# MIT Open Access Articles

## Power law Polya's urn and fractional Brownian motion

**Massachusetts Institute of Technology**

# Power law Pólya's urn and fractional Brownian motion

Alan Hammond[*] and Scott Sheffield[†]

**Abstract**

We introduce a natural family of random walks $S_n$ on $\mathbb{Z}$ that scale to fractional Brownian motion. The increments $X_n := S_n - S_{n-1} \in \{\pm 1\}$ have the property that given $\{X_k : k < n\}$, the conditional law of $X_n$ is that of $X_{n-k_n}$, where $k_n$ is sampled independently from a fixed law $\mu$ on the positive integers. When $\mu$ has a roughly power law decay (precisely, when $\mu$ lies in the domain of attraction of an $\alpha$-stable subordinator, for $0 < \alpha < 1/2$) the walks scale to fractional Brownian motion with Hurst parameter $\alpha + 1/2$. The walks are easy to simulate and their increments satisfy an FKG inequality. In a sense we describe, they are the natural "fractional" analogues of simple random walk on $\mathbb{Z}$.

## 1  Introduction

Fractional Brownian motion is a one-parameter family of stochastic processes, mapping the real line to itself, that are the only stationary-increment Gaussian processes that, for some fixed $H > 0$, are invariant under the space-time rescalings $S \to Y$ of the form $Y_t = c^{-H} S_{ct}$ with $c > 0$. The parameter $H$ is called the Hurst parameter, and may take any value in $(0, 1)$. Fractional Brownian motion $S^H : (0, \infty) \to \mathbb{R}$ with Hurst parameter $H$ satisfies

$$\mathbb{E}(S_s^H S_t^H) = \frac{1}{2}\big(|t|^{2H} + |s|^{2H} - |t - s|^{2H}\big),$$

or, equivalently,

$$\mathbb{E}(|S_t^H - S_s^H|^2) = |t - s|^{2H}, \quad S_0^H = 0.$$

The process satisfies $\mathbb{E}\big(S_t^H\big) = 0$ for all $t \in \mathbb{R}$. When $H = 1/2$, it is Brownian motion, whose increments are of course independent. If $H > 1/2$, increments over two given disjoint intervals are positively correlated, while they have negative correlation if $H < 1/2$.

Fractional Brownian motion was first considered by Kolmogorov [16] as a model of turbulence, and there is a now large literature treating this family of processes, e.g. as models in mathematical finance, and developing their stochastic calculus [19]. The stated characterization of fractional Brownian motion may permit this family of processes to be considered as at least slightly canonical. However, there are only a few examples of members of the family arising as a scaling or continuum limit of a discrete model (other than Brownian motion itself). The fractional Brownian motion with Hurst parameter $1/4$ arises as a scaling limit for the tagged particle in a one-dimensional symmetric exclusion process [2]. For general values of $H \in (0, 1)$, fractional Brownian motion has been exhibited as a scaling limit of an average of a mixture of independent random walks, each walk having a decay rate for the correlation of its increments, which rate is selected independently from a law that depends on $H$ [11]. In this paper, we present a simple discrete random walk that scales to fractional Brownian motion. The process may be considered to be a discrete counterpart to fractional Brownian motion.

In finance applications, as a model for the drift adjusted logarithm of an asset price, fractional Brownian motion retains much of the simplicity of ordinary Brownian motion (stationarity, continuity, Gaussian increments) but dispenses with independence of increments, thereby allowing for "momentum effects" (i.e., increment positive correlations), which have been observed empirically in some markets. One hypothesis is that momentum effects result from market inefficiencies associated with insider trading; if an event occurs of which the market is unaware, insider trading may cause the asset price to change gradually, over a period of time, instead of all at once — initially because of trades by individuals with privileged knowledge, later by a larger number of market participants. These ideas are discussed further in [13], where a model based on shocks of this form is shown to be arbitrage free and and to have fractional Brownian motion as a scaling limit. (We will comment further on the possibility of arbitrage for fractional Brownian motion in Section 1.2.) See also [3], which uses "investor inertia" to explain models in which drift-adjusted logarithmic price is a stochastic integral of fractional Brownian motion. Momentum effects appear naturally in the random walks we introduce: i.e., it will be easy to see why an "event" (the sampling of an increment of the walk at one point in time) has an influence on the expectation of future increments.

An informal description of the walk is as follows. Let $\mu$ be a given law on the natural numbers (which we take to exclude zero). The walk associated with law $\mu$ has increments, each of which is either $-1$ or $1$. Independent samples $k_n$ of $\mu$ are attached to the vertices of $\mathbb{Z}$. The sequence of increments $\{X_k : k \in \mathbb{N}\}$ is such that, given the values $\{X_k : k < n\}$, $X_n$ is set equal to the increment $X_{n-k_n}$ obtained by looking back $k_n$ steps in the sequence. The walk is then defined by adding the successive increments from some fixed number. We remark that the notion of determining the value $X_n$ of a process at time $n$ by looking a random number of steps into the past also appears in some of the urn models and reinforced random walk models studied and surveyed by Pemantle in [20]. Recall that in the traditional *Pólya's*

*urn* process, the value of $X_n$ represents the color of the $n$th ball added to an urn, and the conditional law of $X_n$ given the past is that of a uniform sample from $\{X_1, X_2, \ldots, X_{n-1}\}$. Our $X_n$ differs from the Pólya's urn process in that the past is infinite and the uniform sample is replaced by one having a power law distance in time from the present.

This description of the walk is a rough one, because it would require some initial condition to construct it. We will turn to a Gibbs measure formulation for a precise mathematical description. Fixing the law $\mu$, we write $G_\mu$ for the random directed spanning graph on $\mathbb{Z}$ in which each vertex $z$ has a unique outward pointing edge pointing to $z - k_z$, where $k_z$ is sampled independently according to the measure $\mu$. We call $z - k_z$ the parent of $z$. The ancestral line of $z$ is the decreasing sequence whose first element is $z$ and each of whose terms is the parent of the previous one.

A law $\lambda$ on functions mapping $\mathbb{Z}$ to $\{-1, 1\}$ is said to be a $\mu$-Gibbs measure provided that, for any half-infinite interval $R = \{x, \ldots, \infty\}$, the conditional law of $\lambda$ on $R$, given its values in $R^c$, is given by sampling $G_\mu$ and assigning to any element of $R$ the value assigned to its most recent ancestor in $R^c$. We will write $X = X_\lambda : \mathbb{N} \to \{-1, 1\}$ for a realization of the law $\lambda$. (In general, a shift invariant probability measure on $\{-1, 1\}^{\mathbb{Z}}$ that has a specified conditional law for $X_0$ given $\{X_k : k < 0\}$ is called a $g$-**measure** for that specification [7, 14, 15]. One may view $\mu$-Gibbs measures as $g$-measures for a particular specification. We will not make use of this more general framework here.)

A $\mu$-Gibbs measure is called **extremal** if it cannot be written as an average of two distinct $\mu$-Gibbs measures. There are at least two extremal $\mu$-Gibbs measures, whatever the choice of $\mu$: those that assign unit mass to either the constant function equal to $+1$ or to $-1$. If $\mu$ is such that the great common denominator of values in its support is $1$, then the number of components in $G_\mu$ is either equal almost surely to one or to $\infty$. (We furnish a proof of this assertion in Lemma 2.1.) If $G_\mu$ has infinitely many components, then a one-parameter family $\{\lambda_p : p \in [0, 1]\}$ of extremal Gibbs measures may be defined as follows. To sample from $\lambda_p$, first sample $G_\mu$ and then independently give each component of $G_\mu$ a value of 1 (with probability $p$) or $-1$ (with probability $1 - p$). Then assign all of the vertices in that component the corresponding value.

Perhaps surprisingly, for such laws $\mu$, any Gibbs measure is a mixture of these ones:

**Proposition 1** *Let $\mu$ denote a probability measure on $\mathbb{N} = \{1, 2, \ldots\}$ the greatest common denominator of whose support is equal to one. If $G_\mu$ has one component almost surely, then there are no extremal $\mu$-Gibbs measures other than the two trivial ones, $\lambda_0$ and $\lambda_1$. If $G_\mu$ has infinitely many components almost surely, then the space of extremal Gibbs measures is the family $\{\lambda_p : p \in [0, 1]\}$.*

From this result, it is not hard to conclude that if the greatest common denominator of the support of $\mu$ is some $k \neq 1$, and $\lambda$ is an extremal $\mu$-Gibbs measure, then the restriction of $\lambda$ to points in $k\mathbb{Z} + a$, for each $a \in \{0, 1, \ldots, k - 1\}$, will be an extremal $\mu$-Gibbs measure of the type described in Proposition 1, and that these measures will be independent for

3

different values of $a$ (though $p$ may depend on $a$). Hence, there is no real loss of generality in restricting to the case that the greatest common denominator is $1$, as we do in Proposition 1 (and throughout most of the remainder of this paper).

Next, we define the measures $\mu$ that we will use for most of this paper.

**Definition 1.1** *Let $\mu$ denote a probability measure on $\mathbb{N}$. For $\alpha \in (0, \infty)$, we say that $\mu \in \Gamma_\alpha$ if there exists a slowly varying function $L : (0, \infty) \to (0, \infty)$ for which*

$$\mu\{n, \dots, \infty\} = n^{-\alpha} L(n) \tag{1.1}$$

*for each $n \in \mathbb{N}$. Recall that by slowly varying is meant*

$$\lim_{u \to \infty} \frac{L(u(1 + r))}{L(u)} = 1, \tag{1.2}$$

*for any $r > 0$.*

Note that if we required $L$ to be a constant function, then the measures satisfying the first condition would be simply those for which $\mu\{n, \dots, \infty\}$ is a constant times $n^{-\alpha}$. The generalization to slowly varying $L$ is quite natural, for the following reason. Let $R_\mu$ denote the random set $\left\{ \sum_{i=1}^{j} X_i : j \in \mathbb{N} \right\}$ of values assumed by partial sums of an independent sequence of samples $X_i$ of the law $\mu$. (Clearly, $-1$ times the ancestral line of $0$ has the same law as $R_\mu$.) Then it turns out that when $\alpha \in (0, 1)$, the random set $\epsilon R_\mu$ converges in law as $\epsilon \to 0$ to the range of a stable subordinator with parameter $\alpha$ *if and only if* $\mu \in \Gamma_\alpha$ (Theorem 8.3.1 of [6]). Indeed, there is a sizable literature on probability distributions with power law decays up to a slowly varying function [6]. However, throughout this paper, when a result is stated for all $\mu \in \Gamma_\alpha$, the reader may find it easier on a first reading to focus on the special case that $L$ is constant.

Our next result relates the decay rate of the tail of $\mu$ to the number of components in $G_\mu$.

**Proposition 2** *Let $\mu \in \Gamma_\alpha$ for some $\alpha \in (0, \infty)$. If $\alpha > 1/2$, then $G_\mu$ almost surely has one component, while, if $\alpha < 1/2$, then $G_\mu$ almost surely has infinitely many.*

To any $\mu$-Gibbs measure $\lambda$, we associate a random walk $S = S_\lambda : \mathbb{Z} \to \mathbb{Z}$ by setting $S(0) = 0$ and $S(n) = \sum_{i=1}^{n} X(i)$ for $n > 0$, and $S(n) = -\sum_{i=-n}^{-1} X(i)$ for $n < 0$. We extend the domain of $S$ to $\mathbb{R}$ by linearly interpolating its values between successive integers.

We now ready to state our main result. Essentially, it says that, for $S$ the random walk associated to an extremal $\mu$-Gibbs measure for a choice of $\mu \in \Gamma_\alpha$, the time-scaled process $S(nt)$, further rescaled by subtracting its mean and multiplying by a deterministic $n$-dependent factor, converges to fractional Brownian motion with Hurst parameter $\alpha + 1/2$. This normalizing factor is written below as $\tilde{c} n^{-\frac{1}{2} - \alpha} L(n)$. The explicit form for the constant $\tilde{c}$ will be explained by the proof of Lemma 3.1, in which the asymptotic variance of $S_n$ is determined.

4

**Theorem 1.1** *For $\alpha \in (0, 1/2)$, let $\mu \in \Gamma_\alpha$. Let $L : (0, \infty) \to (0, \infty)$ be given by (1.1). Define $\tilde{c} > 0$ by means of*

$$\tilde{c}^2 = \frac{\sum_{i=0}^\infty q_i^2}{2p(1-p)} \alpha \big(2\alpha + 1\big) \Gamma\big(1 - 2\alpha\big)^2 \Gamma\big(2\alpha\big) \cos\big(\pi\alpha\big), \qquad (1.3)$$

*where $q_i = \mathbb{P}\big(i \in R_\mu\big)$ for $i \geq 1$, and $q_0 = 1$. Then, for each $p \in (0, 1)$, there exists a sequence of couplings $\mathrm{C}_n$ of the process*

$$S_p^n : (0, \infty) \to \mathbb{R} : t \to \tilde{c} n^{-\frac{1}{2} - \alpha} L(n) \Big( S_{\lambda_p}\big(nt\big) - n(2p - 1)t \Big)$$

*and fractional Brownian motion $S_{\alpha+1/2}$ with Hurst parameter $\alpha + 1/2$ such that, for each $T > 0$ and $\epsilon > 0$,*

$$\lim_{n \to \infty} \mathrm{C}_n \Big( ||S_p^n - S_{\alpha+1/2}||_{L_\infty\big([0,T]\big)} > \epsilon \Big) = 0. \qquad (1.4)$$

## 1.1 Discussion regarding Theorem 1.1

In light of Theorem 1.1, it is tempting to argue that the random walks we construct are in some sense *the* canonical fractional analogs of simple random walk on $\mathbb{Z}$. To make this point, we note that any random walk on $\mathbb{Z}$ (viewed as a graph) has increments in $\{-1, 1\}$, which means that its law, conditioned on $\{X_n : n < M\}$ for some fixed $M$, is determined by the value of $E[X_n | \{X_k : k < n\}]$. If we assume stationarity of increments and we further posit that $E[X_0 | \{X_k : k < 0\}]$ has a simple form — say, that it is a monotonically increasing linear function of $\{X_k : k < 0\}$ — then we have

$$E[X_n | \{X_k : k < n\}] = \sum_{i=1}^\infty p_i X_{n-i}$$

for some $p_i \geq 0$ with $\sum p_i \leq 1$. It is not hard to show that if $\sum p_i < 1$, the process scales to ordinary Brownian motion. We are therefore left with the case $\sum p_i = 1$, which corresponds to the walks we consider with $\mu(\{i\}) = p_i$.

It seems plausible that any $\mu$ for which the conclusion of the theorem holds (with the normalizing factor $\tilde{c} n^{-\frac{1}{2} - \alpha} L(n)$ replaced by *some* deterministic function of $n$) must be a member of $\Gamma_\alpha$. Similarly, it seems highly plausible (in light of (3.19)) that the closely related assertion that the variance of $S_n$ is $n^{2\alpha+1}$ (multiplied by a slowly varying function) implies $\mu \in \Gamma_\alpha$. We will not prove either of these statements here.

We see that the model undergoes a phase transition at the value $\alpha = \frac{1}{2}$. We remark that, while there are no non-trivial $\mu$-Gibbs measures for $\mu \in \Gamma_\alpha$ with $\alpha > 1/2$, there is nonetheless a further phase transition at the value $\alpha = 1$, which is the maximal value for which an element $\mu \in \Gamma_\alpha$ may have infinite mean. Indeed, suppose that we define the

walk associated to a measure $\mu$ by instead specifying its domain to be the positive integers and then making the following adjustment to the existing definition. As previously, we take $\{k_n : n \in \mathbb{N}\}$ to be a sequence of independent samples of the measure $\mu$, and, in the case that $n - k_n \geq 0$, we continue to set the increment $X_n$ to be equal to $X_{n-k_n}$. In the other case, we choose this increment to be $+1$ or $-1$ with equal probability, independently of previous such choices. Then it is easily seen that the walk takes infinitely many steps of each type if and only if $\mu$ has infinite mean.

Finally, we mention that the decomposition of $\mathbb{Z}$ into components provided by Proposition 2 with a choice of $\mu \in \Gamma_\alpha$ with $\alpha < 1/2$ has something in common with the following process, discussed in [1]. A system of particles, one at each element of $\mathbb{Z}$, are labelled, each by its location at an initial time $t = 0$. Each pair of adjacent particles consider swapping locations at an independent Poisson sequence of times, but do so only if the higher-labelled particle lies on the right-hand-side in the pair just before the proposed swap. Each particle behaves as a second class particle in a totally asymmetric exclusion process, begun from an initial condition in which only sites to its left are occupied; as such, each has an asymptotic speed, which is a random variable having the uniform law on $[0, 1]$. Defining a convoy to be the collection of locations inhabited at time zero by particles that share a common speed, ([1]) shows that the convoy containing the origin is almost surely infinite, has zero density, and, conditionally on the speed of the particle beginning at the origin, is a renewal process. As such, the collection of convoys provides a partition of $\mathbb{Z}$ that has some similarities with $G_\mu$, when it has infinitely many components.

## 1.2 Further comments

In financial contexts, where fractional Brownian motion may serve as a model of evolving prices or preference strengths, it is particularly natural to aim to simulate its future trajectory given some fixed past. The discrete model that we introduce is very well suited to this problem, because, given the past, $N$ subsequent steps of the process may be sampled in time $O(N)$ (neglecting logarithmic corrections arising due to random number generation). In comparison, fast Fourier transform methods for approximately sampling fractional Brownian motion necessarily take at least $O(N \log N)$ time.

Mandelbrot and Van Ness introduced a representation of fractional Brownian as a weighted average of white noise. The representation of fractional Brownian motion stated in footnote 3 on page 424 of [18] inspires has a natural discrete counterpart, in which a weighted avarge of independent $\pm 1$ coin tosses is used. However, the walks obtained in this way do not have $\pm 1$ increments and are not as simple to generate as those of this paper. In regard to discretizations of fractional Brownian motion, we mention [22], that introduces a random walk approximation and shows its weak convergence to fractional Brownian motion.

The question of whether (and what kind of) arbitrage is admitted in financial models based on fractional Brownian motion or geometric fractional Brownian motion has been

explored at some length in the finance literature. Fractional Brownian motion is not a semi-martingale: if the class of admissible trading strategies is large enough (and allows a potentially unbounded number of trades at arbitrarily small intervals) then one can indeed construct strategies that that admit positive return with probability one [21]. Even if such strategies involve trading that is too frequent to be practical, their existence complicates the construction of consistent derivative pricing models. A good deal of literature has addressed ways in which transaction costs, other strategy restrictions, or minor modifications to the model can be imposed to make the arbitrage opportunites go away. See [4] for a survey of this literature that clarifies how the presence or absence of arbitrage depends on the precise choice of admissible trading strategies. We note that in a simple discrete model such as ours — where the walk goes up or down by a unit increment at every discrete time step and all possible finite-length trajectories have positive probability — one cannot have arbitrage in the strong sense, so these particular subtleties are not relevant to us. It remains interesting, however, to think about the market efficiency implications of our model, and we present some preliminary thoughts on this subject in the appendix.

## 1.3 Structure of the paper

Section 2 is devoted to the analysis of $\mu$-Gibbs measures, and the proofs of Propositions 1 and 2 are presented there. In section 3, we turn to the convergence of the discrete process to fractional Brownian motion, proving Proposition 3, which shows convergence in finite dimensional distributions of the rescaled walk to the continuous process. A useful tool here is an explicit asymptotic formula for the variance of the walk, which is presented in Lemma 3.1. In Section 4, we prove a correlation inequality applicable to any extremal Gibbs measures in Proposition 4, and then apply it to improve the topology of convergence to yield the $L^\infty$-small coupling given in Theorem 1.1.

The discrete processes that we introduce form a natural counterpart to fractional Brownian motion, and we are hopeful that they may serve as a tool for the analysis of the continuous process. They also appear to have various discrete relatives that are candidates for further enquiry. We take the opportunity to discuss this by presenting several open problems, in Section 5. We end with an appendix that describes potential economic and financial applications of the model.

# 2 The space of Gibbs measures

We now prove the first two propositions. We begin with two simple lemmas.

**Lemma 2.1** *Let $\mu$ be a measure on the positive integers, the greatest common denominator of its support being equal to one. Then $G_\mu$ has either one component almost surely, or infinitely many almost surely.*

**Proof.** We write $A_i$ for the ancestral line of $i \in \mathbb{Z}$. Let $m_n = \mathbb{P}\big(A_0 \cap A_n \neq \emptyset\big)$. We now argue that if $\inf_n m_n > 0$, then $G_\mu$ has one component almost surely. If $m := \inf_n m_n > 0$, we may find $g_n \in \mathbb{N}$ such that the ancestral lines from $0$ and $n$ meet at some integer exceeding $-g_n$ with probability at least $m/2$. If the ancestral lines $A_0$ and $A_n$ fail to meet in $\{-g_n, \dots\}$, then we sample them until each first passes $-g_n$, at $f_1$ and $f_2$, say. Without loss of generality, $f_1 < f_2$. Conditional on $A_0 \cap A_n \cap \{-g_n, \dots\} = \emptyset$, and on $f_1$ and $f_2$, the probability that $A_0 \cap A_n \cap \{f_1 - g_{f_2 - f_1}, \dots, f_1\} \neq \emptyset$ is at least $m/2$. Iterating, we construct a sequence of intervals each of which has a probability at least $m/2$ of containing a point in $A_0 \cap A_n$, conditionally on its precedessors not doing so. We see that the ancestral lines indeed meet almost surely.

If $m_n \to 0$ subsequentially, then we will show that there exists a sequence $\{n_i : i \in \mathbb{N}\}$ such that

$$\mathbb{P}\Big(A_{n_i} \cap \bigcup_{j < i} A_{n_j} = \emptyset\Big) \geq 1 - i^{-2}. \tag{2.5}$$

This suffices to show that $G_\mu$ has infinitely many components almost surely, by the Borel-Cantelli lemma. To construct the sequence, for $i \in \mathbb{N}$, set $q_i = \mathbb{P}\big(0 \in A_i\big)$. Then, whenever $n > l > 0$,

$$\mathbb{P}\Big(A_0 \cap A_n \neq \emptyset\Big) \geq \mathbb{P}\Big(\big\{A_0 \cap A_n \neq \emptyset\big\} \cap \big\{n - l \in A_n\big\}\Big)$$
$$= \mathbb{P}\Big(A_0 \cap A_n \neq \emptyset \big| n - l \in A_n\Big)\mathbb{P}\Big(n - l \in A_n\Big) = \mathbb{P}\Big(A_0 \cap A_{n-l} \neq \emptyset\Big)\mathbb{P}\Big(n - l \in A_n\Big)$$

so that

$$m_n \geq m_{n-l} q_l. \tag{2.6}$$

It follows readily from g.c.d.supp$(\mu) = 1$ that $q_j > 0$ for all sufficiently high $j \in \mathbb{N}$. Supposing that we have constructed an increasing sequence $\{n_1, \dots, n_{i-1}\}$ for which $q_{n_j} > 0$ for $j < i$, choose $n_i > n_{i-1}$ satisfying $q_{n_i} > 0$ and $m_{n_i} \sum_{j=1}^{i-1} q_{n_j}^{-1} \leq i^{-2}$. Note then that

$$\mathbb{P}\Big(A_{n_i} \cap \bigcup_{j<i} A_{n_j} \neq \emptyset\Big) \leq \sum_{j=1}^{i-1}\mathbb{P}\Big(A_{n_i} \cap A_{n_j} \neq \emptyset\Big) \leq m_{n_i}\sum_{j=1}^{i-1} q_{n_j}^{-1} \leq i^{-2},$$

the second inequality by (2.6). In this way, we construct a sequence satisfying (2.5). $\square$

**Proof of Proposition 1.** The case that $G_\mu$ has one component almost surely is trivial.

Let $\mu$ be a probability measure on $\mathbb{N}$ whose support has greatest common denominator one, and for which $G_\mu$ has infinitely many components almost surely. Let $\lambda$ denote a $\mu$-Gibbs measure. For $m \in \mathbb{Z}$, set $\sigma_m = \sigma\{X_i : i < m\}$. By the backwards martingale

8

convergence theorem (Section XI.15 of [8]),

$$p_{n,-\infty} = \lim_{m \to -\infty} \lambda\big(X_n = 1 \big| \sigma_m\big) \in \sigma_{-\infty} := \bigcap_{m<0} \sigma_m \qquad (2.7)$$

exists $\lambda$-a.s. Note that it is an almost sure constant, because $\lambda$ is extremal.

We will now argue that

$$p_{n,-\infty} = p_{m,-\infty} \qquad \lambda\text{-a.s.} \qquad (2.8)$$

We will denote the common value by $p$.

To show (2.8), let $a, b \in \text{supp}(\mu)$, $a \neq b$. Consider firstly the case that $|b - a|$ divides $|n - m|$. For $k \in \mathbb{N}$, we enumerate the ancestral line $A_k = \{k = x_0(k), x_1(k), \dots\}$ emanating from $k$ in decreasing order. We define a coupling $\Theta$ of $A_n$ and $A_m$. At any given step $l \in \mathbb{N}$, the initial sequences $\{x_0(n), \dots, x_l(n)\}$ and $\{x_0(m), \dots, x_l(m)\}$ have been formed. If $x_l(m) \neq x_l(n)$, then we set $x_{l+1}(m) - x_l(m) = -q$, where $q$ has law $\mu$. If $q \notin \{a, b\}$, then we take $x_{l+1}(n) - x_l(n) = -q$ also. If $q \in \{a, b\}$, then we take $x_{l+1}(n) - x_l(n)$ to have the conditional distribution of a sample of $\mu$ given that its value belongs to $\{a, b\}$. If $x_l(m) = x_l(n)$, we simply take $x_{l+1}(m) - x_l(m) = x_{l+1}(n) - x_l(n) = -q$.

The sequence of differences $\{x_l(m) - x_l(n) : l \in \mathbb{N}\}$ performs a random walk with independent jumps, a jump equal to zero with probability $1 - 2\mu\{a\}\mu\{b\}$, and, otherwise, with equal probability to be $-|b-a|$ or $|b-a|$, until the walk reaches $0$. The walk beginning at a multiple of $|b - a|$, the sequence of differences reaches zero, and then remains there, almost surely, by the recurrence of one-dimensional simple random walk.

If $|b - a|$ does not divide $|n - m|$, we begin by sampling initial segments of the two ancestral lines $\{x_0(n), \dots, x_j(n)\}$ and $\{x_0(m), \dots, x_j(m)\}$, where $j$ is the stopping time

$$j = \inf\Big\{k \in \mathbb{N} : |b - a| \text{ divides } |x_k(n) - x_k(m)|\Big\},$$

which is almost surely finite, because $\text{g.c.d.supp}(\mu) = 1$. Indeed, this finiteness easily follows from the positivity of $q_j$ for all sufficiently high $j$. The coupling demonstrates that, given $\epsilon > 0$, there exists $l$ sufficiently negative that, $\sigma_l$-a.s.,

$$\left| \lambda\Big(X_n = 1 \Big| \sigma_l\Big) - \lambda\Big(X_m = 1 \Big| \sigma_l\Big) \right| < \epsilon,$$

from which (2.8) follows.

We learn then that, for any $n \in \mathbb{Z}$,

$$\lambda\big(X_n = 1\big) = \mathbb{E}_\lambda \mathbb{P}\big(X_n = 1 \big| \sigma_{-\infty}\big) = \mathbb{E}p_{n,-\infty} = p.$$

Let $L \subseteq \mathbb{N}$ be finite. Set

$$s_L = \lambda\Big(A_i, i \in L, \text{ are not pairwise disjoint}\Big).$$

9

For $m, l \in \mathbb{Z}$, $m < l$, let $v_{l,m} = \sup A_l \cap \{\dots, m\}$ denote the first element in the ancestral line of $l$ that is at most $m$. For $L \subseteq \mathbb{N}$ finite with $m < \inf L$, write $v_{L,m} = \{v_{l,m} : l \in L\}$ for the locations of ancestors of elements of $L$ that are at most $m$ whose child is at least $m + 1$.

For $\epsilon > 0$, set

$$g_{\epsilon,L} = \sup \left\{ m < \inf L : s_{v_{L,m}} < \epsilon \right\}.$$

Noting that $\{|v_{L,m}| : m < \inf L\}$ is a non-decreasing sequence that assumes a common value for all sufficiently high negative choices of $m$, we find that $g_{\epsilon,L} > -\infty$, $\lambda$-a.s.

For $\epsilon > 0$ and $j, k \in \mathbb{Z}$, $k < j$, write

$$\Omega_{k,j}^{(\epsilon)} = \left\{ \omega \in \{-1,1\}^{\{-\infty,\dots,k\}} : \lambda\left(X_j = 1 \big| (\dots, X_k) = \omega\right) \in (p - \epsilon, p + \epsilon) \right\}.$$

Note that, for all $\epsilon > 0$ and $j \in \mathbb{N}$,

$$\lim_{k \to -\infty} \lambda\left(\Omega_{k,j}^{(\epsilon)}\right) = 1, \tag{2.9}$$

by (2.7) and (2.8). Given $L \subseteq \mathbb{N}$ finite, set

$$j_{\epsilon,L} = \sup \left\{ k < \inf L : \lambda\left(\Omega_{k,l}^{(\epsilon)}\right) > 1 - \epsilon \text{ for each } l \in L \right\},$$

so that $j_{\epsilon,L} > -\infty$ for each $\epsilon > 0$ and finite $L \subseteq \mathbb{N}$, by (2.9).

Let $k \in \mathbb{Z}$ and $\omega \in \{-1,1\}^{\{-\infty,\dots,k\}}$. Define $\{Y_{k,l}^{(\omega)} : l \geq k + 1\}$ by constructing an independent collection of ancestral lines (that do not coalesce on meeting) from the elements of $\{l \in \mathbb{Z} : l \geq k + 1\}$, stopping each line on its arrival in $\{l \in \mathbb{Z} : l \leq k\}$. For each $l \geq k + 1$, we set $Y_{k,l}^{(\omega)}$ equal to the $\omega$-value of the vertex at the end of the stopped ancestral line emanating from $l$.

Let $q_{\epsilon,L} = j_{\epsilon,v_{L,g_{\epsilon,L}}}$. Set $\Omega_\epsilon^* = \bigcap_{l \in v_{L,g_{\epsilon,L}}} \Omega_{q_{\epsilon,L},l}^{(\epsilon)}$, and note that $\lambda(\Omega_\epsilon^*) \geq 1 - |L|\epsilon$.

Note further that, conditionally on $v_{L,g_{\epsilon,L}}$ and for any $\omega \in \{-1,1\}^{\{-\infty,\dots,q_{\epsilon,L}\}}$ for which

$$\lambda\left(X_l = 1 \big| (\dots, X_{q_{\epsilon,L}}) = \omega\right) \in (p - \epsilon, p + \epsilon)$$

for each $l \in v_{L,g_{\epsilon,L}}$, the set of values $\{Y_{q_{\epsilon,L},l} : l \in v_{L,g_{\epsilon,L}}\}$ are independent, each being equal to $+1$ with probability at least $p - \epsilon$ and at most $p + \epsilon$.

The labelling $\{X_i : i \in L\}$ is determined by the set of values $\{X_i : i \in v_{L,g_{\epsilon,L}}\}$. However, for each $\omega \in \{-1,1\}^{\{\dots,q_{\epsilon,L}\}}$ and for any finite set $\sigma \subseteq \mathbb{Z}$ for which $v_{L,g_{\epsilon,L}} = \sigma$ is possible,

$$\text{TV}\left(\left(\left(X_l : l \in v_{L,g_{\epsilon,L}}\right)\Big| v_{L,g_{\epsilon,L}} = \sigma, \{\dots, X_{q_{\epsilon,L}}\} = \omega\right), \left(Y_{j_{\epsilon,\sigma},l}^\omega : l \in \sigma\right)\right) \leq s_\sigma.$$

Note that the right-hand-side is at most $\epsilon$, because $v_{L,g_{\epsilon,L}} = \sigma$ is possible. By taking $\epsilon > 0$ arbitrarily small, we obtain the result. $\square$

10

**Definition 2.1** *Let $\mu$ denote a probability distribution on $\mathbb{N}$. Let $R_\mu$ denote the random set $\left\{ \sum_{i=1}^{j} X_i : j \in \mathbb{N} \right\}$ of values assumed by partial sums of an independent sequence of samples $X_i$ of the law $\mu$. When such a measure $\mu$ has been specified, we will write $p_n = \mu\{n\}$ and $q_n = \mathbb{P}(n \in R_\mu)$. We adopt the convention that $q_0 = 1$.*

**Lemma 2.2** *Let $\alpha \in (0,1)$ and let $\mu \in \Gamma_\alpha$.*

1. *We have that*
$$\sum_{n=0}^{\lfloor x \rfloor} q_n \sim \frac{1-\alpha}{\Gamma(2-\alpha)\Gamma(1+\alpha)} x^\alpha L(x)^{-1},$$
   *where $f \sim g$ means $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.*

2. *Moreover, the sum*
$$\sum_{n=0}^{\infty} q_n^2$$
   *converges if $\alpha \in (0, 1/2)$ and diverges if $\alpha \in (1/2, 1)$.*

**Proof.** Let $p, q : [0, \infty) \to [0, \infty]$, given by

$$p(\lambda) = \sum_{n=0}^{\infty} p_n \exp\left\{ -\lambda n \right\}$$

and

$$q(\lambda) = \sum_{n=0}^{\infty} q_n \exp\left\{ -\lambda n \right\},$$

denote the Laplace transforms of these two sequences. We will analyze these transforms, (and, in later arguments, the Fourier transforms) of such sequences by means of some Tauberian theorems. A Tauberian theorem asserts that a sequence (in our case) has a given rate of decay if and only if its Laplace or Fourier transform has a corresponding asymptotic behaviour close to zero. Additional hypotheses, such as regular variation of the sequence (and, correspondingly, the transform), are required.

We will make use of two basic results concerning slowly varying functions. Firstly, the uniform convergence theorem (Theorem 1.2.1 of [6]) states that, if $L : (0, \infty) \to (0, \infty)$ is slowly varying, then

$$\frac{L(\lambda x)}{L(x)} \to 1, \tag{2.10}$$

uniformly on each compact $\lambda$-set in $(0, \infty)$. Secondly, Potter's theorem (Theorem 1.5.6 of [6]) states in part that, for any slowly varying function $L : (0, \infty) \to (0, \infty)$ that is bounded

11

above and below on any compact interval in $(0, \infty)$, and for any $\epsilon > 0$, there exists $C = C_\epsilon$ such that, for all $x, y > 0$,

$$\frac{L(y)}{L(x)} \leq C \max \left\{ (y/x)^\epsilon, (y/x)^{-\epsilon} \right\}. \tag{2.11}$$

From $\{p_n : n \in \mathbb{N}\} \in l_1$, the function $p$ is differentiable on $(0, \infty)$, with

$$\frac{d}{d\lambda} p(\lambda) = -\sum_{n=0}^{\infty} n p_n \exp \left\{ -\lambda n \right\}. \tag{2.12}$$

Setting $U : (0, \infty) \to (0, \infty)$ according to $U(x) = \sum_{n=0}^{\lfloor x \rfloor} n p_n$, it follows from $U(x) = \sum_{j=1}^{\lfloor x \rfloor} \sum_{i=j}^{\lfloor x \rfloor} p_i$, (1.1), (2.10) and (2.11) that $U(x) \sim \frac{\alpha}{1-\alpha} x^{1-\alpha} L(x)$, $x \to \infty$. A special case of Karamata's Tauberian theorem (Theorem 1.7.1 of [6]) states that, if $\{a_n : n \in \mathbb{N}\}$ is a sequence of non-negative numbers, $l : (0, \infty) \to (0, \infty)$ is slowly varying, and $c, \rho \geq 0$, then the following are equivalent:

$$\sum_{n=1}^{\lfloor x \rfloor} a_n \sim \frac{c}{\Gamma(1+\rho)} x^\rho l(x), \qquad x \to \infty,$$

$$\sum_{n=1}^{\infty} a_n e^{-\lambda n} \sim c \lambda^{-\rho} l(\lambda^{-1}), \qquad \lambda \to 0^+.$$

Applying the theorem in the present case, we learn that

$$\sum_{n=1}^{\infty} n p_n e^{-\lambda n} \sim \frac{\alpha}{1-\alpha} \Gamma(2-\alpha) \lambda^{\alpha-1} L(1/\lambda), \qquad \lambda \to 0^+.$$

By (2.12), $1 - p(\lambda) \sim \frac{\Gamma(2-\alpha)}{1-\alpha} \lambda^\alpha L(1/\lambda)$, $\lambda \to 0^+$, since $\int_0^\lambda t^{\alpha-1} L(t^{-1}) dt \sim \alpha^{-1} \lambda^\alpha L(\lambda^{-1})$, by means of (2.10) and (2.11). Noting that the Laplace transforms of the two sequences are related by $q = (1-p)^{-1}$, we learn that

$$q(\lambda) \sim \frac{1-\alpha}{\Gamma(2-\alpha)} \lambda^{-\alpha} L(1/\lambda)^{-1}, \qquad \lambda \to 0^+. \tag{2.13}$$

The reverse implication in Karamata's Tauberian theorem then yields the first statement of the lemma.

From (2.13) and

$$q(\lambda) \leq \left( \sum_{n=0}^{\infty} \exp \left\{ -2\lambda n \right\} \right)^{1/2} \left( \sum_{n=0}^{\infty} q_n^2 \right)^{1/2} \leq \lambda^{-1/2} \left( \sum_{n=0}^{\infty} q_n^2 \right)^{1/2},$$

12

(the first inequality is Cauchy-Schwarz; the second inequality holds for $\lambda > 0$ small enough), we find that $\sum_{n=0}^{\infty} q_n^2$ diverges, if $\alpha > 1/2$.

Set

$$P(t) = \sum_{n=0}^{\infty} p_n \exp\{itn\}$$

and

$$Q(t) = \sum_{n=0}^{\infty} q_n \exp\{itn\}$$

to be the Fourier transforms of the two sequences.

From $\mu \in \Gamma_\alpha$ with $\alpha \in (0,1)$, we learn from Theorem 1 of [12] that

$$1 - \mathrm{Re}P(t) \sim \mu(t, \infty)\Gamma(1 - \alpha) \cos(\alpha\pi/2) \qquad \text{as } t \downarrow 0$$

and

$$\mathrm{Im}P(t) \sim \tan(\alpha\pi/2)\Big(1 - \mathrm{Re}P(t)\Big) \qquad \text{as } t \downarrow 0,$$

whence

$$\Big|1 - P(t)\Big| \sim \Gamma(1 - \alpha)t^\alpha L(t^{-1}) \qquad \text{as } t \downarrow 0, \tag{2.14}$$

if $\mu(t, \infty) = t^{-\alpha}L(t)$. Similarly,

$$\Big|1 - P(2\pi - t)\Big| \sim \Gamma(1 - \alpha)t^\alpha L(t^{-1}) \qquad \text{as } t \downarrow 0.$$

From $Q = (1 - P)^{-1}$, we find that, for any constant $C > \Gamma(1 - \alpha)^{-1}$, there exists $c > 0$ such that

$$|Q(t)| \leq Ct^{-\alpha}L(1/t)^{-1} \tag{2.15}$$

for $0 < t < c$, and, similarly,

$$|Q(2\pi - t)| \leq Ct^{-\alpha}L(1/t)^{-1}. \tag{2.16}$$

From the support of $\mu$ having greatest common denominator one, $|P(\lambda)| < 1$ for $\lambda \in (0, 2\pi)$. By the continuity of $P : (0, \infty) \to \mathbb{C}$,

$$\sup\Big\{|P(\lambda)| : \lambda \in [c, 2\pi - c]\Big\} \leq 1 - c, \tag{2.17}$$

for $c > 0$ small.

By Parseval's identity,

$$\sum_{n=0}^{\infty} q_n^2 = \frac{1}{2\pi} \int_0^{2\pi} |Q(t)|^2 dt,$$

13

which, by (2.15), (2.16) and (2.17) is finite, provided that $\alpha < 1/2$. $\square$

**Proof of Proposition 2.** Let $\alpha > 0$ and $\mu \in \Gamma_\alpha$. Let $n, m \in \mathbb{N}$ satisfy $n < m$. We wish to show that, if $\alpha > 1/2$, then $n$ and $m$ lie in the same component of $G_\mu$, while, in the case that $\alpha < 1/2$, there is a positive probability (depending on $m - n$) that they lie in different components.

Let $R_\mu^{(1)}$ and $R_\mu^{(2)}$ be independent samples of $R_\mu$. It is easily seen that $n$ and $m$ almost surely lie in the same component of $G_\mu$ if and only if $\mathbb{E}\big(\big|\{n - R_\mu^{(1)}\} \cap \{m - R_\mu^{(2)}\}\big|\big) = \infty$. Moreover, if there exists some pair $(m, n) \in \mathbb{Z}^2$ for which this expectation is infinite, then clearly, it is infinite for all such pairs. Hence, there exists one component of $G_\mu$ precisely when $\mathbb{E}\big|R_\mu^1 \cap R_\mu^2\big|$ is infinite. This expectation is equal to $\sum_{n=1}^\infty q_n^2$, so the result follows from Lemma 2.2(ii). $\square$

# 3    Convergence to fractional Brownian motion

In this section, we establish that the walk has a fractional Brownian motion scaling limit in the sense of finite dimensional distributions:

**Proposition 3** *Let $\mu \in \Gamma_\alpha$, for some $\alpha \in (0, 1/2)$. Recall the constant $\tilde{c} = \tilde{c}(p)$ for $p \in (0, 1)$ from (1.3). Then, for each $p \in (0, 1)$, the process $(0, \infty) \to \mathbb{R} : t \to cn^{-\frac{1}{2} - \alpha} L(n)\Big(S_{\lambda_p}(nt) - n(2p-1)t\Big)$ converges weakly (in the sense of finite dimensional distributions) as $n \to \infty$ to fractional Brownian motion with Hurst parameter $\alpha + 1/2$.*

In preparation for the proof, let $X : \mathbb{Z} \to \{-1, 1\}$ be a sample of $\lambda_p$, and let $S : \mathbb{Z} \to \mathbb{Z}$ be the random walk satisfying $S(0) = 0$ and $S_n - S_{n-1} = X_n$ for each $n \in \mathbb{Z}$. It is our aim to show that $S_n$ is approximately normally distributed, when $n$ is chosen to be high. We begin by finding an explicit expression for the variance of $S_n$.

**Lemma 3.1** *We have that*

$$\mathrm{Var}\big(S_n\big) \sim \frac{4p(1-p)K_\alpha}{\big(|Q|^2\big)_0} n^{2\alpha+1} L(n)^{-2},$$

*where $\big(|Q|^2\big)_i = \sum_{j=0}^\infty q_j q_{i+j}$ denotes the $i$-th Fourier coefficient of $|Q|^2$, and where*

$$K_\alpha = \frac{1}{2\alpha(2\alpha + 1)}\Big(\Gamma\big(1 - 2\alpha\big)^2 \Gamma\big(2\alpha\big) \cos\big(\pi\alpha\big)\Big)^{-1}.$$

**Proof.** We begin by showing that

$$\mathrm{Var}S_n = \frac{4p(1-p)}{\big(|Q|^2\big)_0}\Big(2\sum_{i=1}^n (n-i)\big(|Q|^2\big)_i + n\big(|Q|^2\big)_0\Big). \tag{3.18}$$

14

(In fact, (3.18) holds, more generally, for any $\mu$ for which $G_\mu$ has a.s. infinitely many components.) To do so, write $T_1, \ldots T_{r(n)}$ for the trees having non-empty intersection with $\{1, \ldots, n\}$. Then

$$
\begin{aligned}
\mathrm{Var}\big(S_n\big) &= 4p(1-p)\mathbb{E}\sum_{i=1}^{r(n)} \Big|T_i \cap \{1,\ldots,n\}\Big|^2 \\
&= 4p(1-p)\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{P}\Big(A_i \cap A_j \neq \emptyset\Big).
\end{aligned}
$$

Note that, for $i < j$,

$$
\mathbb{E}\Big(A_i \cap A_j\Big) = \sum_{k=0}^{\infty} q_k q_{j-i+k} = \big(|Q|^2\big)_{j-i},
$$

and also that

$$
\mathbb{E}\big(A_i \cap A_j\big) = \mathbb{P}\big(A_i \cap A_j \neq \emptyset\big)\sum_{i=0}^{\infty}q_i^2 = \mathbb{P}\big(A_i \cap A_j \neq \emptyset\big)\big(|Q|^2\big)_0,
$$

since we adopt the convention that $q_0 = 1$. Hence,

$$
\mathrm{Var}\big(S_n\big) = \frac{4p(1-p)}{\big(|Q|^2\big)_0}\sum_{i=1}^{n}\sum_{j=1}^{n}\big(|Q|^2\big)_{|j-i|},
$$

whence, (3.18). By (3.18), it suffices to show that

$$
\sum_{i=1}^{n}(n-i)\big(|Q|^2\big)_i \sim K_\alpha n^{2\alpha+1}L(n)^{-2}, \tag{3.19}
$$

because Lemma 2.2(ii) implies that $\big(|Q|^2\big)_0 < \infty$. To this end, note that, by (2.14) and $Q = (1-P)^{-1}$,

$$
\big|Q(t)\big| \sim \Gamma(1-\alpha)^{-1}t^{-\alpha}L\big(t^{-1}\big)^{-1},
$$

so that

$$
|Q(t)|^2 \sim \Gamma(1-\alpha)^{-2}t^{-2\alpha}L\big(t^{-1}\big)^{-2},
$$

By Theorem 4.10.1(a) of [6], using $|Q(t)|^2 \in \mathbb{R}$,

$$
\sum_{i=1}^{n}\big(|Q(t)|^2\big)_i \sim \Big(\Gamma\big(1-2\alpha\big)^2\Gamma\big(2\alpha\big)\cos\big(\pi\alpha\big)\Big)^{-1}\frac{n^{2\alpha}L(n)^{-2}}{2\alpha}.
$$

15

It follows that

$$\sum_{i=1}^{n} (n-i)\big(|Q|^2\big)_i \sim \Big(\Gamma\big(1-2\alpha\big)^2 \Gamma\big(2\alpha\big) \cos\big(\pi\alpha\big)\Big)^{-1} n^{2\alpha+1} L(n)^{-2} \int_0^1 (1-x) x^{2\alpha-1} dx.$$

Evaluating the integral, we obtain (3.19). □

**Proof of Proposition 3.** There being a unique stationary Gaussian process $S^H : \mathbb{R} \to \mathbb{R}$ with covariances given by $\mathbb{E}|S_t^H|^2 = t^{2\alpha+1}$, it suffices, in light of Lemma 3.1, to establish that $S_n$ has a distribution that is asymyptotically Gaussian. We will make use of the following notation. For $M \in \mathbb{Z}$, set $\sigma_M = \{X_i : i < M\}$. Let $X_n^M$ and $S_n^M$ denote the expected values of $X_n$ and $S_n$ respectively given $\sigma_M$. (Note that $X_n^M$ is the expected value of $X_k$ where $k$ is the first element in the ancestral line of $n$ that is less than $M$.) Clearly, $X_n^M$ and $S_n^M$ are martingales in $M$, and $S_n^M = S_n$ when $n \geq M$. We will establish that $S_n$ has an asymptotically Gaussian law by applying the martingale central limit theorem to $S_n^M$. This will require showing that $S_n^M$ has small increments compared to its total size and a use of a long-range near-independence argument to show that the sum of the conditional variances of the increments is concentrated. We now state the martingale central limit theorem in the form that we require. See Theorem 7.2 in Chapter 7 of [10], and the remark following its proof, for a derivation.

**Definition 3.1** *We say that $X_{n,m}, \mathbb{F}_{n,m}$, $n \geq 1$, $1 \leq m \leq k_n$ is a martingale difference array if $X_{n,m} \in \mathbb{F}_{n,m}$ and $\mathbb{E}\big(X_{n,m}\big|\mathbb{F}_{n,m-1}\big) = 0$ for $1 \leq m \leq k_n$. Let $V_n = \sum_{i=1}^{k_n} \mathbb{E}\big(X_{n,m}^2\big|\mathbb{F}_{n,m-1}\big).$*

**Theorem 3.1** *Suppose that $\big\{X_{n,m}, \mathbb{F}_{n,m}\big\}$ is a martingale difference array. Let $X_{n,0} \in \mathbb{F}_{n,0}$, $n \in \mathbb{N}$. Set $S_n = \sum_{j=0}^{k_n} X_{n,j}$. Assume that, for some sequence $\epsilon_n \to 0$,*

- $\big(\mathbb{E}V_n\big)^{-1}\big|X_{n,m}\big|^2 \leq \epsilon_n$ *for all* $m \in \{1, \ldots, k_n\}$

- $\big(\mathbb{E}V_n\big)^{-1}\mathrm{Var}\big(X_{n,0}\big) \leq \epsilon_n$ *for $n$ sufficiently high,*

- $\frac{V_n}{\mathbb{E}V_n} \to 1$, *in probability.*

*Then $\mathbb{E}(V_n)^{-1/2} S_n$ converges in distribution to the normal distribution of mean zero and unit variance.*

We will apply the result in the following way. Let $\big\{\epsilon_n : n \in \mathbb{N}\big\}$ be a given sequence converging to zero. Let $\big\{k_n : n \in \mathbb{N}\big\}$ be a sequence satisfying

$$\mathrm{Var}\big(S_n^{-k_n}\big) \leq 2^{-1}\epsilon_n \frac{4p(1-p)K_\alpha}{(|Q|^2)_0} n^{2\alpha+1} L(n)^{-2}, \tag{3.20}$$

where the constant $K_\alpha$ is specified in Lemma 3.1. We in addition assume that $k_n \geq Cn$ for each $n \in \mathbb{N}$, and for some large constant $C > 0$. We choose as martingale difference array,

16

$X_{n,0} = S_n^{-k_n}$, $\mathbb{F}_{n,0} = \sigma_{-k_n}$ for each $n \geq 1$, and $X_{n,i} = S_n^{-k_n+i} - S_n^{-k_n+i-1}$, $\mathbb{F}_{n,i} = \sigma_{-k_n+i}$ for each $n \geq 1$ and $1 \leq i \leq k_n + n$.

Clearly, we must verify that the hypotheses of Theorem 3.1 hold. The first column of the martingale difference array has been chosen to meet the second condition listed in the theorem, as we now confirm. Note that

$$\mathbb{E}V_n = \mathrm{Var}\left(S_n - S_n^{k_n}\right). \tag{3.21}$$

By (3.20) and Lemma 3.1, we see that

$$\mathrm{Var}\left(S_n^{-k_n}\right) \leq 2^{-1}\epsilon_n \mathrm{Var}S_n\left(1 + o(1)\right). \tag{3.22}$$

Using $\mathrm{Var}(S_n) = \mathrm{Var}(S_n - S_n^{-k_n}) + \mathrm{Var}(S_n^{-k_n})$ and (3.21), we obtain $\mathrm{Var}S_n^{-k_n} \leq \epsilon_n \mathbb{E}V_n$ for $n$ sufficiently high, which is indeed the second hypothesis of Theorem 3.1.

The general martingale difference term takes the form $S_n^m - S_n^{m-1}$, for some $n \geq 1$ and $m \leq n$. A basic observation is that, for any $M < n$,

$$X_n^{M+1} - X_n^M = \left(X_M - X_M^M\right)q_{n-M}, \tag{3.23}$$

where $q_i$ (as in Definition 2.1) denotes the probability that the vertex $i$ has 0 as an ancestor. As stated earlier, $X_n^M$ is the expected value of the value observed by tracing back the ancestral line in a sample of $G_\mu$ from $n$, until an ancestor with index strictly less than $M$ is reached. The difference in the above equation, then, is due to the event that the ancestral line reaches $M$.

If $n \geq 1$ and $m < 0$, note that $S_n^{m+1} - S_n^m = \sum_{i=1}^n \left(X_i^{m+1} - X_i^m\right)$, so that (3.23) yields

$$S_n^{m+1} - S_n^m = \left(X_m - X_m^m\right)F_n^m, \tag{3.24}$$

where we define $F_n^m = \sum_{i=1}^n q_{i-m}$. In the case where $n \geq 1$ and $m \in \{0, \ldots, n-1\}$, on the other hand,

$$S_n^{m+1} - S_n^m = \sum_{i=m}^{n-1} \left(X_i^{m+1} - X_i^m\right) = \left(X_m - X_m^m\right)\sum_{i=m}^{n-1} q_{i-m},$$

by the convention that $q_0 = 1$. We record this as

$$S_n^{m+1} - S_n^m = \left(X_m - X_m^m\right)\sum_{i=0}^{n-m-1} q_i. \tag{3.25}$$

Note also that, for any $m \in \mathbb{Z}$,

$$\mathrm{Var}\left(X_m - X_m^m \middle| \sigma_m\right) = 4P_m\left(1 - P_m\right),$$

where $P_m$ is defined to be $\mathbb{P}(X_m = 1 | \sigma_m)$ and may be written $(X_m^m + 1)/2$. The definition of $V_n$ and the expressions (3.24) and (3.25) now give the formula

$$V_n = \sum_{M=-k_n}^{-1} (F_M^n)^2 \mathrm{Var}(X_M - X_M^M | \sigma_M) + \sum_{m=1}^{n} \left( \sum_{l=1}^{n-m} q_l \right)^2 \mathrm{Var}\left( X_m - X_m^m \Big| \sigma_M \right),$$

which may be written

$$V_n = 4 \sum_{M=-k_n}^{-1} P_M (1 - P_M) (F_M^n)^2 + 4 \sum_{m=1}^{n} P_m (1 - P_m) \left( \sum_{l=1}^{n-m} q_l \right)^2. \tag{3.26}$$

As such, the following lemma shows that the third hypothesis of Theorem 3.1 is satisfied in the present case.

**Lemma 3.2** *Setting* $c_0 = 4\mathbb{E}(P_M (1 - P_m))$,

$$4 \sum_{M=-k_n}^{-1} P_M (1 - P_M) (F_M^n)^2 = c_0 \sum_{M=-k_n}^{-1} (F_M^n)^2 \left( 1 + E_1(n) \right),$$

*where* $E_1(n) \to 0$ *in probability. We have further that*

$$\sum_{m=1}^{n} P_m (1 - P_m) \left( \sum_{l=1}^{n-m} q_l \right)^2 = c \sum_{m=1}^{n} \left( \sum_{l=1}^{n-m} q_l \right)^2 \left( 1 + E_2(n) \right).$$

*where* $E_2(n) \to 0$ *in probability.*

The ergodic theorem might be used to prove a concentration inequality of this sort. We prefer, however, to derive it directly, from a second moment estimate. We begin by observing the following.

**Lemma 3.3** *Set*
$$Z_i = \mathrm{Var}(X_i - X_i^i | \sigma_i) = 4 P_i (1 - P_i),$$

*and write*
$$\rho_{l,m} = \mathbb{E}(Z_l Z_m) - \mathbb{E}(Z_l) \mathbb{E}(Z_m).$$

*Then, for each* $\epsilon > 0$, *there exists* $K_0 \in \mathbb{N}$ *such that* $|l - m| \geq K_0$ *implies that* $|\rho_{l,m}| < \epsilon$.

**Proof.** For $i \in \mathbb{Z}$ and $k \in \mathbb{N}$, set $X_{i,k}^i = \sum_{j=1}^{k} \mu\{j\} X_{i-j}$. Note that

$$\left| X_i^i - X_{i,k}^i \right| \leq \sum_{j=k+1}^{\infty} \mu\{j\} = (k+1)^{-\alpha} L(k+1). \tag{3.27}$$

18

For $i, j \in \mathbb{Z}$ and $k \in \mathbb{N}$, set

$$R_{i,j,k} = \left\{ \bigcup_{m=i-k}^{i-1} A_m \cap \bigcup_{n=j-k}^{j-1} A_n = \emptyset \right\},$$

where recall that we write $A_n$ for the ancestral line of $n \in \mathbb{Z}$. Shortly, we will show that, for each $\epsilon > 0$ and $k \in \mathbb{N}$, there exists $n_0 \in \mathbb{N}$ such that, if $i, j \in \mathbb{Z}$ satisfy $|i - j| \geq n_0$, then

$$\lambda\left(R_{i,j,k}\right) < \epsilon. \tag{3.28}$$

For now, we show that this suffices for the proof of the lemma.

Note that, for any $i, j \in \mathbb{Z}$ and $k \in \mathbb{N}$ for which $|i - j| \geq k$, if $Y \in \sigma\{X_{i-k}, \ldots, X_{i-1}\}$ and $Z \in \sigma\{X_{j-k}, \ldots, X_{j-1}\}$

$$\mathbb{E}\left(YZ \mid R_{i,j,k}\right) = \mathbb{E}\left(Y \mid R_{i,j,k}\right)\mathbb{E}\left(Z \mid R_{i,j,k}\right). \tag{3.29}$$

Note that

$$\mathbb{E}\left(P_i(1 - P_i)P_j(1 - P_j)\right)$$

$$= \mathbb{E}\left(\frac{1 + X_i^i}{2}\right)\left(\frac{1 - X_i^i}{2}\right)\left(\frac{1 + X_j^j}{2}\right)\left(\frac{1 - X_j^j}{2}\right)$$

$$= \mathbb{E}\left(\frac{1 + X_{i,k}^i}{2}\right)\left(\frac{1 - X_{i,k}^i}{2}\right)\left(\frac{1 + X_{j,k}^j}{2}\right)\left(\frac{1 - X_{j,k}^j}{2}\right) + O\left(k^{-\alpha}L(k)\right)$$

$$= \mathbb{E}\left(\left(\frac{1 + X_{i,k}^i}{2}\right)\left(\frac{1 - X_{i,k}^i}{2}\right)\left(\frac{1 + X_{j,k}^j}{2}\right)\left(\frac{1 - X_{j,k}^j}{2}\right)\Big| R_{i,j,k}\right) + O(\epsilon) + O\left(k^{-\alpha}L(k)\right)$$

$$= \mathbb{E}\left(\left(\frac{1 + X_{i,k}^i}{2}\right)\left(\frac{1 - X_{i,k}^i}{2}\right)\Big| R_{i,j,k}\right)\mathbb{E}\left(\left(\frac{1 + X_{j,k}^j}{2}\right)\left(\frac{1 - X_{j,k}^j}{2}\right)\Big| R_{i,j,k}\right) + O(\epsilon) + O\left(k^{-\alpha}L(k)\right)$$

$$= \left(\mathbb{E}\left(\frac{1 + X_{i,k}^i}{2}\right)\left(\frac{1 - X_{i,k}^i}{2}\right) + O\left(\lambda(R_{i,j,k}^c)\right)\right)$$

$$\qquad \left(\mathbb{E}\left(\frac{1 + X_{j,k}^j}{2}\right)\left(\frac{1 - X_{j,k}^j}{2}\right) + O\left(\lambda(R_{i,j,k}^c)\right)\right) + O(\epsilon) + O\left(k^{-\alpha}L(k)\right)$$

$$= \mathbb{E}\left(\frac{1 + X_i^i}{2}\right)\left(\frac{1 - X_i^i}{2}\right)\mathbb{E}\left(\frac{1 + X_j^j}{2}\right)\left(\frac{1 - X_j^j}{2}\right) + O(\epsilon) + O\left(k^{-\alpha}L(k)\right),$$

the second equality by (3.27), the third by $|i - j| \geq n_0$ and (3.28), the fourth by (3.29), and the sixth by (3.27) and (3.28). Thus, for any $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that, if $i, j \in \mathbb{Z}$ satisfy $|i - j| \geq n_0$, then

$$\mathrm{Cov}\left(P_i(1 - P_i), P_j(1 - P_j)\right) < \epsilon.$$

19

This completes the proof of the lemma, subject to verifying (3.28). Note that, for this, it suffices to show that, for any $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that, if $i, j \in \mathbb{Z}$ satisfy $|i - j| \geq n_0$, then $\lambda\big(A_i \cap A_j \neq \emptyset\big) < \epsilon$. To see this, note that, for $i < j$,

$$\lambda\big(A_i \cap A_j \neq \emptyset\big) \leq \mathbb{E}\big(A_i \cap A_j\big) = \sum_{n=0}^{\infty} q_n q_{n+i-j} \leq \Big(\sum_{n=0}^{\infty} q_n^2\Big)^{1/2}\Big(\sum_{n=0}^{\infty} q_{n+i-j}^2\Big)^{1/2},$$

so that Lemma 2.2(ii) yields the desired conclusion. $\square$

**Proof of Lemma 3.2.** To verify the first claim, we aim to show that

$$\mathrm{Var} Y_n = o\Big(\big(\mathbb{E}Y_n\big)^2\Big), \tag{3.30}$$

where here we write

$$Y_n = 4 \sum_{M=-k_n}^{-1} P_M(1 - P_M)\big(F_M^n\big)^2.$$

We will require

$$\frac{\max_{-k_n < M < 0} \big(F_M^n\big)^2}{\sum_{M=-k_n}^{-1} \big(F_M^n\big)^2} \to 0 \tag{3.31}$$

as $n \to \infty$.

Indeed, by computing the second moment of $Y_n$, (3.30) follows from (3.31) by means of Lemma 3.3.

We turn to proving (3.31). Recall that we supposed that $k_n \geq Cn$, with $C > 0$ a large constant. Note that

$$\sum_{M=-k_n}^{-1} \big(F_M^n\big)^2 = \sum_{m=1}^{k_n} \Big(\sum_{i=1}^{n} q_{i+m}\Big)^2 \geq \sum_{m=1}^{n} \Big(\sum_{i=1}^{n} q_{i+m}\Big)^2 \geq cn^{2\alpha+1} L(n)^{-2},$$

with $c > 0$ a constant satisfying $c < \Big(\frac{2(1-\alpha)}{\Gamma(2-\alpha)\Gamma(1+\alpha)}\Big)^2$, and the latter inequality by Lemma 2.2(i), (2.10) and (2.11). For $0 \leq m \leq \lfloor Cn \rfloor$ and a constant $C_0 > \frac{2(1-\alpha)}{\Gamma(2-\alpha)\Gamma(1+\alpha)}$

$$\sum_{i=1}^{n} q_{i+m} \leq C_0 n^{\alpha} L(n)^{-1},$$

by Lemma 2.2(i) and (2.10). We learn that

$$\frac{\max_{-k_n < M < 0} \big(F_M^n\big)^2}{\sum_{M=-\lfloor Cn \rfloor}^{-1} \big(F_M^n\big)^2} \leq C^2 c^{-1} n^{-1}. \tag{3.32}$$

20

To treat the bound on $F_M^n$ for $-k_n < M < -n$, note that, if $m > n$, then either

$$\sum_{i=1}^{n} q_{i+j} \geq 2^{-1} \sum_{i=1}^{n} q_{i+m} \text{ for each } j \in \{m - n/2, \ldots, m\}$$

or

$$\sum_{i=1}^{n} q_{i+j} \geq 2^{-1} \sum_{i=1}^{n} q_{i+m} \text{ for each } j \in \{m + 1, \ldots, m + n/2\},$$

the first alternative holding provided that at least one half of the sum $\sum_{i=1}^{n} q_{i+m}$ is contained in its first $\lfloor n/2 \rfloor$ terms. From this, we conclude that, for any $M < -n$,

$$\frac{\left(F_M^n\right)^2}{\sum_{M=-\infty}^{-1} \left(F_M^n\right)^2} \leq 8n^{-1}\left(1 + o(1)\right),$$

whence

$$\frac{\left(F_M^n\right)^2}{\sum_{M=-k_n}^{-1} \left(F_M^n\right)^2} \leq 10n^{-1},$$

by increasing the value of $k_n \in \mathbb{N}$ if necessary. This completes the proof of (3.31).

We have verified the first claim of the lemma, since it evidently follows from (3.30), subject to proving the lemma that follows.

The second claim is proved analogously. The analog of (3.32) is

$$\frac{\max_{1 \leq m \leq n} \left(\sum_{l=1}^{n-m} q_l\right)^2}{\sum_{m=1}^{n} \left(\sum_{l=1}^{n-m} q_l\right)^2}, \tag{3.33}$$

which follows from the inequality

$$\left(\sum_{l=1}^{n-1} q_l\right)^2 \leq Cn^{2\alpha} L(n)^{-2},$$

which is a consequence of Lemma 2.2(i), and the bound

$$\sum_{m=1}^{n} \left(\sum_{l=1}^{n-m} q_l\right)^2 \geq cn^{2\alpha+1} L(n)^{-2},$$

which follows from Lemma 2.2(i) and (2.10). $\square$

It remains to show that the first hypothesis of Theorem 3.1 holds. Recalling (3.24), we see that, for $-k_n < M < 0$, $\left|S_n^{M+1} - S_n^M\right|^2 \leq 4\left(F_M^n\right)^2$, whose right-hand-side is $o\left(\mathbb{E}V_n\right)$ by

21

(3.31), (3.26) and Lemma 3.2. For $0 \leq m < n$, we use (3.25) to find that $\left|S_n^{m+1} - S_n^m\right|^2 \leq 4\left(\sum_{l=1}^{n-m} q_l\right)^2$. However, this right-hand-side is at most $4cn^{2\alpha}L(n)^{-2}$ by Lemma 2.2(i). From the discussion after (3.22), we know that $\mathbb{E}V_n \geq \left(1 - \epsilon_n\right)\mathrm{Var}S_n$ for $n$ sufficiently high, so that $4cn^{2\alpha}L(n)^{-2} \leq O\left(n^{-1}\mathbb{E}(V_n)\right)$ by Lemma 3.1. Thus, the first hypothesis of Theorem 3.1 is indeed satisfied. $\square$

# 4    The FKG inequality and convergence in $L^\infty$

This section is devoted to the remaining step in the proof of Theorem 1.1, namely, to improving the topology in our convergence result.

**Proposition 4** *Let $\mu$ be any probability measure on $\mathbb{N}$, and let $\lambda$ be any extremal $\mu$-Gibbs measure. Let $X, Y \in L^2(\lambda)$ be two increasing functions, where $\{-1, 1\}^{\mathbb{Z}}$ is given the natural poset structure. Then $X$ and $Y$ are not negatively correlated under $\lambda$.*

**Proof.** Note that $\mathrm{Cov}(X, Y) \geq 0$ if and only if

$$\mathrm{Var}(X + Y) \geq \mathrm{Var}(X) + \mathrm{Var}(Y). \tag{4.34}$$

For any $Z \in L^2(\lambda)$, we write

$$Z = \sum_{n \in \mathbb{Z}} \left(\mathbb{E}(Z|\sigma_n) - \mathbb{E}(Z|\sigma_{n-1})\right), \tag{4.35}$$

where $\sigma_n$ is the $\sigma$-algebra generated by coordinates with index strictly less than $n$. In regard to (4.35), note that the maps $L^2(\lambda) \to \mathbb{R} : Z \to \mathbb{E}(Z|\sigma_n)$ that project functions onto their mean given $\sigma_n$ form an increasing sequence of projections in the Hilbert space $L^2(\lambda)$. The identity (4.35) expresses Z as a sum of differences of the successive projections. As such, this is a sum of orthogonal vectors in $L^2(\lambda)$. The sum of a countable collection of orthogonal vectors in a Hilbert space has finite norm if and only if the sum of the squares of the norms of these vectors is finite. The norm of the sum has a square given by the sum of the squares of the norms of the constitutent vectors. From this, we obtain

$$\mathrm{Var}(Z) = \sum_{n \in \mathbb{Z}} \mathbb{E}\left(\mathbb{E}(Z|\sigma_n) - \mathbb{E}(Z|\sigma_{n-1})\right)^2.$$

We find then that

$$\mathrm{Var}(Z) = \sum_{n \in \mathbb{Z}} \mathbb{E}_{\sigma_{n-1}} \mathrm{Var}\left(\mathbb{E}(Z|\sigma_n) - \mathbb{E}(Z|\sigma_{n-1})\right). \tag{4.36}$$

Note that, in the summand on the right-hand-side of (4.36), Var denotes a conditional variance: the data in $\sigma_{n-1}$ is fixed, and the relevant randomness arises from the bit with index $n - 1$.

Given $\sigma_{n-1}$, each of $X$ and $Y$ does not decrease if we condition on the value of the $n$-th coordinate to be 1, and does not increase if this value is conditioned to be $-1$. Thus,

$$\mathrm{Var}\Big(\mathbb{E}\big(X+Y|\sigma_n\big)-\mathbb{E}\big(X+Y|\sigma_{n-1}\big)\Big)$$
$$\geq\quad \mathrm{Var}\Big(\mathbb{E}\big(X|\sigma_n\big)-\mathbb{E}\big(X|\sigma_{n-1}\big)\Big)+\mathrm{Var}\Big(\mathbb{E}\big(Y|\sigma_n\big)-\mathbb{E}\big(Y|\sigma_{n-1}\big)\Big),$$

$\sigma_{n-1}$-a.s. Taking expectation over $\sigma_{n-1}$ and summing yields (4.34) by means of (4.36). $\square$

**Definition 4.1** *Recall that $S_0 = 0$, $S_i = \sum_{j=1}^{i} X_j$ if $i > 0$, and $S_i = -\sum_{j=i}^{-1} X_j$ if $i < 0$, where $X_i = X_n \lambda_p(i)$. For $l < m$, write*

$$A_{l,m} = \max_{i\in\{l,\dots,m\}}\big(S_i-(2p-1)i\big)-\big(S_l-(2p-1)l\big)$$

*and*

$$B_{l,m} = \max_{i\in\{l,\dots,m\}}\big(S_i-(2p-1)i\big)-\big(S_m-(2p-1)m\big)$$

**Lemma 4.1** *There exists a constant $C_0 > 0$ such that, for each $\epsilon > 0$, there exists $n_0 = n_0(\epsilon)$ such that, for $n \geq n_0$,*

$$\mathbb{P}\Big(A_{0,n} > \Big(C_0+50\tilde{c}^{-1}\sqrt{\log\big(\epsilon^{-1}\big)}\Big)n^{1/2+\alpha}L(n)^{-1}\Big) < \epsilon,$$

*where the constant $\tilde{c} = \tilde{c}(p)$ was defined in (1.3).*

**Proof.** We write $A_n = A_{0,n}$ and $B_n = B_{0,n}$. For $n \in \mathbb{N}$, set $h(n) = 3\tilde{c}^{-1}\sqrt{\log\big(\epsilon^{-1}\big)}n^{1/2+\alpha}L(n)^{-1}$. We begin by establishing the following statements. For any $\epsilon > 0$, there exists $n_0 = n_0(\epsilon)$ such that, if $n \geq n_0(\epsilon)$, then, for any $K > 0$,

$$\mathbb{P}\Big(B_n > K+h(n)\Big) \geq \epsilon \quad\Longrightarrow\quad \mathbb{P}\big(A_n > K\big) \geq 1-\epsilon \tag{4.37}$$

and

$$\mathbb{P}\big(A_n > K+h(n)\big) \geq \epsilon \quad\Longrightarrow\quad \mathbb{P}\Big(B_n > K\Big) \geq 1-\epsilon. \tag{4.38}$$

Note that $A_n$ is an increasing and $B_n$ a decreasing random variable. As such, supposing that $\mathbb{P}\big(B_n > K+h(n)\big) \geq \epsilon$ and $\mathbb{P}\big(A_n \leq K\big) \geq \epsilon$, then

$$\mathbb{P}\Big(\big\{B_n > K+h(n)\big\}\cap\big\{A_n \leq K\big\}\Big)$$
$$=\quad \mathbb{P}\Big(B_n > K+h(n)\Big)\mathbb{P}\Big(A_n \leq K\big|B_n > K+h(n)\Big)$$
$$\geq\quad \mathbb{P}\Big(B_n > K+h(n)\Big)\mathbb{P}\Big(A_n \leq K\Big) \geq \epsilon^2,$$

23

the first inequality by means of Proposition 4. However, $A_n - B_n = S_n - (2p-1)n$, so that

$$\mathbb{P}\Big(B_n - A_n > h(n)\Big) = \mathbb{P}\Big(n^{-1/2-\alpha}L(n)\tilde{c}\big(S_n - (2p-1)n\big) < -3\sqrt{\log\big(\epsilon^{-1}\big)}\Big) < \epsilon^2,$$

the inequality valid for $n$ sufficiently high, because Proposition 3 implies that $n^{-1/2-\alpha}L(n)\tilde{c}\big(S_n - (2p-1)n\big)$ converges in distribution to a Gaussian random variable of mean zero and unit variance. In this way, we establish (4.37), with (4.38) following similarly.

We now show that there exists $\alpha_n \in (0, \infty)$ such that, for $\epsilon > 0$ and $n \geq n_0(\epsilon)$,

$$\mathbb{P}\Big(A_n \in \big(\alpha_n - 3h(n), \alpha_n + 3h(n)\big)\Big) \geq 1 - 2\epsilon \tag{4.39}$$

and

$$\mathbb{P}\Big(B_n \in \big(\alpha_n - 3h(n), \alpha_n + 3h(n)\big)\Big) \geq 1 - 2\epsilon. \tag{4.40}$$

To this end, set

$$K(n) = \sup\big\{k \in (0, \infty) : \mathbb{P}\big(A_n > k\big) \geq \epsilon\big\} - \epsilon/2. \tag{4.41}$$

Note that $\mathbb{P}\big(A_n > K(n)\big) \geq \epsilon$ and $\mathbb{P}\big(A_n > K(n) + \epsilon\big) < \epsilon$. The first inequality forces $\mathbb{P}\big(B_n > K(n) - h(n)\big) \geq 1 - \epsilon$ by means of (4.38), which gives

$$\mathbb{P}\big(A_n > K(n) - 2h(n)\big) \geq 1 - \epsilon. \tag{4.42}$$

by (4.37). We now set $\alpha_n = K(n) - h(n)$ and note that, if $\epsilon < h(n)$, then (4.39) holds. Indeed, these choices for $\alpha_n$ and $\epsilon$, alongside $P(A_n > K(n) + \epsilon) < \epsilon$, yield

$$P\big(A_n > \alpha_n + 2h(n)\big) < \epsilon. \tag{4.43}$$

The bounds (4.42) and (4.43) imply that

$$P(\alpha_n + 2h(n) \geq A_n \geq \alpha_n - h(n)) \geq 1 - 2\epsilon,$$

whence (4.39). Note now that $\mathbb{P}\big(B_n \geq \alpha_n + 2h(n) + \epsilon\big) < \epsilon$: for otherwise, (4.37) would imply that $\mathbb{P}\big(A_n > \alpha_n + h(n) + \epsilon\big) \geq 1 - \epsilon$, contradicting our assumption. We see that

$$\mathbb{P}\Big(B_n \in \big(\alpha_n, \alpha_n + 2h(n) + \epsilon\big)\Big) \geq 1 - 2\epsilon,$$

whence (4.40).

We set $r_n = \alpha_n n^{-1/2-\alpha}L(n)$. We claim that, for any $K \in \mathbb{N}$, there exists $n_0 \in \mathbb{N}$ such that $n \geq n_0$ implies

$$r_{Kn} \leq 2K^{-1/2-\alpha}r_n + 20\tilde{c}^{-1}\sqrt{\log\big(\epsilon^{-1}\big)} + 24\tilde{c}^{-1}\sqrt{\log\big(\epsilon^{-1}\big)}K^{-1/2-\alpha}. \tag{4.44}$$

From this, we will argue that

$$\sup_{n \in \mathbb{N}} r_n \leq C_0 + 40\tilde{c}^{-1} \sqrt{\log\left(\epsilon^{-1}\right)} \tag{4.45}$$

for some $C_0 > 0$. Indeed, fixing $K \in \mathbb{N}$ in (4.44) high enough that $2K^{-1/2-\alpha} < 1$, we find that the limsup of $r_m$ along the subsequence $\{m = Kn : n \in \mathbb{N}\}$ is at most $(20 + 12 + 1)\left(\tilde{c}^{-1}\sqrt{\log\left(\epsilon^{-1}\right)}\right)$. Note then that $\{A_n : n \in \mathbb{N}\}$ is a random process whose increments are in absolute value at most one. Thus, the definition (4.41) implies that $|\alpha_m - \alpha_n| \leq |m - n|$. From this, we obtain the above bound on the limsup of $r_m$ over all $m \in \mathbb{N}$ (with an arbitrarily small addition to the value 33); and thus (4.45).

In tandem with (4.39), (4.45) yields

$$\mathbb{P}\left(A_n > \left(C_0 + 49\tilde{c}^{-1}\sqrt{\log\left(\epsilon^{-1}\right)}\right)n^{1/2+\alpha}L(n)^{-1}\right) \leq 2\epsilon,$$

from which the statement of the lemma follows. To derive (4.44), note that

$$A_{Kn} \leq \max_{j \in \{1,\ldots,K-1\}}\left(S_{jn} - (2p-1)jn\right) + \max_{j \in \{0,\ldots,K-1\}} A_{jn,(j+1)n}. \tag{4.46}$$

By the convergence for finite-dimensional distributions stated in Proposition 3, we have that, for each $\epsilon > 0$, there exists $n_0 = n_0(\epsilon)$ such that, for $n \geq n_0$,

$$\mathbb{P}\left(\max_{j \in \{1,\ldots,K-1\}}\left(S_{jn} - (2p-1)jn\right) > 3\tilde{c}^{-1}\sqrt{\log\left(\epsilon^{-1}\right)}n^{1/2+\alpha}L(n)^{-1}\right) < (K-1)\epsilon.$$

Note also that

$$\mathbb{P}\left(\max_{j \in \{0,\ldots,K-1\}} A_{jn,(j+1)n} \geq \alpha_n + 3h(n)\right) \leq 2K\epsilon,$$

by (4.39). From (4.46), then,

$$\mathbb{P}\left(A_{Kn} > 3\tilde{c}^{-1}\sqrt{\log\left(\epsilon^{-1}\right)}n^{1/2+\alpha}L(n)^{-1} + \alpha_n + 3h(n)\right) \leq (3K-1)\epsilon \tag{4.47}$$

for $n \geq n_0(\epsilon)$. Choosing $\epsilon > 0$ to satisfy $(3K-1)\epsilon \leq 1 - 2\epsilon$, from (4.47) and the inequality arising from (4.39) by the substitution of $Kn$ for $n$, we see that

$$3\tilde{c}^{-1}\sqrt{\log\left(\epsilon^{-1}\right)}n^{1/2+\alpha}L(n)^{-1} + \alpha_n + 3h(n) \geq \alpha_{Kn} - 3h(Kn)$$

Rearranging, and by $L : (0,\infty) \to (0,\infty)$ being a slowly varying function, we obtain (4.44). $\square$

**Proof of Theorem 1.1.** We will make use of:

**Lemma 4.2** *There exists a constant $C \in (0, \infty)$ such that, for any $\epsilon_0 > 0$ sufficiently small and for $\delta > 0$,*

$$\limsup_{n \to \infty} \mathbb{P}\left( \sup_{k=0,\ldots,\lfloor \epsilon_0^{-1} \rfloor} A_{k \lfloor \epsilon_0 n \rfloor, (k+1) \lfloor \epsilon_0 n \rfloor} > \delta n^{1/2+\alpha} L(n)^{-1} \right) \leq C\left(\epsilon_0^{-1} + 1\right) \exp\left\{ -\frac{\tilde{c}^2}{2 \cdot 50^2} \delta^2 \epsilon_0^{-1/2-2\alpha} \right\}.$$

**Remark.** An analogous bound holds for the counterpart of $A_{l,m}$, in which the maximum taken in the definition is replaced by a minimum.

**Proof.** Note that (2.11) implies that, for $n$ sufficiently high,

$$n^{1/2+\alpha} L(n)^{-1} \geq \epsilon_0^{-1/4-\alpha} \left(\epsilon_0 n\right)^{1/2+\alpha} L\left(\epsilon_0 n\right)^{-1}.$$

The probability that we must estimate is bounded above by

$$\left(\epsilon_0^{-1} + 1\right) \mathbb{P}\left( A_{0, \lfloor \epsilon_0 n \rfloor} \geq \delta \epsilon_0^{-1/4-\alpha} \left(\epsilon_0 n\right)^{1/2+\alpha} L\left(n \epsilon_0\right)^{-1} \right)$$

Setting $\epsilon > 0$ in Lemma 4.1 according to $C_0 + 50 \tilde{c}^{-1} \sqrt{\log\left(\epsilon^{-1}\right)} = \delta \epsilon_0^{-1/4-\alpha}$, we obtain that the last displayed expression is at most

$$\left(\epsilon_0^{-1} + 1\right) \exp\left\{ -\frac{\tilde{c}^2}{50^2} \left(\delta \epsilon_0^{-1/4-\alpha} - C_0\right)^2 \right\} \leq C\left(\epsilon_0^{-1} + 1\right) \exp\left\{ -\frac{\tilde{c}^2}{2 \cdot 50^2} \delta^2 \epsilon_0^{-1/2-2\alpha} \right\}$$

for some constant $C > 0$, for $\epsilon_0$ small, and for $n$ sufficiently high. $\square$

For the proof of the theorem, it suffices to construct, for each $\epsilon > 0$, a sequence $\left\{ C_n^\epsilon : n \in \mathbb{N} \right\}$ of couplings $S_p^n$ and $S_{\alpha+1/2}$ for which

$$\lim_{n \to \infty} C_n^\epsilon\left( \|S_p^n - S_{\alpha+1/2}\|_{L_\infty\left([0,T]\right)} > \epsilon \right) = 0, \tag{4.48}$$

for then we may find an increasing sequence $\left\{ n_i : i \in \mathbb{N} \right\}$ such that

$$\sup_{n \geq n_i} C_n^{2^{-i}}\left( \|S_p^n - S_{\alpha+1/2}\|_{L_\infty\left([0,T]\right)} > 2^{-i} \right) < 2^{-i},$$

for each $i \in \mathbb{N}$, and set $C_m = C_m^{2^{-k_m}}$, where $k_m$ is the maximal $n_i$ that does not exceed $m$; note that $k_m \to \infty$ ensures that this sequence of couplings is as the statement of the theorem demands. A coupling $C_n^\epsilon$ suitable for (4.48) may be obtained by the use of Proposition 3 to couple the values of $S_p^n$ and $S_{\alpha+1/2}$ at points of the form $\left\{ i\epsilon T : i = 0, \ldots, \lfloor \epsilon^{-1} \rfloor + 1 \right\}$, so that the maximum difference $|S_p^n - S_{\alpha+1/2}|$ at such points tends to zero in probability under $C_n^\epsilon$. We then use Lemma 4.2, and the remark following its statement, as well as $S_{\alpha+1/2}$ being uniformly continuous on $[0, T]$ to verify (4.48). $\square$

26

# 5    Open problems

- It would be interesting to find the negatively correlated fractional Brownian motions, with Hurst parameter $H \in (0, 1/2)$, as scaling limits of variants of the discrete processes that we consider. A natural first guess is the walk arising from the model in which each vertex of $\mathbb{Z}$ is given the opposite sign of its parent in $G_\mu$, instead of the same sign. However, it is quite possible that this gives rise to Brownian motion as a scaling limit of the associated walk. A more promising candidate discrete model is one in which the vertices in each component of $G_\mu$ — ordered with respect to the natural ordering of $\mathbb{Z}$ — are given values that alternate between $1$ and $-1$, independently choosing one of the two ways of doing this, each with probability $1/2$.

- As mentioned in the discussion following the statement of Theorem 1.1, we believe that the theorem is sharp. We pose the problem to show that if a measure $\mu$ on $\mathbb{N}$ is such that the conclusion of Theorem 1.1 holds, with some deterministic function of $n$ playing the role of $\tilde{c}n^{-1/2-\alpha}L(n)$, then $\mu \in \Gamma_\alpha$. An important step here would be to show that, if the measure $\mu$ is such that the variance of $S_n$ is a slowly varying multiple of $n^{2\alpha+1}$, then $\mu \in \Gamma_\alpha$.

- Define $\tilde{\Gamma}_\alpha$ according to Definition 1.1, with the alteration that $\mu$ is supported on $\mathbb{Z}$ and is symmetric about $0$, and let $\mu \in \tilde{\Gamma}_\alpha$ for some $\alpha \in (0, \infty)$. Consider a voter model in dimension one, in which voters reside at the elements of $\mathbb{Z}$. Each voter at a given time has an affiliation to one of two political parties. At any given moment $t \in \mathbb{Z}$ of time, each voter selects the resident at a displacement given by an independent sample of $\mu$, and inherits the affiliation that this resident held at time $t - 1$. (A continuous time variant, where opinions are imposed on a voter at a Poisson point process of times, may also be considered.) We propose the problem of considering the set of equilibrium measures of this process as an analogue of the family $\{\lambda_p : p \in [0, 1]\}$ and seeking a counterpart to Theorem 1.1. The two unanimous configurations are always equilibrium measures, and we anticipate that mixtures of these are the only such measures exactly when $\alpha \geq 1$ (and the condition g.c.d.$(\mu) = 1$ is satisfied). This is because the transmission histories of the affiliation held by two voters will almost surely coincide in the distant past precisely when a discrete time random walk with step distribution $\mu$ almost surely reaches zero. The case $\alpha = 1$ corresponds to the step distribution of the $x$-displacement of a two-dimensional simple random walk between its successive visits to the $x$-axis; such a walk is recurrent, but only "marginally so", pointing to the value of $\alpha = 1$ being critical for the problem.

- Let $d \geq 2$. For any measure $\mu$ whose support is contained in the integer lattice $\mathbb{Z}^d$, the random $\mathbb{Z}^d$-spanning graph structure $G_\mu$ may be defined. Consider a law $\mu$ that has a regularly decaying heavy tail. For example, we might insist that $\mu(B_n^c) = n^{-\alpha}L(\|n\|)$,

where $L : (0, \infty) \to (0, \infty)$ is slowly varying. Here, $B_n = \big\{ x \in \mathbb{Z}^d : ||x|| \le n \big\}$. If $\mu$ is chosen to be "symmetric", in the sense that $\mu\{x\} = \mu\{y\}$ whenever $x = y$, it is not hard to show that all the components of $G_\mu$ are finite, each containing a unique cycle. Choices of $\mu$ supported in a half-plane $\big\{ x_1 < 0 \big\}$ will give rise to infinite components, however. This raises the question of finding a phase transition in $\alpha$ for uniqueness of the infinite component for some family of measures $\mu$ supported in the strict half-plane $\big\{ x_1 < 0 \big\}$. In fact, the discussion in the preceding problem already addresses this problem in a certain guise. To obtain the voter model problem, we take $d = 2$, write points in the plane in the form $(t, x)$, take $\mu$ to be supported on the line $\big\{ x = -1 \big\}$, with $\mu\{-1, \cdot\} \in \tilde{\Gamma}_\alpha$. Naturally, we might take $d > 2$. In this case, in seeking a result analogous to Theorem 1.1, we would seek to show that, for regularly varying laws $\mu$ whose decay is slow enough to ensure that the transition histories of distinct voters may be disjoint, the affiliation "white noise" in the time-slice $\big\{ t = \text{constant} \big\}$ at equilibrium is given by a fractional Gaussian field.

- Theorem 1.1 states that fractional simple random walk and fractional Brownian motion may be coupled to be close in $L^\infty$ on compact sets under rescaling of the discrete walk. One might seek to quantify this, investigating how quickly $\epsilon$ may be taken to 0 in the limit of high $n$ in (1.4), for some coupling $C_n$. The analogue for simple random walk and its convergence to Brownian motion is the celebrated Komlos-Major-Tusnady theorem [17], one case of which states that simple random walk may be coupled with asymptotic probability one to a Brownian motion with a uniform error in the first $n$ steps of at most a large constant multiple of $\log n$. Quantifying the rate of convergence could be useful if we would like to compare the stochastic difference equation driven by the fractional discrete process and its continuous analogue driven by fractional Brownian motion, the latter having been extensively studied. (See [5, 9, 19] for treatments of the stochastic calculus of fractional Brownian motion.)

# A    Market interpretation

This section was added in response to referee's question about whether there was a market-based "story" to go with our random walk model. We will not address the empirical question of whether our construction describes actual markets. The literature on market-increment autocorrelations is too vast (searches for *market momentum* or *mean reversion* turn up hundreds or thousands of articles per year) to summarize here. Instead, we offer a few very general observations in response to a simpler question: *if* an asset's price were described by one of our random walks (for some $\mu$), what kind of story would explain it? Could one reconcile the story with (some form of) market efficiency? The first story to spring to mind is the following:

**Story 1: Peer-influenced decisions.** *Many decisions involve the consultation (conscious or otherwise) of people who made a similar decision recently. Jeans or khakis? Buy or rent? State or private school? If the decision probability is a monotone linear function of previous decisions, one can represent this by having individuals copy (with some probability) a randomly chosen previous decision. The time elapsed since that decision is random (and a power law is not unreasonable). The cumulative number of decisions of one type, minus the number of the other type, is thus one of the random walks described in this paper.*

This story might explain why some of a firm's fundamentals (market share, revenue, etc.) would correspond to our model for some $\mu$. It does not explain momentum effects in asset prices, however, since one should be able to foresee these effects and price them in.

In classical finance, asset prices are martingales with respect to the *risk neutral* probability (assuming no interest, a point we return to later). This may in fact be taken as the definition of risk neutral probability. In models with $\pm 1$ increments, the risk neutral probability measure is that of the simple random walk. (Let us assume that the probability of an up step, conditioned on the past, is always strictly between zero and one.) Assuming no arbitrage, the price of a derivative is its expected value in this measure. Because of this simplicity, models with $\pm 1$ increments are especially natural to work with.[1]

The interesting question is the following: *why do risk neutral and true probability differ?*

**Note:** Discrepancies between true probability and risk neutral probability do not necessarily imply market inefficiency. However, they are most plausible if the discrepancy is small. (If a trader could get a 5000 percent return with probability .99 in one year, this would be hard to reconcile with market efficiency.) In our model, this suggests using a $\mu$ that decays slowly. In the classical capital asset pricing model (CAPM) these discrepancies are explained by the asset's correlation with so-called *systemic risk*: when the price goes down, the so-called *market portfolio* is generally likely to go down also; the demand for money in that scenario is greater because people are risk averse. In the simplest CAPM story, all investors hold some combination of risk free assets and the market portfolio.

**Story 2: Inhomogeneous market portfolio.** *Suppose that one population of traders tends to enter and leave the market for apparently irrational reasons. Another population of "savvy traders" holds a time inhomogenous "savvy market portfolio," including long and short positions that vary in time. This group makes money in expectation by holding this portfolio (twelve percent per year, say), while the other traders make less money in expectation.*

The behavior of the savvy traders in the above story makes perfect sense, but we have

---

[1] Any continuous model $X_t$ can be interpreted as having $\pm 1$ increments if one changes time parameterization: simply define times $t_k$ such that $X_{t_k}$ is an integer for each $k$ and $t_{k+1}$ is the first time after $t_k$ that a distinct integer is reached. Then $N_k := X_{t_k}$ has $\pm 1$ increments. One can even approximate a discontinuous jump process by using a probability measure in which $N_k$ is likely to sometimes go up or down several steps in a row.

not explained the actions of the "less savvy" group. Let us suppose that the asset price is 10 today and has gone up recently, so that it now has a .52 chance to be 11 tomorrow, and a .48 chance to be 9 tomorrow. Or suppose the asset price has dropped recently and now has a .48 to chance to be 11 tomorrow and a .52 to be 9. In each case, it is clear why a savvy trader would take the side of the bet with positive expectation. Under certain parameters, it is also reasonable to suppose that such investors are too risk averse or otherwise constrained to bid the price all the way up or down to tomorrow's expected value. But who is taking the negative-expectation side of the bet? One can think of many stories, but here is one that fits our model.

**Story 3: Anti-momentum trading.** *Less savvy traders tend to sell after the price goes up ("profit taking") and buy after the price goes down ("bargain hunting"). In other words, they systematically bet against the momentum strategy. Each less savvy trader decides whether to hold a stock during a period of time by looking at price changes in the recent past. The probability that the less savvy trader takes the "less savvy" position is a linear function of these prior changes. The amount this skews the price is approximately linear in the number of less savvy traders taking the anti-momentum strategy.*

This story suggests a rather paradoxical conclusion. Market returns may exhibit positive autocorrelation (momentum) precisely because the less savvy investors *think* or *feel* that they should exhibit negative autocorrelation (mean reversion). Indeed, the fact that the savvy investors tend to earn higher returns essentially *implies* that whatever strategy the less savvy investors employ tends to be wrong.

We note a couple of other (possibly complementary, possibly second-order) stories:

**Story 4: Variable interest rate.** *The presence of interest affects the definition of risk neutral probability. The interest-discounted asset price (not the asset price itself) is the martingale. Even if the rate of interest earned per time unit is constant or very slowly varying, it may become non-constant when time is parameterized by the number $k$ of integer price changes. Perhaps when the price has gone up recently, people trade more slowly (so that the interest per tick is higher). This leads to momentum affects when one parameterizes time by $k$.*

**Story 5: Non-liquidity/inefficiency.** *There is inefficiency, but the market is too small for the arbitrage opportunities to be very valuable. There may be times at which price history indicates a high likelihood of a rise or fall in future prices, but (for whatever reason) the volume of trading at these times is not sufficient to attract arbitrageurs.*

Finally, we remark that if a market were to exhibit momentum on the scale of days and mean reversion on the scale of several months, this could be modeled with a variant of our walks, where one samples $j$ according to $\mu$ and copies the increment $j$ steps previously (for some range of $j$ values) or the opposite of that increment (for another range of $j$ values). The stories described above make sense for these variants as well. However, we stress, in

conclusion, that the extent to which any of the above stories is, in actual markets, *correct* is an empirical matter well beyond the scope of this work.

# References

[1] Gideon Amir, Omer Angel and Benedek Valko. The TASEP speed process. arXiv:0811.3706v1.

[2] Richard Arratia. The motion of a tagged particle in the simple symmetric exclusion system on **Z**. *Ann. Probab.*, 11(2):362–373, 1983.

[3] Erhan Bayraktar, Ulrich Horst, and Ronnie Sircar. A limit theorem for financial markets with inert investors. *Math. Oper. Res.*, 31(4):789–810, 2006.

[4] C. Bender, T. Sottinen and E. Valkeila. Arbitrage with fractional Brownian motion? Theory Stoch. Process. 13 (2007), no. 1-2, 2334.

[5] Francesca Biagini, Yaozhong Hu, Bernt Oksendal and Tusheng Zhang. Stochastic calculus for fractional Brownian motion and applications. Springer-Verlag London Ltd., London, 2008.

[6] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1987.

[7] M. Bramson, S. Kalikow. Nonuniqueness in $g$ -functions. *Israel J. Math.* 84, 153-160, 1993.

[8] J.L. Doob. *Measure Theory*. Number 143 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1994.

[9] Tyrone E. Duncan, Yaozhong Hu and Bozenna Pasik-Duncan. Stochastic calculus for fractional Brownian motion. I. Theory. *SIAM J. Control Optim.*, 38(2):582–612 (electronic), 2000.

[10] Richard Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, second edition, 1996.

[11] Nathanaël Enriquez. A simple construction of the fractional Brownian motion. *Stochastic Process. Appl.*, 109(2):203–223, 2004.

[12] J. L. Geluk and L. de Haan. Stable probability distributions and their domains of attraction: a direct approach. *Probab. Math. Statist.*, 20(1, Acta Univ. Wratislav. No. 2246):169–188, 2000.

[13] Claudia Klüppelberg and Christoph Kühn. Fractional Brownian motion as a weak limit of Poisson shot noise processes—with applications to finance. *Stochastic Process. Appl.*, 113(2):333–351, 2004.

[14] S. Kalikow. Random Markov processes and uniform martingales. *Israel J. Math.* 71: 33-54, 1990.

[15] M. Keane, Strongly mixing $g$-measures. *Invent. Math.* 16, 309-324, 1972.

[16] A.N. Kolmogorov. Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *Dok. Akad. Nauk SSSR*, 26:115–118, 1940.

[17] J. Komlós, P. Major and G. Tusnády An approximation of partial sums of independent RV's and the sample DF. I. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 32 (1975), 111–131.

[18] B. Mandelbrot and J. Van Ness Fractioanl Brownian motions, fractional noises and applications. SIAM Review 10 (1968) 422–437.

[19] Yuliya S. Mishura. Stochastic calculus for fractional Brownian motion and related processes. Springer-Verlag, Berlin, 2008.

[20] Robin Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 4:1–79 (electronic), 2007.

[21] L. C. G. Rogers. Arbitrage with fractional Brownian motion. Math. Finance 7 (1997), no. 1, 95105.

[22] T. Sottinen. Fractional Brownian motion, random walks and binary market models. Finance Stoch. 5 (2001), no. 3, 343355.