RANDOM MINIMAL DIRECTED SPANNING TREES AND DICKMAN-TYPE DISTRIBUTIONS

MATHEW D. PENROSE, University of Bath
(Postal Address: Department of Mathematical Sciences, University of Bath,
Bath BA2 7AY, England. Email: M.D.Penrose@bath.ac.uk)

ANDREW WADE, University of Durham
(Postal Address: Department of Mathematical Sciences, University of Durham,
South Road, Durham DH1 3LE, England. Email: a.r.wade@durham.ac.uk)

November 2003

Abstract

In Bhatt and Roy's minimal directed spanning tree construction for n random points in the unit square, all edges must be in a southwesterly direction and there must be a directed path from each vertex to the root placed at the origin. We identify the limiting distributions (for large n) for the total length of rooted edges, and also for the maximal length of all edges in the tree. These limit distributions have been seen previously in analysis of the Poisson-Dirichlet distribution and elsewhere; they are expressed in terms of Dickman's function, and their properties are discussed in some detail.

AMS 2003 *Mathematics Subject Classification.* Primary 60D05, 60G70. Secondary 05C80, 60F05.
Key words and phrases. Spanning tree, extreme values, weak convergence, Dickman distribution, Poisson-Dirichlet distribution.

1 Introduction

The probability theory of graphs, generated by randomly placing points in the unit square and connecting nearby points according to some deterministic rule, has recently grown considerably. 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 graphs of this type, which are locally determined in a certain sense, are by now understood. See for example [16, 17, 18, 21, 24].

The *minimal directed spanning tree* (or MDST for short) was introduced by Bhatt and Roy in [3]. In its structure, the MDST resembles both the standard minimal spanning tree and the nearest neighbour graph for point sets in the plane, with the extra twist that all edges must be oriented in a south-westerly direction. This feature gives rise to significant boundary effects and hence to asymptotic properties which are qualitatively different from those for many of the previously considered graphs.

Of interest is the behaviour of the length of the graph, or of various parts of the graph. We consider elsewhere [19] the total length of all edges.

The edges *incident to the origin* were the principal object of analysis in [3], in which Bhatt and Roy established (amongst other things) existence of a weak limit for the total length of such edges, without describing that limit. We use a different method to characterize the limiting distribution as a variant of the *Dickman distribution* which has previously arisen in such fields as probabilistic number theory, population genetics, and the theory of random search trees (see Section 3). We also extend the result to power-weighted edges.

In addition, we derive a weak convergence result for the *maximum* of all edge lengths in the MDST (Bhatt and Roy obtained such a result for maximum length of edges incident to the origin). In this case, the limiting distribution is related to the distribution of the largest component of the Poisson-Dirichlet distribution with parameter 1. The latter distribution has also sometimes been called a 'Dickman distribution' (see [1, 5]) and we shall call it the *max-Dickman* distribution. In Section 3, we shall discuss both types of Dickman distribution in some detail (they are related).

The MDST is defined formally in the next section. Motivation comes from the modelling of communications or drainage networks. The communications model considered in [3] goes as follows. Consider a network of radio masts, each of which can receive signals only from masts to the south-west. Suppose a source transmitter is positioned at the origin of the plane, and a network of masts is positioned in the first quadrant. Then the graph of the transmission network can be viewed as a directed spanning tree. For convenience, the direction of the edges is taken to be from receiver to transmitter, so that all the directed paths eventually meet at the origin. We restrict the model to a single link into each receiver, which we may justify by asserting that once the first connection has been established, further links may be ignored for many purposes. Various characteristics of the resulting graph are then of interest.

The same graph may be considered as a model for drainage networks, following the spirit of Rodriguez-Iturbe and Rinaldo in [20]; again, see [3]. The idea is that water is allowed to run off an inclined bounded field, forming several drainage channels. These channels eventually merge so that all the water flows out of the field at the lowest point on the boundary. Given any particular landscape geometry, this situation is fairly unpleasant to model directly, so we study a model that maintains the essential features of the above system while being much simpler to handle.

In one model proposed in [20], given a fixed number n of points which the stream network (graph) must contain as nodes, the optimal configuration is achieved by minimizing the quantity $\sum_i Q_i^{1/2} L_i$ where L_i is the length and Q_i the discharge of stream (edge) i. If we assume that Q_i is fixed for all i (and so the flows are non-additive), and flow is constrained to be in a south-westerly direction, the optimum configuration on a set of points is given by the construction we consider here. Another viewpoint is to consider the catchment of the network, which will depend on the total length of the channels.

Understanding these networks for large systems may be difficult: by investigating the behaviour of the MDST on random points we hope to shed light on their 'typical' behaviour.

2 Definitions and main results

Suppose V is a finite set endowed with a partial ordering $\preccurlyeq$ (i.e., a reflexive transitive binary relation such that $u \preccurlyeq v$ and $v \preccurlyeq u$ only when $u = v$; see e.g. [13]). The partial ordering induces a directed graph $G = (V, E)$ on V, with vertex set V and with edge set consisting of all ordered pairs (v, u) of distinct elements of V such that $u \preccurlyeq v$. We make the following definitions.

A *minimal element*, or *sink*, is a vertex $v_0 \in V$ for which there exists no $v \in V \setminus \{v_0\}$ such that $v \preccurlyeq v_0$. Observe that any finite partially ordered set will have at least one sink, and V has a single sink if and only if there exists some $v_0 \in V$ such that $v_0 \preccurlyeq v \; \forall \, v \in V$. Let V_0 denote the set of all sinks of V.

A *directed spanning subgraph (DSS)* of G is a subgraph $H = (V_H, E_H)$ of (V, E) such that $V_H = V$ and $E_H \subseteq E$.

A *directed spanning forest (DSF)* T on V is a DSS on V such that 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 where V_0 consists of a single sink, we refer to any DSF on V as a *directed spanning tree (DST)* on V.

It follows from the definitions that if H is a DSF on V, then there is no branching point in H, i.e., there do not exist distinct vertices $u, u', v \in V$ such that (v, u) and (v, u') are both edges of H (for if such a triple existed, the path from v to a sink would not be unique). Hence, if we ignore the orientation of edges then the DSF H is a forest whose components are in one-to-one correspondence with the set of sinks. If there is just one sink, then (ignoring orientation) any DST on V is a tree.

A *weight function* on the edges of a directed graph (V, E) is a function $w : E \to [0, \infty)$

Definition 1 *Suppose V is a partially ordered finite set, and that the induced graph $G = (V, E)$ carries a weight function. A* minimal directed spanning forest (MDSF) *on V (or, equivalently, on G), is a directed spanning forest T on V with edge set $E_T \subseteq E$ such that*

$$w(T) := \sum_{e \in E_T} w(e) = \min \left\{ \sum_{e \in E_{T'}} w(e) : T' = (V, E_{T'}) \text{ a DSF on } V \right\}. \tag{1}$$

If V has a single sink, then any minimal directed spanning forest on V is called a minimal directed spanning tree (MDST) *on V.*

Thus, a MDSF on V is defined as a solution to a global optimization problem. However, the following simple result shows that when all weights are distinct, a MDSF can be constructed in a 'local' manner, reminiscent of Kruskal's greedy algorithm [14] for finding the minimal spanning tree in an undirected graph.

Definition 2 *We say that $u \in V$ is a* directed nearest neighbour *of $v \in V$ if $u \preccurlyeq v$ and $w(v, u) \leq w(v, u')$ for all $u' \in V \setminus \{v\}$ such that $u' \preccurlyeq v$.*

Proposition 1 *Suppose that V is a finite partially ordered set with its set of sinks denoted V_0, and that the induced graph G is endowed with a weight function w. 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). Let $M = (V, E_M)$ be the directed spanning subgraph of V obtained by taking*

$$E_M := \{(v, n_v) : v \in V \setminus V_0\}$$

Then M is a MDSF of V.

Proof. Let T be an arbitrary DSF on G. Then for every $v \in V \setminus V_0$, there exists a unique element of V, denoted u_v, such that $(v, u_v) \in T$ (uniqueness follows from the absence of branching points). Necessarily $u_v \preccurlyeq v$, and by definition of directed nearest neighbours we have

$$w(M) = \sum_{v \in V \setminus V_0} w(v, n_v) \leq \sum_{v \in V \setminus V_0} w(v, u_v) = w(T),$$

for every DSF T. Thus, M is a MDSF of V. □

While the statements above apply to any partially ordered set V with weights defined for all induced edges, in this paper we are are exclusively concerned with the case where V is a randomly generated subset of $\mathbf{R}^2$, and where the partial ordering and weight function are defined as follows.

The partial ordering (for $V \subset \mathbf{R}^2$) is defined coordinate-wise; in other words, for $u = (u_1, u_2)$ and $v = (v_1, v_2)$ in V, we set $u \preccurlyeq v$ if and only if $u_1 \leq v_1$ and $u_2 \leq v_2$.

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

Moreover, we shall assume that $V \subset \mathbf{R}^2$ is given by $V = \mathcal{S} \cup \{\mathbf{0}\}$, where $\mathbf{0}$ is the origin in $\mathbf{R}^2$ 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 process $\mathcal{P}_n$ of intensity n on the unit square $(0, 1]^2$, or a binomial point process $\mathcal{X}_n$ of n uniformly distributed points on $(0, 1]^2$.

In this random setting, with probability one $V_0 = \{\mathbf{0}\}$ and each point of $\mathcal{S}$ has a unique directed nearest neighbour, so that by Proposition 1, V has a unique MDST, which does not depend on the choice of α. We are concerned with the total weight of the edges incident to $\mathbf{0}$ in the MDST on $\mathcal{S} \cup \{\mathbf{0}\}$; denote this length by $\mathcal{L}_0^\alpha(\mathcal{S})$. Then

$$\mathcal{L}_0^\alpha(\mathcal{S}) = \sum_{X \in \mathcal{S}, X \text{ minimal}} \|X\|^\alpha.$$

Our first main result describes the limiting distribution of $\mathcal{L}_0^\alpha(\mathcal{X}_n)$ or $\mathcal{L}_0^\alpha(\mathcal{P}_n)$ more fully in terms of a *Dickman distribution.* Given $\theta > 0$, we shall say a random variable X has a *generalized Dickman* distribution with shape parameter θ (or $X \sim \mathrm{GD}(\theta)$ for short) if it satisfies the distributional fixed-point identity

$$X \stackrel{\mathcal{D}}{=} U^{1/\theta}(1 + X),$$

where U is uniform on $(0, 1]$, and is independent of the X on the right, and where $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution. For further information on Dickman distributions, see Section 3.

Theorem 1 *Let $\alpha > 0$. Let Z have the generalized Dickman distribution with shape parameter $\theta = 2/\alpha$. Then as $n \to \infty$,*

$$\mathcal{L}_0^\alpha(\mathcal{P}_n) \xrightarrow{\mathcal{D}} Z \tag{2}$$

and

$$\mathcal{L}_0^\alpha(\mathcal{X}_n) \xrightarrow{\mathcal{D}} Z. \tag{3}$$

The limiting distribution has Laplace transform

$$\psi_Z(t) = E[e^{-tZ}] = \exp\left(\frac{2}{\alpha}\int_0^t \frac{e^{-s} - 1}{s} \mathrm{d}s\right), \qquad t \in \mathbf{R}. \tag{4}$$

In the special case $\alpha = 1$, the distribution of the limiting variable Z has mean 2 and variance 1, and moments $m_2 = 5$, $m_3 = \frac{44}{3}$, $m_4 = \frac{293}{6}$, ...

Remarks. Perhaps the most natural case is $\alpha = 1$ (i.e., simply take the Euclidean length of edges). By considering the more general case allowing for any $\alpha > 0$, we get the whole range of possible generalized Dickman distributions as limits.

Bhatt and Roy [3] use a different approach based on the method of moments to prove the weak convergence (3) (only for $\alpha = 1$). Their argument is complicated and they give only values for the first two moments of Z, not the higher moments. Nor do they say anything about the density, distribution or moment generating functions of Z. Thus, even for $\alpha = 1$ our approach gives a good deal of extra information beyond that provided in [3]. Conversely, since Bhatt and Roy prove convergence of all moments of $\mathcal{L}_0^1(\mathcal{X}_n)$ to the corresponding moments of Z, this combined with our characterization of Z means we can identify the limit of the k-th moment of $\mathcal{L}_0^1(\mathcal{X}_n)$, for any fixed k, by computing the kth moment of Z recursively using the formula (14) below.

Our second main result concerns the *maximum* edge length of the MDST; when considering maxima we consider only the case with $\alpha = 1$ (results on maxima for other values of α are easily deduced from results for this case). Bhatt and Roy [3] considered the *maximum length of edges joined to the origin*, for the MDST on $\mathcal{X}_n \cup \{\mathbf{0}\}$, and showed that as $n \to \infty$,

$$\max_{X \in \mathcal{X}_n, X \text{ minimal}} \|X\| \xrightarrow{\mathcal{D}} \max\{U_1, U_2\} \stackrel{\mathcal{D}}{=} U_1^{1/2}, \tag{5}$$

where U_1, U_2 are independent uniform random variables on $(0,1)$.

Here, we consider instead the global maximum of *all* Euclidean edge lengths in the MDST on $\mathcal{S} \cup \{\mathbf{0}\}$, not just those joined to the origin. Denote this maximal edge length by $\mathcal{M}(\mathcal{S})$.

The limit variable for maximum edge length is given in terms of what we shall call the *max-Dickman* distribution. We define this to be the (unique) distribution of a random variable M which satisfies the distributional identity

$$M \stackrel{\mathcal{D}}{=} \max(1 - U, UM) \tag{6}$$

where U is uniformly distributed on $(0,1)$ and independent of the M on the right.

Theorem 2 *Suppose M and M' are independent max-Dickman random variables. As $n \to \infty$,*

$$\mathcal{M}(\mathcal{P}_n) \xrightarrow{\mathcal{D}} \max\{M, M'\}. \tag{7}$$

and

$$\mathcal{M}(\mathcal{X}_n) \xrightarrow{\mathcal{D}} \max\{M, M'\}. \tag{8}$$

We prove Theorem 2 in section 5.

The generalized Dickman GD(1) and max-Dickman distributions are more closely related than might at first be apparent. In probabilistic terms, they can both be expressed in terms of a Poisson point process on $(0,1)$ with mean measure μ given by $\mathrm{d}\mu = (1/x)\mathrm{d}x$. Suppose the points of this Poisson point process are listed in decreasing order as $Y_1, Y_2, \ldots$. Then the sum $\sum_i Y_i$ has the GD(1) distribution, while the maximum spacing $\max\{1 - Y_1, Y_1 - Y_2, Y_2 - Y_3, \ldots\}$ has the max-Dickman distribution. The latter is also the distribution of the largest component of the Poisson-Dirichlet distribution; see Section 3.5.

In more analytical terms, both the GD(1) and the max-Dickman probability density functions are defined in terms of the *Dickman function*, which appeared in the 1930 paper of K. Dickman on large prime factors of large integers (for a more recent reference, see [5]). In Section 3, the Dickman function and the generalized Dickman and max-Dickman distributions are described in more detail.

3 Dickman-type distributions

In this section, we review some of the properties of the distributions arising as limits in Theorems 1 and 2, before returning subsequently to the MDST. Some of these properties can be found in the literature (see, e.g., [1, 4, 7, 8, 10, 11, 23]); we endeavour to make most of the current presentation self-contained.

3.1 The Dickman function

Dickman's equation, which appears in analytic number theory, is the differential-difference equation:

$$u\rho'(u) + \rho(u-1) = 0 \qquad (u > 1). \tag{9}$$

The *Dickman function* is defined as the (unique) continuous solution $\rho(u)$ to (9) with $\rho(u) = 1$ for $0 < u \le 1$ and with ρ differentiable on $(1, \infty)$. It is convenient to extend ρ over all of $\mathbf{R}$ by setting $\rho(u) = 0$ for $u \le 0$.

It is known (see [22]) that the Dickman function is positive and decreasing on the whole interval $(1, \infty)$; that it satisfies $\rho(u) \le 1/\Gamma(u+1)$ for $u > 1$; and that it integrates to

$$\int_0^\infty \rho(x)\mathrm{d}x = e^\gamma, \tag{10}$$

where γ denotes Euler's constant, i.e. $\gamma = \lim_{k\to\infty}\left(\sum_{i=1}^k \frac{1}{i} - \log k\right)$. Numerically, $\gamma = 0.57721566\ldots$, so $e^\gamma = 1.78107\ldots$.

3.2 Probabilistic properties of the GD distributions

Proposition 2 *Let $\theta > 0$. The following random variables X are distributionally equivalent.*

(a) A random variable X satisfying the fixed point equation

$$X \stackrel{\mathcal{D}}{=} U^{1/\theta}(1+X), \tag{11}$$

where U is uniform on $(0,1]$ and independent of the X on the right hand side.

(b) A random variable X given by

$$X = \sum_{j=1}^\infty \left(\prod_{i=1}^j U_i^{1/\theta}\right) = U_1^{1/\theta} + (U_1U_2)^{1/\theta} + (U_1U_2U_3)^{1/\theta} + \cdots, \tag{12}$$

where $U_1, U_2, U_3, \ldots$ are independent uniform random variables on $(0,1]$.

(c) A random variable X given by

$$X = \sum_{n=1}^\infty \exp(-T_n)$$

where $T_1, T_2, \ldots$ are the successive arrival times of a homogeneous Poisson process of rate θ on the half-line $(0,\infty)$.

(d) A random variable X given by $X = \sum_{n=1}^\infty Y_n$, where $Y_1, Y_2, Y_3, \cdots$ are the points of a non-homogeneous Poisson point process on $(0,1)$ with mean measure $(\theta/x)\mathrm{d}x$, taken in decreasing order.

We say that a random variable X given by any of the conditions (a), (b), (c) or (d) in Proposition 2 has the *generalized Dickman* distribution with parameter θ (or $X \sim \mathrm{GD}(\theta)$ for short).

The term *Dickman distribution* has previously been used for the GD(1) distribution, i.e. that of a variable X satisfying $X \stackrel{\mathcal{D}}{=} U(1+X)$ (see e.g. [11]), and this is the usage we favour. The same term has also been used [4] for the distribution of a random variable Y satisfying $Y \stackrel{\mathcal{D}}{=} UY + 1$, as well as for other distributions [1]. It is easy to see that such a Y can be obtained by taking $Y = 1 + X$ with $X \sim \mathrm{GD}(1)$.

We shall see later (Corollary 5) that if $X \sim \mathrm{GD}(1)$ then its density function satisfies Dickman's equation.

Remark. The $\mathrm{GD}(\theta)$ distributions (particularly for $\theta = 1$ and $\theta = 2$) also appear as the limits of certain random variables in Hoare's FIND algorithm on random permutations and its variants (see e.g. [11, 15]). They also appear in the study of perpetuities (see [7]).

Proof of Proposition 2. First, suppose that X is given by the sum of the infinite random series (12), which converges almost surely because it has nonnegative terms and finite expectation. By (12),

$$X = U_1^{1/\theta}\left(1 + U_2^{1/\theta} + (U_2U_3)^{1/\theta} + (U_2U_3U_4)^{1/\theta} + \cdots\right). \tag{13}$$

The second factor in the right-hand side of (13) has the same distribution as $1 + X$, and is independent of U_1; hence, X satisfies the distributional identity (11).

Conversely, suppose that X satisfies (11). Suppose $U_1, U_2, \ldots$ are uniform on $(0, 1]$, independent of X and of each other, and set $V_i := U_i^{1/\theta}$, for each i. Then X has the same distribution as $V_1(1 + X) = V_1 + V_1X$, and hence the same distribution as $V_1(1 + V_2(1 + X)) = V_1 + V_1V_2 + V_1V_2X$, and so on. Repeating this process, the term involving X converges in probability to zero and we see that X has the same distribution as $V_1 + V_1V_2 + V_1V_2V_3 + \cdots$.

Next, suppose that X is given by definition (c), i.e. $X = \sum_n e^{-T_n}$ where $\{T_n\}$ are successive arrival times of a Poisson process of rate θ on $(0, \infty)$. Set $Y_1 = T_1$ and $Y_n = T_n - T_{n-1}$ for $n \geq 2$. The inter-arrival times $Y_1, Y_2, \ldots$ are independent and exponentially distributed with parameter θ, so for each i, and for $0 < t \leq 1$,

$$P[e^{-Y_i} \leq t] = P[Y_i \geq -\log(t)] = e^{\theta \log t} = t^\theta$$

so that e^{-Y_i} has the same distribution as $U^{1/\theta}$, where U is uniform on $(0, 1]$. Since

$$X = \sum_{n=1}^{\infty} e^{-T_n} = \sum_{n=1}^{\infty}\left(\prod_{i=1}^{n} e^{-Y_i}\right),$$

it follows that X has the same distribution as given in part (b).

Finally, definition (d) is distributionally equivalent to definition (c) by the Mapping Theorem [12], because the image of the uniform (Lebesgue) measure on $(0, \infty)$ with density θ, under the mapping $x \mapsto e^{-x}$, is the measure on $(0, 1)$ with density (θ/x). □

We now collect some further properties of the generalized Dickman distribution.

Proposition 3 *(a) If $X \sim \mathrm{GD}(\theta)$, then the Laplace transform ψ of the distribution of X is given by*

$$\psi(t) = E\left[e^{-tX}\right] = \exp\left(\theta \int_0^t \frac{e^{-s} - 1}{s} ds\right) = \exp\left(\theta \int_0^1 \frac{e^{-tu} - 1}{u} du\right), \quad t \in \mathbf{R}.$$

(b) For $\theta, \theta' \in (0, \infty)$ if X and Y are independent random variables with $X \sim \mathrm{GD}(\theta)$ and $Y \sim \mathrm{GD}(\theta')$, then $X + Y \sim \mathrm{GD}(\theta + \theta')$.

(c) The $GD(\theta)$ distribution is infinitely divisible.

(d) If $X \sim GD(\theta)$ then the k-th cumulant of X is equal to $\frac{\theta}{k}$.

(e) If $X \sim GD(\theta)$ then the moments $m_k := E[X^k]$ satisfy $m_0 = 1$ and, for integer $k \geq 1$,

$$m_k = \frac{\theta}{k} \sum_{j=0}^{k-1} \binom{k}{j} m_j. \tag{14}$$

In particular, X has expected value θ and variance $\frac{\theta}{2}$.

Proof. Suppose $X \sim \mathrm{GD}(\theta)$ and set $\psi(t) = E[e^{-tX}]$, the Laplace transform of the distribution of X. Then by definition, $X \stackrel{\mathcal{D}}{=} U^{1/\theta}(X+1)$ and so

$$\begin{aligned}
\psi(t) &= E\left[E\left[\exp\left(-tU^{1/\theta}(X+1)\right)|U\right]\right] \\
&= \int_0^1 E\left[e^{-tu^{1/\theta}}e^{-tu^{1/\theta}X}\right]\mathrm{d}u = \int_0^1 e^{-tu^{1/\theta}}\psi(tu^{1/\theta})\mathrm{d}u \\
&= \int_0^t e^{-w}\psi(w)\frac{\theta w^{\theta-1}}{t^\theta}\mathrm{d}w \\
\Rightarrow t^\theta\psi(t) &= \theta\int_0^t e^{-w}\psi(w)w^{\theta-1}\mathrm{d}w \\
\Rightarrow t^\theta\psi'(t) + \theta t^{\theta-1}\psi(t) &= \theta e^{-t}\psi(t)t^{\theta-1} \\
\Rightarrow t\psi'(t) &= \theta(e^{-t}-1)\psi(t).
\end{aligned}$$

We have the initial condition $\psi(0) = 1$ and so we have

$$\log(\psi(t)) = \int_0^t \frac{\psi'(s)}{\psi(s)}\mathrm{d}s = \theta\int_0^t \frac{e^{-s}-1}{s}\mathrm{d}s = \theta\int_0^1 \frac{e^{-tu}-1}{u}\mathrm{d}u.$$

This completes the proof of part (a). Parts (b) and (c) follow at once from (a), or alternatively by a more probabilistic argument based on the Poisson process representation of X in part (d) of Proposition 2.

Since the kth cumulant of X is defined to be the kth derivative of $\log\psi(-t)$, evaluated at $t=0$, part (d) can also be deduced from (a).

To prove part (e), suppose $X \sim \mathrm{GD}(\theta)$, and write m_k for $E[X^k]$. Then by (11),

$$\begin{aligned}
m_k = E[X^k] &= E[U^{k/\theta}]E[(1+X)^k] \\
\Rightarrow m_k &= \frac{\theta}{k+\theta}\left(m_k + \sum_{j=0}^{k-1}\binom{k}{j}m_j\right) \\
\Rightarrow m_k &= \frac{\theta}{k}\cdot\sum_{j=0}^{k-1}\binom{k}{j}m_j. \qquad \square
\end{aligned}$$

3.3 GD probability density and distribution functions

In this section we derive further properties of generalized Dickman distributions, including, among other things, a partially explicit form of the probability density and distribution functions for these distributions.

We show first that the $\mathrm{GD}(\theta)$ distribution has a probability density function that is continuous except at 0, is piecewise differentiable and satisfies a certain differential-difference equation, which generalizes Dickman's equation.

Proposition 4 *The generalized Dickman distribution with parameter $\theta > 0$ has a probability density function g_θ which is identically zero on $(-\infty, 0)$, is continuous on $(0,\infty)$, and is differentiable on $(0,1)\cup(1,\infty)$, satisfying the differential-difference equation*

$$tg'_\theta(t) = (\theta-1)g_\theta(t) - \theta g_\theta(t-1). \tag{15}$$

Proof. Let $X \sim \mathrm{GD}(\theta)$. Let G_θ be the cumulative distribution function of X. By (11), we have that

$$\begin{aligned}
G_\theta(t) &= P[X \le t] = \int_0^1 P\left[u^{1/\theta}(1+X) \le t\right]\mathrm{d}u \\
&= \int_0^1 G_\theta\left(\frac{t}{u^{1/\theta}} - 1\right)\mathrm{d}u. \qquad (16)
\end{aligned}$$

Make the substitution $s = \frac{t}{u^{1/\theta}} - 1$, so that $u = \left(\frac{t}{s+1}\right)^{\theta}$. This gives

$$G_\theta(t) = -\int_{t-1}^{\infty} G_\theta(s)\frac{\mathrm{d}u}{\mathrm{d}s}\mathrm{d}s.$$

Integrating by parts, we obtain

$$G_\theta(t) \quad = \quad G_\theta(t-1) + t^{\theta}\int_{t-1}^{\infty}(s+1)^{-\theta}\mathrm{d}G_\theta(s). \tag{17}$$

By the characterization of X in part (b) of Proposition 2, $P[X > 0] = 1$; hence, $G_\theta(t) = 0$ for $t \leq 0$. By (17),

$$G_\theta(t) = \kappa_\theta t^{\theta}, \qquad 0 \leq t \leq 1, \tag{18}$$

where $\kappa_\theta := E[(X+1)^{-\theta}]$.

By (17) and induction on n, G_θ is continuous on the interval $(-\infty, n)$ and continuously differentiable on the interval $(n-1, n)$ for $n = 1, 2, 3, \ldots$. (the case $n = 1$ is covered by (18)). Setting $g_\theta(t) = G_\theta'(t)$, for non-integer $t > 0$ we may differentiate (17) to obtain

$$g_\theta(t) \quad = \quad \theta t^{\theta-1}\int_{t-1}^{\infty}(s+1)^{-\theta}\mathrm{d}G_\theta(s). \tag{19}$$

Rearranging (19) and then differentiating once more yields

$$t^{1-\theta}g_\theta(t) = \theta\int_{t-1}^{\infty}\frac{g_\theta(s)}{(s+1)^{\theta}}\mathrm{d}s$$
$$\Rightarrow t^{1-\theta}g_\theta'(t) + (1-\theta)t^{-\theta}g_\theta(t) = -\theta t^{-\theta}g_\theta(t-1),$$

and further rearrangement gives us (15) for non-integer t. Finally, since probability density functions are defined only modulo a set of measure zero we may *define* the density function g_θ by (19) for integer t; with this definition we see from (19) and (15) that g_θ is continuous on the whole interval $(0, \infty)$ and differentiable on the interval $(1, \infty)$. □

Remark. From (15), we see that, for $t > 1$, $g_\theta'(t)$ is negative when $(\theta - 1)g_\theta(t) - \theta g_\theta(t-1) < 0$. This is true for all $t > 1$ if $\theta \leq 1$, and so, for $0 < \theta \leq 1$, g_θ is a decreasing function for $t > 1$. For $\theta > 1$, g_θ is eventually decreasing.

Corollary 5 *The generalized Dickman distribution with parameter $\theta = 1$ has a probability density function given by*

$$g_1(x) = e^{-\gamma}\rho(x) \qquad (x \in \mathbf{R}), \tag{20}$$

where ρ is the Dickman function.

Proof. By the case $\theta = 1$ of Proposition 4, the probability density function g_1 of the GD(1) distribution satisfies Dickman's equation (9), and since g_1 must be normalized to be a probability density function, by (10) it is given by (20), as required. □

Returning to the case of general $\theta > 0$, define the constant κ_θ by

$$\kappa_\theta := E[(1+X)^{-\theta}], \qquad X \sim \mathrm{GD}(\theta).$$

The constant κ_θ, $\theta > 0$, is actually given by

$$\kappa_\theta = \frac{e^{-\theta\gamma}}{\Gamma(\theta+1)}; \tag{21}$$

see, for example, [9] or [23]. In particular, $\kappa_1 = e^{-\gamma}$ and $\kappa_2 = \frac{1}{2}e^{-2\gamma}$. We also note that $\kappa_\theta = \frac{\kappa_1^\theta}{\Gamma(\theta+1)}$.

The next result gives expressions for the GD(θ) density and distribution functions obtained piecewise on the unit intervals of the positive real line, where the piecewise components are given recursively by an integral recursion relation, which can sometimes be solved explicitly.

Proposition 6 *Let g_θ and G_θ denote the probability density and cumulative distribution function, respectively, of the* GD(θ) *distribution. Then $g_\theta(t) = G_\theta(t) = 0$ for $t \leq 0$, and the functions $g_\theta(t)$ and $G_\theta(t)$ can be expressed piecewise over the unit intervals $t \in [n, n+1]$ for $n \in \mathbf{N}$, as*

$$g_\theta(t) = \begin{cases} \theta\kappa_\theta t^{\theta-1} & \text{if } 0 < t \leq 1 \\ \left(\frac{t}{n}\right)^{\theta-1} g_\theta(n) - \theta t^{\theta-1}\int_{n-1}^{t-1} \frac{g_\theta(s)}{(s+1)^\theta}\mathrm{d}s & \text{if } n \leq t \leq n+1 \ (n \in \mathbf{N}) \end{cases} \tag{22}$$

and

$$G_\theta(t) = \begin{cases} \kappa_\theta t^\theta & \text{if } 0 < t \leq 1 \\ G_\theta(t-1) + \frac{t}{\theta}g_\theta(t) & \text{if } t \geq 1 \end{cases} \tag{23}$$

Proof. For both g_θ and G_θ, the case $t \leq 0$ follows from Proposition 4, and the case $0 < t \leq 1$ follows from (18).

Suppose $n \leq t \leq n+1$ for $n \in \mathbf{N}$. Then equation (19) yields

$$t^{1-\theta}g_\theta(t) - n^{1-\theta}g_\theta(n) = -\theta\int_{n-1}^{t-1} \frac{g_\theta(s)}{(s+1)^\theta}\mathrm{d}s.$$

Rearranging this gives us (22).

Substituting in for the integral in equation (17) from equation (19) gives

$$\theta\left(G_\theta(t) - G_\theta(t-1)\right) = t g_\theta(t),$$

and (23) follows. □

The integrals one is required to perform to obtain expressions for $g_\theta(t)$ and $G_\theta(t)$ with $t \in [n, n+1]$ and $n \geq 1$ get successively more complicated as n increases, and appear to be intractable for $n \geq 2$. However, one can make progress in the $n = 1$ case. By (22) we have that for $1 \leq t \leq 2$,

$$g_\theta(t) = \theta\kappa_\theta t^{\theta-1} - \theta t^{\theta-1}\int_0^{t-1} \frac{\theta\kappa_\theta s^{\theta-1}}{(s+1)^\theta}\mathrm{d}s = \theta\kappa_\theta t^{\theta-1}\left(1 - \theta\int_1^t \frac{(u-1)^{\theta-1}}{u^\theta}\mathrm{d}u\right). \tag{24}$$

In particular, for $\theta = 1$ we see that equation (24) reduces to

$$g_1(t) = \kappa_1(1 - \log t), \qquad 1 \leq t \leq 2 \tag{25}$$

and using (23) we obtain

$$G_1(t) = \kappa_1(2t - t\log t - 1), \qquad 1 \leq t \leq 2, \tag{26}$$

while for $\theta = 2$ and $1 \le t \le 2$ we obtain

$$\begin{aligned} g_2(t) &= 2\kappa_2 t \left(1 - 2\int_1^t \frac{u-1}{u^2}\mathrm{d}u\right) = 2\kappa_2 t\left(1 - 2\left(\log t + \frac{1}{t} - 1\right)\right) \\ &= 2\kappa_2 \left(3t - 2t\log t - 2\right), \end{aligned} \tag{27}$$

and then

$$G_2(t) = \kappa_2(4t^2 - 4t - 2t^2 \log t + 1), \quad 1 \le t \le 2. \tag{28}$$

For general θ, we have that

$$\int \frac{s^{\theta-1}}{(s+1)^\theta}\mathrm{d}s = \frac{s^\theta}{\Gamma(\theta)}\sum_{k=0}^{\infty}\frac{\Gamma(\theta+k)(-s)^k}{(\theta+k)k!},$$

so that for $1 \le t \le 2$,

$$g_\theta(t) = \theta\kappa_\theta t^{\theta-1} - \theta^2\kappa_\theta t^{\theta-1}\left[\frac{(t-1)^\theta}{\Gamma(\theta)}\sum_{k=0}^{\infty}\frac{\Gamma(\theta+k)(-(t-1))^k}{(\theta+k)k!}\right]. \tag{29}$$

3.4 A generalization of Dickman's function

The density function g_θ also appears in connection with the Poisson-Dirichlet distribution with parameter $\theta > 0$, and a generalization of Dickman's function. See e.g. [10]. Define the function ρ_θ such that $\rho_\theta(t) = 1$ for $0 \le t \le 1$ and it satisfies the differential-difference equation

$$t^\theta \rho_\theta'(t) + \theta(t-1)^{\theta-1}\rho_\theta(t-1) = 0, \quad (t > 1). \tag{30}$$

Then

$$g_\theta(t) = \frac{e^{-\gamma\theta}}{\Gamma(\theta)}t^{\theta-1}\rho_\theta(t) = \theta\kappa_\theta t^{\theta-1}\rho_\theta(t), \tag{31}$$

where one can check that $g_\theta(t)$ is indeed the probability density function of our GD(θ) random variable, as it satisfies the Dickman-type equation (15). Also, notice that if we integrate (30) between 1 and ∞, making use of (31) we obtain $G_\theta(1) = \kappa_\theta$ (compare Proposition 6). One can often deduce results about $\rho_\theta(x)$ by studying $g_\theta(x)$, which is often easier to handle.

As Holst remarks [10], g_θ is the density of an infinitely divisible distribution with Lévy-Khinchine measure $\theta\mathbf{1}\{0 < x < 1\}(1/x)\mathrm{d}x$. See also Section 6.3 of Goldie and Grübel [7], which is concerned with the tail behaviour of a class distributions obtained as sums of products, including the GD distributions.

In fact, the largest component of the Poisson-Dirichlet distribution with parameter θ has distribution function $\rho_\theta(1/x)$. We return to this in section 3.5, where we discuss this distribution when $\theta = 1$ (which we call the max-Dickman distribution), since it turns out to describe the limiting distribution of the maximum edge length in the MDST.

3.5 The max-Dickman distribution

As in the case of the $GD(\theta)$ distributions, there are many characterizations of the max-Dickman distribution.

Proposition 7 *The following random variables are distributionally equivalent.*

(a) *A random variable M satisfying the fixed point equation*

$$M \stackrel{\mathcal{D}}{=} \max\{1-U, UM\}, \tag{32}$$

where U is uniform on $(0,1)$ and independent of the M on the right hand side.

(b) *A random variable M given by*

$$M = \max\{1-U_1, U_1(1-U_2), U_1U_2(1-U_3), U_1U_2U_3(1-U_4), \ldots\}, \tag{33}$$

where U_i, $i = 1,2,3,\ldots$ are i.i.d. uniform random variables on $(0,1)$.

(c) *A random variable M given by $M = \max\{1-Y_1, Y_1-Y_2, Y_2-Y_3, \ldots\}$, where $Y_1, Y_2, Y_3, \ldots$ are the points of a Poisson point process on $(0,1)$ whose intensity measure has a density $1/x$ (taken in decreasing order).*

(d) *A random variable M given by the largest (and first) component of the Poisson-Dirichlet distribution with parameter 1.*

(e) *A random variable M with distribution function $P[M \le x] = \rho(1/x)$, where ρ is the Dickman function.*

(f) *A random variable M with the size-biased distribution of $1/(Z+1)$, where Z is GD(1).*

We shall say that a random variable given by any of the conditions (a) – (f) in Proposition 7 has the *max-Dickman distribution.* Like the $GD(\theta)$ distribution, the max-Dickman distribution on $(0,1)$ has arisen in various contexts. See, for example, [5, 10].

Proof of Proposition 7. The proof of equivalence of (a) and (b) is similar to that given in the proof of Proposition 2, and is omitted this time round.

Let $Y_1, Y_2, Y_3, \ldots$ be the points of a Poisson point process on $(0,1)$ whose intensity measure has a density $1/x$ (taken in decreasing order). We have seen in the proof of Proposition 2 that the variables $Y_1, Y_2/Y_1, Y_3/Y_2, \ldots$ are independent and uniform on $(0,1)$. If we set $U_1 := Y_1$ and $U_i := Y_i/Y_{i-1}$ for $i \ge 2$, then the U_i are independent $U(0,1)$ variables, and with this definition of the U_is the definitions (b) and (c) are identical.

The equivalence of (c) and (d) follows from the fact that the vector of variables $1-Y_1, Y_1-Y_2, Y_2-Y_3, \ldots$, rearranged in decreasing order, has the Poisson-Dirichlet distribution with parameter 1. See e.g. [5].

Suppose now that M is given by the definition in part (e). Then, following [10], we have for $0 \le t \le 1$ that if U is uniform on $(0,1)$ and independent of M, then

$$\begin{aligned} P[\max\{1-U, UM\} \le t] &= \int_{1-t}^{1} P\left[M \le \frac{t}{u}\right] \mathrm{d}u = \int_{1-t}^{1} \rho\left(\frac{u}{t}\right) \mathrm{d}u \\ &= t\int_{(1/t)-1}^{1/t} \rho(y)\mathrm{d}y = \rho(1/t), \end{aligned}$$

where the last equality follows from (23) and Corollary 5. Thus, M satisfies (32).

To check the equivalence of definitions (f) and (e), let $Y = (Z+1)^{-1}$ with $Z \sim \mathrm{GD}(1)$, and let f_Y denote the probability density function of Y. Then for $0 < t \le 1$, using Dickman's equation we have

$$\begin{aligned} P[Y \le t] &= 1 - G_1(t^{-1}-1) \\ \Rightarrow f_Y(t) &= t^{-2}g_1(t^{-1}-1) = -t^{-3}g_1'(t^{-1}), \end{aligned}$$

so that the size-biased distribution of Y has a probability density function on $(0,1)$ proportional to $-t^{-2}g_1'(t^{-1})$.

On the other hand, M given by definition (e) has probability density function $-x^{-2}\rho'(1/x)$. These two distributions are the same. □

Let h and H respectively denote the probability density and distribution functions of the max-Dickman distribution. We can obtain expressions for h and H from the GD(1) density function g_1. Again, we obtain a piecewise description of the functions, but now the intervals are $\left[\frac{1}{n+1}, \frac{1}{n}\right]$, $n \in \mathbf{N}$. Note that the cumulative distribution of the limiting variable in Theorem 2, namely that of the maximum of two independent max-Dickman variables, is given by $H(\cdot)^2$, so the next result provides some partial information about this distribution function.

Proposition 8 *The max-Dickman density and distribution functions h and H are given in terms of the* GD(1) *density function g_1 as follows:*

$$h(x) = \begin{cases} 0 & \text{if } x \geq 1 \\ 1/x & \text{if } \frac{1}{2} \leq x < 1 \\ \frac{1}{x} + \frac{1}{x}\log\left(\frac{x}{1-x}\right) & \text{if } \frac{1}{3} \leq x < \frac{1}{2} \\ \frac{e^\gamma}{x} g_1\left(\frac{1-x}{x}\right) & \text{if } \frac{1}{n+1} \leq x < \frac{1}{n} \end{cases} \quad \text{and } H(x) = \begin{cases} 1 & \text{if } x \geq 1 \\ 1 + \log x & \text{if } \frac{1}{2} \leq x < 1 \\ e^\gamma g_1(1/x) & \text{if } \frac{1}{n+1} \leq x < \frac{1}{n} \end{cases}$$

for all $n \in \mathbf{N}$, and with $h(x) = H(x) = 0$ for $x \leq 0$.

Proof. By part (e) of Proposition 7, $H(x) = \rho(1/x)$. Differentiating, we obtain

$$h(x) = -\frac{1}{x^2}\rho'(1/x) = \frac{1}{x}\rho\left(\frac{1}{x} - 1\right) = \frac{e^\gamma}{x} g_1\left(\frac{1-x}{x}\right),$$

where the second equality follows from Dickman's equation. Using the fact that $g_1(x) = e^{-\gamma}$ for $0 \leq x \leq 1$ and $g_1(x) = e^{-\gamma}(1 - \log x)$ for $1 \leq x \leq 2$ then yields

$$h(x) = \frac{1}{x} \quad (1 \geq x \geq 1/2); \; h(x) = \frac{1}{x}\left(1 - \log\frac{1-x}{x}\right), \quad (1/2 \geq x \geq 1/3),$$

and

$$H(x) = 1 - \log(1/x) = 1 + \log x, \quad (1 \geq x \geq 1/2).$$

This completes the proof. □

Remarks. Also of interest is the largest component M of the Poisson-Dirichlet distribution with parameter θ, for general $\theta > 0$. See, for example, [8, 10, 23]. Then $P[M \leq x] = \rho_\theta(1/x)$, where the function ρ_θ, related to g_θ, is as introduced in Section 3.4.

Let $E_1(y)$ denote the exponential integral function,

$$E_1(y) = \int_y^\infty \frac{e^{-x}}{x}\mathrm{d}x = \int_1^\infty \frac{e^{-yx}}{x}\mathrm{d}x.$$

Then, Proposition 2.2 of [10] (with a minor correction to the denominator there) shows that for $k = 1, 2, 3, \ldots$,

$$E[M^k] = \frac{\Gamma(\theta)}{\Gamma(\theta + k)}\int_0^\infty y^{k-1}\exp(-y - \theta E_1(y)). \tag{34}$$

In particular, for the $\theta = 1$ case this leads to $E[M] = \int_0^\infty e^{-y-E_1(y)} dy$, which can be evaluated numerically to give $E[M] = 0.6243299..$ (see e.g. [23]). Griffiths [8] tabulates values for $P[M > x]$ for several values of θ.

Returning to the case with $\theta = 1$, we note that one can show that $E[(M+1)^{-1}] = E[M]$, and that $E[M^{-k}] = ke^{\gamma} m_{k-1}$ for $k \in \mathbf{N}$, where $(m_k)_{k \geq 1}$ are the moments of the GD(1) distribution. Thus, using (14) one can recursively generate the moments of the distribution of M^{-1}, which is yet another distribution that has on occasion been given the term 'Dickman distribution' (see [1]).

4 Proof of Theorem 1

The intuition behind Theorem 1 goes as follows. If there exists a minimal point of $\mathcal{P}_n$ (or $\mathcal{X}_n$) near to the origin, then there is no minimal point lying to the north-east of that point. Hence, the minimal points are likely to all lie near to either the x-axis or the y-axis, and the contributions from these two axes are nearly independent. Near the x-axis, the x-coordinates of successive minimal points (taken in order of increasing y-coordinate) form a sequence of products of uniforms $U_1, U_1U_2, U_1U_2U_3, \ldots$ and summing these gives a Dickman distribution. Similarly for the y-axis.

In the course of the proof we use the notation card($\mathcal{X}$) for the number of elements (i.e., the cardinality) of any finite point set $\mathcal{X}$. We shall also use *Slutsky's Theorem* (see e.g. [6, 16]). This says that if $(\xi_n, \zeta_n)_{n \geq 1}$ is a sequence of random pairs with $\xi_n \xrightarrow{\mathcal{D}} \xi$ and $\zeta_n \xrightarrow{P} 0$ as $n \to \infty$ for some ξ, then $\xi_n + \zeta_n \xrightarrow{\mathcal{D}} \xi$.

We shall also use the following coupling lemma relating the point processes $\mathcal{X}_n$ and $\mathcal{P}_n$.

Lemma 1 *There exist point processes $\mathcal{X}'_n, \mathcal{P}'_n$ defined on the same probability space as each other for each n, such that:*

- *$\mathcal{X}'_n$ has the same distribution as $\mathcal{X}_n$.*
- *$\mathcal{P}'_n$ has the same distribution as $\mathcal{P}_n$.*
- *With probability tending to 1 as $n \to \infty$, the set of minimal elements of $\mathcal{P}'_n$ is identical to the set of minimal elements of $\mathcal{X}'_n$.*

Proof. Let $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \ldots$ be independent and uniformly distributed on $(0,1]^2$, let $N(n)$ be Poisson with parameter n and independent of $(\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \ldots)$, let $\mathcal{P}'_n := \{\mathbf{U}_1, \ldots, \mathbf{U}_{N(n)}\}$, and for $m \in \mathbf{N}$ set $\mathcal{X}'_m := \{\mathbf{U}_1, \ldots, \mathbf{U}_m\}$. Then $\mathcal{P}'_n \stackrel{\mathcal{D}}{=} \mathcal{P}_n$ and $\mathcal{X}'_n \stackrel{\mathcal{D}}{=} \mathcal{X}_n$.

Let A_m be the event that $\mathbf{U}_m$ is a minimal element of $\mathcal{X}'_m$, and let δ_m be the number of minimal elements of $\mathcal{X}'_m$. By exchangeability, each point $\mathbf{U}_i, i \leq m$ is equally likely to be minimal in $\mathcal{X}'_m$, so that $E[\delta_m] = mP[A_m]$. By [2], or by the proof of Theorem 1.1(a) of [3]

$$E[\delta_m] = \sum_{i=1}^{m} (1/i) \sim \log(m) \quad \text{as } m \to \infty.$$

Hence, $P[A_m] \sim (\log m)/m$ as $m \to \infty$, and therefore

$$P\left[\cup_{n-n^{3/4} \leq m \leq n+n^{3/4}} A_m\right] \leq 3n^{3/4}(\log n)/n \to 0. \tag{35}$$

Let E_n denote event that the set of minimal points in $\mathcal{X}'_n$ differs from the set of minimal points of $\mathcal{P}'_n$. By the coupling of $\mathcal{X}'_m$ $(m \geq 1)$ and $\mathcal{P}'_n$, E_n occurs only if A_m occurs for some m with $N(n) < m \leq n$ (if $N(n) < n$) or with $n < m \leq N(n)$ (if $N(n) > n$). Hence,

$$P[E_n] \leq P[|N(n) - n| \geq n^{3/4}] + P[\cup_{n-n^{3/4} \leq m \leq n+n^{3/4}} A_m].$$

In the right hand side, the first probability tends to zero by Chebyshev's inequality while the second tends to zero by (35), and hence $P[E_n] \to 0$ as asserted. $\square$

We now work towards a proof of (2). Let $\mathcal{Y}_n$ be the set of minimal elements of the point set $\mathcal{P}_n$, i.e., the set of elements of $\mathcal{P}_n$ which are joined to $\mathbf{0}$ in the MDST on $\mathcal{P}_n \cup \{\mathbf{0}\}$.

Lemma 2 *As $n \to \infty$, we have* $(\log n)^{-1}\mathrm{card}(\mathcal{Y}_n) \xrightarrow{P} 1$.

Proof. The corresponding result for the number of minimal points of binomial point process $\mathcal{X}_n$ (actually with almost sure convergence) is Theorem 1.1(a) of [3]. Using Lemma 1, we can deduce the result asserted for the Poisson point process $\mathcal{P}_n$. $\square$

Fix a constant δ lying in the range $(0, 1/2)$ but otherwise arbitrary. Define the point sets

$$\mathcal{Y}_n^x := \mathcal{Y}_n \cap \left((0,1] \times (0, n^{-\delta}]\right); \qquad \mathcal{Y}_n^y := \mathcal{Y}_n \cap \left((0, n^{-\delta}] \times (0,1]\right).$$

Fix $\alpha > 0$, as given in the statement of Theorem 1. Define the variables

$$L_n^x := \textstyle\sum_{\mathbf{X} \in \mathcal{Y}_n^x} \|\mathbf{X}\|^\alpha; \qquad L_n^y := \displaystyle\sum_{\mathbf{X} \in \mathcal{Y}_n^y} \|\mathbf{X}\|^\alpha; \tag{36}$$

$$N_n^x := \mathrm{card}(\mathcal{Y}_n^x); \qquad N_n^y := \mathrm{card}(\mathcal{Y}_n^y).$$

Thus, L_n^x is the total weight of α-power-weighted edges of the MDST on $\mathcal{P}_n$ which are incident to the origin and lie entirely in the horizontal strip $(0,1] \times (0, n^{-\delta}]$, while N_n^x is the number of such edges; L_n^y and N_n^y are defined analogously in terms of a vertical strip.

Proposition 9 *Let $S \sim \mathrm{GD}(1/\alpha)$, i.e. let S be a generalized Dickman random variable with parameter $\theta = 1/\alpha$. Then as $n \to \infty$,*

$$L_n^x \xrightarrow{\mathcal{D}} S, \quad \text{and} \quad L_n^y \xrightarrow{\mathcal{D}} S.$$

Proof. We give the proof only for L_n^x; the argument for L_n^y is entirely analogous.

List the minimal points $\mathcal{Y}_n^x$, in order of increasing y-coordinate, as $\mathbf{X}_1^x, \mathbf{X}_2^x, \ldots, \mathbf{X}_{N_n^x}^x$. In co-ordinates we set $\mathbf{X}_j^x = (X_j^x, Y_j^x)$. Since the points $\mathbf{X}_j^x$ are minimal, we have that

$$Y_1^x < Y_2^x < \cdots < Y_{N_n^x}^x, \quad \text{and} \quad X_1^x > X_2^x > \cdots > X_{N_n^x}^x.$$

Then $L_n^x = \sum_{j=1}^{N_n^x} \|\mathbf{X}_j^x\|^\alpha$. For each n, let S_n^x be the estimate for L_n^x obtained by counting only the projections of the edge lengths onto the x-axis, i.e., set

$$S_n^x = \sum_{j=1}^{N_n^x} (X_j^x)^\alpha.$$

If $(x,y) \in (0,1]^2$ then $\|(x,y)\| \le x + y$, and by the Mean Value Theorem,

$$\|(x,y)\|^\alpha - x^\alpha \le (x+y)^\alpha - x^\alpha \le \alpha 2^{\alpha-1} y \qquad (\alpha \ge 1)$$

whereas by the concavity of the function $t \mapsto t^\alpha$ for $\alpha < 1$,

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

Hence, there is a constant $C(\alpha)$ such that with probability 1,

$$\begin{aligned} 0 \leq L_n^x - S_n^x &\leq C(\alpha) \sum_{j=1}^{N_n^x} (Y_j^x)^{\min(1,\alpha)} \\ &\leq C(\alpha) n^{-\delta \min(1,\alpha)} N_n^x. \end{aligned} \tag{37}$$

Since $N_n^x = O(\log(n))$ in probability by Lemma 2, it follows that $n^{-\delta \min(1,\alpha)} N_n^x$ converges in probability to zero as $n \to \infty$, and hence so does $L_n^x - S_n^x$. Therefore, by Slutsky's theorem it suffices to prove that

$$S_n^x \xrightarrow{\mathcal{D}} S \quad \text{as } n \to \infty. \tag{38}$$

We prove this by a coupling argument in which we construct (copies of) the random variables S_n^x $(n \geq 1)$ on a common probability space.

Let $\mathcal{H}$ be a homogeneous Poisson process of unit intensity on the infinite strip $(0,1] \times (0,\infty)$. Let $\mathcal{H}_n$ be the image of $\mathcal{H}$ under the linear mapping $\tau_n : \mathbf{R}^2 \to \mathbf{R}^2$ given by

$$\tau_n((x,y)) = (x, n^{-1}y). \tag{39}$$

By the Mapping Theorem [12], $\mathcal{H}_n$ is a homogeneous Poisson process of intensity n on the same strip $(0,1] \times (0,\infty)$. Since we are interested only in proving a convergence in distribution result (38), we may assume without loss of generality that $\mathcal{P}_n$ is the restriction of the Poisson process $\mathcal{H}_n$ to the unit square $(0,1] \times (0,1]$.

List the minimal elements of $\mathcal{H}$ in order of increasing y-coordinate as $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3 \ldots$, with coordinate representation $\mathbf{X}_j = (X_j, Y_j)$. Then $Y_1 < Y_2 < Y_3 < \cdots$, and $X_1 > X_2 > \cdots$. Define $U_1 = X_1$, and set

$$U_j = \frac{X_j}{X_{j-1}}, \quad j = 2, 3, \ldots.$$

It is not hard to see that $U_1, U_2, \ldots$ are mutually independent and are each uniformly distributed over $(0,1)$. Therefore, setting

$$S := \sum_{j=1}^{\infty} X_j^\alpha = \sum_{j=1}^{\infty} \left(\prod_{i=1}^{j} U_i^\alpha \right), \tag{40}$$

we see from Proposition 2 that S has a generalized Dickman distribution $\mathrm{GD}(1/\alpha)$.

The set of minimal elements of a point set in $\mathbf{R}^2$ is invariant under the linear transformation $\tau_n(\cdot)$ defined at (39), as is the relative order of the y-coordinates of the minimal elements. Therefore, under our assumption that $\mathcal{P}_n$ is the restriction of $\tau_n(\mathcal{H})$ to the unit square, we see that $\mathbf{X}_j^x = \tau_n(\mathbf{X}_j)$ for $1 \leq j \leq N_n^x$. Hence, since the mapping τ_n leaves x-coordinates unchanged,

$$S_n^x = \sum_{j=1}^{N_n^x} X_j^\alpha.$$

Since N_n^x is the number of minimal elements in the restriction of $\mathcal{H}$ to the set $(0,1] \times (0, n^{1-\delta}]$, it is the case that $N_n^x \to \infty$ almost surely as $n \to \infty$. Therefore, with this particular coupled construction of the point processes $\mathcal{P}_n, n \geq 1$, the variables S_n^x converge to S as $n \to \infty$, almost surely and hence in distribution. In other words, (38) holds as required. □

The random variables L_n^x and L_n^y are not quite independent since they both depend on the configuration of points of $\mathcal{P}_n$ in $(0, n^{-\delta}]^2$. Our argument to deal with this fact requires some further terminology.

Fix a further constant β with $0 < \beta < \delta < 1/2$, and define the following rectangular regions, as shown in Figure 1 below:

$$R_2^x(n) := (n^{-\beta}, 1] \times (0, n^{-\delta}]; \qquad R_2^y(n) := (0, n^{-\delta}] \times (n^{-\beta}, 1];$$
$$R_1^x(n) := (n^{-\delta}, n^{-\beta}] \times (0, n^{-\delta}]; \qquad R_1^y(n) := (0, n^{-\delta}] \times (n^{-\delta}, n^{-\beta}];$$
$$R_0(n) := (0, n^{-\delta}]^2; \qquad R_3(n) := (n^{-\delta}, 1]^2.$$

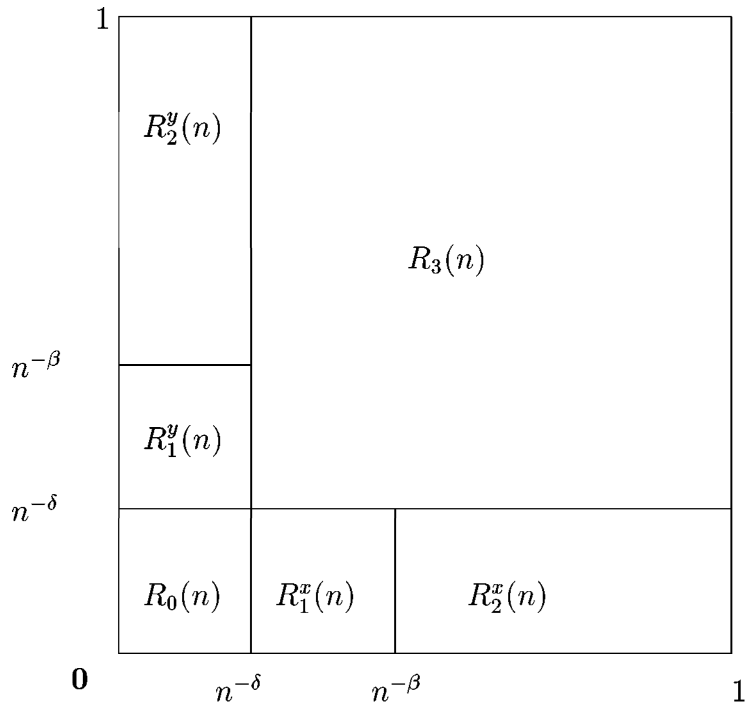

Figure 1: The regions of $(0,1]^2$.

Let $N_2^x(n)$, $N_2^y(n)$, $N_1^x(n)$, $N_1^y(n)$, $N_0(n)$, and $N_3(n)$ be the number of elements of $\mathcal{Y}_n$ that fall in the regions $R_2^x(n)$, $R_2^y(n)$, $R_1^x(n)$, $R_1^y(n)$, $R_0(n)$ and $R_3(n)$ respectively.

Similarly, let $L_2^x(n)$, $L_2^y(n)$, $L_1^x(n)$, $L_1^y(n)$, $L_0(n)$, and $L_3(n)$ be the total weights of edges that are incident to the origin in the MDST on $\mathcal{P}_n \cup \{\mathbf{0}\}$ and start from points that fall in the regions $R_2^x(n)$, $R_2^y(n)$, $R_1^x(n)$, $R_1^y(n)$, $R_0(n)$ and $R_3(n)$ respectively, i.e., set

$$L_2^x(n) := \sum_{\mathbf{X} \in \mathcal{Y}_n \cap R_2^x(n)} \|\mathbf{X}\|^\alpha, \quad L_1^x(n) := \sum_{\mathbf{X} \in \mathcal{Y}_n \cap R_1^x(n)} \|\mathbf{X}\|^\alpha, \quad L_0(n) := \sum_{\mathbf{X} \in \mathcal{Y}_n \cap R_0(n)} \|\mathbf{X}\|^\alpha, \tag{41}$$

$$L_2^y(n) := \sum_{\mathbf{X} \in \mathcal{Y}_n \cap R_2^y(n)} \|\mathbf{X}\|^\alpha, \quad L_1^y(n) := \sum_{\mathbf{X} \in \mathcal{Y}_n \cap R_1^y(n)} \|\mathbf{X}\|^\alpha, \quad L_3(n) := \sum_{\mathbf{X} \in \mathcal{Y}_n \cap R_3(n)} \|\mathbf{X}\|^\alpha. \tag{42}$$

Then

$$\mathcal{L}_0^\alpha(\mathcal{P}_n) = L_2^x(n) + L_2^y(n) + L_1^x(n) + L_1^y(n) + L_0(n) + L_3(n). \tag{43}$$

The next result shows that most of the terms in (43) are asymptotically negligible.

Lemma 3 *As $n \to \infty$,*

$$L_1^x(n) + L_1^y(n) + L_0(n) + L_3(n) \xrightarrow{P} 0.$$

Proof. Observe first that

$$L_1^x(n) + L_1^y(n) + L_0(n) \leq (2n^{-\beta})^\alpha (N_1^x(n) + N_1^y(n) + N_0(n))$$

and since $N_1^x(n) + N_1^y(n) + N_0(n) = O(\log n)$ in probability by Lemma 2,

$$L_1^x(n) + L_1^y(n) + L_0(n) \xrightarrow{P} 0. \tag{44}$$

If $\mathrm{card}(\mathcal{P}_n \cap R_0(n)) > 0$ then $L_3(n) = N_3(n) = 0$. However, $\mathrm{card}(\mathcal{P}_n \cap R_0(n))$ is Poisson with parameter $n^{1-2\delta}$, which tends to infinity since we assume $\delta < 1/2$. Hence, $P[L_3(n) \neq 0] \to 0$, so that $L_3(n) \xrightarrow{P} 0$. Combined with (44), this gives us the result. □

Define $\tilde{\mathcal{P}}_n$ to be the point process $\mathcal{P}_n$ with all points in the corner region $R_0(n)$ removed, i.e., set

$$\tilde{\mathcal{P}}_n := \mathcal{P}_n \backslash R_0(n).$$

Let $\tilde{\mathcal{Y}}_n$ be the set of minimal elements of $\tilde{\mathcal{P}}_n$. Define the point sets

$$\mathcal{Z}_n^x := \mathcal{Y}_n \cap R_2^x(n); \quad \tilde{\mathcal{Z}}_n^x := \tilde{\mathcal{Y}}_n \cap R_2^x(n);$$
$$\mathcal{Z}_n^y := \mathcal{Y}_n \cap R_2^y(n); \quad \tilde{\mathcal{Z}}_n^y := \tilde{\mathcal{Y}}_n \cap R_2^y(n).$$

Then $\mathcal{Z}_n^x \subseteq \tilde{\mathcal{Z}}_n^x$, since adding the points in $R_0(n)$ cannot cause any new minimal points in $R_2^x(n)$ to be created, although it can cause previously minimal points in $R_2^x(n)$ to cease to be minimal. Using the convention $\min\{\} = +\infty$, set

$$Y_0^-(n) := \min\{Y : \mathbf{X} = (X, Y) \in \mathcal{P}_n \cap R_0(n)\},$$

which is the y-coordinate of the lowest point of $\mathcal{P}_n$ in $R_0(n)$ (or $+\infty$ if no such point exists). Let

$$Y_1^-(n) := \min\{Y : \mathbf{X} = (X, Y) \in \mathcal{P}_n \cap R_1^x(n)\},$$

which is the y-coordinate of the lowest point of $\mathcal{P}_n$ in $R_1^x(n)$ (or $+\infty$ if there are no such points).

Lemma 4 *If $Y_1^-(n) < Y_0^-(n)$, then $\mathcal{Z}_n^x = \tilde{\mathcal{Z}}_n^x$.*

Proof. If $\mathbf{X} = (X, Y) \in R_2^x(n)$ and $\mathbf{X}' = (X', Y') \in R_0(n) \cup R_1^x(n)$, then $\mathbf{X}' \preccurlyeq \mathbf{X}$ if and only if $Y' \leq Y$. Hence, $\tilde{\mathcal{Z}}_n^x$ consists of those minimal elements of $\mathcal{P}_n \cap R_2^x(n)$ that have a lower y-coordinate than $Y_1^-(n)$. Likewise, $\mathcal{Z}_n^x$ consists of those minimal elements of $\mathcal{P}_n \cap R_2^x(n)$ that have a lower y-coordinate than $\min(Y_1^-(n), Y_0^-(n))$. Thus, if $Y_1^-(n) < Y_0^-(n)$, then the sets $\tilde{\mathcal{Z}}_n^x$ and $\mathcal{Z}_n^x$ must be identical. □

Lemma 5 *As $n \to \infty$, $P[Y_1^-(n) < Y_0^-(n)] \to 1$.*

Proof. List the points of $\mathcal{P}_n \cap (0, n^{-\beta}] \cap (0, \infty)$ in order of increasing y-coordinate as $\mathbf{V}_1^n, \mathbf{V}_2^n, \mathbf{V}_3^n \ldots$. In coordinates, write $\mathbf{V}_1^n = (V_1^n, W_1^n)$. Then V_1^n is uniform on $(0, n^{-\beta}]$ and is independent of W_1^n. Also W_1^n is exponential with parameter $n^{1-\beta}$. Since $\beta < \delta$ and $\delta < 1/2 < 1 - \beta$,

$$P[\{V_1 \in (n^{-\delta}, n^{-\beta}]\} \cap \{W_1 < n^{-\delta}\}] \to 1 \quad \text{as } n \to \infty.$$

However, if this event occurs then $Y_1^-(n) < Y_0^-(n)$ so the proof is complete. □

Define the random variables

$$\tilde{L}_2^x(n) := \sum_{\mathbf{X} \in \tilde{\mathcal{Z}}_n^x} \|\mathbf{X}\|^\alpha, \quad \tilde{L}_2^y(n) := \sum_{\mathbf{X} \in \tilde{\mathcal{Z}}_n^y} \|\mathbf{X}\|^\alpha.$$

In other words, $\tilde{L}_2^x(n)$, $\tilde{L}_2^y(n)$ are the total weight of edges from points in $R_2^x(n)$, $R_2^y(n)$ respectively joined to the origin in the MDST on $\tilde{\mathcal{P}}_n \cup \{\mathbf{0}\}$.

We assert that $\tilde{L}_2^x(n)$ and $\tilde{L}_2^y(n)$ are independent. This follows because $\tilde{L}_2^x(n)$ is determined by the configuration of $\mathcal{P}_n \cap (R_1^x(n) \cup R_2^x(n))$, whereas $\tilde{L}_n^y$ is determined by the configuration of $\mathcal{P}_n \cap (R_1^y(n) \cup R_2^y(n))$. Since the regions $R_1^x(n) \cup R_2^x(n)$ and $R_1^y(n) \cup R_2^y(n)$ are disjoint, the independence asserted follows from the standard spatial independence properties of the Poisson process.

Proof of Theorem 1. By the earlier definitions at (36) and (41), $L_n^x = L_2^x(n) + L_1^x(n) + L_0(n)$. Hence,

$$\tilde{L}_2^x(n) = (L_n^x - L_1^x(n) - L_0(n)) + (\tilde{L}_2^x(n) - L_2^x(n)).$$

By Lemma 3, $L_1^x(n) + L_0(n) \xrightarrow{P} 0$. Also, by Lemmas 4 and 5, $\tilde{L}_2^x(n) - L_2^x(n) \xrightarrow{P} 0$. Hence, by Proposition 9, and Slutsky's theorem,

$$\tilde{L}_2^x(n) \xrightarrow{\mathcal{D}} S,$$

where $S \sim \mathrm{GD}(1/\alpha)$, and by an analogous argument we obtain $\tilde{L}_2^y(n) \xrightarrow{\mathcal{D}} S$.

Let S and S' be independent $\mathrm{GD}(1/\alpha)$ variables and let $Z \sim GD(2/\alpha)$. Since $\tilde{L}_2^x(n)$ and $\tilde{L}_2^y(n)$ are independent, we obtain

$$\tilde{L}_2^x(n) + \tilde{L}_2^y(n) \xrightarrow{\mathcal{D}} S + S' \stackrel{\mathcal{D}}{=} Z, \tag{45}$$

where the last distributional equality follows from Proposition 3 (b). By (43), we have

$$\begin{aligned} \mathcal{L}_0^\alpha(\mathcal{P}_n) - (\tilde{L}_2^x(n) + \tilde{L}_2^y(n)) = &(L_2^x(n) - \tilde{L}_2^x(n)) + (L_2^y(n) - \tilde{L}_2^y(n)) \\ &+ L_1^x(n) + L_1^y(n) + L_0(n) + L_3(n). \end{aligned}$$

In this expression, the right hand side tends to zero in probability by Lemmas 3, 4 and 5. Hence, by (45) and Slutsky's theorem we obtain (2).

Next we prove (3). To do this we use the coupled copies $\mathcal{X}_n'$ and $\mathcal{P}_n'$ of $\mathcal{X}_n$, $\mathcal{P}_n$ respectively, given by Lemma 1. That result gives us

$$\mathcal{L}_0^\alpha(\mathcal{P}_n') - \mathcal{L}_0^\alpha(\mathcal{X}_n') \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty. \tag{46}$$

Since $\mathcal{P}_n' \stackrel{\mathcal{D}}{=} \mathcal{P}_n$, we see from (2) that $\mathcal{L}_0^\alpha(\mathcal{P}_n')$ converges in distribution to the $\mathrm{GD}(2/\alpha)$ variable Z. By (46) and Slutsky's theorem, the same is true of $\mathcal{L}_0^\alpha(\mathcal{X}_n')$, and (3) follows since $\mathcal{X}_n' \stackrel{\mathcal{D}}{=} \mathcal{X}_n$. □

In the case $\alpha = 1$, the limiting variable Z is GD(2); its moments and moment generating function are obtained by application of Proposition 3 . □

5 Proof of Theorem 2

The intuition behind Theorem 2 is that the longest edge is likely to be near either the x-axis or y-axis. Near the x-axis, the x-coordinates of the points of $\mathcal{P}_n$ (or $\mathcal{X}_n$), taken in order of increasing y-coordinate, form a sequence of uniforms with each uniform joined to its nearest predecessor lying to its left. Similarly for the y-coordinate.

The proof of Theorem 2 follows similar lines to that of Theorem 1 (see Section 4). Fix a constant $\delta \in (1/2, 1)$. Define the point sets

$$\mathcal{P}_n^x := \mathcal{P}_n \cap \left((0,1] \times (0, n^{-\delta}]\right); \qquad \mathcal{P}_n^y := \mathcal{P}_n \cap \left((0, n^{-\delta}] \times (0,1]\right).$$

For $\mathbf{X} \in \mathcal{P}_n$, if $\mathbf{X}'$ is the directed nearest neighbour of $\mathbf{X}$ in $\mathcal{P}_n$, write $d(\mathbf{X})$ for the length of the edge from $\mathbf{X}$ in the MDST, i.e. $d(\mathbf{X}) = \|\mathbf{X} - \mathbf{X}'\|$. Define

$$M_n^x := \max_{\mathbf{X} \in \mathcal{P}_n^x} d(\mathbf{X}); \qquad M_n^y := \max_{\mathbf{X} \in \mathcal{P}_n^y} d(\mathbf{X}). \tag{47}$$

Thus, M_n^x is the length of the longest edge in the MDST on $\mathcal{P}_n$ from points in the horizontal strip $(0,1] \times (0, n^{-\delta}]$; M_n^y is defined analogously in terms of a vertical strip.

Proposition 10 *Let M have the max-Dickman distribution given by (32). Then as $n \to \infty$,*

$$M_n^x \xrightarrow{\mathcal{D}} M, \quad \text{and} \quad M_n^y \xrightarrow{\mathcal{D}} M.$$

Proof. We give the proof only for M_n^x; the argument for M_n^y is entirely analogous.

Define the random variable $\nu(n) := \mathrm{card}(\mathcal{P}_n^x)$. List the points of $\mathcal{P}_n^x$, in order of increasing y-coordinate, as $\mathbf{X}_1^x, \mathbf{X}_2^x, \mathbf{X}_3^x, \ldots, \mathbf{X}_{\nu(n)}^x$. In co-ordinates we set $\mathbf{X}_j^x = (X_j^x, Y_j^x)$. Then $Y_1^x < Y_2^x < \cdots < Y_{\nu(n)}^x$.

For each n, let ξ_n^x be the estimate for M_n^x obtained by considering only the projections of the edge lengths onto the x-axis, i.e., set

$$\xi_n^x = \max_{1 \le i \le \nu(n)} \left\{ X_i^x - \max_{0 \le j < i} \left(X_j^x \mathbf{1}_{\{X_j^x < X_i^x\}} \right) \right\}. \tag{48}$$

where we set $X_0^x := 0$.

By construction of the MDST and the triangle inequality, with probability 1,

$$0 \le M_n^x - \xi_n^x \le n^{-\delta},$$

so that $M_n^x - \xi_n^x$ converges to 0 almost surely. Therefore, by Slutsky's theorem it suffices to prove that

$$\xi_n^x \xrightarrow{\mathcal{D}} M \quad \text{as } n \to \infty. \tag{49}$$

As in the proof of Proposition 9, let $\mathcal{H}$ be a homogeneous Poisson process of unit intensity on the infinite strip $(0,1] \times (0,\infty)$, and let $\mathcal{H}_n$ be the image of $\mathcal{H}$ under the linear mapping τ_n defined at (39). Again, we may assume without loss of generality that $\mathcal{P}_n$ is the restriction of the Poisson process $\mathcal{H}_n$ to the unit square $(0,1]^2$.

List the elements of $\mathcal{H}$ in order of increasing y-coordinate as $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3 \ldots$, with coordinate representation $\mathbf{X}_j = (X_j, Y_j)$. Since the linear mapping τ_n preserves x-coordinates and the relative order of y-coordinates, our coupling of $\mathcal{P}_n$ to $\mathcal{H}$ means that the sequence $X_1^x, \ldots X_{\nu(n)}^x$ is identical to the first $\nu(n)$ terms in the infinite sequence $(X_1, X_2, \ldots)$.

A *record value* in the sequence $X_1, X_2, X_3, \ldots$ is a value X_i which exceeds $\max\{X_1, \ldots, X_{i-1}\}$ (the first value X_1 is also included as a record value). Let $j(1), j(2), j(3), \ldots$ be the values of $i \in \{1,2,3,\ldots\}$ such that X_i is a record value, arranged in increasing order so that $1 = j(1) < j(2) < j(3) < \cdots$. Let $R_n := \max\{k : j(k) \le \nu(n)\}$ be the number of record values in the finite sequence $(X_1, X_2 \ldots, X_{\nu(n)})$.

Since each non-record X_i lies in an interval between preceding record values, the first maximum in the definition at (48) is achieved at a record value, so that

$$\xi_n^x = \max_{1 \le i \le R_n} X_{j(i)} - X_{j(i-1)}, \tag{50}$$

where we set $j(0) = 0$ and $X_0 = 0$. Define $U_1 = 1 - X_1$, and set

$$U_i = \frac{1 - X_{j(i)}}{1 - X_{j(i-1)}}, \quad i = 2, 3, \ldots.$$

It is not hard to see that $U_1, U_2, \ldots$ are mutually independent and are each uniformly distributed over $(0, 1)$. Therefore, setting

$$M := \max\{1 - U_1, U_1(1 - U_2), U_1U_2(1 - U_3), U_1U_2U_3(1 - U_4), \ldots\}, \tag{51}$$

we see that M indeed has the max-Dickman distribution as described in Proposition 7 (b). Further,

$$(1 - U_k)\prod_{i=1}^{k-1} U_i = \frac{X_{j(k)} - X_{j(k-1)}}{1 - X_{j(k)}} \prod_{i=1}^{k-1}\left(\frac{1 - X_{j(i)}}{1 - X_{j(i-1)}}\right) = X_{j(k)} - X_{j(k-1)}, \tag{52}$$

for $k = 2, 3, \ldots$.

With our chosen coupling of $\mathcal{P}_n$ to $\mathcal{H}$, $\nu(n) := \text{card}(\mathcal{P}_n^x)$ is the number of points in the restriction of $\mathcal{H}$ to the set $(0, 1] \times (0, n^{1-\delta}]$, so that $\nu(n) \to \infty$ almost surely as $n \to \infty$. Therefore, since there are almost surely infinitely many records, $R_n \to \infty$ almost surely as $n \to \infty$. Hence by (50), (51) and (52), $\xi_n^x \to M$ as $n \to \infty$, almost surely with this coupling. Hence, (49) holds as required. □

Let $M_3(n)$ denote the maximum edge length of edges of the MDST on $\mathcal{P}_n$ starting in $(n^{-\delta}, 1]^2$, i.e., set

$$M_3(n) := \max\{\|d(\mathbf{X})\| : \mathbf{X} \in \mathcal{P}_n \cap (n^{-\delta}, 1]^2\}.$$

Lemma 6 *It is the case that* $M_3(n) \xrightarrow{P} 0$ *as* $n \to \infty$.

Proof. Recall that $(1/2) < \delta < 1$. Choose a second constant $\varepsilon \in (0, 1 - \delta)$. Consider a collection of overlapping horizontal and vertical rectangles of the form

$$((i-1)n^{-\varepsilon}, in^{-\varepsilon}] \times ((j-1)n^{-\delta}, jn^{-\delta}], \qquad (i,j) \in \mathbf{N} \times \mathbf{N}, i \le \lfloor n^{\varepsilon} \rfloor, j \le \lfloor n^{\delta} \rfloor,$$
$$((i-1)n^{-\delta}, in^{-\delta}] \times ((j-1)n^{-\varepsilon}, jn^{-\varepsilon}], \qquad (i,j) \in \mathbf{N} \times \mathbf{N}, i \le \lfloor n^{\delta} \rfloor, j \le \lfloor n^{\varepsilon} \rfloor.$$

For each rectangle, the number of points of $\mathcal{P}_n$ in the rectangle is Poisson with parameter $n^{1-\delta-\varepsilon}$, so that the probability that at least one subsquare contains no point of $\mathcal{P}_n$ is bounded by

$$2n^{\delta+\varepsilon} \exp(-n^{1-\delta-\varepsilon}) \to 0.$$

However, if each rectangle contains at least one point of $\mathcal{P}_n$ then $M_3(n)$ is bounded by $3n^{-\varepsilon}$, and the result follows. □

Proof of Theorem 2 It is a little easier to deal with the non-independence of M_n^x and M_n^y than with the corresponding problem in the proof of Theorem 1. Define $\tilde{M}_n^x$ to be the maximal edge-length of edges starting in $(0, 1] \times (0, n^{-\delta}]$ for the MDST on the point set

$$(\mathcal{P}_n \cap ((n^{-\delta}, 1] \times (0, n^{-\delta}])) \cup \{\mathbf{0}\}.$$

In other words, $\tilde{M}_n^x$ is the same as M_n^x except that Poisson points in $(0, n^{-\delta}]^2$ are ignored in defining $\tilde{M}_n^x$. By independence properties of the Poisson process, $\tilde{M}_n^x$ is independent of M_n^y.

It is not hard to see that

$$|M_n^x - \tilde{M}_n^x| \le 2n^{-\delta}, \qquad \text{almost surely.} \tag{53}$$

Let M, M' be independent random variables both having the max-Dickman distribution. By Proposition 10, equation (53), and Slutsky's theorem,

$$\tilde{M}_n^x \xrightarrow{\mathcal{D}} M \quad \text{and} \quad M_n^y \xrightarrow{\mathcal{D}} M,$$

and since $\tilde{M}_n^x$ and M_n^y are independent,

$$\max(\tilde{M}_n^x, M_n^y) \xrightarrow{\mathcal{D}} \max(M, M'). \tag{54}$$

By (53), with probability 1,

$$|\max(M_n^x, M_n^y) - \max(\tilde{M}_n^x, M_n^y)| \leq 2n^{-\delta},$$

so by (54) and Slutsky's theorem,

$$\max(M_n^x, M_n^y) \xrightarrow{\mathcal{D}} \max(M, M'). \tag{55}$$

Also,

$$\mathcal{M}(\mathcal{P}_n) = \max(M_n^x, M_n^y, M_3(n)),$$

so that

$$0 \leq \mathcal{M}(\mathcal{P}_n) - \max(M_n^x, M_n^y) \leq M_3(n),$$

which tends to zero in probability by Lemma 6. Hence, a further application of Slutsky's theorem to (55) shows that $\mathcal{M}(\mathcal{P}_n) \xrightarrow{\mathcal{D}} \max(M, M')$, i.e., (7) holds.

To deduce (8) from (7), consider the coupled point processes $\mathcal{X}'_n$ and $\mathcal{P}'_n$ described in Lemma 1, given in terms of a sequence of independent uniform points $\mathbf{U}_i$ in $(0,1]^2$ and an independent Poisson variable $N(n)$ as given in the proof of Lemma 1. Let B_n be the event that at least one point of the symmetric difference $\mathcal{X}'_n \triangle \mathcal{P}'_n$ lies in $(0,1]^2 \setminus (n^{-\delta}, 1]^2$. Then

$$P[B_n] \leq P[|N(n) - n| > n^{(1/4)+(\delta/2)}] + 2n^{(1/4)+(\delta/2)} P[\mathbf{U}_1 \in (0,1]^2 \setminus (n^{-\delta}, 1]^2] \to 0, \tag{56}$$

where the convergence follows by Chebyshev's inequality and the fact that we took $\delta > 1/2$.

Recall that $M_3(n)$ denotes the maximum length for edges of the MDST on $\mathcal{P}_n$ starting in $(n^{-\delta}, 1]^2$; similarly, let $M'_3(n)$, respectively $\tilde{M}_3(n)$, denote the maximum edge length for edges of the MDST on $\mathcal{P}'_n$, respectively on $\mathcal{X}'_n$, starting in $(n^{-\delta}, 1]^2$. Then $M'_3(n) \xrightarrow{P} 0$ by Lemma 6, and a similar proof shows that $\tilde{M}_3(n) \xrightarrow{P} 0$ as well. Using also (56) we obtain

$$|\mathcal{M}(\mathcal{X}'_n) - \mathcal{M}(\mathcal{P}'_n)| \leq 2\mathbf{1}_{B_n} + M'_3(n) + \tilde{M}_3(n) \xrightarrow{P} 0,$$

and since $\mathcal{M}(\mathcal{X}_n) \stackrel{\mathcal{D}}{=} \mathcal{M}(\mathcal{X}'_n)$ and $\mathcal{M}(\mathcal{P}_n) \stackrel{\mathcal{D}}{=} \mathcal{M}(\mathcal{P}'_n)$, eqn (8) follows from (7) by yet another application of Slutsky's theorem. □

Acknowledgements

The first author began this work while at the University of Durham, and was also supported by the Isaac Newton Institute for Mathematical Sciences, Cambridge. The second author was supported by the EPSRC.

References

[1] R. Arratia (1998) On the central role of scale invariant Poisson processes on $(0, \infty)$. Microsurveys in discrete probability (Princeton, NJ, 1997), 21–41, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* **41**, AMS, Providence, RI.

[2] O. Barndorff-Nielsen and M. Sobel (1966), On the distribution of the number of admissible points in a vector random sample, *Theory Probab. Appl.* **11**, 249–269.

[3] A.G. Bhatt and R. Roy (1999), On a random directed spanning tree (preprint, to appear in *Adv. Appl. Probab.*).

[4] L. Devroye and R. Neininger (2002), Density approximation and exact simulation of random variables that are solutions of fixed-point equations, *Adv. Appl. Probab.*, **34**, 441–468.

[5] P. Donnelly, and G. Grimmett (1993), On the asymptotic distribution of large prime factors. *J. London Math. Soc.* **47**, 395–404.

[6] R. Durrett, (1991) Probability: Theory and Examples, Wadsworth and Brooks/Cole, Pacific Grove.

[7] C.M. Goldie and R. Grübel (1996), Perpetuities with thin tails, *Adv. Appl. Probab.* **28**, 463–480.

[8] R.C. Griffiths (1988), On the distribution of points in a Poisson process. *J. Appl. Probab.* **25**, 336–345.

[9] D. Hensley (1986), The convolution powers of the Dickman function, *J. London Math. Soc.*, **33**, 395–406.

[10] L. Holst (2001), The Poisson-Dirichlet distribution and its relatives revisited. Preprint, electronically available at http://www.math.kth.se/matstat

[11] H-K. Hwang and T-H. Tsai (2002), Quickselect and Dickman function, *Combinatorics, Probability and Computing*, **11**, 353–371.

[12] J.F.C. Kingman (1993), Poisson Processes, *Oxford Studies in Probability*, **3**, Oxford University Press, Oxford.

[13] A.N. Kolmogorov and S.V. Fomin (trans. R.A. Silverman) (1975), Introductory Real Analysis, Dover, New York.

[14] J.B. Kruskal (1956), On the shortest spanning subtree of a graph and the travelling salesman problem, *Proc. Amer. Math. Soc.*, **7**, 48–50.

[15] H.M. Mahmoud and R.T. Smythe (1998), Probabilistic Analysis of MULTIPLE QUICK SELECT, *Algorithmica*, **22**, 569–584.

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

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

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

[19] M.D. Penrose and A.R. Wade, On the total length of the random minimal directed spanning tree (provisional title, in preparation).

[20] I. Rodriguez-Iturbe and A. Rinaldo (1997), Fractal River Basins: Chance and Self-Organization, Cambridge University Press, Cambridge.

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

[22] G. Tenenbaum (1995), Introduction to Analytic and Probabilistic Number Theory, Cambridge University Press, Cambridge.

[23] G.A. Watterson (1976) The stationary distribution of the infinitely-many alleles diffusion model. *J. Appl. Prob.* **13**, 639–651.

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