area that contributes to the link directly, and not via its descendant joint) is a function of the link length. This allows us to examine the average total contributing area A_i of a stream of order i. In particular, the order-independent link lengths imply order-independent local areas. The following result establishes Hack's law in a critical Tokunaga tree.

Theorem 6 (Hack's law in a critical Tokunaga tree). Consider a critical Tokunaga tree (Definition 3). Then the average lengths Λ_i of the longest stream and the average total contributing area A_i of a basin are related as

$$\Lambda_i \sim \text{const} \times (A_i)^{\mathbf{h}}, \quad \text{where}$$
$$\mathbf{h} = \mathbf{d}^{-1} = -\frac{\log \zeta}{\log w_0} = \frac{\log R_L}{\log R_B}. \tag{32}$$

Proof is given in Appendix F.

The Hack's law of Eq. (32) also holds in more general self-similar RAM trees (which may not be critical Tokunaga) as is shown in Appendix H. Figure 4(b) shows a map of $\mathbf{h}$ as a function of the Horton exponents R_B and R_L. The critical Tokunaga case corresponds to $R_B = 2R_L = 2c$ and hence $\mathbf{d} = \log_c(2c)$ and $\mathbf{h} = \log_{2c} c$; this case is depicted by a green line in Fig. 4.

Combining our results, we obtain the following summary of the Horton exponents in a critical Tokunaga tree.

Corollary 3 (Horton exponents in a critical Tokunaga tree). The Horton exponents in the critical Tokunaga model are given by

$$2c = R_B = R_M = R_A > R_S = R_L = R_\Lambda = c. \tag{33}$$

Proof is given in Appendix F.

Corollary 3 reveals that the essential Horton exponents in critical Tokunaga trees only assume two distinct values (c and $2c$). The inequality $R_S < R_B$, which is a part of Eq. (33), has been conjectured by Peckham [8] for trees with a well-defined Tokunaga sequence.

Notably, the critical Galton-Watson process with exponential edge lengths [6], which is equivalent to Shreve's random topology model after conditioning on the basin magnitude, is a special case of the critical Tokunaga model with $c = 2$ [32, Theorem 15]. In other words, the critical Tokunaga model offers a natural extension of the critical binary Galton-Watson process to a similarly versatile family of processes with a wide range of Horton exponents, fractal dimensions, Hack's exponents, and other parameters. As such, this model may be useful for multiple fields beyond hydrology.

Results of Chunikhina [51,52] imply that the critical Tokunaga model with $c = 2$ maximizes the entropy rate among the trees that satisfy the Horton's law of stream numbers, and that the critical Tokunaga model with a fixed c maximizes the entropy rate among the trees that satisfy the Horton's law for stream numbers with $R_B = 2c$.

Some additional scaling properties of the critical Tokunaga trees are collected in Appendix G.

F. Critical Tokunaga model closely fits observations

The critical Tokunaga model provides a very close fit to the data and scaling relations reported in river studies over the past decades. Table I summarizes the values of the key scaling exponents in the critical Tokunaga model and compares them with exponents in river network observations and the well-established OCN model [12]. The table uses the results of Corollary 3, Eq. (33) (Horton exponents), Theorem 4, Eq. (29) (fractal dimension **d**), Theorem 6, Eq. (32) (Hack's exponent **h**), and Eq. (6) (scaling exponents β_Λ and β_A).

The critical Tokunaga model's fit to the observed data is further illustrated in the Beaver Creek basin of Fig. 1. Figure 2 shows seven Horton's laws fit by the critical Tokunaga model with $c = 2.3$. Specifically, we consider the following stream attributes averaged over streams of order $K = 1, \dots, 6$: the stream number N_K (panel A), the average magnitude M_K and the average number S_K of links in a stream (panel B), the average total contributing area A_K and the average total upstream channel length $L_K^{\rm tot}$ (panel C), and the average stream length L_K and the average length Λ_K of the longest channel to the divide (panel D). The fitting lines correspond to the critical Tokunaga model predictions, which impressively agree with observations of all examined stream attributes (see figure caption).

III. DISCUSSION

A solid body of observational, modeling, and theoretical studies connects Horton's laws and power-law distributions of and power-law relations among river stream attributes to the self-similar structure of stream networks [2,5,8,17–19,21,36,37,39,49,53–56]. We suggest a rigorous treatment of the appearance and parametrization of Horton's laws in river networks using a recently formulated theory of random self-similar trees [32]. The proposed framework unifies the existing results and contributes to explaining the ubiquity of Horton's laws in dendritic systems of arbitrary origin.

The main technical contribution of our paper is a rigorous treatment of the appearance of Horton's laws in self-similar trees [Eqs. (18), (19), (22), (28), and (30)]. We show that the two fundamental properties—coordination and Horton prune invariance—necessarily lead to the Horton's laws for stream numbers and magnitudes (Theorem 2). Additional mild assumptions, like those in the RAM, yield the Horton's laws for multiple other attributes (Theorem 5; Corollaries 2 and 3), which in turn imply basin fractal dimensions (Theorem 4), Hack's law (Theorem 6), and other power-law scaling relations [Eqs. (4) and (6)]. Our results can be easily extended to other stream attributes such as stream slope, width, depth, and velocity, which are known to be proportional to a power of the upstream magnitude [2,17]. The developed framework may also facilitate analysis of the width function [57] or scaling of hydrologic fluxes [18,58] in self-similar basins. Such analyses can be done either analytically, or using ensemble simulation that is facilitated by a fast simulation algorithm for RAM trees.

The self-similarity is defined here (Definition 1) as invariance of a coordinated tree distribution with respect to the operation of Horton pruning, which is in accord with the empirical and modeling evidence of the past decades [5,8,20,34,44,49]. This approach unifies three seemingly distinct definitions of self-similarity that existed in the literature [5,8,17,20]. Importantly, each self-similar tree distribution corresponds to a unique Tokunaga sequence T_k that quantifies merging of branches of distinct orders (Theorem 1, Corollary 1). This provides a fundamental connection between an abstract measure-theoretic prune-invariance property and the tangible Tokunaga coefficients that can be statistically estimated in a single tree.

The family of self-similar distributions (Definition 1) rigorously reproduces the key geomorphic scalings discovered and reconfirmed during the past 80 years for river basins and summarized by [2,14,18,19,53], with a close fit to the examined exponents (Table I, Fig. 2). Interestingly, this fit is achieved within a one-parameter family of critical Tokunaga trees [Definition 3, Eq. (31)]. Although trees that satisfy Eq. (31) (and commonly referred to as Tokunaga trees) have been known for a long time [8,48,49], only very recently a rigorous understanding has been gained of the theoretical importance of this constraint within the general framework of branching processes [32,44,50]. In addition, neither the order distribution nor link lengths distribution [specified by π_K and λ_j of Eq. (27)] has been examined in Tokunaga trees. The critical Tokunaga model provides a natural parametric extension of the critical binary Galton-Watson branching process (and includes it as a special case with $c = 2$), which proved to be an indispensable model in many areas and remains at the forefront of theoretical and applied research nearly 150 years after its discovery [59,60]. This hints at deep and not fully understood symmetries in the structure of river networks. The theory of random self-similar trees explains the mathematical origin of these symmetries and offers tools for future exploration.

The presented results might advance applied statistical analysis of river stream attributes, via mapping all quantities of interest to a single master parameter c of Eq. (31). Statistical estimation of this parameter can be designed more effectively than that for a range of distinct yet possibly related quantities (e.g., multiple Horton exponents). This in turn facilitates global mapping of river network features and studying possible effects of hydroclimatic variables on landscape dissection. Corollary 3 shows that multiple Horton's laws examined in this paper hold with only two distinct Horton exponents: $R_B = R_M = R_A = 2c$ and $R_L = R_S = R_\Lambda = c$. This substantial reduction of observed quantities is well supported by data (Table I, Fig. 2) and might inform a range of modeling and theoretical efforts.

The critical Tokunaga model presents an ultimately symmetric class of trees characterized by coordination, Horton prune invariance, criticality, time invariance, identically distributed link lengths, and identically distributed local contributing areas. Despite these multiple constraints, this class is surprisingly rich, extending from perfect binary trees ($c = 1$) to the famous Shreve's random topology model ($c = 2$) to the structures reminiscent of the observed river networks ($c \approx 2.3$) and beyond. While offering a convenient theoretical and modeling paradigm, the critical Tokunaga model is merely a subclass of a much broader family of self-similar trees that might better accommodate various problem-specific data features. For instance, Fig. 4 suggests that the observed stream networks tend to cluster around the critical Tokunaga line $R_B = 2R_L$ in the (R_B, R_L) space. An applied study can use the self-similar theory to either focus on the symmetries of the critical Tokunaga family, or explore deviations from this stiff parametrization, both of which may have physical underpinnings.

Multiple properties of the critical Tokunaga family are well justified by the empirical evidence. We already mentioned that the coordination means that the basin structure is determined by its Horton-Strahler order, and the Horton prune invariance implies that the fundamental scaling laws remain the same after changing data resolution. Criticality ensures that the stream networks uniformly fill the space, instead of exploding (supercritical case) or rapidly fading off (subcritical case). The time invariance (invariance of basin order frequencies at different distances to the outlet) might reflect a physical process of formation of a stream network from sources downstream, so that a link only "knows" the information about the upstream part of the basin, yet remains unaware of how far it is from the outlet. Deviations from this invariance might point to anthropogenic changes in a basin by which various downstream alterations (dam construction, sediment aggradation, etc.) impose upstream changes that deviate from the natural organization of a left-alone erosional landscape. In the same vein, it would be interesting to find a hydrogeomorphological explanation for the joint distribution of the merging sub-basins (25).

The understanding of the hierarchical organization and scaling in convergent (tributary) river networks gained here can be extended to other geomorphological processes. Important examples include dynamic reorganization of landscapes and stream networks [61–63] and scaling of peak flows [64]. Our results can also be extended to study the divergent (distributary) networks of river deltas that are commonly represented by a directed acyclic graph [65]—a next step in complexity after trees examined in this paper. Quantifying the structure, self-similarity, and scaling of such graphs contributes to a still-missing unifying theory explaining how deltaic river networks self-organize to distribute water and sediment fluxes to the shoreline [66].

The self-similar family extends beyond the hydrological constraints, allowing one to study self-similar trees with edge lengths that depend on the position within the hierarchy, arbitrary fractal dimension $\mathbf{d} > 1$, and arbitrary Horton exponents $R_B > 2$ and $R_L > 1$. For instance, the RAM might be a suitable model for dendritic structures generated by diffusion limited aggregation (DLA). We recall that the geometric form of the Tokunaga coefficients $T_k \propto c^k$ with $c \approx 2.72 \pm 0.22$ has been known in DLA for a long time [20,67]. It is noteworthy that the independently estimated fractal dimension of DLA clusters, $\mathbf{d} = 1.7 \pm 0.05$ [68], coincides with the fractal dimension of a critical Tokunaga tree with $c = 2.72 \pm 0.22$ according to our Eq. (29): $\mathbf{d} = \log_c(2c) = 1.7 \pm 0.05$.

In this paper we examined a very particular class of graphs justified by hydrogeomorphological applications—reduced binary trees. Each internal vertex in such a tree has a fixed degree 3, being connected to a single downstream link and exactly two upstream links, while the root and leaves have degree 1. The theory of random self-similar trees, however, readily applies to trees with multiple branching and extends to general loopy graphs (networks).

For instance, all self-similar Galton-Watson (nonbinary) trees have been recently described in [69]. The self-similar family includes the critical binary Galton-Watson tree (using the terminology of this paper—the critical Tokunaga tree with $c = 2$), which is the only member with a finite branching. All other family members have a power-law degree distribution with tail $\propto k^{-\alpha}$ and index α spanning the interval (2,3). Natural applications for nonbinary self-similar trees include phylogenetics [70] and statistical seismology, where trees represent temporal evolution of earthquake clusters [24–27].

More generally, many observed and simulated networks are known to preserve their key statistical properties under coarse graining—a renormalization group transformation that merges selected vertices of a network $\mathcal{G}$ according to their proximity [71]. Such a transformation is uniquely represented by a tree $T(\mathcal{G})$ the edges of which correspond to mergers among the network's nodes. The leaves of $T(\mathcal{G})$ correspond one to one to the nodes of $\mathcal{G}$, internal vertices correspond to the coarse-grained supernodes, and the root corresponds to the single ancestral supernode that represents the entire network. Informally, coarse graining of a graph $\mathcal{G}$ corresponds to cutting selected peripheral subtrees of its merger tree $T(\mathcal{G})$. Different rules of coarse graining correspond to different types of tree pruning, for example those discussed in [72,73]. Horton pruning is an important member of this larger class of tree reductions. Statistical invariance of a network $\mathcal{G}$ under coarse graining suggests that the corresponding tree $T(\mathcal{G})$ is invariant with respect to the associated pruning. This assertion is supported by numerical analysis of merger trees that represents two-dimensional site percolation [74].

ACKNOWLEDGMENTS

We would like to thank two anonymous reviewers for their constructive feedback. The work of Y.K. is supported by NSF Grant No. DMS-1412557. The work of I.Z. is supported by NSF Grant No. EAR-1723033. The work of E.F.-G. is supported by NSF Grants No. DMS-1839336 and No. EAR-1811909.

APPENDIX A: HORTON-STRAHLER ORDERS AND HORTON PRUNING

The importance of links and junctions in the basin hierarchy is measured by the Horton-Strahler order $K \geqslant 1$ [1,33]. Each link and its upstream junction have the same order. The order assignment is done in a hierarchical fashion, from the sources downstream. Each source is assigned order 1. When two links of the same order K merge at a junction, the junction is assigned order $K + 1$. When two links with different orders $K_1 > K_2$ merge at a junction, the largest order prevails and the junction is assigned order K_1. The connected sequence of links and their upstream junctions of the same order K is called a *stream* (*branch*) of order K. We denote by $N_K = N_K[T]$ the number of streams of order K in a finite tree T. The Horton-Strahler order $\text{ord}(T)$ of a tree T is the maximal order of its links (junctions, streams). The Horton-Strahler ordering is illustrated in Fig. 1.

The Horton-Strahler orders are closely related to the operation of Horton pruning $\mathcal{R}$ that removes the source links from a basin. This relation has been first recognized by Melton [34] and proved valuable in rigorous statistical analyses of tree self-similarity [5,8,21,32,44]. Formally, we consider the map $\mathcal{R} : \mathcal{L} \to \mathcal{L}$ that removes the source links from a tree T. This may create nonbranching chains of links connected by degree 2 junctions—every such chain is merged into a single link. The Horton pruning $\mathcal{R}$ reduces the order of each surviving stream, and hence the basin order, by one. Accordingly, the order of a tree is the minimal number of Horton prunings that completely eliminates it, as in Eq. (1). We emphasize that the pruning cannot cut a stream in the middle—it can only eliminate the entire stream after a finite number of iterations [34]. Figure 1 illustrates the Horton pruning for the stream network of Beaver Creek, KY—the order of this basin is $K = 6$ because it is eliminated in six Horton prunings.

APPENDIX B: POWER-LAW DISTRIBUTION OF LINK ATTRIBUTES

Consider the value $Z_{(i)}$ of an attribute Z calculated at link i in a large basin. The average number of links of order K is given by $N_K S_K$, where S_K denotes the average number of links within a stream of order K. One can heuristically approximate the frequencies of $\{Z_{(i)}\}$ by using the same average value Z_K for any link of order K. Then, assuming Horton's laws for N_K and S_K with exponents R_B and R_S, respectively, and considering the limit of an infinitely large basin we find

$$\left|i : Z_{(i)} \geqslant R_Z^K\right| \approx \sum_{j=K}^{\infty} N_j S_j \propto \sum_{j=K}^{\infty} \left(\frac{R_S}{R_B}\right)^j \propto \left(\frac{R_S}{R_B}\right)^K. \tag{B1}$$

This is a punctured (by discrete order) version of a general power-law relation of Eq. (6).

APPENDIX C: ASYMPTOTIC BEHAVIOR OF A SEQUENCE—THE GENERATING FUNCTION APPROACH

This section summarizes the basic facts about generating functions that are the main tool in establishing asymptotic behavior of stream attributes in a self-similar basin.

The generating function $f(z)$ of a sequence $a_k \geqslant 0$, $k = 0, 1, \dots,$ of non-negative real numbers is defined as a formal power series:

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \quad z \in \mathbb{C}. \tag{C1}$$

It is known [75] that there exists such a real number $r \geqslant 0$ that the series in the right hand side of (C1) converges to the function $f(z)$ for any $|z| < r$ and diverges for any $|z| > r$. The number r is called the *radius of convergence* of the sequence a_k; it provides notable constraints on the asymptotic behavior of a_k. The smaller the radius of convergence, the faster the growth of the sequence coefficients. Informally, $0 < r < 1$ implies that the coefficients a_k increase geometrically, $r > 1$ implies that the coefficients decrease geometrically, and $r = 1$ implies that the coefficients vary at a rate slower than geometric (e.g., polynomially). The values $r = 0$ and ∞ imply a faster than geometric growth or decay, respectively.

The Cauchy-Hadamard theorem [75] expresses the radius of convergence in terms of the series coefficients:

$$\frac{1}{r} = \limsup_{k\to\infty} a_k^{1/k}. \tag{C2}$$

Often, the radius of convergence for a_k can be easily found from the explicit form of $f(z)$. Specifically, if $r > 0$, then the function $f(z)$ is analytic within the disk $|z| < r$ and has at least one singularity on the circle $|z| = r$, that is, it has to diverge for at least one point on that circle [76, Theorem 2.4.2]. Thus, the radius of convergence equals to the modulus of a singularity closest to the origin. Furthermore, recalling that $a_k \geqslant 0$ we have

$$|f(z)| = \left|\sum_{k=0}^{\infty} a_k\, z^k\right| \leqslant \sum_{k=1}^{\infty} a_k\, |z|^k = f(|z|), \tag{C3}$$

where the equality is only achieved for $z = |z|$. This means that the singularity closest to the origin lies on the real axis (although there might be other singularities with the same modulus.) This makes the search for such a singularity much easier: one can only consider the restriction of the function $f(z)$ on the real axis. In other words, despite the use of complex analysis in establishing some of our results, the applied treatment of suitable generating functions can be done in the real domain. Furthermore, if the singularity of $f(z)$ nearest to the origin is a simple pole, then the coefficients a_k asymptotically form a geometric series, which we refer to as *Horton's law*.

Proposition 1 (Horton's Law for a simple pole sequence). Suppose $f(z) = \sum_{i=1}^{\infty} a_k z^k$ is analytic in the disk $|z| < \rho$ except for a single pole of multiplicity one at a positive real value $r < \rho$. Then the sequence a_k obeys Horton's law

$$\lim_{k\to\infty} a_k\, r^k = \alpha \tag{C4}$$

for some $0 < \alpha < \infty$. Furthermore, if we define $g(z) = f(z)(z - r)$, then $\alpha = -g(r)/r$.

Proof. We have, for any $\Delta \in (0, r)$ [75],

$$a_k = \frac{1}{2\pi i} \oint_{|z|=\Delta} \frac{f(z)dz}{z^{k+1}}. \tag{C5}$$

By the residue theorem [75], we obtain, for any $\gamma \in (r, \rho)$,

$$\frac{1}{2\pi i} \oint_{|z|=\gamma} \frac{f(z)dz}{z^{k+1}} = \text{res}\left(\frac{f(z)}{z^{k+1}}; 0\right) + \text{res}\left(\frac{f(z)}{z^{k+1}}; r\right) \tag{C6}$$

$$= a_k + \text{res}\left(\frac{f(z)}{z^{k+1}}; r\right). \tag{C7}$$

Therefore,

$$a_k = \frac{1}{2\pi i} \oint_{|z|=\gamma} \frac{f(z)dz}{z^{k+1}} - \text{res}\left(\frac{f(z)}{z^{k+1}}; r\right), \tag{C8}$$

where

$$\left|\oint_{|z|=\gamma} \frac{f(z)dz}{z^{k+1}}\right| \leqslant \frac{\max_{|z|=\gamma} |f(z)|}{\gamma^k} = o(r^{-k}). \tag{C9}$$

Consider $g(z) = (z - r)f(z)$. It is known that [75]

$$\text{res}(f(z); r) = g(r), \tag{C10}$$

and hence

$$\mathrm{res}\left(\frac{f(z)}{z^{k+1}};r\right) = \frac{g(r)}{r^{k+1}} = \frac{g(r)}{r}r^{-k}. \tag{C11}$$

Accordingly, we obtain

$$a_k = -\frac{g(r)}{r}r^{-k} + o(r^{-k}), \tag{C12}$$

which completes the proof. ■

Proposition 1 is used in Appendices I and J to establish Horton's laws for Λ_k and A_k.

APPENDIX D: PROOF OF THEOREM 2

The average magnitude M_k is the mean number of sources upstream of an order k stream. It can be represented as the sum of the magnitudes of two order $k-1$ streams that formed this stream, plus the magnitudes of all its side tributaries. Hence $M_1 = 1$, and

$$M_k = 2\,M_{k-1} + \sum_{i=1}^{k-1} M_i\, T_{k-i}, \quad \text{for } k > 1. \tag{D1}$$

The generating function for the average magnitudes M_k is obtained by multiplying both sides in (D1) by z^k and summing over $k \geqslant 1$:

$$M(z) = \sum_{k=1}^{\infty} M_k\, z^k = z + 2zM(z) + M(z)\,T(z).$$

Thus,

$$M(z) = \frac{z}{1-2z-T(z)} = -\frac{z}{\hat{t}(z)}, \tag{D2}$$

where, according to Eq. (16) of the main text,

$$\hat{t}(z) = -1 + 2z + T(z), \quad T(z) = \sum_{i=1}^{\infty} T_k z^k.$$

The function $M(z)$ is analytic with the exception of zeros and singularities of $\hat{t}(z)$. Observe that $t(0) = -1$, and since $T_k \geqslant 0$ we have $t(1/2) = T(1/2) \geqslant 0$. Furthermore, since

$$\frac{d}{dz}\hat{t}(z) = 2 + \sum_{k=1}^{\infty} kT_k\, z^{k-1} > 0 \quad \text{for all } z \in (0,\infty),$$

the equation $\hat{t}(z) = 0$ has a unique real root w_0 of multiplicity in the interval $(0, 1/2]$. Let r_T be the radius of convergence for $T(z)$, and hence for $\hat{t}(z)$. We notice that $r_T > w_0$, so the radius of convergence of $M(z)$ coincides with the root of $\hat{t}(z)$ closest to the origin. We claim that this root is w_0. Assuming otherwise, there exists $w \in \mathbb{C}$ such that $\hat{t}(w) = 0$ and $|w| < w_0$. Since w_0 is the unique real root of $\hat{t}(z)$ within $(0, 1/2]$, w must have a nonzero imaginary part. This means that the singularity of $M(z)$ closest to the origin is not on the real axis, which contradicts (C3). Hence the radius of convergence of $M(z)$ is w_0, and w_0 is a simple pole of $M(z)$. Proposition 1 now establishes the result. ■

APPENDIX E: EXACT HORTON'S LAW

Assume that the Horton's law for stream numbers $\mathcal{N}_1[K]$, and hence for magnitudes M_K, holds exactly (recall that $M_1 = 1$):

$$M_K = R_M^{K-1}. \tag{E1}$$

Then,

$$M(z) = \frac{z}{1-R_M z},$$

which leads to

$$\hat{t}(z) = -\frac{z}{M(z)} = -1 + R_M z \quad \text{and} \quad T(z) = (R_M - 2)z.$$

This implies that the only self-similar model with exact Horton's law corresponds to the Tokunaga sequence

$$T_1 = R_M - 2, \quad T_k = 0 \quad \text{for} \quad k > 1.$$

APPENDIX F: PROOFS OF COROLLARY 2, THEOREM 6, AND COROLLARY 3

Proof of Corollary 2. By definition of the RAM, the length of a stream of order j is an exponential random variable with rate λ_j. In a self-similar tree, the rate is given by Eq. (31): $\lambda_j = \gamma\,\zeta^{1-j}$. This implies $L_j = \lambda_j^{-1} = \gamma^{-1}\,\zeta^{j-1}$, which is equivalent to the statement of Theorem 6 [Eq. (32)]. ■

Proof of Theorem 6. Recall the Horton's law for the average magnitude [Theorem 2, Eq. (19)] that holds in any tree with a tamed Tokunaga sequence ($\limsup_{k\to\infty} T_k^{1/k} < \infty$) and the Horton's law for the average length of the longest stream [Theorem 5, Eq. (30)] that holds in any self-similar RAM tree. These laws apply to a critical Tokunaga tree of the current statement. Furthermore, the asymptotic equivalence between the average basin contributing area and average basin magnitude [Eq. (G3)] implies the Horton's law for the average basin areas with Horton exponent R_M. Finally, we use the general result of Eq. (5) to establish Hack's law [Eq. (32)] in a critical Tokunaga tree. ■

Proof of Corollary 3. Using the definition of $\hat{t}(z)$ [Eq. (16)] and the geometric form of the Tokunaga coefficients [Eq. (31)] we obtain $\hat{t}(z) = (1-2cz)(z-1)/(1-cz)$. The only real root of $\hat{t}(z)$ within $(0, 1/2]$ is $w_0 = (2c)^{-1}$. By Theorem 2, Eq. (20) we have $R_B = R_M = 2c$, and Eq. (26) implies $S_K = c^{K-1}$, which corresponds to $R_S = c$. The equality $R_S = R_L$ follows from independence of the distribution of link lengths of their position within a basin. Finally, $R_\Lambda = c$ is established in Theorem 5, Eq. (30). ■

APPENDIX G: SCALINGS IN A CRITICAL TOKUNAGA TREE

The critical Tokunaga model (Definition 3) satisfies Horton's law for the original stream counts N_i [[32], Corollary 5]:

$$\frac{N_i}{N_{i+1}} \xrightarrow{p} R_B \quad \text{as } i \to \infty, \tag{G1}$$

where $\xrightarrow{p}$ denotes convergence in probability [77]. This result strengthens the statement of Theorem 2, Eq. (22) that is formulated for the respective averages.

The weak law of large numbers holds for the tree order. Formally, denote by $T[K]$ a critical Tokunaga tree of order K and write $|T[K]|$ for the number of links in this tree. Then [32, Corollary 6]

$$\frac{\log_{R_B}|T[K]|}{K} \xrightarrow{p} 1 \quad \text{as } K \to \infty. \tag{G2}$$

Informally, this means that the tree order grows as a logarithm base R_B of the tree size.

The identically distributed link lengths imply identically distributed local areas, which in turn establishes the Horton's law for A_i. Specifically, in a critical Tokunaga tree we have (Appendix J)

$$A_i \sim \text{const} \times M_i, \tag{G3}$$

where $x_i \sim y_i$ stands for $\lim_{i\to\infty} x_i/y_i = 1$. The same approach shows that the asymptotic of Eq. (G3) holds also for the average total channel length L_i^{tot} upstream of a stream of order i, with different proportionality constant. The asymptotic of Eq. (G3) formalizes one of the key empirical observations [2] that connects physical (area A_i) and combinatorial (magnitude M_i) attributes of a river basin. This asymptotic may not hold in a general self-similar (not critical Tokunaga) tree.

APPENDIX H: FRACTAL DIMENSION OF A SELF-SIMILAR RAM TREE

Consider a self-similar RAM tree T (Theorem 3) with a Tokunaga sequence $\{T_k\}$ satisfying $\limsup_{k\to\infty} T_k^{1/k} < \infty$, and parameters $\gamma > 0$ and $\zeta > 1$. Below we construct a Markov tree process $\{\Upsilon_K\}_{K=1,2,\dots}$ corresponding to T following [32] and use it to find the fractal dimension of the resulting tree in the limit of infinite tree order. The construction below closely reproduces that of the RAM (see the main text), but scales the edge lengths so that an infinitely large tree has a proper fractal dimension.

Let Υ_1 be an I-shaped tree of Horton-Strahler order 1, with the edge length distributed as an exponential random variable with parameter γ. Conditioned on Υ_K, the tree Υ_{K+1} is constructed according to the following transition rules. We attach new leaf edges to Υ_K at the points sampled by an inhomogeneous Poisson point process with the intensity $\rho_{j,K} = \gamma\zeta^{K-j}T_j$ along the edges of order $j \leqslant K$ in Υ_K. We also attach a pair of new leaf edges to each of the leaves in Υ_K. The lengths of all the newly attached leaf edges are independent and identically distributed exponential random variables with parameter $\gamma\zeta^K$ that are independent of the combinatorial shape and the edge lengths in Υ_K. Finally, we let the tree Υ_{K+1} consist of Υ_K and all the attached leaves and leaf edges.

By construction, a branch of order j in Υ_K becomes a branch of order $j+1$ in Υ_{K+1} after the attachment of new leave edges. The length of order j branch in Υ_K (and therefore, the length of order $j+1$ branch in Υ_{K+1}) is an exponential random variable with parameter $\gamma\zeta^{K-j}$. Therefore, in a tree Υ_{K+1}, the number $n_{1,j+1}(K+1)$ of side branches of order 1 in a branch of order $j+1$ has geometric distribution:

$$\begin{aligned} P\big(n_{1,j+1}(K+1) = r\big) &= \frac{\gamma\zeta^{K-j}}{\gamma\zeta^{K-j} + \rho_{j,K}}\left(\frac{\rho_{j,K}}{\gamma\zeta^{K-j} + \rho_{j,K}}\right)^r \\ &= \frac{1}{1+T_j}\left(\frac{T_j}{1+T_j}\right)^r \end{aligned} \tag{H1}$$

for $r = 0, 1, 2, \dots$, with the mean value

$$E[n_{1,j+1}(K+1)] = \frac{\rho_{j,K}}{\gamma\zeta^{K-j}} = T_j.$$

After $i \geqslant 1$ rounds of attachments the mean number $n_{i,j+i}(M)$ of side branches of order i in a branch of order $j+i$ in a tree Υ_M (where $M = K + i$ and $K \geqslant j$) is

$$E[n_{i,j+i}(M)] = T_j.$$

Each tree Υ_K is distributed as a self-similar RAM tree [32] with Tokunaga sequence $\{T_k\}$ and parameters (γ, ζ), conditioned on its Horton-Strahler order being equal to K, and with its edge lengths scaled by ζ^{1-K}.

Observe that by construction $\Upsilon_K \subset \Upsilon_{K+1}$. Accordingly, there exists the limit space

$$\Upsilon_\infty = \lim_{K\to\infty} \Upsilon_K = \bigcup_{K=1}^{\infty} \Upsilon_K.$$

The self-similarity of the RAM process suggests that the limit space does not change its statistical properties after rescaling, which corresponds here to the Horton pruning. Let $\mathbf{d}$ denote its fractal dimension. That the limit space includes at least the root branch Υ_1 implies $\mathbf{d} \geqslant 1$. Assume that $\mathbf{d} > 1$. Then, denoting the mean $\mathbf{d}$-dimensional volume of Υ_∞ by vol, we have

$$\text{vol} = \sum_{k=1}^{\infty} t_k \frac{\text{vol}}{\zeta^{\mathbf{d}k}}. \tag{H2}$$

This equation is obtained by splitting a tree Υ_∞ into the subtrees attached to its highest-order branch Υ_1. There is an average of $t_1 = T_1 + 2$ subtrees distributed as Υ_∞ scaled by ζ^{-1}. In general, for each k, there will be an average of t_k subtrees distributed as Υ_∞ scaled by ζ^{-k}. Scaling the lengths by ζ^{-k} in the $\mathbf{d}$-dimensional space results in scaling the volume by $\zeta^{-\mathbf{d}k}$. The vol term in (H2) can be canceled out, yielding

$$\hat{t}(\zeta^{-\mathbf{d}}) = 0, \tag{H3}$$

and hence $\zeta^{-\mathbf{d}} = w_0 = R_B^{-1}$. This leads to (29).

APPENDIX I: HORTON'S LAW FOR Λ_k

If T is the tree representing a stream network, then the length of the longest stream is the height of the tree T, denoted by T_H [6,32].

Consider a tree T generated by a self-similar RAM with a Tokunaga sequence $\{T_k\}$ satisfying $\limsup_{k\to\infty} T_k^{1/k} < \infty$, and parameters $\gamma > 0$ and $\zeta > 1$. Let

$$\Lambda_k = E[T_H \,|\, \text{ord}(T) = k] \tag{I1}$$

that represents the mean length of the longest river stream in a basin with the Horton-Strahler order k. Notice that, since $[\mathcal{R}(T)]_H \leqslant T_H$,

$$\begin{aligned} \zeta\,\Lambda_{k-1} &= E\{[\mathcal{R}(T)]_H \,|\, \text{ord}(T) = k\} \\ &\leqslant E[T_H \,|\, \text{ord}(T) = k] = \Lambda_k. \end{aligned} \tag{I2}$$

Hence, since $\Lambda_1 = \gamma^{-1}$, we have $\Lambda_k \geqslant \gamma^{-1}\zeta^{k-1}$. Next, let

$$Y_1, Y_2, \dots, Y_{N_1[T]}$$

denote the leaf lengths in the tree T. Then, since

$$T_H \leqslant [\mathcal{R}(T)]_H + \max_{j=1,\dots,N_1[T]} Y_j,$$

we have

$$\begin{aligned}
\Lambda_k &\leqslant E\{[\mathcal{R}(T)]_H \,|\, \mathrm{ord}(T)=k\} \\
&\quad + E\Big[\max_{j=1,\dots,N_1[T]} Y_j \,\Big|\, \mathrm{ord}(T)=k\Big] \\
&= \zeta\,\Lambda_{k-1} + \gamma^{-1} E\left[\sum_{j=1}^{N_1[T]} \frac{1}{j} \,\middle|\, \mathrm{ord}(T)=k\right] \\
&\leqslant \zeta\,\Lambda_{k-1} + \gamma^{-1} E\{1+\log(N_1[T]) \,|\, \mathrm{ord}(T)=k\} \\
&\leqslant \zeta\,\Lambda_{k-1} + \gamma^{-1} + \gamma^{-1}\log\{E[N_1[T]\,|\mathrm{ord}(T)=k]\}
\end{aligned} \tag{I3}$$

by Wald's equation, the coupon collector problem, and finally Jensen's inequality. Recall (Theorem 2) the Horton's law for the leaf count in a self-similar process:

$$\mathcal{N}_1[k] = M_k = M\,R_B^k + o\big(R_B^k\big).$$

Hence, Eqs. (I2) and (I3) imply

$$0 \leqslant \Lambda_k - \zeta\,\Lambda_{k-1} \leqslant \gamma^{-1} k \log R_B + \beta$$

for some constant β, and

$$0 \leqslant \frac{\Lambda_k}{\Lambda_{k-1}} - \zeta \leqslant \gamma^{-1}\frac{k\log R_B + \beta}{\Lambda_{k-1}} \leqslant \frac{k\log R_B+\beta}{\zeta^{k-2}} \to 0 \tag{I4}$$

as $k\to\infty$. Accordingly,

$$\begin{aligned}
\log\Lambda_k &= \sum_{j=2}^{k} \log\left(\frac{\Lambda_k}{\Lambda_{k-1}}\right) + \log\Lambda_1 \\
&= (k-1)\log\zeta + \sum_{j=2}^{k}\log(1+\mathcal{E}_j) - \log\gamma,
\end{aligned} \tag{I5}$$

where $0 \leqslant \mathcal{E}_j \leqslant (k\log R_B + \beta)\zeta^{1-k}$, and therefore $\sum_{j=2}^{\infty}\log(1+\mathcal{E}_j)$ converges to a constant. We therefore conclude that the strong Horton's law holds for Λ_k with Horton exponent $R_\Lambda = R_L = \zeta$:

$$\Lambda_k \sim \mathrm{const}\times\zeta^k. \tag{I6}$$

APPENDIX J: HORTON LAW FOR A_k

Assume that the mean local contributing area of a link of order k equals α_k. Then the total mean contributing area A_K of a tree of order $K \geqslant 1$ is

$$A_K = \sum_{i=1}^{K} \alpha_i S_i \mathcal{N}_i[K], \tag{J1}$$

where $S_k\mathcal{N}_k[K]$ is the mean number of links of order k in a tree of order K. A convenient recursive expression is obtained by noticing that $A_1 = \alpha_1$ and

$$A_K = 2A_{K-1} + \alpha_K S_K + \sum_{i=1}^{K-1} A_i\,T_{K-i} \quad \text{for } K \geqslant 2. \tag{J2}$$

The generating function for A_k is given by

$$A(z) = \sum_{k=1}^{\infty} A_k z^k = 2zA(z) + \sum_{k=1}^{\infty}\alpha_k S_k z^k + A(z)T(z),$$

which yields

$$A(z) = \frac{\sum_{k=1}^{\infty}\alpha_k S_k z^k}{1-2z-T(z)} = -\frac{D(z)}{\hat{t}(z)} = M(z)\frac{D(z)}{z}. \tag{J3}$$

Here $D(z)$ is the generating function for the mean local contributing areas $\alpha_k\,S_k$ of streams of order k. Suppose that the radius of convergence of $D(z)$ is larger than w_0. Then, by Proposition 1,

$$A_k \sim \mathrm{const}\times w_0^{-k} \sim \mathrm{const}\times M_k. \tag{J4}$$

Consider the critical Tokunaga model. Here $\alpha_k = \alpha$, $S_k = c^{k-1}$, and hence

$$D(z) = \frac{\alpha\, z}{1-cz}$$

with the radius of convergence c^{-1} that coincides with that of $\hat{t}(z)$. Observe that the radius of convergence of $\hat{t}(z)$ must be greater than its zero, hence $w_0 < c^{-1}$, and so the asymptotic of (J4) holds.

---

[1] R. E. Horton, Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology, Bull. Geophys. Soc. Am. **56**, 275 (1945).

[2] I. Rodriguez-Iturbe and A. Rinaldo, *Fractal River Basins: Chance and Self-Organization* (Cambridge University, Cambridge, England, 2001).

[3] J. W. Kirchner, Statistical inevitability of Horton's laws and the apparent randomness of stream channel networks, Geology **21**, 591 (1993).

[4] R. L. Shreve, Statistical law of stream numbers, J. Geol. **74**, 17 (1966).

[5] G. Burd, E. C. Waymire, and R. D. Winn, A self-similar invariance of critical binary Galton-Watson trees, Bernoulli **6**, 1 (2000).

[6] J. Pitman, *Combinatorial Stochastic Processes: Ecole d'Été de Probabilités de Saint-Flour XXXII-2002*, Lecture Notes in Mathematics Vol. 1875 (Springer, New York, 2006).

[7] H. De Vries, T. Becker, and B. Eckhardt, Power law distribution of discharge in ideal networks, Water Resour. Res. **30**, 3541 (1994).

[8] S. D. Peckham, New results for self-similar trees with applications to river networks, Water Resour. Res. **31**, 1023 (1995).

[9] A. E. Scheidegger, A stochastic model for drainage patterns into an intramontane treinch, Hydrological Sci. J. **12**, 15 (1967).

[10] H. Takayasu, I. Nishikawa, and H. Tasaki, Power-law mass distribution of aggregation systems with injection, Phys. Rev. A **37**, 3110 (1988).

[11] A. Rinaldo, I. Rodriguez-Iturbe, R. Rigon, R. L. Bras, E. Ijjasz-Vasquez, and A. Marani, Minimum energy and fractal structures of drainage networks, Water Resour. Res. **28**, 2183 (1992).

[12] R. Rigon, A. Rinaldo, I. Rodriguez-Iturbe, R. L. Bras, and E. Ijjasz-Vasquez, Optimal channel networks: A framework for the study of river basin morphology, Water Resour. Res. **29**, 1635 (1993).

[13] A. Rinaldo, I. Rodriguez-Iturbe, R. Rigon, E. Ijjasz-Vasquez, and R. L. Bras, Self-organized Fractal River Networks, Phys. Rev. Lett. **70**, 822 (1993).

[14] A. Maritan, A. Rinaldo, R. Rigon, A. Giacometti, and I. Rodriguez-Iturbe, Scaling laws for river networks, Phys. Rev. E **53**, 1510 (1996).

[15] A. Rinaldo, R. Rigon, J. R. Banavar, A. Maritan, and I. Rodriguez-Iturbe, Evolution and selection of river networks: Statics, dynamics, and complexity, Proc. Natl. Acad. Sci. USA **111**, 2417 (2014).

[16] P. Balister, J. Balogh, E. Bertuzzo, B. Bollobás, G. Caldarelli, A. Maritan, R. Mastrandrea, R. Morris, and A. Rinaldo, River landscapes and optimal channel networks, Proc. Natl. Acad. Sci. USA **115**, 6548 (2018).

[17] V. K. Gupta and E. C. Waymire, Statistical self-similarity in river networks parameterized by elevation, Water Resour. Res. **25**, 463 (1989).

[18] V. K. Gupta, B. M. Troutman, and D. R. Dawdy, Towards a nonlinear geophysical theory of floods in river networks: An overview of 20 years of progress, in *Nonlinear Dynamics in Geosciences* (Springer, New York, 2007), pp. 121–151.

[19] D. L. Turcotte, *Fractals and Chaos in Geology and Geophysics* (Cambridge University, Cambridge, England, 1997).

[20] W. I. Newman, D. L. Turcotte, and A. M. Gabrielov, Fractal trees with side-branching, Fractals **5**, 603 (1997).

[21] S. D. Peckham and V. K. Gupta, A reformulation of Horton's laws for large river networks in terms of statistical self-similarity, Water Resour. Res. **35**, 2763 (1999).

[22] P. Flajolet, J.-C. Raoult, and J. Vuillemin, The number of registers required for evaluating arithmetic expressions, Theor. Comput. Sci. **9**, 99 (1979).

[23] M. Drmota and H. Prodinger, The register function for t-ary trees, ACM Transactions on Algorithms **2**, 318 (2006).

[24] M. Baiesi and M. Paczuski, Scale-free networks of earthquakes and aftershocks, Phys. Rev. E **69**, 066106 (2004).

[25] J. R. Holliday, D. L. Turcotte, and J. B. Rundle, Self-similar branching of aftershock sequences, Physica A **387**, 933 (2008).

[26] M. R. Yoder, J. Van Aalsburg, D. L. Turcotte, S. G. Abaimov, and J. B. Rundle, Statistical variability and Tokunaga branching of aftershock sequences utilizing BASS model simulations, Pure Appl. Geophys. **170**, 155 (2013).

[27] I. Zaliapin, and Y. Ben-Zion, Earthquake clusters in southern California I: Identification and stability, J. Geophys. Res.: Solid Earth **118**, 2847 (2013).

[28] G. S. Kassab, The coronary vasculature and its reconstruction, Ann. Biomed. Eng. **28**, 903 (2000).

[29] F. Cassot, F. Lauwers, C. Fouard, S. Prohaska, and V. Lauwers-Cances, A novel three-dimensional computer-assisted method for a quantitative study of microvascular networks of the human cerebral cortex, Microcirculation **13**, 1 (2006).

[30] E. H. Campbell Grant, W. H. Lowe, and W. F. Fagan, Living in the branches: population dynamics and ecological processes in dendritic networks, Ecology Lett. **10**, 165 (2007).

[31] D. L. Turcotte, J. D. Pelletier, and W. I. Newman, Networks with side-branching in biology, J. Theor. Biol. **193**, 577 (1998).

[32] Y. Kovchegov and I. Zaliapin, Random Self-Similar Trees: A mathematical theory of Horton's laws, Probab. Surveys **17**, 1 (2020).

[33] A. N. Strahler, Quantitative analysis of watershed geomorphology, Trans. Am. Geophys. Un. **38**, 913 (1957).

[34] M. A. Melton, A derivation of Strahler's channel-ordering system, J. Geology **67**, 345 (1959).

[35] L. Leopold, M. G. Wolman, and J. Miller, *Fluvial Processes in Geomorphology* (Dover, New York, 1992).

[36] D. G. Tarboton, Fractal river networks, Horton's laws and Tokunaga cyclicity, J. Hydrology **187**, 105 (1996).

[37] V. K. Gupta and E. C. Waymire, Some mathematical aspects of rainfall, landforms and floods, in *Rainfall, Landforms and Floods*, edited by O. E. Barndorff-Nielsen, V. K. Gupta, V. Perez-Abreu, and E. C. Waymire (World Scientific, Singapore, 1998).

[38] S. Zanardo, I. Zaliapin, and E. Foufoula-Georgiou, Are American rivers Tokunaga self-similar? (in New results on fluvial network topology and its climatic dependence), J. Geophys. Res.: Earth Surface **118**, 166 (2013).

[39] O. J. Mesa, Cuatro modelos de redes de drenaje, Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales **42**, 379 (2018).

[40] J. T. Hack, Studies of longitudinal stream profiles in Virginia and Maryland, U.S. Government Printing Office 294 (1957).

[41] R. Rigon, I. Rodriguez-Iturbe, A. Maritan, A. Giacometti, D. Tarboton, and A. Rinaldo, On Hack's law, Water Resour. Res. **32**, 3367 (1996).

[42] D. G. Tarboton, R. L. Bras, and I. Rodriguez-Iturbe, The fractal nature of river networks, Water Resour. Res. **24**, 1317 (1988).

[43] I. Rodriguez-Iturbe, E. J. Ijjasz-Vasquez, R. L. Bras, and D. G. Tarboton, Power law distributions of discharge mass and energy in river basins, Water Resour. Res. **28**, 1089 (1992).

[44] Y. Kovchegov and I. Zaliapin, Horton law in self-similar trees, Fractals **24**, 1650017 (2016).

[45] http://rivix.com.

[46] P. La Barbera and R. Rosso, On the fractal dimension of stream networks, Water Resour. Res. **25**, 735 (1989).

[47] D. G. Tarboton, R. L. Bras, and I. Rodriguez-Iturbe, Scaling and elevation in river networks, Water Resour. Res. **25**, 2037 (1989).

[48] E. Tokunaga, Consideration on the composition of drainage networks and their evolution, Geographical Reports of Tokyo Metropolitan University **13**, 1 (1978).

[49] S. A. Veitzer and V. K. Gupta, Random self-similar river networks and derivations of generalized Horton's Laws in terms of statistical simple scaling, Water Resour. Res. **36**, 1033 (2000).

[50] Y. Kovchegov and I. Zaliapin, Tokunaga self-similarity arises naturally from time invariance, Chaos **28**, 041102 (2018).

[51] E. V. Chunikhina, Entropy rates for Horton self-similar trees, Chaos **28**, 081104 (2018).

[52] E. V. Chunikhina, Information theoretical analysis of self-similar trees, Ph.D. thesis, Oregon State University, 2018.

[53] P. Dodds and D. Rothman, Scaling, universality, and geomorphology, Ann. Rev. Earth Planet. Sci. **28**, 571 (2000).

[54] V. K. Gupta, S. L. Castro, and T. M. Over, On scaling exponents of spatial peak flows from rainfall and river network geometry, J. Hydro. **187**, 81 (1996).

[55] M. Cieplak, A. Giacometti, A. Maritan, A. Rinaldo, I. Rodriguez-Iturbe, and J. R. Banavar, Models of fractal river basins, J. Stat. Phys. **91**, 1 (1998).

[56] J. D. Pelletier and D. L. Turcotte, Shapes of river networks and leaves: Are they statistically similar, Phil. Trans. R. Soc. B **355**, 307 (2000).

[57] B. Lashermes and E. Foufoula-Georgiou, Area and width functions of river networks: New results on multifractal properties, Water Resour. Res. **43**, W09405 (2007).

[58] V. K. Gupta and E. C. Waymire, Multiscaling properties of spatial rainfall and river flow distributions, J. Geo. Research: Atmospheres **95**, 1999 (1990).

[59] H. W. Watson and F. Galton, On the probability of the extinction of families, J. Anthropological Institute of Great Britain and Ireland **4**, 138 (1875).

[60] R. Bhattacharya and E. C. Waymire, Bienaymé-Galton-Watson simple branching process and extinction, in *Random Walk, Brownian Motion, and Martingales*, Graduate Texts in Mathematics Vol. 292 (Springer, New York, 2021), p. 113.

[61] E. Foufoula-Georgiou, V. Ganti, and W. E. Dietrich, A non-local theory for sediment transport on hillslopes, J. Geophys. Res.: Earth Surface **115**, F00A16 (2010).

[62] A. Singh, L. Reinhardt, and E. Foufoula-Georgiou, Landscape reorganization under changing climatic forcing: Results from an experimental landscape, Water Resour. Res. **51**, 4320 (2015).

[63] S. D. Willett, S. W. McCoy, J. T. Perron, L. Goren, and C. Y. Chen, Dynamic reorganization of river basins, Science **343**, 6175 (2014).

[64] R. Mantilla, V. K. Gupta, and O. J. Mesa, Role of coupled flow dynamics and real network structures on Hortonian scaling of peak flows, J. Hydrology **322**, 155 (2006).

[65] A. Tejedor, A. Longjas, D. A. Edmonds, I. Zaliapin, T. T. Georgiou, A. Rinaldo, and E. Foufoula-Georgiou, Entropy and optimality in river deltas, Proc. Natl. Acad. Sci. USA **114**, 11651 (2017).

[66] A. Tejedor, A. Longjas, I. Zaliapin, and E. Foufoula-Georgiou, Delta channel networks: 1. A graph-theoretic approach for studying connectivity and steady state transport on deltaic surfaces, Water Resour. Res. **51**, 3998 (2015).

[67] P. Ossadnik, Branch order and ramification analysis of large diffusion-limited-aggregation clusters, Phys. Rev. A **45**, 1058 (1992).

[68] G. Daccord, J. Nittmann, and H. E. Stanley, Radial Viscous Fingers and Diffusion-Limited Aggregation: Fractal Dimension and Growth Sites, Phys. Rev. Lett. **56**, 336 (1986).

[69] Y. Kovchegov and I. Zaliapin, Invariance and attraction properties of Galton-Watson trees, Bernoulli **27**, 1789 (2021).

[70] D. J. Aldous, Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today, Stat. Sci. **16**, 23 (2001).

[71] M. Boguna, I. Bonamassa, M. De Domenico, S. Havlin, D. Krioukov, and M. A. Serrano, Network geometry, Nat. Rev. Phys. **3**, 114 (2021).

[72] T. Duquesne and M. Winkel, Hereditary tree growth and Lévy forests, Stochastic Processes and their Applications **129**, 3690 (2019).

[73] Y. Kovchegov, and I. Zaliapin, Dynamical pruning of rooted trees with applications to 1-D ballistic annihilation, J. Stat. Phys. **181**, 618 (2020).

[74] I. Zaliapin, H. H. L. Wong, and A. Gabrielov, Inverse cascade in a percolation model: Hierarchical description of time-dependent scaling, Phys. Rev. E **71**, 066118 (2005).

[75] L. V. Ahlfors, *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Complex Variable*, International series in pure and applied mathematics (McGraw-Hill, New York, 1953), Vol. 177.

[76] H. S. Wilf, *Generatingfunctionology* (Academic Press, Boston, 1990).

[77] R. N. Bhattacharya and E. C. Waymire, *A Basic Course in Probability Theory, Universitext* (Springer, New York, 2007), Vol. 69.