# ON THE TOTAL LENGTH OF THE RANDOM MINIMAL DIRECTED SPANNING TREE

MATHEW D. PENROSE * ** AND
ANDREW R. WADE,* *** *University of Bath*

## Abstract

In Bhatt and Roy's minimal directed spanning tree construction for a random, partially ordered set of points in the unit square, all edges must respect the 'coordinatewise' partial order and there must be a directed path from each vertex to a minimal element. We study the asymptotic behaviour of the total length of this graph with power-weighted edges. The limiting distribution is given by the sum of a normal component away from the boundary plus a contribution introduced by the boundary effects, which can be characterized by a fixed-point equation, and is reminiscent of limits arising in the probabilistic analysis of certain algorithms. As the exponent of the power weighting increases, the distribution undergoes a phase transition from the normal contribution being dominant to the boundary effects being dominant. In the critical case in which the weight is simple Euclidean length, both effects contribute significantly to the limit law.

*Keywords:* Spanning tree; nearest-neighbour graph; weak convergence; fixed-point equation; phase transition; fragmentation process

2000 Mathematics Subject Classification: Primary 60D05; 60F05
Secondary 05C80; 05C05

## 1. Introduction

Recent interest in graphs, generated over random point sets consisting of independent uniform points in the unit square by connecting nearby points according to some deterministic rule, has been considerable. Such graphs include the geometric graph, the nearest-neighbour graph and the minimal-length spanning tree. Many aspects of the large-sample asymptotic theory for such graphs, when they are locally determined in a certain sense, are by now quite well understood. See, for example, [8], [11], [16], [17], [21], [22], and [23].

One such graph is the *minimal directed spanning tree* (MDST), which was introduced by Bhatt and Roy in [5]. In the MDST, each point $x$ of a finite (random) subset $\mathscr{S}$ of $(0, 1]^2$ is connected by a directed edge to the nearest point $y \in \mathscr{S} \cup \{(0, 0)\}$ such that $y \neq x$ and $y \preccurlyeq^* x$, where $y \preccurlyeq^* x$ means that each component of $x - y$ is nonnegative. See Figure 1 for a realization of the MDST on simulated random points.

Motivation comes from the modelling of communications or drainage networks (see [5], [14], and [18]). For example, consider the problem of designing a set of canals to connect a set of hubs so as to minimize their total length subject to the constraint that all canals must flow downhill. The mathematical formulation given above for this constraint can lead to significant boundary
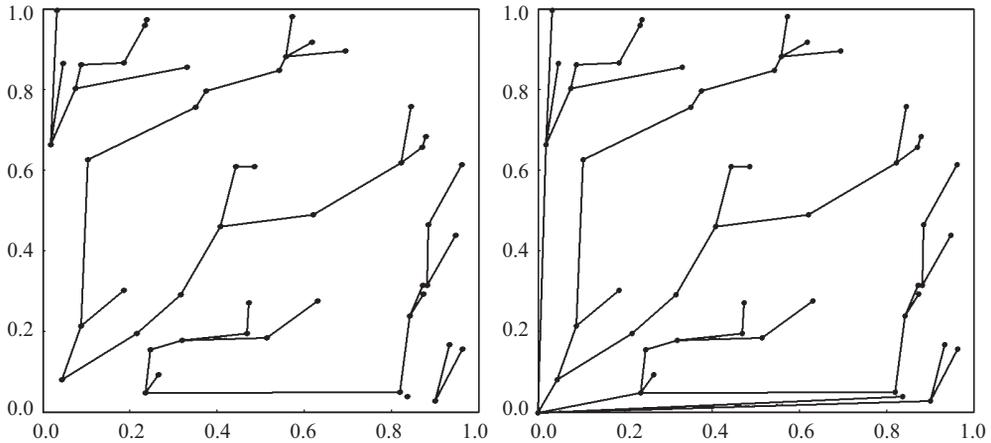
FIGURE 1: Realizations of the MDSF (*left*) and MDST (*right*) on 50 simulated uniform random points in the unit square, under the partial ordering '$\preccurlyeq^*$'.

effects due to the possibility of long edges occurring near the lower and left-hand boundaries of the unit square; these boundary effects distinguish the MDST qualitatively from the standard minimal spanning tree and the nearest-neighbour graph for point sets in the plane. Another difference is the fact that there is no uniform upper bound on vertex degrees in the MDST.

In the present work, we consider the total length, with power-weighted edges, of the MDST on random points in $(0, 1]^2$, as the number of points becomes large. We also consider the total power-weighted length of the *minimal directed spanning forest* (MDSF), which is the MDST with edges incident at the origin removed (see Figure 1 for an example). In [5], Bhatt and Roy mentioned that the total length is an object of considerable interest, although they restricted their analysis to the length of the edges joined to the origin (subsequently also examined in [14]). A first-order result for the total power-weighted length of the MDST or MDSF is a law of large numbers; this was given in [13] for a family of MDSFs indexed by partial orderings on $\mathbb{R}^2$, which include '$\preccurlyeq^*$' as a special case.

In this paper we are mainly concerned with establishing second-order results, i.e. weak convergence results for the distribution of the total power-weighted length, suitably centred and scaled, when the partial order is '$\preccurlyeq^*$'. For the length of edges from points in the region away from the boundary, we prove a central limit theorem. The boundary effects are significant, and near the boundary the MDST can be described in terms of a one-dimensional, on-line (i.e. sequentially generated) version of the MDST which we call the directed linear tree (DLT), and which we examine in Section 3. In the DLT, each point in a sequence of independent uniform random points in an interval is joined to its nearest neighbour to the left, amongst those points arriving earlier in the sequence. This DLT is of separate interest in relation to, for example, network modelling and molecular fragmentation (see [4], [3], and references therein).

In Theorem 3.1 we establish that the limiting distribution of the centred total length of the DLT is characterized by a distributional fixed-point equation which resembles those encountered in the probabilistic analysis of algorithms such as Quicksort [6]. Such fixed-point distributional equalities, and the so-called 'divide and conquer' or recursive algorithms from which they arise, have received considerable attention recently; see, for example, [7], [10], [19], and [20].

Our weak convergence results (Theorem 2.1) demonstrate that, depending on the value chosen for the weight exponent of the edges, there are two regimes – one in which the boundary

effects dominate and one in which those edges away from the boundary are dominant – and that there is a critical value (when we take simple Euclidean length as the weight) for which neither effect dominates.

In the related paper [14], we give results dealing with the weight of the edges joined to the origin, including weak convergence results in which the limiting distributions are given in terms of some generalized Dickman distributions. Subsequently, it has been shown [1] that this two-dimensional case is rather special: in higher dimensions the corresponding limits are normally distributed. Reference [14] also deals with the maximum edge length of the MDST (the maximum length of those edges incident at the origin was dealt with in [5]).

In the next section we give formal definitions of the MDST and MDSF and state our main results (Theorem 2.1) on the total lengths of the MDST and MDSF. The results on the DLT which we present in Section 3 and the general central limit theorems which we present in Section 4 are of some independent interest.

## 2. Definitions and main results

We work in the same framework as [14]. Here we briefly recall the relevant terminology. See [14] for more details.

Suppose that $V$ is a finite set endowed with a partial ordering '$\preccurlyeq$'. A *minimal element*, or *sink*, of $V$ is a vertex $v_0 \in V$ for which there exists no $v \in V \setminus \{v_0\}$ such that $v \preccurlyeq v_0$. Let $V_0$ denote the set of all sinks of $V$.

The partial ordering induces a directed graph $G = (V, E)$, with vertex set $V$ and with edge set $E$ consisting of all ordered pairs, $(v, u)$, of distinct elements of $V$ such that $u \preccurlyeq v$. A *directed spanning forest (DSF)* on $V$ is a subgraph $T = (V_T, E_T)$ of $(V, E)$ such that (i) $V_T = V$ and $E_T \subseteq E$ and, (ii) for each vertex $v \in V \setminus V_0$, there exists a unique directed path in $T$ that starts at $v$ and ends at some sink $u \in V_0$. In the case in which $V_0$ consists of a single sink, we refer to any DSF on $V$ as a *directed spanning tree (DST)* on $V$. If we ignore the orientation of edges then (see [14]) a DSF on $V$ is indeed a forest and if there is just one sink then any DST on $V$ is a tree.

Suppose that the directed graph $(V, E)$ carries a *weight function* $w: E \to [0, \infty)$ on its edges. If $T$ is a DSF on $V$ then we set $w(T) := \sum_{e \in E_T} w(e)$. A *minimal directed spanning forest (MDSF)* on $V$ is a directed spanning forest $T$ on $V$ such that $w(T) \leq w(T')$ for every DSF $T'$ on $V$. If $V$ has a single sink then a minimal directed spanning forest on $V$ is called a *minimal directed spanning tree (MDST)* on $V$.

For $v \in V$, we say that $u \in V \setminus \{v\}$ is a *directed nearest neighbour* of $v$ if $u \preccurlyeq v$ and $w(v, u) \leq w(v, u')$ for all $u' \in V \setminus \{v\}$ such that $u' \preccurlyeq v$. For each $v \in V \setminus V_0$, let $n_v$ denote a directed nearest neighbour of $v$ (chosen arbitrarily if $v$ has more than one directed nearest neighbour). Then (see [14]) the subgraph $(V, E_M)$ of $(V, E)$, obtained by taking $E_M := \{(v, n_v): v \in V \setminus V_0\}$, is an MDSF of $V$. Thus, if all edge weights are distinct, the MDSF is unique and is obtained by connecting each nonminimal vertex to its directed nearest neighbour.

The partial order that we consider here is the same as in [5], and we denote it by '$\preccurlyeq^*$'. In this case, $u \preccurlyeq^* v$ for $u = (u_1, u_2)$ and $v = (v_1, v_2)$ if and only if $u_1 \leq v_1$ and $u_2 \leq v_2$. This is sometimes called the 'coordinatewise' partial order. The symbol '$\preccurlyeq$' will denote a general partial order on $\mathbb{R}^2$.

The weight function is given by power-weighted Euclidean distance; i.e. to the edge $(u, v) \in E$ we assign weight $w(u, v) = \|u - v\|^\alpha$, where $\| \cdot \|$ denotes the Euclidean norm on $\mathbb{R}^2$ and $\alpha > 0$ is an arbitrary fixed parameter. Thus, when $\alpha = 1$ the weight of an edge is

simply its Euclidean length. Moreover, we shall assume that $V \subset \mathbb{R}^2$ is given by $V = \mathcal{S}$ or $V = \mathcal{S}^0 := \mathcal{S} \cup \{\mathbf{0}\}$, where $\mathbf{0}$ is the origin of $\mathbb{R}^2$ and $\mathcal{S}$ is generated in a *random* manner. The random point set $\mathcal{S}$ will usually be either the set of points given by a homogeneous Poisson point process, $\mathcal{P}_n$, of intensity $n$ on the unit square $(0, 1]^2$, or a binomial point process, $\mathcal{X}_n$, consisting of $n$ independent, uniformly distributed points on $(0, 1]^2$.

Note that in this random setting, each point of $\mathcal{S}$ almost surely has a unique directed nearest neighbour, meaning that $V$ has a unique MDSF, which does not depend on the choice of $\alpha$. Denote by $\mathcal{L}^\alpha(\mathcal{S})$ the total weight of all the edges in the MDSF on $\mathcal{S}$, and let $\tilde{\mathcal{L}}^\alpha(\mathcal{S}) := \mathcal{L}^\alpha(\mathcal{S}) - \mathrm{E}[\mathcal{L}^\alpha(\mathcal{S})]$, the centred total weight.

Our main result (Theorem 2.1) presents convergence in distribution for the partial order '$\preccurlyeq^*$'; the limiting distributions are of different types in the three cases $\alpha = 1$ (the same situation as [5]), $0 < \alpha < 1$, and $\alpha > 1$. We define these limiting distributions in Theorem 2.1, in terms of distributional fixed-point equations. These fixed-point equations are of the form

$$X \overset{\mathrm{D}}{=} \sum_{r=1}^{k} A_r X^{\{r\}} + B, \tag{2.1}$$

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

$$\mathrm{E}\left[ \sum_{r=1}^{k} |A_r|^2 \right] < 1, \qquad \mathrm{E}[B] = 0, \qquad \mathrm{E}[B^2] < \infty. \tag{2.2}$$

(By '$\overset{\mathrm{D}}{=}$' we denote equality in distribution.) Theorem 3 of Rösler [19] (proved using the contraction mapping theorem; see also [10] and [20]) says that if (2.2) holds, there is a unique square-integrable distribution with mean 0 satisfying the fixed-point equation (2.1); this will guarantee uniqueness of solutions to all the distributional fixed-point equalities considered in the sequel.

We define our random variables of interest as (unique) solutions to distributional fixed-point equations. In each of these equations, $U$ denotes a uniform random variable, and all the different random variables on the right-hand sides are independent.

Define $\tilde{D}_1$ by

$$\tilde{D}_1 \overset{\mathrm{D}}{=} U \tilde{D}_1^{\{1\}} + (1 - U) \tilde{D}_1^{\{2\}} + U \log U + (1 - U) \log(1 - U) + U. \tag{2.3}$$

In Section 3.4, we give a plot (Figure 2) of the probability density function of this distribution, estimated by simulation. We shall see later (in Propositions 3.5 and 3.6) that $\mathrm{E}[\tilde{D}_1] = 0$ and $\mathrm{var}[\tilde{D}_1] = 2 - \pi^2/6$; higher-order moments may be obtained recursively from (2.3). For example, $\mathrm{E}[\tilde{D}_1^3] \approx 0.154\,11$, which shows that $\tilde{D}_1$ is not Gaussian and is consistent with the skewness of the plot in Figure 2.

For $\alpha > 1$, define $\tilde{D}_\alpha$ by

$$\tilde{D}_\alpha \overset{\mathrm{D}}{=} U^\alpha \tilde{D}_\alpha^{\{1\}} + (1 - U)^\alpha \tilde{D}_\alpha^{\{2\}} + \frac{\alpha}{\alpha - 1} U^\alpha + \frac{1}{\alpha - 1} (1 - U)^\alpha - \frac{1}{\alpha - 1}. \tag{2.4}$$

Also for $\alpha > 1$, let $\tilde{F}_\alpha$ be defined by

$$\tilde{F}_\alpha \overset{\mathrm{D}}{=} U^\alpha \tilde{F}_\alpha + (1 - U)^\alpha \tilde{D}_\alpha + \frac{U^\alpha}{\alpha(\alpha - 1)} + \frac{(1 - U)^\alpha}{\alpha - 1} - \frac{1}{\alpha(\alpha - 1)}, \tag{2.5}$$

where $\tilde{D}_\alpha$ has the distribution given by (2.4). In Section 3 we shall see that for $\alpha > 1$ the random variables $\tilde{D}_\alpha$ and $\tilde{F}_\alpha$ arise as centred versions of random variables (denoted by $D_\alpha$ and $F_\alpha$, respectively) satisfying somewhat simpler fixed-point equations. Thus, $\tilde{D}_\alpha$ and $\tilde{F}_\alpha$ both have mean 0; their variances are given by (3.22) and (3.24), below.

Let $\mathcal{N}(0, s^2)$ denote the normal distribution with mean 0 and variance $s^2$, and let '$\xrightarrow{\mathrm{D}}$' denote convergence in distribution.

**Theorem 2.1.** *For $\alpha > 0$ and partial order '$\preccurlyeq^*$', there exist constants $t_\alpha$ and $s_\alpha$, $0 < t_\alpha^2 \le s_\alpha^2$, such that, for normal random variables $Y_\alpha \sim \mathcal{N}(0, s_\alpha^2)$ and $W_\alpha \sim \mathcal{N}(0, t_\alpha^2)$, the following statements hold.*

(i) *As $n \to \infty$,*

$$n^{(\alpha-1)/2} \tilde{\mathcal{L}}^\alpha(\mathcal{P}_n^0) \xrightarrow{\mathrm{D}} Y_\alpha \quad and \quad n^{(\alpha-1)/2} \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n^0) \xrightarrow{\mathrm{D}} W_\alpha, \qquad 0 < \alpha < 1, \ (2.6)$$

$$\tilde{\mathcal{L}}^1(\mathcal{P}_n^0) \xrightarrow{\mathrm{D}} \tilde{D}_1^{\{1\}} + \tilde{D}_1^{\{2\}} + Y_1 \quad and \quad \tilde{\mathcal{L}}^1(\mathcal{X}_n^0) \xrightarrow{\mathrm{D}} \tilde{D}_1^{\{1\}} + \tilde{D}_1^{\{2\}} + W_1, \qquad (2.7)$$

$$\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n^0) \xrightarrow{\mathrm{D}} \tilde{D}_\alpha^{\{1\}} + \tilde{D}_\alpha^{\{2\}} \quad and \quad \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n^0) \xrightarrow{\mathrm{D}} \tilde{D}_\alpha^{\{1\}} + \tilde{D}_\alpha^{\{2\}}, \qquad \alpha > 1. \quad (2.8)$$

*Here all the random variables in the limits are independent, and $\tilde{D}_\alpha^{\{i\}}$, $i = 1, 2$, are independent copies of $\tilde{D}_\alpha$ as defined in (2.3) for $\alpha = 1$ and (2.4) for $\alpha > 1$.*

(ii) *As $n \to \infty$,*

$$n^{(\alpha-1)/2} \tilde{\mathcal{L}}^\alpha(\mathcal{P}_n) \xrightarrow{\mathrm{D}} Y_\alpha \quad and \quad n^{(\alpha-1)/2} \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n) \xrightarrow{\mathrm{D}} W_\alpha, \qquad 0 < \alpha < 1, \ (2.9)$$

$$\tilde{\mathcal{L}}^1(\mathcal{P}_n) \xrightarrow{\mathrm{D}} \tilde{D}_1^{\{1\}} + \tilde{D}_1^{\{2\}} + Y_1 \quad and \quad \tilde{\mathcal{L}}^1(\mathcal{X}_n) \xrightarrow{\mathrm{D}} \tilde{D}_1^{\{1\}} + \tilde{D}_1^{\{2\}} + W_1, \qquad (2.10)$$

$$\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n) \xrightarrow{\mathrm{D}} \tilde{F}_\alpha^{\{1\}} + \tilde{F}_\alpha^{\{2\}} \quad and \quad \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n) \xrightarrow{\mathrm{D}} \tilde{F}_\alpha^{\{1\}} + \tilde{F}_\alpha^{\{2\}}, \qquad \alpha > 1. \quad (2.11)$$

*Here all the random variables in the limits are independent, $\tilde{D}_1^{\{i\}}$, $i = 1, 2$, are independent copies of $\tilde{D}_1$ as defined in (2.3), and, for $\alpha > 1$, $\tilde{F}_\alpha^{\{i\}}$, $i = 1, 2$, are independent copies of $\tilde{F}_\alpha$ as defined in (2.5).*

**Remarks 2.1.** The normal random variables $Y_\alpha$ and $W_\alpha$ arise from the edges away from the boundary (see Section 5). The nonnormal variables (the $\tilde{D}_\alpha$s and $\tilde{F}_\alpha$s) arise from the edges very close to the boundary, where the MDSF is asymptotically close to the 'directed linear forest' discussed below in Section 3.

Theorem 2.1 indicates a phase transition in the character of the limit law as $\alpha$ increases. The normal contribution (from the points away from the boundary) dominates for $0 < \alpha < 1$, while the boundary contributions dominate for $\alpha > 1$. In the critical case, $\alpha = 1$, neither effect dominates and both terms contribute significantly to the asymptotic behaviour.

Noteworthy in the case $\alpha = 1$ is the fact that, by (2.7) and (2.10), the limiting distribution is the same for $\tilde{\mathcal{L}}^1(\mathcal{P}_n)$ as for $\tilde{\mathcal{L}}^1(\mathcal{P}_n^0)$, and the same for $\tilde{\mathcal{L}}^1(\mathcal{X}_n)$ as for $\tilde{\mathcal{L}}^1(\mathcal{X}_n^0)$. Note, however, that the difference $\tilde{\mathcal{L}}^1(\mathcal{P}_n) - \tilde{\mathcal{L}}^1(\mathcal{P}_n^0)$ is the (centred) total length of edges incident at the origin, which is not negligible, but itself converges in distribution (see [14]) to a nondegenerate random variable, namely a centred generalized Dickman random variable with parameter 2 (see (3.14), below). As an extension of Theorem 2.1, it should be possible to show that the joint distribution of $(\tilde{\mathcal{L}}^1(\mathcal{P}_n), \tilde{\mathcal{L}}^1(\mathcal{P}_n^0))$ converges to that of two coupled random variables, both having the distribution of $\tilde{D}_1$, whose difference has the centred generalized Dickman distribution with parameter 2; likewise for the joint distribution of $(\tilde{\mathcal{L}}^1(\mathcal{X}_n), \tilde{\mathcal{L}}^1(\mathcal{X}_n^0))$.

The remainder of the paper is organized as follows. After a discussion of the DLT in Section 3, in Section 4 we present general limit theorems in geometric probability, which we shall use in obtaining our main results for the MDST. The proof of Theorem 2.1 is prepared in Sections 5 and 6 and completed in Section 7. In these proofs, we repeatedly use *Slutsky's theorem* (see, for example, [11]) which says that if $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} 0$ (i.e. $Y_n$ converges to 0 in probability), then $X_n + Y_n \xrightarrow{D} X$. For brevity, we omit details of some proofs. More details can be found in the longer version of this paper, which is available electronically [13].

## 3. The directed linear forest and tree

The directed linear forest and directed linear tree are for us tools for the analysis of the limiting behaviour of the contribution to the total weight of the random MDSF or MDST from edges near the boundary of the unit square. In this section we derive the properties of the DLF that we need (in particular Theorem 3.1); subsequently, in Theorem 6.1, we shall see that the total weight of edges from the points near the boundaries converges in distribution to the limit of the total weight of the DLF, as $n \to \infty$.

The DLT is also of some intrinsic interest. It is a one-dimensional directed analogue of the so-called 'on-line nearest-neighbour graph', which is of interest in the study of networks such as the World Wide Web (see, e.g. [4]; also see [12] and [15] for more on the on-line nearest-neighbour graph). Moreover, the DLT is constructed via a fragmentation process similar to those seen in, for example, [3]; the tree provides a historical representation of the fragmentation process.

For any finite sequence $\mathcal{T}_m = (x_1, x_2, \ldots, x_m) \in (0, 1]^m$, we construct the directed linear forest as follows. We insert the points $x_i$ in order, one at a time, starting with $i = 1$. At the insertion of each point, we join the new point to its nearest neighbour among those points already present that lie to the *left* of the point (provided that such a point exists). In other words, for each point $x_i$, $i \geq 2$, we join $x_i$ by a directed edge to the point $\max\{x_j : 1 \leq j < i, \ x_j < x_i\}$. If $\{x_j : 1 \leq j < i, \ x_j < x_i\}$ is empty, we do not add any directed edge from $x_i$. In this way we construct the so-called directed linear forest, which we denote by $\mathrm{DLF}(\mathcal{T}_m)$. We denote the total weight (under weight function with exponent $\alpha$) of $\mathrm{DLF}(\mathcal{T}_m)$ by $D^\alpha(\mathcal{T}_m)$, that is, we set

$$D^\alpha(\mathcal{T}_m) := \sum_{i=2}^m (x_i - \max\{x_j : 1 \leq j < i, \ x_j < x_i\})^\alpha 1\{\min\{x_j : 1 \leq j < i\} < x_i\}.$$

Here we will take $\mathcal{T}_m$ to be random. In this case, set $\tilde{D}^\alpha(\mathcal{T}_m) := D^\alpha(\mathcal{T}_m) - \mathrm{E}[D^\alpha(\mathcal{T}_m)]$, the centred total weight of the DLF. In particular, let $(X_1, X_2, X_3, \ldots)$ be a sequence of independent, uniformly distributed random variables in $(0, 1]$, and for $m \in \mathbb{N}$ set $\mathcal{U}_m := (X_1, X_2, \ldots, X_m)$ and $\mathcal{U}_m^0 := (0, X_1, X_2, \ldots, X_m)$. We consider $D^\alpha(\mathcal{U}_m)$ and $D^\alpha(\mathcal{U}_m^0)$. Note that the DLF on $\mathcal{U}_m^0$ will always be a tree rooted at 0, and in this case we call it the directed linear tree.

For the random variables $D^\alpha(\mathcal{U}_m)$ and $D^\alpha(\mathcal{U}_m^0)$ we establish asymptotic behaviour of the mean value in Propositions 3.1 and 3.2, along with the following convergence results, which are the principal results of this section.

For $\alpha > 1$, let $D_\alpha$ denote a random variable with distribution characterized by

$$D_\alpha \overset{\mathrm{D}}{=} U^\alpha D_\alpha^{\{1\}} + (1 - U)^\alpha D_\alpha^{\{2\}} + U^\alpha, \tag{3.1}$$

where $U$ is uniform on $(0, 1)$ and independent of the other variables on the right-hand side. Also for $\alpha > 1$, let $F_\alpha$ denote a random variable with distribution characterized by

$$F_\alpha \overset{\mathrm{D}}{=} U^\alpha F_\alpha + (1 - U)^\alpha D_\alpha, \tag{3.2}$$

where $U$ is uniform on $(0, 1)$, $D_\alpha$ has the distribution given by (3.1), and $U$, the $D_\alpha$, and the $F_\alpha$ on the right-hand side are independent. The corresponding centred random variables, $\tilde{D}_\alpha := D_\alpha - E[D_\alpha]$ and $\tilde{F}_\alpha := F_\alpha - E[F_\alpha]$, respectively satisfy (2.4) and (2.5). The solutions to (2.4) and (2.5) are unique by the criterion given in (2.2), and hence the solutions to (3.1) and (3.2) are also unique. For $p = 1, 2$, let '$\to_{L^p}$' denote convergence in $p$th mean.

**Theorem 3.1.** (i) *As* $m \to \infty$ *we have* $\tilde{D}^1(\mathcal{U}_m^0) \to_{L^2} \tilde{D}_1$ *and* $\tilde{D}^1(\mathcal{U}_m) \to_{L^2} \tilde{F}_1$, *where* $\tilde{D}_1$ *has the distribution given by (2.3) and* $\tilde{F}_1$ *has the same distribution as* $\tilde{D}_1$. *Also, the variance of* $\tilde{D}_1$ *(and hence also of* $\tilde{F}_1$*) is* $2 - \pi^2/6 \approx 0.355\,066$. *Finally,* $\mathrm{cov}(\tilde{D}_1, \tilde{F}_1) = 7/4 - \pi^2/6 \approx 0.105\,066$.

(ii) *For* $\alpha > 1$, *as* $m \to \infty$ *we have* $D^\alpha(\mathcal{U}_m^0) \to D_\alpha$ *almost surely and in* $L^2$, *and* $D^\alpha(\mathcal{U}_m) \to F_\alpha$ *almost surely and in* $L^2$, *where the distributions of* $D_\alpha$ *and* $F_\alpha$ *are respectively given by (3.1) and (3.2). Also,* $E[D_\alpha] = (\alpha - 1)^{-1}$ *and* $E[F_\alpha] = (\alpha(\alpha - 1))^{-1}$, *while* $\mathrm{var}(D_\alpha)$ *and* $\mathrm{var}(F_\alpha)$ *are respectively given by (3.22) and (3.24), below.*

*Proof.* Part (i) follows from Propositions 3.5, 3.6, and 3.7, below. Part (ii) follows from Propositions 3.3 and 3.4. We prove these results in the following sections.

An interesting property of the DLT, which we use in establishing fixed-point equations for limit distributions, is its *self-similarity* (scaling property). In terms of the total weight, this says that, for any $t \in (0, 1)$, if $Y_1, \ldots, Y_n$ are independent and uniformly distributed on $(0, t]$, then the distribution of $D^\alpha(Y_1, \ldots, Y_n)$ is the same as that of $t^\alpha D^\alpha(X_1, \ldots, X_n)$.

### 3.1. The mean total weight of the DLF and DLT

First we consider the rooted case, i.e. the DLT on $\mathcal{U}_m^0$. For $m \in \mathbb{N}$, denote by $Z_m$ the random variable given by the gain in length of the tree upon the addition of one point, $X_m$, to an existing $m - 1$ points in the DLT on a sequence of uniform random variables $\mathcal{U}_{m-1}^0$; i.e. with the conventions $D^1(\mathcal{U}_0^0) = 0$ and $X_0 = 0$, we set

$$Z_m := D^1(\mathcal{U}_m^0) - D^1(\mathcal{U}_{m-1}^0) = X_m - \max\{X_j : 0 \le j < m, X_j < X_m\}. \qquad (3.3)$$

Thus, with weight exponent $\alpha$, the $m$th edge to be added has weight $Z_m^\alpha$.

**Lemma 3.1.** (i) $Z_m$ *has distribution function* $F_m$, *given by* $F_m(t) = 0$ *for* $t < 0$, $F_m(t) = 1$ *for* $t > 1$, *and* $F_m(t) = 1 - (1 - t)^m$ *for* $0 \le t \le 1$.

(ii) *For* $\beta > 0$, $Z_m$ *has moments*

$$E[Z_m^\beta] = \frac{\Gamma(m + 1)\Gamma(1 + \beta)}{\Gamma(1 + \beta + m)}. \qquad (3.4)$$

*In particular,*

$$E[Z_m] = \frac{1}{m + 1}, \qquad \mathrm{var}[Z_m] = \frac{m}{(m + 1)^2(m + 2)}. \qquad (3.5)$$

(iii) *For* $\beta > 0$, *as* $m \to \infty$ *we have*

$$E[Z_m^\beta] \sim \Gamma(1 + \beta)m^{-\alpha}. \qquad (3.6)$$

(iv) *As* $m \to \infty$, $mZ_m$ *converges in distribution to an exponential random variable with parameter* 1.

*Proof.* For $0 \leq t \leq 1$, we have

$$P[Z_m > t] = P[X_m > t \text{ and none of } X_1, \ldots, X_{m-1} \text{ lies in } (X_m - t, X_m)] = (1-t)^m,$$

and (i) follows. We then obtain (ii), since for any $\beta > 0$ we have

$$E[Z_m^\beta] = \int_0^1 P[Z_m > t^{1/\beta}]\, dt = \int_0^1 (1 - t^{1/\beta})^m \, dt = \frac{\Gamma(m+1)\Gamma(1+\beta)}{\Gamma(1+\beta+m)}.$$

Then (iii) follows by Stirling's formula. For (iv), we have from (i) that, for $t \in [0, \infty)$ and $m$ large enough that $t/m \leq 1$,

$$P[mZ_m \leq t] = F_m\left(\frac{t}{m}\right) = 1 - \left(1 - \frac{t}{m}\right)^m \to 1 - e^{-t}, \quad \text{as } m \to \infty.$$

However, $1 - e^{-t}$, $t \geq 0$, is the exponential distribution function with parameter 1.

Note that $Z_m$ has the same distribution as the spacing $S_1^n$ (see Section 3.2). The following result gives the asymptotic behaviour of the expected total weight of the DLT. Let $\gamma$ denote Euler's constant, whence

$$\sum_{i=1}^k \frac{1}{i} - \log k = \gamma + O(k^{-1}). \tag{3.7}$$

**Proposition 3.1.** *As $m \to \infty$, the expected total weight of the DLT on $\mathcal{U}_m^0$ satisfies*

$$E[D^\alpha(\mathcal{U}_m^0)] \sim \frac{\Gamma(\alpha+1)}{1-\alpha} m^{1-\alpha}, \qquad 0 < \alpha < 1, \tag{3.8}$$

$$E[D^1(\mathcal{U}_m^0)] - \log m \to \gamma - 1, \tag{3.9}$$

$$E[D^\alpha(\mathcal{U}_m^0)] = \frac{1}{\alpha-1} + O(m^{1-\alpha}), \qquad \alpha > 1. \tag{3.10}$$

*Proof.* We have

$$E[D^\alpha(\mathcal{U}_m^0)] = \sum_{i=1}^m (E[D^\alpha(\mathcal{U}_i^0)] - E[D^\alpha(\mathcal{U}_{i-1}^0)]) = \sum_{i=1}^m E[Z_i^\alpha].$$

For $\alpha = 1$, $E[Z_i] = (i+1)^{-1}$ by (3.5), and (3.9) follows from (3.7). For a general $\alpha > 0$, with $\alpha \neq 1$, from (3.4) we have

$$E[D^\alpha(\mathcal{U}_m^0)] = \Gamma(1+\alpha) \sum_{i=1}^m \frac{\Gamma(i+1)}{\Gamma(1+\alpha+i)} = \frac{1}{\alpha-1} - \frac{\Gamma(1+\alpha)\Gamma(m+2)}{(\alpha-1)\Gamma(m+1+\alpha)}, \tag{3.11}$$

which can be proved by induction on $m$. By Stirling's formula, the last term satisfies

$$-\frac{\Gamma(1+\alpha)\Gamma(m+2)}{(\alpha-1)\Gamma(m+1+\alpha)} = -\frac{\Gamma(1+\alpha)}{\alpha-1} m^{1-\alpha}(1 + O(m^{-1})), \tag{3.12}$$

which, for $\alpha > 1$, tends to 0 as $m \to \infty$, to give us (3.10). For $\alpha < 1$, we obtain (3.8) from (3.11) and (3.12).

Now consider the unrooted case, i.e. the DLF. For $\mathcal{U}_m$ as above, the total weight of the DLF is denoted by $D^\alpha(\mathcal{U}_m)$, and the centred total weight by $\tilde{D}^\alpha(\mathcal{U}_m) := D^\alpha(\mathcal{U}_m) - \mathrm{E}[D^\alpha(\mathcal{U}_m)]$. We then see that

$$D^\alpha(\mathcal{U}_m^0) = D^\alpha(\mathcal{U}_m) + \mathcal{L}_0^\alpha(\mathcal{U}_m^0), \tag{3.13}$$

where $\mathcal{L}_0^\alpha(\mathcal{U}_m^0)$ is the total weight of edges incident at 0 in the DLT on $\mathcal{U}_m^0$.

The following lemma says that $\mathcal{L}_0^\alpha(\mathcal{U}_m^0)$ converges to a random variable that has the generalized Dickman distribution with parameter $1/\alpha$ (see [14]), that is, the distribution of a random variable $X$ which satisfies the distributional fixed-point equation

$$X \stackrel{\mathrm{D}}{=} U^\alpha(1 + X), \tag{3.14}$$

where $U$ is uniform on $(0, 1)$ and independent of the $X$ on the right-hand side. We recall from Proposition 3 of [14] that if $X$ satisfies (3.14), then

$$\mathrm{E}[X] = \frac{1}{\alpha} \quad \text{and} \quad \mathrm{E}[X^2] = \frac{\alpha + 2}{2\alpha^2}. \tag{3.15}$$

**Lemma 3.2.** *Let* $\alpha > 0$. *There is a random variable* $\mathcal{L}_0^\alpha$, *with the generalized Dickman distribution with parameter* $1/\alpha$, *such that as* $m \to \infty$ *we have* $\mathcal{L}_0^\alpha(\mathcal{U}_m^0) \to \mathcal{L}_0^\alpha$ *almost surely and in* $L^2$.

*Proof.* Let $\delta_D(\mathcal{U}_m^0)$ denote the degree of the origin in the directed linear tree on $\mathcal{U}_m^0$, implying that $\delta_D(\mathcal{U}_m^0)$ is the number of lower records in the sequence $(X_1, \ldots, X_m)$. Then

$$\mathcal{L}_0^\alpha(\mathcal{U}_m^0) = U_1^\alpha + (U_1 U_2)^\alpha + \cdots + (U_1 \cdots U_{\delta_D(\mathcal{U}_m^0)})^\alpha,$$

where $(U_1, U_2, \ldots)$ is a certain sequence of independent, uniform random variables on $(0, 1)$, namely the ratios between successive lower records of the sequence $(X_n)$. The sum $U_1^\alpha + (U_1 U_2)^\alpha + (U_1 U_2 U_3)^\alpha + \cdots$ has nonnegative terms and finite expectation, so it converges almost surely to a limit, which we denote $\mathcal{L}_0^\alpha$. Then $\mathcal{L}_0^\alpha$ has the generalized Dickman distribution with parameter $1/\alpha$ (see Proposition 2 of [14]).

Since $\delta_D(\mathcal{U}_m^0)$ tends to $\infty$ almost surely as $m \to \infty$, we have $\mathcal{L}_0^\alpha(\mathcal{U}_m^0) \to \mathcal{L}_0^\alpha$ almost surely. Also, $\mathrm{E}[(\mathcal{L}_0^\alpha)^2] < \infty$ by (3.15), and $(\mathcal{L}_0^\alpha - \mathcal{L}_0^\alpha(\mathcal{U}_m^0))^2 \leq (\mathcal{L}_0^\alpha)^2$ for all $m$. Thus, $\mathrm{E}[(\mathcal{L}_0^\alpha(\mathcal{U}_m^0) - \mathcal{L}_0^\alpha)^2] \to 0$ by the dominated convergence theorem, from which the required $L^2$-convergence also follows.

**Proposition 3.2.** *As* $m \to \infty$, *the expected total weight of the DLF on* $\mathcal{U}_m$ *satisfies*

$$\mathrm{E}[D^\alpha(\mathcal{U}_m)] \sim \frac{\Gamma(\alpha + 1)}{1 - \alpha} m^{1-\alpha}, \qquad 0 < \alpha < 1, \tag{3.16}$$

$$\mathrm{E}[D^1(\mathcal{U}_m)] - \log m \to \gamma - 2, \tag{3.17}$$

$$\mathrm{E}[D^\alpha(\mathcal{U}_m)] \to \frac{1}{\alpha(\alpha - 1)}, \qquad \alpha > 1. \tag{3.18}$$

*Proof.* By (3.13) we have $\mathrm{E}[D^\alpha(\mathcal{U}_m)] = \mathrm{E}[D^\alpha(\mathcal{U}_m^0)] - \mathrm{E}[\mathcal{L}_0^\alpha(\mathcal{U}_m^0)]$, and by Lemma 3.2 and (3.15) we have

$$\mathrm{E}[\mathcal{L}_0^\alpha(\mathcal{U}_m^0)] \longrightarrow \mathrm{E}[\mathcal{L}_0^\alpha] = 1/\alpha.$$

We then obtain (3.16), (3.17), and (3.18) from Proposition 3.1.

In the following sections we present the limiting behaviour of the DLT or DLF in the cases $\alpha = 1$ and $\alpha > 1$. The case $\alpha < 1$ does not concern us here. However, a divide and conquer approach, as used in [15] to prove a limit theorem for the total weight of the on-line nearest-neighbour graph on $(0, 1)$ when $\frac{1}{2} < \alpha < 1$, can be used to give a similar result for the DLT or DLF.

### 3.2. Orthogonal increments for $\alpha = 1$

In this section we shall show (in Lemma 3.5) that when $\alpha = 1$, the variables $Z_i$, $i \geq 1$, are mutually orthogonal, in the sense of having zero covariances, which fact will be used later on to establish convergence of the (centred) total length of the DLT. To prove this, we first need some further notation.

Given $X_1, \ldots, X_m$, let us denote the order statistics of $X_1, \ldots, X_m$, taken in increasing order, as $X_{(1)}^m, X_{(2)}^m, \ldots, X_{(m)}^m$. Thus, $(X_{(1)}^m, X_{(2)}^m, \ldots, X_{(m)}^m)$ is a nondecreasing sequence forming a permutation of the original, $(X_1, \ldots, X_m)$. Denote the existing $m + 1$ intervals between points by $I_j^m := (X_{(j-1)}^m, X_{(j)}^m)$, $j = 1, 2, \ldots, m + 1$, where we set $X_{(0)}^m := 0$ and $X_{(m+1)}^m := 1$. Let the widths of these intervals (the spacings) be

$$S_j^m := |I_j^m| = X_{(j)}^m - X_{(j-1)}^m, \qquad 1 \leq j \leq m + 1.$$

Then $0 \leq S_j^m < 1$ for $1 \leq j \leq m + 1$, and $\sum_{j=1}^{m+1} S_j^m = 1$, i.e. the vector $(S_1^m, S_2^m, \ldots, S_{m+1}^m)$ belongs to the $m$-dimensional simplex $\Delta_m$. Note that only $m$ of the $S_j^m$ are required to specify the vector.

We can arrange the spacings themselves (i.e. $S_j^m$, $1 \leq j \leq m+1$) into increasing order to give $S_{(1)}^m, S_{(2)}^m, \ldots, S_{(m+1)}^m$. Then let $\mathcal{F}_S^m$ denote the $\sigma$-field generated by these ordered spacings, i.e. $\mathcal{F}_S^m = \sigma(S_{(1)}^m, \ldots, S_{(m+1)}^m)$. The following interpretation of $\mathcal{F}_S^m$ may be helpful. The set $(0, 1) \setminus \{X_1, \ldots, X_m\}$ consists almost surely of $m + 1$ connected components ('fragments') of total length 1, and $\mathcal{F}_S^m$ is the $\sigma$-field generated by the collection of lengths of these fragments, ignoring the order in which they appear.

By definition, the value of $Z_m$ must be one of the (ordered) spacings $S_{(1)}^m, \ldots, S_{(m+1)}^m$. The next result says that, given the values of these spacings, each of the possible values for $Z_m$ is equally likely.

**Lemma 3.3.** *For $m \geq 1$, we have*

$$P[Z_m = S_{(i)}^m \mid \mathcal{F}_S^m] = \frac{1}{m + 1} \quad \text{almost surely, for } i = 1, \ldots, m + 1. \qquad (3.19)$$

*Hence,*

$$E[Z_m \mid \mathcal{F}_S^m] = \frac{1}{m + 1} \sum_{i=1}^{m+1} S_{(i)}^m = \frac{1}{m + 1}. \qquad (3.20)$$

*Proof.* We know that $(S_1^m, \ldots, S_{m+1}^m)$ is uniform over the $m$-dimensional simplex $\Delta_m$. In particular, the $S_j^m$ are exchangeable. Thus, given $S_{(1)}^m, \ldots, S_{(m+1)}^m$, i.e. $\mathcal{F}_S^m$, the actual values of $S_1^m, \ldots, S_{m+1}^m$ are equally likely to be any permutation of $S_{(1)}^m, \ldots, S_{(m+1)}^m$, and, given $S_1^m, \ldots, S_{m+1}^m$, the value of $Z_m$ is equally likely to be any one of $S_1^m, \ldots, S_m^m$ (but cannot be $S_{m+1}^m$).

Hence, given $S_{(1)}^m, \ldots, S_{(m+1)}^m$, the probability that $Z_m = S_{(i)}^m$ is $(1/m) \times m/(m + 1) = 1/(m + 1)$, yielding (3.19); (3.20) then follows from $\sum_{j=1}^{m+1} S_{(j)}^m = 1$.

**Lemma 3.4.** *Let $1 \leq m < \ell$. Given $\mathcal{F}_S^m$, $Z_\ell$ and $Z_m$ are conditionally independent.*

*Proof.* Given $\mathcal{F}_S^m$, we have $S_{(1)}^m, \ldots, S_{(m+1)}^m$, and by (3.19) the (conditional) distribution of $Z_m$ is uniform on $\{S_{(1)}^m, \ldots, S_{(m+1)}^m\}$. Furthermore, given $\mathcal{F}_S^m$, the conditional distribution of $Z_\ell$, $\ell > m$, depends only on $S_{(1)}^m, \ldots, S_{(m+1)}^m$, and not which one of them $Z_m$ happens to be. Hence, $Z_m$ and $Z_\ell$ are conditionally independent.

**Lemma 3.5.** *For $1 \leq m < \ell$, the random variables $Z_m$ and $Z_\ell$ satisfy $\mathrm{cov}(Z_m, Z_\ell) = 0$.*

*Proof.* From Lemmas 3.4 and 3.3,

$$\mathrm{E}[Z_m Z_\ell \mid \mathcal{F}_S^m] = \mathrm{E}[Z_m \mid \mathcal{F}_S^m] \, \mathrm{E}[Z_\ell \mid \mathcal{F}_S^m] = \frac{1}{m+1} \, \mathrm{E}[Z_\ell \mid \mathcal{F}_S^m],$$

and by taking expectations we obtain

$$\mathrm{E}[Z_m Z_\ell] = \frac{1}{m+1} \, \mathrm{E}[Z_\ell] = \frac{1}{m+1} \frac{1}{\ell+1} = \mathrm{E}[Z_m] \, \mathrm{E}[Z_\ell].$$

Hence, the covariance of $Z_m$ and $Z_\ell$ is 0.

### 3.3. Limit behaviour for $\alpha > 1$

We now consider the limit distribution of the total weight of the DLT and DLF. In this section we consider the case of $\alpha$-power-weighted edges with $\alpha > 1$; that is, we prove part (ii) of Theorem 3.1. To describe the moments of the limiting distribution of $D^\alpha(\mathcal{U}_m^0)$ and $D^\alpha(\mathcal{U}_m)$, we introduce the notation

$$J(\alpha) := \int_0^1 u^\alpha (1-u)^\alpha \, \mathrm{d}u = 2^{-1-2\alpha} \sqrt{\pi} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+\frac{3}{2})}. \tag{3.21}$$

We start with the rooted case, $D^\alpha(\mathcal{U}_m^0)$, and subsequently consider the unrooted case, $D^\alpha(\mathcal{U}_m)$.

**Proposition 3.3.** *Let $\alpha > 1$. Then there exists a random variable $D_\alpha$ such that, as $m \to \infty$, we have $D^\alpha(\mathcal{U}_m^0) \to D_\alpha$ almost surely and in $L^2$. Also, $D_\alpha$ satisfies the distributional fixed-point equality (3.1). Furthermore, $\mathrm{E}[D_\alpha] = 1/(\alpha-1)$ and*

$$\mathrm{var}(D_\alpha) = \frac{\alpha(\alpha - 2 + 2(2\alpha+1)J(\alpha))}{(\alpha-1)^2(2\alpha-1)}. \tag{3.22}$$

*Proof.* Let $Z_i$ be the length of the $i$th edge of the DLT, as defined in (3.3). Let $D_\alpha := \sum_{i=1}^\infty Z_i^\alpha$. The sum converges almost surely, since it has nonnegative terms and, by (3.6), has finite expectation for $\alpha > 1$. By (3.6) and the Cauchy–Schwarz inequality, there exists a constant, $C$, $0 < C < \infty$, such that

$$\mathrm{E}[D_\alpha^2] = \sum_{i=1}^\infty \sum_{j=1}^\infty \mathrm{E}[Z_i^\alpha Z_j^\alpha] \leq C \sum_{i=1}^\infty \sum_{j=1}^\infty i^{-\alpha} j^{-\alpha} < \infty,$$

since $\alpha > 1$. The stated $L^2$-convergence then follows from the dominated convergence theorem.

Taking $U = X_1$ here, by the self-similarity of the DLT we have

$$D^\alpha(\mathcal{U}_m^0) \overset{\mathrm{D}}{=} U^\alpha D_{\{1\}}^\alpha(\mathcal{U}_N^0) + (1-U)^\alpha D_{\{2\}}^\alpha(\mathcal{U}_{m-1-N}^0) + U^\alpha, \tag{3.23}$$

where $N \sim \text{Bin}(m-1, U)$, given $U$, and, given $U$ and $N$, $D^\alpha_{\{1\}}(\mathcal{U}^0_N)$ and $D^\alpha_{\{2\}}(\mathcal{U}^0_{m-1-N})$ are independent and have the distributions of $D^\alpha(\mathcal{U}^0_N)$ and $D^\alpha(\mathcal{U}^0_{m-1-N})$, respectively. As $m \to \infty$, $N$ and $m - N$ both tend to $\infty$ almost surely; thus, by taking $m \to \infty$ in (3.23), we obtain the fixed-point equation (3.1).

The identity $\text{E}[D_\alpha] = (\alpha - 1)^{-1}$ is obtained either from (3.10) or by taking expectations in (3.1). Next, if we set $\tilde{D}_\alpha = D_\alpha - \text{E}[D_\alpha]$, (3.1) yields (2.4). Then, using the definition, (3.21), of $J(\alpha)$, the fact that $\text{E}[\tilde{D}_\alpha] = 0$, and independence, from (2.4) we obtain

$$\text{E}[\tilde{D}^2_\alpha] = \frac{2\,\text{E}[\tilde{D}^2_\alpha]}{2\alpha + 1} + \frac{\alpha^2 + 1}{(\alpha - 1)^2(2\alpha + 1)} + \frac{2\alpha J(\alpha)}{(\alpha - 1)^2} - \frac{1}{(\alpha - 1)^2}.$$

Rearranging this gives (3.22).

Recall from Lemma 3.2 that $\mathcal{L}^\alpha_0$ is the limiting weight of edges attached to the origin in the DLT on uniform points. Combining this fact with Proposition 3.3, we obtain a similar result to the latter for the unrooted case, as follows.

**Proposition 3.4.** *Let $\alpha > 1$. There is a random variable $F_\alpha$, satisfying the distributional fixed-point equality (3.2), such that as $n \to \infty$, $D^\alpha(\mathcal{U}_m) \to F_\alpha$ almost surely and in $L^2$. Furthermore, $\text{E}[F_\alpha] = 1/(\alpha(\alpha - 1))$ and*

$$\text{var}(F_\alpha) = \frac{1}{2\alpha}\,\text{var}(D_\alpha) + \frac{\alpha + 2(2\alpha + 1)J(\alpha) - 2}{2\alpha^2(\alpha - 1)^2}, \tag{3.24}$$

*where $J(\alpha)$ is given by (3.21) and $\text{var}(D_\alpha)$ by (3.22).*

*Proof.* By Lemma 3.2 and Proposition 3.3, there are random variables $D_\alpha$ and $\mathcal{L}^\alpha_0$ such that, as $m \to \infty$, we have $D^\alpha(\mathcal{U}^0_m) \to_{L^2} D_\alpha$ and $\mathcal{L}^\alpha_0(\mathcal{U}^0_m) \to_{L^2} \mathcal{L}^\alpha_0$, with almost-sure convergence in both cases. Hence, with $F_\alpha := D_\alpha - \mathcal{L}^\alpha_0$, by (3.13) we have

$$D^\alpha(\mathcal{U}_m) = D^\alpha(\mathcal{U}^0_m) - \mathcal{L}^\alpha_0(\mathcal{U}^0_m) \to F_\alpha \quad \text{almost surely and in } L^2. \tag{3.25}$$

Next we show that $F_\alpha$ satisfies the distributional fixed-point equality (3.2). The self-similarity of the DLT implies that

$$D^\alpha(\mathcal{U}_m) \overset{\text{D}}{=} U^\alpha D^\alpha(\mathcal{U}_N) + (1 - U)^\alpha D^\alpha(\mathcal{U}^0_{m-1-N}), \tag{3.26}$$

where $N \sim \text{Bin}(m-1, U)$, given $U$, and $D^\alpha(\mathcal{U}_N)$ and $D^\alpha(\mathcal{U}^0_{m-1-N})$ are independent, given $U$ and $N$. As $m \to \infty$, $N$ and $m - N$ both tend to $\infty$ almost surely, so by taking $m \to \infty$ in (3.26) and using Proposition 3.3 and (3.25), we obtain the fixed-point equation (3.2).

The identity $\text{E}[F_\alpha] = \alpha^{-1}(\alpha - 1)^{-1}$ is obtained either from (3.18) or by taking expectations in (3.2) and using the formula for $\text{E}[D_\alpha]$ of Proposition 3.3. Then, with $\tilde{F}_\alpha := F_\alpha - \text{E}[F_\alpha]$, we obtain (2.5) from (3.2), and, using independence and the fact that $\text{E}[\tilde{F}_\alpha] = \text{E}[\tilde{D}_\alpha] = 0$, we obtain

$$\frac{2\alpha}{2\alpha + 1}\,\text{E}[\tilde{F}^2_\alpha] = \frac{\text{E}[\tilde{D}^2_\alpha]}{2\alpha + 1} + \frac{2\alpha J(\alpha) - 1}{\alpha^2(\alpha - 1)^2} + \frac{\alpha^2 + 1}{\alpha^2(\alpha - 1)^2(2\alpha + 1)},$$

which yields (3.24).

### 3.4. Limit behaviour for $\alpha = 1$

Unlike in the case $\alpha > 1$, for $\alpha = 1$ the mean of the total weight $D^1(\mathcal{U}_m^0)$ diverges as $m \to \infty$ (see Proposition 3.1), so clearly there is no limiting distribution for $D^1(\mathcal{U}_m^0)$. Nevertheless, by using the orthogonality of the increments of the sequence $(D^1(\mathcal{U}_m^0), m \geq 1)$, we are able to show that the *centred* total weight, $\tilde{D}^1(\mathcal{U}_m^0)$, does converge in distribution (in fact, in $L^2$) to a limiting random variable, and likewise in the unrooted case; this is our next result.

Subsequently, we shall characterize the distribution of the limiting random variable (in both the rooted and unrooted cases) by a fixed-point identity, and thereby complete the proof of Theorem 3.1(i).

**Proposition 3.5.** (i) *As $m \to \infty$, the random variable $\tilde{D}^1(\mathcal{U}_m^0)$ converges in $L^2$ to a limiting random variable $\tilde{D}_1$ with* $\mathrm{E}[\tilde{D}_1] = 0$ *and* $\mathrm{var}(\tilde{D}_1) = 2 - \pi^2/6$. *In particular,* $\mathrm{var}(D^1(\mathcal{U}_m^0)) \to 2 - \pi^2/6$ *as $m \to \infty$.*

(ii) *As $m \to \infty$, $\tilde{D}^1(\mathcal{U}_m)$ converges in $L^2$ to the limiting random variable* $\tilde{F}_1 := \tilde{D}_1 - \mathcal{L}_0^1 + 1$.

*Proof.* Adopt the convention $D^1(\mathcal{U}_0^0) = 0$. By the orthogonality of the $Z_j$ (Lemma 3.5) and (3.5), for $0 \leq \ell < m$ we have

$$
\mathrm{var}(\tilde{D}^1(\mathcal{U}_m^0) - \tilde{D}^1(\mathcal{U}_\ell^0)) = \mathrm{var}\left( \sum_{j=\ell+1}^{m} (Z_j - \mathrm{E}[Z_j]) \right)
$$

$$
= \sum_{j=\ell+1}^{m} \frac{j}{(j+1)^2(j+2)}
$$

$$
\to 0 \quad \text{as } m, \ell \to \infty.
$$

Hence, $\tilde{D}_1(\mathcal{U}_m^0)$ is a Cauchy sequence in $L^2$ and, so, converges in $L^2$ to a limiting random variable, which we denote $\tilde{D}_1$. Then $\mathrm{E}[\tilde{D}_1] = \lim_{m \to \infty} \mathrm{E}[\tilde{D}_1(\mathcal{U}_m^0)] = 0$ and

$$
\mathrm{var}(\tilde{D}_1) = \lim_{m \to \infty} \mathrm{var}(\tilde{D}^1(\mathcal{U}_m^0))
$$

$$
= \sum_{j=1}^{\infty} \frac{j}{(j+1)^2(j+2)}
$$

$$
= \sum_{j=1}^{\infty} \left[ \frac{2}{j+1} - \frac{2}{j+2} \right] - \sum_{j=1}^{\infty} \frac{1}{(j+1)^2}
$$

$$
= 1 - \left( \frac{\pi^2}{6} - 1 \right)
$$

$$
= 2 - \frac{\pi^2}{6}.
$$

It remains to prove part (ii), the convergence of the centred total length of the DLF $\tilde{D}^1(\mathcal{U}_m)$. By (3.13) we have

$$
\tilde{D}^1(\mathcal{U}_m) = \tilde{D}^1(\mathcal{U}_m^0) - \mathcal{L}_0^1(\mathcal{U}_m^0) + \mathrm{E}[\mathcal{L}_0^1(\mathcal{U}_m^0)] \to_{L^2} \tilde{D}_1 - \mathcal{L}_0^1 + 1,
$$

where the convergence follows by Lemma 3.2 and part (i). Thus, $\tilde{D}^1(\mathcal{U}_m)$ converges in $L^2$ as $m \to \infty$.

For the next few results it is more convenient to consider the DLF defined on a Poisson-distributed number of points. Let $(X_1, X_2, \ldots)$ be a sequence of independent, uniformly distributed random variables in $(0, 1]$, and let $(N(t), t \geq 0)$ be the counting process of a homogeneous Poisson process of unit rate in $(0, \infty)$, independent of $(X_1, X_2, \ldots)$. Thus, $N(t)$ is a Poisson variable with parameter $t$. As before, let $\mathcal{U}_m = (X_1, \ldots, X_m)$, and (in this section only) let $\mathcal{P}_t := \mathcal{U}_{N(t)}$. Let $\mathcal{P}_t^0 := \mathcal{U}_{N(t)}^0$, whence $\mathcal{P}_t^0 = (0, X_1, X_2, \ldots, X_{N(t)})$.

We construct the DLF and DLT on $X_1, X_2, \ldots, X_{N(t)}$ as before. Let $\tilde{D}^1(\mathcal{P}_t^0) = D^1(\mathcal{P}_t^0) - \mathrm{E}[D^1(\mathcal{P}_t^0)]$ and $\tilde{D}^1(\mathcal{P}_t) = D^1(\mathcal{P}_t) - \mathrm{E}[D^1(\mathcal{P}_t)]$. We aim to show that the limit distribution for $\tilde{D}^1(\mathcal{P}_t^0)$ is the same as that for $\tilde{D}^1(\mathcal{U}_m^0)$, and likewise in the unrooted case. We shall need the following result.

**Lemma 3.6.** *As $t \to \infty$,*

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{E}[D^1(\mathcal{P}_t)] = \frac{1}{t} + O(t^{-2}) \quad and \quad \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{E}[D^1(\mathcal{P}_t^0)] = \frac{1}{t} + O(t^{-2}).$$

*Proof.* The point set $\{X_1, \ldots, X_{N(t)}\}$ is a homogeneous Poisson point process in $(0, 1)$, so we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{E}[D^1(\mathcal{P}_t)] = \mathrm{E}[\text{length of new arrival}]$$

$$= \int_0^1 \mathrm{d}u \, \mathrm{E}[\text{distance to next point to the left of } u \text{ in } \mathcal{P}_t]$$

$$= \int_0^1 \mathrm{d}u \int_0^u st\mathrm{e}^{-ts} \, \mathrm{d}s$$

$$= \frac{1}{t} + \frac{2}{t^2}(\mathrm{e}^{-t} - 1) + \frac{\mathrm{e}^{-t}}{t}$$

$$= \frac{1}{t} + O(t^{-2}).$$

Similarly,

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{E}[D^1(\mathcal{P}_t^0)] = \int_0^1 \mathrm{d}u \, \mathrm{E}[\text{distance to next point to the left of } u \text{ in } \mathcal{P}_t \cup \{0\}]$$

$$= \int_0^1 \mathrm{d}u \int_0^u \mathrm{e}^{-ts} \, \mathrm{d}s$$

$$= \frac{1}{t} + \frac{\mathrm{e}^{-t} - 1}{t^2}$$

$$= \frac{1}{t} + O(t^{-2}).$$

**Lemma 3.7.** (i) *As $t \to \infty$, $\tilde{D}^1(\mathcal{P}_t^0)$ converges in distribution to $\tilde{D}_1$, the $L^2$ large-m limit of $\tilde{D}^1(\mathcal{U}_m^0)$.*

(ii) *As $t \to \infty$, $\tilde{D}^1(\mathcal{P}_t)$ converges in distribution to $\tilde{F}_1$, the $L^2$ large-m limit of $\tilde{D}^1(\mathcal{U}_m)$.*

*Proof.* (i) From Proposition 3.5, we have $\tilde{D}^1(\mathcal{U}_m^0) \to_{L^2} \tilde{D}_1$ as $m \to \infty$. Let $a_t := \mathrm{E}[D^1(\mathcal{P}_t^0)]$ and $\mu_m := \mathrm{E}[D^1(\mathcal{U}_m^0)]$. Since

$$\mu_m = \mathrm{E}\left[\sum_{i=1}^m Z_i\right] = \sum_{i=1}^m (1+i)^{-1},$$

by (3.5), for any positive integers $\ell$ and $m$ we have

$$|\mu_m - \mu_\ell| = \sum_{j=\min(m,\ell)+1}^{\max(m,\ell)} \frac{1}{j+1} \leq \log\left(\frac{\max(m,\ell)+1}{\min(m,\ell)+1}\right) = \left|\log\left(\frac{m+1}{\ell+1}\right)\right|. \quad (3.27)$$

Note the distributional equalities

$$\mathcal{L}(D^1(\mathcal{P}_t^0) \mid N(t) = m) = \mathcal{L}(D^1(\mathcal{U}_m^0)),$$
$$\mathcal{L}(D^1(\mathcal{P}_t^0) - \mu_{N(t)} \mid N(t) = m) = \mathcal{L}(\tilde{D}^1(\mathcal{U}_m^0)). \quad (3.28)$$

First we aim to show that $a_t - \mu_{\lfloor t \rfloor} \to 0$ as $t \to \infty$, where $\lfloor x \rfloor$ denotes the integer part of $x$. Set $p_m(t) := e^{-t} t^m / m!$. Then we can write

$$a_t - \mu_{\lfloor t \rfloor} = \sum_{m=0}^{\infty} p_m(t)(\mu_m - \mu_{\lfloor t \rfloor})$$

$$= \sum_{\{m : |m-\lfloor t \rfloor| \leq t^{3/4}\}} p_m(t)(\mu_m - \mu_{\lfloor t \rfloor}) + \sum_{\{m : |m-\lfloor t \rfloor| > t^{3/4}\}} p_m(t)(\mu_m - \mu_{\lfloor t \rfloor}). \quad (3.29)$$

We examine these two sums separately. First consider the sum for $|m - \lfloor t \rfloor| \leq t^{3/4}$. By (3.27), we have

$$\sup_{\{m : |m-\lfloor t \rfloor| \leq t^{3/4}\}} |\mu_m - \mu_{\lfloor t \rfloor}| \leq \max\left(\log\left(\frac{\lfloor t \rfloor + 1 + t^{3/4}}{\lfloor t \rfloor + 1}\right), \log\left(\frac{\lfloor t \rfloor + 1}{\lfloor t \rfloor + 1 - t^{3/4}}\right)\right)$$

$$= O(t^{-1/4})$$

$$\to 0 \quad \text{as } t \to \infty.$$

Hence, the first sum in (3.29) tends to 0 as $t \to \infty$. To estimate the second sum, observe that

$$\sum_{\{m : |m-\lfloor t \rfloor| > t^{3/4}\}} p_m(t)(\mu_m - \mu_{\lfloor t \rfloor}) \leq \sum_{\{m : |m-\lfloor t \rfloor| > t^{3/4}\}} p_m(t)(m + t)$$

$$= E[(N(t) + t)\mathbf{1}\{|N(t) - \lfloor t \rfloor| > t^{3/4}\}]$$

$$\leq (E[(N(t) + t)^2] P[|N(t) - \lfloor t \rfloor| > t^{3/4}])^{1/2}. \quad (3.30)$$

By Chernoff bounds on the tail probabilities of a Poisson random variable (see, e.g. Lemma 1.4 of [11]), (3.30) is $O(t \exp(-t^2/18))$ and, so, tends to 0. Hence, the second sum in (3.29) tends to 0 and, thus,

$$a_t - \mu_{\lfloor t \rfloor} \to 0 \quad \text{as } t \to \infty. \quad (3.31)$$

Now we show that $\tilde{D}^1(\mathcal{P}_t^0) \xrightarrow{\text{D}} \tilde{D}_1$ as $t \to \infty$. We have

$$\tilde{D}^1(\mathcal{P}_t^0) = (D^1(\mathcal{P}_t^0) - \mu_{N(t)}) + (\mu_{N(t)} - \mu_{\lfloor t \rfloor}) + (\mu_{\lfloor t \rfloor} - a_t). \quad (3.32)$$

The final bracket tends to 0, by (3.31). Also, by (3.28) and the fact that $N(t) \to \infty$ almost surely as $t \to \infty$, we have

$$D^1(\mathcal{P}_t^0) - \mu_{N(t)} \xrightarrow{\text{D}} \tilde{D}_1.$$

Finally, using (3.27), we have

$$\left| \mu_{N(t)} - \mu_{\lfloor t \rfloor} \right| \leq \left| \log\left( \frac{N(t)+1}{\lfloor t \rfloor + 1} \right) \right| \xrightarrow{\text{P}} 0 \quad \text{as } t \to \infty,$$

since $N(t)/\lfloor t \rfloor \xrightarrow{\text{P}} 1$ in the same limit. Slutsky's theorem applied to (3.32) thus yields $\tilde{D}^1(\mathcal{P}_t^0) \xrightarrow{\text{D}} \tilde{D}_1$ as $t \to \infty$, completing the proof of (i).

The proof of (ii) follows in the same way as that of (i), except that in (3.27) the first equality is replaced by the inequality '$\leq$'. This does not affect the rest of the proof.

The next two propositions complete the proof of Theorem 3.1.

**Proposition 3.6.** *The limiting random variable $\tilde{D}_1$ of Proposition 3.5(i) satisfies the fixed-point equation (2.3).*

*Proof.* For integer $n > 0$, let $T_n := \min\{s\colon N(s) \geq n\}$, the $n$th arrival time of the Poisson process with counting process $N(\cdot)$. Set $T := T_1$ and set $U := X_1$ (the latter of which is uniform on $(0, 1)$).

By the marking theorem for Poisson processes [9, p. 55], the two-dimensional point process $\mathcal{Q} := \{(X_n, T_n)\colon n \geq 1\}$ is a homogeneous Poisson process of unit intensity on $(0, 1) \times (0, \infty)$. Given the value of $(U, T)$, the restriction of $\mathcal{Q}$ to $(0, U] \times (T, \infty)$ and the restriction of $\mathcal{Q}$ to $(U, 1] \times (T, \infty)$ are independent, homogeneous Poisson processes on these regions. Hence, by scaling properties of the Poisson process (see the mapping theorem of [9, pp. 17–19]) and of the DLT, and writing $D^1_{\{i\}}(\cdot)$, $i = 1, 2$, for independent copies of $D^1(\cdot)$, we have

$$D^1(\mathcal{P}_t^0) \stackrel{\text{D}}{=} (U D^1_{\{1\}}(\mathcal{P}_{U(t-T)}^0) + (1-U) D^1_{\{2\}}(\mathcal{P}_{(1-U)(t-T)}^0) + U)\mathbf{1}\{t > T\}. \tag{3.33}$$

Let $a_s = 0$ for $s \leq 0$ and $a_s = \mathrm{E}[D^1(\mathcal{P}_s^0)]$ for $s > 0$. Then $\tilde{D}^1(\mathcal{P}_t^0) = D^1(\mathcal{P}_t^0) - a_t$, whence, by (3.33),

$$\tilde{D}^1(\mathcal{P}_t^0) \stackrel{\text{D}}{=} (U \tilde{D}^1_{\{1\}}(\mathcal{P}_{U(t-T)}^0) + (1-U) \tilde{D}^1_{\{2\}}(\mathcal{P}_{(1-U)(t-T)}^0) + U)\mathbf{1}\{t > T\}$$
$$+ U(a_{U(t-T)} - a_t) + (1-U)(a_{(1-U)(t-T)} - a_t). \tag{3.34}$$

From Lemma 3.6 we have $\mathrm{d}a_t/\mathrm{d}t = 1/t + O(t^{-2})$. Hence, if $T < t$ then

$$a_t - a_{U(t-T)} = \int_{U(t-T)}^t \frac{\mathrm{d}a_s}{\mathrm{d}s}\,\mathrm{d}s = \log t - \log(U(t-T)) + O((U(t-T))^{-1})$$

and, hence, as $t \to \infty$,

$$a_t - a_{U(t-T)} \to -\log U \quad \text{almost surely.} \tag{3.35}$$

Since $\mathrm{P}[T < t]$ tends to 1, by taking $t \to \infty$ in (3.34) and using Slutsky's theorem we obtain (2.3).

**Proposition 3.7.** *The limiting random variable $\tilde{F}_1$ of Proposition 3.5(ii) satisfies the fixed-point equation (2.3) and, so, has the same distribution as $\tilde{D}_1$. Also, $\mathrm{cov}(\tilde{F}_1, \tilde{D}_1) = 7/4 - \pi^2/6$.*
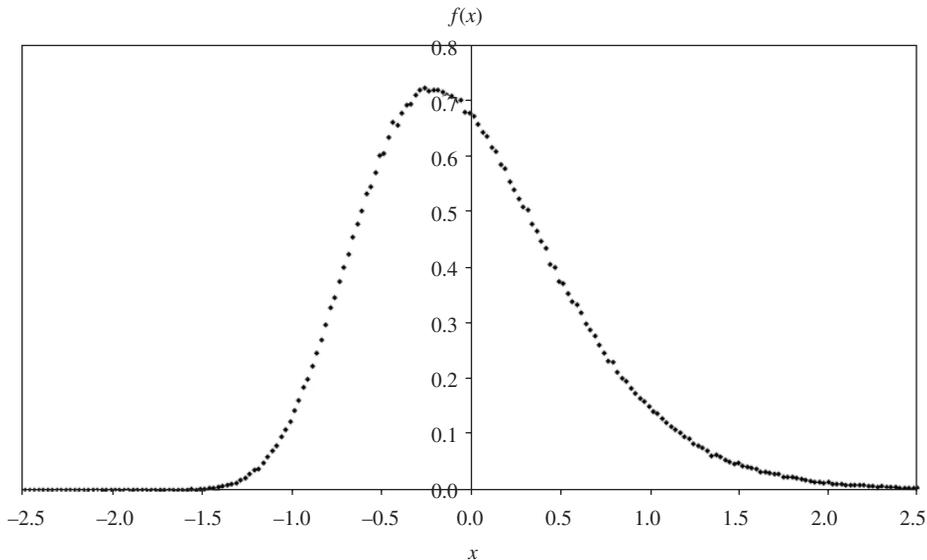
FIGURE 2: Estimated probability density function of $\tilde{D}_1$.

*Proof.* The proof follows similar lines to that of Proposition 3.6. Once more let $a_s = \mathrm{E}[D^1(\mathcal{P}_s^0)]$ for $s \geq 0$ and $a_s = 0$ for $s < 0$. Let $b_s = \mathrm{E}[D^1(\mathcal{P}_s)]$ for $s > 0$ and $b_s = 0$ for $s \leq 0$, and let $T := \min\{t : N(t) \geq 1\}$. Then

$$D^1(\mathcal{P}_t) \stackrel{\mathrm{D}}{=} (U D_{\{1\}}^1(\mathcal{P}_{U(t-T)}) + (1 - U)D_{\{2\}}^1(\mathcal{P}_{(1-U)(t-T)}^0))1\{t > T\}, \qquad (3.36)$$

where $D_{\{1\}}^1(\cdot)$ and $D_{\{2\}}^1(\cdot)$ are independent copies of $D^1(\cdot)$. Then $\tilde{D}^1(\mathcal{P}_t) = D^1(\mathcal{P}_t) - b_t$ and $\tilde{D}^1(\mathcal{P}_t^0) = D^1(\mathcal{P}_t^0) - a_t$, by which (3.36) yields

$$\tilde{D}^1(\mathcal{P}_t) \stackrel{\mathrm{D}}{=} (U \tilde{D}_{\{1\}}^1(\mathcal{P}_{U(t-T)}) + (1 - U)\tilde{D}_{\{2\}}^1(\mathcal{P}_{(1-U)(t-T)}^0))1\{t > T\}$$
$$+ U(b_{U(t-T)} - b_t) + (1 - U)(a_{(1-U)(t-T)} - b_t). \qquad (3.37)$$

From Lemma 3.6 we have $\mathrm{d}b_t/\mathrm{d}t = 1/t + O(t^{-2})$. Hence, by the same argument as used to obtain (3.35),

$$b_t - b_{U(t-T)} \to -\log U \quad \text{almost surely.}$$

Also, $a_t - b_t = \mathrm{E}[\mathcal{L}_0^1(\mathcal{P}_t^0)]$ by (3.13), whence $\lim_{t\to\infty}(a_t - b_t) = 1$ by Lemma 3.2 and the fact that $\mathrm{E}[\mathcal{L}_0^1] = 1$ (by (3.15)). From (3.35) we also find that as $t \to \infty$,

$$a_{(1-U)(t-T)} - b_t = (a_{(1-U)(t-T)} - a_t) + (a_t - b_t) \to 1 + \log(1 - U) \quad \text{almost surely.}$$

By taking $t \to \infty$ in (3.37) and using Slutsky's theorem we obtain

$$\tilde{F}_1 \stackrel{\mathrm{D}}{=} U \tilde{F}_1 + (1 - U)\tilde{D}_1 + U \log U + (1 - U) \log(1 - U) + (1 - U). \qquad (3.38)$$

The change of variable $(1 - U) \mapsto U$ then shows that $\tilde{D}_1$, as defined in (2.3), satisfies (3.38); thus, by the uniqueness of solution, $\tilde{F}_1$ has the same distribution as $\tilde{D}_1$ and satisfies (2.3).

To obtain the covariance of $\tilde{F}_1$ and $\tilde{D}_1$, observe from Proposition 3.5(ii) that $\mathcal{L}_0^1 = \tilde{D}_1 - \tilde{F}_1 + 1$. Therefore, by (3.15), we have

$$\tfrac{1}{2} = \mathrm{var}(\mathcal{L}_0^1) = \mathrm{var}(\tilde{D}_1) + \mathrm{var}(\tilde{F}_1) - 2\,\mathrm{cov}(\tilde{D}_1, \tilde{F}_1). \tag{3.39}$$

Since $\mathrm{var}(\tilde{F}_1) = \mathrm{var}(\tilde{D}_1) = 2 - \pi^2/6$ by Proposition 3.5(i), by rearranging (3.39) we find that $\mathrm{cov}(\tilde{D}_1, \tilde{F}_1) = 7/4 - \pi^2/6$.

**Remark 3.1.** Figure 2 is a plot of the estimated probability density function of $\tilde{D}_1$. This was obtained by performing $10^6$ repeated simulations of the DLT on a sequence of $10^3$ uniform (simulated) random points in $(0, 1]$. For each simulation, the expected value of $D^1(\mathcal{U}_{10^3})$ (which is precisely $\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{1001}$, by Lemma 3.1) was subtracted from the total length of the simulated DLT to give an approximate realization of $\tilde{D}_1$. The density function was then estimated from the sample of $10^6$ approximate realizations of $\tilde{D}_1$, using a window width of 0.0025. The simulated sample from which the density estimate for $\tilde{D}_1$ was taken had approximate sample mean $-2 \times 10^{-4}$ and approximate sample variance 0.3543, which are reasonably close to the expectation and variance of $\tilde{D}_1$.

## 4. General central limit theorems in geometric probability

Notions of *stabilizing* functionals of point sets have recently proved to be a useful basis for a general methodology for establishing limit theorems for functionals of random point sets in $\mathbb{R}^d$. In particular, Penrose and Yukich [16] provided general central limit theorems for stabilizing functionals. One might hope to apply these results in the case of the MDSF weight. However, to obtain the central limit theorem for edges away from the boundary in the MDSF and MDST, we need an extension of the general result of [16], which we describe in this section.

To describe our general results we use the following notation. Let $d \geq 1$ be an integer. For $\mathcal{X} \subset \mathbb{R}^d$, a constant $a > 0$, and $\mathbf{y} \in \mathbb{R}^d$, let $\mathbf{y} + a\mathcal{X}$ denote the transformed set $\{\mathbf{y} + a\mathbf{x} : \mathbf{x} \in \mathcal{X}\}$. Let $\mathrm{diam}(\mathcal{X}) := \sup\{\|\mathbf{x}_1 - \mathbf{x}_2\| : \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}\}$ and let $\mathrm{card}(\mathcal{X})$ denote the cardinality (the number of elements) of $\mathcal{X}$ (when finite).

For $\mathbf{x} \in \mathbb{R}^d$ and $r > 0$, let $B(\mathbf{x}; r)$ denote the closed Euclidean ball with centre $\mathbf{x}$ and radius $r$, and let $Q(\mathbf{x}; r)$ denote the corresponding $l_\infty$-ball, i.e. the $d$-cube $\mathbf{x} + [-r, r]^d$. For a bounded measurable set $R \subset \mathbb{R}^d$, let $|R|$ denote the Lebesgue measure of $R$, let $\partial R$ denote the topological boundary of $R$, and, for $r > 0$, set $\partial_r R := \bigcup_{\mathbf{x} \in \partial R} Q(\mathbf{x}; r)$, the $r$-neighbourhood of the boundary of $R$.

Let $\xi(\mathbf{x}; \mathcal{X})$ be a measurable, $\mathbb{R}_+$-valued function defined for all pairs $(\mathbf{x}, \mathcal{X})$, where $\mathcal{X} \subset \mathbb{R}^d$ is finite and $\mathbf{x} \in \mathcal{X}$. Assume that $\xi$ is translation invariant, that is, $\xi(\mathbf{y} + \mathbf{x}; \mathbf{y} + \mathcal{X}) = \xi(\mathbf{x}; \mathcal{X})$ for all $\mathbf{y} \in \mathbb{R}^d$. When $\mathbf{x} \notin \mathcal{X}$, we abbreviate the notation $\xi(\mathbf{x}; \mathcal{X} \cup \{\mathbf{x}\})$ to $\xi(\mathbf{x}; \mathcal{X})$. For $\tau \in (0, \infty)$, let $\mathcal{H}_\tau$ be a homogeneous Poisson process of intensity $\tau$ on $\mathbb{R}^d$.

A translation-invariant, real-valued functional $\xi(\mathbf{x}; \mathcal{X})$ defined for finite point sets $\mathcal{X} \subset \mathbb{R}^d$ and points $\mathbf{x} \in \mathcal{X}$ induces a translation-invariant functional $H(\mathcal{X}; S)$ defined on all $\mathcal{X} \subset \mathbb{R}^d$ and all Borel-measurable regions $S \subseteq \mathbb{R}^d$ by

$$H(\mathcal{X}; S) := \sum_{\mathbf{x} \in \mathcal{X} \cap S} \xi(\mathbf{x}; \mathcal{X}). \tag{4.1}$$

It is this 'restricted' functional that interests us here; in [16] the authors were concerned rather with the global functional $H(\mathcal{X}; \mathbb{R}^d)$. In our particular application (the length of the edges of the MDST on random points in a square), the global functional fails to satisfy the conditions

of the central limit theorems of [16], owing to boundary effects. Here we generalize the result of [16] to the 'restricted' functional $H(\mathcal{X}; S)$. It is this generalized result that we can apply to the MDST, when we take $S$ to be a region 'away from the boundary' of the square in which the random points are placed.

We use a notion of *stabilization* for $H$. Loosely speaking, $\xi$ is stabilizing if, when a point is inserted at the origin into a homogeneous Poisson process, only nearby Poisson points affect the inserted point; for $H$ to be stabilizing we require also that the inserted point affects only nearby points.

For $B \subseteq \mathbb{R}^d$, let $\Delta(\mathcal{X}; B)$ denote the 'add-one cost' of the functional $H$ on the insertion of a point at the origin:

$$\Delta(\mathcal{X}; B) := H(\mathcal{X} \cup \{\mathbf{0}\}; B) - H(\mathcal{X}; B).$$

Let $\mathcal{P} := \mathcal{H}_1$, a homogeneous Poisson point process of unit intensity on $\mathbb{R}^d$. Adapting the ideas of [16], we make the following definitions.

**Definition 4.1.** We say the functional $H$ is *strongly stabilizing* if there exist almost surely finite random variables $R$ (a *radius of stabilization*) and $\Delta(\infty)$ such that with probability 1, for any $B \supseteq B(\mathbf{0}; R)$,

$$\Delta(\mathcal{P} \cap B(\mathbf{0}; R) \cup \mathcal{A}; B) = \Delta(\infty) \quad \text{for all finite sets } \mathcal{A} \subset \mathbb{R}^d \setminus B(\mathbf{0}; R).$$

We say that the functional $H$ is *polynomially bounded* if for all $B \ni \mathbf{0}$ there exists a constant, $\beta$, such that for all finite sets $\mathcal{X} \subset \mathbb{R}^d$,

$$|H(\mathcal{X}; B)| \leq \beta(\operatorname{diam}(\mathcal{X}) + \operatorname{card}(\mathcal{X}))^{\beta}. \tag{4.2}$$

We say that $H$ is *homogeneous of order $\gamma$* if for all finite sets $\mathcal{X} \subset \mathbb{R}^d$, all Borel sets $B \subseteq \mathbb{R}^d$, and all $a \in \mathbb{R}$, it is the case that

$$H(a\mathcal{X}; aB) = a^{\gamma} H(\mathcal{X}; B).$$

Let $(R_n, S_n)$, $n = 1, 2, \ldots$, be a sequence of ordered pairs of bounded Borel subsets of $\mathbb{R}^d$ such that $S_n \subseteq R_n$ for all $n$. Assume that, for all $r > 0$, $n^{-1}|\partial_r R_n| \to 0$ and $n^{-1}|\partial_r S_n| \to 0$ (the *vanishing relative boundary condition*). Assume also that $|R_n| = n$ for all $n$ and that $|S_n|/n \to 1$ as $n \to \infty$; that $S_n$ tends to $\mathbb{R}^d$ in the sense that $\bigcup_{n \geq 1} \bigcap_{m \geq n} S_m = \mathbb{R}^d$; and that there exists a constant, $\beta$, such that $\operatorname{diam}(R_n) \leq \beta n^{\beta}$ for all $n$ (the *polynomial boundedness* condition on $(R_n, S_n)_{n \geq 1}$). Subject to these conditions, the choice of $(R_n, S_n)_{n \geq 1}$ is arbitrary.

Let $\boldsymbol{U}_{1,n}, \boldsymbol{U}_{2,n}, \ldots$ be independent, identically distributed uniform random vectors on $R_n$. Let

$$\mathcal{U}_{m,n} = \{\boldsymbol{U}_{1,n}, \ldots, \boldsymbol{U}_{m,n}\}$$

(a binomial point process), and for Borel sets $A \subseteq \mathbb{R}^d$ with $0 < |A| < \infty$, let $\mathcal{U}_{m,A}$ be the binomial point process of $m$ independent, identically distributed uniform random vectors on $A$.

Let $\mathcal{R}$ be the collection of all pairs $(A, B)$, with $A, B \subset \mathbb{R}^d$, of the form

$$(A, B) = (\boldsymbol{x} + R_n, \boldsymbol{x} + S_n),$$

with $\boldsymbol{x} \in \mathbb{R}^d$ and $n \in \mathbb{N}$. That is, $\mathcal{R}$ is the collection of all the $(R_n, S_n)$ and their translates.

We say that the functional $H$ satisfies the *uniform bounded moments condition* on $\mathcal{R}$ if

$$\sup_{\{(A,B) \in \mathcal{R} : \, \mathbf{0} \in A\}} \left( \sup_{|A|/2 \leq m \leq 3|A|/2} \mathrm{E}[\Delta(\mathcal{U}_{m,A}; B)^4] \right) < \infty. \tag{4.3}$$

Our first general result extends Theorem 2.1 of [16]. We omit the proof here; it can be found in [13]. Let $\mathcal{Q}_n := \mathcal{P} \cap R_n$, the restriction of $\mathcal{P}$ to $R_n$

**Theorem 4.1.** *Suppose that H is strongly stabilizing, is polynomially bounded (see 4.2)), and satisfies the uniform bounded moments condition, (4.3), on $\mathcal{R}$. Then there exist constants $s^2$ and $t^2$, with $0 \leq t^2 \leq s^2$, such that as $n \to \infty$,*

(i) $n^{-1} \operatorname{var}(H(\mathcal{Q}_n; S_n)) \to s^2,$

(ii) $n^{-1/2}(H(\mathcal{Q}_n; S_n) - \mathrm{E}[H(\mathcal{Q}_n; S_n)]) \xrightarrow{\mathrm{D}} \mathcal{N}(0, s^2),$

(iii) $n^{-1} \operatorname{var}(H(\mathcal{U}_{n,n}; S_n)) \to t^2,$

(iv) $n^{-1/2}(H(\mathcal{U}_{n,n}; S_n) - \mathrm{E}[H(\mathcal{U}_{n,n}; S_n)]) \xrightarrow{\mathrm{D}} \mathcal{N}(0, t^2).$

*Also, $s^2$ and $t^2$ are independent of the choice of the $(R_n, S_n)$. Furthermore, if the distribution of $\Delta(\infty)$ is nondegenerate, then $s^2 \geq t^2 > 0$.*

Our second general result generalizes Corollary 2.1 of [16]. Let $R_0$ be a fixed, bounded Borel subset of $\mathbb{R}^d$ with $|R_0| = 1$ and $|\partial R_0| = 0$. Let $(S_{0,n}, n \geq 1)$ be a sequence of Borel sets with $S_{0,n} \subseteq R_0$, such that both $|S_{0,n}| \to 1$ as $n \to \infty$ and, for all $r > 0$, we have $|\partial_{n^{-1/d}r} S_{0,n}| \to 0$ as $n \to \infty$.

Let $\mathcal{R}_0$ be the collection of all pairs of the form

$$(\boldsymbol{x} + n^{1/d} R_0, \boldsymbol{x} + n^{1/d} S_{0,n})$$

with $n \geq 1$ and $\boldsymbol{x} \in \mathbb{R}^d$. Let $\mathcal{X}_n$ be the binomial point process of $n$ independent, identically distributed uniform random vectors on $R_0$, and let $\mathcal{P}_n$ be a homogeneous Poisson point process of intensity $n$ on $R_0$.

**Corollary 4.1.** *Suppose that H is strongly stabilizing, satisfies the uniform bounded moments condition on $\mathcal{R}_0$, is polynomially bounded, and is homogeneous of order $\gamma$. Then with $s^2$ and $t^2$ as in Theorem 4.1, as $n \to \infty$ we have*

(i) $n^{2\gamma/d-1} \operatorname{var}(H(\mathcal{P}_n; S_{0,n})) \to s^2,$

(ii) $n^{\gamma/d-1/2}(H(\mathcal{P}_n; S_{0,n}) - \mathrm{E}[H(\mathcal{P}_n; S_{0,n})]) \xrightarrow{\mathrm{D}} \mathcal{N}(0, s^2),$

(iii) $n^{2\gamma/d-1} \operatorname{var}(H(\mathcal{X}_n; S_{0,n})) \to t^2,$

(iv) $n^{\gamma/d-1/2}(H(\mathcal{X}_n; S_{0,n}) - \mathrm{E}[H(\mathcal{X}_n; S_{0,n})]) \xrightarrow{\mathrm{D}} \mathcal{N}(0, t^2).$

*Proof.* The corollary follows from Theorem 4.1 by taking $R_n = n^{1/d} R_0$ and $S_n = n^{1/d} S_{0,n}$ (or suitable translates thereof), and scaling, since $H$ is homogeneous of order $\gamma$.

## 5. Central limit theorem away from the boundary

While it should be possible to adapt the argument of this section to more general partial orders, from now on we take the partial order on $\mathbb{R}^2$ to be '$\preccurlyeq^*$'. For each $n$, define the region $S_{0,n} := (n^{\varepsilon-1/2}, 1]^2$, where $\varepsilon \in (0, \frac{1}{2})$ is a small constant to be chosen later. In this section, we use the general central limit theorems of Section 4 to demonstrate a central limit theorem for the contribution to the total weight of the MDSF, under '$\preccurlyeq^*$', from edges away from the boundary, that is, from points in the region $S_{0,n}$.

Given $\alpha > 0$, consider the MDSF total weight functional $H = \mathcal{L}^\alpha$ on point sets in $\mathbb{R}^2$. We take $\xi(\boldsymbol{x}; \mathcal{X})$ to be $d(\boldsymbol{x}; \mathcal{X})^\alpha$, where $d(\boldsymbol{x}; \mathcal{X})$ is the distance from point $\boldsymbol{x}$ to its directed nearest neighbour in $\mathcal{X}$ under '$\preccurlyeq^*$', if such a neighbour exists, and is 0 otherwise. Thus, in our case,

$$\xi(\boldsymbol{x}; \mathcal{X}) = d(\boldsymbol{x}; \mathcal{X})^\alpha \quad \text{with} \quad d(\boldsymbol{x}; \mathcal{X}) := \min\{\|\boldsymbol{x} - \boldsymbol{y}\| : \boldsymbol{y} \in \mathcal{X} \setminus \{\boldsymbol{x}\}, \ \boldsymbol{y} \preccurlyeq^* \boldsymbol{x}\},$$

with the convention that $\min \varnothing = 0$. For $R \subseteq \mathbb{R}^2$, let

$$\mathcal{L}^\alpha(\mathcal{X}; R) = \sum_{\boldsymbol{x} \in \mathcal{X} \cap R} \xi(\boldsymbol{x}; \mathcal{X}) \tag{5.1}$$

and $\mathcal{L}^\alpha(\mathcal{X}) := \mathcal{L}^\alpha(\mathcal{X}; \mathbb{R}^2)$.

Let $\mathcal{X}_n$ be the binomial point process of $n$ independent, identically distributed uniform random vectors in $(0, 1]^2$, and let $\mathcal{P}_n$ be a homogeneous Poisson process of intensity $n$ on $(0, 1]^2$. The main result of this section is the following theorem.

**Theorem 5.1.** *Suppose that $\alpha > 0$ and the partial order is '$\preccurlyeq^*$'. Then there exist constants $t_\alpha$ and $s_\alpha$, $0 < t_\alpha \leq s_\alpha$, not depending on the choice of $\varepsilon$, such that as $n \to \infty$,*

(i) $n^{\alpha-1} \operatorname{var}(\mathcal{L}^\alpha(\mathcal{X}_n; S_{0,n})) \to t_\alpha^2$,

(ii) $n^{(\alpha-1)/2} \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n; S_{0,n}) \xrightarrow{\text{D}} \mathcal{N}(0, t_\alpha^2)$,

(iii) $n^{\alpha-1} \operatorname{var}(\mathcal{L}^\alpha(\mathcal{P}_n; S_{0,n})) \to s_\alpha^2$,

(iv) $n^{(\alpha-1)/2} \tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; S_{0,n}) \xrightarrow{\text{D}} \mathcal{N}(0, s_\alpha^2)$.

The following corollary states that Theorem 5.1 is also true in the rooted cases, i.e. with $\mathcal{X}_n$ replaced by $\mathcal{X}_n^0$ and $\mathcal{P}_n$ replaced by $\mathcal{P}_n^0$.

**Corollary 5.1.** *Suppose that $\alpha > 0$ and the partial order is '$\preccurlyeq^*$'. Then, with $t_\alpha$ and $s_\alpha$ as given in Theorem 5.1, as $n \to \infty$ we have*

(i) $n^{\alpha-1} \operatorname{var}(\mathcal{L}^\alpha(\mathcal{X}_n^0; S_{0,n})) \to t_\alpha^2$,

(ii) $n^{(\alpha-1)/2} \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n^0; S_{0,n}) \xrightarrow{\text{D}} \mathcal{N}(0, t_\alpha^2)$,

(iii) $n^{\alpha-1} \operatorname{var}(\mathcal{L}^\alpha(\mathcal{P}_n^0; S_{0,n})) \to s_\alpha^2$,

(iv) $n^{(\alpha-1)/2} \tilde{\mathcal{L}}^\alpha(\mathcal{P}_n^0; S_{0,n}) \xrightarrow{\text{D}} \mathcal{N}(0, s_\alpha^2)$.

Theorem 5.1 and Corollary 5.1 were proved in [13]. The proof of the theorem relies on showing that the functional $\mathcal{L}^\alpha$ satisfies suitable versions of the conditions of Theorem 4.1 and Corollary 4.1, namely that $\mathcal{L}^\alpha$ is polynomially bounded (see (4.2)), homogeneous of order $\alpha$, and strongly stabilizing (see Definition 4.1). Also, the distribution of $\Delta(\infty)$ is nondegenerate. Finally, with $R_0 := (0, 1]^2$, recalling that $S_{0,n} := (n^{\varepsilon-1/2}, 1]^2$ throughout this section, and with $\mathcal{R}_0$ as defined just before Corollary 4.1, $\mathcal{L}^\alpha$ satisfies the uniform bounded moments condition, (4.3), on $\mathcal{R}_0$. For the details, see [13].

## 6. The edges near the boundary

Next in our analysis of the MDST on random points in the unit square, we consider the length of the edges close to the boundary of the square. The limiting structure of the MDSF

and MDST near the boundaries is described by the directed linear forest model discussed in Section 3.

We initially consider the 'rooted' case, in which we insert a point at the origin. Later we shall analyse the multiple sink (or 'unrooted') case, in which we do not insert a point at the origin, in a similar way.

Fix $\sigma \in (\frac{1}{2}, \frac{2}{3})$. Let $B_n$ denote the L-shaped boundary region $(0, 1]^2 \setminus (n^{-\sigma}, 1]^2$. Recall from (5.1) that $\mathcal{L}^\alpha(\mathcal{X}; R)$ denotes the contribution to the total weight of the MDST on $\mathcal{X}$ from edges starting at points of $\mathcal{X} \cap R$. When $\mathcal{X}$ is a random point set, set $\tilde{\mathcal{L}}^\alpha(\mathcal{X}; R) := \mathcal{L}^\alpha(\mathcal{X}; R) - \mathrm{E}[\mathcal{L}^\alpha(\mathcal{X}; R)]$.

**Theorem 6.1.** *Suppose that the partial order is '$\preccurlyeq^*$'. Then as $n \to \infty$,*

$$\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n^0; B_n) \xrightarrow{\mathrm{D}} \tilde{D}_\alpha^{\{1\}} + \tilde{D}_\alpha^{\{2\}}, \qquad \alpha \geq 1, \tag{6.1}$$

$$\tilde{\mathcal{L}}^\alpha(\mathcal{X}_n^0; B_n) \xrightarrow{\mathrm{D}} \tilde{D}_\alpha^{\{1\}} + \tilde{D}_\alpha^{\{2\}}, \qquad \alpha \geq 1, \tag{6.2}$$

*where $\tilde{D}_\alpha^{\{1\}}$ and $\tilde{D}_\alpha^{\{2\}}$ are independent random variables with the distribution of $\tilde{D}_\alpha$ given by the fixed-point equation (2.3) for $\alpha = 1$ and by (2.4) for $\alpha > 1$. As $n \to \infty$ we also have*

$$\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; B_n) \xrightarrow{\mathrm{D}} \tilde{F}_\alpha^{\{1\}} + \tilde{F}_\alpha^{\{2\}}, \qquad \alpha \geq 1, \tag{6.3}$$

$$\tilde{\mathcal{L}}^\alpha(\mathcal{X}_n; B_n) \xrightarrow{\mathrm{D}} \tilde{F}_\alpha^{\{1\}} + \tilde{F}_\alpha^{\{2\}}, \qquad \alpha \geq 1, \tag{6.4}$$

*where $\tilde{F}_\alpha^{\{1\}}$ and $\tilde{F}_\alpha^{\{2\}}$ are independent random variables with the same distribution as $\tilde{D}_1$ for $\alpha = 1$, and with the distribution given by the fixed-point equation (2.5) for $\alpha > 1$. Finally, as $n \to \infty$ we additionally have*

$$n^{(\alpha-1)/2}\mathcal{L}^\alpha(\mathcal{P}_n; B_n) \to_{L^1} 0, \qquad 0 < \alpha < 1, \tag{6.5}$$

$$n^{(\alpha-1)/2}\mathcal{L}^\alpha(\mathcal{P}_n^0; B_n) \to_{L^1} 0, \qquad 0 < \alpha < 1. \tag{6.6}$$

The idea behind the proof of Theorem 6.1 is to show that the MDSF near each of the two boundaries is close to a DLF system defined on a sequence of uniform random variables coupled to the points of the MDSF. To do this, we produce two explicit sequences of random variables on which we construct the DLF coupled to $\mathcal{P}_n$, the Poisson process of intensity $n$ on $(0, 1]^2$, on which the MDSF is constructed.

Let $B_n^x$ be the rectangle $(n^{-\sigma}, 1] \times (0, n^{-\sigma}]$, let $B_n^y$ be the rectangle $(0, n^{-\sigma}] \times (n^{-\sigma}, 1]$, and let $B_n^0$ be the square $(0, n^{-\sigma}]^2$; see Figure 3. Then $B_n = B_n^0 \cup B_n^x \cup B_n^y$. Define the point processes

$$\mathcal{V}_n^x := \mathcal{P}_n \cap (B_n^x \cup B_n^0), \qquad \mathcal{V}_n^y := \mathcal{P}_n \cap (B_n^y \cup B_n^0), \quad \text{and} \quad \mathcal{V}_n^0 := \mathcal{P}_n \cap B_n^0, \tag{6.7}$$

and let $N_n^x := \mathrm{card}(\mathcal{V}_n^x)$, $N_n^y := \mathrm{card}(\mathcal{V}_n^y)$, and $N_n^0 := \mathrm{card}(\mathcal{V}_n^0)$. List $\mathcal{V}_n^x$, in order of increasing $y$-coordinate, as $X_i^x$, $i = 1, 2, \ldots, N_n^x$. In coordinates, set $X_i^x = (X_i^x, Y_i^x)$ for each $i$. Similarly, list $\mathcal{V}_n^y$, in order of increasing $x$-coordinate, as $X_i^y = (X_i^y, Y_i^y)$, $i = 1, \ldots, N_n^y$. Let $\mathcal{U}_n^x = (X_i^x, i = 1, 2, \ldots, N_n^x)$ and $\mathcal{U}_n^y = (Y_i^y, i = 1, 2, \ldots, N_n^y)$. Then $\mathcal{U}_n^x$ and $\mathcal{U}_n^y$ are sequences of uniform random variables in $(0, 1]$, on which we may construct a DLF. Also, we write $\mathcal{U}_n^{x,0}$ for the sequence $(0, X_1^x, X_2^x, \ldots, X_{N_n^x}^x)$, and $\mathcal{U}_n^{y,0}$ for the sequence $(0, Y_1^y, Y_2^y, \ldots, Y_{N_n^y}^y)$.

With the total DLF or DLT weight functional $D^\alpha(\cdot)$ defined in Section 3 for random finite sequences in $(0, 1)$, the DLF weight $D^\alpha(\mathcal{U}_n^x)$ is coupled in a natural way to the MDSF contribution $\mathcal{L}^\alpha(\mathcal{V}_n^x)$, and likewise for $D^\alpha(\mathcal{U}_n^y)$ and $\mathcal{L}^\alpha(\mathcal{V}_n^y)$, for $D^\alpha(\mathcal{U}_n^{x,0})$ and $\mathcal{L}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\})$, and for $D^\alpha(\mathcal{U}_n^{y,0})$ and $\mathcal{L}^\alpha(\mathcal{V}_n^y \cup \{\mathbf{0}\})$.
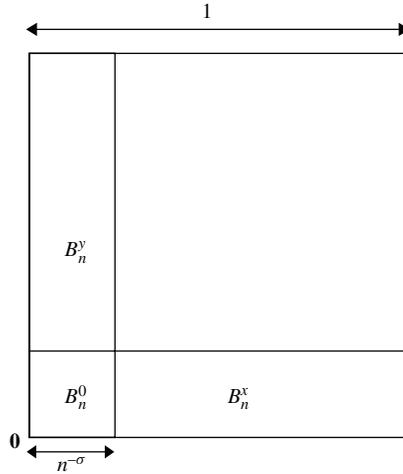
FIGURE 3: The boundary regions.

**Lemma 6.1.** *For any $\alpha \geq 1$, as $n \to \infty$,*

$$\mathcal{L}^\alpha(\mathcal{V}_n^x) - D^\alpha(\mathcal{U}_n^x) \to_{L^2} 0 \quad and \quad \mathcal{L}^\alpha(\mathcal{V}_n^y) - D^\alpha(\mathcal{U}_n^y) \to_{L^2} 0, \tag{6.8}$$

$$\mathcal{L}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\}) - D^\alpha(\mathcal{U}_n^{x,0}) \to_{L^2} 0 \quad and \quad \mathcal{L}^\alpha(\mathcal{V}_n^y \cup \{\mathbf{0}\}) - D^\alpha(\mathcal{U}_n^{y,0}) \to_{L^2} 0. \tag{6.9}$$

*Furthermore, for $0 < \alpha < 1$, as $n \to \infty$,*

$$\mathrm{E}[|\mathcal{L}^\alpha(\mathcal{V}_n^x) - D^\alpha(\mathcal{U}_n^x)|^2] = O(n^{2-2\sigma-2\alpha\sigma}), \tag{6.10}$$

*and a corresponding result holds for $\mathcal{V}_n^y$ and $\mathcal{U}_n^y$, and for the rooted cases (with the addition of the origin).*

*Proof.* We approximate the MDSF in the region $B_n$ by two DLFs, coupled to the MDSF. Consider $\mathcal{V}_n^x$; the argument for $\mathcal{V}_n^y$ is entirely analogous.

We have the set of points $\mathcal{V}_n^x = \{(X_i^x, Y_i^x),\ i = 1, \ldots, N_n^x\}$. We construct the MDSF on these points and construct the DLF on the $x$-coordinates, $\mathcal{U}_n^x = (X_i^x,\ i = 1, \ldots, N_n^x)$. Consider any point $(X_i^x, Y_i^x)$. For any single point, either an edge exists from that point in both constructions or in neither. Suppose that an edge exists; that is, suppose that $X_i^x$ is joined to a point $X_{D(i)}^x$, $D(i) < i$, in the DLF model, and $(X_i^x, Y_i^x)$ to a point $(X_{N(i)}^x, Y_{N(i)}^x)$ in the MDST (we do not necessarily have $N(i) = D(i)$). By construction, we know that $|X_i^x - X_{D(i)}^x| \leq |X_i^x - X_{N(i)}^x|$, since $N(i) < i$ according to the ordering of our points. It then follows that

$$\|(X_i^x, Y_i^x) - (X_{N(i)}^x, Y_{N(i)}^x)\|^\alpha \geq |X_i^x - X_{N(i)}^x|^\alpha \geq |X_i^x - X_{D(i)}^x|^\alpha,$$

and so we have established that for all $\alpha > 0$,

$$D^\alpha(\mathcal{U}_n^x) \leq \mathcal{L}^\alpha(\mathcal{V}_n^x) \quad \text{and} \quad D^\alpha(\mathcal{U}_n^{x,0}) \leq \mathcal{L}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\}).$$

Now, by the construction of the MDST, we have

$$\|(X_i^x, Y_i^x) - (X_{N(i)}^x, Y_{N(i)}^x)\| \leq \|(X_i^x, Y_i^x) - (X_{D(i)}^x, Y_{D(i)}^x)\|. \tag{6.11}$$

If $(x, y) \in (0, 1]^2$ then $\|(x, y)\| \leq x + y$, and, by the mean value theorem for the function $t \mapsto t^\alpha$, for $\alpha \geq 1$ we have

$$\|(x, y)\|^\alpha - x^\alpha \leq (x + y)^\alpha - x^\alpha \leq \alpha 2^{\alpha-1} y, \qquad \alpha \geq 1.$$

Hence, for $\alpha \geq 1$,

$$\|(X_i^x, Y_i^x) - (X_{D(i)}^x, Y_{D(i)}^x)\|^\alpha - (X_i^x - X_{D(i)}^x)^\alpha \leq \alpha 2^{\alpha-1}(Y_i^x - Y_{D(i)}^x). \qquad (6.12)$$

Then, for $\alpha \geq 1$, (6.11) and (6.12) yield

$$\|(X_i^x, Y_i^x) - (X_{N(i)}^x, Y_{N(i)}^x)\|^\alpha - (X_i^x - X_{D(i)}^x)^\alpha \leq \alpha 2^{\alpha-1}(Y_i^x - Y_{D(i)}^x).$$

Hence, for $\alpha \geq 1$,

$$0 \leq \mathcal{L}^\alpha(\mathcal{V}_n^x) - D^\alpha(\mathcal{U}_n^x) \leq \alpha 2^{\alpha-1} \sum_{i=1}^{N_n^x}(Y_i^x - Y_{D(i)}^x)$$

and, thus,

$$0 \leq \mathcal{L}^\alpha(\mathcal{V}_n^x) - D^\alpha(\mathcal{U}_n^x) \leq \alpha 2^{\alpha-1} N_n^x n^{-\sigma},$$
$$0 \leq \mathcal{L}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\}) - D^\alpha(\mathcal{U}_n^{x,0}) \leq \alpha 2^{\alpha-1} N_n^x n^{-\sigma}. \qquad (6.13)$$

We have $N_n^x \sim \text{Po}(n^{1-\sigma})$, whence, since $\sigma > \frac{1}{2}$,

$$\mathrm{E}[(\mathcal{L}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\}) - D^\alpha(\mathcal{U}_n^{x,0}))^2] \leq \alpha^2 2^{2\alpha-2} n^{-2\sigma} \mathrm{E}[(N_n^x)^2] \to 0, \qquad \alpha \geq 1.$$

An entirely analogous argument leads to the same statement for $\mathcal{U}_n^y$ and $\mathcal{V}_n^y$, and we obtain (6.8), and (6.9) in identical fashion.

Now let $0 < \alpha < 1$. By the concavity of the function $t \mapsto t^\alpha$ for $\alpha < 1$, for $x > 0$ and $y > 0$ we have

$$\|(x, y)\|^\alpha - x^\alpha \leq (x + y)^\alpha - x^\alpha \leq y^\alpha, \qquad 0 < \alpha < 1.$$

Then, by an argument similar to that yielding (6.13) in the $\alpha \geq 1$ case, we obtain

$$0 \leq \mathcal{L}^\alpha(\mathcal{V}_n^x) - D^\alpha(\mathcal{U}_n^x) \leq N_n^x n^{-\alpha\sigma}.$$

Then (6.10) follows since $N_n^x \sim \text{Po}(n^{1-\sigma})$, and the rooted case is similar.

**Lemma 6.2.** *Suppose that $\tilde{D}_1$ has distribution given by (2.3), $\tilde{D}_\alpha$, $\alpha > 1$, has distribution given by (2.4), and $\tilde{F}_\alpha$, $\alpha > 1$, has distribution given by (2.5). Then as $n \to \infty$,*

$$\tilde{\mathcal{L}}^1(\mathcal{V}_n^x \cup \{\mathbf{0}\}) \overset{\mathrm{D}}{\to} \tilde{D}_1 \quad and \quad \tilde{\mathcal{L}}^1(\mathcal{V}_n^x) \overset{\mathrm{D}}{\to} \tilde{D}_1, \qquad (6.14)$$
$$\tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\}) \overset{\mathrm{D}}{\to} \tilde{D}_\alpha \quad and \quad \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x) \overset{\mathrm{D}}{\to} \tilde{F}_\alpha, \qquad \alpha > 1. \qquad (6.15)$$

*Moreover, (6.14) and (6.15) also hold with $\mathcal{V}_n^x$ replaced by $\mathcal{V}_n^y$.*

*Proof.* As usual we present the argument for $\mathcal{V}_n^x$ only, since the result for $\mathcal{V}_n^y$ follows in the same manner. First consider the $\alpha > 1$ case. We have the distributional equalities

$$\mathcal{L}(D^\alpha(\mathcal{U}_n^{x,0}) \mid N_n^x = m) = \mathcal{L}(D^\alpha(\mathcal{U}_m^0)), \qquad \mathcal{L}(D^\alpha(\mathcal{U}_n^x) \mid N_n^x = m) = \mathcal{L}(D^\alpha(\mathcal{U}_m)).$$

However, $N_n^x$ is Poisson with mean $n^{1-\sigma}$ and, so, tends to $\infty$ almost surely. Thus, by Theorem 3.1(ii),

$$D^\alpha(\mathcal{U}_n^{x,0}) \overset{D}{\to} D_\alpha \quad \text{and} \quad D^\alpha(\mathcal{U}_n^x) \overset{D}{\to} F_\alpha \quad \text{as } n \to \infty,$$

and by Lemma 6.1 and Slutsky's theorem,

$$\mathcal{L}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\}) \overset{D}{\to} D_\alpha \quad \text{and} \quad \mathcal{L}^\alpha(\mathcal{V}_n^x) \overset{D}{\to} F_\alpha \quad \text{as } n \to \infty. \tag{6.16}$$

Also, $\mathrm{E}[D^\alpha(\mathcal{U}_n^{x,0})] \to (\alpha - 1)^{-1}$ by (3.10), whence, by Lemma 6.1 and Proposition 3.3,

$$\mathrm{E}[\mathcal{L}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\})] \to (\alpha - 1)^{-1} = \mathrm{E}[D_\alpha].$$

Similarly, by (3.18), Lemma 6.1, and Proposition 3.4,

$$\mathrm{E}[\mathcal{L}^\alpha(\mathcal{V}_n^x)] \to (\alpha(\alpha - 1))^{-1} = \mathrm{E}[F_\alpha].$$

Hence, (6.16) also holds with the centred variables (i.e. (6.15) holds).

Now suppose that $\alpha = 1$. Since $N_n^x$ is Poisson with parameter $n^{1-\sigma}$, Lemma 3.7(i) with $t = n^{1-\sigma}$ shows that $\tilde{D}^1(\mathcal{U}_n^{x,0}) \overset{D}{\to} \tilde{D}_1$ as $n \to \infty$. Slutsky's theorem with Lemma 6.1 then implies that $\tilde{\mathcal{L}}^1(\mathcal{V}_n^x \cup \{\mathbf{0}\}) \overset{D}{\to} \tilde{D}_1$. In the same way we obtain $\tilde{\mathcal{L}}^1(\mathcal{V}_n^x) \overset{D}{\to} \tilde{D}_1$, this time using part (ii) of Lemma 3.7 instead of part (i), along with Proposition 3.7.

Note that $D^\alpha(\mathcal{U}_n^x)$ and $D^\alpha(\mathcal{U}_n^y)$ are not independent. To deal with this, we define

$$\tilde{\mathcal{V}}_n^x := \mathcal{P}_n \cap B_x^n \quad \text{and} \quad \tilde{\mathcal{V}}_n^y := \mathcal{P}_n \cap B_y^n.$$

(Also, recall the definition of $\mathcal{V}_n^0$ in (6.7).) Let $\tilde{N}_n^x := \mathrm{card}(\tilde{\mathcal{V}}_n^x)$ and $\tilde{N}_n^y := \mathrm{card}(\tilde{\mathcal{V}}_n^y)$. Since $B_x^n$ and $B_n^y$ are disjoint, $\mathcal{L}^\alpha(\tilde{\mathcal{V}}_n^x)$ and $\mathcal{L}^\alpha(\tilde{\mathcal{V}}_n^y)$ are independent, by the spatial independence property of the Poisson process $\mathcal{P}_n$.

Now we make the following observation. Using notation from Section 4, for $k \in \mathbb{N}$ and for $a < b$ and $c < d$, let $\mathcal{U}_{k,(a,b]\times(c,d]}$ denote the point process consisting of $k$ independent random vectors uniformly distributed on the rectangle $(a, b] \times (c, d]$. Before proceeding further, we recall that if $M(\mathcal{X})$ denotes the number of minimal elements of a point set $\mathcal{X} \subset \mathbb{R}^2$ (under the ordering '$\preceq^*$'), then

$$\mathrm{E}[M(\mathcal{U}_{k,(a,b]\times(c,d]})] = \mathrm{E}[M(\mathcal{X}_k)] = 1 + \frac{1}{2} + \cdots + \frac{1}{k} \le 1 + \log k. \tag{6.17}$$

The first equality in (6.17) comes from some obvious scaling showing that the distribution of $M(\mathcal{U}_{k,(a,b]\times(c,d]})$ does not depend on $a$, $b$, $c$, and $d$. For the second equality in (6.17), see [2] or the proof of Theorem 1.1(a) of [5].

**Lemma 6.3.** *Suppose that $\alpha > 0$.*

(i) *As $n \to \infty$,*

$$\mathcal{L}^{\alpha}(\mathcal{V}_n^x) - \mathcal{L}^{\alpha}(\tilde{\mathcal{V}}_n^x) \to_{L^1} 0 \quad \text{and} \quad \mathcal{L}^{\alpha}(\mathcal{V}_n^y) - \mathcal{L}^{\alpha}(\tilde{\mathcal{V}}_n^y) \to_{L^1} 0, \tag{6.18}$$

$$\mathcal{L}^{\alpha}(\mathcal{V}_n^x \cup \{\mathbf{0}\}) - \mathcal{L}^{\alpha}(\tilde{\mathcal{V}}_n^x \cup \{\mathbf{0}\}) \to_{L^1} 0 \quad \text{and} \quad \mathcal{L}^{\alpha}(\mathcal{V}_n^y \cup \{\mathbf{0}\}) - \mathcal{L}^{\alpha}(\tilde{\mathcal{V}}_n^y \cup \{\mathbf{0}\}) \to_{L^1} 0. \tag{6.19}$$

(ii) *As $n \to \infty$, we have $\mathcal{L}^{\alpha}(\mathcal{V}_n^0) \to_{L^1} 0$ and $\mathcal{L}^{\alpha}(\mathcal{V}_n^0 \cup \{\mathbf{0}\}) \to_{L^1} 0$.*

*Proof.* We first prove (i). We give only the argument for $\mathcal{V}_n^x$, as that for $\mathcal{V}_n^y$ is analogous. Let $\Delta := \mathcal{L}^{\alpha}(\mathcal{V}_n^x) - \mathcal{L}^{\alpha}(\tilde{\mathcal{V}}_n^x)$ and $\beta = (\sigma + \frac{1}{2})/2$. Then $\frac{1}{2} < \beta < \sigma$.

Assume without loss of generality that $\mathcal{P}_n$ is the restriction to $(0, 1]^2$ of a homogeneous Poisson process, $\mathcal{H}_n$, of intensity $n$ on $\mathbb{R}^2$. Let $X^- = (X^-, Y^-)$ be the point of $\mathcal{H}_n \cap ((0, n^{-\beta}] \times (0, \infty))$ with minimal $y$-coordinate. Then $X^-$ is uniform on $(0, n^{-\beta}]$. Let $E_n$ be the event that $X^- > 3n^{-\sigma}$; then $P[E_n^c] = 3n^{\beta - \sigma}$ for large enough $n$.

Let $\Delta_1$ be the contribution to $\Delta$ from edges starting at points in $(0, n^{-\beta}] \times (0, n^{-\sigma}]$. The absolute value of $\Delta_1$ is then bounded by the product of $(\sqrt{2}n^{-\beta})^{\alpha}$ and the number of points of $\mathcal{P}_n$ in $(0, n^{-\beta}] \times (0, n^{-\sigma}]$. Hence, for any $\alpha > 0$,

$$\begin{aligned}
E[|\Delta_1|] &\leq (\sqrt{2}n^{-\beta})^{\alpha} \, E[\text{card}(\mathcal{P}_n \cap ((0, n^{-\beta}] \times (0, n^{-\sigma}]))] \\
&= 2^{\alpha/2} n^{1-\beta-\sigma-\alpha\beta} \\
&\to 0. \tag{6.20}
\end{aligned}$$

Let $\Delta_2 := \Delta - \Delta_1$, the contribution to $\Delta$ from edges starting at points in $(n^{-\beta}, 1] \times (0, n^{-\sigma}]$. Then, by the triangle inequality, if $E_n$ occurs these edges are unaffected by points in $B_n^0$ and, so, $\Delta_2 = 0$. Also, only minimal elements of $\mathcal{P}_n \cap (n^{-\beta}, 1] \times (0, n^{-\sigma}]$ can possibly have their directed nearest neighbour in $(0, n^{-\sigma}] \times (0, n^{-\sigma}]$; hence, if $M_n$ denotes the number of such minimal elements, then $|\Delta_2|$ is bounded by $2^{\alpha/2} M_n$. Hence, using (6.17), we obtain

$$E[|\Delta_2|] \leq 2^{\alpha/2} P[E_n^c] E[M_n] = O(n^{\beta-\sigma} \log n),$$

which tends to 0. Combined with (6.20), this gives us (6.18). The same argument gives us (6.19).

To prove part (ii), note that

$$E[\mathcal{L}^{\alpha}(\mathcal{V}_n^0)] \leq (\sqrt{2}n^{-\sigma})^{\alpha} \, E[N_n^0] = 2^{\alpha/2} n^{1-2\sigma-\sigma\alpha} \to 0 \quad \text{as } n \to \infty,$$

for any $\alpha > 0$. Thus, $\mathcal{L}^{\alpha}(\mathcal{V}_n^0) \to_{L^1} 0$ and, similarly, $\mathcal{L}^{\alpha}(\mathcal{V}_n^0 \cup \{\mathbf{0}\}) \to_{L^1} 0$.

In proving our next lemma (and again later on) we use the following elementary fact. If $N(n)$ is Poisson-distributed with parameter $n$, then as $n \to \infty$,

$$E[|N(n) - n| \log \max(N(n), n)] = O(n^{1/2} \log n). \tag{6.21}$$

To see this, let $Y_n := |N(n) - n| \log \max(N(n), n)$. Then

$$Y_n 1\{N(n) \leq 2n\} \leq |N(n) - n| \log(2n),$$

and the expectation of this is $O(n^{1/2} \log n)$ by Jensen's inequality, since $\text{var}(N(n)) = n$. As the Cauchy–Schwarz inequality shows that $E[Y_n 1\{N(n) > 2n\}] \to O(n^{1/2})$, (6.21) thus follows.

We now state a lemma for coupling $\mathcal{X}_n$ and $\mathcal{P}_n$. The result for $\alpha \geq 1$ will be used in the proof of Theorem 6.1. The result for $0 < \alpha < 1$ will be needed later, in the proof of Theorem 2.1. As in Section 5, let $S_{0,n}$ denote the 'inner' region $(n^{\varepsilon-1/2}, 1]^2$, with $\varepsilon \in (0, \frac{1}{2})$ a constant. The boundary region $B_n$ is disjoint from $S_{0,n}$; let $C_n$ denote the intermediate region $(0, 1]^2 \setminus (B_n \cup S_{0,n})$, whence $B_n \cup C_n = (0, 1]^2 \setminus S_{0,n}$.

**Lemma 6.4.** *There exists a coupling of $\mathcal{X}_n$ and $\mathcal{P}_n$ such that the following statements hold.*

(i) *For $0 < \alpha < 1$, provided that $\varepsilon < (1 - \alpha)/2$, as $n \to \infty$ we have*

$$n^{(\alpha-1)/2} \, \mathrm{E}[|\mathcal{L}^\alpha(\mathcal{X}_n; B_n \cup C_n) - \mathcal{L}^\alpha(\mathcal{P}_n; B_n \cup C_n)|] \to 0 \tag{6.22}$$

*and*

$$n^{(\alpha-1)/2} \, \mathrm{E}[|\mathcal{L}^\alpha(\mathcal{X}_n^0; B_n \cup C_n) - \mathcal{L}^\alpha(\mathcal{P}_n^0; B_n \cup C_n)|] \to 0. \tag{6.23}$$

(ii) *For $\alpha \geq 1$, as $n \to \infty$ we have*

$$\mathrm{E}[|\mathcal{L}^\alpha(\mathcal{X}_n; B_n) - \mathcal{L}^\alpha(\mathcal{P}_n; B_n)|] \to 0 \tag{6.24}$$

*and*

$$\mathrm{E}[|\mathcal{L}^\alpha(\mathcal{X}_n^0; B_n) - \mathcal{L}^\alpha(\mathcal{P}_n^0; B_n)|] \to 0. \tag{6.25}$$

*Proof.* We couple $\mathcal{X}_n$ and $\mathcal{P}_n$ in the following standard way. Let $X_1, X_2, X_3, \ldots$ be independent, uniform random vectors in $(0, 1]^2$, and let $N(n) \sim \mathrm{Po}(n)$ be independent of $(X_1, X_2, \ldots)$. For $m \in \mathbb{N}$ (in particular for $m = n$), set $\mathcal{X}_m := \{X_1, \ldots, X_m\}$, and set $\mathcal{P}_n := \{X_1, \ldots, X_{N(n)}\}$.

For each $m \in \mathbb{N}$, let $Y_m$ denote the in-degree of vertex $X_m$ in the MDST on $\mathcal{X}_m$. Suppose that $X_m = x$. An upper bound for $Y_m$ is then provided by the number of minimal elements of the restriction of $\mathcal{X}_{m-1}$ to the rectangle $\{y \in (0, 1]^2 : x \preccurlyeq^* y\}$. Hence, conditional on $X_m = x$ and on there being $k$ points of $\mathcal{X}_{m-1}$ in this rectangle, the expected value of $Y_m$ is bounded by the expected number of minimal elements in a random uniform sample of $k$ points in this rectangle, and, hence (see (6.17)), by $1 + \log k$. Given the value of $X_m$, the conditional expectation of $Y_m$ is thus bounded by $1 + \log m$.

We first prove the statements in part (i), with $0 < \alpha < 1$. Suppose that $\varepsilon < (1 - \alpha)/2$. Then

$$|\mathcal{L}^\alpha(\mathcal{X}_m; B_n \cup C_n) - \mathcal{L}^\alpha(\mathcal{X}_{m-1}; B_n \cup C_n)| \leq 2^{\alpha/2}(Y_m + 1)\mathbf{1}\{X_m \in B_n \cup C_n\}. \tag{6.26}$$

Since $B_n \cup C_n$ has area $2n^{\varepsilon-1/2} - n^{2\varepsilon-1}$, we obtain

$$\mathrm{E}[(Y_m + 1)\mathbf{1}\{X_m \in B_n \cup C_n\}] \leq (2 + \log m)2n^{\varepsilon-1/2}.$$

Hence, by (6.26) there exists a constant, $C$, such that

$$n^{(\alpha-1)/2} \, \mathrm{E}[|\mathcal{L}^\alpha(\mathcal{P}_n; B_n \cup C_n) - \mathcal{L}^\alpha(\mathcal{X}_n; B_n \cup C_n)| \mid N(n)]$$
$$\leq C|N(n) - n| \log(\max(N(n), n))n^{(\alpha+2\varepsilon-2)/2},$$

and, since we assume that $\alpha + 2\varepsilon < 1$, by (6.21) the expected value of the right-hand side tends to 0 as $n \to \infty$, and we obtain (6.22). A similar argument yields (6.23) in the rooted case.

Now we prove part (ii). For $\alpha \geq 1$, we have

$$|\mathcal{L}^\alpha(\mathcal{X}_m; B_n) - \mathcal{L}^\alpha(\mathcal{X}_{m-1}; B_n)| \leq 2^{\alpha/2}(Y_m + 1)\mathbf{1}\{X_m \in B_n\}. \tag{6.27}$$

Since $B_n$ has area $2n^{-\sigma} - n^{-2\sigma}$, by (6.27) there is a constant, $C$, such that

$$\mathrm{E}[|\mathcal{L}^\alpha(\mathcal{P}_n; B_n) - \mathcal{L}^\alpha(\mathcal{X}_n; B_n)| \mid N(n)] \leq C|N(n) - n|\log(\max(N(n), n))n^{-\sigma},$$

and, since $\sigma > \frac{1}{2}$, by (6.21) the expected value of the right-hand side tends to 0 as $n \to \infty$, and we obtain (6.24). We obtain (6.25) similarly.

*Proof of Theorem 6.1.* Suppose that $\alpha \geq 1$. Then

$$\tilde{\mathcal{L}}^\alpha(\tilde{\mathcal{V}}_n^x) = \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x) + (\tilde{\mathcal{L}}^\alpha(\tilde{\mathcal{V}}_n^x) - \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x)).$$

The final bracket converges to 0 in probability, by Lemma 6.3(i). Thus, by Lemma 6.2 and Slutsky's theorem, we obtain $\tilde{\mathcal{L}}^\alpha(\tilde{\mathcal{V}}_n^x) \xrightarrow{\mathrm{D}} \tilde{F}_\alpha$ (where $\tilde{F}_1 \stackrel{\mathrm{D}}{=} \tilde{D}_1$). Now,

$$\tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x) + \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^y) = \tilde{\mathcal{L}}^\alpha(\tilde{\mathcal{V}}_n^x) + \tilde{\mathcal{L}}^\alpha(\tilde{\mathcal{V}}_n^y) + (\tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x) - \tilde{\mathcal{L}}^\alpha(\tilde{\mathcal{V}}_n^x)) + (\tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^y) - \tilde{\mathcal{L}}^\alpha(\tilde{\mathcal{V}}_n^y)).$$

The last two brackets converge to 0 in probability, by Lemma 6.3(i). Then the independence of $\tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x)$ and $\tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^y)$ and another application of Slutsky's theorem yield

$$\tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x) + \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^y) \xrightarrow{\mathrm{D}} \tilde{F}_\alpha^{\{1\}} + \tilde{F}_\alpha^{\{2\}},$$

where $\tilde{F}_\alpha^{\{1\}}$ and $\tilde{F}_\alpha^{\{2\}}$ are independent copies of $\tilde{F}_\alpha$. Similarly,

$$\tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\}) + \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^y \cup \{\mathbf{0}\}) \xrightarrow{\mathrm{D}} \tilde{D}_\alpha^{\{1\}} + \tilde{D}_\alpha^{\{2\}}.$$

Finally, since $\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; B_n) = \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^x) + \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^y) - \tilde{\mathcal{L}}^\alpha(\mathcal{V}_n^0)$ (with a similar statement including the origin), Lemma 6.3(ii) and Slutsky's theorem complete the proof of (6.1) and (6.3).

To deduce (6.2) and (6.4), assume without loss of generality that $\mathcal{X}_n$ and $\mathcal{P}_n$ are coupled in the manner of Lemma 6.4. Then $\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; B_n) - \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n; B_n)$ tends to 0 in probability by (6.24), and $\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n^0; B_n) - \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n^0; B_n)$ tends to 0 in probability by (6.25). Hence, by Slutsky's theorem, the convergence results (6.1) and (6.3) carry through to the binomial point process case, i.e. (6.2) and (6.4) hold.

Now suppose that $0 < \alpha < 1$. Then (6.10) gives us

$$\mathrm{E}[|n^{(\alpha-1)/2}(\mathcal{L}^\alpha(\mathcal{V}_n^x) - D^\alpha(\mathcal{U}_n^x))|^2] = O(n^{(\alpha+1)(1-2\sigma)}), \tag{6.28}$$

which tends to 0 as $n \to \infty$, since $\sigma > \frac{1}{2}$. Likewise, for the rooted case,

$$\mathrm{E}[|n^{(\alpha-1)/2}(\mathcal{L}^\alpha(\mathcal{V}_n^x \cup \{\mathbf{0}\}) - D^\alpha(\mathcal{U}_n^{x,0}))|^2] = O(n^{(\alpha+1)(1-2\sigma)}). \tag{6.29}$$

By Proposition 3.2, we have

$$\mathrm{E}[n^{(\alpha-1)/2}D^\alpha(\mathcal{U}_n^x)] = O(n^{(\alpha-1)/2}\mathrm{E}[(N_n^x)^{1-\alpha}]) = O(n^{(\alpha-1)(\sigma-1/2)}) \to 0,$$

and combined with (6.28) this completes the proof of (6.5). Similarly, by Proposition 3.1,

$$\mathrm{E}[n^{(\alpha-1)/2}D^\alpha(\mathcal{U}_n^{x,0})] = O(n^{(\alpha-1)/2}\mathrm{E}[(N_n^x)^{1-\alpha}]) = O(n^{(\alpha-1)(\sigma-1/2)}) \to 0,$$

and combined with (6.29) this gives us (6.6).

## 7. Proof of Theorem 2.1

Let $\sigma \in (\frac{1}{2}, \frac{2}{3})$, and let $\varepsilon > 0$ with

$$\varepsilon < \min\left(\frac{1}{2}, \frac{1-\sigma}{3}, \frac{3-4\sigma}{10}, \frac{2-3\sigma}{8}\right). \tag{7.1}$$

In addition, if $0 < \alpha < 1$ we impose the further condition that $\varepsilon < (1-\alpha)/2$. As in Section 5, denote by $S_{0,n}$ the region $(n^{\varepsilon-1/2}, 1]^2$. As in Section 6, let $B_n$ denote the region $(0, 1]^2 \setminus (n^{-\sigma}, 1]^2$, and let $C_n$ denote the region $(0, 1]^2 \setminus (B_n \cup S_{0,n})$.

We know from Sections 5 and 6 that, for large $n$, the weight of edges starting in $S_{0,n}$ satisfies a central limit theorem, and the weight of edges starting in $B_n$ can be approximated by the directed linear forest. We shall show in Lemmas 7.2 and 7.3 that (with a suitable scaling factor for $\alpha < 1$) the contribution to the total weight from points in $C_n$ has a variance converging to 0. To complete the proof of Theorem 2.1 in the Poisson case, we shall show that the lengths from $B_n$ and $S_{0,n}$ are asymptotically independent by virtue of the fact that the configuration of points in $C_n$ is (with probability approaching 1) sufficient to ensure that the configuration of points in $B_n$ has no effect on the edges from points in $S_{0,n}$. To extend the result to the binomial point process case, we shall use a de-Poissonization argument related to that used in [16].

First consider the region $C_n$. We divide this naturally into three regions. Let

$$C_n^x := (n^{\varepsilon-1/2}, 1] \times (n^{-\sigma}, n^{\varepsilon-1/2}],$$
$$C_n^y := (n^{-\sigma}, n^{\varepsilon-1/2}] \times (n^{\varepsilon-1/2}, 1],$$
$$C_n^0 := (n^{-\sigma}, n^{\varepsilon-1/2}]^2.$$

Also, as in Section 6, let

$$B_n^x := (n^{-\sigma}, 1] \times (0, n^{-\sigma}], \qquad B_n^y := (0, n^{-\sigma}] \times (n^{-\sigma}, 1], \qquad B_n^0 := (0, n^{-\sigma}]^2.$$

We divide the $C_n$ and $B_n$ into rectangular cells as follows (see Figure 4), leaving $C_n^0$ undivided. We set

$$k_n := \lfloor n^{1-\sigma-2\varepsilon} \rfloor \tag{7.2}$$

and divide $C_n^x$ lengthways into $k_n$ cells. For each cell,

$$\text{width} = (1 - n^{\varepsilon-1/2})/k_n \sim n^{2\varepsilon+\sigma-1}, \qquad \text{height} = n^{\varepsilon-1/2} - n^{-\sigma} \sim n^{\varepsilon-1/2}. \tag{7.3}$$

Label these cells $\Gamma_i^x$, $i = 1, 2, \ldots, k_n$, from left to right. For each cell $\Gamma_i^x$, define the adjoining cell of $B_n^x$, formed by extending the vertical edges of $\Gamma_i^x$, to be $\beta_i^x$. Each cell $\beta_i^x$ then has width $(1 - n^{\varepsilon-1/2})/k_n \sim n^{2\varepsilon+\sigma-1}$ and height $n^{-\sigma}$.

In a similar way, we divide $C_n^y$ into $k_n$ cells, $\Gamma_i^y$, of height $(1 - n^{\varepsilon-1/2})/k_n$ and width $n^{\varepsilon-1/2} - n^{-\sigma}$, and divide $B_n^y$ into the corresponding cells $\beta_i^y$, $i = 1, \ldots, k_n$.

For $i = 2, \ldots, k_n$, let $E_{x,i}$ denote the event that the cell $\beta_{i-1}^x$ contains at least one point of $\mathcal{P}_n$, and let $E_{y,i}$ denote the event that $\beta_{i-1}^y$ contains at least one point of $\mathcal{P}_n$.

**Lemma 7.1.** *For sufficiently large $n$, and for $1 \le j < i \le k_n$ with $i - j > 3$, if $E_{x,i}$ occurs then no point in the cell $\Gamma_i^x$ has a directed nearest neighbour in the cell $\Gamma_j^x$ or $\beta_j^x$. Similarly, under the same conditions, if $E_{y,i}$ occurs then no point in the cell $\Gamma_i^y$ has a directed nearest neighbour in the cell $\Gamma_j^y$ or $\beta_j^y$.*
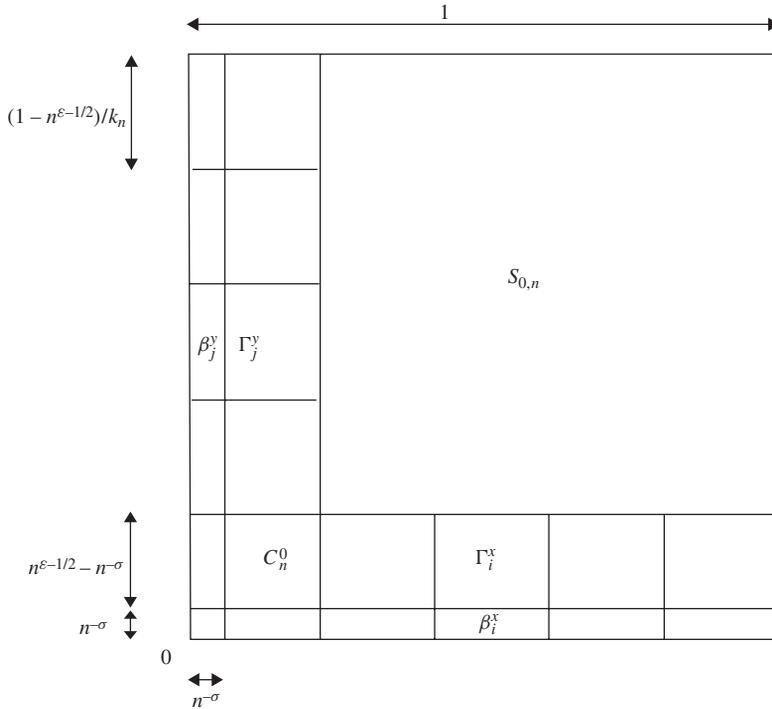
FIGURE 4: The regions of $[0, 1]^2$.

*Proof.* Consider a point $X$, say, in cell $\Gamma_i^x$ in $C_n^x$. Given $E_{x,i}$, we know that there is a point $Y$, say, in the cell, $\beta_{i-1}^x$, to the left of the cell, $\beta_i^x$, immediately below $\Gamma_i^x$, such that $Y \preccurlyeq^* X$ yet the difference in $x$-coordinates between $X$ and $Y$ is no more than twice the width of a cell. Thus, by the triangle inequality, we have

$$\|X - Y\| \leq 2(1 - n^{\varepsilon-1/2})/k_n + n^{\varepsilon-1/2} \sim 2n^{2\varepsilon+\sigma-1}, \qquad (7.4)$$

since $\sigma > \frac{1}{2}$. Now consider a point $Z$, say, in a cell $\Gamma_j^x$ or $\beta_j^x$ with $j \leq i - 4$. In this case, the difference in $x$-coordinates between $X$ and $Z$ is at least the width of three cells, meaning that

$$\|X - Z\| \geq 3(1 - n^{\varepsilon-1/2})/k_n \sim 3n^{2\varepsilon+\sigma-1}. \qquad (7.5)$$

By comparing (7.4) and (7.5), we see that $X$ is not connected to $Z$, which completes the proof.

Recall from (5.1) that for a point set $\mathscr{S} \subset \mathbb{R}^2$ and a region $R \subseteq \mathbb{R}^2$, $\mathcal{L}^\alpha(\mathscr{S}; R)$ denotes the total weight of the edges of the MDSF on $\mathscr{S}$ which originate in the region $R$.

**Lemma 7.2.** *As $n \to \infty$, we have*

$$\mathrm{var}(\mathcal{L}^\alpha(\mathcal{P}_n; C_n)) \to 0 \quad and \quad \mathrm{var}(\mathcal{L}^\alpha(\mathcal{P}_n^0; C_n)) \to 0, \qquad \alpha \geq 1, \qquad (7.6)$$

$$\mathrm{var}(n^{(\alpha-1)/2}\mathcal{L}^\alpha(\mathcal{P}_n; C_n)) \to 0, \qquad 0 < \alpha < 1, \qquad (7.7)$$

$$\mathrm{var}(n^{(\alpha-1)/2}\mathcal{L}^\alpha(\mathcal{P}_n^0; C_n)) \to 0, \qquad 0 < \alpha < 1. \qquad (7.8)$$

*Proof.* For ease of notation, we write $X_i = \mathcal{L}^\alpha(\mathcal{P}_n; \Gamma_i^x)$ and $Y_i = \mathcal{L}^\alpha(\mathcal{P}_n; \Gamma_i^y)$, for $i = 1, 2, \ldots, k_n$. Also, let $Z = \mathcal{L}^\alpha(\mathcal{P}_n; C_n^0)$. Then

$$\text{var}(\mathcal{L}^\alpha(\mathcal{P}_n; C_n)) = \text{var}\left(Z + \sum_{i=1}^{k_n} X_i + \sum_{i=1}^{k_n} Y_i\right). \tag{7.9}$$

Let $N_i^x$, $N_i^y$, and $N_0$ respectively denote the number of points of $\mathcal{P}_n$ in $\Gamma_i^x$, $\Gamma_i^y$, and $C_n^0$. Then, by (7.3), $N_i^x$ is Poisson-distributed with parameter asymptotic to $n^{3\varepsilon+\sigma-1/2}$, while $N_1^x + N_1^y + N_0$ is Poisson-distributed with parameter asymptotic to $2n^{3\varepsilon+\sigma-1/2}$; hence, as $n \to \infty$ we have

$$\text{E}[(N_i^x)^2] \sim n^{6\varepsilon+2\sigma-1}, \qquad \text{E}[(N_1^x + N_1^y + N_0)^2] \sim 4n^{6\varepsilon+2\sigma-1}. \tag{7.10}$$

Edges from points in $\Gamma_1^x \cup \Gamma_1^y \cup C_n^0$ are of length at most $2n^{2\varepsilon+\sigma-1}$, and, hence,

$$\text{var}(X_1 + Y_1 + Z) \leq (2n^{2\varepsilon+\sigma-1})^{2\alpha} \, \text{E}[(N_1^x + N_1^y + N_0)^2]$$
$$\sim 2^{2+2\alpha} n^{6\varepsilon+2\sigma-1+2\alpha(2\varepsilon+\sigma-1)}. \tag{7.11}$$

For $\alpha \geq 1$, since $\varepsilon$ is small (see (7.1)), the expression (7.11) is $O(n^{10\varepsilon+4\sigma-3})$ and in fact tends to 0, implying that

$$\text{var}(X_1 + Y_1 + Z) \to 0, \qquad \alpha \geq 1. \tag{7.12}$$

By Lemma 7.1 and (7.4), given $E_{x,i}$, an edge from a point of $\Gamma_i^x$ can be of length no more than $3n^{2\varepsilon+\sigma-1}$. Thus, using (7.10), we have

$$\text{var}(X_i 1\{E_{x,i}\}) \leq \text{E}[X_i^2 1\{E_{x,i}\}] \leq (3n^{2\varepsilon+\sigma-1})^{2\alpha} \, \text{E}[(N_i^x)^2]$$
$$= O(n^{6\varepsilon+2\sigma-1+2\alpha(2\varepsilon+\sigma-1)}). \tag{7.13}$$

Next observe that $\text{cov}(X_i 1\{E_{x,i}\}, X_j 1\{E_{x,j}\}) = 0$ for $i - j > 3$, since, by Lemma 7.1, $X_i 1\{E_{x,i}\}$ is determined by the restriction of $\mathcal{P}_n$ to the union of the regions $\Gamma_\ell^x \cup \beta_\ell^x$, $i - 3 \leq \ell \leq i$. Thus, by (7.2), the Cauchy–Schwarz inequality, and (7.13), we obtain

$$\text{var}\left(\sum_{i=2}^{k_n} X_i 1\{E_{x,i}\}\right) = \sum_{i=2}^{k_n} \text{var}(X_i 1\{E_{x,i}\}) + \sum_{i=2}^{k_n} \sum_{\{j:\, 1\leq|j-i|\leq 3\}} \text{cov}(X_i 1\{E_{x,i}\}, X_j 1\{E_{x,j}\})$$
$$= O(n^{4\varepsilon+\sigma+2\alpha(2\varepsilon+\sigma-1)}). \tag{7.14}$$

For $\alpha \geq 1$, the bound in (7.14) tends to 0 as $n \to \infty$, since $\frac{1}{2} < \sigma < \frac{2}{3}$ and $\varepsilon$ is small (see (7.1)).

By (7.2), each cell $\beta_i^x$, $i = 1, \ldots, k_n$, has width asymptotic to $n^{2\varepsilon+\sigma-1}$ and height $n^{-\sigma}$, implying that the mean number of points of $\mathcal{P}_n$ in one of these cells is asymptotic to $n^{2\varepsilon}$; hence, for any cell $\beta_i^x$ or $\beta_i^y$, $i = 1, \ldots, k_n$, the probability that the cell contains no point of $\mathcal{P}_n$ is given by $\exp(-n^{2\varepsilon}(1 + o(1)))$. Hence, for large enough $n$, and $i = 2, \ldots, k_n$, we have $\text{P}[E_{x,i}^c] \leq \exp(-n^\varepsilon)$ and, thus, by (7.10),

$$\text{var}(X_i 1\{E_{x,i}^c\}) \leq \text{E}[X_i^2 \mid E_{x,i}^c] \, \text{P}[E_{x,i}^c] \leq 2^\alpha \, \text{E}[(N_i^x)^2] \, \text{P}[E_{x,i}^c]$$
$$= O(n^{6\varepsilon+2\sigma-1} \exp(-n^\varepsilon)).$$

Hence, by the Cauchy–Schwarz inequality, we have

$$\text{var}\left(\sum_{i=2}^{k_n} X_i 1\{E^c_{x,i}\}\right) = \sum_{i=2}^{k_n} \text{var}(X_i 1\{E^c_{x,i}\}) + \sum_{i \neq j} \text{cov}(X_i 1\{E^c_{x,i}\}, X_j 1\{E^c_{x,j}\})$$

$$= O(k_n^2 n^{6\varepsilon + 2\sigma - 1} \exp(-n^\varepsilon)) \to 0 \quad \text{as } n \to \infty. \tag{7.15}$$

Then, by (7.14), (7.15), and the analogous estimates for $Y_i$, along with the Cauchy–Schwarz inequality, for $\alpha \geq 1$ we obtain

$$\text{var}\left(\sum_{i=2}^{k_n} X_i 1\{E_{x,i}\} + \sum_{i=2}^{k_n} Y_i 1\{E_{y,i}\} + \sum_{i=2}^{k_n} X_i 1\{E^c_{x,i}\} + \sum_{i=2}^{k_n} Y_i 1\{E^c_{y,i}\}\right) \to 0 \quad \text{as } n \to \infty. \tag{7.16}$$

From (7.9) with (7.12), (7.16), and the Cauchy–Schwarz inequality, we obtain the first part of (7.6). The argument for $\mathcal{P}_n^0$ is the same as for $\mathcal{P}_n$, which yields the second part of (7.6).

Now suppose that $0 < \alpha < 1$. We obtain (7.7) and (7.8) in a similar way to (7.6), since (7.11) implies that

$$\text{var}(n^{(\alpha-1)/2}(X_1 + Y_1 + Z)) = O(n^{6\varepsilon + 2\sigma - 2 + \alpha(4\varepsilon + 2\sigma - 1)})$$

and (7.14) implies that

$$\text{var}\left(n^{(\alpha-1)/2} \sum_{i=2}^{k_n} X_i 1\{E_{x,i}\}\right) = O(n^{4\varepsilon + \sigma - 1 + \alpha(4\varepsilon + 2\sigma - 1)}),$$

and both of these bounds tend to 0 when $0 < \alpha < 1$, $\frac{1}{2} < \sigma < \frac{2}{3}$, and $\varepsilon$ is small (see (7.1)).

To prove those parts of Theorem 2.1 that refer to the binomial process $\mathcal{X}_n$, we need further results comparing the processes $\mathcal{X}_n$ and $\mathcal{P}_n$ when they are coupled as in Lemma 6.4.

**Lemma 7.3.** *Suppose that $\alpha \geq 1$. With $\mathcal{X}_n$ and $\mathcal{P}_n$ coupled as in Lemma 6.4, as $n \to \infty$ we have*

$$\mathcal{L}^\alpha(\mathcal{X}_n; C_n) - \mathcal{L}^\alpha(\mathcal{P}_n; C_n) \to_{L^1} 0 \quad \text{and} \quad \mathcal{L}^\alpha(\mathcal{X}_n^0; C_n) - \mathcal{L}^\alpha(\mathcal{P}_n^0; C_n) \to_{L^1} 0. \tag{7.17}$$

*Proof.* Let $\mathcal{P}_n$ and $\mathcal{X}_m$, $m \in \mathbb{N}$, be coupled as described in Lemma 6.4. Given $n$, for $m \in \mathbb{N}$ define the event

$$E_{m,n} := \bigcap_{1 \leq i \leq k_n} (\{\mathcal{X}_{m-1} \cap \beta_i^x \neq \varnothing\} \cap \{\mathcal{X}_{m-1} \cap \beta_i^y \neq \varnothing\}),$$

with the subcells $\beta_i^x$ and $\beta_i^y$ of $B_n$ as defined near the start of Section 7. Then, by arguments similar to those for $P[E^c_{x,i}]$ above, we have

$$P[E^c_{m,n}] = O(n^{1-\sigma-2\varepsilon} \exp(-n^\varepsilon/2)), \qquad m \geq n/2 + 1.$$

As in the proof of Lemma 6.4, let $Y_m$ denote the in-degree of vertex $X_m$ in the MDST on $\mathcal{X}_m$. Then

$$|\mathcal{L}^\alpha(\mathcal{X}_m; C_n) - \mathcal{L}^\alpha(\mathcal{X}_{m-1}; C_n)| \leq (Y_m + 1)1\{X_m \in C_n\}((3n^{2\varepsilon+\sigma-1})^\alpha + 2^{\alpha/2}1\{E^c_{m,n}\}).$$

Thus, given $N(n)$,

$$|\mathcal{L}^\alpha(\mathcal{X}_n; C_n) - \mathcal{L}^\alpha(\mathcal{P}_n; C_n)|$$
$$\leq \sum_{m=\min(N(n),n)}^{\max(N(n),n)} (Y_m + 1)1\{X_m \in C_n\}(3^\alpha n^{\alpha(2\varepsilon+\sigma-1)} + 2^{\alpha/2}1\{E_{m,n}^c\}).$$

Since $C_n$ has area less than $2n^{\varepsilon-1/2}$, by (6.17) there exists a constant, $C$, such that for $n$ sufficiently large and $N(n) \geq n/2 + 1$,

$$\mathrm{E}[|\mathcal{L}^\alpha(\mathcal{X}_n; C_n) - \mathcal{L}^\alpha(\mathcal{P}_n; C_n)| \mid N(n)]$$
$$\leq 2^{\alpha/2}n1\{N(n) < n/2 + 1\}$$
$$+ C|N(n) - n|\log(\max(N(n), n))n^{\alpha(2\varepsilon+\sigma-1)+\varepsilon-1/2}1\{N(n) \geq n/2 + 1\}. \quad (7.18)$$

By tail bounds for the Poisson distribution, we have $n\,\mathrm{P}[N(n) < n/2 + 1] \to 0$ as $n \to \infty$, and, hence, by taking expectations in (7.18) and using (6.21), we obtain

$$\mathrm{E}[|\mathcal{L}^\alpha(\mathcal{X}_n; C_n) - \mathcal{L}^\alpha(\mathcal{P}_n; C_n)|] = O(n^{\alpha(2\varepsilon+\sigma-1)+\varepsilon}\log n) + o(1),$$

which tends to 0 since $\alpha \geq 1$, $\frac{1}{2} < \sigma < \frac{2}{3}$, and $\varepsilon$ is small (see (7.1)). We thus obtain the unrooted part of (7.17). The argument is the same in the rooted case.

**Lemma 7.4.** *Suppose that $\mathcal{X}_n$ and $\mathcal{P}_n$ are coupled as described in Lemma 6.4, with $N(n) := \mathrm{card}(\mathcal{P}_n)$. Let $\Delta(\infty)$ be given by Definition 4.1 with $H = \mathcal{L}^1$, and let $\alpha_1 := \mathrm{E}[\Delta(\infty)]$. Then as $n \to \infty$,*

$$\mathcal{L}^1(\mathcal{P}_n; S_{0,n}) - \mathcal{L}^1(\mathcal{X}_n; S_{0,n}) - n^{-1/2}\alpha_1(N(n) - n) \to_{L^2} 0, \quad (7.19)$$
$$\mathcal{L}^1(\mathcal{P}_n^0; S_{0,n}) - \mathcal{L}^1(\mathcal{X}_n^0; S_{0,n}) - n^{-1/2}\alpha_1(N(n) - n) \to_{L^2} 0. \quad (7.20)$$

We omit the proof of this lemma; see [13] for more details.

We are now in a position to prove Theorem 2.1. We divide the proof into two cases: $\alpha \neq 1$ and $\alpha = 1$. In the latter case, to prove the result for the Poisson process $\mathcal{P}_n$ we need to show that $\mathcal{L}^1(\mathcal{P}_n; B_n)$ and $\mathcal{L}^1(\mathcal{P}_n; S_{0,n})$ are asymptotically independent; likewise for $\mathcal{P}_n^0$. We shall then obtain the results both for the binomial process $\mathcal{X}_n$ and for $\mathcal{X}_n^0$ from those for $\mathcal{P}_n$ and $\mathcal{P}_n^0$, via the coupling described in Lemma 6.4.

*Proof of Theorem 2.1 for $\alpha \neq 1$.* First suppose that $0 < \alpha < 1$. For the Poisson case, we have

$$n^{(\alpha-1)/2}\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n) = n^{(\alpha-1)/2}\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; S_{0,n}) + n^{(\alpha-1)/2}\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; B_n) + n^{(\alpha-1)/2}\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; C_n). \quad (7.21)$$

The first term on the right-hand side of (7.21) converges in distribution to $\mathcal{N}(0, s_\alpha^2)$ by Theorem 5.1(iv), and the other two terms converge in probability to 0 by (6.5) and (7.7). Thus, Slutsky's theorem yields the first (Poisson) part of (2.9). To obtain the second (binomial) part of (2.9), we use the coupling of Lemma 6.4. We write

$$n^{(\alpha-1)/2}\tilde{\mathcal{L}}^\alpha(\mathcal{X}_n) = n^{(\alpha-1)/2}\tilde{\mathcal{L}}^\alpha(\mathcal{X}_n; S_{0,n}) + n^{(\alpha-1)/2}\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; B_n \cup C_n)$$
$$+ n^{(\alpha-1)/2}(\tilde{\mathcal{L}}^\alpha(\mathcal{X}_n; B_n \cup C_n) - \tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; B_n \cup C_n)). \quad (7.22)$$

The first term on the right-hand side of (7.22) is asymptotically $\mathcal{N}(0, t_\alpha^2)$, by Theorem 5.1(ii). The second term tends to 0 in probability, by (6.5) and (7.7). The third term tends to 0 in probability, by (6.22). Thus, we obtain the binomial case of (2.9).

The rooted case, (2.6), is similar. For the first (Poisson) part of (2.6), we use Corollary 5.1(iv) with (6.6) and (7.8), and Slutsky's theorem. The second part of (2.6) follows from the statement analogous to (7.22) with the addition of the origin, using Corollary 5.1(ii) with (6.6), (7.8), and (6.23), and, again, Slutsky's theorem.

Next suppose that $\alpha > 1$. We have

$$\tilde{\mathcal{L}}^\alpha(\mathcal{P}_n) = \tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; S_{0,n}) + \tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; C_n) + \tilde{\mathcal{L}}^\alpha(\mathcal{P}_n; B_n). \tag{7.23}$$

The first term on the right-hand side converges to 0 in probability, by Theorem 5.1(iii). The second term also converges to 0 in probability, by the first part of (7.6). Then, by (6.3) and Slutsky's theorem, we obtain the first (Poisson) part of (2.11). To obtain the rooted version, i.e. the first part of (2.8), we replace $\mathcal{P}_n$ by $\mathcal{P}_n^0$ in (7.23), combine (6.1) with Corollary 5.1(iii) and the second part of (7.6), and apply Slutsky's theorem again.

To obtain the binomial versions of the results (2.8) and (2.11), we again make use of the coupling described in Lemma 6.4. We have

$$\tilde{\mathcal{L}}^\alpha(\mathcal{X}_n) = \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n; S_{0,n}) + \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n; C_n) + \tilde{\mathcal{L}}^\alpha(\mathcal{X}_n; B_n). \tag{7.24}$$

The first term on the right-hand side converges in probability to 0, by Theorem 5.1(i). The second term also converges in probability to 0, by the first part of (7.6) and the first part of (7.17). The third part converges in distribution to $\tilde{F}_\alpha^{\{1\}} + \tilde{F}_\alpha^{\{2\}}$, by (6.4). Hence, Slutsky's theorem yields the binomial part of (2.11).

Similarly, by replacing $\mathcal{P}_n$ by $\mathcal{P}_n^0$ and $\mathcal{X}_n$ by $\mathcal{X}_n^0$ in (7.24), and using Corollary 5.1(i), the second parts of (7.6) and (7.17), (6.2), and Slutsky's theorem, we obtain the binomial part of (2.8). This completes the proof for $\alpha \neq 1$.

*Proof of Theorem 2.1 for $\alpha = 1$: the Poisson case.* We now prove the first part of (2.7) and the first part of (2.10). Given $n$, let $q_n := 4\lfloor n^{\varepsilon+\sigma-1/2}\rfloor$. Split each cell $\Gamma_i^x$ of $C_n^x$ into a grid of $4q_n$ rectangular subcells by splitting the horizontal edge into $q_n$ segments and the vertical edge into four segments. Similarly, split each cell $\Gamma_i^y$ by splitting the vertical edge into $q_n$ segments and the horizontal edge into four segments. Finally, add a single square subcell in the top right-hand corner of $C_n^0$, of side-length $(1/4)n^{\varepsilon-1/2}$, and denote this the 'corner subcell'.

The total number of all such subcells is $1 + 8k_nq_n \sim 32n^{(1/2)-\varepsilon}$. Each of the subcells has width and height asymptotic to $(1/4)n^{\varepsilon-1/2}$, and so the area of each cell is asymptotic to $(1/16)n^{2\varepsilon-1}$. Therefore, for large $n$, for each of these subcells, the probability that it contains no point of $\mathcal{P}_n$ is bounded by $\exp(-n^\varepsilon)$.

Let $E_n$ be the event that each of the subcells described above contains at least one point of $\mathcal{P}_n$. Then

$$\mathrm{P}[E_n^c] = O(n^{(1/2)-\varepsilon}\exp(-n^\varepsilon)) \to 0.$$

Suppose that $x$ lies on the lower boundary of $S_{0,n}$. Consider the rectangular subcell of $\Gamma_i^x$ lying just to the left of the subcell directly below $x$ (or the corner subcell, if it lies just to the left of the subcell directly below $x$). All points $y$ in this subcell satisfy $y \preccurlyeq^* x$ and, for large $n$, satisfy $\|y - x\| < (3/4)n^{\varepsilon-1/2}$, whereas the nearest point to $x$ in $B_n$ is at a distance at least $(3/4)n^{\varepsilon-1/2}$. Arguing similarly for points $x$ on the left-hand boundary of $S_{0,n}$, and using the

triangle inequality, we see that if $E_n$ occurs, no point in $S_{0,n}$ can be connected to any point in $B_n$, provided that $n$ is sufficiently large.

For simplicity of notation, we write $X_n := \tilde{\mathcal{L}}^1(\mathcal{P}_n; B_n)$ and $Y_n := \tilde{\mathcal{L}}^1(\mathcal{P}_n; S_{0,n})$. We also write $X := \tilde{D}_1^{\{1\}} + \tilde{D}_1^{\{2\}}$ and $Y \sim \mathcal{N}(0, s_1^2)$, independent of $X$, with $s_1$ as given in Theorem 5.1. We know from Theorem 6.1 and Theorem 5.1 that $X_n \overset{\mathrm{D}}{\to} X$ and $Y_n \overset{\mathrm{D}}{\to} Y$ as $n \to \infty$.

We need to show that $X_n + Y_n \overset{\mathrm{D}}{\to} X + Y$, where $X$ and $Y$ are independent random variables. We show this by convergence of the characteristic function,

$$\mathrm{E}[\exp(\mathrm{i}t(X_n + Y_n))] \to \mathrm{E}[\exp(\mathrm{i}t X)]\,\mathrm{E}[\exp(\mathrm{i}t Y)]. \tag{7.25}$$

With $\omega$ denoting the configuration of points in $C_n$, we have

$$\mathrm{E}[\exp(\mathrm{i}t(X_n + Y_n))] = \int_{E_n} \mathrm{E}[\mathrm{e}^{\mathrm{i}t X_n}\mathrm{e}^{\mathrm{i}t Y_n} \mid \omega]\,\mathrm{d}P(\omega) + \mathrm{E}[\mathrm{e}^{\mathrm{i}t(X_n+Y_n)}1\{E_n^{\mathrm{c}}\}]$$

$$= \int_{E_n} \mathrm{E}[\mathrm{e}^{\mathrm{i}t X_n}]\,\mathrm{E}[\mathrm{e}^{\mathrm{i}t Y_n} \mid \omega]\,\mathrm{d}P(\omega) + \mathrm{E}[\mathrm{e}^{\mathrm{i}t(X_n+Y_n)}1\{E_n^{\mathrm{c}}\}],$$

where we have used the facts that $X_n$ and $Y_n$ are conditionally independent, given $\omega \in E_n$, for $n$ sufficiently large, and that $X_n$ is independent of the configuration in $C_n$. It follows that $\mathrm{E}[\mathrm{e}^{\mathrm{i}t(X_n+Y_n)}1\{E_n^{\mathrm{c}}\}] \to 0$ as $n \to \infty$, since $\mathrm{P}[E_n^{\mathrm{c}}] \to 0$. Thus,

$$\mathrm{E}[\exp(\mathrm{i}t(X_n + Y_n))] - \mathrm{E}[\mathrm{e}^{\mathrm{i}t X_n}]\,\mathrm{E}[\mathrm{e}^{\mathrm{i}t Y_n}1\{E_n\}] \to 0,$$

and we obtain (7.25) since $\mathrm{E}[\mathrm{e}^{\mathrm{i}t Y_n}1\{E_n\}] = \mathrm{E}[\mathrm{e}^{\mathrm{i}t Y_n}] - \mathrm{E}[\mathrm{e}^{\mathrm{i}t Y_n}1\{E_n^{\mathrm{c}}\}]$, $\mathrm{E}[\mathrm{e}^{\mathrm{i}t Y_n}1\{E_n^{\mathrm{c}}\}] \to 0$, $\mathrm{E}[\mathrm{e}^{\mathrm{i}t X_n}] \to \mathrm{E}[\mathrm{e}^{\mathrm{i}t X}]$, and $\mathrm{E}[\mathrm{e}^{\mathrm{i}t Y_n}] \to \mathrm{E}[\mathrm{e}^{\mathrm{i}t Y}]$, as $n \to \infty$.

We can now prove the first (Poisson) part of (2.10). Consider the $\alpha = 1$ case of (7.23). The contribution from $C_n$ converges in probability to 0, by the first part of (7.6). Slutsky's theorem and (7.25) then give the first (Poisson) part of (2.10). The rooted Poisson case, (2.7), follows from the rooted version of (7.23), this time applying the argument for (7.25) with $X_n := \tilde{\mathcal{L}}^1(\mathcal{P}_n^0; B_n)$, $Y_n := \tilde{\mathcal{L}}^1(\mathcal{P}_n^0; S_{0,n})$, and $X$ and $Y$ as before, and then using the second part of (7.6) and, again, Slutsky's theorem. Thus, we obtain the first (Poisson) part of (2.7).

*Proof of Theorem 2.1 for $\alpha = 1$: the binomial case.* It remains for us to prove the second part of (2.7) and the second part of (2.10). To do this, we use the coupling of Lemma 6.4 once more. Considering first the unrooted case, we here let $X_n := \mathcal{L}^1(\mathcal{X}_n; B_n)$, $Y_n := \mathcal{L}^1(\mathcal{X}_n; S_{0,n})$, $X_n' := \mathcal{L}^1(\mathcal{P}_n; B_n)$, and $Y_n' := \mathcal{L}^1(\mathcal{P}_n; S_{0,n})$ (note that all these random variables are uncentred).

Let $Y \sim \mathcal{N}(0, s_1^2)$ with $s_1$ as given in Theorem 5.1, and let $X := \tilde{D}_1^{\{1\}} + \tilde{D}_1^{\{2\}}$, independent of $Y$. Then, by (7.25), we have (in our new notation)

$$X_n' - \mathrm{E}[X_n'] + Y_n' - \mathrm{E}[Y_n'] \overset{\mathrm{D}}{\to} X + Y. \tag{7.26}$$

By (6.24), we have $X_n - X_n' \overset{\mathrm{P}}{\to} 0$ and $\mathrm{E}[X_n] - \mathrm{E}[X_n'] \to 0$. Also, with $\alpha_1$ as defined in Lemma 7.4, (7.19) yields

$$Y_n' - Y_n - n^{-1/2}\alpha_1(N(n) - n) \to_{L^2} 0, \tag{7.27}$$

whence $\mathrm{E}[Y_n'] - \mathrm{E}[Y_n] \to 0$. By combining these observations with (7.26) and using Slutsky's theorem, we obtain

$$X_n - \mathrm{E}[X_n] + Y_n - \mathrm{E}[Y_n] + n^{-1/2}\alpha_1(N(n) - n) \overset{\mathrm{D}}{\to} X + Y. \tag{7.28}$$

By Theorem 5.1(iii) we have $\mathrm{var}(Y_n') \to s_1^2$ as $n \to \infty$. By (7.27) and the independence of $N(n)$ and $Y_n$, we have

$$s_1^2 = \lim_{n\to\infty} \mathrm{var}(Y_n + n^{-1/2}\alpha_1(N(n) - n)) = \lim_{n\to\infty}(\mathrm{var}(Y_n) + \alpha_1^2),$$

whence $\alpha_1^2 \leq s_1^2$. Also, $n^{-1/2}\alpha_1(N(n) - n)$ is independent of $X_n + Y_n$ and asymptotically $\mathcal{N}(0, \alpha_1^2)$-distributed. Since the characteristic function of $\mathcal{N}(0, s^2)$ is $\exp(-s^2 t^2/2)$, for all $t \in \mathbb{R}$ we find from (7.28) that

$$\mathrm{E}[\exp(\mathrm{i}t(X_n - \mathrm{E}[X_n] + Y_n - \mathrm{E}[Y_n]))] \to \exp\left(-\frac{(s_1^2 - \alpha_1^2)t^2}{2}\right)\mathrm{E}[\exp(\mathrm{i}tX)],$$

whence

$$X_n - \mathrm{E}[X_n] + Y_n - \mathrm{E}[Y_n] \overset{\mathrm{D}}{\to} X + W, \tag{7.29}$$

where $W \sim \mathcal{N}(0, s_1^2 - \alpha_1^2)$ is independent of $X$.

Consider the $\alpha = 1$ case of (7.24). By the first part of (7.6) and the first part of (7.17), the contribution from $C_n$ tends to 0 in probability. Hence, by (7.29) and Slutsky's theorem, we obtain the second (binomial) part of (2.10).

For the rooted case, we apply the argument for (7.29), now with $X_n := \mathcal{L}^1(\mathcal{X}_n^0; B_n)$, $Y_n := \mathcal{L}^1(\mathcal{X}_n^0; S_{0,n})$, and $X$, $Y$, and $W$ as before. The rooted case of (7.26) follows from the rooted case of (7.25), and now we have $X_n - X_n' \overset{\mathrm{P}}{\to} 0$ and $\mathrm{E}[X_n] - \mathrm{E}[X_n'] \to 0$, by (6.25). In the rooted case (7.27) still holds, by (7.20), and we obtain the rooted case of (7.29) as before.

To obtain the second (binomial) part of (2.7), we start with the rooted version of the $\alpha = 1$ case of (7.24). By the second parts of (7.6) and (7.17), the contribution from $C_n$ tends to 0 in probability. Hence, by the rooted version of (7.29) and Slutsky's theorem, we obtain the second part of (2.7).

This completes the proof of the $\alpha = 1$ case, and, hence, the proof of Theorem 2.1.

## Acknowledgements

## References

[1] BAI, Z.-D., LEE, S. AND PENROSE, M. D. (2006). Rooted edges in a minimal directed spanning tree. *Adv. Appl. Prob.* **38,** 1–30.

[2] BARNDORFF-NIELSEN, O. AND SOBEL, M. (1966). On the distribution of the number of admissible points in a vector random sample. *Theory Prob. Appl.* **11,** 249–269.

[3] BERTOIN, J. AND GNEDIN, A. (2004). Asymptotic laws for nonconservative selfsimilar fragmentations. *Electron. J. Prob.* **9,** 575–593.

[4] BERGER, N. *et al.* (2003). Degree distribution of the FKP model. In *Automata, Languages and Programming* (Proc. 30th Internat. Colloquium, ICALP 2003, Eindhoven; Lecture Notes in Comp. Sci. **2719**), eds J. C. M. Baeten, J. K. Lenstra, J. Parrow and G. J. Woeginger, Springer, Berlin, pp. 725–738.

[5] BHATT, A. G. AND ROY, R. (2004). On a random directed spanning tree. *Adv. Appl. Prob.* **36,** 19–42.

[6] HOARE, C. A. R. (1961). Algorithm 64: Quicksort. *Commun. Assoc. Comput. Mach.* **4,** 321.

[7] HWANG, H.-K. (1998). Asymptotics of divide-and-conquer recurrences: Batcher's sorting algorithm and a minimum Euclidean matching heuristic. *Algorithmica* **22,** 529–546.

[8] KESTEN, H. AND LEE, S. (1996). The central limit theorem for weighted minimal spanning trees on random points. *Ann. Appl. Prob.* **6,** 495–527.

[9] KINGMAN, J. F. C. (1993). *Poisson Processes* (Oxford Stud. Prob. **3**). Clarendon Press, Oxford.

[10] NEININGER, R. AND RÜSCHENDORF, L. (2004). A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Prob.* **14,** 378–418.

[11] PENROSE, M. (2003). *Random Geometric Graphs* (Oxford Stud. Prob. **5**). Oxford University Press.

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

[13] PENROSE, M. D. AND WADE, A. R. (2004). On the total length of the random minimal directed spanning tree. Preprint. Available at http://arxiv.org/abs/math.PR/0409201.

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

[15] PENROSE, M. D. AND WADE, A. R. (2006). Limit theory for the random on-line nearest-neighbour graph. Preprint. Available at http://arxiv.org/abs/math.PR/0603561.

[16] PENROSE, M. D. AND YUKICH, J. E. (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Prob.* **11,** 1005–1041.

[17] PENROSE, M. D. AND YUKICH, J. E. (2003). Weak laws of large numbers in geometric probability. *Ann. Appl. Prob.* **13,** 277–303.

[18] RODRIGUEZ-ITURBE, I. AND RINALDO, A. (1997). *Fractal River Basins: Chance and Self-Organization.* Cambridge University Press.

[19] RÖSLER, U. (1992). A fixed point theorem for distributions. *Stoch. Process. Appl.* **42,** 195–214.

[20] RÖSLER, U. AND RÜSCHENDORF, L. (2001). The contraction method for recursive algorithms. *Algorithmica* **29,** 3–33.

[21] SEPPÄLÄINEN, T. AND YUKICH, J. E. (2001). Large deviation principles for Euclidean functionals and other nearly additive processes. *Prob. Theory Relat. Fields* **120,** 309–345.

[22] STEELE, J. M. (1997). *Probability Theory and Combinatorial Optimization.* Society for Industrial and Applied Mathematics, Philadelphia, PA.

[23] YUKICH, J. E. (1998). *Probability Theory of Classical Euclidean Optimization Problems* (Lecture Notes Math. **1675**). Springer, Berlin.