# ON THE $h$-DIAMETER OF A RANDOM POINT SET

**M J Appel**[1] **and R P Russo**[2]

ABSTRACT. Let $(S,\sum)$ be a measurable space and let $h : S\times S \to \mathbf{R}$ be a real symmetric Borel/$\sum\times\sum$-measurable function on $S\times S$. Let $B$ be a non-empty measurable subset in $S$ and let $\mu$ be a probability measure supported on the restriction of the measurable space $(S,\sum)$ to $B$. Let $B$ have finite *h-diameter*

$$\overline{h} \equiv \text{ess sup}\{h(u,v) : u,v \in B\} < \infty.$$

Let $U, U_1, U_2$,.... be a sequence of independent random points taking values in $B$ according to $\mu$ and let

$$H_n = \max\{h(U_i,U_j) : 1 \le i < j \le n\}$$

denote the $h$-diameter of the set $\{U_i, i = 1,\dots,n\}$, the maximum pair-wise $h$-distance among the first $n$ points.

A theoretical framework is provided from which one may deduce the weak convergence of $H_n$, upon suitable centering and rescaling, to an extreme-value distribution. The sufficient condition provided herein is quite different from that of Appel, *et al.* [1]. Several applications of the theory are provided.

Keywords: symmetric function, limit distribution, Weibull distribution, $h$-diameter, extreme value

2000 *Mathematics Subject Classification*: 60G70, 62G32, 60D05

Suggested running head: *The h-diameter of a random point set*

[1] MORTGAGE GUARANTY INSURANCE CORPORATION
MILWAUKEE WI 53201 USA MARTIN_ APPEL@MGIC.COM (414) 430 0026

[2] DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE, UNIVERSITY OF IOWA
IOWA CITY, IA 52242 USA RRUSSO@STAT.UIOWA.EDU

## 1. INTRODUCTION.

Let $(S,\sum)$ be a measurable space and let $h : S \times S \to \mathbf{R}$ be a real symmetric Borel/$\sum \times \sum$ - measurable function on $S \times S$. Let $B$ be a non-empty measurable subset in $S$. Let $\mu$ be a probability measure supported on the restriction of the measurable space $(S,\sum)$ to $B$ and assume that $B$ has finite *h-diameter*

$$
\begin{aligned}
\overline{h} &\equiv \text{ess sup}\{h(u,v) : u,v \in B\} \\
&= \sup\{x : \mu \times \mu\{h(u,v) > x\} > 0\} \qquad (1.1)\\
&< \infty.
\end{aligned}
$$

Let $U, U_1, U_2$,.... be a sequence of independent random points taking values in $B$ according to $\mu$. Let $P$ denote the product measure constructed from copies of $\mu$, so that $P$ will refer to the joint distribution of the $U_i$'s. Let

$$
H_n = \max\{h(U_i, U_j) : 1 \le i < j \le n\} \qquad (1.2)
$$

denote the $h$-diameter of the set $\{U_1, U_2, \ldots, U_n\}$, the maximum pair-wise $h$-distance among the first $n$ points. A precise description of the almost-sure limiting behavior of $H_n$ appears in Appel, *et al* [2]. We are concerned here with weak asymptotics. Our main result (Theorem 1 below) provides a methodology for inferring extreme-value limit laws for $H_n$, upon suitable (non-random) rescaling. While the sufficient condition of the main result is quite different from that of Appel, *et al* [1], we continue here the line of investigation therein.

As an application of our results, we consider the special case where $U_1, U_2, \ldots, U_n$ is a random sample from the uniform distribution on a closed disk(sphere) $B$ with center $c$ in $\mathbf{R}^2(\mathbf{R}^3)$. Let $h$ denote the ordinary $l_2$ euclidean metric and let $B$ have radius $\rho = h(c,v)$, where $v$ is any point on the boundary $\partial B$. If $c$ is known but $\rho$ is unknown, then the sample "radius"

$$
R_n = \max\{h(c,U_i) : 1 \le i \le n\}
$$

is observable and can be used to construct an exact $p \times 100\%$ confidence interval for the diameter of $B$

$$
\left(\frac{2R_n}{t_2}, \frac{2R_n}{t_1}\right),
$$

where $0 < t_1 < t_2$ are points for which $t_2^{kn} - t_1^{kn} = p$, $k = 2,3$. However, if both center and radius are unknown, one may need to rely on the diameter $H_n$ in order to construct such an interval. The weak limiting behavior of $H_n$ would thus seem to be of interest.

Appel, Najim and Russo [3] studied the weak limiting behavior of diameters over uniform point sets on compact planar regions having finitely many axes with no vertices in common (the unit square, for example). In comparison, the disk in $\mathbf{R}^2$ is interesting as it provides no geometric clue regarding the location of the vertices that define the diameter except that for large $n$ those vertices are likely to be close to the boundary and nearly diametrically opposite to each other. See Dette and Henze [5], Steele and Tierney [11] and, more recently, Appel and Russo [1] for related results on the maximal diameter. In the context of random geometric graphs, $H_n$ is the minimal edge length for which the point set induces a complete graph (a graph possessing all $\binom{n}{2}$ possible edges). Many wireless communication network protocols are based on the properties of random geometric graphs generated on the unit disk;

see Ellis, *et al* [7]. For a comprehensive treatment of the theory of random geometric graphs, see Penrose [9].

## 2. Statement and Proof of Main Result.

We assume that $\overline{h} < \infty$. This entails no loss of generality since we may always apply a monotone transformation (for example $h \to h/(1+h)$) to an unbounded $h$ to produce an equivalent bounded one.

Given a point $u$ in $B$ and $q < \overline{h}$, define $H(u,q) = \{v \in B : h(u,v) > q\}$. Let $\{q_n\}$ be monotone non-decreasing with $q_n \uparrow \overline{h}$ as $n \to \infty$. For each $n$, define $\Gamma_n = \{v \in B : h(w,v) > q_n,$ for some $w \in B\}$. Clearly,

(2.1) $$H\ (u,q_n) \cap \Gamma_n^c = \emptyset, \text{ any } u \text{ in } B.$$

Define random variables

(2.2) $$Y_{k,n} = Y_{k,n}(q_n) = P(U \in H\ (U_k,q_n) \mid U_k).$$

Let $F(x) = P(h(U_1,U_2) \le x)$. Note that $EY_{1,n} = 1 - F(q_n)$.

We now state our main result.

*Theorem 1. If*

*(a) the sequence* $\{n^2 \cdot EY_{1,n}\}$ *satisfies*

(2.3) $$0 < \liminf_n \left[n^2 \cdot EY_{1,n}\right]] \le \limsup_n \left[n^2 \cdot EY_{1,n}\right] < \infty,$$

*and if*

*(b) there is a non-random sequence* $\{m_n\}$ *such that*

(2.4) $$Y_{1,n} \le m_n = o\left(\frac{1}{n}\right), \text{ everywhere, as } n \to \infty,$$

*and if*

*(c) for each* $\delta > 0$ *and* $0 < \varepsilon < 1$

(2.5) $$\lim_{n\to\infty} \sum_{j=[n\varepsilon]}^{n-1} P\left(\left|\sum_{k=1}^{j} Y_{k,n} - jEY_{1,n}\right| \ge j\delta EY_{1,n}\right) = 0,$$

*then*

(2.6) $$\lim_{n\to\infty}\left[\exp\left(\frac{n^2\cdot(1-F(q_n))}{2}\right)\cdot P(H_n \le q_n)\right] = 1.$$

*Corollary 1. If (a) above holds and there is a non-random sequence* $\{m_n\}$ *such that*

$$Y_{1,n} \leq m_n = o\left(\frac{1}{n\log n}\right), \text{ a.s., as } n \to \infty \tag{2.7}$$

*then (2.6) holds.*

REMARKS.

1. For each $t > 0$, take $q_n$ to be the quantile

$$q_n = q_n(t) = \inf\left\{x : F(x) \geq 1 - \frac{2t}{n^2}\right\}. \tag{2.8}$$

If the distribution function $F(\cdot)$ above is continuous in its far-right tail, then we have $EY_{1,n} = 1 - F(q_n) = \frac{2t}{n^2}$ for all $n$ large, in which case the conclusion from (2.6) is that $P(H_n \leq q_n) \to \exp(-t)$ as $n \to \infty$.

2. In general, the condition (2.5) is sufficient but not necessary for (2.6) to obtain. To see this, suppose the distribution of $h(U_1, U_2)$ has a final jump discontinuity in its right tail: for some $z > 0$, $F(x) \leq 1 - z$ for all $x \leq x^*$ and $F(x^*) = 1$. Then, for a given $t$, the quantile $q_n(t)$ in (2.8) takes the value $x^* = \overline{h}$ for all $n$ large enough. With $q_n = q_n(t)$, $EY_{1,n} = 1 - F(x^*) = 0$ and all of the probabilities in the sum in (2.5) are equal to 1; the condition fails to obtain. However, the limit (2.6) exists, trivially.

PROOF OF THEOREM 1. We first state two lemmas that we will use below. The first is Bernstein's inequality (cf. Serfling [10], page 95).

*Lemma 1.* (S.N. Bernstein) *Let* $Y, Y_n, n \geq 1$, *be a sequence of i.i.d. random variables satisfying* $\Pr(|Y - EY| \leq m) = 1$, *where* $m < \infty$. *Then, for any* $t > 0$ *and all* $n \geq 1$

$$\Pr\left(\left|\sum_{j=1}^{n} Y_j - nEY\right| \geq nt\right) \leq 2\exp\left(\frac{-nt^2}{2\mathrm{Var}Y + \frac{2}{3}mt}\right).$$

The next result follows from the convergence $x^{-1}(1 - \exp(-rx)) \to r$ as $x \to 0$.

*Lemma 2. Fix* $r > 1$. *Then for all small* $x$, $\exp(-rx) \leq 1 - x \leq \exp(-x)$.

Let $\delta > 0$ and $0 < \varepsilon < 1$ be fixed but arbitrary. Define events

$$R_{n,q} = \{H_n \leq q\} \tag{2.9}$$

and for each $j = 2, \ldots, n$

$$A_{j,n} = \left\{ U_j \notin \bigcup_{k=1}^{j-1} H\left(U_k, q_n\right) \right\} \tag{2.10}$$

$$E_{j,n} = \left\{ \left| \sum_{k=1}^{j} Y_{k,n} - jEY_{1,n} \right| \geq j\delta EY_{1,n} \right\} \tag{2.11}$$

$$F_{j,n} = \left\{ \sum_{k=1}^{j} Y_{k,n} \leq j(1+\delta)EY_{1,n} \right\}. \tag{2.12}$$

It is easy to check that

$$R_{n,q_n} = \cap_{j=2}^{n} A_{j,n}. \tag{2.13}$$

Since $nEY_{1,n} < \frac{1}{1+\delta}$ for large enough $n$, invoking the double expectation rule and keeping the second-order terms in the Bonferroni bound, we have

$$\begin{aligned} P\left(R_{n,q_n}\right) &= E\left[1_{R_{n-1,q_n}}\left(1 - P\left(A_{n,n}^c \,\middle|\, U_1, \ldots, U_{n-1}\right)\right)\right] \\ &\leq E\left[1_{R_{n-1,q_n}}\left(1 - \sum_{k=1}^{n-1} Y_{k,n} + \sum_{1\leq j\neq k\leq n-1} P\left(U_n \in H\left(U_j, q_n\right) \cap H\left(U_k, q_n\right) \,\middle|\, U_j, U_k\right)\right)\right] \\ &\leq E\left[1_{R_{n-1,q_n}\cap E_{n-1,n}^c}\left(1 - \sum_{k=1}^{n-1} Y_{k,n}\right)\right] + E\left[1_{R_{n-1,q_n}\cap E_{n-1,n}}\left(1 - \sum_{k=1}^{n-1} Y_{k,n}\right)\right] + (n-1)^2 EY_{1,n}^2 \\ &\leq P\left(R_{n-1,q_n}\right)\left(1 - (n-1)(1-\delta)EY_{1,n}\right) + P\left(E_{n-1,n}\right) + (n-1)^2 EY_{1,n}^2. \end{aligned} \tag{2.14}$$

Repeating the above argument $n - [n\varepsilon]$ times gives

$$\begin{aligned} P\left(R_{n,q_n}\right) &\leq P\left(R_{[n\varepsilon],q_n}\right) \cdot \Pi_{j=[n\varepsilon]}^{n-1}\left(1 - j(1-\delta)EY_{1,n}\right) + \sum_{j=[n\varepsilon]}^{n-1} P\left(E_{j,n}\right) + \sum_{j=[n\varepsilon]}^{n-1} j^2 EY_{1,n}^2 \\ &\leq \Pi_{j=[n\varepsilon]}^{n-1}\left(1 - j(1-\delta)EY_{1,n}\right) + \sum_{j=[n\varepsilon]}^{n-1} P\left(E_{j,n}\right) + \sum_{j=[n\varepsilon]}^{n-1} j^2 EY_{1,n}^2. \end{aligned} \tag{2.15}$$

To get an inequality similar to (2.15) which goes in the other direction, by Boole's inequality we may write

$$
\begin{aligned}
P\left(R_{n,q_n}\right) &= E\left[P\left(R_{n,q_n} \mid U_1,\ldots,U_{n-1}\right)\right] \\
&= E\left[1_{R_{n-1,q_n}} P\left(A_{n,n} \mid U_1,\ldots,U_{n-1}\right)\right] \\
&\geq E\left[1_{R_{n-1,q_n}\cap F_{n-1,n}}\left(1-P\left(A^c_{n,n} \mid U_1,\ldots,U_{n-1}\right)\right)\right] \\
&\geq E\left[1_{R_{n-1,q_n}\cap F_{n-1,n}}\left(1-\sum_{k=1}^{n-1} Y_{k,n}\right)\right] \\
&\geq E\left[1_{R_{n-1,q_n}\cap F_{n-1,n}}\left(1-(n-1)(1+\delta)EY_{1,n}\right)\right] \\
&= \left(1-(n-1)(1+\delta)EY_{1,n}\right)\left(P\left(R_{n-1,q_n}\right)-P\left(R_{n-1,q_n}\cap F^c_{n-1,n}\right)\right) \\
&\geq P\left(R_{n-1,q_n}\right)\left(1-(n-1)(1+\delta)EY_{1,n}\right)-P\left(F^c_{n-1,n}\right) \\
&\geq P\left(R_{n-1,q_n}\right)\left(1-(n-1)(1+\delta)EY_{1,n}\right)-P\left(E_{n-1,n}\right).
\end{aligned} \tag{2.16}
$$

Repeating the above argument $n-[n\varepsilon]$ times gives

$$
P\left(R_{n,q_n}\right) \geq P\left(R_{[n\varepsilon],q_n}\right)\cdot \Pi_{j=[n\varepsilon]}^{n-1}\left(1-j(1+\delta)EY_{1,n}\right)-\sum_{j=[n\varepsilon]}^{n-1} P\left(E_{j,n}\right). \tag{2.17}
$$

By Boole's inequality, the i.i.d. assumption and (2.3),

$$
\begin{aligned}
P\left(R_{[n\varepsilon],q_n}\right) &\geq 1-\binom{[n\varepsilon]}{2}\cdot\left(1-F\left(q_n\right)\right) \\
&\approx 1-\frac{\varepsilon^2}{2}\cdot n^2\cdot\left(1-F\left(q_n\right)\right) \\
&\geq 1-c\cdot\varepsilon^2,
\end{aligned} \tag{2.18}
$$

for some constant $c$. Therefore, $\liminf_n P\left(R_{[n\varepsilon],q_n}\right)$ can be made arbitrarily close to 1 if $\varepsilon$ is chosen small enough.

By Lemma 2, we have

$$
\Pi_{j=[n\varepsilon]}^{n-1}\left(1-j(1-\delta)EY_{1,n}\right) \leq \exp\left(-\left(1-\delta\right)EY_{1,n}\sum_{j=[n\varepsilon]}^{n-1} j\right), \tag{2.19}
$$

and for fixed but arbitrary $r>1$

$$
\exp\left(-r\left(1+\delta\right)EY_{1,n}\sum_{j=[n\varepsilon]}^{n-1} j\right) \leq \Pi_{j=[n\varepsilon]}^{n-1}\left(1-j(1+\delta)EY_{1,n}\right), \tag{2.20}
$$

as $n \to \infty$. Note that $\sum_{j=[n\varepsilon]}^{n-1} j \approx (1-\varepsilon)\frac{n^2}{2}$. To complete the proof of Theorem 1, by (2.4) we have

$$\text{(2.21)} \qquad \begin{aligned} \sum_{j=[n\varepsilon]}^{n-1} j^2 EY_{1,n}^2 &\le \frac{2tm_n}{n^2} \cdot O\left(n^3\right) \\ &= o(1), \end{aligned}$$

as $n \to \infty$. By (2.5), the sum $\sum_{j=[n\varepsilon]}^{n-1} P\left(E_{j,n}\right)$ which appears in both (2.15) and (2.17) vanishes, as $n \to \infty$. Theorem 1 now follows from (2.15), (2.17), (2.19), (2.20) and (2.21), multiplying by $\exp\left(n^2 EY_{1,n}\right)$ everywhere. □

PROOF OF COROLLARY 1. From Lemma 1, we have

$$\text{(2.22)} \qquad P\left(E_{j,n}\right) \le 2\exp\left(-\frac{j\delta^2 E^2 Y_{1,n}}{2\text{Var}Y_{1,n} + \frac{2}{3}\delta m_n EY_{1,n}}\right).$$

Now $m_n$ bounds $Y_{1,n}$ and hence $|Y_{1,n} - EY_{1,n}|$, a.s. Thus $\text{Var}Y_{1,n} \le m_n E\,|Y_{1,n} - EY_{1,n}| \le m_n EY_{1,n}$. Since $j \ge [n\varepsilon]$, the exponent in (2.22) is therefore bounded above by

$$\text{(2.23)} \qquad -\frac{[n\varepsilon]\delta^2 EY_{1,n}}{m_n\left(2+\frac{2}{3}\delta\right)} \approx -\frac{c(t)}{nm_n},$$

where

$$\text{(2.24)} \qquad c(t) = \frac{\varepsilon\delta^2 t}{1+\frac{\delta}{3}}.$$

If (2.7) holds, then it is clear that (2.5) holds. □

## 3. APPLICATIONS.

MINIMUM SPACING IN AN EXPONENTIAL RANDOM SAMPLE. Let $U_1, U_2, \ldots, U_n$ be independent random variables with common unit mean exponential distribution. Let $X_{(1)} = \min_{i=1,\ldots,n}\left\{X_{(i)}\right\} < X_{(2)} < \cdots < X_{(n)} = \max_{i=1,\ldots,n}\left\{X_{(i)}\right\}$ be the order statistics and let $s_n = \min_{k=1,\ldots,n}\left\{X_{(k+1)} - X_{(k)}\right\}$ denote the minimum spacing. Set $h(u,v) = \frac{1}{1+|u-v|}$. Then $H_n = \frac{1}{1+s_n} \uparrow \overline{h} \equiv 1$, a.s. To see how fast, we let $0 < x < 1$ and compute $F(x) = P\left(h(U_1,U_2) \le x\right) = P\left(|U_1 - U_2| > \frac{1}{x} - 1\right) = \exp\left(-\frac{1}{x}+1\right)$. For fixed $t > 0$, we take $q_n$ to be the quantile $q_n(t) = \inf\left\{x : F(x) \ge 1 - \frac{2t}{n^2}\right\}$, as above in (2.8). We invert $F$ and solve

$$q_n = \frac{1}{1 - \log\left(1 - \frac{2t}{n^2}\right)}.$$

It is easy to check that the random variable $Y_{k,n}$ is actually a constant here:

$$\begin{aligned} Y_{k,n}(q_n) &= P(h(U_1,U_k) > q_n \mid U_k) \\ &= 1 - F(q_n) \\ &= \frac{2t}{n^2}, \end{aligned}$$

the second equality following from continuity, so that conditions (a), (b) and (trivially) (c) in Theorem 1 (as well as (2.7) in Corollary 1) obtain. It follows from Theorem 1 that (2.6) holds:

$$\begin{aligned} \lim_{n\to\infty}\left[\exp\left(\frac{n^2\cdot(1-F(q_n))}{2}\right)\cdot P(H_n \le q_n)\right] &= \exp(t)\cdot\lim_{n\to\infty} P(H_n \le q_n) \\ &= \exp(t)\cdot\lim_{n\to\infty} P\left(\frac{n^2}{2}\cdot[(1-\exp(-s_n))] > t\right) \quad (3.2) \\ &= 1, \quad (3.3) \end{aligned}$$

thus $\frac{n^2}{2}\cdot[(1-\exp(-s_n))]$ converges weakly to a unit mean exponential distribution. This fact is easily deduced from first principles: by the memoryless property, the spacings between the order statistics are distributed as independent exponentials with respective rates $n-1, n-2, \ldots, 1$ so the minimum spacing is exponential with rate $\sum_{i=1}^{n-1} i = \binom{n}{2}$.

I.I.D. SEQUENCES OF INDEPENDENT BERNOULLI TRIALS. Define a distance measure $d$ on pairs of infinite binary strings $u = (\omega_1(u), \omega_2(u), \ldots)$, $\omega_j(u) \in \{0,1\}$, by

$$d(u,v) = \inf\left\{ j \ge 1 : \omega_j(u) = \omega_j(v)\right\}, \quad (3.4)$$

with $\sup\emptyset = 0$. That is, $d(u,v)$ is the first coordinate-wise agreement between $u$ and $v$. Let $h(u,v) = \frac{d(u,v)}{1+d(u,v)}$. Again, $\overline{h} = 1$. Let $U_1, U_2, \ldots, U_n$ be an independent sequence of infinite sequences of independent Bernoulli trials with common success probability $\frac{1}{3} < p \le \frac{1}{2}$. We set $\varphi = \varphi(p) = 2p(1-p)$, $N(x) = \left[\frac{x}{1-x}\right]$, and compute $F(x) = P(h(U_1,U_2) \le x) = 1 - \varphi^{N(x)}$. Let $\{q_n\}$ be any sequence which satisfies $q_n \uparrow 1$ and

$$\sup_n \left| \frac{q_n}{1-q_n} + \frac{2\log n}{\log\varphi(p)} \right| < \infty. \quad (3.5)$$

Condition (a) in Theorem 1 thus holds. It is easy to see that

$$\begin{aligned} Y_{k,n}(q_n) &= P(h(U,U_k) > q_n \mid U_k) \\ &= \Pi_{j=1}^{N(q_n)} \rho_{j,k}, \quad (3.6) \end{aligned}$$

where

$$\rho_{j,k} = p\cdot 1_{\{\omega_j(U_k)=1\}} + (1-p)\cdot 1_{\{\omega_j(U_k)=0\}}. \quad (3.7)$$

Let $m_n \equiv (1-p)^{N(q_n)}$ and $Z = 2\frac{\log(1-p)}{\log\varphi}$. It is easy to check that $Z > 1$ when $p > \frac{1}{3}$ and therefore that $m_n = o\left(\frac{1}{n\log n}\right)$ as $n \to \infty$. Hence condition (2.7) in Corollary 1 holds. From (2.6) we conclude that

$$\lim_{n\to\infty}\left[\exp\left(\frac{n^2}{2}\cdot\varphi^{N(q_n)}\right)\cdot P\left(H_n \le q_n\right)\right] = 1. \tag{3.8}$$

Likewise, we may redefine

$$d(u,v) = \inf\left\{j \ge 1 : \omega_j(u) \ne \omega_j(v)\right\} \tag{3.9}$$

and show that the maximum first *dis*agreement time among pairs of $n$ such infinite Bernoulli sequences has the limiting distribution (3.8) with $\varphi = p^2 + (1-p)^2$, for any success probability $0 < p < 1$, when the rate $q_n$ satisfies (3.5).

I.I.D. UNIFORM POINTS IN THE EUCLIDEAN UNIT DISK AND UNIT SPHERE. For i.i.d uniform points $U_i$, we consider the convergence of $H_n$ in the closed euclidean disk $B$ of unit radius centered at the origin. Let $u$ denote an arbitrary point in $\Gamma_n$. We will approximate $P\left(U \in \mathcal{H}\left(u, b_n\right) | U_k = u\right)$ for $n$ large and invoke Corollary 1 to obtain the following result.

*Proposition 1.*

$$\lim_{n\to\infty} P\left(n^{4/5}\left(2 - H_n\right) > w\right) = \exp\left(-\frac{4w^{5/2}}{5\pi}\right). \tag{3.10}$$

PROOF. In what follows, if $\lambda_{n,u}$ and $\mu_{n,u}$ are two quantities that depend on both $u$ and $n$ we write $\lambda_{n,u} \sim \mu_{n,u}$ to indicate that $\lambda_{n,u}/\mu_{n,u} \to 1$ as $n \to \infty$, uniformly in $u \in \Gamma_n$. Let $b_n \to 0$. Draw two line segments of length $q_n = 2 - b_n$ from $u$ to points $e_1$ and $e_2$ on the boundary $\partial B$ of $B$. Let $\alpha_u$ denote the angle formed by these line segments at $u$ and let $\theta_u$ denote the angle at 0 formed by the line segments joining $e_1$ to 0 and $e_2$ to 0. Let $y_u = 1 - d(u,0)$ denote the distance between $u$ and $\partial B$. By the law of cosines

$$\begin{aligned}
-\cos\theta_u &= \frac{(2-b_n)^2 - (1-y_u)^2 - 1}{2(1-y_u)} \\
&= \frac{2 - 4b_n + b_n^2 + 2y_u - y_u^2}{2(1-y_u)} \\
&\sim \frac{1 - 2b_n + y_u}{(1-y_u)} \\
&\sim 1 - 2b_n + 2y_u.
\end{aligned} \tag{3.11}$$

Thus, since $\theta_u \to \pi$ as $b_n \to 0$, we have

$$\begin{aligned}
\pi - \theta_u &\sim \sin(\pi - \theta_u) \\
&= \sin\theta_u \\
&= (1 - \cos^2\theta_u)^{1/2} \\
&\sim 2(b_n - y_u)^{1/2}.
\end{aligned} \tag{3.12}$$

Let $S_{u,1}$ denote the sector in $B$ formed by the line segments joining 0 to $e_1$ and $e_2$, respectively. We have

$$P(U \in S_{u,1}) \sim \frac{2(b_n - y_u)^{1/2}}{\pi}. \tag{3.13}$$

Let $S_{u,2}$ denote the sector formed by the line segments joining $u$ to $e_1$ and $e_2$ on the disk with center $u$ and radius $2 - b_n$. By the law of sines, $\alpha_u \sim \sin\alpha_u = (2-b_n)^{-1}\sin\theta_u$, so that

$$P(U \in S_{u,2}) \sim \frac{(4-2b_n)(b_n - y_u)^{1/2}}{\pi}. \tag{3.14}$$

Let $T_{u,1}$ be the triangle with vertices 0, $u$ and $e_1$, and let $T_{u,2}$ be the triangle with vertices 0, $u$ and $e_2$. We have

$$P(U \in T_{u,1} \cup T_{u,2}) \sim \frac{2(1-y_u)(b_n - y_u)^{1/2}}{\pi}. \tag{3.15}$$

Thus we have the non-random bound

$$\begin{aligned} P\left(U \in H\,(u,b_n)\,|U_k = u\right) &= P(U \in S_{u,1}) - P(U \in S_{u,2}) + P(U \in T_{u,1} \cup T_{u,2}) \\ &\sim \frac{2(b_n - y_u)^{3/2}}{\pi} \\ &\le m_n := \frac{2b_n^{3/2}}{\pi}, \end{aligned} \tag{3.16}$$

for all $n$ large, uniformly in $u$. Moreover, since $P(h(U,0) > 2-y) = 2y - y^2$, we compute

$$\begin{aligned} EY_{1,n} &\sim \int_0^{b_n} \frac{2(b_n - y)^{3/2}}{\pi}(2-2y)\,dy \\ &= \frac{8b_n^{5/2}}{5\pi}. \end{aligned} \tag{3.17}$$

By continuity $EY_{1,n} = \frac{2t}{n^2}$. Thus (3.17) implies that $m_n = \frac{2b_n^{3/2}}{\pi} = O\left(n^{-6/5}\right)$, and so $m_n$ satisfies (2.7). Equation (3.17) also allows us to solve directly for $b_n$:

$$b_n = b_n(t) \sim \frac{w}{n^{4/5}}, \text{ as } n \to \infty, \tag{3.18}$$

where $w = \left(\frac{5\pi t}{4}\right)^{2/5}$. Equation (3.10) now follows from Corollary 1. □

Next, let $B$ be the closed unit sphere centered at 0 in $\mathbf{R}^3$. Using arguments analogous to those for the disk, we will prove the following result.

*Proposition 2.*

$$\lim_{n\to\infty} P\left(n^{2/3}(2 - H_n) > w\right) = \exp\left(-\frac{3w^3}{4}\right). \tag{3.19}$$

PROOF. Again, let $b_n \to 0$ and $q_n = 2 - b_n$. Given $u$ in $\Gamma_n$, let

$$C(u) = \{x \in \partial B : h(u,x) = q_n\} \tag{3.20}$$

be the circle embedded in the boundary $\partial B$ whose center $\phi^*$ lies in the interior of $B$. Let $\text{cap}_1$ denote the spherical cap of $B$ whose circular base has circumference $C(u)$. Let $h$ be the height of $\text{cap}_1$: $h = 1 - h(0,\phi^*)$. Then (cf. [8])

$$\text{Vol}\,(\text{cap}_1) = \frac{\pi h^2(3-h)}{3}. \tag{3.21}$$

As in the 2-dimensional case, let $y_u = 1 - h(u,0) =$ denote the distance between $u$ and $\partial B$. From the Pythagorean relation we have $h \sim 2b_n - 2y_u$ so that

$$\text{Vol}\,(\text{cap}_1) \sim 4\pi(b_n - y_u)^2. \tag{3.22}$$

Next, let $S_{u,b_n}$ denote the sphere with center $u$ and radius $2 - b_n$ and let $\text{cap}_2$ denote the spherical cap of $S_{u,b_n}$ contained in $\text{cap}_1$ and with the same base. Again, the Pythagorean relation implies that the height $h$ of $\text{cap}_2$ satisfies $h \sim b_n - y_u$, so that

$$\begin{aligned} \text{Vol}\,(\text{cap}_2) &= \frac{\pi h^2(3-h)}{3} \\ &\sim 2\pi(b_n - y_u)^2. \end{aligned} \tag{3.23}$$

Note that a point $x$ in $B$ satisfies $h(u,x) > 2 - b_n$ if and only if $x$ lies in $\text{cap}_1 \setminus \text{cap}_2$. Thus,

$$\begin{aligned} P(U \in \mathcal{H}\,(u,b_n)) &= \frac{\text{Vol}\,(\text{cap}_1) - \text{Vol}\,(\text{cap}_2)}{\text{Vol}\,(B)} \\ &\sim \frac{3}{2}(b_n - y_u)^2 \end{aligned} \tag{3.24}$$

uniformly in $u$ which lie in $\Gamma_n$, that is, which satisfy $h(0,u) > 1 - b_n$. Hence we have the non-random bound

$$Y_{k,n} \lesssim m_n := \frac{3b_n^2}{2}, \tag{3.25}$$

for all $n$ large enough. Since $P(h(U,0) > 2 - y) = 1 - (1-y)^3$, we compute

$$\begin{aligned} EY_{1,n} &\sim \int_0^{b_n} \frac{9(b_n - y)^2}{2}(1-y)^2 dy \\ &= \frac{3b_n^3}{2}. \end{aligned} \tag{3.26}$$

Again, by continuity $EY_{1,n} = \frac{2t}{n^2}$ and so (3.25) implies that $m_n = O\left(n^{-4/3}\right)$, which shows that $m_n$ satisfies (2.7). Solving for $b_n = \left(\frac{4t}{3}\right)^{1/3} n^{-2/3}$, (3.19) now follows from Corollary 1. □

It is interesting to compare the limit laws (3.10) and (3.19) to the 1-dimensional case of i.i.d. uniform variables in the interval [-1,1]. In this case, since $\frac{H_n}{2}$ is distributed as the range of $n$ variables distributed uniformly on [0,1], it follows from the elementary properties of order statistics (cf. Arnold

[4], page 31) that

$$P(H_n < 2t) = nt^{n-1} - (n-1)t^n, \tag{3.27}$$

from which we obtain

$$\begin{aligned} P(n(2-H_n) > w) &= n\left(1-\frac{w}{2n}\right)^{n-1} - (n-1)\left(1-\frac{w}{2n}\right)^n \\ &\to \left(1+\frac{w}{2}\right)\exp\left(-\frac{w}{2}\right). \end{aligned} \tag{3.28}$$

It is also interesting to compare (3.10) and (3.19) to results obtained recently on the weak behavior of the diameter of $n$ uniform points on the *boundary* of the disk or sphere. From Proposition 1 of Appel and Russo [1], we have for the circle:

$$\lim_{n\to\infty} P\left(n^4(2-H_n) \le w\right) = 1 - \exp\left(-\frac{w^{1/2}}{\pi}\right), \tag{3.29}$$

while for the sphere:

$$\lim_{n\to\infty} P\left(n^2(2-H_n) \le w\right) = 1 - \exp\left(-\frac{w}{2}\right). \tag{3.30}$$

On the boundary sets (circle or sphere surface, respectively), every observation is potentially a vertex of the maximal diameter. The convergence of $H_n$ to 2 is comparatively fast, hence the need for greater magnification of the difference $2 - H_n$ in order to get a weak limit. However, on the entire disk or sphere, since only points observed near the boundary (that is, in a shrinking annulus) can determine the maximal diameter, the convergence of $H_n$ to 2 is comparitively slow.

The appearance of a factor of $\pi$ in the functional form of the limit in the 2- but not in the 3-dimensional case is curious, but is easily explained by noting that the calculations used in the proofs of the 3-dimensional cases compare volumes of the sphere to spherical sub-regions (see, for example (3.24)), whereas the calculations in the 2-dimensional cases do not.

## 4. ACKNOWLEDGMENTS.

Prof. N.D. Shayamalkumar for computational help which provided verification of the correctness of our results in the context of the unit disk and sphere.

## REFERENCES

[1] Appel, M.J., Russo, R.P. (2005). Limiting distributions for the maximum of a symmetric function on a random point set. *J. Theor. Prob.*, to appear.

[2] Appel, M.J.B., Klass, M., Russo, R.P. (1999). Series criteria for the maximum of a symmetric function on a random point set. *J. Theor. Prob.* **12**, 27-47.

[3] Appel, M.J.B., Najim, C.A, Russo, R.P. (2002). Limit laws for the diameter of a random point set. *Adv. Appl. Prob.* **34**, 1-10.

[4] B. Arnold, Balakrishnan, N., and Nagaraja, H.N. (1992). *A First Course in Order Statistics*, Wiley.

[5] Dette, H. and Henze, N. (1989) The limit distribution of the largest nearest-neighbor link in the unit d-cube. *J. Appl. Prob.* **26**, 67-80.

[6] Durrett, R. (2005). *Probability: Theory and Examples*, 3rd edition, Brooks/Cole.

[7] Ellis, R.B., Jia, X., and Yan, C. (2006) On random points in the unit disk. *Rand. Struct. and Alg.*, to appear.

[8] http://mathworld.wolfram.com/SphericalCap.html.

[9] Penrose, M. D., (2003). *Random Geometric Graphs*, Oxford University Press.

[10] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley.

[11] Steele, J.M. and Tierney, L. (1986) Boundary domination and the distribution of the largest nearest-neighbor link. *J. Appl. Prob.* **23**, 524-528.

[12] van der Vaart, A.W., Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*, Springer-Verlag.