Limit theory for the random on-line nearest-neighbour graph

Mathew D. Penrose*
Department of Mathematical Sciences, University of Bath,
Claverton Down, Bath BA2 7AY, England.

Andrew R. Wade†
Department of Mathematics, University of Bristol,
University Walk, Bristol BS8 1TW, England.

In the on-line nearest-neighbour graph (ONG), each point after the first in a sequence of points in $\mathbb{R}^d$ is joined by an edge to its nearest neighbour amongst those points that precede it in the sequence. We study the large-sample asymptotic behaviour of the total power-weighted length of the ONG on uniform random points in $(0,1)^d$. In particular, for $d = 1$ and weight exponent $\alpha > 1/2$, the limiting distribution of the centred total weight is characterized by a distributional fixed-point equation. As an ancillary result, we give exact expressions for the expectation and variance of the standard nearest-neighbour (directed) graph on uniform random points in the unit interval.

Key words and phrases: Nearest-neighbour graph; spatial network evolution; weak convergence; fixed-point equation; divide-and-conquer.

1 Introduction

Spatial graphs, defined on random point sets in Euclidean space, constructed by joining nearby points according to some deterministic rule, have been the subject of considerable recent interest. Examples of such graphs include the geometric graph, the minimal-length spanning tree, and the nearest-neighbour graph and its relatives. Many aspects of the

large-sample asymptotic theory for such graphs, which are locally determined in a certain sense, are by now quite well understood. See for example [11, 14, 15, 17–19, 24, 28].

Many real-world networks share several common features, including spatial structure, local construction (nearby nodes are more likely to be connected), and sequential growth (the network evolves over time via the addition of new nodes). In this paper our main object of interest is the *on-line nearest-neighbour graph*, which is one of the simplest models of network evolution that captures some of these features. We give a detailed description later. Recently, graphs with an 'on-line' structure, i.e. in which vertices are added sequentially and connected to existing vertices via some rule, have been the subject of considerable study in relation to the modelling of real-world networks. The literature is extensive (see for example [8, 13] for surveys), but mostly non-rigorous. Rigorous mathematical results are fewer in number, even for simple models, and the existing results concentrate on graph-theoretic rather than geometric properties (see e.g. [3, 6]).

The on-line nearest-neighbour graph (or ONG for short) is constructed on n points arriving sequentially in $\mathbb{R}^d$ by connecting each point after the first to its (Euclidean) nearest neighbour amongst the preceding points in the sequence. The ONG was apparently introduced in [3] as a simple growth model of the world wide web graph (for $d = 2$); see also [9]. When $d = 1$, the ONG is related to certain fragmentation processes, which are of separate interest in relation to, for example, molecular fragmentation (see e.g. [4], and references therein). The ONG in $d = 1$ is related to the so-called 'directed linear tree' considered in [16]. The higher dimensional ONG has also been studied [15]. Figure 1 shows a realization of the ONG on 50 simulated random points in the unit interval. Figure 2 below shows realizations of the planar and three-dimensional ONG, each on 50 simulated uniform random points.

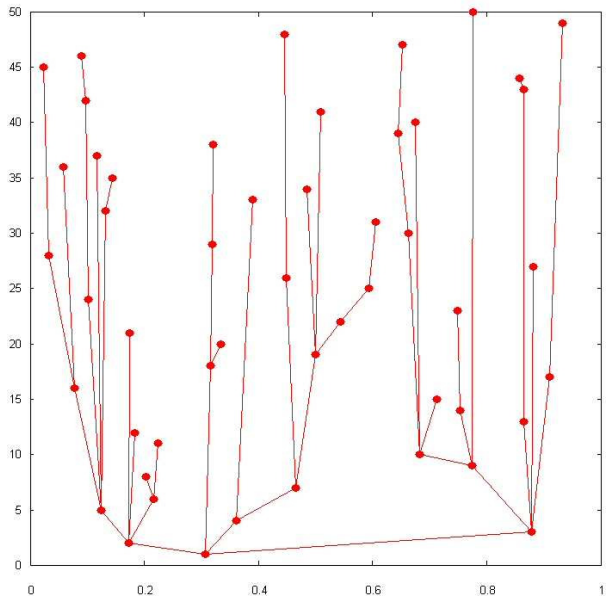

Figure 1: Realization of the ONG on 50 simulated uniform random points in the unit interval. The vertical axis gives the order in which the points arrive, and their position is given by the horizontal axis.

We consider the total power-weighted length of the ONG on uniform random points

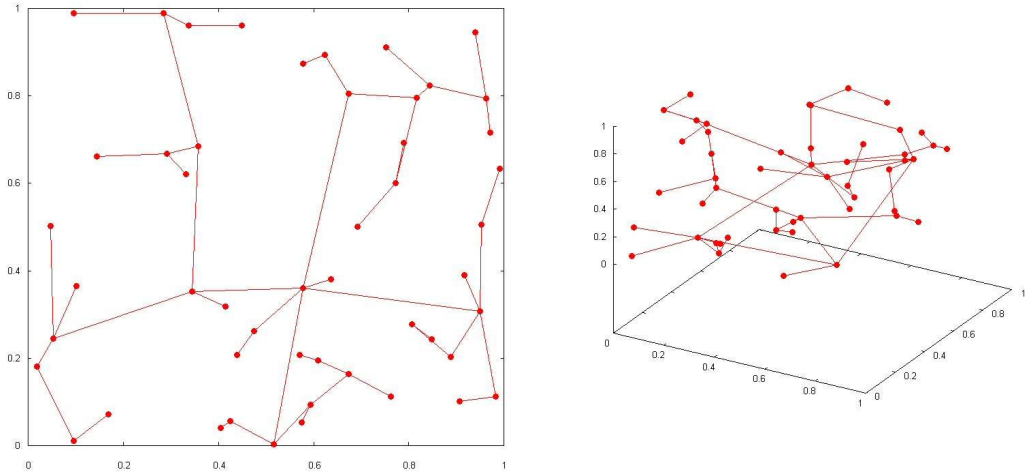

Figure 2: Realizations of the ONG on 50 simulated uniform random points in the unit square (left) and the unit cube (right).

in $(0,1)^d$, $d \in \mathbb{N}$. We are interested in large-sample asymptotics, as the number of points tends to infinity. Explicit laws of large numbers for the random ONG in $(0,1)^d$ are given in [26]. In the present paper we give further results on the limiting behaviour in general dimensions d.

The main part of the present paper is concerned with convergence in distribution results for the ONG. We give detailed properties of the random ONG on uniform random points in the unit interval ($d = 1$), and identify the limiting distribution of the centred total power-weighted length of the graph. When the weight exponent α is greater than $1/2$, this distribution is described in terms of a distributional fixed-point equation reminiscent of those encountered in, for example, the analysis of stochastic 'divide-and-conquer' or recursive algorithms. Such fixed-point distributional equalities, and the recursive algorithms from which they arise, have received considerable attention recently; see, for example, [2, 12, 22, 23].

On the other hand, we believe that for $\alpha \in (0, 1/2]$ the total weight, suitably centred and scaled, satisfies a central limit theorem (CLT). Penrose [15] gave such a result for $\alpha \in (0, 1/4)$. [27] addresses the case $\alpha \in (0, 1/2)$.

In this paper we also give new explicit results on the expectation and variance of the standard one-dimensional nearest-neighbour (directed) graph, in which each point is joined by a directed edge to its nearest neighbour, on uniform random points in the unit interval. This is related to our results on the one-dimensional ONG via the theory of Dirichlet spacings, which we use in our analysis.

2 Definitions and main results

Let $d \in \mathbb{N}$, and let $\|\cdot\|$ be the Euclidean norm on $\mathbb{R}^d$. For $d \in \mathbb{N}$, let

$$v_d := \pi^{d/2} \left[\Gamma\left(1 + (d/2)\right)\right]^{-1}, \tag{1}$$

the volume of the unit d-ball (see e.g. equation (6.50) of [10]).

Define w to be a weight function on edges, assigning weight $w(\mathbf{x}, \mathbf{y})$ to the edge between $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{y} \in \mathbb{R}^d$, such that $w : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty)$. A case of particular interest is when the weight is taken to be power-weighted Euclidean distance. In this case, for some $\alpha \geq 0$, we have the weight function given by, for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$,

$$w_\alpha(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|^\alpha. \tag{2}$$

2.1 The on-line nearest-neighbour graph

We now give a formal definition of the on-line nearest-neighbour graph (ONG). Let $d \in \mathbb{N}$. Suppose $(\mathbf{x}_1, \mathbf{x}_2, \ldots)$ are points in $(0,1)^d$, arriving sequentially; the ONG on vertex set $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is formed by connecting each point $\mathbf{x}_i$, $i = 2, 3, \ldots, n$ to its nearest neighbour (in the Euclidean sense) amongst the preceding points in the sequence (i.e. $\{\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}\}$), using the lexicographic ordering on $\mathbb{R}^d$ to break any ties. We call the resulting tree the ONG on $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$.

From now on we take the sequence of points to be *random*. Let $(\mathbf{U}_1, \mathbf{U}_2, \ldots)$ be a sequence of independent uniform random vectors on $(0,1)^d$. Then for $n \in \mathbb{N}$ take $\mathcal{U}_n = (\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_n)$. The points of sequence $\mathcal{U}_n$ then constitute a binomial point process consisting of n independent uniform random vectors on $(0,1)^d$. Denote the ONG constructed on $\mathcal{U}_n$ by ONG($\mathcal{U}_n$). Note that, with probability one, $\mathcal{U}_n$ has distinct inter-point distances so that the ONG on $\mathcal{U}_n$ is almost surely unique.

The ONG is of interest as a natural growth model for random spatial graphs; in particular it has been used (with $d = 2$) in the context of the world wide web graph (see [3]). In [15], stabilization techniques were used to prove that the total length (suitably centred and scaled) of the ONG on uniform random points in $(0,1)^d$ for $d > 4$ converges in distribution to a normal random variable. In [27] the CLT is extended to cover $d = 3, 4$. It is suspected that a CLT also holds for $d = 2$. On the other hand, when $d = 1$, the limit is not normal, as demonstrated by Theorem 2.2(ii) below.

For $d \in \mathbb{N}$ and $\alpha \geq 0$, let $\mathcal{O}^{d,\alpha}(\mathcal{U}_n)$ denote the total weight, with weight function w_α as given by (2), of ONG($\mathcal{U}_n$). Our results for the ONG in general dimensions are as follows, and constitute a distributional convergence result for $\alpha > d$, and asymptotic behaviour of the mean for $\alpha = d$. For the sake of completeness, we include the law of large numbers for $\alpha < d$ from [26] (Theorem 4 therein) as part (i) of the theorem below. By '$\xrightarrow{L^p}$' we denote convergence in L^p norm.

Theorem 2.1 *Suppose $d \in \mathbb{N}$. We have the following:*

(i) Suppose $0 \leq \alpha < d$. Then, as $n \to \infty$

$$n^{(\alpha-d)/d} \mathcal{O}^{d,\alpha}(\mathcal{U}_n) \xrightarrow{L^1} \frac{d}{d-\alpha} v_d^{-\alpha/d} \Gamma(1 + (\alpha/d)). \tag{3}$$

(ii) Suppose $\alpha > d$. Then, as $n \to \infty$,

$$\mathcal{O}^{d,\alpha}(\mathcal{U}_n) \longrightarrow W(d, \alpha), \tag{4}$$

where the convergence is in L^p, ($p \in \mathbb{N}$), and almost sure, and $W(d, \alpha)$ is a non-negative random variable with $\mathbb{E}[(W(d, \alpha))^k] < \infty$ for $k \in \mathbb{N}$.

(iii) Suppose $\alpha = d$. Then, as $n \to \infty$,

$$\mathbb{E}[\mathcal{O}^{d,d}(\mathcal{U}_n)] = v_d^{-1} \log n + o(\log n). \tag{5}$$

In particular (5) implies that $\mathbb{E}[\mathcal{O}^{1,1}(\mathcal{U}_n)] \sim (1/2) \log n$, a result given more precisely in Proposition 2.1 below. We prove Theorem 2.1(ii) and (iii) in Section 3.

Now we consider the particular case of the ONG in $d = 1$, where $\mathcal{U}_n$ is now a sequence of independent uniform random points in the unit interval $(0, 1)$. Let γ denote Euler's constant, so that $\gamma \approx 0.57721566$ and

$$\left(\sum_{i=1}^{k} \frac{1}{i} \right) - \log k = \gamma + O(k^{-1}). \tag{6}$$

The following result gives the expectation of the total weight of $\mathrm{ONG}(\mathcal{U}_n)$.

Proposition 2.1 *As $n \to \infty$, we have*

$$\begin{aligned}
\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n)] &= \frac{\Gamma(\alpha+1)}{1-\alpha} 2^{-\alpha} n^{1-\alpha} + \frac{2}{\alpha} - \frac{2^{-\alpha}(2-\alpha)}{\alpha(1-\alpha)} + O(n^{-\alpha}); \quad (0 < \alpha < 1) \\
\mathbb{E}[\mathcal{O}^{1,1}(\mathcal{U}_n)] &= \frac{1}{2} \log n + \frac{\gamma}{2} - \frac{1}{4} + o(1); \\
\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n)] &= \frac{2}{\alpha(\alpha+1)} \left(1 + \frac{2^{-\alpha}}{\alpha - 1} \right) + O(n^{1-\alpha}) \quad (\alpha > 1)
\end{aligned}$$

Proof. The proposition follows from Proposition 4.2 and Lemma 4.3 below. □

In Theorem 2.2 below, we present our main convergence in distribution results for the total weight of the ONG (centred, in some cases) in $d = 1$. The limiting distributions are of different types depending on the value of α in the weight function (2). In this paper, we restrict attention to $\alpha > 1/2$, and we define these limiting distributions in Theorem 2.2, in terms of distributional fixed-point equations (sometimes called *recursive distributional equations*, see [2]). Writing '$\stackrel{d}{=}$' for equality in distribution, these fixed-point equations are of the form

$$X \stackrel{d}{=} \sum_{r=1}^{k} A_r X^{\{r\}} + B, \tag{7}$$

where $k \in \mathbb{N}$, $X^{\{r\}}, r = 1, \ldots, k$, are independent copies of the random variable X, and $(A_1, \ldots, A_k, B)$ is a random vector, independent of $(X^{\{1\}}, \ldots, X^{\{k\}})$, satisfying the conditions

$$\mathbb{E} \sum_{r=1}^{k} |A_r|^2 < 1, \qquad \mathbb{E}[B] = 0, \qquad \mathbb{E}[B^2] < \infty. \tag{8}$$

Theorem 3 of Rösler [22] (proved by the contraction mapping theorem; see also [12, 23]) says that if (8) holds, there is a unique square-integrable distribution with mean zero satisfying the fixed-point equation (7), and this will guarantee uniqueness of solutions to all the distributional fixed-point equalities considered in the sequel.

We now define the distributions that will appear as limits in Theorem 2.2, in terms of (unique) solutions to fixed-point equations. In each case, U denotes a uniform random variable on $(0,1)$, independent of the other random variables on the right-hand side of the distributional equality. The fixed-point equations (9)–(12) are all of the form of (7), and hence define unique solutions.

We define $\tilde{J}_1$ by the distributional fixed-point equation

$$\tilde{J}_1 \stackrel{d}{=} \min\{U, 1-U\} + U\tilde{J}_1^{\{1\}} + (1-U)\tilde{J}_1^{\{2\}} + \frac{U}{2}\log U + \frac{1-U}{2}\log(1-U). \tag{9}$$

We shall see later (Proposition 4.4) that $\mathbb{E}[\tilde{J}_1] = 0$. For $\alpha > 1/2$, $\alpha \neq 1$, define $\tilde{J}_\alpha$ by

$$\tilde{J}_\alpha \stackrel{d}{=} U^\alpha \tilde{J}_\alpha^{\{1\}} + (1-U)^\alpha \tilde{J}_\alpha^{\{2\}} + \min\{U^\alpha, (1-U)^\alpha\} + \frac{2^{-\alpha}}{\alpha - 1}\left(U^\alpha + (1-U)^\alpha - 1\right). \tag{10}$$

Define the random variable $\tilde{H}_1$ by

$$\tilde{H}_1 \stackrel{d}{=} U\tilde{J}_1 + (1-U)\tilde{H}_1 + \frac{U}{2} + \frac{U}{2}\log U + \frac{1-U}{2}\log(1-U), \tag{11}$$

where $\tilde{J}_1$ has the distribution given by (9), and is independent of the $\tilde{H}_1$ on the right. We shall see later (Theorem 4.1) that $\mathbb{E}[\tilde{H}_1] = 0$. We give the first three moments of $\tilde{J}_1$ and $\tilde{H}_1$ in Table 2 later in this paper. For $\alpha > 1/2$, $\alpha \neq 1$, define $\tilde{H}_\alpha$ by

$$\tilde{H}_\alpha \stackrel{d}{=} U^\alpha \tilde{J}_\alpha + (1-U)^\alpha \tilde{H}_\alpha + U^\alpha\left(1 + \frac{2^{-\alpha}}{\alpha - 1}\right) + \left((1-U)^\alpha - 1\right)\left(\frac{1}{\alpha} + \frac{2^{-\alpha}}{\alpha(\alpha-1)}\right), \tag{12}$$

where $\tilde{J}_\alpha$ has the distribution given by (10) and is independent of the $\tilde{H}_\alpha$ on the right. We shall see later that, for $\alpha > 1$, the $\tilde{J}_\alpha$ and $\tilde{H}_\alpha$ defined in (10) and (12) arise as centred versions of the random variables J_α and H_α, respectively, satisfying the slightly simpler fixed-point equations (13) and (14) below, so that $\mathbb{E}[\tilde{J}_\alpha] = \mathbb{E}[\tilde{H}_\alpha] = 0$; see Proposition 4.5. For $\alpha > 1$, we have

$$J_\alpha \stackrel{d}{=} U^\alpha J_\alpha^{\{1\}} + (1-U)^\alpha J_\alpha^{\{2\}} + \min\{U^\alpha, (1-U)^\alpha\}. \tag{13}$$

Also for $\alpha > 1$, we have

$$H_\alpha \stackrel{d}{=} U^\alpha + U^\alpha J_\alpha + (1-U)^\alpha H_\alpha, \tag{14}$$

where J_α has distribution given by (13) and is independent of the H_α on the right. The expectations of J_α and H_α are given in Proposition 4.5. Note that the uniqueness of the $\tilde{J}_\alpha$ and $\tilde{H}_\alpha$ implies the uniqueness of J_α and H_α also.

Theorem 2.2 gives our main results for the ONG($\mathcal{U}_n$) in one dimension. Theorem 2.2 will follow as a corollary to Theorem 4.1, which we present later. Let $\tilde{\mathcal{O}}^{d,\alpha}(\mathcal{U}_n) := \mathcal{O}^{d,\alpha}(\mathcal{U}_n) - \mathbb{E}[\mathcal{O}^{d,\alpha}(\mathcal{U}_n)]$ be the centred total weight of the ONG on $\mathcal{U}_n$. For ease of notation, we define the following random variables. As before, U is uniform on $(0,1)$ and independent of the other variables on the right. For $1/2 < \alpha < 1$, let

$$\tilde{G}_\alpha \stackrel{d}{=} U^\alpha \tilde{H}_\alpha^{\{1\}} + (1-U)^\alpha \tilde{H}_\alpha^{\{2\}} + \left(U^\alpha + (1-U)^\alpha - \frac{2}{1+\alpha}\right)\left(\frac{1}{\alpha} - \frac{2^{-\alpha}}{\alpha(1-\alpha)}\right), \tag{15}$$

where $\tilde{H}_\alpha^{\{1\}}, \tilde{H}_\alpha^{\{2\}}$ are independent with distribution given by (12). Also let

$$\tilde{G}_1 \stackrel{d}{=} U\tilde{H}_1^{\{1\}} + (1-U)\tilde{H}_1^{\{2\}} + \frac{U}{2}\log U + \frac{1-U}{2}\log(1-U) + \frac{1}{4}, \tag{16}$$

where $\tilde{H}_1^{\{1\}}, \tilde{H}_1^{\{2\}}$ are independent with distribution given by (11). We write '$\stackrel{d}{\longrightarrow}$' for convergence in distribution. We prove Theorem 2.2 in Section 4.

Theorem 2.2 *(i) For $1/2 < \alpha < 1$, we have that, as $n \to \infty$,*

$$\tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n) \stackrel{d}{\longrightarrow} \tilde{G}_\alpha, \tag{17}$$

where $\tilde{G}_\alpha$ has distribution given by (15), and $\mathbb{E}[\tilde{G}_\alpha] = 0$.

(ii) For $\alpha = 1$, we have that, as $n \to \infty$,

$$\mathcal{O}^{1,1}(\mathcal{U}_n) - \frac{1}{2}(\gamma + \log n) + \frac{1}{4} \stackrel{d}{\longrightarrow} \tilde{G}_1, \tag{18}$$

where $\tilde{G}_1$ has distribution given by (16). Also, $\mathbb{E}[\tilde{G}_1] = 0$, $\mathrm{Var}[\tilde{G}_1] = (19 + 4\log 2 - 2\pi^2)/48 \approx 0.042362$, and $\mathbb{E}[\tilde{G}_1^3] \approx 0.00444287$.

(iii) For $\alpha > 1$, the distribution of the limit $W(1,\alpha)$ of (4) is given by

$$W(1,\alpha) \stackrel{d}{=} U^\alpha H_\alpha^{\{1\}} + (1-U)^\alpha H_\alpha^{\{2\}},$$

where $H_\alpha^{\{1\}}, H_\alpha^{\{2\}}$ are independent with the distribution given by (14).

Remarks. (a) Since the variables $W(d,\alpha)$ are non-negative, Theorem 2.1(ii) shows that when $\alpha > d$, the distributional limit of $\tilde{\mathcal{O}}^{d,\alpha}(\mathcal{U}_n)$ is not Gaussian. On the other hand, in Theorem 3.6 of [15], a CLT for $\tilde{\mathcal{O}}^{d,\alpha}(\mathcal{U}_n)$ is obtained for the case $\alpha \in (0, d/4)$. In the context of Theorem 2.1, the result of [15] implies that, provided $\alpha \in (0, d/4)$, $n^{(\alpha/d)-(1/2)}\tilde{\mathcal{O}}^{d,\alpha}(\mathcal{U}_n)$ is asymptotically normal as $n \to \infty$. In [15], it is remarked that it should be possible to extend the result to the case $\alpha \in (d/4, d/2)$ and perhaps $\alpha = d/2$ also. In [27] (using different techniques from those of the present paper) this extension is addressed, and it is also shown that $\tilde{\mathcal{O}}^{d,\alpha}(\mathcal{U}_n)$ converges in distribution to some limit variable when $\alpha \in (d, 2, d]$. In the special case $d = 1$, Theorem 2.2 above gives such a result, and moreover demonstrates that such a limit is non-Gaussian.

(b) A closely related 'directed' version of the one-dimensional ONG is the 'directed linear tree' (DLT) introduced in [16], in which each point is joined to its nearest neighbour to the *left* amongst those points preceding it in the sequence, if such points exist. In [16], results for the DLT with $\alpha \geq 1$ analogous to parts (ii) and (iii) of Theorem 2.2 were given. Following the methods of the present paper, one can obtain results for the DLT with $1/2 < \alpha < 1$ analogous to part (i) of Theorem 2.2.

(c) With some more detailed calculations (given in [25]), one can replace the error term $o(\log n)$ in (5) by $O(1)$ (see the remark in Section 3).

(d) Figure 3 is a plot of the estimated probability density function of $\tilde{G}_1$ given by (16). This was obtained by performing 10^5 repeated simulations of the ONG on a sequence of 10^3 uniform (simulated) random points on $(0,1)$. For each simulation, the expected value of $\mathcal{O}^{1,1}(\mathcal{U}_{10^3})$ was subtracted from the total length of the simulated ONG to give an approximate realization of the distributional limit. The density function was then estimated from the sample of 10^5 realizations. The simulated sample from which the density estimate was taken had sample mean $\approx 3 \times 10^{-3}$ and sample variance ≈ 0.0425, which are reasonably close to the expectation and variance of $\tilde{G}_1$.

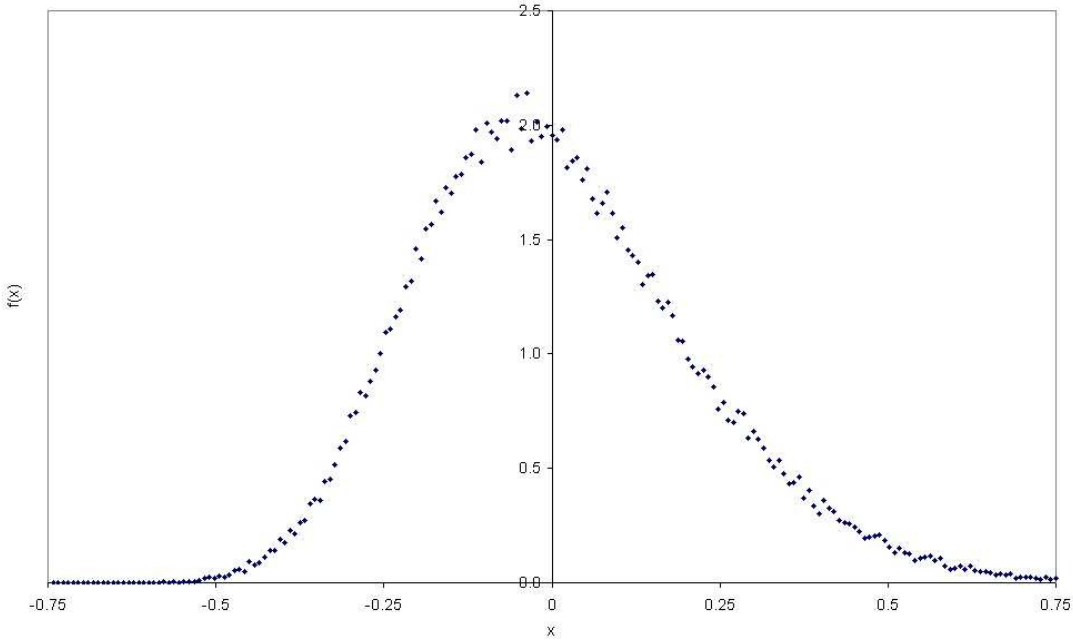

Figure 3: Estimated probability density function for $\tilde{G}_1$.

2.2 The nearest-neighbour (directed) graph

Our next result gives exact expressions for the expectation and variance of the total weight of the nearest-neighbour (directed) graph on n independent uniform random points in the unit interval. The nearest-neighbour (directed) graph on a point set $\mathcal{X}$ places a directed edge from each vertex to its nearest neighbour (in the Euclidean sense).

Let $\mathcal{L}_1^{1,\alpha}(\mathcal{X})$ denote the total weight, with weight function w_α given by (2), of the nearest-neighbour (directed) graph on vertex set $\mathcal{X} \subset (0,1)$. We use this notation to be consistent with [26], which presents explicit laws of large numbers for nearest-neighbour graphs including this one. For this section, let $\mathcal{U}_n$ denote the binomial point process consisting of n independent uniform random points in the unit interval. In this section with give explicit results for the expectation and variance of $\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)$.

Let ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ denote the Gauss hypergeometric function (see e.g. Chapter 15 of [1]) defined for $|z| < 1$ and $c \neq 0, -1, -2, \ldots$ by

$$ {}_2F_1(a,b;c;z) := \sum_{i=0}^{\infty} \frac{(a)_i (b)_i}{(c)_i i!} z^i, $$

where $(a)_i$ is Pochhammer's symbol $(a)_i := \Gamma(a+i)/\Gamma(a)$. For $n \in \{2,3,\ldots\}$, $\alpha > 0$, set

$$ J_{n,\alpha} := 6^{-\alpha-1} \frac{\Gamma(n+1)\Gamma(2+2\alpha)}{(1+\alpha)\Gamma(n+1+2\alpha)} \, {}_2F_1(-\alpha, 1+\alpha; 2+\alpha; 1/3). \tag{19} $$

Also, for $\alpha > 0$, set

$$ j_\alpha := 8 \lim_{n\to\infty} (n^{2\alpha} J_{n,\alpha}) = 8 \cdot 6^{-\alpha-1} \frac{\Gamma(2+2\alpha)}{1+\alpha} \, {}_2F_1(-\alpha, 1+\alpha; 2+\alpha; 1/3). \tag{20} $$

Theorem 2.3 *Suppose $\alpha > 0$. For $n \in \{2, 3, 4, \ldots\}$ we have*

$$\mathbb{E}[\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)] = ((n-2)2^{-\alpha} + 2)\frac{\Gamma(n+1)\Gamma(\alpha+1)}{\Gamma(n+\alpha+1)} \sim 2^{-\alpha}\Gamma(\alpha+1)n^{1-\alpha}, \tag{21}$$

as $n \to \infty$. Also, for $n \in \{4, 5, 6, \ldots\}$

$$\begin{aligned}\mathrm{Var}[\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)] &= \frac{\Gamma(n+1)}{\Gamma(n+2\alpha+1)}\left[\Gamma(2\alpha+1)(2 - 2\cdot 3^{-2\alpha} + 4^{-\alpha}n + 2\cdot 3^{-1-2\alpha}n)\right. \\ &\quad \left. +\Gamma(\alpha+1)^2(4 + 12\cdot 4^{-\alpha} - 12\cdot 2^{-\alpha} + 2^{2-\alpha}n - 7\cdot 4^{-\alpha}n + 4^{-\alpha}n^2)\right] \\ &\quad - \left(\mathbb{E}[\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)]\right)^2 + 8(n-3)J_{n,\alpha}, \end{aligned} \tag{22}$$

where $\mathbb{E}[\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)]$ is given by (21) and $J_{n,\alpha}$ is given by (19). Further, for $\alpha > 0$

$$n^{2\alpha-1}\mathrm{Var}[\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)] \to (4^{-\alpha} + 2\cdot 3^{-1-2\alpha})\Gamma(2\alpha+1) - 4^{-\alpha}(3+\alpha^2)\Gamma(\alpha+1)^2 + j_\alpha, \tag{23}$$

as $n \to \infty$, where j_α is given by (20).

Using (22), with (19), we obtain, for instance

$$\mathrm{Var}[\mathcal{L}_1^{1,1}(\mathcal{U}_n)] = \frac{2n^2 + 17n + 12}{12(n+1)^2(n+2)} = \frac{1}{6}n^{-1} + O(n^{-2}),$$

and

$$\mathrm{Var}[\mathcal{L}_1^{1,2}(\mathcal{U}_n)] = \frac{85n^3 + 3645n^2 + 7154n - 456}{108(n+1)^2(n+2)^2(n+3)(n+4)} = \frac{85}{108}n^{-3} + O(n^{-4}).$$

Table 1 below gives some values of $V_\alpha := \lim_{n\to\infty}(n^{2\alpha-1}\mathrm{Var}[\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)])$. We prove Theorem 2.3 in Section 5. One can obtain analogous explicit results in the case of $\mathcal{L}_1^{1,\alpha}(\mathcal{P}_n)$, where

α	$\frac{1}{2}$	1	2	3	4
V_α	$\frac{1}{2} + \sqrt{2}\arcsin\left(\frac{1}{\sqrt{3}}\right) - \frac{13\pi}{32} \approx 0.094148$	$\frac{1}{6}$	$\frac{85}{108}$	$\frac{149}{18}$	$\frac{135793}{972}$

Table 1: Some values of V_α.

$\mathcal{P}_n$ is a homogeneous Poisson point process of intensity n on $(0,1)$: see [25], where a "Poissonized" version of (23) is given.

The remainder of the present paper is organized as follows. Our results on the ONG in general dimensions (Theorem 2.1(ii) and (iii)) are proved in Section 3. The main body of this paper, Section 4, is devoted to the ONG in one dimension and the proof of Theorem 2.2. In Section 5 we prove Theorem 2.3. Finally, in the Appendix, we give the proofs of some technical lemmas which would otherwise interrupt the flow of the paper.

3 Proof of Theorem 2.1 (ii) and (iii)

Suppose $d \in \mathbb{N}$. For $i \in \mathbb{N}$, let $Z_i(d) := \mathcal{O}^{d,1}(\mathcal{U}_i) - \mathcal{O}^{d,1}(\mathcal{U}_{i-1})$, setting $\mathcal{O}^{d,1}(\mathcal{U}_0) := 0$. That is, $Z_i(d)$ is the gain in length of the ONG on a sequence of independent uniform random points in $(0,1)^d$ on the addition of the ith point. For $\mathbf{x} \in \mathbb{R}^d$ and point set $\mathcal{X} \subset \mathbb{R}^d$, let

$$d_1(\mathbf{x}; \mathcal{X}) := \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in \mathcal{X} \setminus \{\mathbf{x}\}\}$$

denote the (Euclidean) distance between $\mathbf{x}$ and its nearest neighbour in $\mathcal{X} \setminus \{\mathbf{x}\}$.

Lemma 3.1 *For $\alpha > 0$ and $d \in \mathbb{N}$, as $n \to \infty$,*

$$\mathbb{E}[(Z_n(d))^\alpha] = O(n^{-\alpha/d}). \tag{24}$$

Proof. We have

$$\mathbb{E}[(Z_n(d))^\alpha] = \mathbb{E}[(d_1(\mathbf{U}_1; \mathcal{U}_n))^\alpha] = n^{-\alpha/d}\mathbb{E}[(d_1(n^{1/d}\mathbf{U}_1; n^{1/d}\mathcal{U}_n))^\alpha],$$

which is $O(n^{-\alpha/d})$ (see the proof of Lemma 3.3 in [26]). □

Remark. We can obtain, by some more detailed analysis (see [25]),

$$\mathbb{E}[(Z_n(d))^\alpha] = (nv_d)^{-\alpha/d}\Gamma(1 + (\alpha/d)) + o(n^{-(\alpha/d)}).$$

Proof of Theorem 2.1(ii) and (iii). With the definition of $Z_i(d)$ in this section, let

$$W(d, \alpha) = \sum_{i=1}^{\infty} (Z_i(d))^\alpha.$$

The sum converges almost surely since it has non-negative terms and, by (24), has finite expectation for $\alpha > d$. Let $k \in \mathbb{N}$. By (24) and Hölder's inequality, there exists a constant $C \in (0, \infty)$ such that

$$\begin{aligned}
\mathbb{E}[(W(d,\alpha))^k] &= \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty}\cdots\sum_{i_k=1}^{\infty} \mathbb{E}[(Z_{i_1}(d))^\alpha (Z_{i_2}(d))^\alpha \cdots (Z_{i_k}(d))^\alpha] \\
&\le \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty}\cdots\sum_{i_k=1}^{\infty} \mathbb{E}[(Z_{i_1}(d))^{-\alpha k}]^{1/k} \cdots \mathbb{E}[(Z_{i_k}(d))^{-\alpha k}]^{1/k} \\
&\le C\sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty}\cdots\sum_{i_k=1}^{\infty} i_1^{-\alpha/d} i_2^{-\alpha/d} \cdots i_k^{-\alpha/d} < \infty,
\end{aligned}$$

since $\alpha/d > 1$. The L^k convergence then follows from the dominated convergence theorem, and we have part (ii) of Theorem 2.1.

Finally, for (iii) of Theorem 2.1, we have, when $\alpha = d$

$$n(Z_n(d))^d \stackrel{d}{=} (d_1(n^{1/d}\mathbf{U}_1; n^{1/d}\mathcal{U}_n))^d \xrightarrow{d} d_1(\mathbf{0}; \mathcal{H}_1)^d, \tag{25}$$

where $\mathcal{H}_1$ denotes a homogeneous Poisson point process of unit intensity on $(0,1)^d$, by the proof of Lemma 3.2 of [18]. Let $B(\mathbf{x}; r)$ denote the closed (Euclidean) d-ball with centre $\mathbf{x} \in \mathbb{R}^d$ and radius $r > 0$. For $n \in \mathbb{N}$ and $r \ge 0$, by scaling,

$$\mathbb{P}[(d_1(n^{1/d}\mathbf{U}_1; n^{1/d}\mathcal{U}_n))^d > r] = \mathbb{P}[n^{1/d} d_1(\mathbf{U}_1; \mathcal{U}_n) > r^{1/d}] = \mathbb{P}[d_1(\mathbf{U}_1; \mathcal{U}_n) > (r/n)^{1/d}].$$

For all $r > 0$ and $n \ge 2$ this is bounded by

$$\mathbb{P}(B(\mathbf{U}_1; (r/n)^{1/d}) \cap (0,1)^d \cap \mathcal{U}_n = \{\mathbf{U}_1\}) \le (1 - C(r/n))^{n-1},$$

since for some $C \in (0, \infty)$ depending only on d, $|B(\mathbf{U}_1; s) \cap (0,1)^d| \ge Cs^d$ for all $\mathbf{U}_1 \in (0,1)^d$ and all $s \le d$, and the probability in the display is 0 for $(r/n)^{1/d} \ge d$. Thus for some $C' < \infty$

$$\sup_{\mathbf{U}_1 \in (0,1)^d; n \ge 1} \mathbb{P}((d_1(n^{1/d}\mathbf{U}_1; n^{1/d}\mathcal{U}_n))^d > r) \le C' \exp(-r/C').$$

It follows that the sequence $(d_1(n^{1/d}\mathbf{U}_1; n^{1/d}\mathcal{U}_n))^d$ is uniformly integrable, and so by (25)

$$\mathbb{E}[n(Z_n(d))^d] \to \mathbb{E}[(d_1(\mathbf{0}; \mathcal{H}_1))^d] = v_d^{-1},$$

where the last inequality follows by a simple computation, or by equation (19) of [26]. So $\mathbb{E}[(Z_n(d))^d] = n^{-1}(v_d^{-1} + h(n))$ where $h(n) \to 0$ as $n \to \infty$. Thus

$$\mathbb{E}\sum_{i=1}^{n}(Z_i(d))^d = \sum_{i=1}^{n} i^{-1}(v_d^{-1} + h(i)) = v_d^{-1}\log n + o(\log n),$$

and so we have (5), completing the proof of Theorem 2.1. □

4 The ONG in $d = 1$

4.1 Notation and results

In this section we analyse the ONG in the interval $(0, 1)$. Theorem 2.2 will follow from the main result of this section, Theorem 4.1 below. We introduce our notation.

For any finite sequence of points $\mathcal{T}_n = (x_1, x_2, \ldots, x_n) \in [0, 1]^n$ with distinct inter-point distances, we construct the ONG as follows. Insert the points $x_1, x_2, \ldots$ into $[0, 1]$ in order, one at a time. We join a new point by an edge to its nearest neighbour among those already present, provided that such a point exists. In other words, for each point x_i, $i \geq 2$, we join x_i by an edge to the point of $\{x_j : 1 \leq j < i\}$ that minimizes $|x_i - x_j|$. In this way we construct a tree rooted at x_1, which we denote by $\mathrm{ONG}(\mathcal{T}_n)$. Denote the total weight (under weight function w_α given by (2), $\alpha > 0$) of $\mathrm{ONG}(\mathcal{T}_n)$ by $\mathcal{O}^{1,\alpha}(\mathcal{T}_n)$, to be consistent with our previous notation.

For what follows, our main interest is the case in which $\mathcal{T}_n$ is a random vector in $[0, 1]^n$. In this case, set $\tilde{\mathcal{O}}^{1,\alpha}(\mathcal{T}_n) := \mathcal{O}^{1,\alpha}(\mathcal{T}_n) - \mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{T}_n)]$, the centred total weight of the ONG on $\mathcal{T}_n$. Let $(U_1, U_2, U_3, \ldots)$ be a sequence of independent uniformly distributed random variables in $(0, 1)$, and for $n \in \mathbb{N}$ set $\mathcal{U}_n := (U_1, U_2, \ldots, U_n)$. Given $\mathcal{U}_n$, we define the augmented sequences $\mathcal{U}_n^0 = (0, U_1, \ldots, U_n)$ and $\mathcal{U}_n^{0,1} = (0, 1, U_1, \ldots, U_n)$. Notice that $\mathrm{ONG}(\mathcal{U}_n^{0,1})$ and $\mathrm{ONG}(\mathcal{U}_n^0)$ both give a tree rooted at 0, and that in $\mathrm{ONG}(\mathcal{U}_n^{0,1})$ the first edge is from 1 to 0.

We now state the main result of this section, from which Theorem 2.2 will follow. The convergence of joint distribution results in (26) and (27) are given in more detail, complete with joint distribution fixed-point representation, in Propositions 4.3 and 4.4.

Theorem 4.1 *(i) For $1/2 < \alpha < 1$, we have that, as $n \to \infty$,*

$$(\tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}), \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0), \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n)) \xrightarrow{d} (\tilde{J}_\alpha, \tilde{H}_\alpha, \tilde{G}_\alpha), \tag{26}$$

where $\tilde{J}_\alpha$, $\tilde{H}_\alpha$, $\tilde{G}_\alpha$ are jointly distributed random variables with marginal distributions given by (10), (12), (15) respectively.

(ii) For $\alpha = 1$, we have that, as $n \to \infty$,

$$(\tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0,1}), \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^0), \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n)) \xrightarrow{d} (\tilde{J}_1, \tilde{H}_1, \tilde{G}_1), \tag{27}$$

where $\tilde{J}_1$, $\tilde{H}_1$, $\tilde{G}_1$ are jointly distributed random variables with marginal distributions given by (9), (11), (16) respectively. The first three moments of $\tilde{J}_1$, $\tilde{H}_1$ and $\tilde{G}_1$ are given in Table 2. Further, the variables on the right-hand side of (27) satisfy $\mathrm{Cov}(\tilde{J}_1, \tilde{H}_1) = ((9 + 6\log 2)/32) - (\pi^2/24) \approx -1.84204 \times 10^{-5}$, $\mathrm{Cov}(\tilde{G}_1, \tilde{H}_1) = ((35 + 10\log 2)/48) - (\pi^2/24) \approx 0.0255536$, *and* $\mathrm{Cov}(\tilde{G}_1, \tilde{J}_1) = ((7 + 4\log 2)/24) - (\pi^2/24) \approx -4.04232 \times 10^{-3}$.

(iii) For $\alpha > 1$, we have that, as $n \to \infty$,

$$\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1}) \to 1 + J_\alpha; \quad \mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0}) \to H_\alpha,$$

where the convergence is almost sure and in L^p, $p \in \mathbb{N}$, and the distributions of J_α and H_α are given by (13) and (14) respectively.

	$\mathbb{E}[\cdot]$	$\mathrm{Var}[\cdot]$	$\mathbb{E}[(\cdot)^3]$
$\tilde{J}_1$	0	$((1 + \log 2)/4) - (\pi^2/24) \approx 0.012053$	≈ -0.00005733
$\tilde{H}_1$	0	$((3 + \log 2)/8) - (\pi^2/24) \approx 0.050410$	≈ 0.00323456
$\tilde{G}_1$	0	$((19 + 4\log 2)/48) - (\pi^2/24) \approx 0.042362$	≈ 0.00444287

Table 2: First three moments for the random variables $\tilde{J}_1$, $\tilde{H}_1$, $\tilde{G}_1$.

Our method for establishing convergence in distribution results is based on the *recursive* nature of the ONG. Essential is its *self-similarity* (scaling property). In terms of the total weight, this says that for any $t \in (0,1)$, if $V_1, \ldots, V_n$ are independent and uniformly distributed on $(0,t)$, then the distribution of $\mathcal{O}^{1,\alpha}(V_1, \ldots, V_n)$ is the same as that of $t^\alpha \mathcal{O}^{1,\alpha}(U_1, \ldots, U_n)$.

Write $U = U_1$ for the position of the first arrival. For ease of notation, denote

$$Y_n := \mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1}) - 1, \tag{28}$$

where by subtracting 1 we discount the length of the edge from 1 to 0. Using the self-similarity of the ONG, and conditioning on the first arrival, we have the following relations:

$$\mathcal{O}^{1,\alpha}(\mathcal{U}_n) \stackrel{d}{=} U^\alpha \mathcal{O}^{1,\alpha}_{\{1\}}(\mathcal{U}^0_{N(n)}) + (1-U)^\alpha \mathcal{O}^{1,\alpha}_{\{2\}}(\mathcal{U}^0_{n-1-N(n)}), \tag{29}$$

$$\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0) \stackrel{d}{=} U^\alpha \mathcal{O}^{1,\alpha}_{\{1\}}(\mathcal{U}^{0,1}_{N(n)}) + (1-U)^\alpha \mathcal{O}^{1,\alpha}_{\{2\}}(\mathcal{U}^0_{n-1-N(n)}), \tag{30}$$

$$Y_n \stackrel{d}{=} (\min\{U, 1-U\})^\alpha + U^\alpha Y^{\{1\}}_{N(n)} + (1-U)^\alpha Y^{\{2\}}_{n-1-N(n)}, \tag{31}$$

where, given U, $N(n) \sim \mathrm{Bin}(n-1, U)$ gives the number of points of $U_2, U_3, \ldots, U_n$ that arrive to the left of $U_1 = U$. Given U and $N(n)$, $\mathcal{O}^{1,\alpha}_{\{1\}}(\cdot)$ and $\mathcal{O}^{1,\alpha}_{\{2\}}(\cdot)$ are independent copies of $\mathcal{O}^{1,\alpha}(\cdot)$. Also, given U and $N(n)$, $Y^{\{1\}}_{N(n)}$ and $Y^{\{2\}}_{n-1-N(n)}$ are independent with the distribution of $Y_{N(n)}$ and $Y_{n-1-N(n)}$, respectively.

For $\alpha > 1$, we prove almost sure and L^p ($p \in \mathbb{N}$) convergence of $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0)$ and $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1})$, in the same way as in the proof of Theorem 2.1(ii), and thereby obtain the corresponding result for $\mathcal{O}^{1,\alpha}(\mathcal{U}_n)$. The relations (29), (30) and (31) will then enable us to prove the desired results for $\alpha > 1$.

For $1/2 < \alpha \leq 1$, we use a result of Neininger and Rüschendorf [12] on limit theorems for 'divide and conquer' recurrences. However, we cannot apply this directly to (29) to obtain the convergence of $\mathcal{O}^{1,\alpha}(\mathcal{U}_n)$, since (29) is not of the required form; the variables on the right are not of the same type as the variable on the left. On the other hand, we see that (31) is of the desired form. This will be the basis of our analysis for $1/2 < \alpha \leq 1$.

Indeed, by considering a vector defined in terms of all three of $\mathcal{O}^{1,\alpha}(\mathcal{U}_n)$, $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0)$, and $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1})$, we obtain the recurrence relation (72) below. We can then apply the result of [12]. This is why we need to consider $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0)$ and $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1})$ in addition to $\mathcal{O}^{1,\alpha}(\mathcal{U}_n)$.

Since it is the basis of some of our proofs, in Lemma 4.1 we recall the following result of Neininger and Rüschendorf, which is the $s = 2$, dimension 3 case of Theorem 4.1 in [12]. Let (X_n) be a sequence of random 3-vectors satisfying the recurrence

$$X_n \stackrel{d}{=} \sum_{r=1}^{k} A_r(n) X^{\{r\}}_{I_r(n)} + B(n), \tag{32}$$

for $n \geq n_0 \geq 1$, where $k \in \mathbb{N}$, $I(n) := (I_1(n), \ldots, I_k(n))$ is a vector of random cardinalities $I_r(n) \in \{0, \ldots, n\}$, $A_1(n), \ldots, A_k(n)$ are random 3×3 matrices, $B(n)$ is a random 3-vector, $(X_n^{\{1\}}), \ldots, (X_n^{\{k\}})$ are independent copies of (X_n), and the vectors $(A_1(n), \ldots, A_k(n), B(n), I(n))$, $(X_n^{\{1\}})$, $\ldots$, $(X_n^{\{k\}})$ are independent.

For $n \in \mathbb{N}$, let $M_n := \mathbb{E}[X_n]$. We renormalize (32) by setting $\tilde{X}_n := X_n - M_n$. Then $(\tilde{X}_n)$ satisfies the modified recurrence

$$\tilde{X}_n \stackrel{d}{=} \sum_{r=1}^{k} A_r(n) \tilde{X}^{(r)}_{I_r(n)} + \tilde{B}(n), \tag{33}$$

for $n \geq n_0$, where for each $r \in \{1, \ldots, k\}$,

$$\tilde{X}^{\{r\}}_{I_r(n)} = X^{\{r\}}_{I_r(n)} - M_{I_r(n)} = X^{\{r\}}_{I_r(n)} - \mathbb{E}[X_{I_r(n)} | I_r(n)], \tag{34}$$

and

$$\tilde{B}(n) = B(n) - \mathbb{E}[X_n] + \sum_{r=1}^{k} A_r(n) \mathbb{E}[X_{I_r(n)} | I_r(n)]. \tag{35}$$

Let $\| \cdot \|_{\mathrm{op}}$ denote the operator norm; i.e. for a 3×3 matrix A, $\|A\|_{\mathrm{op}} := \sup_{\|z\|=1} \|Az\|$, where the supremum is over (column) 3-vectors with norm 1.

Lemma 4.1 *[12] Suppose (X_n) satisfies (32). Set $\tilde{X}_n := X_n - M_n$, so that $\tilde{X}_n$ satisfies (33). Suppose that $(\tilde{X}_n)$ is square-integrable, and*

$$(A_1(n), \ldots, A_k(n), \tilde{B}(n)) \xrightarrow{L^2} (A_1, \ldots, A_k, B), \text{ as } n \to \infty; \tag{36}$$

$$\mathbb{E} \sum_{r=1}^{k} \|A_r\|_{\mathrm{op}}^2 < 1; \tag{37}$$

$$\mathbb{E}[\mathbf{1}_{\{I_r(n) \leq \ell\} \cup \{I_r(n) = n\}} \|A_r(n)\|_{\mathrm{op}}^2] \to 0, \tag{38}$$

as $n \to \infty$, for all $\ell \in \mathbb{N}$ and $r = 1, \ldots, k$. Then as $n \to \infty$, $\tilde{X}_n$ converges to a limit $\tilde{X}$ in Zolotarev ζ_2 metric, and hence in distribution (see e.g. Chapter 14 of [21]), where X is the unique square-integrable solution to the fixed-point equation

$$X \stackrel{d}{=} \sum_{r=1}^{k} A_r X^{\{r\}} + B,$$

where $(A_1, \ldots, A_k, B), X^{\{1\}}, \ldots, X^{\{k\}}$ are independent and each $X^{\{r\}}$ is distributed as X.

The outline of the remainder of this section is as follows. In Section 4.2 below, we give a discussion of the theory of spacings, which will be very useful in the sequel. In Section 4.3 we begin our analysis of the ONG with some preliminary results, based on the discussion in Section 4.2. Then, in Sections 4.4, 4.5 and 4.6 we give results on $\mathcal{O}^{1,\alpha}(\cdot)$ when $1/2 < \alpha < 1$, $\alpha = 1$, and $\alpha > 1$ respectively. Finally, in Section 4.7 we give proofs of Theorems 4.1 and 2.2.

4.2 Spacings

The one-dimensional models considered in this paper (the ONG and the standard nearest-neighbour graph) are defined in terms of the *spacings* of points in the unit interval. Thus the theory of so-called Dirichlet spacings will be useful. For some general references on spacings, see for example [20]. A large number of statistical tests are based on spacings, see e.g. [7] for a few examples.

Let $U_1, U_2, \ldots$ be independent uniform random variables on $(0, 1)$. Let $n \in \mathbb{N}$. Given $\{U_1, \ldots, U_n\} \subseteq (0, 1)$, denote the order statistics of $U_1, \ldots, U_n$, taken in increasing order, as $U^n_{(1)}, U^n_{(2)}, \ldots, U^n_{(n)}$. Thus $(U^n_{(1)}, \ldots, U^n_{(n)})$ is a nondecreasing sequence, forming a permutation of the original $(U_1, \ldots, U_n)$.

The points $U_1, \ldots, U_n$ divide $[0, 1]$ into $n + 1$ intervals. Denote the intervals between points by $I^n_j := (U^n_{(j-1)}, U^n_{(j)})$ for $j = 1, 2, \ldots, n+1$, where we set $U^n_{(0)} := 0$ and $U^n_{(n+1)} := 1$. Let the widths of these intervals (the spacings) be

$$S^n_j := |I^n_j| = U^n_{(j)} - U^n_{(j-1)},$$

for $j = 1, 2, \ldots, n + 1$. For $n \in \mathbb{N}$, let $\Delta_n \subset \mathbb{R}^n$ denote the n-dimensional simplex, that is

$$\Delta_n := \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \geq 0,\ 1 \leq i \leq n;\ \sum_{i=1}^{n} x_i \leq 1 \right\}.$$

By the definition of S^n_j, we have that $S^n_j \geq 0$ for $j = 1, \ldots, n + 1$ and $\sum_{j=1}^{n+1} S^n_j = 1$. So we see that the vector $(S^n_1, S^n_2, \ldots, S^n_{n+1})$ is completely specified by any n of its $n + 1$ components, and any such n-vector belongs to the simplex Δ_n. It is not hard to show that any such n-vector is, in fact, uniformly distributed over the simplex. Hence $(S^n_1, \ldots, S^n_n)$ is uniform over the simplex Δ_n, and $S^n_{n+1} = 1 - \sum_{i=1}^{n} S^n_i$.

Thus $(S^n_1, S^n_2, \ldots, S^n_{n+1})$ has the symmetric *Dirichlet distribution* with parameter 1 (see, e.g., [5], p. 246), and any n-vector of the S^n_j has the Dirichlet density

$$f(x_1, \ldots, x_n) = n!, \quad (x_1, \ldots, x_n) \in \Delta_n. \tag{39}$$

In particular, the spacings S_j^n, $j = 1, \ldots, n+1$ are *exchangeable* – the distribution of $(S_1^n, S_2^n, \ldots, S_{n+1}^n)$ is invariant under any permutation of its components.

By integrating out over the simplex, from (39) one can readily obtain the marginal distributions for the spacings. Thus, for $n \geq 1$, a single spacing has density

$$f(x_1) = n(1-x_1)^{n-1}, \quad 0 \leq x_1 \leq 1, \tag{40}$$

while for $n \geq 2$, any two spacings have joint density

$$f(x_1, x_2) = n(n-1)(1-x_1-x_2)^{n-2}, \quad (x_1, x_2) \in \Delta_2, \tag{41}$$

and for $n \geq 3$ any three spacings have joint density

$$f(x_1, x_2, x_3) = n(n-1)(n-2)(1-x_1-x_2-x_3)^{n-3}, \quad (x_1, x_2, x_3) \in \Delta_3. \tag{42}$$

In fact, for $1 \leq k < n$ and $x_i \geq 0$, $i = 1, \ldots, k$ such that $\sum_{i=1}^k x_i \leq 1$,

$$\mathbb{P}[S_i^n > x_i, i = 1, \ldots, k] = \left(1 - \sum_{i=1}^k x_i\right)^n; \tag{43}$$

this is $n!$ times the volume of a smaller simplex. Using the fact that (see, e.g., 6.2.1 in [1])

$$\int_0^1 t^{a-1}(1-t)^{b-1}\mathrm{d}t = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \tag{44}$$

for $a > 0$, $b > 0$, it then follows from (40) that, for $\beta > 0$, $n \geq 1$

$$\mathbb{E}\left[(S_1^n)^\beta\right] = \frac{\Gamma(n+1)\Gamma(\beta+1)}{\Gamma(n+\beta+1)}, \tag{45}$$

and from (41) that for $\beta > 0$, $n \geq 2$

$$\mathbb{E}\left[(S_1^n)^\beta (S_2^n)^\beta\right] = \frac{\Gamma(n+1)\Gamma(\beta+1)^2}{\Gamma(n+2\beta+1)}. \tag{46}$$

When considering our nearest-neighbour graphs, we will encounter the *minimum* of two (or more) spacings. The following results will also be needed in Section 5.

Lemma 4.2 *For $n \geq 1$,*

$$\min\{S_1^n, S_2^n\} \stackrel{d}{=} S_1^n/2. \tag{47}$$

For $n \geq 2$,

$$(S_1^n, \min\{S_2^n, S_3^n\}) \stackrel{d}{=} (S_1^n, S_2^n/2). \tag{48}$$

Finally, for $n \geq 3$

$$(\min\{S_1^n, S_2^n\}, \min\{S_3^n, S_4^n\}) \stackrel{d}{=} (S_1^n/2, S_2^n/2); \text{ and} \tag{49}$$

$$\min\{S_1^n, S_2^n, S_3^n\} \stackrel{d}{=} S_1^n/3. \tag{50}$$

Proof. These results follow from the observation (43), or directly; for example we give the direct proof of (47) (similar calculations can be used in the other cases). Suppose $n \geq 2$. From (41), we have, for $0 \leq r \leq 1/2$

$$\begin{aligned}\mathbb{P}[\min\{S_1^n, S_2^n\} > r] &= \mathbb{P}[S_1^n > r,\ S_2^n > r] \\ &= n(n-1)\int_r^{1-r} \mathrm{d}x_1 \int_r^{1-x_1} (1-x_1-x_2)^{n-2}\mathrm{d}x_2 \\ &= (1-2r)^n = \mathbb{P}[S_1^n > 2r],\end{aligned}$$

and so we have (47). □

4.3 Preparatory results

We now return to the ONG. We make use of the discussion of spacings in Section 4.2. For $n \in \mathbb{N}$ let Z_n, H_n and T_n denote the random variables given by the gain in length, on the addition of the point U_n, of the ONG on $\mathcal{U}_{n-1}$, $\mathcal{U}_{n-1}^0$ and $\mathcal{U}_{n-1}^{0,1}$ respectively. That is, with the convention $\mathcal{O}^{1,1}(\mathcal{U}_0) = \mathcal{O}^{1,1}(\mathcal{U}_0^0) = 0$ and $\mathcal{O}^{1,1}(\mathcal{U}_0^{0,1}) = 1$, for $n \in \mathbb{N}$ set

$$\begin{aligned}Z_n &:= \mathcal{O}^{1,1}(\mathcal{U}_n) - \mathcal{O}^{1,1}(\mathcal{U}_{n-1}), \\ H_n &:= \mathcal{O}^{1,1}(\mathcal{U}_n^0) - \mathcal{O}^{1,1}(\mathcal{U}_{n-1}^0), \\ T_n &:= \mathcal{O}^{1,1}(\mathcal{U}_n^{0,1}) - \mathcal{O}^{1,1}(\mathcal{U}_{n-1}^{0,1}).\end{aligned} \tag{51}$$

Thus, for example, in the ONG($\mathcal{U}_n^{0,1}$) with weight function w_α as given by (2), the nth edge to be added has weight T_n^α.

We will make use of the following discussion for the proof of Lemma 4.3 below. For $\alpha > 0$, with the definitions at (51), we have that

$$\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0) - \mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1}) = -1 + \sum_{i=1}^n (H_i^\alpha - T_i^\alpha), \quad \text{and} \tag{52}$$

$$\mathcal{O}^{1,\alpha}(\mathcal{U}_n) - \mathcal{O}^{1,\alpha}(\mathcal{U}_n^0) = \sum_{i=1}^n (Z_i^\alpha - H_i^\alpha) = -H_1^\alpha + \sum_{i=2}^n (Z_i^\alpha - H_i^\alpha), \tag{53}$$

since $Z_1 = 0$. Consider the arrival of the point U_n. For any n, T_n and H_n are the same unless the point U_n falls in the right-hand half of the rightmost interval I_n^{n-1} of width S_n^{n-1}. Denote this latter event by E_n. Given S_n^{n-1}, the probability of E_n is $S_n^{n-1}/2$. Given S_n^{n-1}, and given that E_n occurs, the value of T_n is given by $(1-V_n)S_n^{n-1}/2$ and the value of H_n by $(1+V_n)S_n^{n-1}/2$, where $V_n = 1 + 2(U_n - 1)/S_n^{n-1}$ is uniform on $(0,1)$ given E_n. So we have that, for $n \in \mathbb{N}$, given S_n^{n-1}

$$H_n^\alpha - T_n^\alpha = \mathbf{1}_{E_n} \left(\frac{S_n^{n-1}}{2}\right)^\alpha ((1+V_n)^\alpha - (1-V_n)^\alpha), \tag{54}$$

where E_n is an event with probability $S_n^{n-1}/2$. A similar argument (based this time on the *leftmost* spacing) yields that, for $n \geq 2$

$$Z_n^\alpha - H_n^\alpha = \mathbf{1}_{F_n} \left(\frac{S_1^{n-1}}{2}\right)^\alpha ((1+W_n)^\alpha - (1-W_n)^\alpha), \tag{55}$$

where F_n is an event with probability $S_1^{n-1}/2$ and, given F_n, W_n is uniform on $(0,1)$.

We will need the following asymptotic expansion, which follows from Stirling's formula (see e.g. 6.1.37 in [1]). For any $\beta > 0$, as $n \to \infty$,

$$\frac{\Gamma(n+1)}{\Gamma(n+1+\beta)} = n^{-\beta} - \frac{1}{2}\beta(\beta+1)n^{-\beta-1} + O(n^{-\beta-2}). \tag{56}$$

Lemma 4.3 *For $\alpha > 0$ and $n \geq 2$, we have that*

$$\begin{aligned}\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0) - \mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1})] &= \frac{1-2^{-\alpha}-\alpha}{\alpha} + (2^{-\alpha}-1)\frac{\Gamma(\alpha)\Gamma(n+1)}{\Gamma(n+1+\alpha)} \\ &= \frac{1-2^{-\alpha}-\alpha}{\alpha} + O(n^{-\alpha}),\end{aligned} \tag{57}$$

and

$$\begin{aligned}\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n) - \mathcal{O}^{1,\alpha}(\mathcal{U}_n^0)] &= \frac{1-2^{-\alpha}-\alpha}{\alpha(1+\alpha)} + (2^{-\alpha}-1)\frac{\Gamma(\alpha)\Gamma(n+1)}{\Gamma(n+1+\alpha)} \\ &= \frac{1-2^{-\alpha}-\alpha}{\alpha(1+\alpha)} + O(n^{-\alpha}).\end{aligned} \tag{58}$$

Proof. Suppose $\alpha > 0$. From (54) we have that for $n \in \mathbb{N}$

$$\mathbb{E}[H_n^\alpha - T_n^\alpha | S_n^{n-1}] = (S_n^{n-1})^{1+\alpha}\left(\frac{1-2^{-\alpha}}{1+\alpha}\right).$$

So by (45) we have that

$$\mathbb{E}[H_n^\alpha - T_n^\alpha] = \frac{(1-2^{-\alpha})\Gamma(1+\alpha)\Gamma(n)}{\Gamma(n+1+\alpha)}.$$

Thus, from (52),

$$\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0) - \mathcal{O}^{1,1}(\mathcal{U}_n^{0,1})] = -1 + \mathbb{E}\sum_{i=1}^n (H_i^\alpha - T_i^\alpha) = -1 + \sum_{i=1}^n \frac{(1-2^{-\alpha})\Gamma(1+\alpha)\Gamma(i)}{\Gamma(i+1+\alpha)}.$$

Then the first line of (57) follows by induction on n, with the second line of (57) following by (56). Similarly, from (55)

$$\mathbb{E}[Z_n^\alpha - H_n^\alpha] = \frac{(1-2^{-\alpha})\Gamma(1+\alpha)\Gamma(n)}{\Gamma(n+1+\alpha)},$$

for $n \geq 2$, while $\mathbb{E}[H_1^\alpha] = \mathbb{E}[U_1^\alpha] = (\alpha+1)^{-1}$ and $Z_1 = 0$. With (56), (58) follows. □

Lemma 4.4 *(i) For $n \in \mathbb{N}$, T_n as defined at (51) has distribution function F_n given by $F_n(t) = 0$ for $t < 0$, $F_n(t) = 1$ for $t \geq 1/2$, and $F_n(t) = 1-(1-2t)^n$ for $0 \leq t \leq 1/2$.*

(ii) For $\beta > 0$,

$$\mathbb{E}[T_n^\beta] = 2^{-\beta}\frac{\Gamma(n+1)\Gamma(\beta+1)}{\Gamma(n+\beta+1)}. \tag{59}$$

In particular,

$$\mathbb{E}[T_n] = \frac{1}{2(n+1)}; \quad \mathrm{Var}[T_n] = \frac{n}{4(n+1)^2(n+2)}. \tag{60}$$

(iii) For $\beta > 0$, as $n \to \infty$

$$\mathbb{E}[T_n^\beta] = 2^{-\beta}\Gamma(\beta+1)n^{-\beta} + O(n^{-\beta-1}).$$

(iv) As $n \to \infty$,

$$2nT_n \xrightarrow{d} \mathrm{Exp}(1),$$

where Exp(1) *is an exponential random variable with parameter 1.*

Proof. By conditioning on the number of U_j, $j \le n$ with $U_j \le U_n$, using Lemma 4.2, and by exchangeability of the spacings, we have that for $n \ge 1$, $T_n \stackrel{d}{=} \min\{S_1^n, S_2^n\} \stackrel{d}{=} S_1^n/2$, by (47). Then (i) follows by (40), and (ii) follows by (45). Part (iii) then follows from part (ii) by (56). For (iv), we have that, for $t \in [0,\infty)$, and n large enough so that $t/(2n) \le 1/2$,

$$\mathbb{P}[2nT_n > t] = \mathbb{P}[T_n > t/(2n)] = (1-(t/n))^n \to \mathrm{e}^{-t},$$

as $n \to \infty$, but $1 - \mathrm{e}^{-t}$, $t \ge 0$ is the distribution function of an exponential random variable with parameter 1. □

Proposition 4.1 *Recall that $\gamma \approx 0.57721566$ is Euler's constant, defined at (6). Suppose $\alpha > 0$. As $n \to \infty$, we have*

$$\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1})] = \frac{\Gamma(\alpha+1)}{1-\alpha}2^{-\alpha}n^{1-\alpha} + 1 - \frac{2^{-\alpha}}{1-\alpha} + O(n^{-\alpha}); \quad (0<\alpha<1) \tag{61}$$

$$\mathbb{E}[\mathcal{O}^{1,1}(\mathcal{U}_n^{0,1})] = \frac{1}{2}\log n + \frac{1}{2}(\gamma+1) + O(n^{-1}); \tag{62}$$

$$\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1})] = 1 + \frac{2^{-\alpha}}{\alpha-1} + O(n^{1-\alpha}) \quad (\alpha>1) \tag{63}$$

Proof. Counting the first edge from 1 to 0, we have

$$\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1})] = 1 + \sum_{i=1}^n \left(\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_i^{0,1})] - \mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_{i-1}^{0,1})]\right) = 1 + \sum_{i=1}^n \mathbb{E}[T_i^\alpha].$$

In the case where $\alpha = 1$, $\mathbb{E}[T_i] = (2(i+1))^{-1}$ by (60), and (62) follows by (6). For general $\alpha > 0$, $\alpha \ne 1$, from (59) we have that

$$\begin{aligned}\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1})] &= 1 + 2^{-\alpha}\Gamma(1+\alpha)\sum_{i=1}^n \frac{\Gamma(i+1)}{\Gamma(1+\alpha+i)} \\ &= 1 + \frac{2^{-\alpha}}{\alpha-1} - \frac{2^{-\alpha}\Gamma(1+\alpha)\Gamma(n+2)}{(\alpha-1)\Gamma(n+1+\alpha)},\end{aligned} \tag{64}$$

the final equality proved by induction on n. By Stirling's formula, the last term satisfies

$$-\frac{2^{-\alpha}\Gamma(1+\alpha)\Gamma(n+2)}{(\alpha-1)\Gamma(n+1+\alpha)} = -2^{-\alpha}\frac{\Gamma(1+\alpha)}{\alpha-1}n^{1-\alpha}(1+O(n^{-1})), \tag{65}$$

which tends to zero as $n \to \infty$ for $\alpha > 1$, to give us (63). For $\alpha < 1$, we have (61) from (64) and (65). □

Proposition 4.2 *Suppose $\alpha > 0$. As $n \to \infty$, we have*

$$\begin{aligned}
\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0)] &= \frac{\Gamma(\alpha+1)}{1-\alpha} 2^{-\alpha} n^{1-\alpha} + \frac{1}{\alpha} - \frac{2^{-\alpha}}{\alpha(1-\alpha)} + O(n^{-\alpha}); \quad (0 < \alpha < 1) \\
\mathbb{E}[\mathcal{O}^{1,1}(\mathcal{U}_n^0)] &= \frac{1}{2} \log n + \frac{1}{2}\gamma + O(n^{-1}); \\
\mathbb{E}[\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0)] &= \frac{1}{\alpha} + \frac{2^{-\alpha}}{\alpha(\alpha-1)} + O(n^{1-\alpha}) \quad (\alpha > 1)
\end{aligned} \tag{66}$$

Proof. This follows from Proposition 4.1 with (57). □

Lemma 4.5 *Suppose $\alpha > 1/2$. Then there exists $C \in (0, \infty)$ such that for all $n \in \mathbb{N}$,*

$$\mathrm{Var}[\tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n)] \le C, \quad \mathrm{Var}[\tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0)] \le C, \quad \mathrm{Var}[\tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1})] \le C.$$

Proof. The statement for $\mathcal{U}_n$ for $\alpha > 1$ follows from Theorem 2.1(ii), and for $\alpha > 1/2$ is proved in [27]. From (55) we have that

$$\mathbb{E}[(Z_n^\alpha - H_n^\alpha)^2] \le C\mathbb{E}[(S_n^{n-1})^{2\alpha+1}] \le Cn^{-1-2\alpha},$$

by (45) and (56). So by Cauchy-Schwarz we obtain

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{i=2}^n (Z_i^\alpha - H_i^\alpha)\right)^2\right] &\le \sum_{i=2}^n \sum_{j=2}^n (\mathbb{E}[(Z_i^\alpha - H_i^\alpha)^2])^{1/2} (\mathbb{E}[(Z_j^\alpha - H_j^\alpha)^2])^{1/2} \\
&\le C \sum_{i=2}^n \sum_{j=2}^n i^{-(1/2)-\alpha} j^{-(1/2)-\alpha},
\end{aligned}$$

which is uniformly bounded for $n \ge 2$ since $\alpha > 1/2$. Thus from (53) and Cauchy-Schwarz again we obtain

$$\mathbb{E}[(\mathcal{O}^{1,\alpha}(\mathcal{U}_n) - \mathcal{O}^{1,\alpha}(\mathcal{U}_n^0))^2] \le C < \infty,$$

for all $n \in \mathbb{N}$. The statement for $\mathcal{U}_n^0$ in the lemma then follows. The statement for $\mathcal{U}_n^{0,1}$ follows in a similar way, this time using (52) and (54). □

4.4 Limit theory when $1/2 < \alpha < 1$

Let U be uniform on $(0, 1)$, and given U, let $N(n) \sim \mathrm{Bin}(n-1, U)$. Set

$$D_\alpha(n) := (n-1)^{1/2} \left(U^\alpha \left(\frac{N(n)}{n-1} \right)^{1-\alpha} + (1-U)^\alpha \left(\frac{n-1-N(n)}{n-1} \right)^{1-\alpha} - 1 \right). \tag{67}$$

Lemma 4.6 *Suppose $0 \le \alpha \le 1$. Then, as $n \to \infty$,*

$$D_\alpha(n) \xrightarrow{L^3} 0. \tag{68}$$

We defer the proof of this lemma to the Appendix. Note that for what follows in this paper we will only use L^2 convergence in (68). However, the stronger L^3 version requires little extra work, and we anticipate using it in future work.

Proposition 4.3 *Suppose* $1/2 < \alpha < 1$. *Then as* $n \to \infty$,

$$\begin{pmatrix} \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0}) - \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n) - \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \tilde{J}_\alpha \\ \tilde{R} \\ \tilde{S} \end{pmatrix}, \tag{69}$$

where $(\tilde{J}_\alpha, \tilde{R}, \tilde{S})^\top$ *satisfies the fixed-point equation*

$$\begin{aligned} \begin{pmatrix} \tilde{J}_\alpha \\ \tilde{R} \\ \tilde{S} \end{pmatrix} \stackrel{d}{=} & \begin{pmatrix} U^\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U^\alpha & 0 \end{pmatrix} \begin{pmatrix} \tilde{J}_\alpha^{\{1\}} \\ \tilde{R}^{\{1\}} \\ \tilde{S}^{\{1\}} \end{pmatrix} + \begin{pmatrix} (1-U)^\alpha & 0 & 0 \\ 0 & (1-U)^\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{J}_\alpha^{\{2\}} \\ \tilde{R}^{\{2\}} \\ \tilde{S}^{\{2\}} \end{pmatrix} \\ & + \begin{pmatrix} \min\{U, 1-U\}^\alpha + \frac{2^{-\alpha}}{\alpha - 1}((1-U)^\alpha + U^\alpha - 1) \\ (U^\alpha - (1-U)^\alpha)\mathbf{1}_{\{U>1/2\}} + ((1-U)^\alpha - 1)\frac{1-2^{-\alpha}}{\alpha} \\ (U^\alpha - \frac{1}{1+\alpha})\frac{1-2^{-\alpha}-\alpha}{\alpha} \end{pmatrix}. \end{aligned} \tag{70}$$

In particular, $\tilde{J}_\alpha$ *satisfies the fixed-point equation (10). Also,* $\mathbb{E}[\tilde{J}_\alpha] = \mathbb{E}[\tilde{R}] = \mathbb{E}[\tilde{S}] = 0$.

Proof. We use Lemma 4.1 (from Theorem 4.1 of [12]). Recall the definition of Y_n at (28). Let

$$R_n := \mathcal{O}^{1,\alpha}(\mathcal{U}_n^0) - \mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1}) + 1, \quad S_n := \mathcal{O}^{1,\alpha}(\mathcal{U}_n) - \mathcal{O}^{1,\alpha}(\mathcal{U}_n^0). \tag{71}$$

Write $U = U_1$ for the position of the first arrival. Given U, let $N(n) \sim \mathrm{Bin}(n-1, U)$ be the number of points of $U_2, U_3, \ldots, U_n$ that arrive to the left of $U_1 = U$. Using the self-similarity of the ONG, we have that $X_n = (Y_n, R_n, S_n)^\top$ satisfies, for $\alpha > 0$, and $n \geq 1$

$$\begin{aligned} \begin{pmatrix} Y_n \\ R_n \\ S_n \end{pmatrix} \stackrel{d}{=} & \begin{pmatrix} U^\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U^\alpha & 0 \end{pmatrix} \begin{pmatrix} Y_{N(n)}^{\{1\}} \\ R_{N(n)}^{\{1\}} \\ S_{N(n)}^{\{1\}} \end{pmatrix} \\ & + \begin{pmatrix} (1-U)^\alpha & 0 & 0 \\ 0 & (1-U)^\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} Y_{n-1-N(n)}^{\{2\}} \\ R_{n-1-N(n)}^{\{2\}} \\ S_{n-1-N(n)}^{\{2\}} \end{pmatrix} + \begin{pmatrix} \min\{U, 1-U\}^\alpha \\ (U^\alpha - (1-U)^\alpha)\mathbf{1}_{\{U>1/2\}} \\ -U^\alpha \end{pmatrix}, \end{aligned} \tag{72}$$

where, given U and $N(n)$, $Y_{N(n)}^{\{1\}}$, $Y_{n-1-N(n)}^{\{2\}}$ are independent copies of $Y_{N(n)}$, $Y_{n-1-N(n)}$ respectively, and similarly for the Rs and Ss. This equation is of the form of (32) with $k = 2$. Suppose $1/2 < \alpha < 1$. We now renormalize (72) by taking

$$(\tilde{Y}_n, \tilde{R}_n, \tilde{S}_n)^\top := (Y_n - \mathbb{E}[Y_n], R_n - \mathbb{E}[R_n], S_n - \mathbb{E}[S_n])^\top, \tag{73}$$

so we are now in the situation of Lemma 4.1, with $\tilde{X}_n = (\tilde{Y}_n, \tilde{R}_n, \tilde{S}_n)^\top$. In particular,

$$\tilde{Y}_n = \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}), \quad \tilde{R}_n = \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0) - \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}), \quad \tilde{S}_n = \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n) - \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0). \tag{74}$$

Now using the notation of (34) and (35), and the expressions for the expectations at (61), (57) and (58), from (72) we obtain

$$\tilde{X}_n \stackrel{d}{=} \begin{pmatrix} U^\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U^\alpha & 0 \end{pmatrix} \tilde{X}_{N(n)}^{\{1\}} + \begin{pmatrix} (1-U)^\alpha & 0 & 0 \\ 0 & (1-U)^\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \tilde{X}_{n-1-N(n)}^{\{2\}} + \tilde{B}(n), \tag{75}$$

where

$$\tilde{B}(n) = \begin{pmatrix} \min\{U, 1-U\}^\alpha + C_\alpha (n-1)^{(1/2)-\alpha} D_\alpha(n) + \frac{2^{-\alpha}}{\alpha - 1}\left(U^\alpha + (1-U)^\alpha - 1\right) \\ (U^\alpha - (1-U)^\alpha)\mathbf{1}_{\{U>1/2\}} + \frac{1-2^{-\alpha}}{\alpha}((1-U)^\alpha - 1) \\ (U^\alpha - \frac{1}{1+\alpha})\frac{1-2^{-\alpha}-\alpha}{\alpha} \end{pmatrix} + \begin{pmatrix} U^\alpha h(N(n)) + (1-U)^\alpha h(n-1-N(n)) - h(n) \\ (1-U)^\alpha k(n-1-N(n)) - k(n) \\ U^\alpha k(N(n)) - \ell(n) \end{pmatrix},$$

where $D_\alpha(n)$ is as defined at (67), $h(n)$, $k(n)$, $\ell(n)$ are all $o(1)$ as $n \to \infty$ and C_α is a constant.

The fact that $\tilde{X}_n$ is square-integrable, as required for Lemma 4.1, follows from Lemma 4.5. In order to apply Lemma 4.1, we need to verify the conditions (36), (37) and (38). By Lemma 4.6, $(n-1)^{(1/2)-\alpha} D_\alpha(n)$ tends to zero in L^2 as $n \to \infty$, for $1/2 < \alpha < 1$. Thus, for condition (36), as $n \to \infty$,

$$\tilde{B}(n) \xrightarrow{L^2} \begin{pmatrix} \min\{U, 1-U\}^\alpha + \frac{2^{-\alpha}}{\alpha - 1}\left(U^\alpha + (1-U)^\alpha - 1\right) \\ (U^\alpha - (1-U)^\alpha)\mathbf{1}_{\{U>1/2\}} + \frac{1-2^{-\alpha}}{\alpha}((1-U)^\alpha - 1) \\ (U^\alpha - \frac{1}{1+\alpha})\frac{1-2^{-\alpha}-\alpha}{\alpha} \end{pmatrix}. \tag{76}$$

Also, for condition (37),

$$\mathbb{E}\left[\left\| \begin{pmatrix} U^\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U^\alpha & 0 \end{pmatrix} \right\|_{\text{op}}^2 + \left\| \begin{pmatrix} (1-U)^\alpha & 0 & 0 \\ 0 & (1-U)^\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\|_{\text{op}}^2 \right] = \frac{2}{2\alpha + 1} < 1, \tag{77}$$

for $\alpha > 1/2$. Finally, for condition (38), for $\alpha > 0$ and any $\ell \in \mathbb{N}$, as $n \to \infty$

$$\mathbb{E}\left[\mathbf{1}_{\{N(n) \le \ell\} \cup \{N(n) = n\}} U^{2\alpha}\right] \to 0; \quad \mathbb{E}\left[\mathbf{1}_{\{n-1-N(n) \le \ell\} \cup \{n-1-N(n) = n\}} (1-U)^{2\alpha}\right] \to 0, \tag{78}$$

since, for example for any $\ell \in \mathbb{N}$, as $n \to \infty$

$$\mathbb{E}\left[\mathbf{1}_{\{N(n) \le \ell\} \cup \{N(n) = n\}} U^{2\alpha}\right] \le \mathbb{P}(N(n) \le \ell) \to 0,$$

by standard properties of the binomial distribution. Now with $k = 2$, Lemma 4.1 applied to equation (75), with the conditions (77), (76) and (78), implies that $\tilde{X}_n = (\tilde{Y}_n, \tilde{R}_n, \tilde{S}_n)^\top$ converges in distribution to some $X = (\tilde{Y}, \tilde{R}, \tilde{S})^\top$, where $\mathbb{E}[\tilde{Y}] = \mathbb{E}[\tilde{R}] = \mathbb{E}[\tilde{S}] = 0$ and the distribution of $(\tilde{Y}, \tilde{R}, \tilde{S})^\top$ is characterized by the fixed-point equation

$$\begin{pmatrix} \tilde{Y} \\ \tilde{R} \\ \tilde{S} \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} U^\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U^\alpha & 0 \end{pmatrix} \begin{pmatrix} \tilde{Y}^{\{1\}} \\ \tilde{R}^{\{1\}} \\ \tilde{S}^{\{1\}} \end{pmatrix} + \begin{pmatrix} (1-U)^\alpha & 0 & 0 \\ 0 & (1-U)^\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{Y}^{\{2\}} \\ \tilde{R}^{\{2\}} \\ \tilde{S}^{\{2\}} \end{pmatrix} + \begin{pmatrix} \min\{U, 1-U\}^\alpha + ((1-U)^\alpha + U^\alpha - 1)\frac{2^{-\alpha}}{\alpha - 1} \\ (U^\alpha - (1-U)^\alpha)\mathbf{1}_{\{U>1/2\}} + ((1-U)^\alpha - 1)\frac{1-2^{-\alpha}}{\alpha} \\ \left(U^\alpha - \frac{1}{1+\alpha}\right)\frac{1-2^{-\alpha}-\alpha}{\alpha} \end{pmatrix}. \tag{79}$$

That is, $\tilde{Y}$ satisfies (10), so that $\tilde{Y}$ has the distribution of $\tilde{J}_\alpha$, and setting $\tilde{Y} = \tilde{J}_\alpha$ in (79) gives (70). Then (69) follows by (74). □

4.5 Limit theory when α=1

Proposition 4.4 below is our main convergence result when $\alpha = 1$. First, we need the following result, the proof of which we defer to the Appendix. For $x \geq 0$, set $\log^+ x := \max\{\log x, 0\}$.

Lemma 4.7 *Let U be uniform on $(0,1)$ and, given U, let $N(n) \sim \mathrm{Bin}(n-1, U)$. Then, as $n \to \infty$,*

$$U(\log^+ N(n) - \log n) \xrightarrow{L^2} U \log U; \tag{80}$$

$$(1-U)(\log^+(n-1-N(n)) - \log n) \xrightarrow{L^2} (1-U)\log(1-U). \tag{81}$$

Proposition 4.4 *As $n \to \infty$,*

$$\begin{pmatrix} \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0}) - \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n) - \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \tilde{J}_1 \\ \tilde{R} \\ \tilde{S} \end{pmatrix}, \tag{82}$$

where $(\tilde{J}_1, \tilde{R}, \tilde{S})^\top$ satisfies the fixed-point equation

$$\begin{aligned} \begin{pmatrix} \tilde{J}_1 \\ \tilde{R} \\ \tilde{S} \end{pmatrix} \stackrel{d}{=} & \begin{pmatrix} U & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U & 0 \end{pmatrix} \begin{pmatrix} \tilde{J}_1^{\{1\}} \\ \tilde{R}^{\{1\}} \\ \tilde{S}^{\{1\}} \end{pmatrix} + \begin{pmatrix} 1-U & 0 & 0 \\ 0 & 1-U & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{J}_1^{\{2\}} \\ \tilde{R}^{\{2\}} \\ \tilde{S}^{\{2\}} \end{pmatrix} \\ & + \begin{pmatrix} \frac{U}{2}\log U + \frac{1-U}{2}\log(1-U) + \min\{U, 1-U\} \\ (2U-1)\mathbf{1}_{\{U>1/2\}} - \frac{U}{2} \\ \frac{1}{4} - \frac{U}{2} \end{pmatrix}. \end{aligned} \tag{83}$$

In particular, $\tilde{J}_1$ satisfies the fixed-point equation (9). Also, $\mathbb{E}[\tilde{J}_1] = \mathbb{E}[\tilde{R}] = \mathbb{E}[\tilde{S}] = 0$, $\mathrm{Var}[\tilde{R}] = 1/16$, $\mathrm{Var}[\tilde{S}] = 1/24$, and

$$\mathrm{Var}[\tilde{J}_1] = \frac{1}{4}(1 + \log 2) - \frac{\pi^2}{24} \approx 0.012053, \tag{84}$$

and $\mathbb{E}[\tilde{J}_1^3] \approx -0.00005732546$.

Proof. We follow the proof of Proposition 4.3. Recall the definition of Y_n at (28). Again define R_n and S_n as at (71), this time with $\alpha = 1$. Now the $\alpha = 1$ case of (72) holds, and we are in the situation of (32) with $X_n = (Y_n, R_n, S_n)^\top$. We now renormalize (72), taking $\tilde{X}_n = X_n - M_n$ with the notation of (73) and (33). By (62) we have

$$\mathbb{E}[Y_n] = \mathbb{E}[\mathcal{O}^{1,1}(\mathcal{U}_n^{0,1})] - 1 = \frac{1}{2}\log n + \frac{1}{2}(\gamma - 1) + h(n),$$

where $h(n) = o(1)$, while by the $\alpha = 1$ case of (57) $\mathbb{E}[R_n] = (1/2) + k(n)$, where $k(n) = O(n^{-1})$, and by the $\alpha = 1$ case of (58) $\mathbb{E}[S_n] = -(1/4) + \ell(n)$, where $\ell(n) = O(n^{-1})$. Then the modified version of (72) is

$$\tilde{X}_n \stackrel{d}{=} \begin{pmatrix} U & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U & 0 \end{pmatrix} \tilde{X}_{N(n)}^{\{1\}} + \begin{pmatrix} 1-U & 0 & 0 \\ 0 & 1-U & 0 \\ 0 & 0 & 0 \end{pmatrix} \tilde{X}_{n-1-N(n)}^{\{2\}} + \tilde{B}(n), \tag{85}$$

where

$$\tilde{B}(n) = \begin{pmatrix} Uh(N(n)) + (1-U)h(n-1-N(n)) - h(n) \\ (1-U)k(n-1-N(n)) - k(n) \\ Uk(N(n)) - \ell(n) \end{pmatrix}$$
$$+ \begin{pmatrix} \min\{U, 1-U\} + \frac{U}{2}(\log^+ N(n) - \log n) + \frac{1-U}{2}(\log^+(n-1-N(n)) - \log n) \\ (2U-1)\mathbf{1}_{\{U>1/2\}} - \frac{U}{2} \\ \frac{1}{4} - \frac{U}{2} \end{pmatrix}.$$

By the SLLN and dominated convergence, the first vector on the right of the above equation tends to 0 in L^2. So by Lemma 4.7 we obtain

$$\tilde{B}(n) \xrightarrow{L^2} \begin{pmatrix} \frac{U}{2}\log U + \frac{1-U}{2}\log(1-U) + \min\{U, 1-U\} \\ (2U-1)\mathbf{1}_{\{U>1/2\}} - \frac{U}{2} \\ \frac{1}{4} - \frac{U}{2} \end{pmatrix}, \tag{86}$$

as $n \to \infty$. The conditions of Lemma 4.1 are satisfied, by Lemma 4.5 and equations (77), (78) and (86). Then Lemma 4.1 applied to equation (85) shows that $\tilde{X}_n = (\tilde{Y}_n, \tilde{R}_n, \tilde{S}_n)^\top$ converges in distribution to $X = (\tilde{Y}, \tilde{R}, \tilde{S})^\top$, where $\mathbb{E}[\tilde{Y}] = \mathbb{E}[\tilde{R}] = \mathbb{E}[\tilde{S}] = 0$ and the distribution of $(\tilde{Y}, \tilde{R}, \tilde{S})^\top$ is characterized by the fixed-point equation

$$\begin{pmatrix} \tilde{Y} \\ \tilde{R} \\ \tilde{S} \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} U & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U & 0 \end{pmatrix} \begin{pmatrix} \tilde{Y}^{\{1\}} \\ \tilde{R}^{\{1\}} \\ \tilde{S}^{\{1\}} \end{pmatrix} + \begin{pmatrix} 1-U & 0 & 0 \\ 0 & 1-U & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{Y}^{\{2\}} \\ \tilde{R}^{\{2\}} \\ \tilde{S}^{\{2\}} \end{pmatrix}$$
$$+ \begin{pmatrix} \frac{U}{2}\log U + \frac{1-U}{2}\log(1-U) + \min\{U, 1-U\} \\ (2U-1)\mathbf{1}_{\{U>1/2\}} - \frac{U}{2} \\ \frac{1}{4} - \frac{U}{2} \end{pmatrix}. \tag{87}$$

That is, $\tilde{Y}$ satisfies (9), so that $\tilde{Y}$ has the distribution of $\tilde{J}_1$, and setting $\tilde{Y} = \tilde{J}_1$ in (87) gives (83). By the $\alpha = 1$ case of (74) we then have (82).

It remains to prove the results for the higher moments of $\tilde{J}_1$. For the variance of $\tilde{J}_1$, squaring both sides of (9), taking expectations, and using independence and the fact that $\mathbb{E}[\tilde{J}_1] = 0$, we obtain

$$\begin{aligned} \mathbb{E}[\tilde{J}_1^2] &= \frac{2}{3}\mathbb{E}[\tilde{J}_1^2] + \mathbb{E}[\min\{U, 1-U\}^2] + \frac{1}{2}\mathbb{E}[U^2(\log U)^2] \\ &\quad + \frac{1}{2}\mathbb{E}[U(1-U)\log U \log(1-U)] + 2\mathbb{E}[U \log U \min\{U, 1-U\}]. \end{aligned}$$

The integrals required for the expectations are standard, and we find that $\mathbb{E}[\tilde{J}_1^2] = ((1 + \log 2)/4) - (\pi^2/24)$, which yields (84). Similarly, we obtain the third moment $\mathbb{E}[\tilde{J}_1^3] \approx -0.00005732546$ from (9), although in this case numerical methods are required for some of the integrals. □

4.6 Limit theory for $\alpha > 1$

Proposition 4.5 *Let $\alpha > 1$.*

(i) There exists a r.v. J_α such that as $n \to \infty$ $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^{0,1}) \to 1 + J_\alpha$ a.s. and in L^p, $p \in \mathbb{N}$. Also, J_α satisfies the fixed-point equality (13), and $\mathbb{E}[J_\alpha] = 2^{-\alpha}/(\alpha - 1)$.

(ii) There exists a r.v. H_α such that as $n \to \infty$ $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0) \to H_\alpha$ a.s. and in L^p, $p \in \mathbb{N}$. Also, H_α satisfies the fixed-point equality (14), and $\mathbb{E}[H_\alpha] = (1/\alpha) + 2^{-\alpha}/(\alpha(\alpha-1))$.

Proof. First we prove part (i). Let T_i be the length of the ith edge of the ONG on $\mathcal{U}_n^{0,1}$, as defined at (51). Let $J_\alpha := \sum_{i=1}^\infty T_i^\alpha$. The sum converges almost surely since it has non-negative terms and, by (63), for $\alpha > 1$ the partial sums have uniformly bounded expectation so that the infinite sum has finite expectation. By a similar argument to the proof of Theorem 2.1 (ii) in Section 3, the L^p convergence follows by Hölder's inequality and dominated convergence.

We now identify the limit. We have (31), this time for $\alpha > 1$. As $n \to \infty$, $N(n)$ and $n - N(n)$ both tend to infinity almost surely, and so, by taking $n \to \infty$ in (31), we obtain the fixed-point equation (13).

The identity $E[J_\alpha] = 2^{-\alpha}(\alpha-1)^{-1}$ is obtained either from (63), or by taking expectations in (13). Next, if we set $\tilde{J}_\alpha = J_\alpha - \mathbb{E}[J_\alpha]$, (13) yields (10).

We now prove part (ii). Following the above argument with the H_i replacing the T_i and using (66) in place of (63) gives that $\mathcal{O}^{1,\alpha}(\mathcal{U}_n^0)$ converges a.s. and in L^p, $p \in \mathbb{N}$, to some random variable. Once more, we need to identify the limit.

Consider the $\alpha > 1$ case of (30). As $n \to \infty$, $N(n)$ and $n - N(n)$ both tend to infinity almost surely, and so, by taking $n \to \infty$ in (30), and using the fact that $\mathcal{O}^{1,\alpha}(\mathcal{U}_{N(n)}^{0,1})$ converges almost surely to $1 + J_\alpha$ (by part (i)), and that $\mathcal{O}^{1,\alpha}(\mathcal{U}_{n-1-N(n)}^0)$ converges almost surely to H_α (by the argument above) we obtain the fixed-point equation (14).

The identity $E[H_\alpha] = \alpha^{-1} + 2^{-\alpha}\alpha^{-1}(\alpha-1)^{-1}$ is obtained either from (66), or by taking expectations in (14). Next, if we set $\tilde{H}_\alpha = H_\alpha - \mathbb{E}[H_\alpha]$, (14) yields (12). □

4.7 Proof of Theorems 4.1 and 2.2

Proof of Theorem 4.1. First we prove part (i) of the theorem. For $1/2 < \alpha < 1$ we have that

$$\begin{pmatrix} \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0) \\ \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0) - \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n) - \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \tilde{J}_\alpha \\ \tilde{R} \\ \tilde{S} \end{pmatrix}, \quad (88)$$

as $n \to \infty$, by Proposition 4.3. By (70), the final term in (88) is equal in distribution to

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} U^\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U^\alpha & 0 \end{pmatrix} \begin{pmatrix} \tilde{J}_\alpha^{\{1\}} \\ \tilde{R}^{\{1\}} \\ \tilde{S}^{\{1\}} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} (1-U)^\alpha & 0 & 0 \\ 0 & (1-U)^\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{J}_\alpha^{\{2\}} \\ \tilde{R}^{\{2\}} \\ \tilde{S}^{\{2\}} \end{pmatrix}$$
$$+ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \min\{U, 1-U\}^\alpha + ((1-U)^\alpha + U^\alpha - 1)\frac{2^{-\alpha}}{\alpha - 1} \\ (U^\alpha - (1-U)^\alpha)\mathbf{1}_{\{U > 1/2\}} + \frac{1-2^{-\alpha}}{\alpha}((1-U)^\alpha - 1) \\ (U^\alpha - \frac{1}{1+\alpha})\frac{1-2^{-\alpha}-\alpha}{\alpha} \end{pmatrix}.$$

Multiplying out and noting $(U^\alpha - (1-U)^\alpha)\mathbf{1}_{\{U>1/2\}} = U^\alpha - \min\{U, 1-U\}^\alpha$ we get

$$\begin{pmatrix} \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0) \\ \tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \tilde{J}_\alpha \\ \tilde{J}_\alpha + \tilde{R} \\ \tilde{J}_\alpha + \tilde{R} + \tilde{S} \end{pmatrix} \stackrel{d}{=} U^\alpha \begin{pmatrix} \tilde{J}_\alpha^{\{1\}} \\ \tilde{J}_\alpha^{\{1\}} \\ \tilde{J}_\alpha^{\{1\}} + \tilde{R}^{\{1\}} \end{pmatrix} + (1-U)^\alpha \begin{pmatrix} \tilde{J}_\alpha^{\{2\}} \\ \tilde{J}_\alpha^{\{2\}} + \tilde{R}^{\{2\}} \\ \tilde{J}_\alpha^{\{2\}} + \tilde{R}^{\{2\}} \end{pmatrix}$$

$$+\begin{pmatrix} \min\{U,1-U\}^\alpha + ((1-U)^\alpha + U^\alpha - 1)\frac{2^{-\alpha}}{\alpha-1} \\ U^\alpha(1+\frac{2^{-\alpha}}{\alpha-1}) + ((1-U)^\alpha - 1)(\frac{1}{\alpha} + \frac{2^{-\alpha}}{\alpha(\alpha-1)}) \\ (U^\alpha + (1-U)^\alpha - \frac{2}{1+\alpha})(\frac{1}{\alpha} - \frac{2^{-\alpha}}{\alpha(1-\alpha)}) \end{pmatrix}.$$

So setting $\tilde{H}_\alpha = \tilde{J}_\alpha + \tilde{R}$ and $\tilde{G}_\alpha = \tilde{J}_\alpha + \tilde{R} + \tilde{S}$, we have (26).

Now we prove part (ii) of the theorem. For $\alpha = 1$, as an analogue of (88),

$$\begin{pmatrix} \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0}) \\ \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0}) - \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n) - \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \tilde{J}_1 \\ \tilde{R} \\ \tilde{S} \end{pmatrix}, \quad (89)$$

as $n \to \infty$, by Proposition 4.4. By (83), the final term in (89) is equal in distribution to

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} U & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U & 0 \end{pmatrix} \begin{pmatrix} \tilde{J}_1^{\{1\}} \\ \tilde{R}^{\{1\}} \\ \tilde{S}^{\{1\}} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1-U & 0 & 0 \\ 0 & 1-U & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{J}_1^{\{2\}} \\ \tilde{R}^{\{2\}} \\ \tilde{S}^{\{2\}} \end{pmatrix}$$
$$+ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{U}{2}\log U + \frac{1-U}{2}\log(1-U) + \min\{U,1-U\} \\ (2U-1)\mathbf{1}_{\{U>1/2\}} - \frac{U}{2} \\ \frac{1}{4} - \frac{U}{2} \end{pmatrix}.$$

Multiplying out and using the fact that $(2U-1)\mathbf{1}_{\{U>1/2\}} = U - \min\{U, 1-U\}$ we have

$$\begin{pmatrix} \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0,1}) \\ \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n^{0}) \\ \tilde{\mathcal{O}}^{1,1}(\mathcal{U}_n) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \tilde{J}_1 \\ \tilde{J}_1 + \tilde{R} \\ \tilde{J}_1 + \tilde{R} + \tilde{S} \end{pmatrix} \stackrel{d}{=} U \begin{pmatrix} \tilde{J}_1^{\{1\}} \\ \tilde{J}_1^{\{1\}} \\ \tilde{J}_1^{\{1\}} + \tilde{R}^{\{1\}} \end{pmatrix} + (1-U) \begin{pmatrix} \tilde{J}_1^{\{2\}} \\ \tilde{J}_1^{\{2\}} + \tilde{R}^{\{2\}} \\ \tilde{J}_1^{\{2\}} + \tilde{R}^{\{2\}} \end{pmatrix}$$
$$+ \begin{pmatrix} (U/2)\log U + \frac{1-U}{2}\log(1-U) + \min\{U,1-U\} \\ (U/2)\log U + \frac{1-U}{2}\log(1-U) + \frac{U}{2} \\ (U/2)\log U + \frac{1-U}{2}\log(1-U) + \frac{1}{4} \end{pmatrix}.$$

So setting $\tilde{H}_1 = \tilde{J}_1 + \tilde{R}$ and $\tilde{G}_1 = \tilde{J}_1 + \tilde{R} + \tilde{S}$, we have (27). Proposition 4.4 gives $\mathbb{E}[\tilde{J}_1] = \mathbb{E}[\tilde{R}] = \mathbb{E}[\tilde{S}] = 0$, and so $\mathbb{E}[\tilde{H}_1] = \mathbb{E}[\tilde{G}_1] = 0$ also. Proposition 4.4 also gives $\mathrm{Var}[\tilde{J}_1]$. We obtain the higher moments of $\tilde{H}_1$ and $\tilde{G}_1$ from (11) and (16). The stated covariances follow from (83) and the moments given in Proposition 4.4.

Finally, part (iii) of the theorem is Proposition 4.5. □

Proof of Theorem 2.2. Parts (i) and (ii) of the theorem follow directly from the corresponding parts of Theorem 4.1. It remains to prove part (iii) of the theorem. Suppose $\alpha > 1$. Consider the $\alpha > 1$ case of (29). We use the fact that $N(n)$ and $n - N(n)$ tend to infinity almost surely, the independence given U and $N(n)$, and the convergence in L^p and almost surely of $\tilde{\mathcal{O}}^{1,\alpha}(\mathcal{U}_n^0)$ (for $\alpha > 1$) to obtain the result. □

5 Proof of Theorem 2.3

Proof of Theorem 2.3. We make use of the theory of Dirichlet spacings as discussed in Section 4.2. Since the nearest-neighbour (directed) graph joins each vertex (which sits

at the endpoint of each spacing apart from the points 0 and 1) to its nearest neighbour, we have, for $n \geq 3$

$$\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n) = (S_2^n)^\alpha + (S_n^n)^\alpha + \sum_{i=2}^{n-1} \left(\min\{S_i^n, S_{i+1}^n\}\right)^\alpha. \tag{90}$$

Now, from (90), using exchangeability we have that

$$\mathbb{E}[\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)] = 2\mathbb{E}[(S_1^n)^\alpha] + (n-2)\mathbb{E}[(\min\{S_1^n, S_2^n\})^\alpha],$$

where, from (47) and (45) we have

$$\mathbb{E}\left[(\min\{S_1^n, S_2^n\})^\alpha\right] \;=\; 2^{-\alpha}\mathbb{E}[(S_1^n)^\alpha] = 2^{-\alpha}\frac{\Gamma(\alpha+1)\Gamma(n+1)}{\Gamma(n+\alpha+1)}. \tag{91}$$

Then (21) follows. We now prove (22). Squaring both sides of (90) and taking expectations, we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)\right)^2\right] \\
=\;& \sum_{i=2}^{n-1}\mathbb{E}\left[\left(\min\{S_i^n, S_{i+1}^n\}\right)^{2\alpha}\right] + 2\sum_{i=3}^{n-1}\sum_{j=2}^{i-1}\mathbb{E}\left[\left(\min\{S_i^n, S_{i+1}^n\}\right)^\alpha\left(\min\{S_j^n, S_{j+1}^n\}\right)^\alpha\right] \\
&+\mathbb{E}[(S_2^n)^{2\alpha}] + \mathbb{E}[(S_n^n)^{2\alpha}] + 2\sum_{i=2}^{n-1}\mathbb{E}[(S_2^n)^\alpha(\min\{S_i^n, S_{i+1}^n\})^\alpha] \\
&+2\sum_{i=2}^{n-1}\mathbb{E}[(S_n^n)^\alpha(\min\{S_i^n, S_{i+1}^n\})^\alpha] + 2\mathbb{E}[(S_2^n)^\alpha(S_n^n)^\alpha].
\end{aligned}$$

Then, by exchangeability,

$$\begin{aligned}
\mathbb{E}\left[\left(\mathcal{L}_1^{1,\alpha}(\mathcal{U}_n)\right)^2\right] &= (n-2)\mathbb{E}\left[(\min\{S_1^n, S_2^n\})^{2\alpha}\right] + 2\mathbb{E}[(S_1^n S_2^n)^\alpha] \\
&+(n-3)(n-4)\mathbb{E}\left[(\min\{S_1^n, S_2^n\})^\alpha(\min\{S_3^n, S_4^n\})^\alpha\right] \\
&+2(n-3)\mathbb{E}\left[(\min\{S_1^n, S_2^n\})^\alpha(\min\{S_2^n, S_3^n\})^\alpha\right] + 2\mathbb{E}[(S_1^n)^{2\alpha}] \\
&+4(n-3)\mathbb{E}[(S_1^n)^\alpha(\min\{S_2^n, S_3^n\})^\alpha] + 4\mathbb{E}[(S_1^n)^\alpha(\min\{S_1^n, S_2^n\})^\alpha].
\end{aligned} \tag{92}$$

Now, by (46) and (48) we have

$$\mathbb{E}[(S_1^n)^\alpha(\min\{S_2^n, S_3^n\})^\alpha] = 2^{-\alpha}\frac{\Gamma(n+1)\Gamma(1+\alpha)^2}{\Gamma(n+1+2\alpha)},$$

and, using (46) this time with (49) we obtain

$$\mathbb{E}[(\min\{S_1^n, S_2^n\})^\alpha(\min\{S_3^n, S_4^n\})^\alpha] = 2^{-2\alpha}\frac{\Gamma(n+1)\Gamma(1+\alpha)^2}{\Gamma(n+1+2\alpha)}.$$

Also we have that

$$\begin{aligned}
\mathbb{E}[(S_1^n)^\alpha(\min\{S_1^n, S_2^n\})^\alpha] &= \mathbb{E}[(S_1^n)^{2\alpha}\mathbf{1}_{\{S_1^n < S_2^n\}}] + \mathbb{E}[(S_1^n)^\alpha(S_2^n)^\alpha\mathbf{1}_{\{S_1^n > S_2^n\}}] \\
&= \frac{1}{2}\mathbb{E}[(\min\{S_1^n, S_2^n\})^{2\alpha}] + \frac{1}{2}\mathbb{E}[(S_1^n)^\alpha(S_2^n)^\alpha].
\end{aligned}$$

Hence from (91) and (46) we obtain

$$\mathbb{E}[(S_1^n)^\alpha(\min\{S_1^n, S_2^n\})^\alpha] = \frac{1}{2}\left(2^{-2\alpha}\Gamma(1+2\alpha) + \Gamma(1+\alpha)^2\right)\frac{\Gamma(n+1)}{\Gamma(n+1+2\alpha)}.$$

The final term on the right-hand side of (92) that we need to evaluate is

$$\begin{aligned}\mathbb{E}[(\min\{S_1^n, S_2^n\})^\alpha(\min\{S_2^n, S_3^n\})^\alpha] &= \mathbb{E}[(S_2^n)^{2\alpha}\mathbf{1}_{\{S_2^n<S_1^n,\ S_2^n<S_3^n\}}] \\ &\quad +4\mathbb{E}[(S_1^n)^\alpha(S_2^n)^\alpha\mathbf{1}_{\{S_1^n<S_2^n<S_3^n\}}]. \qquad (93)\end{aligned}$$

For the first term on the right of (93), by (50) we have

$$\begin{aligned}\mathbb{E}[(S_2^n)^{2\alpha}\mathbf{1}_{\{S_2^n<S_1^n,\ S_2^n<S_3^n\}}] &= \frac{1}{3}\mathbb{E}[(\min\{S_1^n, S_2^n, S_3^n\})^{2\alpha}] \\ &= 3^{-1-2\alpha}\frac{\Gamma(1+2\alpha)\Gamma(n+1)}{\Gamma(n+1+2\alpha)}.\end{aligned}$$

Now consider the final term of (93). A direct computation using (42) gives

$$\begin{aligned}&\mathbb{E}[(S_1^n)^\alpha(S_2^n)^\alpha\mathbf{1}_{\{S_1^n<S_2^n<S_3^n\}}] \\ =\ & n(n-1)(n-2)\int_0^{1/3}\mathrm{d}y\int_y^{(1-y)/2}\mathrm{d}x\int_x^{1-x-y}x^\alpha y^\alpha(1-x-y-z)^{n-3}\mathrm{d}z \\ =\ & n(n-1)\int_0^{1/3}\mathrm{d}y\int_y^{(1-y)/2}x^\alpha y^\alpha(1-y-2x)^{n-2}\mathrm{d}x,\end{aligned}$$

which, via the change of variables $w = y + 2x$ and Fubini's theorem is the same as

$$n(n-1)2^{-\alpha-1}\int_0^1(1-w)^{n-2}\mathrm{d}w\int_0^{w/3}y^\alpha(w-y)^\alpha\mathrm{d}y.$$

Setting $t = 3y/w$ reduces this last expression to

$$n(n-1)6^{-\alpha-1}\int_0^1 w^{1+2\alpha}(1-w)^{n-2}\mathrm{d}w\int_0^1 t^\alpha(1-(t/3))^\alpha\mathrm{d}t.$$

Using (44) for the integral involving w, and the fact that (see, e.g., 15.3.1 in [1]) for $a > 0$,

$$\int_0^1 t^{a-1}(1-(t/z))^{-b}\mathrm{d}t = \frac{1}{a}\,{}_2F_1(b, a; a+1; z)$$

for the integral involving t, we obtain the expression for $J_{n,\alpha}$ as given by (19). Then, by (92) and the subsequent calculations, we obtain (22). Finally, (22) and (56) give (23). □

Appendix: technical lemmas

Proof of Lemma 4.6. The result is trivial when $\alpha = 1$ or $\alpha = 0$. Suppose $0 < \alpha < 1$. Suppose $n > 1$. To ease notation, for the duration of this proof, set $m = n - 1$. Then we have that for any $U \in (0, 1)$ and $0 \le N(n) \le m$,

$$-1 \le U^\alpha\left(\frac{N(n)}{m}\right)^{1-\alpha} + (1-U)^\alpha\left(\frac{m-N(n)}{m}\right)^{1-\alpha} - 1 \le 0, \qquad (94)$$

so that in particular $|D_\alpha(n)| \le n^{1/2}$ for $0 \le \alpha \le 1$. Let

$$W_n := \frac{N(n) - mU}{\sqrt{mU(1-U)}},$$

so that $\mathbb{E}[W_n] = 0$, $\mathbb{E}[W_n^2] = 1$, and

$$\frac{N(n)}{mU} = 1 + W_n\sqrt{\frac{1-U}{mU}}; \quad \frac{m - N(n)}{m(1-U)} = 1 - W_n\sqrt{\frac{U}{m(1-U)}}.$$

Then, by Taylor's theorem,

$$\begin{aligned} U^\alpha \left(\frac{N(n)}{m}\right)^{1-\alpha} &= U\left(1 + (1-\alpha)W_n\sqrt{\frac{1-U}{mU}} - R_1(n)W_n^2\frac{1-U}{mU}\right) \quad (95)\\ &= U\left(1 + R_2(n)W_n\sqrt{\frac{1-U}{mU}}\right), \quad (96)\end{aligned}$$

for remainder terms $R_1(n)$, $R_2(n)$ (which depend on W_n and U). Similarly, we have

$$\begin{aligned} &(1-U)^\alpha \left(\frac{m - N(n)}{m}\right)^{1-\alpha}\\ = \;& (1-U)\left(1 - (1-\alpha)W_n\sqrt{\frac{U}{m(1-U)}} - R_3(n)W_n^2\frac{U}{m(1-U)}\right) \quad (97)\\ = \;& (1-U)\left(1 - R_4(n)W_n\sqrt{\frac{U}{m(1-U)}}\right). \quad (98)\end{aligned}$$

By the Lagrange form of the remainder in Taylor's theorem and a continuity argument at $x = 0$ there exists a constant $B \in (0, \infty)$ such that for $\beta = 1 - \alpha$,

$$0 \ge \frac{(1+x)^\beta - 1 - \beta x}{x^2} \ge -B, \text{ and } 0 \le \frac{(1+x)^\beta - 1}{x} \le B,$$

for all $x \ge -1$. Thus we have, for $i \in \{1, 2, 3, 4\}$,

$$0 \le R_i(n) < C, \quad (99)$$

for a finite positive constant C. For $n > 1$, $m = n - 1$, let E_n denote the event $m^{-3/4} < U < 1 - m^{-3/4}$. From (95) and (97) we obtain

$$|D_\alpha(n)\mathbf{1}_{E_n}| = \left|-R_1(n)W_n^2(1-U)m^{-1/2} - R_3(n)W_n^2Um^{-1/2}\right|\mathbf{1}_{E_n} \le Cm^{-1/2}W_n^2\mathbf{1}_{E_n}, \quad (100)$$

for some $C \in (0, \infty)$. By a standard moment generating function calculation,

$$\begin{aligned} \mathbb{E}[(N(n) - mU)^6|U] &= mU(1-U)\left[15m^2U^2(1-U)^2 - 130mU^2(1-U)^2\right.\\ &\quad \left.+25mU(1-U) - 30U(1-U)(1-2U)^2 + 1\right]\\ &\le mU(1-U)(15m^2U^2(1-U)^2 + 25mU(1-U) + 1). \quad (101)\end{aligned}$$

By (101) we have that

$$\mathbb{E}[W_n^6 \mathbf{1}_{E_n}] \leq \mathbb{E}[(N(n) - mU)^6 m^{-3} U^{-3} (1-U)^{-3} | E_n]$$
$$\leq 15 + 25 m^{-1} \mathbb{E}[U^{-1}(1-U)^{-1} | E_n] + m^{-2} \mathbb{E}[U^{-2}(1-U)^{-2} | E_n] = O(1),$$

as $n \to \infty$, so from (100) we have that

$$D_\alpha(n) \mathbf{1}_{E_n} \xrightarrow{L^3} 0. \tag{102}$$

Also, from (96) and (98) we have,

$$\left| D_\alpha(n) \mathbf{1}_{E_n^c} \right| = \left| (R_2(n) - R_4(n)) W_n U^{1/2} (1-U)^{1/2} \right| \mathbf{1}_{E_n^c},$$

and so using (99) we have

$$\left| D_\alpha(n) \mathbf{1}_{E_n^c} \right| \leq C |W_n| U^{1/2} (1-U)^{1/2} \mathbf{1}_{E_n^c}. \tag{103}$$

Now, from (101) we have that

$$\mathbb{E}[(W_n U^{1/2} (1-U)^{1/2})^6] = m^{-3} \mathbb{E}[(N(n) - mU)^6] = O(1),$$

as $n \to \infty$, so by Cauchy-Schwarz we obtain from (103) that

$$\mathbb{E}[|D_\alpha(n) \mathbf{1}_{E_n^c}|^3] \leq (\mathbb{E}[(W_n U^{1/2} (1-U)^{1/2})^6])^{1/2} (\mathbb{P}[E_n^c])^{1/2} \to 0, \tag{104}$$

as $n \to \infty$ by the fact that $\mathbb{P}[E_n^c] = O(n^{-3/4})$. So (102) and (104) complete the proof. □

Proof of Lemma 4.7. For $n \in \mathbb{N}$, let $M_n := \log^+ N(n) - \log n - \log U$. First, suppose $N(n) \geq nU/2$. We have that

$$-\log 2 \leq M_n \mathbf{1}_{\{N(n) \geq nU/2\}} \mathbf{1}_{\{nU \geq 2\}} \leq -\log U.$$

Hence

$$U^2 M_n^2 \mathbf{1}_{\{N(n) \geq nU/2\}} \mathbf{1}_{\{nU \geq 2\}} \leq U^2 \max\{(\log 2)^2, (\log U)^2\}. \tag{105}$$

The expected value of the right-hand side of (105) is finite. Also, $U^2 M_n^2 \to 0$ almost surely as $n \to \infty$, by continuity and the strong law of large numbers for $N(n)$. Hence, by the dominated convergence theorem, as $n \to \infty$,

$$\mathbb{E}[U^2 M_n^2 \mathbf{1}_{\{N(n) \geq nU/2\}} \mathbf{1}_{\{nU \geq 2\}}] \to 0. \tag{106}$$

Also, we have $0 \leq \log^+ N(n) \leq \log n$, so that $-\log n \leq M_n \leq -\log U$. Hence

$$U^4 M_n^4 \leq (\log n)^4 + (\log U)^4, \tag{107}$$

so that $\mathbb{E}[U^4 M_n^4] = O((\log n)^4)$. Since $\mathbb{P}[nU < 2] = 2n^{-1}$, we then obtain, by Cauchy-Schwarz, that there exists a finite positive constant C such that

$$\mathbb{E}[U^2 M_n^2 \mathbf{1}_{\{N(n) \geq nU/2\}} \mathbf{1}_{\{nU < 2\}}] \leq C (\log n)^2 n^{-1/2} \to 0, \tag{108}$$

as $n \to \infty$. Now, suppose $0 \leq N(n) < nU/2$. In this case, from (107), and Cauchy-Schwarz again, for some finite positive constant C

$$\mathbb{E}[U^2 M_n^2 \mathbf{1}_{\{N(n)<nU/2\}}] \leq C(\log n)^2(\mathbb{P}[N(n) < nU/2])^{1/2} \to 0, \tag{109}$$

as $n \to \infty$, since

$$(\log n)^2(\mathbb{P}[N(n)< nU/2])^{1/2} \leq (\log n)^2(\mathbb{P}[U < n^{-1/2}] + \mathbb{P}[U > n^{-1/2}, N(n)< nU/2])^{1/2},$$

which tends to zero as $n \to \infty$, using standard bounds for the tail of a binomial distribution (see, e.g., Lemma 1.1 in [14]) for the final probability. The results (106), (108), and (109) then give (80). The argument for (81) is similar. □

Acknowledgements

Some of this work was done when AW was at the University of Durham, supported by an EPSRC doctoral training account, and subsequently at the University of Bath.

References

[1] Abramowitz, M. and Stegun, I.A. (Eds.) (1965) Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series, **55**. U.S. Government Printing Office, Washington D.C.

[2] Aldous, D.J. and Bandyopadhyay, A. (2005) A survey of max-type recursive distributional equations, *Ann. Appl. Probab.*, **15**, 1047–1110.

[3] Berger, N., Bollobás, B., Borgs, C., Chayes, J., and Riordan, O. (2003) Degree distribution of the FKP model, In *Automata, Languages and Programming*, Lecture Notes in Computer Science, 2719, Springer, Heidelberg, pp. 725–738.

[4] Bertoin, J. and Gnedin, A. (2004) Asymptotic laws for nonconservative selfsimilar fragmentations, *Electr. J. Probab.*, **9**, 575–593.

[5] Billingsley, P. (1999) Convergence of Probability Measures, 2nd edn., Wiley, New York.

[6] Bollobás, B. and Riordan, O.M. (2003) Mathematical results on scale-free random graphs. In *Handbook of Graphs and Networks*, Wiley-VCH, Weinheim, pp. 1–34.

[7] Darling, D.A. (1953), On a class of problems related to the random division of an interval, *Ann. Math. Stats*, **24**, 239–253.

[8] Dorogovstev, S.N. and Medes, J.F.F. (2002), Evolution of networks, *Adv. Phys.*, **51**, 1079–1187.

[9] Fabrikant, A., Koutsoupias, E. and Papadimitriou, C.H. (2002) Heuristically optimized trade-offs: a new paradigm for power laws in the internet. In *Automata, Languages and Programming*, Lecture Notes in Comput. Sci., 2380, Springer, Berlin, pp. 110–122.

[10] Huang, K. (1987) Statistical Mechanics, 2nd edn., Wiley, New York.

[11] Kesten, H. and Lee, S. (1996) The central limit theorem for weighted minimal spanning trees on random points, *Ann. Appl. Probab.*, **6**, 495–527.

[12] Neininger, R. and Rüschendorf, L. (2004) A general limit theorem for recursive algorithms and combinatorial structures, *Ann. Appl. Probab.*, **14**, 378–418.

[13] Newman, M.E.J. (2003) The structure and function of complex networks, *SIAM Rev.*, **45**, 167–256.

[14] Penrose, M. (2003) Random Geometric Graphs, *Oxford Studies in Probability*, **6**, Clarendon Press, Oxford.

[15] Penrose, M.D. (2005) Multivariate spatial central limit theorems with applications to percolation and spatial graphs, *Ann. Probab.*, **33**, 1945–1991.

[16] Penrose, M.D. and Wade, A.R. (2006) On the total length of the random minimal directed spanning tree, *Adv. Appl. Probab.*, **38**, 336–372.

[17] Penrose, M.D. and Yukich, J.E. (2001) Central limit theorems for some graphs in computational geometry, *Ann. Appl. Probab.*, **11**, 1005–1041.

[18] Penrose, M.D. and Yukich, J.E. (2003) Weak laws of large numbers in geometric probability, *Ann. Appl. Probab.*, **13**, 277–303.

[19] Penrose, M.D. and Yukich, J.E. (2005) Normal approximation in geometric probability. In *Stein's Method and Applications*, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, 5, Singapore Univ. Press, Singapore, pp. 37–58.

[20] Pyke, R. (1965) Spacings, *J. Royal. Stat. Soc. Ser. B*, **27**, 395–449.

[21] Rachev, S.T. (1991) Probability Metrics and the Stability of Stochastic Models, Wiley.

[22] Rösler, U. (1992) A fixed point theorem for distributions, *Stochastic Process. Appl.* **42**, 195–214.

[23] Rösler, U. and Rüschendorf, L. (2001) The contraction method for recursive algorithms, *Algorithmica*, **29**, 3–33.

[24] Steele, J.M. (1997) Probability Theory and Combinatorial Optimization, Society for Industrial and Applied Mathematics, Philadelphia.

[25] Wade, A.R. (2005) Limiting behaviour of random spatial graphs and asymptotically homogeneous RWRE, Ph.D. thesis, University of Durham.

[26] Wade, A.R. (2006) Explicit laws of large numbers for random nearest-neighbour type graphs. Preprint available from `http://arxiv.org/abs/math.PR/0603559`.

[27] Wade, A.R. (2006) Central limit theorem for the on-line nearest-neighbour graph. Preprint.

[28] Yukich, J.E. (1998) Probability Theory of Classical Euclidean Optimization Problems, *Lecture Notes in Mathematics*, **1675**, Springer, Berlin.