



Patricia's Bad Distributions

Louigi Addario-Berry  

Department of Mathematics and Statistics, McGill University, Montréal, Canada

Pat Morin  

School of Computer Science, Carleton University, Ottawa, Canada

Ralph Neininger¹  

Institute of Mathematics, Goethe University Frankfurt, Germany

Abstract

The height of a random PATRICIA tree built from independent, identically distributed infinite binary strings with arbitrary diffuse probability distribution μ on $\{0, 1\}^{\mathbb{N}}$ is studied. We show that the expected height grows asymptotically sublinearly in the number of leaves for any such μ , but can be made to exceed any specific sublinear growth rate by choosing μ appropriately.

2012 ACM Subject Classification Theory of computation → Sorting and searching

Keywords and phrases PATRICIA tree, random tree, height of tree, analysis of algorithms

Digital Object Identifier 10.4230/LIPIcs.AofA.2024.25

Related Version *Previous Version:* <https://doi.org/10.48550/arXiv.2403.05269>

Acknowledgements This research was mainly done during the Sixteenth Annual Workshop on Probability and Combinatorics at McGill University's Bellairs Institute in Holetown, Barbados. We thank Bellairs Institute for its hospitality and support. We also thank the referees and Jasmin Straub for comments on a draft of this note.

1 Introduction and results

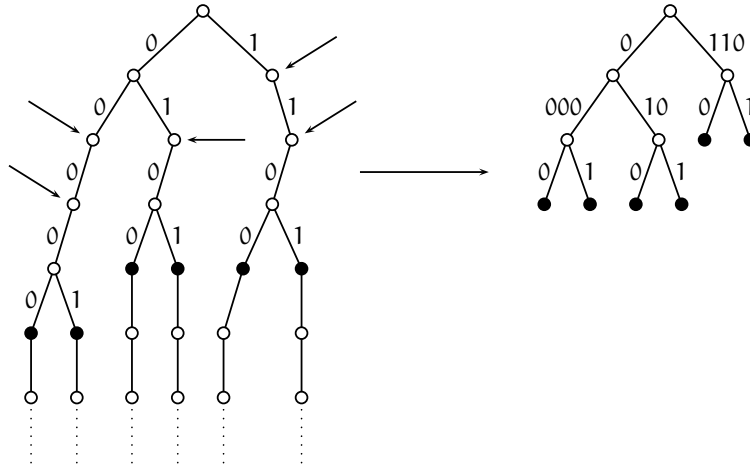
The PATRICIA tree is a space efficient data structure for strings which improves on the trie. For the purpose of this note it is sufficient to introduce these tree structures for binary strings: Label the nodes of the complete infinite rooted binary tree by the elements of $\cup_{k=0}^{\infty} \{0, 1\}^k$, starting at the root with \emptyset and left and right child of a node labelled $v \in \{0, 1\}^k$ with $v0$ and $v1$, respectively. Here, for $v \in \{0, 1\}^k$ with $v = (v_1, \dots, v_k)$ we abbreviate v as $v = v_1 \dots v_k$ and denote $vi := v_1 v_2 \dots v_k i$ for $i = 0, 1$.

The coming definitions are depicted in Figure 1. For distinct infinite binary strings $x_1, \dots, x_n \in \{0, 1\}^{\mathbb{N}}$ a finite tree called a *trie* (or *radix search tree*) to represent the strings x_1, \dots, x_n is constructed by first associating with each x_i the infinite path in $\cup_{k=0}^{\infty} \{0, 1\}^k$ consisting of the nodes whose labels are the prefixes of x_i . The node labelled with the shortest such prefix that is not a prefix of any x_j with $j \in \{1, \dots, n\} \setminus \{i\}$ becomes a leaf in the trie representing string x_i for $i = 1, \dots, n$. The resulting tree, which is a finite binary tree with n leaves, is the trie representing x_1, \dots, x_n . Next, starting from the trie, all vertices with out-degree 1 (i.e. with exactly one child) are deleted and the resulting gaps are closed by merging the two nodes which formed a deleted edge. This results in the PATRICIA tree, which was introduced independently by Morrison [21] and Gwehenberger [12] and first systematically analysed by Knuth [17]. The PATRICIA tree contains all the information needed to retrieve the strings and to perform operations such as sorting, searching and selecting; for broad expositions, see [18, 20, 25].

¹ Corresponding author



PATRICIA trees have been analysed assuming various probabilistic models for the input strings; where usually the infinite strings are assumed to be independent and identically distributed over $\{0, 1\}^{\mathbb{N}}$. Note that atoms of such a distribution result in identical strings with positive probability, and in this case the construction of the trie does not lead to a finite tree. Hence, the law of the strings is usually assumed to be diffuse (non-atomic). Special cases of such diffuse probability distributions have been considered in the analysis of algorithms on strings such as the Bernoulli models, Markov model, dynamical sources or the density model; see [23, 24, 6, 4, 14, 11, 19, 1, 15, 13] and the references given in these papers.



■ **Figure 1** On the left the trie for the strings $00000\dots$, $00001\dots$, $0100\dots$, $0101\dots$, $1100\dots$, and $1101\dots$ is shown. Its leaves are the full black vertices, the indicated children of the full black vertices do not belong to the trie. Vertices with out-degree 1 within the trie are indicated by arrows. On the right the resulting PATRICIA tree by deleting corresponding edges is shown.

In the present note we focus on the height of a PATRICIA tree, which is the maximal (graph) distance of any leaf from the root. The asymptotic behavior of the height of tries and PATRICIA trees under the Bernoulli models is covered by Pittel [23, 24] and Devroye [5, 7]. For example, for the height H_n^{syB} of the PATRICIA tree constructed from n independent strings under the symmetric Bernoulli model, i.e. all bits being independent and Bernoulli($\frac{1}{2}$) distributed, Pittel [23] obtained as $n \rightarrow \infty$ that

$$\frac{H_n^{\text{syB}}}{\log_2 n} \rightarrow 1 \text{ almost surely.}$$

This shows an asymptotic 50% improvement of the PATRICIA tree over the trie, for which the limit constant for the same probabilistic model is 2 instead of 1. For general diffuse laws concentration of the height of PATRICIA trees is studied (assuming only independence of the infinite strings not necessarily identical distribution) by Devroye [8] based on results from [2]; see also [16] for concentration of the height of PATRICIA trees in the Bernoulli model.

While such studies aim to show that the height behaves well with respect to applications from algorithms, Evans and Wakolbinger [9, 10] studied these random tree structures as tree-valued transient Markov chains from the perspective of Doob–Martin boundary theory. They asked (private communication) how high PATRICIA trees can grow for arbitrary diffuse probability distributions of the strings (see [10, Section 5] for specific examples). The subject

of the present note is to answer this question by Theorems 1 and 2: The expected height grows always sublinearly, but can be made to exceed any fixed sublinear growth rate by the choice of an appropriate diffuse law.

For a diffuse probability distribution μ on $\{0, 1\}^{\mathbb{N}}$ and $(\Xi^{(j)})_{j \in \mathbb{N}}$ a sequence of independent and identically distributed random strings with law μ we denote by H_n^μ the height of the PATRICIA tree constructed from $\Xi^{(1)}, \dots, \Xi^{(n)}$.

► **Theorem 1.** *For all diffuse probability distributions μ on $\{0, 1\}^{\mathbb{N}}$ we have, as $n \rightarrow \infty$, that*

$$\frac{\mathbb{E}[H_n^\mu]}{n} \rightarrow 0, \quad \text{and} \quad \frac{H_n^\mu}{n} \rightarrow 0 \text{ almost surely.}$$

► **Theorem 2.** *For any sequence $\alpha = (\alpha_n)_{n \in \mathbb{N}}$ of positive numbers with $\alpha_n \rightarrow \infty$ as $n \rightarrow \infty$ there exists a diffuse probability distribution $\nu = \nu^{(\alpha)}$ on $\{0, 1\}^{\mathbb{N}}$ such that for all n sufficiently large*

$$\frac{\mathbb{E}[H_n^\nu]}{n/\alpha_n} \rightarrow \infty, \quad \text{and} \quad \frac{H_n^\nu}{n/\alpha_n} \rightarrow \infty \text{ almost surely.}$$

We call a law ν on $\{0, 1\}^{\mathbb{N}}$ causing large expected heights $\mathbb{E}[H_n^\nu]$ *bad* since such laws are undesirable from the point of view of the efficiency of algorithms based on PATRICIA trees. The remaining part of the present note contains proofs of these two theorems.

► **Remark 1.** For the density model, which is a subclass of the diffuse distributions on $\{0, 1\}^{\mathbb{N}}$, the asymptotics of Theorem 1 were obtained by Devroye [6, page 419]. There, also bad distributions with asymptotic properties as in our Theorem 2 are constructed for sequences $\alpha_n = n^\varepsilon$ with $0 < \varepsilon < 1$.

2 Proofs

2.1 Proof of Theorem 1

We start with a technical observation:

► **Lemma 3.** *Suppose μ is a diffuse probability distribution on $\{0, 1\}^{\mathbb{N}}$, and let $\Xi = (\xi_i)_{i \in \mathbb{N}}$ be random with law μ . Then for all ε there exists $k = k(\varepsilon) \in \mathbb{N}$ such that for any string $v = v_1 \dots v_k \in \{0, 1\}^k$, $\mathbb{P}(\xi_1 \dots \xi_k = v_1 \dots v_k) < \varepsilon$.*

Proof. Suppose for a contradiction that there exists $\varepsilon > 0$ such that for all $k \in \mathbb{N}$ there is a string $v_1 \dots v_k \in \{0, 1\}^k$ such that $\mathbb{P}(\xi_1 \dots \xi_k = v_1 \dots v_k) \geq \varepsilon$. Then by a compactness argument shown below there exists an infinite string $v = (v_i)_{i \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}}$ such that for all $k \in \mathbb{N}$, $\mathbb{P}(\xi_1 \dots \xi_k = v_1 \dots v_k) \geq \varepsilon$. The events $\{\xi_1 \dots \xi_k = v_1 \dots v_k\}$ are decreasing in k , so this implies that

$$\mathbb{P}(\Xi = v) = \lim_{k \rightarrow \infty} \mathbb{P}(\xi_1 \dots \xi_k = v_1 \dots v_k) \geq \varepsilon,$$

which contradicts the assumption that μ is diffuse.

It remains to show the existence of the infinite string $v = (v_i)_{i \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}}$ such that for all $k \in \mathbb{N}$, $\mathbb{P}(\xi_1 \dots \xi_k = v_1 \dots v_k) \geq \varepsilon$. Consider $\{0, 1\}$ as a topological space with the discrete topology (all subsets being open) and $\{0, 1\}^{\mathbb{N}}$ as the product space with the product topology. As a product of compact spaces $\{0, 1\}^{\mathbb{N}}$ is compact. It is also a Hausdorff space. The projections $\Pi_k : \{0, 1\}^{\mathbb{N}} \rightarrow \{0, 1\}^k$ given by

$$(v_i)_{i \in \mathbb{N}} \xrightarrow{\Pi_k} v_1 \dots v_k$$

25:4 Patricia's Bad Distributions

are continuous for all $k \in \mathbb{N}$. Hence, the set

$$\begin{aligned} V_k &:= \{(v_i)_{i \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}} \mid \mathbb{P}(\xi_1 \dots \xi_k = v_1 \dots v_k) \geq \varepsilon\} \\ &= \bigcup_{\substack{v_1 \dots v_k \in \{0, 1\}^k \\ \mathbb{P}(\xi_1 \dots \xi_k = v_1 \dots v_k) \geq \varepsilon}} \Pi_k^{-1}(\{v_1 \dots v_k\}) \end{aligned}$$

is closed and thus compact in $\{0, 1\}^{\mathbb{N}}$. This implies that $(V_k)_{k \in \mathbb{N}}$ is a nested sequence of non-empty, compact sets. Now, Cantor's intersection theorem implies

$$\bigcap_{k=1}^{\infty} V_k \neq \emptyset.$$

Any element v of $\bigcap_{k=1}^{\infty} V_k$ has the desired property. \blacktriangleleft

Proof of Theorem 1. Fix a diffuse probability distribution μ on $\{0, 1\}^{\mathbb{N}}$. Let $\Xi^{(j)} = (\xi_i^{(j)})_{i \in \mathbb{N}}$ for $j \in \mathbb{N}$ be independent, identically distributed with law μ and denote by T_n the PATRICIA tree built from $\Xi^{(1)}, \dots, \Xi^{(n)}$.

We first show that $H_n^\mu/n \rightarrow 0$ almost surely. Fix any $\varepsilon \in (0, 1/4)$. Let $k = k(\varepsilon)$ be as in Lemma 3, so that for any string $v = v_1 \dots v_k \in \{0, 1\}^k$, if $\Xi = (\xi_i)_{i \in \mathbb{N}}$ has law μ then $\mathbb{P}(\xi_1 \dots \xi_k = v_1 \dots v_k) < \varepsilon$. To prove $H_n^\mu/n \rightarrow 0$ almost surely we first show that

$$\mathbb{P}(\exists n_0 \forall n \geq n_0 : H_n^\mu \leq k + 2\varepsilon n) = 1. \quad (1)$$

Note that if the event

$$E_{n,k} := \bigcup_{v_1 \dots v_k \in \{0, 1\}^k} \{|\{1 \leq j \leq n : \xi_1^{(j)} \dots \xi_k^{(j)} = v_1 \dots v_k\}| \geq 2\varepsilon n\}$$

does not occur then the subtrees of T_n rooted at nodes $v \in \{0, 1\}^k$ all have at most $2\varepsilon n$ leaves and so height less than $2\varepsilon n$; thus if $E_{n,k}$ does not occur then $H_n^\mu \leq k + 2\varepsilon n$. It follows that

$$\begin{aligned} &\mathbb{P}(\exists n_0 \forall n \geq n_0 : H_n^\mu \leq k + 2\varepsilon n) \\ &\geq \mathbb{P}(E_{n,k} \text{ occurs for at most finitely many values } n) \\ &= \mathbb{P}\left(\left(\limsup_{n \rightarrow \infty} E_{n,k}\right)^c\right), \end{aligned}$$

so to prove (1) it suffices to show that the probability of $\limsup_{n \rightarrow \infty} E_{n,k}$ is 0. For this, simply note that

$$\begin{aligned} \mathbb{P}(E_{n,k}) &\leq \sum_{v_1 \dots v_k \in \{0, 1\}^k} \mathbb{P}(|\{1 \leq j \leq n : \xi_1^{(j)} \dots \xi_k^{(j)} = v_1 \dots v_k\}| \geq 2\varepsilon n) \\ &\leq 2^k \mathbb{P}(Y_n \geq 2\varepsilon n), \end{aligned}$$

where Y_n has the Binomial distribution $\text{Bin}(n, \varepsilon)$; the second inequality holds since the events that $\xi_1^{(j)} \dots \xi_k^{(j)} = v_1 \dots v_k$ are independent for distinct $1 \leq j \leq n$, and each has probability at most ε . A Chernoff bound then gives

$$\mathbb{P}(E_{n,k}) \leq 2^k e^{-\varepsilon n/2}.$$

Since this is summable, it follows by the first Borel–Cantelli lemma that

$$\mathbb{P}\left(\limsup_{n \rightarrow \infty} E_{n,k}\right) = 0,$$

hence we obtain (1). Now, note that for any $m_0 \in \mathbb{N}$,

$$\left\{ \frac{H_n^\mu}{n} \rightarrow 0 \right\} = \bigcap_{m=m_0}^{\infty} \bigcup_{n_0=1}^{\infty} \bigcap_{n=n_0}^{\infty} \left\{ \frac{H_n^\mu}{n} \leq \frac{3}{m} \right\}.$$

Thus, for $\varepsilon = \frac{1}{m}$ with fixed $m \geq m_0$ we can choose n sufficiently large so that $k(\varepsilon)/n \leq \varepsilon$ and obtain

$$\{H_n^\mu \leq k(\varepsilon) + 2\varepsilon n\} \subset \left\{ \frac{H_n^\mu}{n} \leq \frac{3}{m} \right\}$$

and see that (1) implies $H_n^\mu/n \rightarrow 0$ almost surely.

Finally, note that by construction of the PATRICIA tree we deterministically have $H_n^\mu \leq n - 1$, thus $H_n^\mu/n \leq 1$. Hence, we obtain from $H_n^\mu/n \rightarrow 0$ almost surely and dominated convergence that $\mathbb{E}[H_n^\mu]/n \rightarrow 0$. ◀

2.2 Proof of Theorem 2

As building blocks for our bad distributions we first define a set of auxiliary probability distributions $(\mu_N, N \in \mathbb{N})$, on $\{0, 1\}^{\mathbb{N}}$ as follows. For fixed $N \in \mathbb{N}$ we choose T uniformly at random from $\{1, \dots, N^2\}$. Independently of T , let $(B_i)_{i \in \mathbb{N}}$ be independent Bernoulli($\frac{1}{2}$)-distributed random variables. Then define a sequence $(\vartheta_i)_{i \in \mathbb{N}}$ by

$$\vartheta_i = \begin{cases} 0, & \text{if } i < T, \\ 1, & \text{if } i = T, \\ B_{i-T}, & \text{if } i > T. \end{cases} \tag{2}$$

Now, μ_N is defined as the law of the string $\Theta = (\vartheta_i)_{i \in \mathbb{N}}$. Note that by definition μ_N is diffuse for all $N \in \mathbb{N}$. We use the notation

$$\langle \Theta \rangle := \min\{i \in \mathbb{N} \mid \vartheta_i = 1\}$$

for the index of the first entry of Θ equal to 1.

► **Lemma 4.** *For any $n \in \{1, \dots, N\}$ we have $\mathbb{E}[H_n^{\mu_N}] \geq n - 2$.*

Proof. Let $1 \leq n \leq N \in \mathbb{N}$ and $\Theta^{(1)}, \dots, \Theta^{(n)}$ be i.i.d. with law μ_N . We consider the set $A := \{\langle \Theta^{(1)} \rangle, \dots, \langle \Theta^{(n)} \rangle\} \subset \{1, \dots, N^2\}$. By construction of the PATRICIA tree we have

$$H_n^{\mu_N} \geq |A| - 1, \tag{3}$$

where $|A|$ denotes the cardinality of A , i.e., the number of distinct elements within the set $\{\langle \Theta^{(1)} \rangle, \dots, \langle \Theta^{(n)} \rangle\}$. For all $1 \leq i < j \leq n$ we have $\mathbb{P}(\langle \Theta^{(i)} \rangle = \langle \Theta^{(j)} \rangle) = 1/N^2$. Hence, we obtain

$$\mathbb{E}[|A|] \geq n - \mathbb{E} \left[\sum_{1 \leq i < j \leq n} \mathbf{1}_{\{\langle \Theta^{(i)} \rangle = \langle \Theta^{(j)} \rangle\}} \right] \geq n - \frac{n^2}{2N^2} \geq n - 1, \tag{4}$$

since $n \leq N$. Now, (3) and (4) imply the assertion. ◀

Proof of Theorem 2. Without loss of generality we may assume that $\alpha_n = o(n)$. There exists an $n_0 \in \mathbb{N}$ such that $\alpha_n \geq 8$ for all $n \geq n_0$. We define $\beta_n := \lfloor \log_2 \alpha_n \rfloor - 2$ and a sequence $(A(n))_{n \in \mathbb{N}}$ as a generalized inverse of $(\beta_n)_{n \in \mathbb{N}}$ by

$$A(n) := \max\{m \in \mathbb{N} \mid \beta_m \leq n\}, \quad n \in \mathbb{N}. \tag{5}$$

First, a probability distribution $\mu^{(\alpha)}$ on $\{0, 1\}^{\mathbb{N}}$ is obtained in two stages. Let G be a random variable with geometric distribution with parameter $\frac{1}{2}$, i.e., with $\mathbb{P}(G = k) = (\frac{1}{2})^k$ for $k \in \mathbb{N}$. Then define a sequence $(\lambda_i)_{i \in \mathbb{N}}$ by

$$\lambda_i = \begin{cases} 0, & \text{if } i < G, \\ 1, & \text{if } i = G, \\ \vartheta_{i-G}, & \text{if } i > G, \end{cases} \quad (6)$$

where $\Theta = (\vartheta_i)_{i \in \mathbb{N}}$, conditional on $\{G = k\}$, has law $\mu_{A(k)}$ defined in (2) with $A(\cdot)$ defined in (5). We then define $\mu = \mu^{(\alpha)}$ as the law of $\Lambda = (\lambda_i)_{i \in \mathbb{N}}$. Since the $\mu_{A(k)}$ are diffuse, we obtain that μ is diffuse.

Now, let $\Lambda^{(j)} = (\lambda_i^{(j)})_{i \in \mathbb{N}}$ for $j \in \mathbb{N}$ be independent with law μ . For $n \geq n_0$, by construction,

$$X_n := \left| \left\{ 1 \leq j \leq n : \left(\lambda_1^{(j)}, \dots, \lambda_{\beta_n}^{(j)} \right) = (0, \dots, 0, 1) \right\} \right|$$

is $\text{Bin}(n, 2^{-\beta_n})$ -distributed. To get rid of the floors in the definition of β_n denote by X'_n a $\text{Bin}(n, 4/\alpha_n)$ -distributed random variable. Note that $2^{-\beta_n} \geq 4/\alpha_n$. By Okamoto's inequality, see [22] or [3, Exercise 2.12] we have

$$\begin{aligned} \mathbb{P}\left(X_n < \frac{2n}{\alpha_n}\right) &\leq \mathbb{P}\left(X'_n - \frac{4n}{\alpha_n} < -\frac{2n}{\alpha_n}\right) \\ &\leq \exp\left(-\frac{n(2/\alpha_n)^2}{2(4/\alpha_n)(1-4/\alpha_n)}\right) \leq \exp\left(-\frac{n}{2\alpha_n}\right). \end{aligned}$$

Hence, with high probability at least $\lceil 2n/\alpha_n \rceil$ of the n strings start with the prefix $(0, \dots, 0, 1)$ of length β_n and thus have suffixes $(\lambda_{\beta_n+1}^{(j)}, \lambda_{\beta_n+2}^{(j)}, \dots)$ drawn independently from $\mu_{A(\beta_n)}$ for the respective j . For all $n \geq n_0$ we have $\lceil 2n/\alpha_n \rceil \leq n \leq A(\beta_n)$. Hence, by Lemma 4, $\lceil 2n/\alpha_n \rceil$ such strings cause an expected height of at least $2n/\alpha_n - 2$. Together we obtain for all sufficiently large n , note also $\alpha_n = o(n)$, that

$$\begin{aligned} \mathbb{E}[H_n^\mu] &\geq \mathbb{P}\left(X_n \geq \frac{2n}{\alpha_n}\right) \mathbb{E}\left[H_n^\mu \mid X_n \geq \frac{2n}{\alpha_n}\right] \\ &\geq \left(1 - \exp\left(-\frac{n}{2\alpha_n}\right)\right) \left(\frac{2n}{\alpha_n} - 2\right) \\ &\geq \frac{n}{\alpha_n}. \end{aligned}$$

Since the sequence $(\log \alpha_n)$ tends to infinity the present proof implies the existence of a diffuse probability distribution $\nu = \nu^{(\alpha)}$ on $\{0, 1\}^{\mathbb{N}}$ such that $\mathbb{E}[H_n^\nu] \geq n/\log \alpha_n$ for all sufficiently large $n \in \mathbb{N}$, hence

$$\frac{\mathbb{E}[H_n^\nu]}{n/\alpha_n} \rightarrow \infty.$$

To prove the second statement of Theorem 2 we use the following bound from Devroye [8, page 21]: for any diffuse probability distribution μ on $\{0, 1\}^{\mathbb{N}}$ and any $t > 0$,

$$\mathbb{P}(H_n^\mu \leq \mathbb{E}[H_n^\mu] - t) \leq \exp\left(-\frac{t^2}{2\mathbb{E}[H_n^\mu + 1]}\right) \leq \exp\left(-\frac{t^2}{2n}\right). \quad (7)$$

We now consider the probabilities

$$\mathbb{P}\left(H_n^\nu \leq \frac{n}{\log^2 \alpha_n}\right) = \mathbb{P}\left(H_n^\nu \leq \mathbb{E}[H_n^\nu] - \left(\mathbb{E}[H_n^\nu] - \frac{n}{\log^2 \alpha_n}\right)\right) \quad (8)$$

and note that for all sufficiently large n we have

$$\mathbb{E}[H_n^\nu] - \frac{n}{\log^2 \alpha_n} \geq \frac{n}{\log \alpha_n} - \frac{n}{\log^2 \alpha_n} = n \frac{\log(\alpha_n) - 1}{\log^2 \alpha_n} \geq \frac{n}{\log^2 n}. \quad (9)$$

Combining (7)–(9) we obtain

$$\mathbb{P}\left(H_n^\nu \leq \frac{n}{\log^2 \alpha_n}\right) \leq \exp\left(-\frac{n}{2\log^4 n}\right)$$

for all sufficiently large n . Since these upper bounds are summable it follows from the first Borel–Cantelli Lemma that $\liminf_{n \rightarrow \infty} H_n^\nu / (n / \log^2 \alpha_n) \geq 1$ almost surely, hence

$$\frac{H_n^\nu}{n/\alpha_n} \rightarrow \infty \text{ almost surely.}$$

Thus, ν has the properties claimed in Theorem 2. ◀

References

- 1 Ali Akhavi, Frédéric Paccaut, and Brigitte Vallée. Building Sources of Zero Entropy: Rescaling and Inserting Delays. In Mark Daniel Ward, editor, *33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022)*, volume 225 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 1:1–1:28, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.AofA.2022.1.
- 2 Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. A sharp concentration inequality with applications. *Random Structures Algorithms*, 16(3):277–292, 2000. doi:10.1002/(SICI)1098-2418(200005)16:3<277::AID-RSA4>3.0.CO;2-1.
- 3 Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux. doi:10.1093/acprof:oso/9780199535255.001.0001.
- 4 Julien Clément, Philippe Flajolet, and Brigitte Vallée. Dynamical sources in information theory: a general analysis of trie structures. *Algorithmica*, 29(1-2):307–369, 2001. Average-case analysis of algorithms (Princeton, NJ, 1998). doi:10.1007/BF02679623.
- 5 Luc Devroye. A note on the probabilistic analysis of patricia trees. *Random Structures & Algorithms*, 3(2):203–214, 1992. doi:10.1002/rsa.3240030209.
- 6 Luc Devroye. A study of trie-like structures under the density model. *Ann. Appl. Probab.*, 2(2):402–434, 1992. URL: [http://links.jstor.org/sici?sici=1050-5164\(199205\)2:2<402:ASOTSU>2.0.CO;2-F&origin=MSN](http://links.jstor.org/sici?sici=1050-5164(199205)2:2<402:ASOTSU>2.0.CO;2-F&origin=MSN).
- 7 Luc Devroye. Universal limit laws for depths in random trees. *SIAM Journal on Computing*, 28(2):409–432, 1998. doi:10.1137/S0097539795283954.
- 8 Luc Devroye. Universal asymptotics for random tries and PATRICIA trees. *Algorithmica*, 42(1):11–29, 2005. doi:10.1007/s00453-004-1137-7.
- 9 Steven N. Evans and Anton Wakolbinger. Radix sort trees in the large. *Electronic Communications in Probability*, 22(none):1–13, 2017. doi:10.1214/17-ECP77.
- 10 Steven N. Evans and Anton Wakolbinger. PATRICIA bridges. In *Genealogies of interacting particle systems*, volume 38 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, pages 231–265. World Sci. Publ., Hackensack, NJ, [2020] ©2020.
- 11 Michael Fuchs, Hsien-Kuei Hwang, and Vytas Zacharovas. An analytic approach to the asymptotic variance of trie statistics and related structures. *ArXiv*, abs/1303.4244, 2013. URL: <https://api.semanticscholar.org/CorpusID:14429040>.
- 12 Gernot Gwehenberger. Anwendung einer binären Verweiskettenmethode beim Aufbau von Listen. *Elektronische Rechenanlagen*, 10:223–226, 1968.

- 13 Jasper Ischebeck. Central limit theorems for fringe trees in patricia tries. *Preprint arXiv:2305.14900 [math.PR]*, 2023. doi:10.48550/arXiv.2305.14900.
- 14 Svante Janson. Renewal theory in the analysis of tries and strings. *Theor. Comput. Sci.*, 416:33–54, January 2012. doi:10.1016/j.tcs.2011.10.007.
- 15 Svante Janson. Central limit theorems for additive functionals and fringe trees in tries. *Electronic Journal of Probability*, 27(none):1–63, 2022. doi:10.1214/22-EJP776.
- 16 Charles Knessl and Wojciech Szpankowski. Limit laws for the height in PATRICIA tries. *J. Algorithms*, 44(1):63–97, 2002. Analysis of algorithms. doi:10.1016/S0196-6774(02)00212-2.
- 17 Donald E. Knuth. *The art of computer programming. Volume 3*. Addison-Wesley Series in Computer Science and Information Processing. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973. Sorting and searching.
- 18 Donald E. Knuth. *The art of computer programming. Vol. 3*. Addison-Wesley, Reading, MA, 1998. Sorting and searching, Second edition [of MR0445948].
- 19 Kevin Leckey, Ralph Neininger, and Wojciech Szpankowski. *Towards More Realistic Probabilistic Models for Data Structures: The External Path Length in Tries under the Markov Model*, pages 877–886. Society for Industrial and Applied Mathematics, 2013. doi:10.1137/1.9781611973105.63.
- 20 Hosam M. Mahmoud. *Evolution of random search trees*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 1992. A Wiley-Interscience Publication.
- 21 Donald R. Morrison. Patricia—practical algorithm to retrieve information coded in alphanumeric. *Journal of the ACM*, 15:514–534, 1968.
- 22 Masashi Okamoto. Some inequalities relating to the partial sum of binomial probabilities. *Ann. Inst. Statist. Math.*, 10:29–35, 1958. doi:10.1007/BF02883985.
- 23 B. Pittel. Asymptotical growth of a class of random trees. *Ann. Probab.*, 13(2):414–427, 1985. doi:10.1214/aop/1176993000.
- 24 Boris Pittel. Paths in a random digital tree: limiting distributions. *Adv. in Appl. Probab.*, 18(1):139–155, 1986. doi:10.2307/1427240.
- 25 Wojciech Szpankowski. *Average case analysis of algorithms on sequences*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2001. With a foreword by Philippe Flajolet. doi:10.1002/9781118032770.