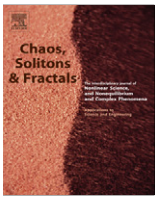
Chaos, Solitons & Fractals 45 (2012) 358–372

Contents lists available at SciVerse ScienceDirect

Chaos, Solitons & Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos



Tokunaga and Horton self-similarity for level set trees of Markov chains

Ilia Zaliapin [a,*], Yevgeniy Kovchegov [b]

[a] Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557-0084, USA
[b] Department of Mathematics, Oregon State University, Corvallis, OR 97331-4605, USA

ARTICLE INFO

Article history:
Received 31 January 2011
Accepted 23 November 2011
Available online 21 December 2011

ABSTRACT

The Horton and Tokunaga branching laws provide a convenient framework for studying self-similarity in random trees. The Horton self-similarity is a weaker property that addresses the *principal branching* in a tree; it is a counterpart of the power-law size distribution for elements of a branching system. The stronger Tokunaga self-similarity addresses so-called *side branching*. The Horton and Tokunaga self-similarity have been empirically established in numerous observed and modeled systems, and proven for two paradigmatic models: the critical Galton–Watson branching process with finite progeny and the finite-tree representation of a regular Brownian excursion. This study establishes the Tokunaga and Horton self-similarity for a tree representation of a finite symmetric homogeneous Markov chain. We also extend the concept of Horton and Tokunaga self-similarity to infinite trees and establish self-similarity for an infinite-tree representation of a regular Brownian motion. We conjecture that fractional Brownian motions are also Tokunaga and Horton self-similar, with self-similarity parameters depending on the Hurst exponent.

1. Introduction and motivation

Hierarchical branching organization is ubiquitous in nature. It is readily seen in river basins, drainage networks, bronchial passages, botanical trees, and snowflakes, to mention but a few (e.g., [1–4]). Empirical evidence reveals a surprising similarity among various natural hierarchies – many of them are closely approximated by so-called *self-similar trees* (SSTs) [1–3,5–16]. An SST preserves its statistical structure, in a sense to be defined, under the operation of *pruning*, i.e., cutting the leaves; this is why the SSTs are sometimes referred to as *fractal* trees [2]. A two-parametric subclass of *Tokunaga* SSTs, introduced by Tokunaga [9] in a hydrological context, plays a special role in theory and applications, as it has been shown to emerge in unprecedented variety of modeled and natural phenomena. The Tokunaga SSTs with a broad range of parameters are seen in studies of river networks [1,5,8–10,15,17], vein structure of botanical leaves [2,3], numerical analyses of diffusion limited aggregation [14,18], two dimensional site percolation [19–22], and nearest-neighbor clustering in Euclidean spaces [23]. The diversity of these processes and models hints at the existence of a universal (not problem-specific) underlying mechanism responsible for the Tokunaga self-similarity and prompts the question: *What probability models may produce Tokunaga self-similar trees*? An important answer to this question was given by Burd et al. [5] who studied Galton–Watson branching processes and have shown that, in this class, the Tokunaga self-similarity is a characteristic property of a *critical binary branching*, that is the discrete-time process that starts with a single progenitor and whose members equiprobably either split in two or die at every step. The critical binary Galton–Watson process is equivalent to the Shreve's random river network model, for which the Tokunaga self-similarity has been known for long time [1,5,8,15]. The Tokunaga self-similarity has also been rigorously established in a general hierarchical coagulation model of Gabrielov et al. [24] introduced in the framework of self-organized criticality, and in a random self-similar network model of Veitzer and Gupta [11] developed as an alternative to the Shreve's random network model for river networks.

---

* Corresponding author. Tel.: +1 775 784 6077; fax: +1 775 784 6378.
E-mail addresses: zal@unr.edu (I. Zaliapin), kovchegy@math.oregonstate.edu (Y. Kovchegov).

0960-0779/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chaos.2011.11.006

Prominently, the results of Burd et al. [5] reveal the Tokunaga self-similarity for any process represented by the finite Galton–Watson critical binary branching. In the context of this paper, the most important example is a regular Brownian motion, whose various connections to the Galton–Watson processes are well-known (see Pitman [25] for a modern review). For instance, the topological structure of the so-called h-excursions of a regular Brownian motion [26] and a Poisson sampling of a Brownian excursion [27] are equivalent to a finite critical binary Galton–Watson tree (Section 3 below explains the tree representation of time series), and hence these processes are Tokunaga self-similar.

This study further explores Tokunaga self-similarity by focusing on trees that describe the topological structure of the level sets of a time series or a real function, so-called *level-set trees*. Our set-up is closely related to the classical Harris correspondence between trees and finite random walks [28], and its later ramifications that include infinite trees with edge lengths [5,17,25,29–33]. The main result of this paper is the Tokunaga and closely related Horton self-similarity for the level-set trees of finite symmetric homogeneous Markov chains (SHMCs) – see Section 5, Theorem 4. Notably, the Tokunaga and Horton self-similarity concepts have been defined so far only for finite trees (e.g., [5,15,34]). We suggest here a natural extension of Tokunaga and Horton self-similarity to infinite trees and establish self-similarity for an infinite-tree representation of a regular Brownian motion. The suggested approach is based on the *forest of trees attached to the floor line* as described by Pitman [25]. Finally, we discuss the strong distributional self-similarity that characterizes Markov chains with exponential jumps.

The paper is organized as follows. Section 2 introduces planar rooted trees, trees with edge lengths, Harris paths, and spaces of random trees with the Galton–Watson distribution. The trees on continuous functions are described in Section 3. Several types of self-similarity for trees – Horton, Tokunaga, and distributional self-similarity – are discussed in Section 4. The main results of the paper are summarized in Section 5. Section 6 addresses special properties of exponential Markov chains that, in particular, enjoy the strong distributional self-similarity. Proofs are collected in Section 7. Section 8 concludes.

2. Trees

We introduce here planar trees, the corresponding Harris paths, and the space of Galton–Watson trees following Burd et al. [5], Ossiander et al. [17] and Pitman [25].

2.1. Planar rooted trees

Recall that a *graph* $\mathcal{G} = (V, E)$ is a collection of vertices (nodes) $V = \{v_i\}$, $1 \leqslant i \leqslant N_V$ and edges (links) $E = \{e_k\}$, $1 \leqslant k \leqslant N_E$. In a *simple* graph each edge is defined as an unordered pair of distinct vertices: $\forall 1 \leqslant k \leqslant N_E$, $\exists! 1 \leqslant i, j \leqslant N_V$, $i \neq j$ such that $e_k = (v_i, v_j)$ and we say that the edge *k connects* vertices v_i and v_j. Furthermore, each pair of vertices in a simple graph may have at most one connecting edge.

A *tree* is a connected simple graph $T = (V, E)$ without cycles, which readily gives $N_E = N_V - 1$. In a *rooted* tree, one node is designated as a root; this imposes a natural *direction* of edges as well as the parent–child relationship between the vertices. Specifically, we follow [5] to represent a labeled (planar) tree T rooted at ϕ by a bijection between the set of vertices V and set of finite integer-valued sequences $\langle i_1, \ldots, i_n \rangle \in T$ such that

(i) $\phi = \langle \emptyset \rangle$,
(ii) if $\langle i_1, \ldots, i_n \rangle \in T$ then $\langle i_1, \ldots, i_k \rangle \in T$ $\forall 1 \leqslant k \leqslant n$, and
(iii) if $\langle i_1, \ldots, i_n \rangle \in T$ then $\langle i_1, \ldots, i_{n-1}, j \rangle \in T$ $\forall 1 \leqslant j \leqslant i_n$.

This representation is illustrated in Fig. 1. If $v = \langle i_1, \ldots, i_n \rangle \in T$ then $u = \langle i_1, \ldots, i_{n-1} \rangle \in T$ is called the *parent* of v, and v is a *child* of u. A *leaf* is a vertex with no children. The number of children of a vertex $u = \langle i_1, \ldots, i_n \rangle \in T$ equals to $c(u) = \max\{j\}$ over such j that $\langle u, j \rangle \equiv \langle i_1, \ldots, i_n, j \rangle \in T$. A *binary* labeled rooted tree is represented by a set of binary sequences with elements $i_k = 1, 2$, where 1, 2 represent the left and right planar directions, respectively. Two trees are called *distinct* if they are represented by distinct sets of the vertex-sequences. We complete each tree T by a special *ghost edge* ϵ attached to the root ϕ, so each vertex in the tree has a single parental edge. A natural direction of edges is from a vertex v to its parent v_p.

In these settings, the total number of distinct trees with n leaves, according to the Cayley's formula, is n^{n-2}. The total number of distinct binary trees with n leaves is given by the $(n-1)$th Catalan number [25]

$$C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}.$$

2.2. Trees with edge-lengths and Harris path

A tree with *edge-lengths* $T = (V, E, W)$ assigns a positive lengths $w(e)$ to each edge e, $W = \{w(e)\}$; such trees are also called *weighted trees* (e.g., [5,17]). The sum of all edge lengths is called the *tree length*; we write $\text{LENGTH}(T) = \sum_e w(e)$. We call the pair (V, E) a *combinatorial tree* and write $(V, E) = \text{SHAPE}(T)$, emphasizing that the lengths are disregarded in this representation.

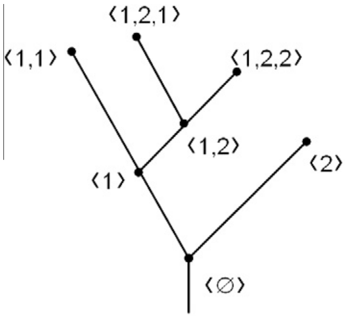


Fig. 1. Representation of a tree via a set of finite sequences $\langle i_1, \ldots, i_n \rangle$: an example.