Extremes on Trees

by

Tailen Hsing [1] and Holger Rootzén [2]

Texas A&M University and Chalmers University of Technology

Abstract. This paper considers the asymptotic distribution of the longest edge of the minimal spanning tree and nearest neighbor graph on $\boldsymbol{X}_1, \dots, \boldsymbol{X}_{N_n}$ where $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ are iid. in $\Re^2$ with distribution F and N_n is independent of the $\boldsymbol{X}_i$ and satisfies $N_n/n \xrightarrow{p} 1$. A new approach based on spatial blocking and a locally orthogonal coordinate system is developed to treat cases for which F has unbounded support. The general results are applied to a number of special cases, including elliptically contoured distributions, distributions with independent Weibull-like margins, and distributions with parallel level curves.

1. Introduction. Recall that the (Euclidean) minimal spanning tree (MST) on a finite set of points $(\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_N)$ in $\Re^2$ is the connected graph with these points as vertices and with the minimum total edge length. The (Euclidean) nearest neighbor graph (NNG) on $(\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_N)$ is the graph on which each point $\boldsymbol{X}_i$ is connected to its nearest neighbor in the set. In this paper the $\boldsymbol{X}_i$ are assumed to be random and we are interested in the asymptotic distribution of the longest edge on these graphs as $N \to \infty$.

Penrose (1997, 1998) considered these problems by assuming that the $\boldsymbol{X}_i$ are uniformly distributed in a unit cube or symmetrically normally distributed in $\Re^d$. The essential ideas are that (a) the lengths of the edges at any location in space depend primarily on the $\boldsymbol{X}_i$ in the vicinity of that location and as a result are asymptically independent of the edges in other parts of the space and (b) the presence of an extremely long edge is a rare event and hence the likelihood of having multiple extremely long edges at any location is asymptotically negligible compared with the likelihood of having one such edge there. Clearly (a) is also essential in proving central limit theorems for the total edge lengths; see Kesten and Lee (1996), Lee (1997) and Penrose (2000). In view of (a) and (b), the asymptotic distribution of the longest edge of MST or NNG can be established through a Poisson convergence of the number of extreme edges, which in Penrose (1997, 1998) is achieved through the Chen-Stein method.

[1]Department of Statistics, Texas A&M University, College Station, Texas, USA, e-mail: "thsing@stat.tamu.edu". Research partially supported by the Alexander von Humboldt Foundation.

[2]Department of Mathematical Statistics, Chalmers University of Technology, SE-41296 Gothenberg, Sweden, e-mail: "rootzen@math.chalmers.se". Research partially supported by the Knut and Alice Wallenberg Foundation.

Specifically, we will consider the case where the random graphs are based on $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{N_n}$ where $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ are iid. with distribution F and N_n is independent of the $\boldsymbol{X}_i$ with $N_n/n \xrightarrow{p} 1$. We are primarily interested in the case where F has an unbounded support (although our methodology works rather generally also if F has bounded support). In particular, we will focus on those F whose density is of the form

$$f(\boldsymbol{x}) = e^{-U(\boldsymbol{x})}$$

where $U(\boldsymbol{x})$ is regularly varying in some sense and satisfies suitable regularity conditions. This covers many elliptically contoured distributions, and in particular correlated normal distributions and distributions with independent Weibull-like marginals as special cases but also large classes of other distributions. Poisson approximation is the key idea. However, we use a more direct approach of spatial blocking as opposed to the Chen-Stein method so as to have a higher degree of flexibility. Computations of integrals of the type $\int_A e^{-nF(S(\boldsymbol{x};r))} dF(\boldsymbol{x})$, where $S(\boldsymbol{x}; r) = \{\boldsymbol{y} : |\boldsymbol{y} - \boldsymbol{x}| \leq r\}$ is the sphere centered at $\boldsymbol{x}$ with radius r, are a key part of the solution for the problem on hand. One of the novelties of our approach is the introduction of a locally orthogonal coordinate system with respect to the levels curves of U, which enables particularly effective handling of such integrals.

This paper is structured as follows. Section 2 introduces the notation and a spatial blocking argument, as well as other preliminaries. The development of a locally orthogonal system is made in Section 3. The main results are given in Section 4. Sections 5 and 6 consider homogeneous level curves and parallel level curves, respectively, and most of the proofs are given in Section 7.

Possible extensions of the results in this paper include (a) allowing the dimension d to be general, (b) allowing the distance measure to be more general (e.g. considering weighted edges on the graphs), (c) considering the k-nearest neighbor graph in which each point is connected to its k nearest neighbors. The solutions for these involve additional technical details but few new important ideas.

2. Fundamentals. For convenience of notation, denote by $\mathrm{MST}(N)$ and $\mathrm{NNG}(N)$, respectively, the MST and NNG on two-dimensional random variables $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N$ for any random variable N defined on the same space as the $\boldsymbol{X}_i$. Also let M_G be the longest edge of G, where in this paper G will be either a MST or NNG.

As outlined in section 1, the graphs of interests will be based on the points $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{N_n}$ where $X_1, X_2, \ldots$ are iid. with distribution F, N_n is positive, integer-valued with

$$N_n/n \xrightarrow{p} 1,$$

and N_n and the X_i are independent. This assumption will hold throughout the paper. The random quantities whose asymptotic distributions we study in this paper are $M_{\mathrm{MST}(N_n)}$ and $M_{\mathrm{NNG}(N_n)}$. We will first assume that N_n is Poisson distributed with mean n, in which case the points $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{N_n}$ can be thought of as the points of a Poisson process with intensity measure nF. The independent-increment property of the Poisson process offers an obvious advantage in proving Poisson convergence. We will later show how the result based on the Poisson assumption can be extended to the general class of point processes described here.

Temporal blocking is a common technique for proving limit theorems for weakly dependent random variables. See Ibragimov and Linnik (1971) and Leadbetter, Lindgren and Rootzén (1983). Based on a spatial blocking argument our first theorem below is a Poisson convergence result for $M_{\mathrm{NNG}(N_n)}$ under the assumption that $N_n \sim \mathrm{Poisson}(n)$. For $r > 0$ and measurable set A, define

$$\mu_n^{(1)}(A, r) = n \int_A e^{-nF(S(\boldsymbol{x};r))} dF(\boldsymbol{x}) \tag{1}$$

so that $\mu_n^{(1)}$ is the expected number of points in A with its nearest neighbor further away than r, and let

$$\mu_n^{(2)}(A, r) = n^2 \int_A \int_A I_{(r<|\boldsymbol{x}-\boldsymbol{y}|\leq 2r)} e^{-nF(S(\boldsymbol{x};r)\cup S(\boldsymbol{y};r))} dF(\boldsymbol{x}) dF(\boldsymbol{y})$$

where $|\cdot|$ denotes the Euclidean norm. In the theorem, the $A_{n,i}$ are the spatial blocks, (a) ensures sufficient separation to make the points in the blocks independent, (b) is the basic norming requirement of convergence of the expected number of points with no nearest neighbor closer than r_n, (c) says that nothing happens on the leftover parts between the blocks, and in particular makes the blocks large enough, (d) ensures that the blocks individually are not too large and (e) prevents clustering.

Theorem 1 *Let N_n be distributed as Poisson with mean n and $\{k_n\}$ a sequence of positive constants tending to ∞. Suppose that $\{r_n\}$ is a sequence of positive constants and for each n, $A_{n,1}, \ldots, A_{n,k_n}$ are disjoint measurable sets in $\Re^2$ such that*

(a) $\min_{1\leq i\neq j\leq k_n} \inf(|\boldsymbol{x}-\boldsymbol{y}| : \boldsymbol{x} \in A_{n,i}, \boldsymbol{y} \in A_{n,j}) > 2r_n$ *for all large n,*

(b) $\lim_{n\to\infty} \mu_n^{(1)}(\Re^2, r_n) =$ *some* $\tau \in (0, \infty)$,

(c) $\lim_{n\to\infty} \mu_n^{(1)}((\cup_{i=1}^{k_n} A_{n,i})^c, r_n) = 0$,

(d) $\lim_{n\to\infty} \max_{1\leq i\leq k_n} \mu_n^{(1)}(A_{n,i}, r_n) = 0$,

(e) $\lim_{n\to\infty} \max_{1\leq i\leq k_n} [\mu_n^{(2)}(A_{n,i}, r_n)/\mu_n^{(1)}(A_{n,i}, r_n)] = 0$.

Then

$$P(M_{\mathrm{NNG}(N_n)} \leq r_n) \to e^{-\tau}. \tag{2}$$

Proof: For convenience write $\mathcal{P}_n = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{N_n}\}$. For any set A, define

$$N(\boldsymbol{x}, A) = \inf(|\boldsymbol{x} - \boldsymbol{y}| : \boldsymbol{y} \in A - \{\boldsymbol{x}\}).$$

We will show the stronger result that $\sum_{\boldsymbol{x} \in \mathcal{P}_n} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r_n)}$ converges in distribution to the Poisson distribution with mean τ, from which (2) follows. Let

$$\tilde{\mu}_n^{(2)}(A, r) = n^2 \int_A \int_A I_{(|\boldsymbol{x} - \boldsymbol{y}| > r)} e^{-nF(S(\boldsymbol{x};r) \cup S(\boldsymbol{y};r))} dF(\boldsymbol{x}) dF(\boldsymbol{y})$$

It is easy to check that

$$E\left(\sum_{\boldsymbol{x} \in \mathcal{P}_n \cap A} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r)} \right) = \mu_n^{(1)}(A, r) \tag{3}$$

and

$$E\left(\sum_{\boldsymbol{x}, \boldsymbol{y} \in \mathcal{P}_n \cap A,\, \boldsymbol{x} \neq \boldsymbol{y}} I_{(N(\boldsymbol{x}, \mathcal{P}_n) \wedge N(\boldsymbol{y}, \mathcal{P}_n) > r)} \right) = \tilde{\mu}_n^{(2)}(A, r). \tag{4}$$

Write

$$\sum_{\boldsymbol{x} \in \mathcal{P}_n} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r_n)} = \sum_{i=1}^{k_n} \sum_{\boldsymbol{x} \in \mathcal{P}_n \cap A_{n,i}} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r_n)} + \sum_{\boldsymbol{x} \in \mathcal{P}_n \cap (\cup_{i=1}^{k_n} A_{n,i})^c} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r_n)}.$$

Roughly speaking, in the following we will establish that the second sum is negligible and for the first sum that it's expectation tends to τ, the summands are infinitesimal and the probability that any summand is bigger than 1 is negligible compared with that it being bigger than 0. More precisely, we will show that

$$\sum_{\boldsymbol{x} \in \mathcal{P}_n \cap A_{n,i}} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r_n)}, 1 \leq i \leq k_n, \text{ are independent random variables for each } n, \tag{5}$$

$$E\left(\sum_{i=1}^{k_n} \sum_{\boldsymbol{x} \in \mathcal{P}_n \cap A_{n,i}} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r_n)} \right) \to \tau, \tag{6}$$

$$E\left(\sum_{\boldsymbol{x} \in \mathcal{P}_n \cap (\cup_{i=1}^{k_n} A_{n,i})^c} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r_n)} \right) \to 0, \tag{7}$$

$$\max_{1 \leq i \leq k_n} P\Big(\sum_{\boldsymbol{x} \in \mathcal{P}_n \cap A_{n,i}} I_{(N(\boldsymbol{x}, \mathcal{P}_n) > r_n)} > 0 \Big) \to 0, \tag{8}$$

$$\max_{1\le i\le k_n} \frac{P\Big(\sum_{\boldsymbol{x}\in\mathcal{P}_n\cap A_{n,i}} I_{(N(\boldsymbol{x},\mathcal{P}_n)>r_n)} > 1\Big)}{P\Big(\sum_{\boldsymbol{x}\in\mathcal{P}_n\cap A_{n,i}} I_{(N(\boldsymbol{x},\mathcal{P}_n)>r_n)} > 0\Big)} \to 0. \tag{9}$$

That $\sum_{\boldsymbol{x}\in\mathcal{P}_n} I_{(N(\boldsymbol{x},\mathcal{P}_n)>r_n)}$ converges in distribution to the Poisson distribution with mean τ follows readily from (5)-(9), which we now proceed to prove. First, it is clear that $\sum_{\boldsymbol{x}\in\mathcal{P}_n\cap A_{n,i}} I_{(N(\boldsymbol{x},\mathcal{P}_n)>r_n)}$ is completely determined by the set $\mathcal{P}_n \cap (A_{n,i})_{r_n}$ where $(A)_\delta = \{\boldsymbol{x} : d(\boldsymbol{x}, A) \le \delta\}$. Hence, by the independent-increment property of the Poisson process and assumption (a) we conclude that (5) holds. In view of (3), that (6) and (7) hold are simple consequences of assumptions (b) and (c). Since

$$P\Big(\sum_{\boldsymbol{x}\in\mathcal{P}_n\cap A} I_{(N(\boldsymbol{x},\mathcal{P}_n)>r)} > 0\Big) \le E\Big(\sum_{\boldsymbol{x}\in\mathcal{P}_n\cap A} I_{(N(\boldsymbol{x},\mathcal{P}_n)>r)}\Big) = \mu_n^{(1)}(A,r),$$

(8) follows from assumption (d). To prove (9) note that by (3), (4) and the Bonferroni inequality,

$$\mu_n^{(1)}(A,r) - \tilde{\mu}_n^{(2)}(A,r) \le P\Big(\sum_{\boldsymbol{x}\in\mathcal{P}_n\cap A} I_{(N(\boldsymbol{x},\mathcal{P}_n)>r)} > 0\Big) \le \mu_n^{(1)}(A,r)$$

and

$$P\left(\sum_{\boldsymbol{x}\in\mathcal{P}_n\cap A} I_{(N(\boldsymbol{x},\mathcal{P}_n)>r_n)} > 1\right) \le P\left(\sum_{\boldsymbol{x},\boldsymbol{y}\in\mathcal{P}_n\cap A,\, \boldsymbol{x}\ne\boldsymbol{y}} I_{(N(\boldsymbol{x},\mathcal{P}_n)\wedge N(\boldsymbol{y},\mathcal{P}_n)>r_n)} > 0\right) \le \tilde{\mu}_n^{(2)}(A,r_n).$$

Hence (9) follows if

$$\lim_{n\to\infty} \max_{1\le i\le k_n} \frac{\tilde{\mu}_n^{(2)}(A_{n,i},r_n)}{\mu_n^{(1)}(A_{n,i},r_n)} = 0. \tag{10}$$

However, note that

$$\begin{aligned}
\tilde{\mu}_n^{(2)}(A,r_n) - \mu_n^{(2)}(A,r_n) &= n^2 \int_A\int_A I_{(|\boldsymbol{x}-\boldsymbol{y}|>2r_n)} e^{-nF(S(\boldsymbol{x};r_n)\cup S(\boldsymbol{y};r_n))} dF(\boldsymbol{x})dF(\boldsymbol{y}) \\
&\le \Big(\mu_n^{(1)}(A,r_n)\Big)^2.
\end{aligned}$$

Hence (10) follows from the assumptions (d) and (e).

□

As a simple example, if F is the uniform distribution on $[0,1]\times[0,1]$ then it is straightforward to verify (a)-(e) of Theorem 1 by taking $r_n = (\frac{\log n - \log\tau}{n\pi})^{1/2}$ (cf. (1) of Penrose, 1997) and the $A_{n,i}$ to be the sets $[(j-1)n^{-1/4}, jn^{-1/4} - n^{-1/3}] \times [(k-1)n^{-1/4}, kn^{-1/4} - n^{-1/3}], j,k = 1\ldots,[n^{1/4}]$. The advantage of this approach will be more obvious for more complicated distributions.

The next result shows how to generalize Theorem 1 by removing the Poisson assumption on N_n.

Theorem 2 *Suppose that*

$$\sup_n n \int \bar{F}^n(S(\boldsymbol{x}; r_n))dF(\boldsymbol{x}) < \infty \tag{11}$$

and

$$\sup_n n^2 \int F(S(\boldsymbol{x}; r_n))\bar{F}^{n-1}(S(\boldsymbol{x}, r_n))dF(\boldsymbol{x}) < \infty \tag{12}$$

where $\bar{F} = 1 - F$. *Then*

$$P(M_{\mathrm{NNG}(N_n)} \le r_n < M_{\mathrm{NNG}(n)}) + P(M_{\mathrm{NNG}(n)} \le r_n < M_{\mathrm{NNG}(N_n)}) \to 0$$

for any sequence of positive, integer-valued random variables N_n *with* $N_n/n \xrightarrow{p} 1$.

Proof: First for any positive constant δ_n,

$$\begin{aligned}
& P(M_{\mathrm{NNG}(N_n)} \le r_n < M_{\mathrm{NNG}(n)}) + P(M_{\mathrm{NNG}(n)} \le r_n < M_{\mathrm{NNG}(N_n)}) \\
\le \quad & P(|N_n - n| > n\delta_n) + \sum_{|j-n| \le n\delta_n} P(N_n = j) \times \\
& \times \Big(P(M_{\mathrm{NNG}(j)} \le r_n < M_{\mathrm{NNG}(n)}) + P(M_{\mathrm{NNG}(n)} \le r_n < M_{\mathrm{NNG}(j)}) \Big).
\end{aligned}$$

Pick δ_n to tend to 0 slowly enough so that the first term tends to 0 and so we only have to deal with the second term. It suffices to show that both $P(M_{\mathrm{NNG}(j)} \le r_n < M_{\mathrm{NNG}(n)})$ and $P(M_{\mathrm{NNG}(n)} \le r_n < M_{\mathrm{NNG}(j)})$ tend to 0 for $j \in [n(1-\delta_n), n(1+\delta_n)]$. Since the same argument applies to both so we focus on the first. Suppose first that $j \in [n(1-\delta_n), n-1]$. Observe that if $M_{\mathrm{NNG}(j)} \le r_n < M_{\mathrm{NNG}(n)}$, then there exists $i \in \{j+1, \ldots, n-1\}$ such that $\wedge_{\substack{1 \le s \le n \\ s \ne i}} |X_i - X_s| > r_n$. Hence

$$P(M_{\mathrm{NNG}(j)} \le r_n < M_{\mathrm{NNG}(n)}) \le n\delta_n \int \bar{F}^{n-1}(S(\boldsymbol{x}, r_n))dF(\boldsymbol{x}) \to 0$$

by (11). Next, let $j \in [n+1, n(1+\delta_n)]$. Suppose that $M_{\mathrm{NNG}(j)} \le r_n < M_{\mathrm{NNG}(n)}$ and that the longest edge on NNG(n) initiates from $\boldsymbol{X}_k$ so that $\wedge_{\substack{1 \le s \le n \\ s \ne k}} |\boldsymbol{X}_k - \boldsymbol{X}_s| > r_n$. Since the longest edge becomes $\le r_n$ by adding points $\boldsymbol{X}_{n+1}, \ldots, \boldsymbol{X}_j$ to the graph, one of those additional points must be within a distance of r_n from $\boldsymbol{X}_k$. Thus,

$$\begin{aligned}
& P(M_{\mathrm{NNG}(j)} \le r_n < M_{\mathrm{NNG}(n)}) \\
\le \quad & P(\text{ for some } k \in \{1, \ldots, n\} \text{ and } \ell \in \{n+1, \ldots, j\}, \\
& \qquad \wedge_{\substack{1 \le s \le n \\ s \ne k}} |X_k - X_s| > r_n \text{ but } |X_k - X_\ell| \le r_n) \\
\le \quad & n^2\delta_n \int F(S(\boldsymbol{x}; r_n))\bar{F}^{n-1}(S(\boldsymbol{x}, r_n))dF(\boldsymbol{x}) \to 0
\end{aligned}$$

by (12). This concludes the proof.

$\square$

We note that (11) follows from readily from (b) of Theorem 1 since $\bar{F}(S(x;r_n))^n \le e^{-nF(S(x;r_n))}$. Indeed, for the cases that we will consider, the conditions (11) and (12) are both naturally satisfied. Thus, once we have the asymptotic distribution of the extreme edge length for the Poisson NNG, we can immediately extend that to a more general class of NNG's.

The following result gives an argument for deriving the asymptotic distribution of $M_{\mathrm{MST}(N_n)}$ from that of $M_{\mathrm{NNG}(N_n)}$, in light of the fact that $M_{\mathrm{MST}(N_n)} \ge M_{\mathrm{NNG}(N_n)}$ (see proof below).

Theorem 3 *Suppose that for any sequence of positive constants m_n with $m_n/n \to 1$, we have*

$$n \int \Big[\bar{F}^{m_n}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|)) - \bar{F}^{m_n}(S(\boldsymbol{x};r_n))\Big] I(|\boldsymbol{x}| > r_n/2) dF(\boldsymbol{x}) \to 0. \tag{13}$$

Then

$$P(M_{\mathrm{NNG}(N_n)} \le r_n < M_{\mathrm{MST}(N_n)}) \to 0$$

for any sequence of positive, integer-valued random variables N_n with $N_n/n \xrightarrow{p} 1$.

Proof: The proof is basically the same as that of Lemma 4 of Penrose (1998). For completeness, the essential ideas are reproduced here. Again write

$$\begin{aligned} & P(M_{\mathrm{NNG}(N_n)} \le r_n < M_{\mathrm{MST}(N_n)}) \\ \le\ & P(|N_n/n - 1| > \delta_n) + \sum_{|m/n-1| \le \delta_n} P(N_n = m) P(M_{\mathrm{NNG}(m)} \le r_n < M_{\mathrm{MST}(m)}), \end{aligned} \tag{14}$$

where δ_n tends to zero slowly enough so that the first term tends to 0. Now make $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m$ a graph by including an edge between each pair of points at a distance at most r. Denote the resulting graph by G_r. Clearly, for a small enough r, G_r comprises connected sub-graphs, called r-clusters, which are disconnected with one another. Observe that

$$\begin{aligned} M_{\mathrm{MST}(m)} &= \inf\{r : \ G_r \text{ is connected}\}, \\ M_{\mathrm{NNG}(m)} &= \inf\{r : \ G_r \text{ does not contain a } r\text{-cluster which is a singleton}\}. \end{aligned}$$

Thus, $M_{\mathrm{MST}(m)} \ge M_{\mathrm{NNG}(m)}$. Suppose $M_{\mathrm{NNG}(m)} \le r_n < M_{\mathrm{MST}(m)}$. Then it means that G_{r_n} is disconnected and each r_n-cluster has at least two points. Take a r_n-cluster and let $\boldsymbol{x}$ be the vertex in the cluster which is closest to 0. Write $\mathcal{I}_m^{\boldsymbol{x}} = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\} - \{\boldsymbol{x}\}$. Clearly

$$\mathcal{I}_m^{\boldsymbol{x}} \cap S(\boldsymbol{x};r_n) \ne \emptyset \text{ and } \mathcal{I}_m^{\boldsymbol{x}} \cap S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|) = \emptyset$$

since $S(\boldsymbol{x}; r_n)$ contains points and only points belonging to the same r_n-cluster while $S(0; |\boldsymbol{x}|) - \{\boldsymbol{x}\}$ does not contain points from the same r_n-cluster. This means that if $M_{\mathrm{NNG}(m)} \le r_n < M_{\mathrm{MST}(m)}$ then there are at least two points belonging to different r_n-clusters having the described property, where one of them must have modulus bigger than $r_n/2$. Thus,

$$
\begin{aligned}
& P(M_{\mathrm{NNG}(m)} \le r_n < M_{\mathrm{MST}(m)}) \\
\le\ & P\Bigg(\sum_{i=1}^{m} I\Big(|\boldsymbol{X}_i| > r_n/2, \mathcal{I}_m^{\boldsymbol{X}_i} \cap S(\boldsymbol{X}_i; r_n) \neq \emptyset \text{ and} \\
& \qquad\qquad\qquad\qquad \mathcal{I}_m^{\boldsymbol{X}_i} \cap S(\boldsymbol{X}_i; r_n) \cap S(0; |\boldsymbol{X}_i|) = \emptyset \Big) \ge 1 \Bigg) \\
\le\ & E\left(\sum_{i=1}^{m} I\Big(|\boldsymbol{X}_i| > r_n/2,\ \mathcal{I}_m^{\boldsymbol{X}_i} \cap S(\boldsymbol{X}_i; r_n) \neq \emptyset \text{ and } \mathcal{I}_m^{\boldsymbol{X}_i} \cap S(\boldsymbol{X}_i; r_n) \cap S(0; |\boldsymbol{X}_i|) = \emptyset \Big) \right) \\
=\ & E\left(\sum_{i=1}^{m} I\Big(|\boldsymbol{X}_i| > r_n/2, \mathcal{I}_m^{\boldsymbol{X}_i} \cap S(\boldsymbol{X}_i; r_n) \cap S(0; |\boldsymbol{X}_i|) = \emptyset \Big) \right) - \\
& \qquad\qquad - E\left(\sum_{i=1}^{m} I\Big(|\boldsymbol{X}_i| > r_n/2, \mathcal{I}_m^{\boldsymbol{X}_i} \cap S(\boldsymbol{X}_i; r_n) = \emptyset \Big) \right) \\
=\ & m \int \Big[\bar{F}^{m-1}(S(\boldsymbol{x}; r_n) \cap S(0; |\boldsymbol{x}|)) - \bar{F}^{m-1}(S(\boldsymbol{x}; r_n)) \Big] I(|\boldsymbol{x}| > r_n/2) dF(\boldsymbol{x}),
\end{aligned}
$$

which tends to 0 by (13). The result follows from this in view of (14).

□

The condition (13) and the logic that leads to it seem to have little bearing on the conclusion they are implying. However, it holds quite naturally for the cases that we study in this paper.

It might also be worth noting that it is easy to find examples where the Poisson convergence which is used for Theorem 1 does not hold. In fact, this is the rule rather than the exception in the one-dimensional case, and for example can be seen to be the case for the one-dimensional density $e^{-|x|}/2$. A simple two-dimensional example is then obtained by letting the two-dimensional distribution be concentrated on the x_1-axis and have this density. If this distribution is mixed with, say, the standard bivariate normal distribution, a more genuine two-dimensional example where Poisson convergence does not occur is obtained. This example can also be simply modified to have a smooth density.

3. Time varying, locally orthogonal coordinate system. We first introduce a "locally orthogonal" coordinate system, which is natural for the problem on hand. In the remaining part of the paper, we will focus on densities of the form

$$f(\boldsymbol{x}) = e^{-U(\boldsymbol{x})}$$

where $U(\boldsymbol{x})$ is continuous and each level curve $(U = u) := \{\boldsymbol{x} : U(\boldsymbol{x}) = u\}$ is a closed and convex curve. This guarantees that f is monotonically decreasing in some sense. Let

$$\nabla U(\boldsymbol{x}) = (U^{(1,0)}(\boldsymbol{x}), U^{(0,1)}(\boldsymbol{x}))$$

be the gradient of $U(\boldsymbol{x})$ at $\boldsymbol{x}$, which is assumed to be continuous. Additional smoothness assumptions on $U(\boldsymbol{x})$ will be described along the way.

For points $\boldsymbol{x}$ for which $U(\boldsymbol{x})$ is large, we define the transformation $\boldsymbol{x} \to (\ell, u)$ where $u = U(x)$ and ℓ is determined in the following manner. We will use $(U = w)$ for some w as the reference curve, in the following sense. If $\boldsymbol{x}$ belongs to the reference level curve $(U = w)$, let $\ell(\boldsymbol{x})$ be the arc length from an arbitrary but fixed point $\boldsymbol{x}_o$ to $\boldsymbol{x}$, measured counterclockwise, say, on the curve $(U = w)$. In general, if $\boldsymbol{x}$ is on the level curve $(U = u)$, where $u \neq w$ then let $\ell(\boldsymbol{x}) = \ell(\boldsymbol{x}')$ where $\boldsymbol{x}'$ lies on $(U(\boldsymbol{x}) = w)$ and $\boldsymbol{x}$ and $\boldsymbol{x}'$ are connected by a curve which is orthogonal to each level curve between $(U = w)$ and $(U = u)$, in the sense that the normal of any intersecting level curve at the point of intersection is in the same direction (modulo π) as the tangent on the connecting curve at the point. The assumption that $\nabla(\boldsymbol{x})$ is continuous assures that the transformation $\boldsymbol{x} \to (\ell, u)$ is one-to-one. The fact that on the $\boldsymbol{x}$-plane, the set of $\boldsymbol{x}$ with ℓ fixed forms a line that cuts through the level curves orthogonally gives the Jacobian a relatively simple form. In what follows we refer to the index n in N_n as "time". For our computations, note that at time n when n is large, all the actions take place in a neighborhood of the level curve $(U = \log n)$ (see proof of 4). Hence at time n it is natural to set $w = \log n$ and use $(U = \log n)$ as the reference curve in defining ℓ. For this reason we shall throughout the rest of this paper adhere to this time-varying, locally orthogonal coordinate system. In doing so, here and elsewhere, the reference of ℓ to time will be suppressed (i.e. instead of ℓ_n we simply write ℓ) for convenience of notation. Also and whenever there is no ambiguity the notation $\ell, \boldsymbol{x}$ will denote the functions $\ell(\boldsymbol{x}), \boldsymbol{x}(\ell, u)$ as well as their possible values.

We now derive the Jacobian from $\boldsymbol{x}$ to (u, ℓ). Note that the unit normal vector at $\boldsymbol{x}$ on the corresponding level curve is

$$\frac{\nabla U(\boldsymbol{x})}{|\nabla U(\boldsymbol{x})|}.$$

By the way in which the transformation is defined,

$$\frac{\partial \boldsymbol{x}}{\partial u} = \frac{1}{|\nabla U(\boldsymbol{x})|} \cdot \frac{\nabla U(\boldsymbol{x})}{|\nabla U(\boldsymbol{x})|} = \frac{\nabla U(\boldsymbol{x})}{|\nabla U(\boldsymbol{x})|^2}. \tag{15}$$

We now derive $(\frac{\partial x_1}{\partial \ell}, \frac{\partial x_2}{\partial \ell})$. Clearly the unit tangent vector at $\boldsymbol{x}$ on the corresponding level curve is

$$\frac{1}{|\nabla U(\boldsymbol{x})|} \left(U^{(0,1)}(\boldsymbol{x}), -U^{(1,0)}(\boldsymbol{x})\right).$$

If $\boldsymbol{x}$ is such that $U(\boldsymbol{x}) = \log n$ then ℓ corresponds to the actual arc length and so

$$\frac{\partial \boldsymbol{x}}{\partial \ell} = \text{ unit tangent at } \boldsymbol{x} = \frac{1}{|\nabla U(\boldsymbol{x})|}\left(U^{(0,1)}(\boldsymbol{x}), -U^{(1,0)}(\boldsymbol{x})\right),$$

in which case the Jacobian is equal to

$$\left|\frac{\partial \boldsymbol{x}}{\partial(\ell, u)}\right| = \left|\frac{\partial x_1}{\partial u}\frac{\partial x_2}{\partial \ell} - \frac{\partial x_1}{\partial \ell}\frac{\partial x_2}{\partial u}\right| = \frac{1}{|\nabla U(\boldsymbol{x})|}.$$

In general, the arc length from $\boldsymbol{x}(\ell_1, u)$ to $\boldsymbol{x}(\ell_2, u)$ is computed as

$$\int_{\ell_1}^{\ell_2} |\boldsymbol{x}^{(1,0)}(\ell, u)| d\ell. \tag{16}$$

Thus, if ℓ changes by $d\ell$ then the arc length changes by $|\boldsymbol{x}^{(1,0)}(\ell, u)|d\ell$ and therefore

$$\frac{\partial \boldsymbol{x}}{\partial \ell} = \frac{|\boldsymbol{x}^{(1,0)}(\ell, u)|}{|\nabla U(\boldsymbol{x})|}\left(U^{(0,1)}(\boldsymbol{x}), -U^{(1,0)}(\boldsymbol{x})\right).$$

Consequently the Jacobian is

$$\left|\frac{\partial \boldsymbol{x}}{\partial(\ell, u)}\right| = \frac{|\boldsymbol{x}^{(1,0)}(\ell, u)|}{|\nabla U(\boldsymbol{x}(\ell, u))|}. \tag{17}$$

For convenience write

$$\xi(\ell, u) = |\nabla U(\boldsymbol{x}(\ell, u))| \tag{18}$$

and

$$\lambda(u) = \text{ length of the level curve } (U = u). \tag{19}$$

Thus, changing variables $\boldsymbol{x} \to (\ell, u)$ gives

$$\int_A e^{-nF(S(\boldsymbol{x};r))} dF(\boldsymbol{x}) = \int_{\tilde{A}} e^{-nF(S(\boldsymbol{x}(\ell,u);r))} e^{-u} \frac{|\boldsymbol{x}^{(1,0)}(\ell, u)|}{\xi(\ell, u)} d\ell du, \tag{20}$$

since $dF(\boldsymbol{x}) = e^{-U(\boldsymbol{x})} d\boldsymbol{x}$, where $\tilde{A}$ is the image of A under the transformation.

4. Main results. In this section we formulate conditions directly in terms of the logarithm of the density $f(\boldsymbol{x}) = e^{-U(\boldsymbol{x})}$ which lead to convergence of the longest edges of the nearest neighbor graph and the minimal spanning tree. As stated in the beginning of Section 3, we focus on functions U for which the gradient is continuous and the level curves are convex. Additional conditions on smoothness and other aspects of U will be stated below in conditions (A1)-(A6).

The core of the problem is to find sequences r_n for which the mean number of points with no r_n-nearest neighbors converges, i.e. which satisfy $\lim_{n\to\infty}\mu_n^{(1)}(\Re^2, r_n) =$ some $\tau\in(0,\infty)$. Denote by $\ell_o=\ell_{n,o}$ any point for which

$$\xi_n := \xi(\ell_o, \log n) = \min_\ell \xi(\ell, \log n).$$

To determine r_n we assume there exists finite positive constants c_1, c_2 and a sequence η_n which satisfies (A5) below, and put

$$r_n := \frac{\eta_n - c_1\log\eta_n - \log(c_2^{-1}\tau\sqrt{2\pi})}{\xi_n}. \tag{21}$$

Condition (A5) is chosen for technical convenience, and in particular is aimed at a second order Taylor expansion of the exponent in the important areas where $\xi(\ell,\log n)$ is close to it's minimum value.

For convenience write $\log_2 = \log\log$ and $\log_3 = \log\log\log$. The following set of assumptions are needed:

(A1) $\frac{U(\boldsymbol{x})}{|\nabla U(\boldsymbol{x})|}$ is uniformly continuous.

(A2) $\limsup_{u\to\infty}\frac{\sup_{\boldsymbol{x}:U(\boldsymbol{x})=u}|\nabla U(\boldsymbol{x})|}{\inf_{\boldsymbol{x}:U(\boldsymbol{x})=u}|\nabla U(\boldsymbol{x})|}<\infty$.

(A3) Both $\lambda(u)$ and $\inf_{\boldsymbol{x}:U(\boldsymbol{x})=u}|\nabla U(\boldsymbol{x})|$ belong to the range $[u^{-\rho}/C, Cu^{\rho}]$ for some $C,\rho\in(0,\infty)$.

(A4) $\lim_{U(\boldsymbol{x})\to\infty}\frac{|U^{(i,j)}(\boldsymbol{x})|}{|\nabla U(\boldsymbol{x})|^2}(\log U(\boldsymbol{x}))^2 = 0$ for any $i,j\geq 0$ with $i+j=2$.

(A5) There exist constants c_1, c_2 and a sequence of constants $\eta_n\to\infty$ with $\eta_n = O(\log_2 n)$ such that (with $\xi(\ell,\log n) = |\nabla U(\boldsymbol{x}(\ell,\log n))|$, as before)

   (a) $\lim_{n\to\infty} e^{-\eta_n}\eta_n^{c_1-1/2}\int_{\ell=0}^{\lambda(\log n)} e^{-[\xi(\ell,\log n)-\xi(\ell_o,\log n)]r_n}[\xi(\ell,\log n)\xi(\ell_o,\log n)]^{1/2}d\ell = c_2$,

   (b) $\lim_{n\to\infty} e^{-\eta_n}\eta_n^{c_1-1/2}\sup_\ell\int_{t=\ell}^{\ell+r_n} e^{-[\xi(t,\log n)-\xi(\ell_o,\log n)]r_n}[\xi(t,\log n)\xi(\ell_o,\log n)]^{1/2}dt = 0$.

(A6) $\lim_{u\to\infty}\inf_{\boldsymbol{x}:U(\boldsymbol{x})>u}\langle \boldsymbol{x}/|\boldsymbol{x}|, \nabla U(\boldsymbol{x})/|\nabla U(\boldsymbol{x})|\rangle > 0$.

In the following theorems let N_n be a positive integer-valued random variable such that $N_n/n \xrightarrow{p} 1$. Also define

$$\mu_n = \frac{\eta_n - c_1\log\eta_n - \log c_2^{-1}\sqrt{2\pi}}{\xi_n} \text{ and } \sigma_n = 1/\xi_n, \tag{22}$$

and denote by

$$\Lambda(x) = e^{-e^{-x}}$$

the Gumbel distribution function.

Theorem 4 *Assume that the conditions (A1)-(A5) hold. Then*

$$(M_{\mathrm{NNG}(N_n)} - \mu_n)/\sigma_n \xrightarrow{d} \Lambda.$$

Theorem 5 *Assume that the conditions (A1)-(A6) hold. Then*

$$(M_{\mathrm{MST}(N_n)} - \mu_n)/\sigma_n \xrightarrow{d} \Lambda.$$

Clearly, if the conclusions of the theorems hold then by weak convergence one has some flexibility in choosing the normalization; indeed, any $\tilde{\eta}_n$ and $\tilde{\xi}_n$ with

$$\tilde{\eta}_n - \eta_n \to 0 \text{ and } \tilde{\xi}_n/\xi_n \to 1$$

can replace η_n and ξ_n.

The set of conditions (A1)-(A6) are designed to meet not only analytic but also geometric considerations in the proofs, but still are quite general. To understand what they mean and how to verify them, we proceed to examples in the next two sections.

5. Homogeneous exponents. The first general example considers distributions with exponents which are homogeneous in the sense that $f(\boldsymbol{x}) = e^{-U(\boldsymbol{x})} = e^{-V(\boldsymbol{x})-c}$ with $V(k\boldsymbol{x}) = k^\alpha V(\boldsymbol{x})$, where the level curves for V are convex. In polar coordinates, $\boldsymbol{x} = (r\cos\theta, r\sin\theta)$ we may then write $V(\boldsymbol{x}) = r^\alpha g(\theta)$, so that

$$f(\boldsymbol{x}) = e^{-U(\boldsymbol{x})} = e^{-r^\alpha g(\theta)-c}. \tag{23}$$

Straightforward computations show that

$$|\nabla U(\boldsymbol{x})| = r^{\alpha-1}[\alpha^2 g(\theta)^2 + g'(\theta)^2]^{1/2} =: r^{\alpha-1}h(\theta). \tag{24}$$

Defining $v_c(n) = \log n - c$, we have that $U(\boldsymbol{x}) = \log n$ is equivalent to $V(\boldsymbol{x}) = v_c(n)$ and hence, with obvious notation, $r(\theta, \log n) = v_c(n)^{1/\alpha} g(\theta)^{-1/\alpha}$, and

$$\xi(\theta, \log n) = |\nabla U(\boldsymbol{x}(\theta, \log n))| = v_c(n)^{1-1/\alpha} g(\theta)^{-1+1/\alpha} h(\theta) =: v_c(n)^{1-1/\alpha} k(\theta). \tag{25}$$

The final ingredient needed to treat distributions of the form (23) is to note that $\frac{\partial \ell(\theta)}{\partial\theta} = |\frac{\partial \boldsymbol{x}(\theta)}{\partial\theta}|$ and use

$$x(\theta, \log n) = r(\theta, \log n)(\cos\theta, \sin\theta) = v_c(n)^{1/\alpha} g(\theta)^{-1/\alpha}(\cos\theta, \sin\theta)$$

to obtain

$$\frac{\partial \ell(\theta)}{\partial \theta} = v_c(n)^{1/\alpha} g(\theta)^{-1/\alpha}\left[\left(\frac{g'(\theta)}{\alpha g(\theta)}\right)^2 + 1\right]^{1/2} =: v_c(n)^{1/\alpha} m(\theta). \tag{26}$$

Theorem 6 *Assume that the density is of the form (23) with $\alpha > 1$ and $g(\theta)$ bounded away from zero and three times continuously differentiable on the torus $[0, 2\pi]$. Suppose further that $k(\theta)$ assumes its minimum value at d distinct points $\theta_o, \ldots, \theta_{d-1}$. Then the conditions (A1)-(A6) hold with $\xi_n/\tilde{\xi}_n \to 1$ for $\tilde{\xi}_n = (\log n)^{1-1/\alpha} k(\theta_0)$, $\eta_n = \log_2 n$, $c_1 = 1$ and*

$$c_2 = \sqrt{2\pi} k(\theta_0)^{3/2} \sum_{i=0}^{d-1} k''(\theta_i)^{-1/2} m(\theta_i).$$

Proof: Together, (23) and (25) can be seen to imply that $U(\boldsymbol{x})/|\nabla U(\boldsymbol{x})|$ has uniformly bounded derivatives, and hence (A1) holds. That (A2) and (A3) hold are immediate consequences of (25). It may be seen that $U^{(i,j)} = V^{(i,j)} = r^{\alpha-(i+j)} g_{i,j}(\theta)$ for $i+j = 2$ and suitable functions $g_{i,j}(\theta)$. Assumption (A4) follows from this and (25). We leave it to the reader to check that (A6) holds.

We now verify (A5). Using in order (26),(25), and Erdelyi (1956),

$$\begin{aligned}
I_n &:= \int_\ell e^{-[\xi(\ell,\log n)-\xi(\ell_o,\log n)]r_n} [\xi(\ell, \log n)\xi(\ell_o, \log n)]^{1/2} d\ell \\
&= \int_0^{2\pi} e^{-[\xi(\theta,\log n)-\xi(\theta_o,\log n)]r_n} [\xi(\theta, \log n)\xi(\theta_0, \log n)]^{1/2} v_c(n)^{1/\alpha} m(\theta) d\theta \\
&= v_c(n) \int_0^{2\pi} e^{-r_n v_c(n)^{1-1/\alpha}[k(\theta)-k(\theta_0)]} [k(\theta)k(\theta_0))]^{1/2} m(\theta) d\theta \\
&\sim v_c(n) k(\theta_0) \sqrt{2\pi} \sum_{i=0}^{d-1} (r_n v_c(n)^{1-1/\alpha} k''(\theta_i))^{-1/2} m(\theta_i).
\end{aligned}$$

Clearly $v_c(n) \sim \log n$ and $r_n v_c(n)^{1-1/\alpha} k(\theta_0) = r_n \xi(\ell_0, \log n) \sim \eta_n$. Hence

$$I_n \sim \sqrt{2\pi} (\log n) k(\theta_0)^{3/2} \eta_n^{-1/2} \sum_{i=0}^{d-1} (k''(\theta_i))^{-1/2} m(\theta_i)$$

and (A5a) follows. Similar considerations prove (A5b).

□

We next apply the result to classes of elliptically contoured distributions and to Weibull-type distributions.

Example 1: Consider elliptically contoured distributions with log density

$$U(\boldsymbol{x}) = r^\alpha((\cos\theta)^2 - 2\rho\cos\theta\sin\theta + (\sin\theta)^2)/d + c.$$

In particular this includes the bivariate normal with standardized marginals and correlation $\rho \in (-1,1)$ which is obtained for $\alpha = 2$, $c = -\log(2\pi(1-\rho^2)^{1/2})$, and $d = 2(1-\rho^2)$. These distributions are of the form (23) with

$$g(\theta) = ((\cos\theta)^2 - 2\rho\cos\theta\sin\theta + (\sin\theta)^2)/d = (1 - 2\rho\cos\theta\sin\theta)/d,$$

and hence with

$$k(\theta) = (1 - 2\rho\cos\theta\sin\theta)^{-1+\frac{1}{\alpha}} \sqrt{\alpha^2\left(1 - 2\rho\cos\theta\sin\theta\right)^2 + \left(2\rho - 4\rho\cos^2\theta\right)^2}/d^{1/\alpha}.$$

The conditions of Theorem 6 are satisfied, and hence the results of Theorems 4 and 5 hold with

$$\mu_n = \frac{\log_2 n - \log_3 n - \log(c_2^{-1}\sqrt{2\pi})}{(\log n)^{1-1/\alpha}k(\theta_0)} \text{ and } \sigma_n = \frac{1}{(\log n)^{1-1/\alpha}k(\theta_0)}. \tag{27}$$

Computation of $k(\theta_0)$ and c_2 require computer algebra, with a slight simplification obtained by noting that the constant c_2 is independent of the value of d. We consider the normal case, which has $\alpha = 2$ and $d = 2(1-\rho^2)$, and present the results for a few values of ρ in the table below. By symmetry, the values for $-\rho$ are the same as for ρ.

ρ	0.1	0.3	0.5	0.7	0.9
$k(\theta_0)$	1.348837	1.240718	1.153867	1.084742	1.026251
c_2	19.2383720	7.9460116	4.0933240	1.9501568	0.5414317

Example 2: In this example we consider independent marginals with Weibull-type densities $const \times e^{-|x_1|^\alpha - |x_2|^\alpha}$. The log density then is

$$U(\boldsymbol{x}) = r^\alpha(|\cos\theta|^\alpha + |\sin\theta|^\alpha) + c$$

and is of the form (23) with

$$g(\theta) = (|\cos\theta|^\alpha + |\sin\theta|^\alpha).$$

Hence,

$$\begin{aligned} k(\theta) \;=\; & \alpha\left(|\cos\theta|^\alpha + |\sin\theta|^\alpha\right)^{1-1/\alpha} \\ & \left\{\left(|\cos\theta|^\alpha + |\sin\theta|^\alpha\right)^2 \left(-\mathrm{sign}(\cos\theta)\sin\theta\,|\cos\theta|^{\alpha-1} + \mathrm{sign}(\sin\theta)\cos\theta\,|\sin\theta|^{\alpha-1}\right)^2\right\}^{1/2}. \end{aligned}$$

To assure three times differentiability we assume that $\alpha > 4$.

Again the conditions of Theorem 6 are satisfied, and hence the results of Theorems 4 and 5 hold with μ_n, σ_n given by (27) in the previous example. A few examples of values of $k(\theta_0), c_2$, obtained with Maple, are given in the table below.

α	5	6	7	8
$k(\theta_0)$	0.769	0.472	0.280	0.162
c_2	0.631	0.307	0.151	0.072

It might be noted that even if the result requires $\alpha > 4$, it can be shown to hold also for $2 < \alpha \leq 4$.

6. Parallel level curves. Consider the situation where the level curves are parallel, namely they are given by

$$\{\boldsymbol{x} : U(\boldsymbol{x}) = u\} = \{\boldsymbol{c}(t) + \omega(u)\boldsymbol{n}(t) : t \in [0, L)\}, \ u \geq \text{ some } u_o > 0, \tag{28}$$

where L is a finite positive constant, $\boldsymbol{c}(t)$ is a closed, strictly convex curve parameterized clockwise by length, $\omega(u)$ is an increasing function with $\omega(u_o) = 0$ and $\omega(\infty) = \infty$, and $\boldsymbol{n}(t)$ is the unit normal of $\boldsymbol{c}$ at the parameter t. Assume that that $\omega(u)$ is differentiable and $\boldsymbol{c}(t)$ is twice continuously differentiable with $|\ddot{\boldsymbol{c}}(t)| \in (0, \infty)$, where "$\cdot$" refers to differentiation with respect to t. Since $|\dot{\boldsymbol{c}}(t)| \equiv 1$, it is easy to see that

$$\dot{\boldsymbol{n}}(t) = |\ddot{\boldsymbol{c}}(t)|\dot{\boldsymbol{c}}(t).$$

Also it is clear that

$$\xi(u, \ell) = |\nabla U(\boldsymbol{x}(u, \ell))| \equiv \frac{1}{\omega'(u)} \tag{29}$$

so that the asymptotic results here are different in nature from what was considered in Theorem 6. Since $\xi(u, \ell)$ does not depend on ℓ we will denote it by $\xi(u)$ henceforth. Define $t(\ell)$ to be inverse function of

$$\begin{aligned} \ell(t) &= \text{arc length of the level curve } (u = \log n) \\ &\qquad\qquad \text{from } \boldsymbol{c}(0) + \omega(\log n)\boldsymbol{n}(0) \text{ to } \boldsymbol{c}(t) + \omega(\log n)\boldsymbol{n}(t) \\ &= \int_0^t |\dot{\boldsymbol{c}}(v) + \omega(\log n)\dot{\boldsymbol{n}}(v)|dv = \int_0^t [1 + \omega(\log n)|\ddot{\boldsymbol{c}}(v)|]dv. \end{aligned}$$

Hence,

$$\begin{aligned} \boldsymbol{x}^{(1,0)}(\ell, u) &= [\dot{\boldsymbol{c}}(t(\ell)) + \omega(u)\dot{\boldsymbol{n}}(t(\ell))]t'(\ell) \\ &= [1 + \omega(u)|\ddot{\boldsymbol{c}}(t(\ell))|]t'(\ell)\dot{\boldsymbol{c}}(t(\ell)) \\ &= \frac{1 + \omega(u)|\ddot{\boldsymbol{c}}(t(\ell))|}{1 + \omega(\log n)|\ddot{\boldsymbol{c}}(t(\ell))|}\dot{\boldsymbol{c}}(t(\ell)) \end{aligned} \tag{30}$$

Also

$$\begin{aligned} \lambda(u) &= \int_0^L [1 + \omega(u)|\ddot{\boldsymbol{c}}(t)|]dt \\ &= L + \omega(u)\int_0^L |\ddot{\boldsymbol{c}}(t)|dt \sim \omega(u)\int_0^L |\ddot{\boldsymbol{c}}(t)|dt \end{aligned} \tag{31}$$

Theorem 7 *Assume that* $\omega(u) = u^\alpha \exp\left(\int_{y_o}^u \frac{a(y)}{y} dy\right)$ *where* $\alpha > 0, a(y) \to 0$ *and* $ya'(y) \to 0$ *as* $y \to \infty$. *Then the condition (A1)-(A6) hold with* $\eta_n = \log[\xi(\log n)\lambda(\log n)], c_1 = 1/2$ *and* $c_2 = 1$.

Proof: Due to (29), all the conditions are easily verified, except perhaps (A4) which we now prove. Solving $\left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}\right)$ and $\left(\frac{\partial t}{\partial x_1}, \frac{\partial t}{\partial x_2}\right)$ in the following in terms of x_1, x_2,

$$\begin{aligned}
1 &= \frac{\partial x_1}{\partial x_1} = (\dot{c}_1 + \omega \dot{n}_1)\frac{\partial t}{\partial x_1} + \omega' n_1 \frac{\partial u}{\partial x_1} \\
0 &= \frac{\partial x_1}{\partial x_2} = (\dot{c}_1 + \omega \dot{n}_1)\frac{\partial t}{\partial x_2} + \omega' n_1 \frac{\partial u}{\partial x_2} \\
0 &= \frac{\partial x_2}{\partial x_1} = (\dot{c}_2 + \omega \dot{n}_2)\frac{\partial t}{\partial x_1} + \omega' n_2 \frac{\partial u}{\partial x_1} \\
1 &= \frac{\partial x_2}{\partial x_2} = (\dot{c}_2 + \omega \dot{n}_2)\frac{\partial t}{\partial x_2} + \omega' n_2 \frac{\partial u}{\partial x_2},
\end{aligned}$$

we get

$$\left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}\right) = \frac{1}{\omega'}(-\dot{c}_2, \dot{c}_1) \quad \text{and} \quad \left(\frac{\partial t}{\partial x_1}, \frac{\partial t}{\partial x_2}\right) = \frac{1}{1 + \omega|\ddot{\boldsymbol{c}}|}(\dot{c}_1, \dot{c}_2).$$

Hence

$$\begin{aligned}
\frac{\partial^2 u}{\partial x_1^2} &= -\frac{\omega^{(2)}\dot{c}_2(\partial u/\partial x_1) - \omega' \ddot{c}_2(\partial t/\partial x_1)}{(\omega')^2} = \frac{\omega^{(2)}\dot{c}_2^2}{(\omega')^3} + \frac{\dot{c}_1\ddot{c}_2}{\omega'(1 + \omega|\ddot{\boldsymbol{c}}|)} \\
\frac{\partial^2 u}{\partial x_2^2} &= \frac{\omega^{(2)}\dot{c}_1(\partial u/\partial x_2) - \omega' \ddot{c}_1(\partial t/\partial x_2)}{(\omega')^2} = \frac{\omega^{(2)}\dot{c}_1^2}{(\omega')^3} + \frac{\ddot{c}_1\dot{c}_2}{\omega'(1 + \omega|\ddot{\boldsymbol{c}}|)} \\
\frac{\partial^2 u}{\partial x_1 \partial x_2} &= \frac{\omega^{(2)}\dot{c}_1(\partial u/\partial x_1) - \omega' \ddot{c}_1(\partial t/\partial x_1)}{(\omega')^2} = -\frac{\omega^{(2)}\dot{c}_1\dot{c}_2}{(\omega')^3} - \frac{\dot{c}_1\ddot{c}_1}{\omega'(1 + \omega|\ddot{\boldsymbol{c}}|)}.
\end{aligned}$$

To show (A4), by the computations above and the fact that $|\dot{\boldsymbol{c}}|$ and $|\ddot{\boldsymbol{c}}|$ are bounded, it suffices to show that

$$\lim_{u\to\infty} \left(\frac{|\omega^{(2)}(u)|}{[\omega'(u)]^3} + \frac{1}{\omega(u)\omega'(u)}\right)(\omega'(u)\log u)^2 = 0. \tag{32}$$

The assumption of the theorem implies that both ω and ω' are regularly varying and in fact

$$\omega'(u) = (\alpha + a(y_o))u^{\alpha-1} \exp\left(\int_{y_o}^u \frac{\tilde{a}(y)}{y} dy\right)$$

where

$$\tilde{a}(y) = a(y) + \frac{ya'(y)}{\alpha + a(y)} \to 0 \text{ as } y \to \infty.$$

It is easily seen that

$$\frac{\omega'(u)}{\omega(u)} = \frac{\alpha + a(u)}{u} \text{ and } \frac{\omega^{(2)}(u)}{\omega'(u)} = \frac{\alpha - 1 + \tilde{a}(u)}{u},$$

from which (32) is straightforward.

□

Example 3: An example of such a distribution is the bivariate normal distribution with independent standard normal marginals, in which case we can take $\boldsymbol{c}(t) = \boldsymbol{n}(t) = (\cos t, \sin t), t \in [0, 2\pi)$, and

$$\omega(u) = \sqrt{2[u - \log(2\pi)] - 1},\ u \geq \log(2\pi) + 1/2.$$

The conditions of Theorem 7 are satisfied and so the conditions (A1)-(A6) hold with

$$\xi_n = (2\log n)^{1/2}, \eta_n = \log 4\pi + \log_2 n, c_1 = 1/2, \text{ and } c_2 = 1.$$

Hence the results of Theorems 4 and 5 hold with

$$\mu_n = \frac{\log_2 n - (1/2)\log_3 n + \log(2\sqrt{2\pi})}{(2\log n)^{1/2}} \text{ and } \sigma_n = \frac{1}{(2\log n)^{1/2}}.$$

This is consistent with the normalization obtained in Penrose (1998).

□

7. Proofs. The proofs of Theorems 4 and 5 are based on three lemmas which are given at the end. The first one, Lemma 8 deduces three consequences of Assumption (A4) which are more directly usable in the proofs of the theorems. Lemma 9 takes one step towards evaluating integrals like (1) and the last one, Lemma 10, approximates $F(S(\boldsymbol{x}, r_n))$, and in particular shows that appropriate sectors of $S(\boldsymbol{x}, r_n)$ may be neglected asymptotically.

In addition to previous notation, we will use in the proofs

$$u_{n,-b} = \log n - b\log_2 n, \quad u_{n,b} = \log n + b\log_2 n,\ b > 0. \tag{33}$$

Proof of Theorem 4. First consider the case $N_n \sim \text{Poisson}(n)$. We start by defining the sets $A_{n,1}, \dots, A_{n,k_n}$ in Theorem 1. For convenience write

$$I_n(\ell) = \frac{\tau \int_{t=0}^{\ell} e^{-[\xi(t,\log n) - \xi(\ell_o,\log n)]r_n} \xi^{1/2}(t,\log n)dt}{\int_{t=0}^{\lambda(\log n)} e^{-[\xi(t,\log n) - \xi(\ell_o,\log n)]r_n} \xi^{1/2}(t,\log n)dt}, \quad \ell \in [0, \lambda(\log n)).$$

Thus, by (53) of Lemma 9 with $k = 0$ and $\delta_n \equiv 1$, for some $b > 0$ we have uniformly for $0 \le \ell_{n,1} < \ell_{n,2} < \lambda(\log n)$,

$$\mu_n^{(1)}\{\boldsymbol{x}(\ell, u) : \ell \in [\ell_{n,1}, \ell_{n,2}), u \in [u_{n,-b}, u_{n,b}]\} \sim I_n(\ell_{n,2}) - I_n(\ell_{n,1}). \tag{34}$$

By (A5a),

$$\varepsilon_n := \sup_\ell [I_n(\ell + r_n) - I_n(\ell)] \to 0.$$

Let k_n be a sequence of integers such that

$$k_n \to \infty \text{ and } k_n\varepsilon_n \to 0,$$

and let j_i be the smallest positive integer j such that $I_n(jr_n)$ exceeds i/k_n, for $i = 1, \dots, k_n$. Define

$$A_{n,i} = \{\boldsymbol{x}(\ell, u) : \ell \in [j_{i-1}r_n, (j_i - 3)r_n), u \in [u_{n,-b}, u_{n,b}]\} \quad 1 \le i \le k_n$$

and

$$A_{n,k_n+1} = A_{n,1}.$$

Also let

$$\bar{A}_{n,i} = \{\boldsymbol{x}(\ell, u) : \ell \in [(j_i - 3)r_n, j_i r_n), u \in [u_{n,-b}, u_{n,b}]\} \quad 1 \le i \le k_n - 1$$

and

$$\bar{A}_{n,k_n} = \{\boldsymbol{x}(\ell, u) : \ell \in [(j_{k_n} - 3)r_n, \lambda(\log n)), u \in [u_{n,-b}, u_{n,b}]\}.$$

By the choice of the sets, it follows from (34) that

$$1/k_n - 4\varepsilon_n \le \mu_n^{(1)}(A_{n,i}, r_n) \le 1/k_n + \varepsilon_n, \quad 1 \le i \le k_n, \tag{35}$$

$$\max_{1 \le i \le k_n - 1} \mu_n^{(1)}(\bar{A}_{n,i}, r_n) \le 3\varepsilon_n \quad \text{and} \quad \mu_n^{(1)}(\bar{A}_{n,k_n}, r_n) \le 3\varepsilon_n + 1/k_n, \tag{36}$$

and hence that

$$\max_{1 \le i \le k_n - 1} \frac{\mu_n^{(1)}(\bar{A}_{n,i}, r_n)}{\mu_n^{(1)}(A_{n,i}, r_n)} \le \frac{3\varepsilon_n}{1/k_n - 4\varepsilon_n} \to 0. \tag{37}$$

We now proceed to verify the conditions (a)-(e) of Theorem 1.

By convexity,

$$c_{n,i} := \inf(|\boldsymbol{x} - \boldsymbol{y}| : \boldsymbol{x} \in A_{n,i}, \boldsymbol{y} \in A_{n,i+1}) = |\boldsymbol{x}((j_i - 3)r_n, u_{n,-b}) - \boldsymbol{x}(j_i r_n, u_{n,-b})|.$$

By (16) and the condition (A4c) of Lemma 8 the length $a_{n,i}$ of the arc that connects $\boldsymbol{x}((j_i - 3)r_n, u_{n,-b})$ with $\boldsymbol{x}(j_i r_n, u_{n,-b})$ on $\{\boldsymbol{x} : U(\boldsymbol{x}) = u_{n,-b}\}$ is asymptotically $3r_n$. Observe that $1 \leq a_{n,i}/c_{n,i} \leq 1/\cos\theta_{n,i}$ where $\theta_{n,i}$ is the angle between $\frac{\nabla U(\boldsymbol{x}((j_i-3)r_n,u_{n,-b}))}{|\nabla U(\boldsymbol{x}((j_i-3)r_n,u_{n,-b}))|}$ and $\frac{\nabla U(\boldsymbol{x}(j_i r_n,u_{n,-b}))}{|\nabla U(\boldsymbol{x}(j_i r_n,u_{n,-b}))|}$. The condition (A1) guarantees that $\max_i \theta_{n,i} \to 0$. Consequently, $\min_i c_{n,i} > 2r_n$ for large n. Hence we have shown that adjacent $A_{n,i}$'s are at least $2r_n$ apart for large n. That non-adjacent $A_{n,i}$'s are at least $2r_n$ apart is also easy to establish by convexity. Hence the condition (a) of Theorem 1 is proved.

It follows from (53) of Lemma 9 with $\delta_n \equiv 1$, $k = 0$, $\ell_{n,1} = 0$ and $\ell_{n,2} = \lambda(\log n)$ that the condition (b) of Theorem 1 holds. To prove (c) of Theorem 1, note that it suffices to show that

$$\mu_n^{(1)}\{\boldsymbol{x}(\ell, u) : u \notin [u_{n,-b}, u_{n,b}]\} \to 0 \text{ and } \sum_{j=1}^{k_n} \mu_n^{(1)}(\bar{A}_{n,i}, r_n) \to 0.$$

The first convergence follows from (52) of Lemma 9 with $\delta_n \equiv 1$ and $k = 0$, while the the second convergence follows from (36) and (37), in conjunction with the condition (b) already proved above. Condition (d) of Theorem 1 holds by (35).

Finally we prove (e) of Theorem 1. By the simple inequality,

$$P(A \cup B) \geq (P(A) + P(B))/2$$

we have

$$\mu_n^{(2)}(A, r_n) \leq \left(n \int_A e^{-nF(S(\boldsymbol{x};r))/2} dF(\boldsymbol{x})\right)^2$$

so that

$$\frac{\mu_n^{(2)}(A_{n,i}, r_n)}{\mu_n^{(1)}(A_{n,i}, r_n)} \leq \frac{\left(n \int_{A_{n,i}} e^{-nF(S(\boldsymbol{x};r))/2} dF(\boldsymbol{x})\right)^2}{n \int_{A_{n,i}} e^{-nF(S(\boldsymbol{x};r))} dF(\boldsymbol{x})}.$$

By (53) of Lemma 9 with $\delta_n \equiv 1/2, 1$ and $k = 0$, and (35),

$$\frac{\left(n \int_{A_{n,i}} e^{-nF(S(\boldsymbol{x};r))/2} dF(\boldsymbol{x})\right)^2}{n \int_{A_{n,i}} e^{-nF(S(\boldsymbol{x};r))} dF(\boldsymbol{x})} \sim 4\tau\mu_n^{(1)}(A_{n,i}, r_n) \to 0$$

uniformly in i. Hence (e) of Theorem 1 is proved. Replacing τ by e^{-x} completes the proof for the case $N_n \sim$ Poisson (n).

Next,

$$\begin{aligned}
& P(M_{\mathrm{NNG}(N_n)} \leq r_n) \\
\leq\; & P(M_{\mathrm{NNG}(n)} \leq r_n) + P(M_{\mathrm{NNG}(N_n)} \leq r_n < M_{\mathrm{NNG}(n)}) \\
\leq\; & P(M_{\mathrm{NNG}(N_n)} \leq r_n) + P(M_{\mathrm{NNG}(n)} \leq r_n < M_{\mathrm{NNG}(N_n)}) \\
& + P(M_{\mathrm{NNG}(N_n)} \leq r_n < M_{\mathrm{NNG}(n)}).
\end{aligned}$$

It follows from (52) of Lemma 9 with $k = 0, 1$ and $\delta_n \equiv 1$ that (11) and (12) hold. Hence, it follows from Theorems 2 and the first part of the proof with $N_n \sim$ Poisson (n) by taking limits throughout the preceding inequalities that

$$\lim_{n\to\infty} P(M_{\mathrm{NNG}(n)} \le r_n) = \lim_{n\to\infty} P(M_{\mathrm{NNG}(N_n)} \le r_n) = e^{-\tau}.$$

Applying the above argument again gives

$$\lim_{n\to\infty} P(M_{\mathrm{NNG}(N_n)} \le r_n) = \lim_{n\to\infty} P(M_{\mathrm{NNG}(n)} \le r_n) = e^{-\tau}$$

for any general N_n satisfying $N_n/n \xrightarrow{p} 1$. Replacing τ by e^{-x} completes the proof. $\square$

Proof of Theorem 5. If follows from Theorem 3 and Theorem 4 that

$$\lim_{n\to\infty} P(M_{\mathrm{MST}(N_n)} \le r_n) = \lim_{n\to\infty} P(M_{\mathrm{NNG}(N_n)} \le r_n) = e^{-\tau}$$

provided we show (13). To do that, it suffices to show that

$$n \int \bar{F}^{\delta_n n}(S(\boldsymbol{x}; r_n)) I(|\boldsymbol{x}| > r_n/2) dF(\boldsymbol{x}) \to \tau, \tag{38}$$

and

$$n \int \bar{F}^{\delta_n n}(S(\boldsymbol{x}; r_n) \cap S(0; |\boldsymbol{x}|)) I(|\boldsymbol{x}| > r_n/2) dF(\boldsymbol{x}) \to \tau \tag{39}$$

for any sequence $\delta_n \to 1$. By convexity and the non-intersection of level curves,

$$|\boldsymbol{x}(\ell, u_{n,-b})| \ge |\boldsymbol{x}(\ell, u_{n,-b}) - \boldsymbol{x}(\ell, u_{n,-b/2})|. \tag{40}$$

By (50) of Lemma 8 and arguments there, uniformly in ℓ,

$$|\boldsymbol{x}(\ell, u_{n,-b}) - \boldsymbol{x}(\ell, u_{n,-b/2})| = \left| \int_{u_{n,-b/2}}^{u_{n,-b}} \frac{\nabla U(\boldsymbol{x}(\ell, s))}{|\nabla U(\boldsymbol{x}(\ell, s))|^2} ds \right| \sim \frac{b \log_2 n}{2\xi(\ell, \log n)}. \tag{41}$$

Since $r_n \le \eta_n/\xi_n$ for large n where $\eta_n = O(\log_2 n)$, we conclude by (40), (41) and (A2) that

$$\liminf_{n\to\infty} \inf_{\ell} |\boldsymbol{x}(\ell, u_{n,-b})|/r_n \ge cb$$

for some finite constant c independent of b. Hence we can choose b sufficiently large to ensure that $\inf_\ell |\boldsymbol{x}(\ell, u_{n,-b})|/r_n \ge 1$. By this and (52) of Lemma 9 that for a large enough b,

$$n \int \bar{F}^{\delta_n n}(S(\boldsymbol{x}; r_n)) I(|\boldsymbol{x}| \le r_n/2) dF(\boldsymbol{x}) \le n \int_{\boldsymbol{x}: U(\boldsymbol{x}) < u_{n,-b}} \bar{F}^{\delta_n n}(S(\boldsymbol{x}; r_n)) dF(\boldsymbol{x}) \to 0$$

so that (38) follows from Lemma 9.

We next consider the proof of (39). We will do so by proving that for a $b > 0$, sufficiently large:

$$n \int_{\boldsymbol{x}:U(\boldsymbol{x})>u_{n,b}} \bar{F}^{\delta_n n}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|))I(|\boldsymbol{x}| > r_n/2)dF(\boldsymbol{x}) \to 0, \tag{42}$$

$$n \int_{\boldsymbol{x}:U(\boldsymbol{x})<u_{n,-b}} \bar{F}^{\delta_n n}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|))I(|\boldsymbol{x}| > r_n/2)dF(\boldsymbol{x}) \to 0, \tag{43}$$

and

$$n \int_{\boldsymbol{x}:u_{n,-b}\le U(\boldsymbol{x})\le u_{n,b}} \bar{F}^{\delta_n n}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|))I(|\boldsymbol{x}| > r_n/2)dF(\boldsymbol{x}) \to \tau. \tag{44}$$

Clearly (42) follows from (57) in the proof of Lemma 9. To deal with (43), first write

$$\begin{aligned} & n \int_{\boldsymbol{x}:U(\boldsymbol{x})<u_{n,-b}} \bar{F}^{\delta_n n}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|))I(|\boldsymbol{x}| > r_n/2)dF(\boldsymbol{x}) \\ = \; & n \int_{\boldsymbol{x}:U(\boldsymbol{x})<\frac{u_{n,-b}}{2}} \bar{F}^{\delta_n n}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|))I(|\boldsymbol{x}| > r_n/2)dF(\boldsymbol{x}) + \\ & + n \int_{\boldsymbol{x}:\frac{u_{n,-b}}{2}\le U(\boldsymbol{x})<u_{n,-b}} \bar{F}^{\delta_n n}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|))I(|\boldsymbol{x}| > r_n/2)dF(\boldsymbol{x}). \end{aligned}$$

We will show that both terms on the right tend to 0. Since $f(\boldsymbol{x}) = e^{-U(\boldsymbol{x})}$, by (A2), (A3) and the mean value theorem, there exists some $b_1 \in (0, \infty)$ such that

$$\inf_{\boldsymbol{y}\in S(\boldsymbol{x};r_n),U(\boldsymbol{x})<\frac{u_{n,-b}}{2}} f(\boldsymbol{y}) \ge e^{-\frac{u_{n,-b}}{2} - b_1\xi(\ell_o,\log n)r_n} \ge e^{-\frac{u_{n,-b}}{2} - b_1\eta_n}.$$

Also observe that for $\boldsymbol{x}$ such that $|\boldsymbol{x}| > r_n/2$,

$$\text{area}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|)) \ge b_2\pi r_n^2$$

for some $b_2 \in (0, 1)$. Hence, for all large n,

$$\inf_{\boldsymbol{x}:|\boldsymbol{x}|>r_n/2,U(\boldsymbol{x})<\frac{u_{n,-b}}{2}} F(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|)) \ge b_2\pi r_n^2 e^{-\frac{u_{n,-b}}{2} - b_1\eta_n}$$

and hence

$$\begin{aligned} & n \int_{\boldsymbol{x}:U(\boldsymbol{x})<\frac{u_{n,-b}}{2}} \bar{F}^{\delta_n n}(S(\boldsymbol{x};r_n) \cap S(0;|\boldsymbol{x}|))I(|\boldsymbol{x}| > r_n/2)dF(\boldsymbol{x}) \\ \le \; & n e^{-\delta_n n b_2 \pi r_n^2 e^{-\frac{u_{n,-b}}{2} - b_1\eta_n}} \to 0. \end{aligned} \tag{45}$$

By (A6), there exists some constant $\zeta \in (-1,0)$ such that for all $\boldsymbol{x}$ with $U(\boldsymbol{x}) > \frac{u_{n,-b}}{2}$,

$$S(\boldsymbol{x}; r_n) \cap \{\boldsymbol{y} : \langle \boldsymbol{y} - \boldsymbol{x}, \tfrac{\nabla U(\boldsymbol{x})}{|\nabla U(\boldsymbol{x})|} \rangle \le \zeta r_n\} \subset S(\boldsymbol{x}; r_n) \cap S(0; |\boldsymbol{x}|).$$

Hence

$$\begin{aligned}
& \inf_{\boldsymbol{x}: \frac{u_{n,-b}}{2} \le U(\boldsymbol{x}) < u_{n,-b}} F(S(\boldsymbol{x}; r_n) \cap S(0; |\boldsymbol{x}|)) \\
\ge & \inf_{\boldsymbol{x}: \frac{u_{n,-b}}{2} \le U(\boldsymbol{x}) < u_{n,-b}} F(S(\boldsymbol{x}; r_n) \cap \{\boldsymbol{y} : \langle \boldsymbol{y} - \boldsymbol{x}, \tfrac{\nabla U(\boldsymbol{x})}{|\nabla U(\boldsymbol{x})|} \rangle \le \zeta r_n\}) \qquad (46) \\
\ge & \inf_{\ell} F(S(\boldsymbol{x}(\ell, u_{n,-b}); r_n) \cap \{\boldsymbol{y} : \langle \boldsymbol{y} - \boldsymbol{x}(\ell, u_{n,-b}), \tfrac{\nabla U(\boldsymbol{x}(\ell, u_{n,-b}))}{|\nabla U(\boldsymbol{x}(\ell, u_{n,-b}))|} \rangle \le \zeta r_n\}).
\end{aligned}$$

It follows from Lemma 10 that uniformly for $u \in [u_{n,-b}, u_{n,b}]$ and ℓ,

$$F(S(\boldsymbol{x}(\ell, u); r_n) \cap \{\boldsymbol{x} : \langle \boldsymbol{x} - \boldsymbol{x}(\ell, u), \tfrac{\nabla U(\boldsymbol{x}(\ell,u))}{|\nabla U(\boldsymbol{x}(\ell,u))|} \rangle \le \zeta r_n\}) \sim F(S(\boldsymbol{x}(\ell, u); r_n)). \qquad (47)$$

By (46) and (47), the same proof that leads to (56) in Lemma 9 now proves

$$n \int_{\boldsymbol{x}: \frac{u_{n,-b}}{2} \le U(\boldsymbol{x}) < u_{n,-b}} \bar{F}^{\delta_n n}(S(\boldsymbol{x}; r_n) \cap S(0; |\boldsymbol{x}|)) I(r_n/2 < |\boldsymbol{x}| < u_{n,-b}) dF(\boldsymbol{x}) \to 0. \qquad (48)$$

Hence (43) follows from (45) and (48). Making use of (47), the proof of (44) mirrors that of (53) of Lemma 9 below and is omitted.

□

Lemma 8 *Assume (A4) holds, and consider the coordinate system based on the level curve* $U(\boldsymbol{x}) = \log n$. *Then, for any finite constant* $b > 0$,

(A4a)

$$\lim_{n \to \infty} \sup_{\boldsymbol{x}: |U(\boldsymbol{x}) - \log n| \le b \log_2 n} \frac{|U^{(i,j)}(\boldsymbol{x})|}{|\nabla U(\boldsymbol{x})|^2} |(\log_2 n)^2 = 0 \textit{ for any } i, j \textit{ with } i + j = 2,$$

(A4b)

$$\lim_{n \to \infty} \sup_{|v - \log n| \le b \log_2 n} \left| \frac{|\nabla U(\ell, v)|}{|\nabla U(\ell, \log n)|} - 1 \right| \log_2 n = 0,$$

(A4c)

$$\lim_{n \to \infty} \sup_{|v - \log n| \le b \log_2 n} ||\boldsymbol{x}^{(1,0)}(\ell, v)| - 1| = 0.$$

Proof: (A4a) is an immediate consequence of (A4) since $\frac{\log U(\boldsymbol{x})}{\log_2 n} \to 1$ in the indicated range.

The proof of parts (A4b) and (A4c) uses the relation (15), that $\frac{\partial \boldsymbol{x}}{\partial u} = \frac{\nabla U(\boldsymbol{x})}{|\nabla U(\boldsymbol{x})|^2}$, in two ways. First, this relation implies that

$$\left|\frac{\partial \boldsymbol{x}_1(\ell, v)}{\partial v}\right| \le \frac{1}{|\nabla U(\boldsymbol{x})|}, \text{ and } \left|\frac{\partial \boldsymbol{x}_2(\ell, v)}{\partial v}\right| \le \frac{1}{|\nabla U(\boldsymbol{x})|}, \tag{49}$$

and secondly it is equivalent to

$$\boldsymbol{x}(\ell, v) - \boldsymbol{x}(\ell, \log n) = \int_{\log n}^{v} \frac{\nabla U(\boldsymbol{x}(\ell, s))}{|\nabla U(\boldsymbol{x}(\ell, s))|^2} ds. \tag{50}$$

Let

$$M_n = \sup_{|v - \log n| \le b \log_2 n, i+j=2} \frac{|U^{(i,j)}(\boldsymbol{x}(\ell, v)|}{|\nabla U(\boldsymbol{x}(\ell, v)|^2}.$$

Now, to prove (A4b), note that

$$\begin{aligned} & \left|\frac{\partial |\nabla U(\boldsymbol{x}(\ell, v)|}{\partial v}\right| \\ = & \left|\frac{\partial}{\partial \boldsymbol{x}_1} |\nabla U(\boldsymbol{x}(\ell, v)| \frac{\partial \boldsymbol{x}_1}{\partial v} + \frac{\partial}{\partial \boldsymbol{x}_2} |\nabla U(\boldsymbol{x}(\ell, v)| \frac{\partial \boldsymbol{x}_2}{\partial v}\right| \\ = & \frac{| < \nabla U, (U^{(2,0)}, U^{(1,1)}) > \frac{\partial \boldsymbol{x}_1}{\partial v} + < \nabla U, (U^{(1,1)}, U^{(0,2)}) > \frac{\partial \boldsymbol{x}_2}{\partial v}|}{|\nabla U|.} \end{aligned} \tag{51}$$

Using (49), for $|v - \log n| \le b \log_2 n$, we hence have that $\frac{\partial \log |\nabla U(\boldsymbol{x}(\ell, v)|}{\partial v} \le 4M_n$, and by integration,

$$(e^{-4M_n b \log_2 n} - 1) \log_2 n \le \left|\frac{|\nabla(\boldsymbol{x}(\ell, v))|}{|\nabla(\boldsymbol{x}(\ell, \log n))|} - 1\right| \log_2 n \le (e^{4M_n b \log_2 n} - 1) \log_2 n$$

and (A4b) follows from (A4).

Next, to prove (A4c), we note that similar calculations as for (51) give that

$$\left|\frac{\partial}{\partial \ell} \frac{\nabla U(\boldsymbol{x}(\ell, v)}{|\nabla U(\boldsymbol{x}(\ell, v)|^2}\right| \le 5M_n |\boldsymbol{x}^{(1,0)}(\ell, v)|.$$

Hence, interchanging differentiation and integration in (50), we obtain

$$|\boldsymbol{x}^{(1,0)}(\ell, v) - \boldsymbol{x}^{(1,0)}(\ell, \log n)| \le 5M_n \int_{\log n}^{v} |\boldsymbol{x}^{(1,0)}(\ell, s)| ds.$$

By definition (cf. Section 3), $|\boldsymbol{x}^{(1,0)}(\ell,\log n)| = 1$, and hence

$$||\boldsymbol{x}^{(1,0)}(\ell,v)|-1| \le 5M_n \int_{\log n}^{v} (||\boldsymbol{x}^{(1,0)}(\ell,s)|-1|+1)ds.$$

It then follows from a Grönvall inequality that for $|v-\log n| \le b\log_2 n$,

$$||\boldsymbol{x}^{(1,0)}(\ell,v)|-1| \le 5M_n b\log_2 n \exp\{5M_n b\log_2 n\} \to 0,$$

and (A4c) follows.

□

Lemma 9 *Assume that the conditions (A1)-(A5) hold. Let constants $k\in[0,\infty)$ and $\delta_n \in [\underline{\delta},\overline{\delta}] \subset (0,\infty)$. Then there exists some $b\in(0,\infty)$ such that*

$$n\int_{\boldsymbol{x}} I(U(\boldsymbol{x}) \notin [u_{n,-b},u_{n,b}])e^{-\delta_n nF(S(\boldsymbol{x};r_n))}[nF(S(\boldsymbol{x};r_n)]^k dF(\boldsymbol{x}) \to 0, \tag{52}$$

and uniformly for $\ell_{n,1} < \ell_{n,2} \in [0,\lambda(\log n)]$,

$$\begin{aligned} & n\int_{\boldsymbol{x}} I(U(\boldsymbol{x}) \in [u_{n,-b},u_{n,b}],\ell(\boldsymbol{x})\in[\ell_{n,1},\ell_{n,2}])e^{-\delta_n nF(S(\boldsymbol{x};r_n))}[nF(S(\boldsymbol{x};r_n)]^k dF(\boldsymbol{x}) \\ \sim\quad & \delta_n^{-1}\tau\frac{\int_{\ell=\ell_{n,1}}^{\ell_{n,2}} e^{-[\xi(\ell,\log n)-\xi(\ell_o,\log n)]r_n}\xi^{1/2}(\ell,\log n)d\ell}{\int_{\ell=0}^{\lambda(\log n)} e^{-[\xi(\ell,\log n)-\xi(\ell_o,\log n)]r_n}\xi^{1/2}(\ell,\log n)d\ell}. \end{aligned} \tag{53}$$

Proof: In this proof we will assume for convenience of notation that $k=0$ and $\delta_n\equiv 1$, since the extension to the general case is straightforward. First note that by the condition (A2), (A4b) of Lemma 8 and the definition of r_n,

$$\sup_{u\in[u_{n,-b},u_{n,b}],\ell} |\xi(\ell,u)-\xi(\ell,\log n)|r_n \le \sup_{u\in[u_{n,-b},u_{n,b}],\ell} \left|\frac{\xi(\ell,u)}{\xi(\ell,\log n)}-1\right| O(\log_2 n) \to 0 \tag{54}$$

and

$$\inf_{u\in[u_{n,-b},u_{n,b}],\ell} \xi(\ell,u)r_n \ge \xi(\ell_o,\log n)r_n \to \infty,\ b>0.$$

By the conditions (A4) and (A2),

$$\lim_{n\to\infty} \sup_{\boldsymbol{x}:U(\boldsymbol{x})\in[u_{n,-b},u_{n,b}]} \sup_{\boldsymbol{y}\in S(\boldsymbol{x};r_n)} |U(\boldsymbol{y})-U(\boldsymbol{x})-\langle\nabla U(\boldsymbol{x}),\boldsymbol{y}-\boldsymbol{x}\rangle| = 0,\ b>0.$$

Hence, by Lemma 10 with $\zeta = 1$, (54) and (A4b) of Lemma 8, we have uniformly for all ℓ and $u \in [u_{n,-b}, u_{n,b}]$,

$$
\begin{aligned}
& F(S(\boldsymbol{x}(\ell, u); r_n)) \\
\sim\ & \sqrt{2\pi} e^{-u} \xi^{-3/2}(\ell, u) r_n^{1/2} e^{\xi(\ell,u) r_n} \\
\sim\ & \sqrt{2\pi} e^{-u} \xi^{-3/2}(\ell, \log n) r_n^{1/2} e^{\xi(\ell_o, \log n) r_n} e^{[\xi(\ell,\log n) - \xi(\ell_o, \log n)] r_n} \\
=\ & e^{-u} \chi_n(\ell)
\end{aligned}
\tag{55}
$$

where

$$\chi_n(\ell) = c_2 \tau^{-1} \xi^{-3/2}(\ell, \log n) r_n^{1/2} e^{\eta_n} \eta_n^{-c_1} e^{[\xi(\ell,\log n) - \xi(\ell_o,\log n)] r_n}.$$

By the non-intersection of level curves, the condition (A2), and (55), for all large n we have, for some constant $b_1, b_2 > 0$,

$$
\begin{aligned}
& \inf_{\boldsymbol{x}: U(\boldsymbol{x}) < u_{n,-b}} F(S(\boldsymbol{x}; r_n)) \\
\geq\ & \inf_{\ell} F(S(\boldsymbol{x}(\ell, u_{n,-b}); r_n)) \\
\geq\ & b_1 \inf_{\ell} e^{-u_{n,-b}} \xi^{-3/2}(\ell, \log n) r_n^{1/2} e^{\eta_n} \eta_n^{-c_1} \\
\geq\ & b_2 n^{-1} (\log n)^b \xi^{-2}(\ell_o, \log n) e^{\eta_n} \eta_n^{1/2 - c_1}.
\end{aligned}
$$

It then follows from the condition (A3) that

$$\inf_{\boldsymbol{x}: U(\boldsymbol{x}) < u_{n,-b}} F(S(\boldsymbol{x}; r_n)) \geq b_2 n^{-1} (\log n)^{b_3}$$

for some constants b_2, b_3, where b_3 can be picked large provided that b is large. Hence if b and b_3 are large enough we have

$$n \int_{\boldsymbol{x}: U(\boldsymbol{x}) < u_{n,-b}} e^{-nF(S(\boldsymbol{x}; r_n))} dF(\boldsymbol{x}) \leq n e^{-b_2 (\log n)^{b_3}} \to 0. \tag{56}$$

Next, by (16) and the assumption (A3),

$$
\begin{aligned}
n \int_{\boldsymbol{x}: U(\boldsymbol{x}) > u_{n,b}} dF(\boldsymbol{x}) &= n \int_{u > u_{n,b}} \int_{\ell=0}^{\lambda(\log n)} e^{-u} \frac{|\boldsymbol{x}^{(1,0)}(\ell, u)|}{\xi(\ell, u)} d\ell du \\
&\leq n \int_{u > u_{n,b}} e^{-u} \frac{\lambda(u)}{\xi(\ell_o, u)|} du \\
&\leq n \int_{u > u_{n,b}} e^{-u} u^{2\rho} du \\
&\sim n e^{-u_{n,b}} u_{n,b}^{2\rho} \to 0
\end{aligned}
\tag{57}
$$

Hence

$$n \int_{\boldsymbol{x}: U(\boldsymbol{x}) > u_{n,b}} e^{-nF(S(\boldsymbol{x}; r_n))} dF(\boldsymbol{x}) \to 0 \tag{58}$$

for a sufficiently large b. Now, (56) and (58) imply (52).

It follows from (20), (55), and (A4c) of Lemma 8 that

$$\begin{aligned} & n\int I(U(\boldsymbol{x})\in[u_{n,-b},u_{n,b}],\ell(\boldsymbol{x})\in[\ell_{n,1},\ell_{n,2}])e^{-nF(S(\boldsymbol{x};r))}dF(\boldsymbol{x}) \\ \sim\ & n\int_{u=u_{n,-b}}^{u_{n,b}}\int_{\ell=\ell_{n,1}}^{\ell_{n,2}} e^{-n(1+o(1))e^{-u}\chi_n(\ell)}e^{-u}\frac{1}{\xi(\ell,\log n)}d\ell du. \end{aligned}$$

Make a change of variables in the above integral with

$$v = u - \log n - \log\chi_n(\ell).$$

By the conditions (A2), (A3) and (A4), $\log\chi_n(\ell) = O(\log_2 n)$. As a result, it is possible to choose large enough b such that $v_{n,-b}\to-\infty$ and $v_{n,b}\to\infty$ where $v_{n,\pm b}$ corresponds to $u_{n,\pm b}$ with respect to the above variable change. Since $\int_{-\infty}^{\infty} e^{-e^{-v}}e^{-v}dv = 1$, we have

$$\begin{aligned} & n\int I(U(\boldsymbol{x})\in[u_{n,-b},u_{n,b}],\ell(\boldsymbol{x})\in[\ell_{n,1},\ell_{n,2}])e^{-nF(S(\boldsymbol{x};r))}dF(\boldsymbol{x}) \\ \sim\ & \int_{v=v_{n,-b}}^{v_{n,b}}\int_{\ell=\ell_{n,1}}^{\ell_{n,2}} e^{-(1+o(1))e^{-v}}e^{-v}\frac{1}{\xi(\ell,\log n)\chi_n(\ell)}d\ell dv \\ =\ & c_2^{-1}\tau r_n^{-1/2}e^{-\eta_n}\eta_n^{c_1}\int_{\ell=\ell_{n,1}}^{\ell_{n,2}} e^{-[\xi(\ell,\log n)-\xi(\ell_o,\log n)]r_n}\xi^{1/2}(\ell,\log n)d\ell \\ \sim\ & c_2^{-1}\tau e^{-\eta_n}\eta_n^{c_1-1/2}\int_{\ell=\ell_{n,1}}^{\ell_{n,2}} e^{-[\xi(\ell,\log n)-\xi(\ell_o,\log n)]r_n}[\xi(\ell,\log n)\xi(\ell_o,\log n)]^{1/2}d\ell \\ \to\ & \tau\frac{\int_{\ell=\ell_{n,1}}^{\ell_{n,2}} e^{-[\xi(\ell,\log n)-\xi(\ell_o,\log n)]r_n}\xi^{1/2}(\ell,\log n)d\ell}{\int_{\ell=0}^{\lambda(\log n)} e^{-[\xi(\ell,\log n)-\xi(\ell_o,\log n)]r_n}\xi^{1/2}(\ell,\log n)d\ell} \end{aligned}$$

by (A5a). This proves (53).

□

Lemma 10 *Suppose that U is differentiable with gradient ∇U and let, for some $\boldsymbol{x}$,*

$$\varepsilon = \sup_{\boldsymbol{y}\in S(\boldsymbol{x};r)} |U(\boldsymbol{y}) - U(\boldsymbol{x}) - \langle\nabla U(\boldsymbol{x}),\boldsymbol{y}-\boldsymbol{x}\rangle|. \tag{59}$$

Let $\zeta\in(-1,1]$. Then, with $\xi = |\nabla U(\boldsymbol{x})|$, there exist constants $\theta_1 = \theta_1(\boldsymbol{x},r)\in[-1,1]$ and $\frac{2}{1+\zeta}+\frac{1}{2}\le\theta_2=\theta_2(\boldsymbol{x},r)\le 0$ and such that

$$\begin{aligned} & F(S(\boldsymbol{x};r)\cap\{\boldsymbol{y}:\langle\boldsymbol{y}-\boldsymbol{x},\nabla U(\boldsymbol{x})/|\nabla U(\boldsymbol{x})|\rangle\le\zeta r\}) \\ =\ & (2\pi r)^{1/2}e^{-U(\boldsymbol{x}_o)}\xi^{-3/2}e^{\xi r}e^{\theta_1\varepsilon}[1+\theta_2/(\xi r)]. \end{aligned}$$

Proof: Let $B = \{\boldsymbol{y} : \langle \boldsymbol{y} - \boldsymbol{x}, \nabla U(\boldsymbol{x})/|\nabla U(\boldsymbol{x})|\rangle \leq \zeta r\}$. By (59), we can write

$$F(S(\boldsymbol{x};r) \cap B) = e^{-U(\boldsymbol{x})} e^{\theta_1 \varepsilon} \int_{S(\boldsymbol{x};r)\cap B} e^{-\langle \nabla U(\boldsymbol{x}), \boldsymbol{y}-\boldsymbol{x}\rangle} d\boldsymbol{y}. \tag{60}$$

Then change variables with $\boldsymbol{y} - \boldsymbol{x} = A\boldsymbol{v}$ where

$$A = \begin{bmatrix} b & a \\ -a & b \end{bmatrix}$$

with $(a,b)' = \nabla U(\boldsymbol{x})/|\nabla U(\boldsymbol{x})|$. It is easy to see that

$$\begin{aligned} & \int_{S(\boldsymbol{x};r)\cap B} e^{-\langle \nabla U(\boldsymbol{x}), \boldsymbol{y}-\boldsymbol{x}\rangle} d\boldsymbol{y} \\ = & \int_{S(0;r)} I(v_2 \leq \zeta r) e^{-\xi v_2} d\boldsymbol{v} \\ = & \ \xi^{-2} \int_{S(0;\xi r)} I(v_2 \leq \xi\zeta r) e^{-v_2} d\boldsymbol{v} \\ = & \ 2\xi^{-2} \int_{-\xi r}^{\xi\zeta r} e^{-v} [(\xi r)^2 - v^2]^{1/2} dv. \end{aligned}$$

Next, letting $z = v + \xi r$, the previous expression is equal to

$$\begin{aligned} & 2\xi^{-2} e^{\xi r} \int_0^{\xi(1+\zeta)r} e^{-z} (2\xi r z - z^2)^{1/2} dz \\ = & \ 2\xi^{-2} e^{\xi r} (2\xi r)^{1/2} \int_0^{\xi(1+\zeta)r} e^{-z} z^{1/2} (1 - z/(2\xi r))^{1/2} dz. \end{aligned}$$

Since $1 - x/2 \leq (1-x)^{1/2} \leq 1$ for $x \in [0,1]$, we have

$$\begin{aligned} & \int_0^{\xi(1+\zeta)r} e^{-z} z^{1/2} dz - \frac{1}{4\xi r} \int_0^{\xi(1+\zeta)r} e^{-z} z^{3/2} dz \\ \leq & \int_0^{\xi(1+\zeta)r} e^{-z} z^{1/2} (1 - z/(2\xi r))^{1/2} dz \\ \leq & \int_0^{\infty} e^{-z} z^{1/2} dz = \Gamma(3/2). \end{aligned}$$

The lower bound is

$$\Gamma(3/2) - \int_{\xi(1+\zeta)r}^{\infty} e^{-z} z^{1/2} dz - \frac{1}{4\xi r} \int_0^{\xi(1+\zeta)r} e^{-z} z^{3/2} dz,$$

and since $\Gamma(3/2) = \pi^{1/2}/2$, we obtain that

$$\begin{aligned} & -\int_{\xi(1+\zeta)r}^{\infty} e^{-z} z^{1/2} dz - \frac{1}{4\xi r}\int_0^{\xi(1+\zeta)r} e^{-z} z^{3/2} dz \\ \leq \quad & \int_0^{\xi(1+\zeta)r} e^{-z} z^{1/2}(1 - z/(2\xi r))^{1/2} dz - \pi^{1/2}/2 \leq 0. \end{aligned}$$

Now,

$$\int_{\xi(1+\zeta)r}^{\infty} e^{-z} z^{1/2} dz \leq \frac{1}{(\xi(1+\zeta)r)}\int_{\xi(1+\zeta)r}^{\infty} e^{-z/2} z^{3/2} dz \leq \frac{1}{\xi(1+\zeta)r}\Gamma(5/2)$$

and

$$\frac{1}{4\xi r}\int_0^{\xi(1+\zeta)r} e^{-z} z^{3/2} dz \leq \frac{1}{4\xi r}\Gamma(5/2).$$

Since $\Gamma(5/2) = 3\pi^{1/2}/4$ and $3\pi^{1/2}/8 \leq 1$, this concludes the proof.

□

References

[1] Erdelyi, A. (1956). *Asymptotic Expansions.* Dover.

[2] Ibragimov, I. A. and Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff.

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

[4] Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer.

[5] Lee, S. (1997). The central limit theorem for Euclidean minimal spanning trees. I. *Ann. Appl. Probab. 7*, 996–1020.

[6] Penrose, M. D. (1997). The longest edge of the minimal spanning tree. *Ann. Appl. Probab. 7*, 340–361.

[7] Penrose, M. D. (1998). Extremes for the minimal spanning tree on normally distributed points. *Adv. Appl. Prob. 30*, 628-639.

[8] Penrose, M. (2000). Central limit theorems for k-nearest neighbour distances. Stochastic Process. Appl. 85 (2000), no. 2, 295–320.