CONVERGENCE IN A MULTIDIMENSIONAL RANDOMIZED KEYNESIAN BEAUTY CONTEST

MICHAEL GRINFELD,* *University of Strathclyde*

STANISLAV VOLKOV,** *Lund University and University of Bristol*

ANDREW R. WADE,*** *Durham University*

Abstract

We study the asymptotics of a Markovian system of $N \geq 3$ particles in $[0, 1]^d$ in which, at each step in discrete time, the particle farthest from the current centre of mass is removed and replaced by an independent $U[0, 1]^d$ random particle. We show that the limiting configuration contains $N - 1$ coincident particles at a random location $\xi_N \in [0, 1]^d$. A key tool in the analysis is a Lyapunov function based on the squared radius of gyration (sum of squared distances) of the points. For $d = 1$, we give additional results on the distribution of the limit ξ_N, showing, among other things, that it gives positive probability to any nonempty interval subset of $[0, 1]$, and giving a reasonably explicit description in the smallest nontrivial case, $N = 3$.

Keywords: Keynesian beauty contest; radius of gyration; rank-driven process; sum of squared distances

1. Introduction, model, and results

In a *Keynesian beauty contest*, N players each guess a number, the winner being the player whose guess is closest to the mean of all the N guesses; the name marks Keynes's discussion of 'those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole' [7, Chapter 12, Section V]. Moulin [10, page 72] formalized a version of the game played on a real interval, the '*p*-beauty contest', in which the target is p ($p > 0$) times the mean value. See, e.g. [2] and the references therein for some recent work on game-theoretic aspects of such 'contests' in economics.

In this paper we study a stochastic process based on an iterated version of the game, in which players *randomly* choose a value in $[0, 1]$, and at each step the worst performer (that is, the player whose guess is farthest from the mean) is replaced by a new player; each player's guess is fixed as soon as they enter the game, so a single new random value enters the system at each step. Analysis of this model was posed as an open problem in [4, page 390]. The natural setting

for our techniques is in fact a generalization in which the values live in $[0, 1]^d$ and the target is the barycentre (centre of mass) of the values. We now formally describe the model and state our main results.

Let $d \in \mathbb{N} := \{1, 2, \ldots\}$. We use the notation $\mathcal{X}_n = (x_1, x_2, \ldots, x_n)$ for a vector of n points $x_i \in \mathbb{R}^d$. We write $\mu_n(\mathcal{X}_n) := n^{-1} \sum_{i=1}^n x_i$ for the barycentre of $\mathcal{X}_n$, and $\| \cdot \|$ for the Euclidean norm on $\mathbb{R}^d$. Let $\text{ord}(\mathcal{X}_n) = (x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ denote the *barycentric order statistics* of $x_1, \ldots, x_n$, so that $\|x_{(1)} - \mu_n(\mathcal{X}_n)\| \leq \|x_{(2)} - \mu_n(\mathcal{X}_n)\| \leq \cdots \leq \|x_{(n)} - \mu_n(\mathcal{X}_n)\|$; any ties are broken randomly. We call $\mathcal{X}_n^* := x_{(n)}$ the *extreme* point of $\mathcal{X}_n$, a point of $x_1, \ldots, x_n$ farthest from the barycentre. We define the *core* of $\mathcal{X}_n$ as $\mathcal{X}_n' := (x_{(1)}, \ldots, x_{(n-1)})$, the vector of $x_1, \ldots, x_n$ with the extreme point removed.

The Markovian model that we study is defined as follows. Start with $X_1(0), \ldots, X_N(0)$ distinct points in $[0, 1]^d$, and write $\mathcal{X}_N(0) := (X_{(1)}(0), \ldots, X_{(N)}(0))$ for the corresponding ordered vector. One possibility is to start with a *uniform random* initial configuration, by taking $X_1(0), \ldots, X_N(0)$ to be independent $U[0, 1]^d$ random variables; here and elsewhere, $U[0, 1]^d$ denotes the uniform distribution on $[0, 1]^d$. In this uniform random initialization, all N points are indeed distinct with probability 1. Given $\mathcal{X}_N(t)$, replace $\mathcal{X}_N^*(t) = X_{(N)}(t)$ by an independent $U[0, 1]^d$ random variable U_{t+1}, so that $\mathcal{X}_N(t + 1) = \text{ord}(X_{(1)}(t), \ldots, X_{(N-1)}(t), U_{t+1})$.

The interesting case is when $N \geq 3$: the $N = 1$ case is trivial, and the $N = 2$ case is also uninteresting since at each step either point is replaced with probability $\frac{1}{2}$ by a $U[0, 1]$ variable, so that, regardless of the initial configuration, after a finite number of steps we will have two independent $U[0, 1]$ points. Our main result, Theorem 1.1, shows that, for $N \geq 3$, all but the most extreme point of the configuration converge to a common limit.

Theorem 1.1. *Let $d \in \mathbb{N}$ and $N \geq 3$. Let $\mathcal{X}_N(0)$ consist of N distinct points in $[0, 1]^d$. There exists a random $\xi_N := \xi_N(\mathcal{X}_N(0)) \in [0, 1]^d$ such that*

$$\mathcal{X}_N'(t) \to (\xi_N, \xi_N, \ldots, \xi_N) \quad \text{and} \quad \mathcal{X}_N^*(t) - U_t \to 0 \quad \text{almost surely (a.s.)} \tag{1.1}$$

as $t \to \infty$. In particular, for $U \sim U[0, 1]^d$, as $t \to \infty$,

$$\mathcal{X}_N(t) \xrightarrow{\text{D}} (\xi_N, \xi_N, \ldots, \xi_N, U).$$

Remark 1.1. Under the conditions of Theorem 1.1, despite the fact that $\mathcal{X}_N^*(t) - U_t \to 0$ a.s., we will see below that $\mathcal{X}_N^*(t) \neq U_t$ infinitely often a.s.

Theorem 1.1 is proved in Section 2. Then, Section 3 is devoted to the one-dimensional case, where we obtain various additional results on the limit ξ_N. Finally, in Appendices A and B we collect some results on uniform spacings and continuity of distributional fixed points that we use in parts of the analysis in Section 3.

2. Proof of convergence

Intuitively, the evolution of the process is as follows. If, on replacement of the extreme point, the new point is the next extreme point (measured with respect to the new centre of mass), then the core is unchanged. However, if the new point is not extreme, it typically penetrates the core significantly, while a more extreme point is thrown out of the core, reducing the size of the core in some sense (we give a precise statement below). Tracking the evolution of the core, by following its centre of mass, we see increasingly long periods of inactivity, since, as the size of the core decreases, changes occur less often, and, moreover, the magnitude of the changes decreases in step with the size of the core. The dynamics are nontrivial, but bear

some resemblance to random walks with decreasing steps (see, e.g. [3], [8], and the references therein) as well as processes with reinforcement, such as the Pólya urn (see, e.g. [12] for a survey). The process is also reminiscent in some ways of iterated interval division [6] or sequential adsorption with interaction [11].

Our analysis will rest on a 'Lyapunov function' for the process, that is, a function of the configuration that possesses pertinent asymptotic properties. We may initially hope, for example, that the diameter of the point set $\mathcal{X}_N(t)$ would decrease over time, but this cannot be the case because the newly added point can be anywhere in $[0, 1]^d$. What then about the diameter of $\mathcal{X}'_N(t)$, for which the extreme point is ignored? We will show later in this section that this quantity is in fact well behaved, but we have to argue somewhat indirectly: the diameter of $\mathcal{X}'_N(t)$ can increase (at least for large enough N; see Remark 2.2 below). However, there *is* a monotone decreasing function associated with the process, based on the sum of squared distances of a configuration, which we will use as our Lyapunov function. To define this function (see (2.9) below), we need some notation.

For $n \in \mathbb{N}$ and $\mathcal{X}_n = (x_1, x_2, \dots, x_n) \in (\mathbb{R}^d)^n$, write

$$G_n(\mathcal{X}_n) := G_n(x_1, \dots, x_n) := n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{i-1} \|x_i - x_j\|^2 = \sum_{i=1}^{n} \|x_i - \mu_n(\mathcal{X}_n)\|^2; \quad (2.1)$$

a detailed proof of the (elementary) final equality in (2.1) may be found on pages 95–96 of [5] for example. We remark that $(1/n)G_n$ is the squared *radius of gyration* of $x_1, \dots, x_n$; see, e.g. [5, page 95]. Note also that calculus verifies the useful variational formula

$$G_n(x_1, \dots, x_n) = \inf_{y \in \mathbb{R}^d} \sum_{i=1}^{n} \|x_i - y\|^2. \quad (2.2)$$

For $n \geq 2$, define

$$F_n(\mathcal{X}_n) := F_n(x_1, \dots, x_n) := G_{n-1}(\mathcal{X}'_n) = G_{n-1}(x_{(1)}, \dots, x_{(n-1)}).$$

Lemma 2.1. *Let $n \geq 2$ and $\mathcal{X}_n = (x_1, x_2, \dots, x_n) \in (\mathbb{R}^d)^n$. Then, for any $x \in \mathbb{R}^d$,*

$$F_n(x_{(1)}, \dots, x_{(n-1)}, x) \leq F_n(\mathcal{X}_n).$$

Proof. For ease of notation, we write simply $(x_1, \dots, x_n)$ for $(x_{(1)}, \dots, x_{(n)})$, i.e. we relabel so that x_j is the jth closest point to $\mu_n(\mathcal{X}_n)$. Then $\mathcal{X}_n^* = x_n$, $\mathcal{X}'_n = (x_1, \dots, x_{n-1})$, and

$$F_{\text{old}} := F_n(\mathcal{X}_n) = G_{n-1}(x_1, \dots, x_{n-1}) = \sum_{i=1}^{n-1} \|x_i - \mu'_{\text{old}}\|^2, \quad (2.3)$$

where $\mu'_{\text{old}} := \mu_{n-1}(\mathcal{X}'_n)$. We compare F_{old} to F_n evaluated on the set of points obtained by removing x_n and replacing it with some $x \in \mathbb{R}^d$.

Write $y := \{x_1, \dots, x_{n-1}, x\}^*$ for the new extreme point. Then

$$F_{\text{new}} := F_n(x_1, \dots, x_{n-1}, x) = \sum_{i=1}^{n-1} \|x_i - \mu'_{\text{new}}\|^2 + \|x - \mu'_{\text{new}}\|^2 - \|y - \mu'_{\text{new}}\|^2, \quad (2.4)$$

where

$$\mu'_{\text{new}} := \frac{1}{n-1}\left(\sum_{i=1}^{n-1} x_i + x - y\right) = \mu'_{\text{old}} + \frac{x-y}{n-1}. \tag{2.5}$$

Define $\mu_{\text{new}} := \mu_n(x_1, \ldots, x_{n-1}, x)$, so

$$\mu'_{\text{new}} = \frac{n\mu_{\text{new}}}{n-1} - \frac{y}{n-1}. \tag{2.6}$$

From (2.3), (2.4), and (2.5), we obtain

$$F_{\text{new}} - F_{\text{old}} = \sum_{i=1}^{n-1}(\|x_i - \mu'_{\text{new}}\|^2 - \|x_i - \mu'_{\text{old}}\|^2) + \|x - \mu'_{\text{new}}\|^2 - \|y - \mu'_{\text{new}}\|^2. \tag{2.7}$$

For the sum on the right-hand side of (2.7), we have

$$\begin{aligned}
\sum_{i=1}^{n-1}(\|x_i - \mu'_{\text{new}}\|^2 - \|x_i - \mu'_{\text{old}}\|^2) &= \sum_{i=1}^{n-1}(2x_i(\mu'_{\text{old}} - \mu'_{\text{new}}) + \|\mu'_{\text{new}}\|^2 - \|\mu'_{\text{old}}\|^2) \\
&= (n-1)(2\mu'_{\text{old}}(\mu'_{\text{old}} - \mu'_{\text{new}}) + \|\mu'_{\text{new}}\|^2 - \|\mu'_{\text{old}}\|^2) \\
&= (n-1)(\|\mu'_{\text{old}}\|^2 - 2(\mu'_{\text{old}}\mu'_{\text{new}}) + \|\mu'_{\text{new}}\|^2).
\end{aligned}$$

Simplifying this last expression and substituting back into (2.7) gives

$$F_{\text{new}} - F_{\text{old}} = (n-1)\|\mu'_{\text{old}} - \mu'_{\text{new}}\|^2 + \|x - \mu'_{\text{new}}\|^2 - \|y - \mu'_{\text{new}}\|^2.$$

Thus, using (2.5) and then (2.6),

$$\begin{aligned}
F_{\text{new}} - F_{\text{old}} &= \frac{\|x-y\|^2}{n-1} + \|x\|^2 - \|y\|^2 - 2\mu'_{\text{new}}(x-y) \\
&= \frac{\|x\|^2 + \|y\|^2 - 2xy}{n-1} + \|x\|^2 - \|y\|^2 - 2\left(\frac{n\mu_{\text{new}}}{n-1} - \frac{y}{n-1}\right)(x-y).
\end{aligned}$$

Hence, we conclude that

$$\begin{aligned}
F_{\text{new}} - F_{\text{old}} &= \frac{n}{n-1}(\|x\|^2 - \|y\|^2 - 2\mu_{\text{new}}(x-y)) \\
&= \frac{n}{n-1}(\|x - \mu_{\text{new}}\|^2 - \|y - \mu_{\text{new}}\|^2) \\
&\leq 0, \tag{2.8}
\end{aligned}$$

since y is, by definition, the farthest point from μ_{new}.

Now we return to the stochastic model. Define

$$F(t) := F_N(\mathcal{X}_N(t)). \tag{2.9}$$

Lemma 2.1 has the following immediate consequence.

Corollary 2.1. *Let $N \geq 2$. Then $F(t + 1) \leq F(t)$.*

Corollary 2.1 shows that our Lyapunov function $F(t)$ is nonincreasing; later we show that $F(t) \to 0$ a.s. (see Lemma 2.4 below). First, we need to relate $F(t)$ to the *diameter* of the point set $\mathcal{X}'_N(t)$. For $n \geq 2$ and $x_1, \ldots, x_n \in \mathbb{R}^d$, write

$$D_n(x_1, \ldots, x_n) := \max_{1 \leq i, j \leq n} \|x_i - x_j\|.$$

Lemma 2.2. *Let $n \geq 2$ and $x_1, \ldots, x_n \in \mathbb{R}^d$. Then*

$$\tfrac{1}{2} D_n(x_1, \ldots, x_n)^2 \leq G_n(x_1, \ldots, x_n) \leq \tfrac{1}{2}(n - 1) D_n(x_1, \ldots, x_n)^2.$$

Remark 2.1. The lower bound in Lemma 2.2 is sharp, and is attained by collinear configurations with two diametrically opposed points x_i, x_j and all the other $n - 2$ points at the midpoint $\mu_2(x_i, x_j) = \mu_n(x_1, \ldots, x_n)$. The upper bound in Lemma 2.2 is not, in general, sharp; determining the sharp upper bound is a nontrivial problem. The bound $G_n(x_1, \ldots, x_n) \leq (n/2)(d/(d + 1)) D_n(x_1, \ldots, x_n)^2$ from [16] is also not always sharp. Witsenhausen [16] conjectured that the maximum is attained if and only if the points are distributed as evenly as possible among the vertices of a regular d-dimensional simplex of edge length $D_n(x_1, \ldots, x_n)$; this was proved relatively recently [1], [15].

Proof of Lemma 2.2. Fix $x_1, \ldots, x_n \in \mathbb{R}^d$. For ease of notation, write $\mu = \mu_n(x_1, \ldots, x_n)$. First we prove the lower bound. For $n \geq 2$, using the second form of G_n in (2.1),

$$G_n(x_1, \ldots, x_n) = \sum_{i=1}^{n} \|x_i - \mu\|^2 \geq \|x_i - \mu\|^2 + \|x_j - \mu\|^2$$

for (x_i, x_j) a diameter, i.e. $D_n(x_1, \ldots, x_n) = \|x_i - x_j\|$. By the $n = 2$ case of (2.2),

$$\|x_i - \mu\|^2 + \|x_j - \mu\|^2 \geq 2\|x_i - \mu_2(x_i, x_j)\|^2 = \tfrac{1}{2}\|x_i - x_j\|^2.$$

This gives the lower bound. For the upper bound, from the first form of G_n in (2.1),

$$G_n(x_1, \ldots, x_n) \leq \frac{1}{n} \sum_{i=1}^{n} (i - 1) D_n(x_1, \ldots, x_n)^2,$$

by the definition of D_n, which yields the result.

Let $D(t) := D_{N-1}(\mathcal{X}'_N(t))$.

Remark 2.2. By Lemma 2.2 (or (2.1)), $G_2(\mathcal{X}'_3(t)) = \tfrac{1}{2} D_2(\mathcal{X}'_3(t))^2$, so when $N = 3$, Lemma 2.1 implies that $D(t + 1) \leq D(t)$ a.s. as well. If $d = 1$, it can be shown that $D(t)$ is nonincreasing also when $N = 4$. In general, however, $D(t)$ can increase.

Let $\mathcal{F}_t := \sigma(\mathcal{X}_N(0), \mathcal{X}_N(1), \ldots, \mathcal{X}_N(t))$, the σ-algebra generated by the process up to time t. Let $B(x; r)$ denote the closed Euclidean d-ball with centre $x \in \mathbb{R}^d$ and radius $r > 0$. Define the events

$$A_{t+1} := \{U_{t+1} \in B(\mu_{N-1}(\mathcal{X}'_N(t)); 3D(t))\},$$
$$A'_{t+1} := \{U_{t+1} \in B(\mu_{N-1}(\mathcal{X}'_N(t)); \tfrac{1}{4}D(t))\}.$$

Lemma 2.3. *There is an absolute constant $\gamma > 0$ for which, for all $N \geq 3$ and all t,*

$$A'_{t+1} \subseteq \{F(t+1) - F(t) \leq -\gamma N^{-1} F(t)\} \subseteq \{F(t+1) - F(t) < 0\} \subseteq A_{t+1}. \quad (2.10)$$

Moreover, there exist constants $c > 0$ and $C < \infty$, depending only on d, for which, for all $N \geq 3$ and all t, a.s.,

$$\mathbb{P}[F(t+1) - F(t) \leq -\gamma N^{-1} F(t) \mid \mathcal{F}_t] \geq \mathbb{P}[A'_{t+1} \mid \mathcal{F}_t] \geq c N^{-d/2} (F(t))^{d/2}, \quad (2.11)$$

$$\mathbb{P}[F(t+1) - F(t) < 0 \mid \mathcal{F}_t] \leq \mathbb{P}[A_{t+1} \mid \mathcal{F}_t] \leq C (F(t))^{d/2}. \quad (2.12)$$

Proof. Here and elsewhere, $\mathbf{1}(A)$ denotes the indicator random variable of the event A. For simplicity, we write $X_1, \ldots, X_{N-1}$ instead of $X_{(1)}(t), \ldots, X_{(N-1)}(t)$ and D instead of $D(t) = D_{N-1}(X_1, \ldots, X_{N-1})$. By the definition of D, there exists some $i \in \{1, \ldots, N-1\}$ such that $\|\mu'_{\text{old}} - X_i\| \geq \frac{1}{2} D$, where $\mu'_{\text{old}} = \mu_{N-1}(X_1, \ldots, X_{N-1})$. Given $\mathcal{F}_t$, the event A'_{t+1} that the new point $U := U_{t+1}$ falls in $B(\mu'_{\text{old}}; \frac{1}{4} D)$ has probability bounded below by $\theta_d D^d$, where $\theta_d > 0$ depends only on d. Let $\mu_{\text{new}} := \mu_N(X_1, \ldots, X_{N-1}, U)$. Suppose that A'_{t+1} occurs. Then,

$$\|\mu_{\text{new}} - \mu'_{\text{old}}\| = \frac{1}{N} \|U - \mu'_{\text{old}}\| \leq \frac{D}{4N} \leq \frac{D}{12}, \quad (2.13)$$

since $N \geq 3$. Hence, by (2.13) and the triangle inequality,

$$\|U - \mu_{\text{new}}\| \leq \|U - \mu'_{\text{old}}\| + \|\mu_{\text{new}} - \mu'_{\text{old}}\| \leq \frac{D}{4} + \frac{D}{12} = \frac{4D}{12}. \quad (2.14)$$

On the other hand, by another application of the triangle inequality and (2.13),

$$\|\mu_{\text{new}} - X_i\| \geq \|\mu'_{\text{old}} - X_i\| - \|\mu_{\text{new}} - \mu'_{\text{old}}\| \geq \frac{D}{2} - \frac{D}{12} = \frac{5D}{12}.$$

Then, by definition, the extreme point $Y := \{X_1, \ldots, X_{N-1}, U\}^*$ satisfies

$$\|Y - \mu_{\text{new}}\| \geq \|\mu_{\text{new}} - X_i\| \geq \frac{5D}{12}. \quad (2.15)$$

Hence, from the $x = U$ case of (2.8) with the bounds (2.14) and (2.15), we conclude that

$$F(t+1) - F(t) \leq \frac{N}{N-1} \left(\left(\frac{4D}{12}\right)^2 - \left(\frac{5D}{12}\right)^2 \right) \mathbf{1}(A'_{t+1}) \leq -\frac{9}{144} D^2 \mathbf{1}(A'_{t+1}), \quad (2.16)$$

for all $N \geq 3$; the first inclusion in (2.10) follows (with $\gamma = \frac{9}{72}$) from (2.16) together with the fact that, by the second inequality in Lemma 2.2, $D^2 \geq 2N^{-1} F(t)$. This in turn implies (2.11), using the fact that $\mathbb{P}[A'_{t+1} \mid \mathcal{F}_t] \geq \theta_d D^d$.

Next we consider the event A_{t+1}. Using the same notation as above, we have

$$\|\mu_{\text{new}} - U\| \geq \|\mu'_{\text{old}} - U\| - \|\mu_{\text{new}} - \mu'_{\text{old}}\| = \left(1 - \frac{1}{N}\right) \|\mu'_{\text{old}} - U\|,$$

by the equality in (2.13). Also, for any $k \in \{1, \ldots, N-1\}$,

$$\|\mu_{\text{new}} - X_k\| \leq \|\mu'_{\text{old}} - X_k\| + \|\mu'_{\text{old}} - \mu_{\text{new}}\| \leq D + \frac{1}{N} \|\mu'_{\text{old}} - U\|,$$

by (2.13) again. Combining these estimates we obtain, for any $k \in \{1, \ldots, N-1\}$,

$$\|\mu_{\text{new}} - U\| - \|\mu_{\text{new}} - X_k\| \geq \left(1 - \frac{2}{N}\right)\|\mu'_{\text{old}} - U\| - D \geq \frac{1}{3}\|\mu'_{\text{old}} - U\| - D$$

for $N \geq 3$. So, in particular, $\|\mu_{\text{new}} - U\| > \|\mu_{\text{new}} - X_k\|$ for all $k \in \{1, \ldots, N-1\}$ provided $\|\mu'_{\text{old}} - U\| > 3D$, i.e. $U \notin B(\mu'_{\text{old}}; 3D)$. In this case, U is the extreme point among $U, X_1, \ldots, X_{N-1}$, i.e.

$$A^c_{t+1} \subseteq \{\mathcal{X}^*_N(t+1) = U_{t+1}\}. \tag{2.17}$$

In particular, on A^c_{t+1}, $F(t+1) = F(t)$, and $F(t+1) < F(t)$ only if A_{t+1} occurs, giving the final inclusion in (2.10). Since $\mathbb{P}[A_{t+1} \mid \mathcal{F}_t]$ is bounded above by $C_d D^d$ for a constant $C_d < \infty$ depending only on d, (2.12) follows from the first inequality in Lemma 2.2. This completes the proof. $\quad\blacksquare$

Lemma 2.4. *Suppose that $N \geq 3$. Then, as $t \to \infty$, $F(t) \to 0$ a.s. and in L^2.*

Proof. Let $\varepsilon > 0$, and let $\sigma := \min\{t \in \mathbb{Z}_+ : F(t) \leq \varepsilon\}$, where $\mathbb{Z}_+ := \{0, 1, 2, \ldots\}$. Then, by (2.11), there exists $\delta > 0$ (depending on d, ε, and N) such that, a.s.,

$$\mathbb{P}[F(t+1) - F(t) \leq -\delta \mid \mathcal{F}_t] \geq \delta \mathbf{1}\{t < \sigma\}.$$

Hence, since $F(t+1) - F(t) \leq 0$ a.s. by Corollary 2.1,

$$\mathbb{E}[F(t+1) - F(t) \mid \mathcal{F}_t] \leq -\delta^2 \mathbf{1}\{t < \sigma\}. \tag{2.18}$$

By Corollary 2.1, $F(t)$ is nonnegative and nonincreasing, and, hence, $F(t)$ converges a.s. as $t \to \infty$ to some nonnegative limit $F(\infty)$; the convergence also holds in L^2 since $F(t)$ is uniformly bounded. In particular, $\mathbb{E}[F(t)] \to \mathbb{E}[F(\infty)]$. So taking expectations in (2.18) and letting $t \to \infty$ we obtain

$$\limsup_{t \to \infty} \delta^2 \mathbb{P}[\sigma > t] \leq 0,$$

which implies that $\mathbb{P}[\sigma > t] \to 0$ as $t \to \infty$. Thus, $\sigma < \infty$ a.s., which together with the monotonicity of $F(t)$ (Corollary 2.1) implies that $F(t) \leq \varepsilon$ for all sufficiently large t. Since $\varepsilon > 0$ was arbitrary, the result follows. $\quad\blacksquare$

Recall the definitions of A_t and A'_t from before Lemma 2.3. Define $(\mathcal{F}_t)$ stopping times $\tau_0 := 0$ and, for $n \in \mathbb{N}$, $\tau_n := \min\{t > \tau_{n-1} : A_t \text{ occurs}\}$. Then $F(t) < F(t-1)$ can only occur if $t = \tau_n$ for some n. Since $\mathbb{P}[A_{t+1} \mid \mathcal{F}_t]$ is bounded below by $cF(t)^{d/2}$ for some $c > 0$, it is not hard to see that, provided $F(0) > 0$, A_t occurs infinitely often, a.s. Indeed, suppose that $F(0) > 0$ and A_t occurs only finitely often. Then $F(t)$ has a nonzero limit. On the other hand, $\sum_t cF(t)^{d/2} \leq \sum_t \mathbb{P}[A_{t+1} \mid \mathcal{F}_t] < \infty$, by Lévy's extension of the Borel–Cantelli lemma, so that $F(t) \to 0$ a.s., which is a contradiction. Hence, $\tau_n < \infty$ a.s. for all n.

Lemma 2.5. *Let $N \geq 3$. There exists $\alpha > 0$ such that, a.s., $D(\tau_n) \leq e^{-\alpha n}$ for all sufficiently large n.*

Proof. It follows from (2.16) and the second inequality in Lemma 2.2 that

$$F(\tau_n) - F(\tau_n - 1) \leq -\delta F(\tau_n - 1)\mathbf{1}(A'_{\tau_n})$$

for some $\delta > 0$. Note also that, by definition of the stopping times τ_n, $F(\tau_n - 1) = F(\tau_{n-1})$. Hence,

$$\mathbb{P}[F(\tau_n) - F(\tau_{n-1}) \leq -\delta F(\tau_{n-1}) \mid \mathcal{F}_{\tau_{n-1}}] \geq \mathbb{P}[A'_{\tau_n} \mid \mathcal{F}_{\tau_{n-1}}] \geq \delta,$$

taking $\delta > 0$ small enough, since, using the fact that $\mathbf{1}(A_{\tau_n}) = 1$ a.s.,

$$\mathbb{P}[A'_{\tau_n} \mid \mathcal{F}_{\tau_{n-1}}] = \mathbb{E}[\mathbb{P}[A'_{\tau_n} \mid \mathcal{F}_{\tau_n}]\mathbf{1}(A_{\tau_n}) \mid \mathcal{F}_{\tau_{n-1}}] = \mathbb{E}[\mathbb{P}[A'_{\tau_n} \mid A_{\tau_n}] \mid \mathcal{F}_{\tau_{n-1}}],$$

where, by the definitions of A_t and A'_t, $\mathbb{P}[A'_{\tau_n} \mid A_{\tau_n}]$ is uniformly positive. Since (by Corollary 2.1) $F(t+1) - F(t) \leq 0$ a.s., it follows that

$$\mathbb{E}[F(\tau_n) - F(\tau_{n-1}) \mid \mathcal{F}_{\tau_{n-1}}] \leq -\delta^2 F(\tau_{n-1}).$$

Taking expectations, we obtain $\mathbb{E}[F(\tau_n)] \leq (1-\delta^2)\mathbb{E}[F(\tau_{n-1})]$, which implies that $\mathbb{E}[F(\tau_n)] = O(e^{-cn})$ for some $c > 0$ depending on δ. Then, by Markov's inequality, $\mathbb{P}[F(\tau_n) \geq e^{-cn/2}] = O(e^{-cn/2})$, which implies that $F(\tau_n) = O(e^{-cn/2})$ a.s. by the Borel–Cantelli lemma. Then the first inequality in Lemma 2.2 gives the result.

Remark 2.3. The proof of Lemma 2.5 shows that $\mathbb{P}[A'_{\tau_n} \mid \mathcal{F}_{\tau_{n-1}}]$ is uniformly positive, so Lévy's version of the Borel–Cantelli lemma, with the fact that $\tau_n < \infty$ a.s. for all n, shows that A'_t occurs for infinitely many t a.s. With the proof of Lemma 2.3, this shows that $\mathcal{X}^*_N(t) \neq U_t$ infinitely often, as claimed in Remark 1.1.

We are almost ready to complete the proof of Theorem 1.1. We state the main step in the remaining argument as the first part of the next lemma, while the second part of the lemma we will need in Section 3.3 below. For $\varepsilon > 0$, define the stopping time $\nu_\varepsilon := \min\{t \in \mathbb{N} : F(t) < \varepsilon^2\}$; for any $\varepsilon > 0$, $\nu_\varepsilon < \infty$ a.s. by Lemma 2.4.

Lemma 2.6. *Let $N \geq 3$. Then there exists $\xi_N \in [0, 1]^d$ such that $\mu_{N-1}(\mathcal{X}'_N(t)) \to \xi_N$ a.s. and in L^2 as $t \to \infty$. Moreover, there exists an absolute constant C such that, for any $\varepsilon > 0$, and any $t_0 \in \mathbb{N}$, on $\{\nu_\varepsilon \leq t_0\}$, a.s.,*

$$\mathbb{E}\left[\max_{t \geq t_0} \|\mu_{N-1}(\mathcal{X}'_N(t)) - \mu_{N-1}(\mathcal{X}'_N(t_0))\| \mid \mathcal{F}_{t_0}\right] \leq C\varepsilon.$$

Proof. Let $\mu'(t) := \mu_{N-1}(\mathcal{X}'_N(t))$. Observe that, for $N \geq 3$, $\mathcal{X}'_N(t)$ and $\mathcal{X}'_N(t-1)$ have at least one point in common; choose one such point, and call it $Z(t)$. Then, $\mu'(t) \in \text{hull } \mathcal{X}'_N(t) \subseteq \text{hull } \mathcal{X}_N(t)$, where hull $\mathcal{X}$ denotes the convex hull of the point set $\mathcal{X}$. So, $\|Z(t) - \mu'(t)\| \leq D(t)$. Similarly, $\|Z(t) - \mu'(t-1)\| \leq D(t-1)$. By the definition of τ_n, $\mu'(t) = \mu'(t-1)$ and $D(t) = D(t-1)$ unless $t = \tau_n$ for some n, in which case $\mu'(\tau_n - 1) = \mu'(\tau_{n-1})$ and $D(\tau_n - 1) = D(\tau_{n-1})$. Hence,

$$\sum_{t \geq 1} \|\mu'(t) - \mu'(t-1)\| = \sum_{n \geq 1} \|\mu'(\tau_n) - \mu'(\tau_{n-1})\|$$
$$\leq \sum_{n \geq 1} (\|\mu'(\tau_n) - Z(\tau_n)\| + \|\mu'(\tau_{n-1}) - Z(\tau_n)\|), \qquad (2.19)$$

by the triangle inequality. Then the preceding remarks imply that

$$\sum_{t \geq 1} \|\mu'(t) - \mu'(t-1)\| \leq \sum_{n \geq 1} (D(\tau_n) + D(\tau_{n-1})) < \infty \quad \text{a.s.}$$

by Lemma 2.5. Hence, there is some (random) $\xi_N \in [0, 1]^d$ for which $\mu'(t) \to \xi_N$ a.s. as $t \to \infty$, and L^2 convergence follows by the bounded convergence theorem.

For the final statement in the lemma, we use a variation of the preceding argument. Let $M := \max\{n \in \mathbb{Z}_+ : \tau_n \le t_0\}$. Then, $F(t_0) = F(\tau_M)$ and $\mu'(\tau_M) = \mu'(t_0)$, so that on $\{v_\varepsilon \le t_0\}$ we have $\{v_\varepsilon \le \tau_M\}$ as well. Hence (by Corollary 2.1), $F(\tau_M) < \varepsilon^2$. A similar argument to that used in the proof of Lemma 2.5 shows that, for $m \ge 0$,

$$\mathbb{E}[F(\tau_{M+m}) \mid \mathcal{F}_{t_0}] \le \mathrm{e}^{-cm}\mathbb{E}[F(\tau_M) \mid \mathcal{F}_{t_0}] \le \varepsilon^2 \mathrm{e}^{-cm}$$

on $\{v_\varepsilon \le t_0\}$, where $c > 0$ depends on N but not on m or ε. Thus, by Lemma 2.2, on $\{v_\varepsilon \le t_0\}$, $\mathbb{E}[D(\tau_{M+m})^2 \mid \mathcal{F}_{t_0}] \le 2\varepsilon^2 \mathrm{e}^{-cm}$. Also, similarly to (2.19),

$$\max_{t \ge \tau_M} \|\mu'(t) - \mu'(\tau_M)\|^2 \le \sum_{t \ge \tau_M} \|\mu'(t) - \mu'(t-1)\|^2 \le \sum_{m \ge 1} (D(\tau_{M+m}) + D(\tau_{M+m-1}))^2.$$

Taking expectations and using the Cauchy–Schwarz inequality, we obtain, on $\{v_\varepsilon \le t_0\}$,

$$\mathbb{E}\Big[\max_{t \ge t_0} \|\mu'(t) - \mu'(t_0)\|^2 \mid \mathcal{F}_{t_0}\Big] = \mathbb{E}\Big[\max_{t \ge \tau_M} \|\mu'(t) - \mu'(\tau_M)\|^2 \mid \mathcal{F}_{t_0}\Big] \le 8\varepsilon^2 \mathrm{e}^c \sum_{m \ge 1} \mathrm{e}^{-cm},$$

a constant times ε^2. Jensen's inequality now gives the result, with $C^2 = 8/(1 - \mathrm{e}^{-c})$.

Proof of Theorem 1.1. Again, let $\mu'(t) := \mu_{N-1}(\mathcal{X}'_N(t))$. It follows from Lemma 2.6 that $\mu'(t) \to \xi_N$ a.s. Now, for any $j \in \{1, \ldots, N-1\}$, by the triangle inequality,

$$\|X_{(j)}(t) - \xi_N\| \le \|X_{(j)}(t) - \mu'(t)\| + \|\mu'(t) - \xi_N\| \le D(t) + \|\mu'(t) - \xi_N\|,$$

which tends to 0 a.s. as $t \to \infty$, since $D(t) \to 0$ a.s. by Lemma 2.5. This establishes the first statement in (1.1). Moreover, by (2.17), $\mathcal{X}^*_N(t+1) \ne U_{t+1}$ only if A_{t+1} occurs. On A_{t+1}, $\mathcal{X}^*_N(t+1)$ is one of the points of $\mathcal{X}'_N(t)$, and so $\|\mathcal{X}^*_N(t+1) - \mu'(t)\| \le D(t)$. In addition, on A_{t+1}, we have $\|U_{t+1} - \mu'(t)\| \le 3D(t)$. So, by the triangle inequality,

$$\|\mathcal{X}^*_N(t+1) - U_{t+1}\| \le 4D(t)\mathbf{1}(A_{t+1}),$$

which tends to 0 a.s., again by Lemma 2.5. This gives the final part of (1.1).

3. The limit distribution in one dimension

3.1. Overview and simulations

Throughout this section we restrict attention to $d = 1$. Of interest is the distribution of the limit ξ_N in (1.1), and its behaviour as $N \to \infty$. Simulations suggest that ξ_N is highly dependent on the initial configuration. In Figure 1 we present histogram estimates for ξ_N from repeated simulations with a deterministic initial condition. In more detail, 10^8 runs of each simulation were performed, each starting from the same initial condition; each run was terminated when $D(t) < 0.0001$ for the first time, and the value of $\mu_{N-1}(\mathcal{X}'_N(t))$ was output as an approximation to ξ_N (cf. Theorem 1.1). Note that, by (2.10), in the simulations one may take the new points not $U[0, 1]$ but uniform on a typically much smaller interval, which greatly increases the rate of updates to the core configuration.

In Figure 2 we present sample results obtained with an initial condition of N independent and identically distributed (i.i.d.) $U[0, 1]$ random points. Now the histograms appear much simpler, although, of course, they can be viewed as mixtures of complicated multimodal histograms

TABLE 1: Here $\kappa(\beta)$ is the Kolmogorov–Smirnov distance between a beta(β, β) distribution and the empirical distribution from the samples of size 10^8 plotted in Figure 2, minimized over β in each case.

N	β	$\kappa(\beta)$
3	1.256	0.0010
10	1.392	0.0016
50	1.509	0.0018
100	1.539	0.0019

similar to those in Figure 1. In the uniform case, it is natural to ask whether ξ_N converge in distribution to some limit distribution as $N \to \infty$.

The form of the histograms in Figure 2 might suggest a beta distribution (this is one sense in which the randomized beauty contest is 'reminiscent of a Pólya urn' [4, page 390]). An ad hoc Kolmogorov–Smirnov analysis (see Table 1) suggests that the distributions are indeed 'close' to beta distributions, but different enough for the match to be unconvincing. Simulations for large N are computationally intensive. We note that it is not unusual for beta or 'approximate beta' distributions to appear as limits of schemes that proceed via iterated procedures on intervals; see, for instance, [6] and the references therein.

In the rest of this section we study ξ_N and its distribution. Our results on the limit distribution, in particular, leave several interesting open problems, including a precise description of the phenomena displayed by the simulations reported above. In Section 3.2 we give an alternative (one might say 'phenomenological') characterization of the limit ξ_N, and contrast this with an appropriate rank-driven process in the sense of [4]. In Section 3.3 we show that the distribution of ξ_N is fully supported on (0, 1) and assigns positive probability to any proper interval, using a construction permitting transformations of configurations. Finally, Section 3.4 is devoted to the $N = 3$ case, for which some explicit computations for the distribution of ξ_N (in particular, its moments) are carried out.
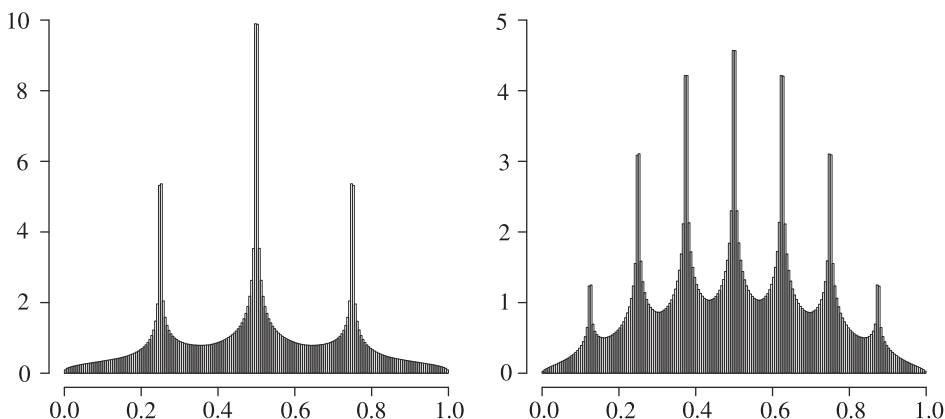

FIGURE 1: Normalized histograms each based on 10^8 simulations, with $N = 3$ and initial points $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$ (*left*), and $N = 7$ and initial points $k/8$, $k \in \{1, \ldots, 7\}$ (*right*).
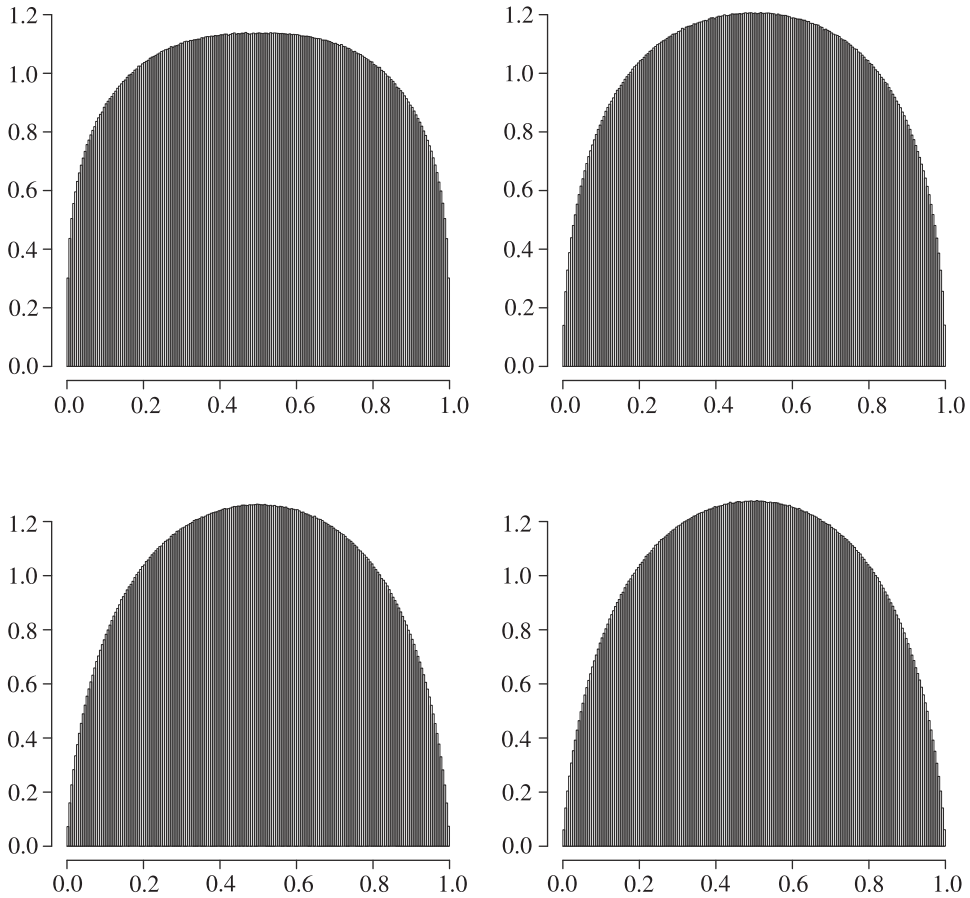
FIGURE 2: Normalized histograms for 10^8 simulations with random i.i.d. uniform initial conditions, with $N = 3, 10$ (*top*) and $N = 50, 100$ (*bottom*).

3.2. A characterization of the limit

Let

$$\pi_N(t) := \frac{1}{t}\#\{s \in \{1, 2, \ldots, t\} : \mathcal{X}_N^*(s) < \mu_N(\mathcal{X}_N(s))\},$$

the proportion of times up to time t for which the extreme point was the *leftmost* point (as opposed to the rightmost). The next result shows that $\pi_N(t)$ converges to the (random) limit ξ_N given by Theorem 1.1; we give the proof after some additional remarks.

Proposition 3.1. *Let $d = 1$ and $N \geq 3$. Then $\lim_{t \to \infty} \pi_N(t) = \xi_N$ a.s.*

It is instructive to contrast this behaviour with a suitable *rank-driven process* (cf. [4]). Namely, fix a parameter $\pi \in (0, 1)$. Take N points in $[0, 1]$, and at each step in discrete time replace either the leftmost point (with probability π) or else the rightmost point (with probability $1 - \pi$), independently at each step; inserted points are independent $U[0, 1]$ variables. For this process, the results in [4] show that the marginal distribution of a typical point converges (as $t \to \infty$ and then $N \to \infty$) to a unit point mass at π (cf. Remark 3.2 of [4]).

This leads us to one sense in which the randomized beauty contest is, to a limited extent, 'reminiscent of a Pólya urn' [4, page 390]. Recall that a Pólya urn consists of an increasing number of balls, each of which is either red or blue; at each step in discrete time, a ball is drawn uniformly at random from the urn and put back into the urn together with an extra ball of the same colour. The stochastic process of interest is the proportion of red balls, say; it converges to a random limit π', which has a beta distribution. The beauty contest can be viewed as occupying a similar relation to the rank-driven process described above as the Pólya urn process does to the simpler model in which, at each step, independently, either a red ball is added to the urn (with probability π') or else a blue ball is added (with probability $1 - \pi'$).

Proof of Proposition 3.1. Given $\tau_0, \tau_1, \tau_2, \ldots, \xi_N = \lim_{n \to \infty} \mu_{N-1}(\mathcal{X}'_N(\tau_n))$ is independent of U_t, $t \notin \{\tau_0, \tau_1, \ldots\}$, since, by (2.17), any such U_t is replaced at time $t + 1$. Let $\varepsilon > 0$. By Theorem 1.1, there exists a random $T < \infty$ a.s. for which $\max_{1 \leq i \leq N-1} |\xi_N - X_{(i)}(t)| \leq \varepsilon$ for all $t \geq T$.

Since $\mu_N(\mathcal{X}_N(t + 1)) = ((N - 1)/N)\mu_{N-1}(\mathcal{X}'_N(t)) + (1/N)U_{t+1}$, then, for $t \geq T$, using the triangle inequality, we have for any $i \in \{1, \ldots, N - 1\}$,

$$|\mu_N(\mathcal{X}_N(t + 1)) - X_{(i)}(t)| \leq \varepsilon + |\mu_N(\mathcal{X}_N(t + 1)) - \xi_N|$$
$$\leq \varepsilon + \frac{N-1}{N}|\mu_{N-1}(\mathcal{X}'_N(t)) - \xi_N| + \frac{1}{N}|U_{t+1} - \xi_N|.$$

Hence, for $t \geq T$,

$$\max_{1 \leq i \leq N-1} |X_{(i)}(t) - \mu_N(\mathcal{X}_N(t + 1))| \leq \frac{1}{N}|\xi_N - U_{t+1}| + 2\varepsilon. \tag{3.1}$$

On the other hand, for $t \geq T$, $\mu_N(\mathcal{X}_N(t + 1)) \geq ((N - 1)/N)(\xi_N - \varepsilon) + (1/N)U_{t+1}$, so that, for $i \in \{1, \ldots, N - 1\}$,

$$\mu_N(\mathcal{X}_N(t + 1)) - U_{t+1} \geq \frac{N-1}{N}(\xi_N - U_{t+1} - \varepsilon). \tag{3.2}$$

Suppose that $U_{t+1} < \xi_N - K\varepsilon$ for some $K \in (1, \infty)$. Then, from (3.1) and (3.2),

$$|\mu_N(\mathcal{X}_N(t + 1)) - U_{t+1}| - \max_{1 \leq i \leq N-1} |X_{(i)}(t) - \mu_N(\mathcal{X}_N(t + 1))|$$
$$\geq \frac{N-2}{N}(\xi_N - U_{t+1}) - \frac{3N-1}{N}\varepsilon$$
$$> \frac{\varepsilon}{N}((N - 2)K - 3N + 1).$$

This last expression is positive provided $K \geq (3N - 1)/(N - 2)$, which is the case for all $N \geq 3$ with the choice $K = 8$, say. Hence, with this choice of K, $\{U_{t+1} < \xi_N - 8\varepsilon\}$ implies that U_{t+1} is farther from $\mu_{N+1}(\mathcal{X}_N(t + 1))$ than any of the points left over from $\mathcal{X}'_N(t)$. Write $L_t := \{U_t < \xi_N - 8\varepsilon\}$. Then we have shown that, for $t \geq T$, the event L_t implies that $U_t = \mathcal{X}^*_N(t)$, and, moreover, $U_t < \mu_N(\mathcal{X}_N(t))$. Hence, for $t \geq T$,

$$\pi_N(t) \geq \frac{1}{t} \sum_{s=T}^{t} \mathbf{1}(L_s) \geq \frac{1}{t} \sum_{s=T, \, s \notin \{\tau_0, \tau_1, \ldots\}}^{t} \mathbf{1}(L_s).$$

Given $\tau_0, \tau_1, \ldots, U_s$, $s \notin \{\tau_0, \tau_1, \ldots\}$, are independent of T and ξ_N. For such an s, U_s is uniform on $I_s := [0, 1] \setminus B(\mu_{N-1}(\mathcal{X}'_N(s)); 3D(s))$, and, for $s \geq T$, $D(s) \leq 2\varepsilon$ so that $I_s \supseteq [0, \max\{\xi_N - 8\varepsilon, 0\}] \cup [\min\{\xi_N + 8\varepsilon, 1\}, 1]$. Hence, given $s \notin \{\tau_0, \tau_1, \ldots\}$ and $s \geq T$,

$$\mathbb{P}[L_s] = \mathbb{P}[U_s < \xi_N - 8\varepsilon \mid U_s \in I_s] \geq \xi_N - 8\varepsilon.$$

Hence, considering separately the cases $\xi_N > 9\varepsilon$ and $\xi_N \leq 9\varepsilon$, the strong law of large numbers implies that

$$\frac{1}{t} \sum_{s=T,\, s \notin \{\tau_0, \tau_1, \ldots\}}^{t} \mathbf{1}(L_s) \geq \xi_N - 9\varepsilon$$

for all sufficiently large t; here we have used the facts that $t - T \to \infty$ a.s. as $t \to \infty$ and $\#\{n \in \mathbb{Z}_+ : \tau_n \leq t\} = o(t)$ a.s., which follows from (2.12) and Lemma 2.4 with a standard concentration argument using, e.g. the Azuma–Hoeffding inequality. Since $\varepsilon > 0$ was arbitrary, it follows that $\liminf_{t \to \infty} t^{-1} \pi_N(t) \geq \xi_N$ a.s. The symmetrical argument considering events of the form $R_t := \{U_t > \xi_N + 8\varepsilon\}$ shows that $\liminf_{t \to \infty} (1 - t^{-1} \pi_N(t)) \geq 1 - \xi_N$ a.s., so $\limsup_{t \to \infty} t^{-1} \pi_N(t) \leq \xi_N$ a.s. Combining the two bounds gives the result.

3.3. The limit has full support

In this section we prove that ξ_N is fully supported on $(0, 1)$ in the sense that $\operatorname{ess\,inf} \xi_N = 0$, $\operatorname{ess\,sup} \xi_N = 1$, and ξ_N assigns positive probability to any non-null interval. Let

$$m_N(0) := \min\{\|X_i(0) - X_j(0)\| : i, j \in \{0, 1, \ldots, N+1\}, i \neq j\}, \tag{3.3}$$

where we use the conventions $X_0(0) := 0$ and $X_{N+1}(0) := 1$. For $\rho > 0$, let S_ρ denote the $\mathcal{F}_0$-event $S_\rho := \{m_N(0) \geq \rho\}$ that no point of $\mathcal{X}_N(0)$ is closer than distance ρ to any other point of $\mathcal{X}_N(0)$ or to either of the ends of the unit interval.

Proposition 3.2. *Let $d = 1$ and $N \geq 3$. Let $\rho \in (0, 1)$. For any non-null interval subset I of $[0, 1]$, there exists $\delta > 0$ (depending on N, I, and ρ) for which*

$$\mathbb{P}[\xi_N \in I \mid \mathcal{X}_N(0)] \geq \delta \mathbf{1}(S_\rho) \quad a.s. \tag{3.4}$$

In particular, in the case where $\mathcal{X}_N(0)$ consists of N independent $U[0, 1]$ points, $\mathbb{P}[\xi_N \in I] > 0$ for any non-null interval $I \subseteq [0, 1]$.

We suspect, but have not been able to prove, that ξ_N has a density f_N with respect to the Lebesgue measure, i.e. ξ_N is absolutely continuous in the sense that, for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\mathbb{P}[\xi_N \in A] < \varepsilon$ for every A with Lebesgue measure less than δ. Note that $\mathbb{P}[\xi_N \in A \mid \mathcal{X}_N(0)]$ may be 0 if $\mathcal{X}_N(0)$ contains nondistinct points; e.g. if $N \geq 3$ and $\mathcal{X}_N(0) = (x, x, \ldots, x, y)$, then $\mathcal{X}'_N(t) = (x, x, \ldots, x)$ for all t.

For $a \in [0, 1]$, $\varepsilon > 0$, and $t \in \mathbb{N}$, define the event

$$E_{a,\varepsilon}(t) := \bigcap_{i=1}^{N} \{|X_i(t) - a| < \varepsilon\}.$$

The main new ingredient needed to obtain Proposition 3.2 is the following result.

Lemma 3.1. *Let $N \geq 3$. For any $\rho \in (0, 1)$ and $\varepsilon > 0$, there exist $t_0 \in \mathbb{N}$ and $\delta_0 > 0$ (depending on N, ρ, and ε) for which, for all $a \in [0, 1]$,*

$$\mathbb{P}[E_{a,\varepsilon}(t_0) \mid \mathcal{X}_N(0)] > \delta_0 \mathbf{1}(S_\rho) \quad a.s.$$

Proof. Fix $a \in [0, 1]$. Let $\rho \in (0, 1)$ and $\varepsilon > 0$. It suffices to suppose that $\varepsilon \in (0, \rho)$, since $E_{a,\varepsilon}(t) \subseteq E_{a,\varepsilon'}(t)$ for $\varepsilon' \geq \varepsilon$. Suppose that S_ρ occurs, so that $m_N(0) \geq \rho$ with $m_N(0)$ defined at (3.3). For ease of notation, we list the points of $\mathcal{X}_N(0)$ in increasing order as $0 < X_1 < X_2 < \cdots < X_N < 1$. Let $M = \lfloor N/2 \rfloor$.

Let $\nu = \varepsilon/N^2$. The following argument shows how one can arrive at a configuration at a finite (deterministic) time t_0 where all of $X_1(t_0), \ldots, X_N(t_0)$ lie inside $(a - \varepsilon, a + \varepsilon)$ with a positive (though possibly very small) probability.

Let us call the points which are present at time 0 *old points*; the points which will gradually replace this set will be called *new points*. We will first describe an event by which all the old points are removed and replaced by new points arranged approximately equidistant in the interval $[X_M, X_{M+1}]$, and then we will describe an event by which such a configuration can migrate to the target interval.

Step 1. Starting from time 0, iterate the following procedure until a new point becomes an extreme point. The construction is such that at each step, the extreme point is one of the old points, either at the extreme left or right of the configuration. At each step, the extreme old point is removed and replaced by a new $U[0, 1]$ point to form the configuration at the next time unit. We describe an event of positive probability by requiring the successive new arrivals to fall in particular intervals, as follows. The first old point removed from the *right* is replaced by a new point in $(X_M + \nu, X_M + \nu + \delta)$, where $\delta \in (0, \nu)$ will be specified later. Subsequently, the ith point ($i \geq 2$) removed from the right is replaced by a new point in $(X_M + i\nu, X_M + i\nu + \delta)$. We call this subset of new points the *accumulation on the left*. On the other hand, the ith extreme point removed from the *left* ($i \in \mathbb{N}$) is replaced by a new point in $(X_{M+1} - i\nu, X_{M+1} - i\nu + \delta)$. This second subset of new points will be called the *accumulation on the right*.

During the first M steps of this procedure, the new points are necessarily internal points of the configuration and so are never removed. Therefore, there will be a time $t_1 \in [M, N]$ at which, for the first time, one of the new points becomes either the leftmost or rightmost point of $\mathcal{X}_N(t_1)$; suppose that it is the rightmost, since the argument in the other case is analogous. If at time t_1 the accumulation on the right is nonempty, we continue to perform the procedure described in step 1, but now allowing ourselves to remove new points from the accumulation on the right. So we continue putting extra points on the accumulation on the left whenever the rightmost point is removed, and similarly putting extra points to the accumulation on the right whenever the leftmost point is removed, as described for step 1. Eventually we will have either (a) a configuration where all the new points of the left or the right accumulation are completely removed, and there are still some of the old points left, or (b) a configuration where all old points are removed. In the next step we describe these two possibilities separately.

Step 2(a). Without loss of generality, suppose that the accumulation on the right is empty, so the configuration consists of k points of the left accumulation and $N - k$ old points remaining to the left of X_M (including X_M itself). Note that step 1 produces at least M new points, so $M \leq k \leq N - 1$, since by assumption we have at least one old point remaining. Let us now denote the points of the configuration $x_1 < x_2 < \cdots < x_N$ so that $x_{N-k} = X_M$, and by the construction in step 1, $x_{N-k+i} \in (X_M + i\nu, X_M + i\nu + \delta)$ for $i = 1, 2, \ldots, k$. Provided that $k \leq N - 2$, so that there are at least two old points, we will show that x_1 is necessarily the extreme point of the configuration. Indeed, writing $\mu = \mu_N(x_1, \ldots, x_N)$, using the facts that

$x_{N-k+i} \geq x_{N-k} + i\nu$ for $1 \leq i \leq k$ and $x_N \leq x_{N-k} + k\nu + \delta$, we have

$$\mu - \frac{x_1 + x_N}{2}$$
$$\geq \frac{x_1 + \cdots + x_{N-k} + kx_{N-k} + (1/2)\nu k(k+1)}{N} - \frac{x_1 + x_{N-k} + \nu k + \delta}{2}$$
$$= \frac{1}{2N}(2x_1 + \cdots + 2x_{N-k} + (2k-N)x_{N-k} - Nx_1 + \nu k(k+1-N) - \delta N).$$

The old points all have separation at least ρ, so, for $1 \leq i \leq N-k$, $x_i \geq x_1 + (i-1)\rho$, and, hence,

$$2x_1 + \cdots + 2x_{N-k} + (2k-N)x_{N-k} \geq Nx_1 + \rho(N-k-1)(N-k) + \rho(2k-N)(N-k-1).$$

It follows, after simplification, that

$$\mu - \frac{x_1 + x_N}{2} \geq \frac{1}{2N}(k(N-k-1)(\rho-\nu) - \delta N) \geq \frac{1}{2N}((N-2)(\rho-\nu) - \delta N),$$

provided $1 \leq k \leq N-2$. By the choice of ν, we have $\nu \leq \frac{1}{9}\rho$ and it follows that the last displayed expression is positive provided δ is small enough compared to ρ ($\delta < \frac{1}{4}\rho$, say). Hence, $|x_1 - \mu| > |x_N - \mu|$. Thus, next we remove x_1. We replace it similarly to the procedure in step 1, but now building up the accumulation on the *left*. We can thus iterate this step, removing old points from the left and building up the accumulation on the left, while keeping the accumulation on the right empty, until we get just one old point remaining (i.e. until $k = N-1$); this last old point will be X_M. At this stage, after a finite number of steps, we arrive at a configuration where the set of points $x_1 < x_2 < \cdots < x_N$ satisfies $x_i \in [X_M + (i-1)\nu, X_M + (i-1)\nu + \delta]$, $i = 1, 2, \ldots, N$.

Step 2(b). Suppose that the configuration is such that all old points have been removed but both left and right accumulations are nonempty. Repeating the procedure of step 1, replacing rightmost points by building the left accumulation and leftmost points by building the right accumulation, we will also, in a finite number of steps, obtain a set of points x_i such that $x_i \in [b + (i-1)\nu, b + (i-1)\nu + \delta]$, $i = 1, 2, \ldots, N$, for some $b \in [0, 1]$.

Step 3. Now we will show how we can get to the situation where all points lie inside the interval $(a - \varepsilon, a + \varepsilon)$ starting from any configuration in which

$$x_i \in [b + i\nu, b + i\nu + \delta], \qquad i = 1, 2, \ldots, N-1, \tag{3.5}$$

where $b \in [0, 1]$ and $x_1 < \cdots < x_{N-1}$ are the core points of the configuration (i.e. with the extreme point removed). We have shown in step 1 and step 2 how we can achieve such a configuration in a finite time with a positive probability. Suppose that $a > b$; the argument for the other case is entirely analogous. We describe an event of positive probability by which the entire configuration can be moved to the right. See Figure 3.

Having just removed the extreme point, we stipulate that the new point y_1 belongs to $(b + N\nu - 6\delta, b + N\nu - 5\delta)$, so $y_1 > x_{N-1}$ is the new rightmost point provided $\delta < \nu/7$. Then to ensure that x_1, and not y_1, is the most extreme point, we need

$$\frac{x_1 + y_1}{2} - \left[b + \nu\frac{N+1}{2}\right] < \frac{x_1 + \cdots + x_{N-1} + y_1}{N} - \left[b + \nu\frac{N+1}{2}\right].$$

The left-hand side of the last inequality is less than -2δ while the right-hand side is more than $-6\delta/N$, so the inequality is indeed satisfied provided $N \geq 3$.
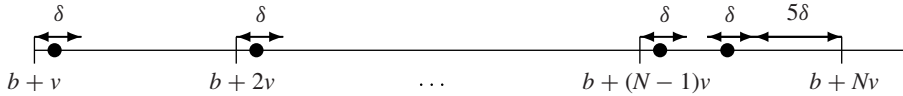
FIGURE 3: Schematic of a configuration at the start of step 3. The black circles represent the points $x_1, x_2, \ldots, x_{N-1}$ and, on the extreme right, the new point y_1.

Hence, at the next step x_1 is removed. Our new collection of core points is $x_2 < \cdots < x_{N-1} < y_1$. We stipulate that the next new point y_2 arrives in $(b + (N + 1)v - 18\delta, b + (N + 1)v - 17\delta)$. So, again, for small enough δ ($\delta < v/13$ suffices), $y_2 > y_1$ and the newly added point (y_1) becomes the rightmost point in the configuration. Again, to ensure that the leftmost point (x_2) is now the extreme point, we require that

$$\frac{x_2 + y_2}{2} - \left[b + v\frac{N+3}{2}\right] < \frac{x_2 + \cdots + x_{N-1} + y_1 + y_2}{N} - \left[b + v\frac{N+3}{2}\right].$$

The left-hand side of the last inequality is less than -8δ, while the right-hand side is more than $-24\delta/N$, and so the displayed inequality is true provided $N \geq 3$.

We repeat this process until we remove the rightmost core point present at the start of step 3, namely, x_{N-1}, located in $[b + (N - 1)v, b + (N - 1)v + \delta]$. We will demonstrate how we can do this, in succession removing points from the left of the configuration and at each step replacing them by points on the right with a careful choice of locations for the new points. We consecutively put new points y_k at locations in intervals

$$\Delta_k := (b + (N - 1 + k)v - 2 \cdot 3^k\delta, b + (N - 1 + k)v - 2 \cdot 3^k\delta + \delta)$$

for $k = 1, 2, \ldots, N - 1$. We have just shown that, for $k = 1, 2$, this procedure will maintain the leftmost point (x_k) as the extreme point. Let us show that this is true for all $1 \leq k \leq N - 1$ by an inductive argument. Indeed, suppose that the original points $x_1, x_2, \ldots, x_{k-1}$ have been removed, the successive new points y_j are located in Δ_j, $j = 1, 2, \ldots, k - 1$, and that the replacement for the most recently removed point x_{k-1} is the new point y_k. Place the new point y_k in Δ_k. Provided $\delta < v/(4 \cdot 3^{k-1} + 1)$, $y_k > y_{k-1}$ and y_k is the rightmost point of the new configuration, while the leftmost point is $x_k \in [b + vk, b + vk + \delta]$. Since $N \geq 3$, we have

$$\frac{x_k + y_k}{2} \leq b + \left[\frac{N-1}{2} + k\right]v - [3^k - 1]\delta$$

$$\leq b + \left[\frac{N-1}{2} + k\right]v - \frac{2(3 + 3^2 + \cdots + 3^k)\delta}{N}$$

$$\leq \frac{x_k + \cdots + x_{N-1} + y_1 + \cdots + y_k}{N},$$

thus ensuring that the leftmost point x_k, and not y_k, is the farthest from the centre of mass. Thus, provided $\delta < 3^{-N}v$, say, we proceed to remove all the points x_k and arrive at a new collection of points $x'_1, \ldots, x'_{N-1}$ satisfying the property

$$x'_i \in [b' + iv, b' + iv + \delta'], \qquad i = 1, 2, \ldots, N - 1,$$

where $b' = b + (N - 1)v - \delta'$ and $\delta' := 3^N\delta$ ($> 2 \cdot 3^{N-1}\delta$). Thus, the situation is similar to that given in (3.5) but with b replaced by $b' > b + v$, so the whole 'grid' is shifted to the right.

Hence, provided δ is small enough, and δ' and its subsequent analogues remain such that $\delta' < 3^{-N}v$, we can repeat the above procedure and move points to the right again, etc., a finite number of times (depending on $|b - a|/v$) until the moment when all the new points are indeed in $(a - \varepsilon, a + \varepsilon)$, and the probability of making all those steps is strictly positive. In particular, we can check that taking $\delta < 3^{-2N/v}v$ will suffice.

All in all, we have performed a finite number of steps, which can be bounded above in terms of N, ρ, and ε, but independently of a, and each of which required a $U[0, 1]$ variable to be placed in a small interval (of width less than $3^{-N}v$) and so has positive probability, which can be bounded below in terms of N, ρ, and ε. So, overall, the desired transformation of the configuration has positive probability depending on N, ρ, and ε, but not on a. This completes the proof of the lemma.

Proof of Proposition 3.2. Write $\mu'(t) := \mu_{N-1}(\mathcal{X}'_N(t))$. Let $I \subseteq [0, 1]$ be a non-null interval. We can (and do) choose $a \in (0, 1)$ and $\varepsilon' > 0$ such that $I' := [a - \varepsilon', a + \varepsilon'] \subseteq I$. Also, take $I'' := [a - \varepsilon, a + \varepsilon] \subset I'$ for $\varepsilon = (4BC)^{-1}N^{-1/2}\varepsilon'$, where C is the constant in Lemma 2.6 and $B \geq 1$ is an absolute constant chosen so that $\varepsilon < \varepsilon'/4$ for all $N \geq 3$. Fix $\rho \in (0, 1)$. It follows from Lemma 3.1 that, for some $\delta_0 > 0$ and $t_0 \in \mathbb{N}$, depending on ε,

$$\mathbb{P}[\{\mu'(t_0) \in I''\} \cap \{D(t_0) \leq 2\varepsilon\} \mid \mathcal{X}_N(0)] \geq \delta_0 \mathbf{1}(S_\rho).$$

By Lemma 2.2, $D(t_0) \leq 2\varepsilon$ implies that $F(t_0) \leq 2N\varepsilon^2 < (\varepsilon'/(2BC))^2$, so that $t_0 \geq v_{\varepsilon'/(2BC)}$, where $v.$ is as defined just before Lemma 2.6. Applying Lemma 2.6 with this choice of t_0 and with the ε there equal to $\varepsilon'/(2BC)$, we obtain, by Markov's inequality,

$$\mathbb{P}\left[\max_{t \geq t_0} |\mu'(t) - \mu'(t_0)| \leq \tfrac{3}{4}\varepsilon' \,\middle|\, \mathcal{F}_{t_0}\right] \geq \tfrac{1}{3} \quad \text{a.s.}$$

It follows that, given $\mathcal{X}_N(0)$, the event

$$\{\mu'(t_0) \in I''\} \cap \{D(t_0) \leq 2\varepsilon\} \cap \{|\xi_N - \mu'(t_0)| \leq \tfrac{3}{4}\varepsilon'\}$$

has probability at least $\frac{1}{3}\delta_0 \mathbf{1}(S_\rho)$, and on this event we have $|\xi_N - a| \leq \varepsilon + (3\varepsilon'/4) < \varepsilon'$, so $\xi_N \in I$. Hence, (3.4) follows.

For the final statement in the proposition, suppose that $\mathcal{X}_N(0)$ consists of independent $U[0, 1]$ points. In this case $m_N(0)$ defined at (3.3) is the minimal spacing in the induced partition of $[0, 1]$ into $N + 1$ segments, which has the same distribution as $1/(N + 1)$ times a single spacing, and in particular has density $f(x) = N(N + 1)(1 - (N + 1)x)^{N-1}$ for $x \in [0, 1/(N + 1)]$ (cf. Appendix A). Hence, for any $\rho \in [0, 1/(N + 1)]$, we have $\mathbb{P}[S_\rho] = (1 - (N + 1)\rho)^N$, which is positive for $\rho = 1/2N$, say. Thus, taking expectations in (3.4) yields the final statement in the proposition.

3.4. Explicit calculations for $N = 3$

In this section we take $N = 3$, the smallest nontrivial example. In this case we can perform some explicit calculations to obtain information about the distribution of ξ_3. In fact, we work with a slightly modified version of the model, avoiding certain 'boundary effects', to ease computation. Specifically, we do not use $U[0, 1]$ replacements but, given $\mathcal{X}_3(t)$, we take U_{t+1} to be uniform on the interval $[\min \mathcal{X}'_3(t) - D(t), \max \mathcal{X}'_3(t) + D(t)]$. If this interval is contained in $[0, 1]$ for all t, this modification would have no effect on the value of ξ_3 realized (only speeding up the convergence), but the fact that now U_{t+1} might be outside $[0, 1]$ *does* change the model.

For this modified model, the argument for Theorem 1.1 follows through with minor changes, although we essentially reprove the conclusion of Theorem 1.1 in this case when we prove the following result, which gives an explicit description of the limit distribution. Here and subsequently, '$\overset{D}{=}$' denotes equality in distribution.

Proposition 3.3. *Let $d = 1$ and $N = 3$, and work with the modified version of the process just described. Let $\mathcal{X}_3(0)$ consist of three distinct points in $[0, 1]$. Write $\mu := \mu_2(\mathcal{X}_3'(0))$ and $D := D_2(\mathcal{X}_3'(0))$. There exists a random $\xi_3 := \xi_3(\mathcal{X}_3(0)) \in \mathbb{R}$ such that $\mathcal{X}_3'(t) \to (\xi_3, \xi_3)$ a.s. as $t \to \infty$. The distribution of ξ_3 can be characterized via $\xi_3 \overset{D}{=} \mu + DL$, where L is independent of (μ, D), $L \overset{D}{=} -L$, and the distribution of L is determined by the distributional solution to the fixed-point equation*

$$
L \overset{D}{=}
\begin{cases}
-\dfrac{1 + U}{2} + UL & \text{with probability } \tfrac{1}{3}, \\[2mm]
-\dfrac{2 - U}{4} + \dfrac{U}{2}L & \text{with probability } \tfrac{1}{6}, \\[2mm]
\dfrac{2 - U}{4} + \dfrac{U}{2}L & \text{with probability } \tfrac{1}{6}, \\[2mm]
\dfrac{1 + U}{2} + UL & \text{with probability } \tfrac{1}{3},
\end{cases}
\tag{3.6}
$$

where $\mathbb{E}[|L|^k] < \infty$ for all k and $U \sim U[0, 1]$. Writing $\theta_k := \mathbb{E}[L^k]$, we have $\theta_k = 0$ for odd k, and $\theta_2 = \frac{7}{12}$, $\theta_4 = \frac{375}{368}$, and $\theta_6 = \frac{76\,693}{22\,080}$. In particular,

$$
\mathbb{E}[\xi_3] = \mathbb{E}[\mu], \qquad \mathbb{E}[\xi_3^2] = \mathbb{E}[\mu^2] + \tfrac{7}{12}\mathbb{E}[D^2], \quad and \quad \mathbb{E}[\xi_3^3] = \mathbb{E}[\mu^3] + \tfrac{7}{4}\mathbb{E}[\mu D^2]. \tag{3.7}
$$

In the case where $\mathcal{X}_3(0)$ contains three independent $U[0, 1]$ points, $\mathbb{E}[\xi_3^k] = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ for $k = 1, 2, 3$, respectively. If $\mathcal{X}_3(0) = (\frac{1}{4}, \frac{1}{2}, \frac{3}{4})$ then $\mathbb{E}[\xi_3^k] = \frac{1}{2}, \frac{29}{96}, \frac{13}{64}, \frac{873}{5888}$ for $k = 1, 2, 3, 4$.

We give the proof of Proposition 3.3 at the end of this section. First we state one consequence of the fixed-point representation (3.6).

Proposition 3.4. *L given by (3.6) has an absolutely continuous distribution.*

Proof. It follows from (3.6) that

$$
\mathbb{P}[L = \tfrac{1}{2}] = \tfrac{1}{3}\mathbb{P}[U(L - \tfrac{1}{2}) = 1] + \tfrac{1}{6}\mathbb{P}[U(L + \tfrac{1}{2}) = 2]
$$
$$
+ \tfrac{1}{6}\mathbb{P}[U(L - \tfrac{1}{2}) = 0] + \tfrac{1}{3}\mathbb{P}[U(L + \tfrac{1}{2}) = 0].
$$

The first two terms on the right-hand side of the above display are 0, by an application of the first part of Lemma B.1 with $X = U$, $Y = L \pm \frac{1}{2}$, and $a = 1, 2$. Also, since $U > 0$ a.s., $\mathbb{P}[U(L \mp \frac{1}{2}) = 0] = \mathbb{P}[L = \pm\frac{1}{2}]$, and, by symmetry, $\mathbb{P}[L = \frac{1}{2}] = \mathbb{P}[L = -\frac{1}{2}]$. Thus, we obtain

$$
\mathbb{P}[L = \tfrac{1}{2}] = \tfrac{1}{6}\mathbb{P}[L = \tfrac{1}{2}] + \tfrac{1}{3}\mathbb{P}[L = -\tfrac{1}{2}] = \tfrac{1}{2}\mathbb{P}[L = \tfrac{1}{2}].
$$

Hence, $\mathbb{P}[L = \frac{1}{2}] = \mathbb{P}[L = -\frac{1}{2}] = 0$.

Each term on the right-hand side of (3.6) is of the form $\pm\frac{1}{2} + V(L \pm \frac{1}{2})$, where V is an absolutely continuous random variable, independent of L (namely, U or $\frac{1}{2}U$). The final statement in Lemma B.1 with the fact that $\mathbb{P}[L = \pm\frac{1}{2}] = 0$ shows that each such term is absolutely continuous. Finally, Lemma B.2 completes the proof.

TABLE 2: Empirical kth moment values (to four decimal places) computed from the simulations in Figure 4.

k	1	2	3	4	5	6
$\pm\frac{1}{2}$ core	0.0001	0.5833	0.0000	1.0192	−0.0005	3.4765
U[0,1]	0.5000	0.3333	0.2500	0.2029	0.1739	0.1561

In principle, characterization (3.6) can be used to recursively determine all the moments $\mathbb{E}[L^k] = \theta_k$, and the moments of ξ_3 may then be obtained by expanding $\mathbb{E}[\xi_3^k] = \mathbb{E}[(\mu + DL)^k]$. However, the calculations soon become cumbersome, particularly as μ and D are, typically, not independent: we give some distributional properties of (μ, D) in the case of a uniform random initial condition in Appendix A.

Before giving the proof of Proposition 3.3, we comment on some simulations. In Figure 4 we present histogram estimates for the distribution of ξ_3 for two initial distributions (one deterministic and the other uniform random), and in Table 2 we present corresponding moment estimates, which may be compared to the theoretical values given in Proposition 3.3. In the uniform case, we only computed the first three moments analytically, namely, $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ as quoted in Proposition 3.3; it is a curiosity that these coincide with the first three moments of the $U[0, 1]$ distribution.

Proof of Proposition 3.3. Let $\mu'(t) := \frac{1}{2}(X_{(1)}(t) + X_{(2)}(t))$ and $D(t) := |X_{(1)}(t) - X_{(2)}(t)|$ denote the mean and diameter of the core configuration, repeating our notation from above.

Consider separately the events that U_{t+1} falls into one of the four intervals

$$[\min \mathcal{X}'_3(t) - D(t), \min \mathcal{X}'_3(t)), \qquad [\min \mathcal{X}'_3(t), \min \mathcal{X}'_3(t) + \tfrac{1}{2}D(t)),$$
$$[\min \mathcal{X}'_3(t) + \tfrac{1}{2}D(t), \max \mathcal{X}'_3(t)), \qquad [\max \mathcal{X}'_3(t), \max \mathcal{X}'_3(t) + D(t)],$$

which have probabilities $\frac{1}{3}, \frac{1}{6}, \frac{1}{6}$, and $\frac{1}{3}$, respectively. Given $(\mu'(t), D(t))$, we see that, for
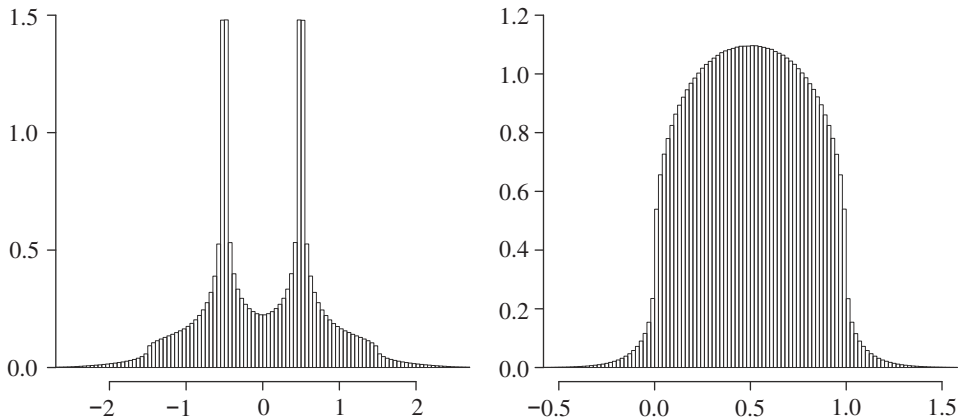

FIGURE 4: Normalized histograms based on 10^8 simulations of the modified $N = 3$ model with a fixed $\{-\frac{1}{2}, \frac{1}{2}, 100\}$ initial condition (*left*) and an i.i.d. $U[0, 1]$ initial condition (*right*).

V_{t+1} a $U[0, 1]$ variable, independent of $(\mu'(t), D(t))$,

$$(\mu'(t+1), D(t+1)) = \begin{cases} \left(\mu'(t) - \frac{1}{2}(1 + V_{t+1})D(t), V_{t+1}D(t)\right) & \text{with probability } \frac{1}{3}, \\ \left(\mu'(t) - \frac{1}{4}(2 - V_{t+1})D(t), \frac{1}{2}V_{t+1}D(t)\right) & \text{with probability } \frac{1}{6}, \\ \left(\mu'(t) + \frac{1}{4}(2 - V_{t+1})D(t), \frac{1}{2}V_{t+1}D(t)\right) & \text{with probability } \frac{1}{6}, \\ \left(\mu'(t) + \frac{1}{2}(1 + V_{t+1})D(t), V_{t+1}D(t)\right) & \text{with probability } \frac{1}{3}. \end{cases}$$
(3.8)

Writing $m_k(t) = \mathbb{E}[D(t)^k \mid \mathcal{X}_3(0)]$ we obtain from the second coordinates in (3.8):

$$m_k(t+1) = \frac{2}{3}\mathbb{E}[V_{t+1}^k]m_k(t) + \frac{1}{3}2^{-k}\mathbb{E}[V_{t+1}^k]m_k(t).$$

This implies that

$$m_k(t) = \left(\frac{1}{3(k+1)}(2 + 2^{-k})\right)^t D(0)^k.$$

For example, $m_1(t) = (\frac{5}{12})^t D(0)$ and $m_2(t) = (\frac{1}{4})^t D(0)^2$.

Next we show that $\mu'(t)$ converges. From (3.8), we have $|\mu'(t+1) - \mu'(t)| \le D(t)$ a.s., so to show that $\mu'(t)$ converges, it suffices to show that $\sum_{t=0}^{\infty} D(t) < \infty$ a.s. But this can be seen from essentially the same argument as Lemma 2.5, or directly from the fact that the sum has nonnegative terms and $\mathbb{E}\sum_{t=0}^{\infty} D(t) = \sum_{t=0}^{\infty} \mathbb{E}[m_1(t)]$, which is finite. Hence, $\mu'(t)$ converges a.s. to some limit, ξ_3, say. Extending this argument a little, it follows from (3.8) that $|\mu'(t+1)| \le |\mu'(t)| + D(t)$ a.s. and $D(t+1) \le V_{t+1}D(t)$ a.s. Hence, for $V_1, V_2, \ldots$ i.i.d. $U[0, 1]$ random variables, we have $D(t) \le V_1 \ldots V_t D(0)$ and

$$|\mu'(t)| - |\mu'(0)| = \sum_{s=0}^{t-1}(|\mu'(s+1)| - |\mu'(s)|) \le \left(1 + \sum_{s=1}^{\infty}\prod_{r=1}^{s} V_r\right)D(0) =: (1 + Z)D(0).$$

Here Z has the so-called *Dickman distribution* (see, e.g. [13, Section 3]), which has finite moments of all orders. Hence, $\mathbb{E}[|\mu'(t)|^p \mid \mathcal{X}_3(0)]$ is bounded independently of t, so, for any $p \ge 1$, $(\mu'(t))^p$ is uniformly integrable, and, hence, $\lim_{t\to\infty}\mathbb{E}[(\mu'(t))^k \mid \mathcal{X}_3(0)] = \mathbb{E}[\xi_3^k \mid \mathcal{X}_3(0)]$ for any $k \in \mathbb{N}$.

We now want to compute the moments of ξ_3; by the previous argument, we can first work with the moments of $\mu'(t)$. Note that, from (3.8),

$$\mathbb{E}[(\mu'(t+1) - \mu'(t))^k \mid \mathcal{X}_3(t)]$$
$$= \frac{1 + (-1)^k}{3}D(t)^k\mathbb{E}\left[\left(\frac{1 + V_{t+1}}{2}\right)^k\right] + \frac{1 + (-1)^k}{6}D(t)^k\mathbb{E}\left[\left(\frac{1 + V_{t+1}}{4}\right)^k\right]$$
$$= \frac{1 + (-1)^k}{6}(2^{1-k} + 2^{-2k})\frac{2^{k+1} - 1}{k+1}D(t)^k,$$

using the fact that $\mathbb{E}[(1 + V_{t+1})^k] = (2^{k+1} - 1)/(k+1)$. In particular, $\mathbb{E}[(\mu'(t+1) - \mu'(t))^k \mid \mathcal{X}_3(t)] = 0$ for odd k, so $\mathbb{E}[\mu'(t) \mid \mathcal{X}_3(0)] = \mu'(0)$, and, hence, $\mathbb{E}[\xi_3] = \lim_{t\to\infty}\mathbb{E}[\mu'(t)] = \mathbb{E}[\mu'(0)]$, giving the first statement in (3.7). In addition,

$$\mathbb{E}[(\mu'(t+1))^2 - (\mu'(t))^2 \mid \mathcal{X}_3(t)]$$
$$= 2\mu'(t)\mathbb{E}[\mu'(t+1) - \mu'(t) \mid \mathcal{X}_3(t)] + \mathbb{E}[(\mu'(t+1) - \mu'(t))^2 \mid \mathcal{X}_3(t)]$$
$$= \frac{7}{16}D(t)^2.$$

Hence,

$$\mathbb{E}[(\mu'(t))^2 - (\mu'(0))^2 \mid \mathcal{X}_3(0)] = \sum_{s=0}^{t-1} \mathbb{E}[(\mu'(s+1))^2 - (\mu'(s))^2 \mid \mathcal{X}_3(0)]$$

$$= \frac{7}{16} \sum_{s=0}^{t-1} m_2(s)$$

$$\rightarrow \frac{7}{16} \sum_{s=0}^{\infty} 4^{-s} D(0)^2$$

as $t \rightarrow \infty$, and the limit evaluates to $\frac{7}{12} D(0)^2$, so that $\mathbb{E}[\xi_3^2] = \lim_{t\to\infty} \mathbb{E}[(\mu'(t))^2] = \mathbb{E}[(\mu'(0))^2] + \frac{7}{12}\mathbb{E}[D(0)^2]$, giving the second statement in (3.7).

Write $L(\mu'(0), D(0)) = \xi_3(\mathcal{X}_3(0))$, emphasizing the dependence on the initial configuration through $\mu'(0)$ and $D(0)$. Then, by translation and scaling properties,

$$L(\mu'(0), D(0)) \stackrel{\mathrm{D}}{=} \mu'(0) + D(0)L(0, 1). \tag{3.9}$$

So we work with $L := L(0, 1)$ (which has the initial core points at $\pm\frac{1}{2}$).

We will derive a fixed-point equation for L. The argument is closely related to that for (3.8). Conditioning on the first replacement and using the transformation relation (3.9), we obtain (3.6). From (3.6) we see that $|L|$ is stochastically dominated by $1 + U|L|$; iterating this, similarly to the argument involving the Dickman distribution above, we find that $|L|$ is stochastically dominated by $1 + Z$, where Z has the Dickman distribution, which is determined by its moments. Hence, (3.6) determines a unique distribution for L with $\mathbb{E}[|L|^k] < \infty$ for all k.

Writing (3.6) in functional form, $L \stackrel{\mathrm{D}}{=} \Psi(L)$, we see that, by symmetry of the form of Ψ, $\Psi(L) \stackrel{\mathrm{D}}{=} -\Psi(-L)$ also. Hence, $-L \stackrel{\mathrm{D}}{=} -\Psi(L) \stackrel{\mathrm{D}}{=} \Psi(-L)$, so $-L$ satisfies the same distributional fixed-point equation as does L. Hence, $L \stackrel{\mathrm{D}}{=} -L$.

Writing $\theta_k := \mathbb{E}[L^k]$, which we know is finite, we obtain

$$\theta_k = \frac{1}{3} \sum_{j=0}^{k} \binom{k}{j}(1 + (-1))^j \theta_{k-j} \mathbb{E}\left[\left(\frac{1+U}{2}\right)^j U^{k-j}\right]$$

$$+ \frac{1}{6} \sum_{j=0}^{k} \binom{k}{j}(1 + (-1))^j \theta_{k-j} \mathbb{E}\left[\left(\frac{2-U}{4}\right)^j \left(\frac{U}{2}\right)^{k-j}\right].$$

Here,

$$\mathbb{E}\left[\left(\frac{1+U}{2}\right)^j U^{k-j}\right] = 2^{-j} \sum_{\ell=0}^{j} \binom{j}{\ell} \frac{1}{k-\ell+1} =: a(k, j),$$

$$\mathbb{E}\left[\left(\frac{2-U}{4}\right)^j \left(\frac{U}{2}\right)^{k-j}\right] = 2^{-k} \sum_{\ell=0}^{j} \binom{j}{\ell} \frac{(-1/2)^{j-\ell}}{k-\ell+1} =: b(k, j).$$

So we obtain

$$\theta_k = \frac{1}{3} \sum_{j \text{ even}, j \le k} \binom{k}{j} \theta_{k-j}(2a(k, j) + b(k, j)). \tag{3.10}$$

In particular, as can be seen either directly by symmetry or by an inductive argument using (3.10), $\theta_k = 0$ for odd k. For even k, we can use (3.10) recursively to find θ_k, obtaining, for example, the values quoted in the proposition.

Note that, by (3.9), $\mathbb{E}[\xi_3^3] = \mathbb{E}[(\mu'(0) + LD(0))^3]$, which, on expansion, gives the final statement in (3.7). The first three moments in the case of the uniform initial condition follow from (3.7) and Lemma A.1. For the initial condition with points $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$, we have $D(0) = \frac{1}{4}$ and $\mu'(0) = \chi \frac{3}{8} + (1 - \chi)\frac{5}{8} = \frac{5}{8} - \chi/4$, where χ is the tiebreaker random variable taking values 0 or 1, each with probability $\frac{1}{2}$. It follows that $\mathbb{E}[\mu'(0)^k] = \frac{1}{2}8^{-k}(3^k + 5^k)$. Then, using (3.9),

$$\mathbb{E}[\xi_3^k] = \mathbb{E}\left[\left(\mu'(0) + \frac{L}{4}\right)^k\right] = \frac{1}{2}8^{-k}\sum_{j=0}^{k}\binom{k}{j}2^j\theta_j(3^{k-j} + 5^{k-j}).$$

We can now compute the four moments given in the proposition.

Appendix A. Uniform spacings

In this appendix we collect some results about uniform spacings which allow us to obtain distributional results about our uniform initial configurations. The basic results that we build on here can be found in Section 4.2 of [14]; see the references therein for a fuller treatment of the theory of spacings.

Let $U_1, U_2, \ldots, U_n$ be independent $U[0, 1]$ points. Denote the corresponding increasing order statistics $U_{[1]} \leq \cdots \leq U_{[n]}$, and define the induced *spacings* by $S_{n,i} := U_{[i]} - U_{[i-1]}$, $i = 1, \ldots, n + 1$, with the conventions $U_{[0]} := 0$ and $U_{[n+1]} := 1$. We collect some basic facts about the $S_{n,i}$. The spacings are exchangeable, and any n-vector, such as $(S_{n,1}, \ldots, S_{n,n})$, has the uniform density on the simplex $\Delta_n := \{(x_1, \ldots, x_n) \in [0, 1]^n : \sum_{i=1}^{n} x_i \leq 1\}$.

We need some joint properties of up to three spacings. Any three spacings have density $f(x_1, x_2, x_3) = n(n - 1)(n - 2)(1 - x_1 - x_2 - x_3)^{n-3}$ on Δ_3. We note that

$$\min\{S_{n,1}, S_{n,2}\} \overset{\text{D}}{=} \tfrac{1}{2}S_{n,1}, \qquad n \geq 1, \tag{A.1}$$

$$(S_{n,1}, \min\{S_{n,2}, S_{n,3}\}) \overset{\text{D}}{=} (S_{n,1}, \tfrac{1}{2}S_{n,2}), \qquad n \geq 2; \tag{A.2}$$

see, for example, Lemma 4.1 of [14]. Finally, for any $n \geq 1$, and $\alpha \geq 0$ and $\beta \geq 0$,

$$\mathbb{E}[S_{n,1}^\alpha S_{n,2}^\beta] = \frac{\Gamma(n + 1)\Gamma(\alpha + 1)\Gamma(\beta + 1)}{\Gamma(n + 1 + \alpha + \beta)}. \tag{A.3}$$

In particular, $\mathbb{E}[S_{n,1}^k] = n!\,k!/(n + k)!$ for $k \in \mathbb{N}$.

Our main application in the present paper of the results on spacings collected above is to obtain the following result, which we use in Section 3.4.

Lemma A.1. *Let $d = 1$ and $N = 3$. Suppose that $\mathcal{X}_3(0)$ consists of three independent $U[0, 1]$ points. Then*

$$(\mu_2(\mathcal{X}_3'(0)), D_2(\mathcal{X}_3'(0))) \overset{\text{D}}{=} \left((S_1 + \tfrac{1}{4}S_2)\zeta + (1 - S_1 - \tfrac{1}{4}S_2)(1 - \zeta), \tfrac{1}{2}S_2\right),$$

where ζ is a Bernoulli random variable with $\mathbb{P}[\zeta = 0] = \mathbb{P}[\zeta = 1] = \frac{1}{2}$. For $k \in \mathbb{Z}_+$,

$$\mathbb{E}[(D_2(\mathcal{X}_3'(0)))^k] = 2^{-k}\frac{6}{(k + 1)(k + 2)(k + 3)}, \tag{A.4}$$

$$\mathbb{E}[(\mu_2(\mathcal{X}_3'(0)))^k] = \frac{4(3k - 5 + (3^{k+3} - 1)4^{-(k+1)})}{(k + 1)(k + 2)(k + 3)}. \tag{A.5}$$

So, for example, the first three moments of $D_2(\mathcal{X}_3'(0))$ are $\frac{1}{8}$, $\frac{1}{40}$, and $\frac{1}{160}$, while the first three moments of $\mu_2(\mathcal{X}_3'(0))$ are $\frac{1}{2}$, $\frac{51}{160}$, and $\frac{73}{320}$. Finally, $\mathbb{E}[\mu_2(\mathcal{X}_3'(0))(D_2(\mathcal{X}_3'(0)))^2] = \frac{1}{80}$.

Proof. The three points of $\mathcal{X}_3(0)$ induce a partition of the interval $[0, 1]$ into uniform spacings S_1, S_2, S_3, S_4, enumerated left to right (for this proof, we suppress the first index in the notation above). For ease of notation, write $D := D_2(\mathcal{X}_3'(0))$ and $\mu := \mu_2(\mathcal{X}_3'(0))$ for the duration of this proof. Then $D = \min\{S_2, S_3\} \overset{D}{=} \frac{1}{2}S_1$ by (A.1). Moreover, $\min\{S_2, S_3\}$ is equally likely to be either S_2 or S_3. In the former case, $\mu = S_1 + \frac{1}{2}\min\{S_2, S_3\}$, while in the latter case $\mu = 1 - S_4 - \frac{1}{2}\min\{S_2, S_3\}$. Using (A.2), we obtain the following characterization of the joint distribution of μ and D:

$$(\mu, D) \overset{D}{=} \begin{cases} \left(S_1 + \frac{1}{4}S_2, \frac{1}{2}S_2\right) & \text{with probability } \frac{1}{2}, \\ \left(1 - S_1 - \frac{1}{4}S_2, \frac{1}{2}S_2\right) & \text{with probability } \frac{1}{2}. \end{cases} \tag{A.6}$$

Hence, $\mathbb{E}[D^k] = 2^{-k}\mathbb{E}[S_1^k]$, which gives (A.4) by the $n = 3$, $\alpha = k$, $\beta = 0$ case of (A.3).

For the moments of μ, it follows from (A.6) that μ has the distribution $W := S_1 + \frac{1}{4}S_2$ with probability $1/2$ or $1 - W$ with probability $\frac{1}{2}$. So we have

$$\mathbb{E}[\mu^k] = \frac{1}{2}\mathbb{E}[W^k] + \frac{1}{2}\mathbb{E}[(1 - W)^k] = \frac{1}{2}w_k + \frac{1}{2}\sum_{j=0}^{k}\binom{k}{j}(-1)^j w_j,$$

where $w_k := \mathbb{E}[W^k]$. Since $w_k = \mathbb{E}[(S_1 + \frac{1}{4}S_2)^k]$, we compute

$$w_k = \sum_{j=0}^{k}\binom{k}{j}4^{-j}\mathbb{E}[S_1^{k-j}S_2^j] = 6\frac{k!}{(k+3)!}\sum_{j=0}^{k}4^{-j},$$

by the $n = 3$, $\alpha = k - j$, $\beta = j$ case of (A.3). Thus, we obtain

$$w_k = \frac{8(1 - 4^{-(k+1)})}{(k+1)(k+2)(k+3)}.$$

It follows that

$$\mathbb{E}[\mu^k] = \frac{1}{2}w_k + 4\sum_{j=0}^{k}\frac{k!}{(j+3)!\,(k-j)!}(-1)^j - \sum_{j=0}^{k}\frac{k!}{(j+3)!(k-j)!}\left(-\frac{1}{4}\right)^j.$$

We deduce (A.5), after simplification, from the claim that, for any $z \in \mathbb{R}$,

$$S(z) := \sum_{j=0}^{k}\frac{k!}{(j+3)!\,(k-j)!}(-z)^j$$

$$= \frac{k!}{2z^3(k+3)!}[z^2(k+2)(k+3) + 2 - 2z(k+3) - 2(1-z)^{k+3}]. \tag{A.7}$$

Thus, it remains to verify (A.7). To this end, note that

$$
S(z) = \frac{k!}{(k+3)!} \sum_{j=0}^{k} \binom{k+3}{j+3}(-z)^j
$$

$$
= \frac{k!}{(k+3)!} \left[-z^{-3} \sum_{j=0}^{k+3} \binom{k+3}{j}(-z)^j + z^{-1}\binom{k+3}{2} - z^{-2}\binom{k+3}{1} + z^{-3}\binom{k+3}{0} \right]
$$

$$
= \frac{k!}{z^3(k+3)!} \left[-(1-z)^{k+3} + \frac{1}{2}z^2(k+2)(k+3) - z(k+3) + 1 \right],
$$

which gives claim (A.7).

For the final statement in the lemma, from (A.6) we have

$$
\mathbb{E}[\mu D^2] = \tfrac{1}{2}\mathbb{E}\big[\big(S_1 + \tfrac{1}{4}S_2\big)\big(\tfrac{1}{4}S_2^2\big)\big] + \tfrac{1}{2}\mathbb{E}\big[\big(1 - S_1 - \tfrac{1}{4}S_2\big)\big(\tfrac{1}{4}S_2^2\big)\big] = \tfrac{1}{8}\mathbb{E}[S_2^2] = \tfrac{1}{80}
$$

by (A.3). This completes the proof of the lemma.

We can also obtain explicit expressions for the densities of D and μ. Since $D \stackrel{\mathrm{D}}{=} \tfrac{1}{2}S_1$, the density of D is $f_D(r) = 3(1 - 2r)^2$ for $r \in [0, \tfrac{1}{2}]$. In addition, μ has density f_μ given by

$$
f_\mu(r) = \begin{cases} 4r[3(1-r) - 4r] & \text{if } r \in [0, \tfrac{1}{4}], \\ 2 - 4r(1-r) & \text{if } r \in [\tfrac{1}{4}, \tfrac{3}{4}], \\ 4(1-r)[3r - 4(1-r)] & \text{if } r \in [\tfrac{3}{4}, 1]. \end{cases} \tag{A.8}
$$

Indeed, with the representation of μ as either W or $1 - W$ with probability $\tfrac{1}{2}$ of each, we have $\mathbb{P}[\mu \le r] = \tfrac{1}{2}\mathbb{P}[W \le r] + \tfrac{1}{2}(1 - \mathbb{P}[W < 1 - r])$. Assuming that W has a density f_W (which indeed it has, as we will show below), we obtain

$$
f_\mu(r) = \tfrac{1}{2}f_W(r) + \tfrac{1}{2}f_W(1 - r). \tag{A.9}
$$

Using the fact that $W \stackrel{\mathrm{D}}{=} S_1 + \tfrac{1}{4}S_2$, we use the joint distribution of (S_1, S_2) to calculate

$$
\mathbb{P}[W \le r] = \int_0^1 dx_1 \int_0^{1-x_1} dx_2 6(1 - x_1 - x_2)\mathbf{1}\left\{x_1 + \tfrac{1}{4}x_2 \le r\right\}
$$

$$
= \int_0^r dx_1 \int_0^{(4(r-x_1))\wedge(1-x_1)} dx_2 6(1 - x_1 - x_2).
$$

After some routine calculation, we then obtain

$$
\mathbb{P}[W \le r] = \begin{cases} 1 - \tfrac{4}{3}(1-r)^3 + \tfrac{1}{3}(1-4r)^3 & \text{if } r \in [0, \tfrac{1}{4}], \\ 1 - \tfrac{4}{3}(1-r)^3 & \text{if } r \in [\tfrac{1}{4}, 1]. \end{cases}
$$

Hence, W has density f_W given by

$$
f_W(r) = \begin{cases} 4(1-r)^2 - 4(1-4r)^2 & \text{if } r \in [0, \tfrac{1}{4}], \\ 4(1-r)^2 & \text{if } r \in [\tfrac{1}{4}, 1]. \end{cases}
$$

Then (A.8) follows from (A.9).

Appendix B. Continuity of random variables

In this appendix we give some results that will allow us to deduce the absolute continuity of certain distributions specified as solutions to fixed-point equations; specifically, we use these results in the proof of Proposition 3.4. The results in this section may well be known, but we were unable to find a reference for them in a form directly suitable for our application, and so we include the (short) proofs.

Lemma B.1. *Let X and Y be independent random variables such that X has an absolutely continuous distribution. Then, for any $a \neq 0$, we have $\mathbb{P}[XY = a] = 0$. Moreover, if $\mathbb{P}[Y = 0] = 0$ then XY is an absolutely continuous random variable.*

Proof. For the moment, assume that $\mathbb{P}[X < 0]$, $\mathbb{P}[X > 0]$, $\mathbb{P}[Y < 0]$, and $\mathbb{P}[Y > 0]$ are all positive. Take some $0 < c < d$. Then

$$
\begin{aligned}
&\mathbb{P}[XY \in (c, d)] \\
&\quad = \mathbb{P}[\log(X) + \log(Y) \in (\log c, \log d) \mid X > 0,\, Y > 0]\mathbb{P}[X > 0]\mathbb{P}[Y > 0] \\
&\qquad + \mathbb{P}[\log(-X) + \log(-Y) \in (\log c, \log d) \mid X < 0,\, Y < 0]\mathbb{P}[X < 0]\mathbb{P}[Y < 0].
\end{aligned}
$$

Note that conditioning X on the event $X > 0$ (or $X < 0$) preserves the continuity of X, and the independence of X and Y. Then, since the sum of two independent random variables at least one of which is absolutely continuous is also absolutely continuous (see [9, Theorem 5.9, page 230]), we have

$$
\mathbb{P}[\log(X) + \log(Y) \in (\log c, \log d) \mid X > 0,\, Y > 0] = \int_{\log c}^{\log d} f_+(x)\, \mathrm{d}x
$$

and

$$
\mathbb{P}[\log(-X) + \log(-Y) \in (\log c, \log d) \mid X < 0,\, Y < 0] = \int_{\log c}^{\log d} f_-(x)\, \mathrm{d}x
$$

for suitable probability densities f_+ and f_-. After the substitution $u = \mathrm{e}^x$, this yields

$$
\mathbb{P}[XY \in (c, d)] = \int_c^d \frac{\mathbb{P}[X > 0]\mathbb{P}[Y > 0]f_+(\log u) + \mathbb{P}[X < 0]\mathbb{P}[Y < 0]f_-(\log u)}{u}\, \mathrm{d}u.
$$

This expression is also valid if some of the probabilities for X and Y in the numerator of the integrand are 0. Therefore, we have, for any $0 < c < d$,

$$
\mathbb{P}[XY \in (c, d)] = \int_c^d f(u)\, \mathrm{d}u \tag{B.1}
$$

for some function $f(u)$ defined for $u > 0$. A similar argument applies to the case $c < d < 0$; then (B.1) is valid for any $c < d < 0$ as well, extending $f(u)$ for strictly negative u. In particular, it follows that $\mathbb{P}[XY = a] = 0$ for $a \neq 0$.

Now if $\mathbb{P}[Y = 0] = 0$ then $\mathbb{P}[XY \neq 0] = 1$. Then we can set $f(0) = 0$ so that (B.1) holds for *all* $c, d \in \mathbb{R}$. This completes the proof. $\quad\blacksquare$

Lemma B.2. *Suppose that a random variable L satisfies the distributional equation*

$$
L \overset{\mathrm{D}}{=}
\begin{cases}
Z_1 & \text{with probability } p_1, \\
\;\vdots & \\
Z_n & \text{with probability } p_n,
\end{cases}
$$

where $n \in \mathbb{N}$, $\sum_{i=1}^{n} p_i = 1$, $p_i > 0$, and each Z_i is an absolutely continuous random variable. Then L is absolutely continuous.

Proof. Suppose that Z_i has a density f_i. Then, for any $-\infty \le a < b \le +\infty$, we have

$$\mathbb{P}[L \in (a,b)] = \sum_{i=1}^{n} p_i \mathbb{P}[Z_i \in (a,b)] = \sum_{i=1}^{n} p_i \int_{a}^{b} f_i(x) \, \mathrm{d}x = \int_{a}^{b} \left[\sum_{i=1}^{n} p_i f_i(x) \right] \mathrm{d}x,$$

which yields the statement of lemma.

Acknowledgement

Parts of this work were done at the University of Strathclyde, where the third author was also employed, during a couple of visits by the second author, who is grateful for the hospitality of that institution.

References

[1] BENASSI, C. AND MALAGOLI, F. (2008). The sum of squared distances under a diameter constraint, in arbitrary dimension. *Arch. Math. (Basel)* **90,** 471–480.

[2] DE GIORGI, E. AND REIMANN, S. (2008). The α-beauty contest: choosing numbers, thinking intervals. *Games Econom. Behav.* **64,** 470–486.

[3] ERDŐS, P. (1940). On the smoothness properties of a family of Bernoulli convolutions. *Amer. J. Math.* **62,** 180–186.

[4] GRINFELD, M., KNIGHT, P. A. AND WADE, A. R. (2012). Rank-driven Markov processes. *J. Statist. Phys.* **146,** 378–407.

[5] HUGHES, B. D. (1995). *Random Walks and Random Environments; Vol. 1, Random Walks.* Clarendon Press, New York.

[6] JOHNSON, N. L. AND KOTZ, S. (1995). Use of moments in studies of limit distributions arising from iterated random subdivisions of an interval. *Statist. Prob. Lett.* **24,** 111–119.

[7] KEYNES, J. M. (1957). *The General Theory of Employment, Interest and Money.* Macmillan, London.

[8] KRAPIVSKY, P. L. AND REDNER, S. (2004). Random walk with shrinking steps. *Amer. J. Phys.* **72,** 591–598.

[9] MORAN, P. A. P. (1968). *An Introduction to Probability Theory.* Clarendon Press, Oxford.

[10] MOULIN, H. (1982). *Game Theory for the Social Sciences.* New York University Press.

[11] MURATOV, A. AND ZUYEV, S. (2013). LISA: locally interacting sequential adsorption. *Stoch. Models* **29,** 475–496.

[12] PEMANTLE, R. (2007). A survey of random processes with reinforcement. *Prob. Surv.* **4,** 1–79.

[13] PENROSE, M. D. AND WADE, A. R. (2004). Random minimal directed spanning trees and Dickman-type distributions. *Adv. Appl. Prob.* **36,** 691–714.

[14] PENROSE, M. D. AND WADE, A. R. (2008). Limit theory for the random on-line nearest-neighbor graph. *Random Structures Algorithms* **32,** 125–156.

[15] PILLICHSHAMMER, F. (2000). On the sum of squared distances in the Euclidean plane. *Arch. Math. (Basel)* **74,** 472–480.

[16] WITSENHAUSEN, H. S. (1974). On the maximum of the sum of squared distances under a diameter constraint. *Amer. Math. Monthly* **81,** 1100–1101.