Subtractive random forests with two choices *

Francisco Calvillo
LPSM, Sorbonne Université, 4 Place Jussieu, 75005 Paris, France

Luc Devroye
School of Computer Science, McGill University,
Montreal, Canada

Gábor Lugosi
Department of Economics and Business,
Pompeu Fabra University, Barcelona, Spain
ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain
Barcelona Graduate School of Economics

May 16, 2024

Abstract

Recommendation systems are pivotal in aiding users amid vast online content. Broutin, Devroye, Lugosi, and Oliveira proposed Subtractive Random Forests (SURF), a model that emphasizes temporal user preferences. Expanding on SURF, we introduce a model for a multi-choice recommendation system, enabling users to select from two independent suggestions based on past interactions. We evaluate its effectiveness and robustness across diverse scenarios, incorporating heavy-tailed distributions for time delays. By analyzing user topic evolution, we assess the system's consistency. Our study offers insights into the performance and potential enhancements of multi-choice recommendation systems in practical settings.

*Gábor Lugosi acknowledges the support of Ayudas Fundación BBVA a Proyectos de Investigación Científica 2021 and the Spanish Ministry of Economy and Competitiveness grant PID2022-138268NB-I00, financed by MCIN/AEI/10.13039/501100011033, FSE+MTM2015-67304-P, and FEDER, EU.

1 Introduction

Today's online platforms have a record number of items available for their users. In many cases presenting them as a catalogue is no longer a good option and can make it difficult for the user to find what they are interested in. Recommendation systems have become an essential tool to cope with the overload of information available on the web.

A well-known way to approach recommendation systems today is through deep learning, and many of the most effective recommendation systems use this principle, see [14]. However, the mechanism of these systems is notable for its opacity. In the realm of online sequential recommendation systems, Broutin, Devroye, Lugosi, and Oliveira [6] present an approach, offering a framework that considers the temporal aspect of user preferences by a simple mechanism. They define a model recommending topics based on a random time delay. The topic that is recommended at time n is the same that was recommended at time $n-Z_n$, where $(Z_k)_{k\geq 1}$ is a sequence of i.i.d. random variables identically distributed on $\{1,2,\ldots\}$. This approach led them to study a family of random forests called subtractive random forests (SURF), allowing a detailed structural study. However, one expects a recommendation system to make several recommendations at a time, not just one.

To model this more realistic scenario, our paper proposes a two-choice recommendation system inspired by the same idea. We envision a scenario where users are presented with two independent recommendations, drawn from the same temporal recommendation mechanism and allowing users to select the most appealing option. The resulting model has some intriguing mathematical properties and the main goal of this paper is to analyze the model in order to understand the long-term behavior of such recommendation systems.

1.1 A two-choice recommendation model

We assume that the initial pool of topics is infinite and represented by the set of non-positive integers $\{0,-1,-2,\ldots\}$. We now consider independent and identically distributed random variables Z, W, $(Z_n)_{n\geq 1}$ and $(W_n)_{n\geq 1}$ on the set of positive integers $\mathbb{N}$. Define

$$q_i = \mathbb{P}(Z = i) \quad \text{and} \quad p_i = \mathbb{P}(Z \geq i), \quad i \geq 1 \ .$$

Each topic $i \leq 0$ is assigned a preference value U_i within the range $[0,1]$, where we assume that $(U_n)_{n\leq 0}$ is a sequence of random variables independent of the sequences $(Z_n)_{n\geq 1}$ and $(W_n)_{n\geq 1}$. These preference values are user-dependent.

Following [6], we can define a sequential random colouring of the positive integers as follows. For each non-positive integer $i \leq 0$, define its colour $C_i = i$

and its preference value $V_i = U_i$. We assign colours and preference values to the positive integers $n \in \mathbb{N}$ by the recursion

$$\begin{cases} C_n = C_{n-Z_n} \text{ and } V_n = V_{n-Z_n} & \text{if } V_{n-Z_n} < V_{n-W_n} \\ C_n = C_{n-W_n} \text{ and } V_n = V_{n-W_n} & \text{otherwise} \end{cases} .$$

Thus, by identifying the colour C_n as the topic recommended at time $n \geq 1$, this definition means that at the time instance $n \geq 1$, the user receives two recommendations (the topic C_{n-Z_n}, and the topic C_{n-W_n}), and chooses the one with the lowest (best) preference value.

This process naturally defines a random directed graph whose vertex set is $\mathbb{Z}$ by drawing an edge from any positive integer $n \geq 1$ to any integer $m < n$ if and only if $m = n - Z_n$ or $m = n - W_n$. Vertices with negative index are called *leaves*, as is customary for final nodes in dags (i.e., directed acyclic graphs). Let T_n denote the set of vertices that are reached from the vertex $n \geq 1$. This set can be defined recursively by saying that $T_n = \{n\}$ for any $n \leq 0$, and

$$T_n = \{n, n - Z_n, n - W_n\} \cup T_{n-Z_n} \cup T_{n-W_n}$$

for any $n \geq 1$. Define

$$\mathcal{L}_n = T_n \cap \{0, -1, -2, \ldots\}$$

as the set of vertices with a non-positive index that is reached from n. In other words, $\mathcal{L}_n$ is the set of leaves in the subgraph of vertices reached from n. Note that

$$C_n = \operatorname*{argmin}_{i \in \mathcal{L}_n} U_i \quad \text{and} \quad V_n = \min_{i \in \mathcal{L}_n} U_i .$$

One can assess the system's long-term performance by analyzing the asymptotic behavior of the sequence $(V_n)_{n \geq 1}$. If V_n approaches $\inf_{i \leq 0} U_i$ as $n \to \infty$, then this means that after waiting a sufficient amount of time, the topics recommended to the user tend to align more closely with their preferences. This gives us a consistency criterion for our model. We consider the three following configurations of the preference values:

(i) $U_0 < U_{-1} < U_{-2} < U_{-3} \ldots$

(ii) $U_0 > U_{-1} > U_{-2} > U_{-3} \ldots$

(iii) $(U_n)_{n \leq 0}$ i.i.d. and uniformly distributed in $[0, 1]$.

These assumptions correspond to natural scenarios and allow us to study long-term consistency of the recommendation system. In the first case, topics have a monotone preference with the most recent (corresponding to $i = 0$) being the

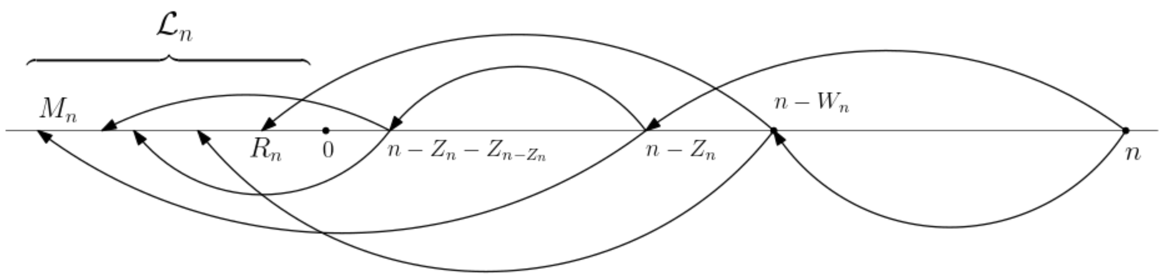

Figure 1: An illustration of the definitions of M_n, R_n, and $\mathcal{L}_n$.

most preferred one. In the second case, older topics are preferred, while in the third case topics have a random order of preference. We say that the system is consistent if, on the long run, it offers near-optimal recommendations to the user in terms of the their preference. In Section 1.3 we rigorously define various notions of consistency, corresponding to the scenarios above. These definitions lead us to the study of the sequence $(V_n)_{n\geq 1}$ by studying the set of leaves $\mathcal{L}_n$. We will pay particular attention to the three following random variables:

- $R_n = \max\{i : i \in \mathcal{L}_n\}$ (the rightmost leaf in $\mathcal{L}_n$)
- $M_n = \min\{i : i \in \mathcal{L}_n\}$ (the leftmost leaf in $\mathcal{L}_n$)
- $L_n = |\mathcal{L}_n|$ (the number of leaves reached from n).

Each of these random variables helps us understand the long-term behavior of V_n in each of the three initial configurations described above. One would ideally hope that $V_n \to \inf_{i\leq 0} U_i$. In configuration (i) this is equivalent to $R_n \to 0$, in configuration (ii) it is equivalent to $M_n \to -\infty$, while in case (iii) to $L_n \to \infty$. However, one cannot expect to have $R_n \to 0$ in a general case. To introduce a more reasonable consistency criterion for case (i), we require boundedness of the sequence $(R_n)_{n\geq 1}$.

Let us define the subsets $T_n^Z = T_n^W = S_n = \{n\}$ for any $n \leq 0$, and let

$$T_n^Z = \{n\} \cup T_{n-Z_n}^Z ,$$

$$T_n^W = \{n\} \cup T_{n-W_n}^W ,$$

$$\text{and } S_n = \{n\} \cup S_{n-\min(Z_n, W_n)}$$

for any $n \geq 1$, so that $T_n^Z \subset T_n$ (resp. $T_n^W \subset T_n$) is the chain of vertices that can be reached from n by following only the Z-edges (resp. W-edges), and $S_n \subset T_n$ is the chain of vertices that can be reached from n by following only the shortest edges.

Finally, for any $n \geq 1$, let $r_n = \mathbb{P}\left(0 \in T_n^Z\right) = \mathbb{P}\left(0 \in T_n^W\right)$ denote the probability that vertex 0 belongs to the Z-chain starting at vertex n. We set $r_0 = 1$ so that r_n satisfies the recursion

$$r_n = \sum_{i=1}^{n} q_i r_{n-i} . \tag{1.1}$$

1.2 Related work

This paper builds on the work of Broutin, Devroye, Lugosi and Oliveira [6], who studied a one-choice version of this model by examining the properties of a graph whose vertex set is $\mathbb{Z}$, and whose edges are $\{n, n - Z_n\}$ for every $n \geq 1$, given some sequence $(Z_n)_{n\geq 1}$ of i.i.d. random variables distributed in $\{1, 2, 3, \ldots\}$. This defines a random forest where each non-positive integer is the root of a tree. They also define a random coloring by setting $C_i = i$ for all $i \leq 0$, and $C_n = C_{n-Z_n}$ for all $n \geq 1$. Each tree in the forest corresponds to a colour (i.e., a topic). The graph that we consider in our paper is nothing else than the superposition of two independent copies of this forest, and the sets T_n^Z and T_n^W defined above are the chains that connect the vertex n to the root of its tree in each of these two copies.

One of the main results in [6] shows that the process has two distinct behaviors: if $\mathbb{E}Z = \infty$, then, almost surely, all trees in the forest are finite, meaning that no topic will be recommended infinitely often, while if $\mathbb{E}Z < \infty$, after some random amount of time, all vertices connect to the same tree almost surely, meaning that all topics recommended to the user become the same. In the two-choice model however, there are three distinct regimes: $\mathbb{E}Z < \infty$; $\mathbb{E}Z = \infty$ yet $\mathbb{E}\min(Z, W) < \infty$; and $\mathbb{E}\min(Z, W) = \infty$. We call these the light-, moderate-, and heavy-tailed regimes.

The random forest in [6] appears as a subgraph of a random graph model that was previously studied by Hammond and Sheffield in [10]. Given a sequence $(Z_n)_{n\in\mathbb{Z}}$ of independent random variables, identically distributed in $\{1, 2, 3, \ldots\}$, they consider a graph with vertex set $\mathbb{Z}$ such that vertices $n, m \in \mathbb{Z}$, with $m < n$, are connected by an edge if and only if $m = n - Z_n$. One can obtain the random forest in [6] by removing all edges between any two vertices with a non-positive index. They show that the random graph has almost surely a unique component when the sum $\sum_{n=1}^{\infty} r_n^2$ converges, and the graph almost surely has infinitely many connected components when the sum diverges.

Other variants of the (single-choice) random subtractive process were studied, in different contexts, by Blath, Jochen, González, Kurt, and Spano [4], Chierichetti, Kumar, and Tomkins [8], Baccelli and Sodre [3], Baccelli, Haji-Mirsadeghi, and Khezeli [2], Baccelli, Haji-Mirsadeghi, and Khaniha [1], and Igelbrink and Wakolbinger [12].

1.3 Results

The goal of this paper is to understand the behavior of the two-choice model by studying its consistency in each of the three configurations (i),(ii), and (iii), for three types of tails, light, moderate, and heavy:

1. $\mathbb{E}Z < \infty$ (light tails)
2. $\mathbb{E}Z = \infty$ and $\mathbb{E}\min(Z, W) < \infty$ (moderate tails)
3. $\mathbb{E}\min(Z, W) = \infty$ (heavy tails)

We now introduce the main definitions of consistency.

Definition 1. *We say that the system is consistent in configuration (ii) if $M_n \to -\infty$; while it is consistent in configuration (iii) if $L_n \to \infty$. We differentiate between strong and weak consistency based on whether convergence occurs almost surely or in probability. In the case of configuration (i), we say that the system is weakly consistent if $(R_n)_{n\geq 1}$ is a tight sequence. It is strongly consistent if the sequence $(R_n)_{n\geq 1}$ is almost surely bounded.*

Remark 1. *Pareto tails. Consider a distribution with Pareto tails, that is, $q_n = \Theta(1/n^{1+\alpha})$, with $\alpha > 0$. This implies that $p_n = \Theta(1/n^{\alpha})$. When $\alpha > 1$, we see that $\mathbb{E}Z < \infty$. The more interesting situation is when $\alpha \in (0, 1]$. For $\alpha \in (1/2, 1]$, we have $\mathbb{E}Z = \infty$, yet $\mathbb{E}\min(Z, W) < \infty$. Finally, for $\alpha \in (0, 1/2]$, we have $\mathbb{E}\min(Z, W) = \infty$. For light-tailed Z, we recall from Broutin, Devroye, Lugosi, Oliveira [6] that $r_n \to 1/\mathbb{E}Z$. For moderate and heavy tails, however, we have $r_n \to 0$. When $\alpha \in (0, 1)$, then Z is in the domain of attraction of the extremal stable law with parameter α, which itself has a tail distribution function that decays at the rate $1/n^{\alpha}$ (see, e.g., Ibragimov and Linnik [11], Zolotarev [15] or Malevich [13]). One can then deduce that r_n decays at the rate $1/n^{1-\alpha}$. In particular, $\sum_n r_n^2 < \infty$ when $\alpha \in (0, 1/2)$, which roughly corresponds to the case of heavy tails.*

Next we describe the main results of the paper. For a summary, see the table at the end of this section.

Light tails

When $\mathbb{E}Z < \infty$, time jumps are so short that T_n can't reach distant leaves, thus forcing the set $\mathcal{L}_n$ to be bounded. This case does not require much work, since one can quickly determine consistency or inconsistency in each of the three configurations just by looking at the sequence $(M_n)_{n\geq 1}$, which happens to be bounded almost surely.

Theorem 2. *Let* $M_\infty = \inf_{n\in\mathbb{N}} M_n$. *If* $\mathbb{E}Z < \infty$, *then* M_∞ *is finite almost surely. Thus, the sequences* $(R_n)_{n\in\mathbb{N}}, (M_n)_{n\in\mathbb{N}}$ *and* $(L_n)_{n\in\mathbb{N}}$ *are almost surely bounded.*

Proof. Note that

$$\mathbb{P}(M_\infty \le -x) \le 2\sum_{n=1}^{\infty} \mathbb{P}(n - Z_n \le -x) = 2\sum_{n=1}^{\infty} p_{n+x} .$$

Since $\mathbb{E}Z < \infty$, we know that $\sum_{n=1}^{\infty} p_{n+x}$ is finite and goes to 0 as $x \to +\infty$. Thus, by continuity of measure,

$$\mathbb{P}(M_\infty = -\infty) = \lim_{x\to+\infty} \mathbb{P}(M_\infty \le -x) = 0 .$$

Since $|M_n| \ge L_n$ and $|M_n| \ge |R_n|$, this also implies that $(L_n)_{n\in\mathbb{N}}$ and $(R_n)_{n\in\mathbb{N}}$ are almost surely bounded. ∎

Moderate and heavy tails

When $\mathbb{E}\min(Z, W) = \infty$, a different behavior emerges:

Theorem 3. *Let* Z, W, $(Z_n)_{n\ge1}$ *and* $(W_n)_{n\ge1}$ *be i.i.d. random variables. Assume furthermore that* $\mathbb{E}\min(Z, W) = \infty$. *Then, with probability one, the number of leaves satisfies*

$$\liminf_{n\to\infty} L_n < \infty .$$

In other words, the system is not strongly consistent in the heavy-tailed regime.

Proof. Note that if $\min(Z_n, W_n) \ge n$, then $L_n \in \{1, 2\}$. Thus, to prove Theorem 3, it suffices to show that with probability one, $\min(Z_n, W_n) \ge n$ happens infinitely often. This follows by the second Borel-Cantelli lemma since the events

$$(\{\min(Z_n, W_n) \ge n\})_{n\in\mathbb{N}}$$

are independent and

$$\sum_{n\ge1} \mathbb{P}(\min(Z_n, W_n) \ge n) = \mathbb{E}(\min(Z, W)) = \infty .$$

∎

The behavior of the set $\mathcal{L}_n$ for moderate and heavy tails is better understood by examining its extreme points M_n and R_n. The following two theorems are proved in Section 3.

Theorem 4. *If $q_i > 0$ for all $i \in \mathbb{N}$ and $\mathbb{E}Z = \infty$, then, with probability one,*

$$M_n \to -\infty \ .$$

Thus, in configuration (ii) the system is strongly consistent in the moderate and heavy-tailed regimes.

Strong (in-)consistency in configuration (i) is established in the next result:

Theorem 5. *Let Z, W, $(Z_n)_{n\geq 1}$ and $(W_n)_{n\geq 1}$ be i.i.d. random variables.*

1. *If $\mathbb{E}\min(Z,W) < \infty$, then $\sup_{n\in\mathbb{N}} |R_n| < \infty$ almost surely.*
2. *If $\mathbb{E}\min(Z,W) = \infty$, then $\sup_{n\in\mathbb{N}} |R_n| = \infty$ almost surely.*

Given that the leftmost leaf in $\mathcal{L}_n$ goes to $-\infty$ and the rightmost leaf remains bounded in the moderate heavy-tail regime (i.e., when $\mathbb{E}Z = \infty$ and $\mathbb{E}\min(Z,W) < \infty$), one can expect that the total number of leaves reached by the vertex n diverges under these two assumptions. Thus, strong consistency is observed in each of the three configurations (i), (ii), and (iii).

On the other hand, only configuration (i) leads to strong consistency when the distribution of Z has a heavy-tail (i.e., when $\mathbb{E}\min(Z,W) = \infty$). However, when $\sum_{n=1}^{\infty} r_n^2 < \infty$, the model remains weakly consistent in configuration (iii), as established by the next two results. Theorems 6 and 7 are proved in Sections 4.1 and 4.2, respectively.

Theorem 6. *Assume that the distribution of Z is not supported by any proper additive subgroup of the integers and that Z exhibits a moderate-sized tail (i.e., $\mathbb{E}\min(Z,W) < \infty$ and $\mathbb{E}Z = \infty$). Then*

$$L_n \to \infty \text{ almost surely.}$$

Theorem 7. *When Z has a heavy tail (i.e., $\mathbb{E}\min(Z,W) = \infty$), and $\sum_{n\geq 0} r_n^2 < \infty$, then*

$$L_n \to \infty \text{ in probability.}$$

We recall from Remark 1 that the summability of r_n^2 is assured for nearly all heavy-tailed Z.

1.4 Summary

The following table summarizes our findings. Observe that weak and strong behavior coincide in most cases. They only differ in the extreme heavy tail case ($\mathbb{E}\min(Z,W) = \infty$). For an optimal user experience, one would need $L_n \to \infty$ almost surely, and the only case that assures this is when we have moderate tails. In addition, such moderate tails also guarantee strong consistency for the decreasing and increasing input scenarios (as is apparent from the strong consistency of M_n and R_n).

	$\mathbb{E}Z < \infty$	$\mathbb{E}Z = \infty$ and $\mathbb{E}\min(Z,W) < \infty$	$\mathbb{E}\min(Z,W) = \infty$
	Strong consistency		
$M_n \to -\infty$ a.s.	no (Thm. 2)	yes (Thm. 4)	yes (Thm. 4)
$L_n \to \infty$ a.s.	no (Thm. 2)	yes (Thm. 6)	no (Thm. 3)
R_n bounded a.s.	yes (Thm. 2)	yes (Thm. 5)	no (Thm. 3)
	Weak consistency		
$M_n \xrightarrow{\mathbb{P}} -\infty$	no (Thm. 2)	yes (Thm. 4)	yes (Thm. 4)
$L_n \xrightarrow{\mathbb{P}} \infty$	no (Thm. 2)	yes (Prop. 11)	yes if $\sum r_n^2 < \infty$ (Thm. 7)
R_n is tight	yes (Thm. 2)	yes (Thm. 5)	

The rest of the paper contains the proofs of the results stated above. In Section 2 we recall some properties of the single-choice model that are useful in our analysis. In Section 3, Theorems 4 and 5 are proven. The main technical content of the paper is presented in Section 4 where the number of leaves is examined in both the moderate-, and heavy-tailed cases, culminating in the proofs of Theorems 6 and 7.

2 The one-choice model

Some of our proofs use the observation that the graph contains several copies of the subtractive random forest studied in [6]. For example, Theorem 1 in [6] states that there is a unique infinite tree in the forest when $\mathbb{E}Z < \infty$:

Theorem 8 (Broutin, Devroye, Lugosi, Oliveira [6])**.** *Assume $\mathbb{E}Z < \infty$ and $q_1 > 0$. Then there exists a positive random variable N with $\mathbb{P}(N < \infty) = 1$ such that, with probability one,*

$$N \in T_n^Z \quad \text{for all } n \geq N .$$

Remark 2. *Note that the actual statement of the theorem in [6] is slightly weaker since it only asserts that there is a random variable N, finite almost surely, such that for all $n \geq N$, the vertex n belongs to the same tree as the vertex N. This does not necessarily imply that $N \in T_n^Z$. Nevertheless, one can add this property of N without changing the proof provided in [6].*

Define $Z_n' = \min(Z_n, W_n)$ and $Z' = \min(Z, W)$ and note that $S_n = T_n^{Z'}$. By Theorem 8 we obtain the following result for moderate and light tails.

Corollary 9. *If* $\mathbb{E}\min(Z,W) < \infty$ *and* $q_1 > 0$*, then there exists a positive random variable N with* $\mathbb{P}(N < \infty) = 1$ *such that, with probability one,*

$$N \in S_n \quad \textit{for all } n \geq N$$

When $\mathbb{E}Z = \infty$, Theorem 3 in [6] states that every tree in the forest is finite. This translates as follows.

Theorem 10 (Broutin, Devroye, Lugosi, Oliveira [6])**.** *If* $\mathbb{E}Z = \infty$ *and* $q_i > 0$ *for all* $i \geq 1$*, then*

$$\mathbb{P}(\cup_{i\leq 0}\left[\left|\{n \geq 1 : i \in T_n^Z\}\right| = \infty\right]) = 0 .$$

3 Rightmost and leftmost leaves

As stated in Theorem 4, the leftmost leaf M_n reached by vertex n diverges to $-\infty$ whenever $\mathbb{E}Z = \infty$. This can be seen as a consequence of the behavior of the one-choice model, as one only needs to check the divergence to negative infinity of the leaf $\min T_n^Z$ reached by vertex n by following only the Z-edges.

Proof. [Proof of Theorem 4] From Theorem 10 we know that with probability one, for all $k \leq 0$, there are at most a finite number of integers $n \geq 1$ such that $k \in T_n^Z$. Thus, the sequence $(\min T_n^Z)_{n\in\mathbb{N}}$, taking values in $\{0,-1,-2,\ldots\}$, cannot take the same value an infinite number of times, so it goes to $-\infty$ as $n \to \infty$ with probability one. This concludes the proof, since $M_n \leq \min T_n^Z$ for all $n \geq 1$.

∎

Proof. [Proof of Theorem 5] Assume that $\mathbb{E}\min(Z,W) < \infty$. From Corollary 9, we know that there exists some random variable N with $\mathbb{P}(N < \infty) = 1$ such that, with probability one, $N \in S_n$ for all $n \geq N$. In particular, for any $n \geq N$ we have $T_N \subset T_n$, which implies that $|R_N| \geq |R_n|$. Thus, with probability one, we have

$$\sup_{n\in\mathbb{N}} |R_n| \leq \max_{1\leq k\leq N} |R_k| < \infty .$$

When $\mathbb{E}\min(Z,W) = \infty$, we have

$$\sum_{n\geq 1} \mathbb{P}(\min(Z_n, W_n) \geq 2n) = \sum_{n\geq 1} p_{2n}^2 = \infty .$$

Hence, by the second Borel-Cantelli lemma, we know that with probability one we have $\min(Z_n, W_n) \geq 2n$ infinitely often, which implies that $R_n \leq -n$ infinitely often and proves the second statement of the theorem.

∎

4 Number of leaves

4.1 Moderate tails

Consider the number L_n of leaves reached by vertex n, when the distribution of Z is such that $\mathbb{E}Z = \infty$ and $\mathbb{E}\min(Z, W) < \infty$. We first show weak divergence, implying weak consistency in configuration (iii).

Proposition 11. *Assume that the distribution of Z is not supported by any proper additive subgroup of the integers. If $\mathbb{E}Z = \infty$ and $\mathbb{E}\min(Z, W) < \infty$, then*

$$L_n \to \infty \text{ in probability as } n \to \infty.$$

Lemma 12. *Define*

$$J_n = \sum_{m=1}^{n} \mathbb{1}_{m \in S_n} \mathbb{1}_{\max(Z_m, W_m) \geq m} .$$

Then $\mathbb{E}J_n \to \infty$ as $n \to \infty$.

Proof. J_n counts the number of vertices in S_n with a positive index that are connected to a leaf through its longest outgoing edge. Note that for $n \geq m \geq 1$, the events $\{m \in S_n\}$ and $\{\max(Z_m, W_m) \geq m\}$ are independent. It follows that

$$\mathbb{E}J_n = \sum_{m=1}^{n} \mathbb{P}(m \in S_n)\mathbb{P}(\max(Z_m, W_m) \geq m) = \sum_{m=1}^{n} v_{n-m}\mathbb{P}(\max(Z_m, W_m) \geq m)$$

where $v_m := \mathbb{P}(0 \in S_m)$ for all $m \geq 1$, and $v_0 = 1$.
Let $f(z) = \sum_{n \geq 0} v_n z^n$, $\tilde{q}_m = \mathbb{P}(\min(Z, W) = m)$ and $g(z) = \sum_{n \geq 1} \tilde{q}_n z^n$. In view of

$$v_n = \sum_{m=0}^{n-1} \tilde{q}_{n-m} v_m,$$

we have

$$f(z) = 1 + \sum_{n=1}^{\infty} z^n \sum_{m=0}^{n} \tilde{q}_{n-m} v_m = 1 + g(z)f(z) = \frac{1}{1 - g(z)}.$$

Since $\min(Z, W)$ is not supported by any proper additive subgroup of the integers, the Erdős-Feller-Pollard theorem [9] implies that

$$\lim_{n \to \infty} v_n = \frac{1}{\mathbb{E}\min(Z, W)} . \tag{4.1}$$

Thus, for some constant $c > 0$, we have

$$\mathbb{E}J_n \geq c \sum_{m=1}^{n} p_m \to \infty \quad \text{as } n \to \infty . \tag{4.2}$$

■

We use the second-moment method to show that $J_n \to \infty$ in probability.

Lemma 13. *Assume that $\mathbb{E}Z = \infty$ and $\mathbb{E}\min(Z, W) < \infty$, and let $\mathbb{V}J_n$ denote the variance of the random variable J_n. Then*

$$\frac{\mathbb{V}J_n}{(\mathbb{E}J_n)^2} \to 0 \quad \text{as } n \to \infty . \tag{4.3}$$

In particular, we have:

$$\mathbb{P}\left(J_n \leq \frac{\mathbb{E}J_n}{2}\right) = \mathbb{P}\left(J_n - \mathbb{E}J_n \leq -\frac{\mathbb{E}J_n}{2}\right) \leq 4\frac{\mathbb{V}J_n}{(\mathbb{E}J_n)^2} \to 0 ,$$

and therefore $J_n \to \infty$ in probability as $n \to \infty$.

Proof. [Proof of Lemma 13] To show (4.3), we prove that $\mathbb{E}[J_n^2]/(\mathbb{E}J_n)^2 \to 1$ as $n \to \infty$. We first define

$$p_n^* = \mathbb{P}(\max(Z, W) \geq n) = 2p_n - p_n^2 .$$

Note that for $n \geq 1$,

$$\begin{aligned} (\mathbb{E}J_n)^2 &= 2 \sum_{1 \leq k < m \leq n} p_m^* p_k^* v_{n-m} v_{n-k} + \sum_{1 \leq m \leq n} (p_m^*)^2 v_{n-m}^2 \\ &= 2 \sum_{1 \leq k < m \leq n} p_m^* p_k^* v_{n-m} v_{n-k} + O(1) \end{aligned} \tag{4.4}$$

since $\sum_{1 \leq m \leq n} (p_m^*)^2 v_{n-m}^2$ is bounded by $4\mathbb{E}\min(Z, W) < \infty$.

Now, define for any $m \geq 1$ the random variable $X_m = \max(Z_m, W_m)$ and observe that for any $1 \leq k < m \leq n$, the events $\{X_k \geq k\}$ and $\{X_m \geq m, m \in S_n, k \in S_n\}$ are independent. Thus, for any fixed $n \geq 1$, we may write

$$\begin{aligned} \mathbb{E}\left[J_n^2\right] &= 2 \sum_{1 \leq k < m \leq n} \mathbb{P}(m \in S_n, k \in S_n, X_m \geq m, X_k \geq k) + \mathbb{E}J_n \\ &= 2 \sum_{1 \leq k < m \leq n} p_k^* \mathbb{P}(m \in S_n, k \in S_n, X_m \geq m) + \mathbb{E}J_n \\ &= 2 \sum_{1 \leq k < m \leq n} p_k^* \mathbb{P}[k \in S_n, X_m \geq m | m \in S_n] \mathbb{P}(m \in S_n) + \mathbb{E}J_n \end{aligned}$$

If $m \in S_n$, the event $\{k \in S_n, X_m \geq m\}$ can only happen if one of the two random variables Z_m and W_m is greater or equal to m and the other one is smaller than $m-k$, for otherwise k could not belong to S_n when $1 \leq k < m \leq n$. Thus, using the independence of Z_m with respect to W_m, $\{k \in S_n\}$ and $\{m \in S_n\}$, and using the fact that $p_m^* + p_m^2 = 2p_m$,

$$
\begin{aligned}
&\mathbb{E}\left[J_n^2\right] && (4.5)\\
&= 2 \sum_{1 \leq k < m \leq n} p_k^* \times 2\mathbb{P}[Z_m \geq m, W_m \leq m-k, k \in S_n | m \in S_n]\,\mathbb{P}(m \in S_n) + \mathbb{E}J_n \\
&= 2 \sum_{1 \leq k < m \leq n} p_k^* (p_m^* + p_m^2) v_{n-m} \mathbb{P}[W_m \leq m-k, k \in S_n | m \in S_n] + \mathbb{E}J_n \\
&= 4 \sum_{1 \leq k < m \leq n} p_k^* (p_m^* + p_m^2) v_{n-m} \sum_{i=1}^{k-m} \mathbb{P}(k \in S_{m-i})\,\mathbb{P}(W_m = i) + \mathbb{E}J_n \\
&= 2 \sum_{1 \leq k < m \leq n} p_k^* (p_m^* + p_m^2) v_{n-m} \sum_{i=1}^{m-k} v_{m-k-i} q_i + \mathbb{E}J_n \,. && (4.6)
\end{aligned}
$$

Observe that

$$
\begin{aligned}
\sum_{1 \leq k < m \leq n} p_k^* p_m^2 v_{n-m} \sum_{i=1}^{m-k} q_i v_{m-k-i} &\leq \sum_{1 \leq k < m \leq n} p_k^* p_m^2 v_{n-m} \\
&\leq \sum_{1 \leq k < m \leq n} p_k^* p_m^2 \\
&\leq \left(\sum_{m=1}^{\infty} p_m^2\right)\left(\sum_{k=1}^{n} p_k^*\right) \\
&= O(\mathbb{E}J_n)
\end{aligned}
$$

by using the inequality in (4.2), which leads to

$$
\mathbb{E}[J_n^2] = 2 \sum_{1 \leq k < m \leq n} p_k^* p_m^* v_{n-m} \sum_{i=1}^{m-k} v_{m-k-i} q_i + O(\mathbb{E}J_n)\,. \tag{4.7}
$$

Fix some $\epsilon > 0$. By (4.1), we know that there exists some constant $x > 0$ such that for all $m \geq x$, we have $|v_m - \lambda| < \epsilon$, where $\lambda = 1/(\mathbb{E}\min(Z, W))$.

If $m \geq x$, see that

$$
\begin{aligned}
\sum_{i=1}^{m} q_i v_{m-i} &\leq \sum_{i=1}^{m-x} q_i (\lambda + \epsilon) + \sum_{m-x < i \leq m} q_i \\
&\leq (\lambda + \epsilon) + p_{m-x}\,.
\end{aligned}
$$

Thus, there exists some constant $y > x$ such that for all $m \geq y$,

$$\sum_{i=1}^{m} q_i v_{m-i} \leq \lambda + 2\epsilon \ .$$

Moreover,

$$\sum_{\substack{1\leq k<m\leq n \\ s.t.\ m-k<y}} p_m^* v_{n-m} \underbrace{p_k^* \sum_{i=1}^{m-k} v_{m-k-i} q_i}_{\leq 1} \leq \sum_{1\leq m\leq n} p_m^* v_{n-m} \sum_{k=m-y}^{m} 1$$
$$= (y+1)\mathbb{E}J_n = O\left(\mathbb{E}J_n\right) \ ,$$

and

$$\sum_{\substack{1\leq k<m\leq n \\ s.t.\ n-m<y}} p_k^* \underbrace{p_m^* v_{n-m} \sum_{i=1}^{m-k} v_{m-k-i} q_i}_{\leq 1} \leq \sum_{n-y\leq m\leq n} 1 \sum_{k=1}^{n} p_k^*$$
$$\leq (y+1) \sum_{k=1}^{n} p_k^* = O\left(\mathbb{E}J_n\right) \ .$$

Putting everything together, we have that

$$\mathbb{E}[J_n^2] = \sum_{m=1}^{n-y} \sum_{k=1}^{m-y} p_m^* p_k^* v_{n-k} \sum_{i=1}^{m-k} q_i v_{m-k-i} + O\left(\mathbb{E}J_n\right) \ . \tag{4.8}$$

Similarly, we have

$$(\mathbb{E}J_n)^2 = \sum_{s=1}^{n-y} \sum_{k=1}^{m-y} p_m^* p_k^* v_{n-m} v_{n-k} + O\left(\mathbb{E}J_n\right) \ . \tag{4.9}$$

Finally, observe that for any $n > y$ we have:

$$\sum_{m=1}^{n-y} \sum_{k=1}^{m-y} p_m^* p_k^* \underbrace{v_{n-m} \sum_{i=1}^{m-k} q_i v_{m-k-i}}_{\leq (\lambda+2\epsilon)^2} \leq (\lambda + 2\epsilon)^2 \sum_{m=1}^{n-y} \sum_{k=1}^{m-y} p_m^* p_k^* \tag{4.10}$$

and

$$\sum_{m=1}^{n-y} \sum_{k=1}^{m-y} p_m^* p_k^* \underbrace{v_{n-m} v_{n-k}}_{\geq (\lambda-\epsilon)^2} \geq (\lambda - \epsilon)^2 \sum_{m=1}^{n-y} \sum_{k=1}^{m-y} p_m^* p_k^* \ . \tag{4.11}$$

From (4.8), (4.9), (4.10) and (4.11) it follows that

$$\limsup_{n\to\infty} \frac{\mathbb{E}[J_n^2]}{(\mathbb{E}J_n)^2} \le \left(\frac{\lambda+2\epsilon}{\lambda-\epsilon}\right)^2 .$$

Since this is true for any $\epsilon > 0$ small enough, and $\mathbb{E}[J_n^2] \ge (\mathbb{E}J_n)^2$, we conclude that

$$\frac{\mathbb{E}[J_n^2]}{(\mathbb{E}J_n)^2} \to 1$$

as $n \to \infty$. ∎

Proposition 11 follows from the fact that $J_n \to \infty$ in probability. In order to see this, we only need to check that the fact that $J_n \to \infty$ in probability implies that the number of leaves connected to T_n goes to infinity as well.

Proof. [Proof of Proposition 11] Let us define

$$\mathcal{L}_n^* = \{k \le 0 : \textit{there exists an } m \in S_n \textit{ such that } \max(Z_m, W_m) = m-k\} .$$

In other words, $\mathcal{L}_n^*$ is the subset of vertices in $\mathcal{L}_n$ that are connected to some vertex in S_n with positive index through the longest edge (i.e., given by $\max(Z,W)$).

Given $n \ge 1$ and $k \in \mathbb{Z}$, we introduce the random variable

$$D_{n,k} = \sum_{m=1}^{n} \mathbb{1}_{\max(Z_m,W_m)=m-k}.$$

Note that for any fixed $x > 0$,

$$\mathbb{P}(|\mathcal{L}_n^*| < x) \tag{4.12}$$

$$\le \mathbb{P}\left(\left[\max_{k\in\mathcal{L}_n^*} \sum_{m\in S_n} \mathbb{1}_{\max(Z_m,W_m)=m-k} < \frac{\mathbb{E}J_n}{2x}\right] \cap [|\mathcal{L}_n^*| < x]\right)$$

$$+ \mathbb{P}\left(\cup_{k\le 0}\left[D_{n,k} \ge \frac{\mathbb{E}J_n}{2x}\right]\right) . \tag{4.13}$$

Since J_n can be written as

$$J_n = \sum_{k\in\mathcal{L}_n^*} \sum_{m\in S_n} \mathbb{1}_{\max(Z_m,W_m)=m-k} ,$$

we have that

$$\mathbb{P}\left(\left[\max_{k\in\mathcal{L}_n^*} \sum_{m\in S_n} \mathbb{1}_{\max(Z_m,W_m)=m-k} < \frac{\mathbb{E}J_n}{2x}\right] \cap [|\mathcal{L}_n^*| < x]\right)$$

$$\leq \mathbb{P}\left(J_n \leq \frac{\mathbb{E}J_n}{2}\right)$$
$$\leq 4\frac{\mathbb{V}J_n}{(\mathbb{E}J_n)^2} \to 0 \,. \tag{4.14}$$

Moreover, for any $k \leq 0$ and $y > 0$, by Chernoff's bound [7], [5], we have that

$$\begin{aligned}\mathbb{P}(D_{n,k} > y) &= \mathbb{P}\left(\sum_{m=1}^{n} \mathbb{1}_{\max(Z_m,W_m)=m-k} > y\right) \\ &\leq \exp\left(y - p^*_{1-k} - y\log\left(\frac{y}{p^*_{1-k}}\right)\right) \\ &\leq \left(\frac{e}{y}\right)^y (p^*_{1-k})^y \,.\end{aligned}$$

Let $y_n = \mathbb{E}J_n/(2x)$ and assume that n is large enough so we have $y_n \geq 2$ and

$$\sum_{m=1}^{\infty}(p^*_m)^{y_n} \leq \sum_{m=1}^{\infty}(p^*_m)^2 < \infty \text{ (since } \mathbb{E}\min(Z,W) < \infty \text{) .}$$

Thus, by the union bound,

$$\begin{aligned}\mathbb{P}\left(\cup_{k\leq 0}\left[D_{n,k} \geq \frac{\mathbb{E}J_n}{2x}\right]\right) &\leq \sum_{k\leq 0}\left(\frac{e}{y_n}\right)^{y_n} (p^*_{1-k})^{y_n} \\ &\leq \left(\frac{e}{y_n}\right)^{y_n} \sum_{k=1}^{\infty}(p^*_k)^2 \to 0 \,.\end{aligned} \tag{4.15}$$

By putting (4.13), (4.14), and (4.15) together we have that $\mathbb{P}(|\mathcal{L}^*_n| < x) \to 0$ as $n \to \infty$ for every fixed $x > 0$, which proves Proposition 11. ■

Proof. [Proof of Theorem 6] For any $n \geq 1$, define $Y_n = \min\{m - \min(Z_m, W_m) : m \geq n\}$. A simple union bound gives, for arbitrary k:

$$\mathbb{P}(Y_n \leq k) \leq \sum_{m\geq n}\mathbb{P}(\min(Z_m, W_m) > m - k).$$

Since $\mathbb{E}\min(Z,W) < \infty$, the sum above is finite and goes to 0 as n goes to infinity. Hence, $Y_n \to \infty$ in probability.

Let us fix $n \geq 0$ and apply Corollary 9 to the model obtained by shifting by n the set of integers so that leaves are indexed by the set $\{k : k \leq n\}$ and the edges are given by the sequences of random variables $(Z_k)_{k\geq n+1}$ and $(W_k)_{k\geq n+1}$. This gives us a random variable $N_n > n$, finite almost surely, such that $N_n \in S_k$ for all $k \geq N_n$. Let $I_n \leq n\}$ be the root of this tree and note that

$$Y_n \leq I_n \leq n < N_n \,.$$

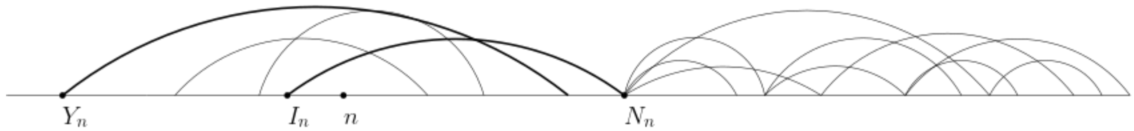

Figure 2: Illustration of the proof of Theorem 6

Since $Y_n \to \infty$ in probability, $I_n \to \infty$ in probability as well. Moreover, for all $n \geq 1$, one can choose N_n such that $N_n - n$ is distributed as N_0. The fact that all vertices $k \geq N_n$ belong to the tree rooted at I_n in this single-choice model implies that $I_n \in T_k$ for all $k \geq N_n$, and in particular $L_{I_n} \leq L_k$ for all $k \geq N_n$. From Proposition 11 we know that $L_n \to \infty$ in probability. Hence, $L_{I_n} \to \infty$ in probability. Moreover, for any fixed $m \geq 1$ we have

$$\begin{aligned}\mathbb{P}(\cup_{k\geq 2n}[L_k \leq m]) &\leq \mathbb{P}(N_n > 2n) + \mathbb{P}\Big(N_n \leq 2n, L_{I_n} \leq m\Big) \\ &\leq \mathbb{P}(N_0 > n) + \mathbb{P}\Big(L_{I_n} \leq m\Big)\ .\end{aligned}$$

Thus, for any $m \geq 1$, $\mathbb{P}(\cup_{k\geq 2n}[L_k \leq m]) \to 0$ as $n \to \infty$. By continuity of measure, this implies that for any $m \geq 1$,

$$\mathbb{P}(\cap_{n\geq 1} \cup_{k\geq 2n} [L_k \leq m]) = 0.$$

By the union bound,

$$\mathbb{P}(\cup_{m\geq 1} \cap_{n\geq 1} \cup_{k\geq 2n} [L_k \leq m]) = 0\ ,$$

which implies that $L_n \to \infty$ almost surely. ■

4.2 Heavy tails

In this section we study the heavy-tailed case, that is, $\mathbb{E}\min(Z,W) = \infty$. From Theorem 3 we already know that L_n cannot go to infinity with probability one. Here we show that, in contrast to this, $L_n \to \infty$ in probability.

Lemma 14. *For any integers $k \in \mathbb{Z}$ and $n \geq \max(1,k)$,*

$$\mathbb{P}\Big(k \in T_n^Z\Big) \leq r_{n-k}\ .$$

Proof. Clearly, we have $\mathbb{P}\Big(k \in T_n^Z\Big) = r_{n-k}$ whenever $k \geq 0$. Moreover, r_n satisfies the recursion

$$r_n = \sum_{i=1}^{n} q_i r_{n-i}\ .$$

Thus, for any $k > 0$, we have

$$\begin{aligned}\mathbb{P}\left(-k \in T_n^Z\right) &= \sum_{i=1}^{n} \mathbb{P}(Z_i = i + k)\mathbb{P}(i \in T_n) \\ &= \sum_{i=1}^{n} q_{k+i} r_{n-i} \\ &= \sum_{i=k+1}^{n+k} q_i r_{n+k-i} \\ &\leq \sum_{i=1}^{n+k} q_i r_{n+k-i} = r_{n+k} \ .\end{aligned}$$

■

Proof. [Proof of Theorem 7] Fix a constant $w > 0$ and let I_n^w denote the number of intervals of the form $(kw, (k+1)w]$, with $k \in \mathbb{Z}_{\geq 0}$, that are intersected by the chain T_n^Z. Note that, for any $x > 0$, if $I_n^w < x$ then $T_n^Z < wx$, thus

$$\begin{aligned}\mathbb{P}(I_n^w < x) &\leq \mathbb{P}\left(T_n^Z < wx\right) \\ &\leq \mathbb{P}\left(Z_1 + \cdots + Z_{\lfloor wx \rfloor} > n\right) \to 0 \ . \end{aligned} \tag{4.16}$$

Let $\ell \geq 1$ be an arbitrary integer and observe that, if the event $I_n^w \geq 2\ell$ occurs, then one can define a random set of vertices A_n such that

1. A_n is independent of $(W_k)_{k\geq 1}$,
2. $A_n \subset T_n^Z$,
3. $|A_n| = \ell$,
4. for all $k \neq m \in A_n$, $|k - m| \geq w$.

Conditionally on $I_n^w \geq 2\ell$, we have

$$\begin{aligned}&\mathbb{P}\left(\cup_{k\neq m, k\in A_n} T_k^W \cap T_m^W \neq \emptyset\right) \\ &\leq \mathbb{E} \sum_{\substack{k,m\in A_n \\ s.t.\ k\neq m}} \mathbb{1}_{T_k^W \cap T_m^W \neq \emptyset} \\ &\leq \mathbb{E} \sum_{\substack{k,m\in A_n \\ s.t.\ k\neq m}} \mathbb{P}\left(T_k^W \cap T_m^W \neq \emptyset\right)\end{aligned} \tag{4.17}$$

$$\leq \binom{\ell}{2} \max_{\substack{k,m\leq n \\ s.t.\ |k-m|\geq w}} \mathbb{P}\Big(T_k^W \cap T_m^W \neq \emptyset\Big) \tag{4.18}$$

by using the first moment method and the independence of A_n with respect to $(W_k)_{k\geq 1}$. Let $k < m$ be a pair of integers such that $|k-m| \geq w$ and observe that

$$\mathbb{P}\Big(T_k^W \cap T_m^W \neq \emptyset\Big) = \mathbb{P}\Big(T_k^W \cap T_m^Z \neq \emptyset\Big) \ .$$

Thus, the union bound, Lemma 14 and the Cauchy-Schwarz inequality yield

$$\begin{aligned}
\mathbb{P}\Big(T_k^W \cap T_m^W \neq \emptyset\Big) &\leq \sum_{i=-\infty}^{k} \mathbb{P}\Big(i \in T_k^W\Big)\mathbb{P}\Big(i \in T_m^Z\Big) \\
&\leq \sum_{i=-\infty}^{k} r_{k-i} r_{m-i} \\
&\leq \sqrt{\sum_{i=0}^{\infty} r_i^2} \times \sqrt{\sum_{i=m-k}^{\infty} r_i^2}.
\end{aligned} \tag{4.19}$$

From (4.18) and (4.19), it follows that

$$\mathbb{P}\Big(\cup_{k\neq m, k\in A_n} T_k^W \cap T_m^W \neq \emptyset\Big) \leq \binom{\ell}{2} \sqrt{\sum_{i=0}^{\infty} r_i^2} \times \sqrt{\sum_{i=w}^{\infty} r_i^2}. \tag{4.20}$$

Fix $\epsilon > 0$. From (4.20) we know that we can choose w big enough so we have for $n \geq 1$,

$$\mathbb{P}\Big(\cup_{k\neq m, k\in A_n} T_k^W \cap T_m^W \neq \emptyset\Big) \leq \epsilon. \tag{4.21}$$

Finally, observe that if $L_n < \ell$ and $I_n^w \geq 2\ell$, then there have to be at least two distinct vertices k, m in A_n such that the chains T_k^W and T_m^W meet each other: otherwise, each chain T_k^W with $k \in A_n$ would lead to a distinct leaf in T_n, contradicting the event $L_n < \ell$. Thus, by (4.20) we have

$$\mathbb{P}(L_n < \ell) \leq \mathbb{P}(I_n^w < 2\ell) + \epsilon \ .$$

From (4.16) it follows that for any $\epsilon > 0$ and any integer $\ell \geq 1$

$$\limsup_{n\to\infty} \mathbb{P}(L_n < \ell) \leq \epsilon \ ,$$

which concludes the proof of Theorem 7. ∎

5 Conclusion

We introduced and studied a simple mathematical model for recommendation systems based on giving each user two random choices, which include a mixture of past choices and untried options. In this setup, we identify three regimes based on the tail behavior of the sizes of the jumps in the past. A natural generalization of the model allows $k > 2$ choices, left for further study. We anticipate that in that case there are $k + 1$ regimes, but even the case $k = 3$ seems interesting. The eagle-eyed reader must have noticed that the bottom right corner of the table summarizing the results is left unfilled. Indeed, the question of the tightness of the rightmost leaf in the extremely heavy-tailed case remains an open problem.

References

[1] Baccelli, F., Haji-Mirsadeghi, M.-O., and Khaniha, S. (2022). Coupling from the past for the null recurrent markov chain. *arXiv preprint arXiv:2203.13585*.

[2] Baccelli, F., Haji-Mirsadeghi, M.-O., and Khezeli, A. (2018). Eternal family trees and dynamics on unimodular random graphs. *Unimodularity in randomly generated graphs*, 719:85–127.

[3] Baccelli, F. and Sodre, A. (2019). Renewal processes, population dynamics, and unimodular trees. *Journal of Applied Probability*, 56(2):339–357.

[4] Blath, J., González Casanova, A., Kurt, N., and Spano, D. (2013). The ancestral process of long-range seed bank models. *Journal of Applied Probability*, 50(3):741–759.

[5] Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, Oxford.

[6] Broutin, N., Devroye, L., Lugosi, G., and Oliveira, R. I. (2024). Subtractive random forests. *ALEA Latin American Journal of Probability and Mathematical Statistics*, 21:575–591.

[7] Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, 23(4):493–507.

[8] Chierichetti, F., Kumar, R., and Tomkins, A. (2020). Asymptotic behavior of sequence models. In *Proceedings of The Web Conference 2020*, pages 2824–2830.

[9] Erdős, P., Feller, W., and Pollard, H. (1949). A theorem on power series. *Bulletin of the American Mathematical Society*, 55:201–204.

[10] Hammond, A. and Sheffield, S. (2013). Power law pólya's urn and fractional brownian motion. *Probability Theory and Related Fields*, 157(3-4):691–719.

[11] Ibragimov, I. and Linnik, Y. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff Publishing, Groningen.

[12] Igelbrink, J. L. and Wakolbinger, A. (2022). Asymptotic gaussianity via coalescence probabilites in the Hammond-Sheffield urn. *arXiv e-prints*, pages arXiv–2201.

[13] Malevich, N. (2015). *Approximations and Asymptotic Expansions for Sums of Heavy-tailed Random Variables*.

[14] Zhang, S., Yao, L., Sun, A., and Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys*, 52(1):1–38.

[15] Zolotarev, V. (1986). *One-dimensional Stable Distributions*. American Mathematical Society, Providence.