MAXIMA IN HYPERCUBES

Zhi-Dong Bai[1]
College of Mathematics and Statistics
Northeast Normal University
130024, Changchun
Jilin, PRC
and
Department of Statistics & Applied Probability
National University of Singapore
3 Science Drive 2, 117543
Singapore

Luc Devroye
School of Computer Science
McGill University
Montreal, H3A 2K6
Canada

Hsien-Kuei Hwang[2], Tsung-Hsi Tsai[3]
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan

August 17, 2004

Abstract

We derive a Berry-Esseen bound, essentially of the order of the square of the standard deviation, for the number of maxima in random samples from $(0,1)^d$. The bound is, although not optimal, the first of its kind for the number of maxima in dimensions higher than two. The proof uses Poisson processes and Stein's method. We also propose a new method for computing the variance and derive an asymptotic expansion. The methods of proof we propose are of some generality and applicable to other regions such as d-dimensional simplex.

1 Introduction

Maxima. A point $\mathbf{p}$ in $\mathbb{R}^d$ is said to *dominate* another point $\mathbf{q}$ if the difference $\mathbf{p}-\mathbf{q}$ has only nonnegative coordinates. We write $\mathbf{q} \prec \mathbf{p}$ or $\mathbf{p} \succ \mathbf{q}$. The nondominated points in a set of points are called *maxima*.

[1]Supported in part by NSFC Grant 201471000 and NUS Grant R-155-000-040-112.
[2]Partially supported by a Research Award of the Alexander von Humboldt Foundation.
[3]Partially supported by NSC under the Grant `NSC-92-2118-M-001-037`.

The aim of this paper is to (i) derive an asymptotic expansion for the variance of the number of maxima in random samples independently and identically distributed (iid) in the hypercube $(0,1)^d$, and (ii) derive a central limit theorem with convergence rate for the number of maxima.

The interest of studying dominance and maxima is multifold. First, dominance represents one of the most natural partial orders for multidimensional points, and has been widely used in many scientific disciplines; see Bai et al. (1998, 2001), Chen et al. (2003) for more information and a large number of references. Second, the number of maxima is itself encountered in many applications like analysis of linear programming and of maxima-finding algorithms; see Blair (1986), Devroye (1986), Golin (1994), Dyer and Walker (1998), Chen et al. (2003). Finally, not much is known as far as probabilistic properties of the number of maxima in dimensions higher than two is concerned. Asymptotic estimates for the mean are usually straightforward, but those for the variance are highly nontrivial even in the simplest case of hypercubes; see Bai et al. (1998). Baryshnikov (2000) indicated that the number of maxima in hypercubes is asymptotically normally distributed but without a complete proof; see also Barbour and Xia (2001). While the asymptotic normality for the number of maxima in high-dimensional regions is quite expected, the technicalities required for a rigorous proof may not be easy to accomplish.

Maxima in hypercubes. Let $\mathbf{p}_1, \dots, \mathbf{p}_n$ be a sequence of iid points chosen uniformly at random from $(0,1)^d$, $d \geq 2$. Denote by $K_n = K_{n,d}$ the number of maxima in $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$.

Distributional properties of K_n were first investigated by Barndorff-Nielsen and Sobel (1966), where they showed that

$$\mathbb{E}[K_n] = \frac{n}{(d-1)!} \int_0^1 (1-x)^{n-1} (-\log x)^{d-1} \, \mathrm{d}x. \tag{1}$$

From this, one obtains

$$\mathbb{E}[K_n] = \frac{(\log n)^{d-1}}{(d-1)!} + O((\log n)^{d-2}), \tag{2}$$

as $n \to \infty$ and d fixed; the asymptotic approximation (2) was first derived in Kuksa and Šor (1972) for $d = 2, 3$, and for general d in Berezovskiĭ and Travkin (1975) (with an asymptotic expansion); Ivanin (1975) considered $\mathbb{E}[K_n]$ under more general model allowing correlations between coordinates.

Similar results were later rederived in Vout (1973), Calpine and Golding (1976), Bentley et al. (1978), Devroye (1980), O'Neill (1980), Buchta (1989). See also Hwang (2002) for a uniform Poisson-type estimate for $\mathbb{E}[K_n]$ when d varies with n. We also collect about a dozen of different expressions for $\mathbb{E}[K_n]$ in Appendix.

The variance of K_n was first studied by Barndorff-Nielsen and Sobel (1966) for $d = 3$, and then by Ivanin (1976), Bai et al. (1998), Devroye (1999) for general d. It satisfies (see Bai et al., 1998)

$$\frac{\mathbb{V}[K_n]}{(\log n)^{d-1}} = \frac{1}{(d-1)!} + \kappa_d + O\left((\log n)^{-1}\right), \tag{3}$$

where

$$\kappa_d = \frac{1}{(d-1)!} \sum_{m \geq 1} \frac{1}{m^2} \sum_{1 \leq j,k \leq m} \binom{m}{j} \binom{m}{k} (-1)^{j+k} jk \left(\left(j^{-1} + k^{-1}\right)^{d-1} - j^{1-d} - k^{1-d} \right),$$

for $d \geq 2$. Recently, Carlsund (2003) provides more terms for (3) for $d = 3$, the main motivation of obtaining more terms being that the convergence of $\mathbb{V}[K_{n,3}]/(\log n)^2$ towards κ_3 is numerically very slow.

We give in this paper a simple method for deriving the following general approximation.

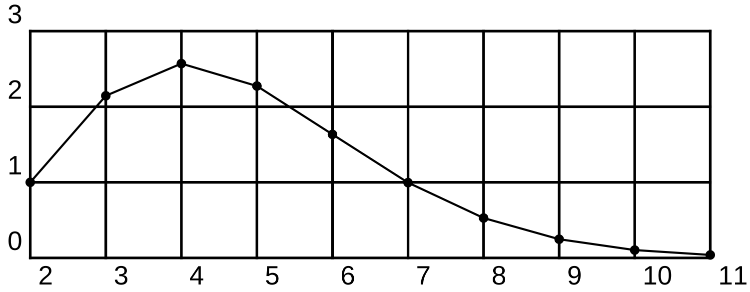

Figure 1: A plot of the values of $\kappa_d + 1/(d-1)!$ for d from 2 to 11.

Theorem 1. *For $d \geq 2$, the variance of K_n satisfies*

$$\frac{\mathbb{V}[K_n]}{(\log n)^{d-1}} = \sum_{0 \leq j \leq d-1} \frac{1}{(d-1-j)!} \left(\frac{(-1)^j}{j!} \Gamma^{(j)}(1) + c_{dj} \right) (\log n)^{-j} + O(n^{-1}(\log n)^{d-1}), \tag{4}$$

where Γ denotes the Gamma function and c_{dj} is defined as the coefficient of u^j in the Taylor expansion of

$$\begin{aligned} &- \frac{2\Gamma(2-u)}{(d-1)!} \int_0^1 \frac{(-\log x)^{d-1}}{(1+x)^{2-u}} \, \mathrm{d}x + \sum_{1 \leq k < d} \frac{\binom{d}{k} \Gamma(2-u)}{(k-1)!(d-1-k)!} \\ &\quad \times \int_0^1 \int_0^1 \left(\frac{(-\log x)^{k-1} (-\log z)^{d-1-k}}{(x+z-xz)^{2-u}} - \frac{(-\log x)^{k-1} (-\log z)^{d-1-k}}{(x+z)^{2-u}} \right) \mathrm{d}x \, \mathrm{d}z. \end{aligned} \tag{5}$$

The first few values of $\kappa_d = c_{d0}/(d-1)!$ are given below in (14). Numerically, the limiting constant of $\mathbb{V}[K_n]/(\log n)^{d-1}$ first increases and then decreases, with a maximum reached at $d = 4$; see Figure 1. One can show, using (13) below, that κ_d goes to zero factorially fast as d increases.

The error term in (4) is not optimal and can be improved if required, and the same method of proof can be extended to other regions.

While numerical calculations of the integrals in (5) are very time-consuming, the expansion (4) represents the first of such approximations for the variance. It remains open how the integrals in (5) may be further simplified.

A CLT with a rate for K_n. We next derive a Berry-Esseen bound for K_n. The proof given here is not only self-contained, but also the first complete one; cf. Baryshnikov (2000).

Let $\Phi(x)$ denote the distribution function of the standard normal distribution

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \, \mathrm{d}t.$$

Suppose that $Y_1, Y_2, \ldots$ is a sequence of random variables. Write $\{Y_n\} \in \mathrm{CLT}(r_n)$, if

$$\sup_x \left| \mathbb{P}\left(\frac{Y_n - \mathbb{E}[Y_n]}{\sqrt{\mathbb{V}[Y_n]}} < x \right) - \Phi(x) \right| = O(r_n),$$

where $r_n \to 0$. A sequence r_n is referred to as a *convergent sequence* if $r_n \to 0$.

Theorem 2. *The sequence of random variables K_n satisfies*

$$\{K_n\} \in \mathrm{CLT}\left((\log n)^{-(d-1)/4} (\log \log n)^d \right). \tag{6}$$

The idea of the proof consists in constructing a sequence of random variables K_{W_n} satisfying the following two estimates.

Proposition 1. *For a convergent sequence* $r_n \geq (\log n)^{-(d-1)/2}$,

$$\{K_n\} \in \mathrm{CLT}\,(r_n) \quad \textit{iff} \quad \{K_{W_n}\} \in \mathrm{CLT}\,(r_n)\,.$$

Proposition 2. *The sequence of the normalized random variables* $K^*_{W_n} := (K_{W_n} - \mathbb{E}[K_{W_n}])/\sqrt{\mathbb{V}[K_{W_n}]}$ *converges to the standard normal distribution with a rate*

$$d_1\left(K^*_{W_n}, N(0,1)\right) = O\left((\log\log n)^{2d}\,(\log n)^{-(d-1)/2}\right),$$

where $N(0,1)$ *denotes a standard normal random variable and*

$$d_1(X,Y) := \sup_h \left\{ |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]| : \sup_x |h(x)| + \sup_x |h'(x)| \leq 1 \right\}.$$

We then show that the the Kolmogorov distance is of the order of $\sqrt{d_1}$, and this will complete the proof of (6).

The main trick used in the paper is the log-transformation first suggested by Baryshnikov (2000). It allows us to observe that nearly all maxima occur in a thin strip sandwiched between two parallel hyperplanes. Switching to a Poisson sample size introduces just enough independence to apply Stein's method. Similar ideas have been explored by Barbour and Xia (2001).

2 Mean and variance of K_n

Define

$$D_n = \{(\mathbf{x},\mathbf{y}) \in [0,1]^d \times [0,1]^d \,:\, \mathbf{x} \text{ incomparable with } \mathbf{y}\},$$

where two points are said to be *incomparable* if none dominates the other. For notational convenience, $\Pi\, x_i$ stands for $\prod_{1\leq i\leq d} x_i$; similarly, $\Sigma\, x_i = \sum_{1\leq i\leq d} x_i$.

Mean. The mean of K_n is easily derived and simplified as follows.

$$\begin{aligned}
\mathbb{E}[K_n] &= n\int_{[0,1]^d} (1-\Pi\, x_i)^{n-1}\ \mathrm{d}\mathbf{x} &(7)\\
&= n\int_{[0,1]^d} e^{-n\Pi\, x_i}\,\mathrm{d}\mathbf{x} + T_1(n)\\
&= \int_{[0,n^{1/d}]^d} e^{-\Pi\, u_i}\,\mathrm{d}\mathbf{u} + T_1(n) \quad (x_i \mapsto n^{-1/d}u_i)\\
&= \int_{[-d^{-1}\log n,\infty)^d} \exp\left(-e^{-\Sigma\, z_i} - \Sigma\, z_i\right)\,\mathrm{d}\mathbf{z} + T_1(n) \quad (z_i \mapsto -\log u_i)\\
&= \frac{1}{(d-1)!}\int_{-\log n}^{\infty} (\log n + x)^{d-1}\exp(-x-e^{-x})\,\mathrm{d}x + T_1(n) \quad (x \mapsto \Sigma\, z_i)\\
&= \frac{1}{(d-1)!}\int_0^n (\log n - \log y)^{d-1} e^{-y}\,\mathrm{d}y + T_1(n) \quad (y \mapsto e^{-x})\\
&= \frac{(\log n)^{d-1}}{(d-1)!}\sum_{0\le j<d}\binom{d-1}{j}\frac{(-1)^j}{(\log n)^j}\int_0^\infty (\log y)^j e^{-y}\,\mathrm{d}y + T_1(n) + O\left(e^{-n}(\log n)^{d-1}\right)\\
&= (\log n)^{d-1}\sum_{0\le j<d}\frac{(-1)^j\Gamma^{(j)}(1)}{j!(d-1-j)!}(\log n)^{-j} + T_1(n) + O\left(e^{-n}(\log n)^{d-1}\right); &(8)
\end{aligned}$$

see also Berezovskiĭ and Travkin (1975) for more involved expressions for the coefficients, and Appendix for other expressions for (7).

Here $T_1(n)$ accounts for two types of errors introduced: (i) when replacing $n-1$ by n, which yields an error of order $O(n^{-1}(\log n)^{d-1})$, and (ii) when replacing $(1-x)^n$ by e^{-nx}, which produces an error of the form

$$\varepsilon_n := n^2\int_{[0,1]^d} e^{-n\Pi\, x_i}\Pi\, x_i^2\,\mathrm{d}\mathbf{x},$$

(by the elementary inequality $e^{-nx}(1-nx^2) \le (1-x)^n \le e^{-nx}$ for $n \ge 0$ and $0 \le x \le 1$); see Lemma 5 in Bai et al. (2001). By the same analysis as above, $\varepsilon_n = O(n^{-1}(\log n)^{d-1})$. Thus the total errors are are of order $O(n^{-1}(\log n)^{d-1})$.

Note that the right order of the error in approximating $\mathbb{E}(K_n)$ by the partial sum on the right-hand side of (8) is indeed $O(n^{-1}(\log n)^{d-2})$. The preceding derivations are written in the given forms so that the same procedure can be easily amended for other integrals.

For simplicity, the symbol $a_n \simeq b_n$ denotes $a_n = b_n + O(n^{-1}(\log n)^{2d-2})$.

Variance. For the second moment, we have

$$\begin{aligned}
\mathbb{E}[K_n^2] - \mathbb{E}[K_n] &= n(n-1)\int_{D_n} \left(1 - \Pi\, x_i - \Pi\, y_i + \Pi\,(x_i \wedge y_i)\right)^{n-2}\, \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \\
&\simeq n^2 \int_{D_n} e^{-n(\Pi\, x_i + \Pi\, y_i - \Pi\,(x_i\wedge y_i))}\, \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \\
&= n^2 \int_{D_n} e^{-n(\Pi\, x_i + \Pi\, y_i)}\, \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} + n^2 \int_{D_n} e^{-n(\Pi\, x_i + \Pi\, y_i)} \left(e^{n\Pi\,(x_i\wedge y_i)} - 1\right)\, \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \\
&\simeq (\mathbb{E}[K_n])^2 - I_{n0} + \sum_{1\le k<d} \binom{d}{k} I_{nk}, \qquad (9)
\end{aligned}$$

where (using the abbreviations $\Pi'\, x_i = \prod_{i=1}^k x_i$ and $\Pi''\, x_i = \prod_{i=k+1}^d x_i$)

$$\begin{aligned}
I_{nk} &:= n^2 \int\limits_{\substack{x_i>y_i,\, 1\le i\le k\\ x_i<y_i,\, k<i\le d}} e^{-n(\Pi\, x_i + \Pi\, y_i)} \left(e^{n\Pi'\, y_i \Pi''\, x_i} - 1\right)\, \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}, \\
I_{n0} &= 2n^2 \int\limits_{\substack{[0,1]^d\times[0,1]^d\\ \mathbf{y}\prec\mathbf{x}}} e^{-n(\Pi\, x_i + \Pi\, y_i)}\, \mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}.
\end{aligned}$$

The main difference between the current proof and that in Bai et al. (1998) is that we subtract the square of the mean at this early step, so that the large amount of cancellations caused by subtracting $(\mathbb{E}[K_n])^2$ is easier to manage.

By the changes of variables

$$x_i \mapsto u_i,\ y_i \mapsto u_i v_i \quad \text{for } i \le k \qquad \text{and} \qquad x_i \mapsto u_i v_i,\ y_i \mapsto u_i \quad \text{for } i > k,$$

we have

$$\begin{aligned}
I_{nk} &= n^2 \int_{[0,1]^d\times[0,1]^d} e^{-n\Pi\, u_i(\Pi'\, v_i + \Pi''\, v_i)} \left(e^{n\Pi\, u_i v_i} - 1\right) \Pi\, u_i\, \mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v} \\
I_{n0} &= 2n^2 \int_{[0,1]^d\times[0,1]^d} e^{-n\Pi\, u_i(1+\Pi\, v_i)} \Pi\, u_i\, \mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v}.
\end{aligned}$$

By applying the same procedure as for $\mathbb{E}[K_n]$, we have

$$\begin{aligned}
I_{n0} &= \frac{2}{(d-1)!} \int_0^n (\log n - \log y)^{d-1} y \int_{[0,1]^d} e^{-y(1+\Pi\, v_i)}\, \mathrm{d}\mathbf{v}\,\mathrm{d}y \\
&\simeq 2\frac{(\log n)^{d-1}}{(d-1)!^2} \sum_{0\le j<d} \binom{d-1}{j} \frac{1}{(\log n)^j} \int_0^\infty \int_0^1 y(-\log y)^j (-\log z)^{d-1} e^{-y(1+z)}\, \mathrm{d}z\,\mathrm{d}y.
\end{aligned}$$

Denote by $[u^j]f(u)$ the coefficient of u^j in the Taylor expansion of f. Then

$$\int_0^\infty y(-\log y)^j e^{-y(1+x)}\, \mathrm{d}y = j![u^j]\frac{\Gamma(2-u)}{(1+x)^{2-u}},$$

and

$$I_{n0} \simeq 2\frac{(\log n)^{d-1}}{(d-1)!}\sum_{0\le j<d}\frac{(\log n)^{-j}}{(d-1-j)!}[u^j]\Gamma(2-u)\int_0^1 \frac{(-\log x)^{d-1}}{(1+x)^{2-u}}\,\mathrm{d}x. \tag{10}$$

Note that

$$j![u^j]\frac{\Gamma(2-u)}{(1+x)^{2-u}} = \sum_{0\le \ell\le j}\binom{j}{\ell}(-1)^{j-\ell}\Gamma^{(j-\ell)}(2)\frac{\log^{\ell}(1+x)}{(1+x)^2}.$$

Similarly, for $1 \le k \le d-1$, we have

$$\begin{aligned}
I_{nk} &= \frac{1}{(d-1)!}\int_0^n\int_{[0,1]^d}(\log n-\log y)^{d-1} y e^{-y(\Pi' v_i+\Pi'' v_i)}\left(e^{y\Pi v_i}-1\right)\,\mathrm{d}\mathbf{v}\,\mathrm{d}y\\
&\simeq \frac{(\log n)^{d-1}}{(d-1)!}\sum_{0\le j<d}\binom{d-1}{j}\frac{(\log n)^{-j}}{(k-1)!(d-1-k)!}\\
&\quad\times\int_0^\infty\int_0^1\int_0^1(-\log y)^j(-\log x)^{k-1}(-\log z)^{d-1-k} y e^{-y(x+z)}\left(e^{yxz}-1\right)\,\mathrm{d}x\,\mathrm{d}z\,\mathrm{d}y\\
&= (\log n)^{d-1}\sum_{0\le j<d}\frac{(\log n)^{-j}}{(d-1-j)!(k-1)!(d-1-k)!}\\
&\quad\times [u^j]\Gamma(2-u)\int_0^1\int_0^1\left(\frac{(-\log x)^{k-1}(-\log z)^{d-1-k}}{(x+z-xz)^{2-u}}-\frac{(-\log x)^{k-1}(-\log z)^{d-1-k}}{(x+z)^{2-u}}\right)\,\mathrm{d}x\,\mathrm{d}z.
\end{aligned} \tag{11}$$

Thus we obtain (4) for the variance of K_n, and the expression (5) follows from (9), (10) and (11). This completes the proof of (4). □

A quick check. Take $d = 2$ in (4), we obtain

$$\mathbb{V}[K_{n,2}] = (1+c_{20})\log n + \gamma + c_{21} + O\left(n^{-1}(\log n)^2\right),$$

where

$$\begin{aligned}
c_{20} &= 2\int_0^1\frac{\log x}{(1+x)^2}\,\mathrm{d}x + 2\int_0^1\int_0^1\left(\frac{1}{(x+z-xz)^2}-\frac{1}{(x+z)^2}\right)\,\mathrm{d}x\,\mathrm{d}z,\\
c_{21} &= 2\int_0^1\frac{\log x}{(1+x)^2}\left(\log(1+x)-1+\gamma\right)\,\mathrm{d}x\\
&\quad + 2\int_0^1\int_0^1\left(\frac{\log(x+z-xz)-1+\gamma}{(x+z-xz)^2}-\frac{\log(x+z)-1+\gamma}{(x+z)^2}\right)\,\mathrm{d}x\,\mathrm{d}z.
\end{aligned}$$

It is then straightforward to check that $c_{20} = 0$ and $c_{21} = -\pi^2/6$.

The leading constant κ_d. By comparing the two equations (3) and (4), we obtain an alternative expression for κ_d

$$\begin{aligned}
\kappa_d &= -\frac{2}{(d-1)!^2}\int_0^1\frac{(-\log x)^{d-1}}{(1+x)^2}\,\mathrm{d}x + \frac{1}{(d-1)!}\sum_{1\le k<d}\binom{d}{k}\frac{1}{(k-1)!(d-1-k)!}\\
&\quad\times\int_0^1\int_0^1(-\log x)^{k-1}(-\log z)^{d-1-k}\left(\frac{1}{(x+z-xz)^2}-\frac{1}{(x+z)^2}\right)\,\mathrm{d}x\,\mathrm{d}z.
\end{aligned} \tag{12}$$

Yet another expression for κ_d can be derived from the main theorem in Bai et al. (1998) ($\mu_{n,d} := \mathbb{E}[K_n]$)

$$\begin{aligned}\kappa_d &= \sum_{1\le k\le d-2} \frac{1}{k!(d-1-k)!} \sum_{m\ge 1} \frac{\mu_{m,k}\,\mu_{m,d-1-k}}{m^2} \\ &= \sum_{1\le k\le d-2} \frac{1}{k!(d-1-k)!(k-1)!(d-2-k)!} \int_0^1\!\!\int_0^1 \frac{(-\log x)^{k-1}(-\log z)^{d-2-k}}{x+z-xz}\,\mathrm{d}x\,\mathrm{d}z, \qquad (13)\end{aligned}$$

where we use the integral representation (1).

A natural question then is how to prove directly the identity implied by equating (13) to (12). Note that numerically the use of (13) is preferable to (12).

The exact values of κ_d for $d \le 8$. The first few κ_d's can be explicitly expressed in terms of Riemann's zeta function $\zeta(s)$ as follows.

$$\begin{cases} \kappa_2 = 0, \\ \kappa_3 = \zeta(2), \\ \kappa_4 = 2\zeta(3), \\ \kappa_5 = \frac{33}{16}\zeta(4), \\ \kappa_6 = \frac{5}{4}\zeta(5) + \frac{1}{6}\zeta(2)\zeta(3), \\ \kappa_7 = \frac{1451}{1728}\zeta(6) + \frac{7}{72}\zeta(3)^2, \\ \kappa_8 = \frac{1729}{5760}\zeta(7) + \frac{13}{360}\zeta(2)\zeta(5) + \frac{181}{1440}\zeta(3)\zeta(4). \end{cases} \qquad (14)$$

The values of $\kappa_2, \ldots, \kappa_6$ are already given in Bai et al. (1998).

To see how these values are obtained from (12), we start from the integral

$$\begin{aligned}-\frac{2}{(d-1)!^2}\int_0^1 \frac{(-\log x)^{d-1}}{(1+x)^2}\,\mathrm{d}x &= -\frac{2}{(d-1)!}\sum_{k\ge 1}(-1)^{k-1}k^{1-d} \\ &= -\frac{2}{(d-1)!}(1-2^{2-d})\zeta(d-1),\end{aligned}$$

for $d \ge 2$, where $(1-2^{2-d})\zeta(d-1) = \log 2$ for $d = 2$.

Similarly, the integrals with $k = 1$ in (12) satisfy

$$\begin{aligned}&\frac{d}{(d-1)!(d-2)!}\int_0^1\!\!\int_0^1 (-\log z)^{d-2}\left(\frac{1}{(x+z-xz)^2} - \frac{1}{(x+z)^2}\right)\mathrm{d}x\,\mathrm{d}z \\ &= \frac{d}{(d-1)!(d-2)!}\int_0^1 \frac{(-\log z)^{d-2}}{1+z}\,\mathrm{d}z \\ &= \frac{d}{(d-1)!}(1-2^{2-d})\zeta(d-1),\end{aligned}$$

and the integrals corresponding to $k = d-1$ in κ_d give the same value.

For the integrals with $k = 2$ and $k = d-2$ in (12), we have

$$\begin{aligned}&\frac{d}{2(d-2)!(d-3)!}\int_0^1\!\!\int_0^1 (-\log x)(-\log z)^{d-3}\left(\frac{1}{(x+z-xz)^2} - \frac{1}{(x+z)^2}\right)\mathrm{d}x\,\mathrm{d}z \\ &= \frac{d}{2(d-2)!}(d-3+2^{2-d})\zeta(d-1);\end{aligned}$$

and for $k=3$ and $k=d-4$ ($\mathrm{Li}_2(z) := \sum_{m\ge1} z^m/m^2$)

$$\frac{d}{12(d-3)!(d-4)!}\int_0^1\!\!\int_0^1 (-\log x)^2(-\log z)^{d-4}\left(\frac{1}{(x+z-xz)^2}-\frac{1}{(x+z)^2}\right)\mathrm{d}x\,\mathrm{d}z$$
$$=\frac{d}{12(d-3)!}\Bigg(\left((d-2)(d-3)+2-2^{3-d}\right)\zeta(d-1)+2\zeta(2)\zeta(d-3)$$
$$-2\sum_{m\ge1}\frac{H_m^{(2)}}{m^{d-3}}-2(d-3)\sum_{m\ge1}\frac{H_m}{m^{d-2}}\Bigg)\qquad (d\ge5),$$

where $H_m^{(2)} := \sum_{1\le j\le m} 1/j$ and $H_m^{(2)} := \sum_{1\le j\le m} 1/j^2$.

These relations, together with the identities (see Flajolet and Salvy, 1998)

$$\sum_{m\ge1}\frac{H_m}{m^5}=\tfrac{7}{2}\zeta(6)-\zeta(2)\zeta(4)-\tfrac12\zeta(3)^2,\qquad \sum_{m\ge1}\frac{H_m^{(2)}}{m^4}=\zeta(3)^2-\tfrac13\zeta(6),$$

give the values of κ_d, $2\le d\le 7$.

Finally, the value of κ_8 is obtained by a lengthy *ad hoc* calculation via Euler sums (see Flajolet and Salvy, 1998).

Remark. The same approach is applicable to the number of maxima $\hat{K}_n$ in d-simplex $\{\mathbf{x} : x_i > 0, x_1 + \cdots + x_d \le 1\}$ for which we have

$$\mathbb{E}[\hat{K}_n] \sim \Gamma\left(\tfrac1d\right) n^{(d-1)/d},$$
$$\mathbb{V}[\hat{K}_n] \sim C_d n^{(d-1)/d},$$

where

$$C_d := \Gamma\left(\tfrac1d\right) - 2\Gamma\left(\tfrac1d\right)\int_0^1 \frac{(1-x)^{d-1}}{(1+x^d)^{1+1/d}}\,\mathrm{d}x$$
$$+2(d-1)\Gamma\left(\tfrac1d\right)\sum_{1\le k<d}\binom{d}{k}\binom{d-2}{k-1}\int_0^1(1-x)^{k-1}\int_0^1(1-xz)^{d-1-k}z^k$$
$$\times\left(\frac{1}{(1+z^d-x^dz^d)^{1+1/d}}-\frac{1}{(1+z^d)^{1+1/d}}\right)\mathrm{d}z\,\mathrm{d}x.$$

3 A Berry-Esseen bound for K_n

The proof of Proposition 1 is divided into several steps.

A log-transformation

Assume now that $\mathbf{p}_1,\dots,\mathbf{p}_n$ are iid points uniformly distributed in the cube $(-1,0)^d$. A crucial step in our analysis is to apply the log-transformation first introduced by Baryshnikov (2000): $\mathbf{x}=(x_1,\dots,x_d)\to\mathbf{y}=(y_1,\dots,y_d)$, where

$$y_i = -\log(-x_i),\quad i=1,\dots,d,$$

from $(-1,0)^d$ to $\mathbb{R}_+^d = \{\mathbf{x} : x_i > 0 \text{ for all } i = 1, \ldots, d\}$. Such a transformation preserves the dominance relation, and the maximal points are thus unchanged. Denote by $\mathbf{q}_1, \ldots, \mathbf{q}_n$ the images of $\mathbf{p}_1, \ldots, \mathbf{p}_n$ under such a transformation. Then the components of $\mathbf{q}_1$ are iid with exponential distribution ($\lambda = 1$). We define $\|\mathbf{x}\| = x_1 + \cdots + x_d$ for $\mathbf{x} \in \mathbb{R}_+^d$. Then $\|\mathbf{q}_1\|$ has a gamma distribution with parameter $(d, 1)$, that is, the density function of $\|\mathbf{q}_1\|$ is $e^{-x}x^{d-1}/(d-1)!$ for $x > 0$ and zero otherwise.

Approximation of K_n by the number of maxima in a strip

Let $B_\alpha = \{\mathbf{x} : \|\mathbf{x}\| > \alpha\} \cap \mathbb{R}_+^d$ and $B_\alpha^c = \{\mathbf{x} : \|\mathbf{x}\| \le \alpha\} \cap \mathbb{R}_+^d$. Take

$$\alpha = \log n - \log\left(4(d-1)\log\log n\right),$$
$$\beta = \log n + 4(d-1)\log\log n.$$

Let V_n be the event that there is no point of $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ falling in B_β. Let $\overline{K}_n$ be the number of maxima falling in B_α when conditioning on V_n. We first prove that for a convergent sequence $r_n \ge (\log n)^{-(d-1)/2}$,

$$\{K_n\} \in \mathrm{CLT}\,(r_n) \quad \text{iff} \quad \{\overline{K}_n\} \in \mathrm{CLT}\,(r_n)\,. \tag{15}$$

The following Lemma is needed to prove (15).

Lemma 1. *Let X_n, Y_n be two sequences of random variables and r_n be a convergent sequence. Suppose that (i) the total variation distance $d(X_n, Y_n)$ between X_n and Y_n is bounded above by $O(r_n)$, (ii)*

$$|\mathbb{E}[X_n] - \mathbb{E}[Y_n]| = O(r_n\sqrt{\mathbb{V}[X_n]}),$$

and (iii)

$$|\mathbb{V}[X_n] - \mathbb{V}[Y_n]| = O(r_n\sqrt{\mathbb{V}[X_n]}).$$

Then $\{X_n\} \in \mathrm{CLT}(r_n)$ iff $\{Y_n\} \in \mathrm{CLT}(r_n)$.

Let $N_n(A)$ denote the number of points of $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ falling in A and $K_n(A)$ denote the number of maxima of $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ falling in A. Clearly, $K_n(A) \le N_n(A)$.

To prove (15), we first decompose K_n into three parts

$$K_n = K_n(B_\alpha)1_{V_n} + K_n(B_\alpha)1_{V_n^c} + K_n(B_\alpha^c).$$

Note that

$$\overline{K}_n = K_n(B_\alpha)1_{V_n} \text{ conditioned on } V_n.$$

To apply Lemma 1, we first derive the three estimates:

$$d(K_n, \overline{K}_n) \le \mathbb{P}(K_n(B_\alpha)1_{V_n^c} \ge 1) + \mathbb{P}(K_n(B_\alpha^c) \ge 1) + \mathbb{P}(V_n^c),$$
$$\left|\mathbb{E}[K_n] - \mathbb{E}[\overline{K}_n]\right| \le \left|1 - \frac{1}{\mathbb{P}(V_n)}\right|\mathbb{E}[K_n] + \frac{\mathbb{E}[K_n(B_\alpha)1_{V_n^c}] + \mathbb{E}[K_n(B_\alpha^c)]}{\mathbb{P}(V_n)},$$
$$\left|\mathbb{V}[K_n] - \mathbb{V}[\overline{K}_n]\right| \le \left|\mathbb{E}[K_n^2] - \mathbb{E}[\overline{K}_n^2]\right| + \left|\mathbb{E}[K_n] - \mathbb{E}[\overline{K}_n]\right|\left(\mathbb{E}[K_n] + \mathbb{E}[\overline{K}_n]\right),$$

where

$$\left|\mathbb{E}[K_n^2] - \mathbb{E}[\overline{K}_n^2]\right| \le \left|1 - \frac{1}{\mathbb{P}(V_n)}\right| \mathbb{E}[K_n^2] + \frac{\mathbb{E}[K_n^2(B_\alpha)1_{V_n^c}] + \mathbb{E}[K_n^2(B_\alpha^c)] + 2\mathbb{E}[K_n^2(B_\alpha)]\mathbb{E}[K_n^2(B_\alpha^c)]}{\mathbb{P}(V_n)}.$$

Recall that $\mathbb{E}[K_n] \asymp (\log n)^{d-1}$ and $\mathbb{E}[K_n^2] \asymp (\log n)^{2(d-1)}$; see (2) and (3). By Chebyshev's inequality $\mathbb{P}(V_n^c) \le E[N_n(B_\beta)]$.

We claim that

(i) $\mathbb{E}[N_n(B_\beta)] = O((\log n)^{-3(d-1)})$,

(ii) $\mathbb{E}[K_n(B_\alpha^c)] = O((\log n)^{-(d-1)})$,

(iii) $\mathbb{E}[K_n(B_\alpha)1_{V_n^c}] = O((\log n)^{-(d-1)})$,

(iv) $\mathbb{E}[K_n^2(B_\alpha)1_{V_n^c}] = O(1)$, and

(v) $\mathbb{E}[K_n^2(B_\alpha^c)] = O((\log n)^{-2(d-1)})$.

It follows from these claims that

$$\begin{aligned} d(K_n, \overline{K}_n) &= O((\log n)^{-(d-1)}), \\ \left|\mathbb{E}[K_n] - \mathbb{E}[\overline{K}_n]\right| &= O((\log n)^{-(d-1)}), \\ \left|\mathbb{V}[K_n] - \mathbb{V}[\overline{K}_n]\right| &= O(1). \end{aligned}$$

Thus the proof of (15) is reduced, by Lemma 1, to proving the five claims.

Proof of (i).

$$\begin{aligned} \mathbb{E}[N_n(B_\beta)] &= n\mathbb{P}\left(\|\mathbf{q}_1\| \ge \beta\right) \\ &= n\int_\beta^\infty \frac{x^{d-1}}{(d-1)!}e^{-x}\,\mathrm{d}x \\ &= O\left(n\beta^{d-1}e^{-\beta}\right) \\ &= O\left((\log n)^{-3(d-1)}\right). \end{aligned}$$

Proof of (ii). Let $U_{\mathbf{y}} = \{\mathbf{z} : z_i > y_i, i = 1, \dots, d\}$ be the first quadrant of $\mathbf{y}$. Then the probability that $\mathbf{q}_1$ falls in $U_{\mathbf{y}}$ is given by

$$\int_{\|\mathbf{y}\|}^\infty \frac{(x - \|\mathbf{y}\|)^{d-1}}{(d-1)!}e^{-x}\,\mathrm{d}x = e^{-\|\mathbf{y}\|}.$$

Thus, given $\mathbf{q}_1$, the conditional probability that $\mathbf{q}_1$ is a maximal point satisfies $(1-e^{-\|\mathbf{q}_1\|})^{n-1} \leq e^{-(n-1)e^{-\|\mathbf{q}_1\|}}$. Therefore,

$$\begin{aligned}
\mathbb{E}[K_n(B_\alpha^c)] &= n\mathbb{P}\left(\mathbf{q}_1 \text{ is a maximal point falling in } B_\alpha^c\right) \\
&\leq n\int_0^\alpha \frac{x^{d-1}}{(d-1)!} e^{-(n-1)e^{-x}-x}\,\mathrm{d}x \\
&\leq \frac{n(\log n)^{d-1}}{(d-1)!}\int_0^\alpha e^{-(n-1)e^{-x}-x}\,\mathrm{d}x \\
&\leq \frac{n(\log n)^{d-1}}{(d-1)!}\cdot\frac{e^{-(n-1)e^{-\alpha}}}{n-1} \\
&= O\left((\log n)^{-3(d-1)}\right).
\end{aligned}$$

Proof of (iii). Let A_i be the event that $\mathbf{q}_i$ lies in B_β. Then $V_n^c \subset \cup_{i=1}^n A_i$. The number of maxima in A_i is less than the number of maxima in $\{\mathbf{q}_1, \ldots, \mathbf{q}_{i-1}, \mathbf{q}_{i+1}, \ldots, \mathbf{q}_n\} + 1$. Note that $\mathbb{P}(A_i) = \mathbb{P}\left(\|\mathbf{q}_1\| \geq \beta\right) = O\left(n^{-1}(\log n)^{-3(d-1)}\right)$. Thus

$$\mathbb{E}[K_n 1_{A_i}] \leq \mathbb{P}(A_i)(\mathbb{E}[K_{n-1}] + 1) = O(n^{-1}(\log n)^{-2(d-1)}),$$

and then

$$\mathbb{E}[K_n(B_\alpha)1_{V_n^c}] \leq \mathbb{E}[K_n 1_{V_n^c}] = O((\log n)^{-2(d-1)}).$$

Proof of (iv). Similarly,

$$\mathbb{E}[K_n^2 1_{A_i}] \leq \mathbb{P}(A_i)(\mathbb{E}[K_{n-1}^2] + 2\mathbb{E}[K_{n-1}] + 1) = O(n^{-1}(\log n)^{-(d-1)}).$$

Thus

$$\mathbb{E}[K_n^2(B_\alpha)1_{V_n^c}] \leq \mathbb{E}[K_n^2 1_{V_n^c}] = O((\log n)^{-(d-1)}).$$

Proof of (v). Given $\mathbf{q}_1, \mathbf{q}_2$, the conditional probability that $\mathbf{q}_3$ falls in $U_{\mathbf{q}_1} \cup U_{\mathbf{q}_2}$ is

$$\mathbb{P}(U_{\mathbf{q}_1}) + \mathbb{P}(U_{\mathbf{q}_2}) - \mathbb{P}(U_{\mathbf{q}_1} \cap U_{\mathbf{q}_2}) \geq \frac{1}{2}\left(e^{-\|\mathbf{q}_1\|} + e^{-\|\mathbf{q}_2\|}\right);$$

the conditional probability that both $\mathbf{q}_1$ and $\mathbf{q}_2$ are maxima is less than

$$\left(1 - \frac{1}{2}\left(e^{-\|\mathbf{q}_1\|} + e^{-\|\mathbf{q}_2\|}\right)\right)^{n-2} \leq e^{-\frac{1}{2}(n-2)\left(e^{-\|\mathbf{q}_1\|}+e^{-\|\mathbf{q}_2\|}\right)}.$$

We thus have

$$\begin{aligned}
\mathbb{E}[K_n^2(B_\alpha^c)] &= \mathbb{E}\left[\sum_{1\leq i\leq n} 1_{\mathbf{q}_i \text{ is maximal and } \|\mathbf{q}_i\|\leq\alpha}\right]^2 \\
&= \mathbb{E}[K_n(B_\alpha^c)] + n(n-1)\mathbb{P}(\text{both } \mathbf{q}_1 \text{ and } \mathbf{q}_2 \text{ are maxima falling in } B_\alpha^c) \\
&\leq \mathbb{E}[K_n(B_\alpha^c)] + \frac{n^2}{[(d-1)!]^2}\int_0^\alpha\int_0^\alpha (xy)^{d-1}e^{-\frac{1}{2}(n-2)[e^{-x}+e^{-y}]-x-y}\,\mathrm{d}x \\
&\leq \mathbb{E}[K_n(B_\alpha^c)] + \frac{n^2(\log n)^{2(d-1)}}{[(n-2)(d-1)!]^2}e^{-(n-2)e^{-\alpha}} \\
&= O\left((\log n)^{-2(d-1)}\right).
\end{aligned}$$

A Poisson process approximation

Construct a Poisson process $\{\mathbf{W}_n\}$ on $S_n = B_\alpha \cap B_\beta^c$ with intensity function

$$\lambda_n = \frac{ne^{-\|\mathbf{w}\|}}{\mathbb{P}(\|\mathbf{q}_1\| \leq \beta)},$$

Denote by N_w the number of points of the Poisson process falling in S_n. Also, let K_{W_n} denote the number of maxima of the Poisson process and $\overline{N}_n$ be the number of points that falls in S_n when conditioning on V_n. It is easy to see that the conditional distribution of $\overline{K}_n$ given $\overline{N}_n = m$ is identical to the conditional distribution of K_{W_n} given $N_w = m$. Thus, the total variation distance between $\overline{K}_n$ and K_{W_n} satisfies (see Prohorov, 1953)

$$\begin{aligned}
&\sup_A \left|\mathbb{P}(\overline{K}_n \in A) - \mathbb{P}(K_{W_n} \in A)\right| \\
&= \sup_A \left| \sum_{0\leq m\leq n} \mathbb{P}(\overline{N}_n = m)\mathbb{P}(\overline{K}_n \in A|\overline{N}_n = m) - \sum_{m\geq 0} \mathbb{P}(N_w = m)\mathbb{P}(K_{W_n} \in A|N_w = m) \right| \\
&\leq \sum_{m\geq 0} \left|\mathbb{P}(\overline{N}_n = m) - \mathbb{P}(N_w = m)\right| \\
&= O(p_n),
\end{aligned}$$

(the implied constant can be taken to be 2; see Barbour et al., 1992), where

$$\begin{aligned}
p_n &:= \mathbb{P}(\mathbf{q}_1 \in S_n | \|\mathbf{q}_1\| \leq \beta) \\
&= \frac{1}{\mathbb{P}(\|\mathbf{q}_1\| \leq \beta)} \int_\alpha^\beta \frac{x^{d-1}}{(d-1)!} e^{-x} \,\mathrm{d}x \\
&= O\left(\frac{(\log n)^{d-1} \log\log n}{n}\right).
\end{aligned}$$

Similarly, we have the two estimates

$$\begin{aligned}
\left|\mathbb{E}[\overline{K}_n] - \mathbb{E}[K_{W_n}]\right| &\leq np_n^2, \\
\left|\mathbb{E}[\overline{K}_n(\overline{K}_n - 1)] - \mathbb{E}[K_{W_n}(K_{W_n} - 1)]\right| &\leq n(n-1)p_n^3.
\end{aligned}$$

The above three estimates imply, by Lemma 1, that for a convergent sequence $r_n \geq (\log n)^{-(d-1)/2}$

$$\{\overline{K}_n\} \in \text{CLT}\,(r_n) \quad \text{iff} \quad \{K_{W_n}\} \in \text{CLT}\,(r_n)\,.$$

A central limit theorem for K_{W_n}

We prove in this section Proposition 2 by applying Stein's method.

Split $\mathbb{R}_+^d$ into cubes $G_{n,m,v}$ of edge-length $1/2^m$, where $m \geq 0$. Let

$$Z_{n,m,v} := \min\left(1, \text{ the number of maxima of } \mathbf{W}_n \text{ falling in the cell } G_{n,m,v}\right).$$

Observe that $\sum_v Z_{n,m,v}$ is nondecreasing in m, where the sum runs over all possible indices for cells, and

$$\lim_{m\to\infty} \mathbb{P}\left(\sum_v Z_{n,m,v} = K_{W_n}\right) = 1,$$

for fixed n. Thus

$$\lim_{m\to\infty} d_1\left(K^*_{W_n}, \frac{\sum_v (Z_{n,m,v} - \mathbb{E}[Z_{n,m,v}])}{\sqrt{\mathbb{V}[\sum_v Z_{n,m,v}]}}\right) = 0;$$

the same result also holds under the total variation distance.

For convenience, we write $G_v = G_{n,m,v}$ and $Z_v = Z_{n,m,v}$. In the following proof, n is a large integer and m is suitably chosen (whose value depends on n). To prove Proposition 2, it is sufficient to prove the same convergence rate for $\{\sum_v Z_v\}$ (normalized) to $N(0,1)$ for sufficiently large m, and for that purpose, we apply Stein's method (as formulated in Theorem 6.33, Janson et al., 2000), where the Stein remainder term is expressed in terms of the d_1-distance. It suffices to verify that

$$\frac{M_n Q_n^2}{(\mathbb{V}[\sum_v Z_v])^{3/2}} = O\left((\log\log n)^{2d} (\log n)^{-(d-1)/2}\right),$$

where

$$\begin{aligned} M_n &= \sum_v \mathbb{E}[Z_v] \le \mathbb{E}[K_{W_n}] = O\left((\log n)^{d-1}\right), \\ Q_n &= \max_{j,k} \sum_{Z_v \text{ dependent on } Z_j \text{ or } Z_k} \mathbb{E}[Z_v|Z_j, Z_k] = O\left((\log\log n)^d\right). \end{aligned} \tag{16}$$

For large enough m, $\mathbb{V}[\sum_v Z_v] \sim \mathbb{V}[K_{W_n}] \asymp (\log n)^{d-1}$; see (3).

Proof of (16). Let N_v be the number of points of $\mathbf{W}_n$ falling in G_v. We choose m so large that $\min_{j,k} \mathbb{P}(N_j = 0, N_k = 0) \ge 1/2$. Note that $\mathbb{E}[Z_v|Z_j, Z_k] = Z_v \le 1$ for $v = j$ or $v = k$.

We claim that

$$\mathbb{E}[Z_v|Z_j, Z_k] \le 2\mathbb{E}[N_v], \tag{17}$$

for all $v \neq j, k$. As one can see from Figure 2, for any given G_k and G_j, Z_v is dependent on Z_j or Z_k only when the overlapping region of G_v and the shaded area is nonempty. From this it follows that

$$\begin{aligned} \sum_{Z_v \text{ dependent on } Z_j \text{ or } Z_k} \mathbb{E}[Z_v|Z_j, Z_k] &\le 2 + 2 \sum_{Z_v \text{ dependent on } Z_j \text{ or } Z_k} \mathbb{E}[N_v] \\ &= O\left((\beta-\alpha)^{d-1}\int_\alpha^\beta n e^{-t}\,\mathrm{d}t\right) \\ &= O\left((\log\log n)^d\right), \end{aligned}$$

proving (16). Thus the proof of (16) is reduced to that of (17), which is split into three cases conditioning on the possible values of $(Z_j, Z_k) = (0,0), (0,1)$ or $(1,1)$.

The first case when both Z_j and Z_k are zeros is estimated as follows.

$$\begin{aligned} \mathbb{E}[Z_v|Z_j = 0, Z_k = 0] &= \frac{\mathbb{P}(Z_v = 1, Z_j = 0, Z_k = 0)}{\mathbb{P}(Z_j = 0, Z_k = 0)} \\ &\le \frac{\mathbb{E}[N_v]}{\mathbb{P}(N_j = 0, N_k = 0)} \\ &\le 2\mathbb{E}[N_v]. \end{aligned}$$

For the remaining cases, we observe that

$$\begin{aligned}\mathbb{P}\left(Z_k=1|N_i=0\right)&\geq\mathbb{P}\left(Z_k=1\right)\ \text{for } i\neq k,\\ \mathbb{P}\left(Z_j=1,Z_k=1|N_i=0\right)&\geq\mathbb{P}\left(Z_j=1,Z_k=1\right)\ \text{for } i\neq j,k;\end{aligned}\tag{18}$$

and thus

$$\mathbb{P}\left(Z_k=1|N_i\geq 1\right)\leq\mathbb{P}\left(Z_k=1\right)\ \text{for } i\neq k,\tag{19}$$

$$\mathbb{P}\left(Z_j=1,Z_k=1|N_i\geq 1\right)\leq\mathbb{P}\left(Z_j=1,Z_k=1\right)\ \text{for } i\neq j,k.\tag{20}$$

Consequently, when both Z_j and Z_k are 1, we have, by (20),

$$\begin{aligned}\mathbb{E}[Z_v|Z_j=1,Z_k=1]&=\frac{\mathbb{P}\left(Z_v=1,Z_j=1,Z_k=1\right)}{\mathbb{P}\left(Z_j=1,Z_k=1\right)}\\ &\leq\frac{\mathbb{P}\left(N_v\geq 1,Z_j=1,Z_k=1\right)}{\mathbb{P}\left(Z_j=1,Z_k=1\right)}\\ &=\frac{\mathbb{P}\left(Z_j=1,Z_k=1|N_v\geq 1\right)\mathbb{P}\left(N_v\geq 1\right)}{\mathbb{P}\left(Z_j=1,Z_k=1\right)}\\ &\leq\mathbb{P}\left(N_v\geq 1\right)\\ &\leq\mathbb{E}[N_v].\end{aligned}$$

Similarly, from (18) and (19) it follows that

$$\begin{aligned}\mathbb{E}[Z_v|Z_j=0,Z_k=1]&=\frac{\mathbb{P}\left(Z_v=1,Z_j=0,Z_k=1\right)}{\mathbb{P}\left(Z_j=0,Z_k=1\right)}\\ &\leq\frac{\mathbb{P}\left(N_v\geq 1,Z_k=1\right)}{\mathbb{P}\left(N_j=0,Z_k=1\right)}\\ &=\frac{\mathbb{P}\left(Z_k=1|N_v\geq 1\right)\mathbb{P}\left(N_v\geq 1\right)}{\mathbb{P}\left(Z_k=1|N_j=0\right)\mathbb{P}\left(N_j=0\right)}\\ &\leq\frac{\mathbb{P}\left(N_v\geq 1\right)}{\mathbb{P}\left(N_j=0\right)}\\ &\leq 2\mathbb{E}[N_v].\end{aligned}$$

for $j\neq k$. This completes the proof of (17).

From d_1-distance to Kolmogorov distance

With Propositions 1 and 2 at hand, we need only to apply the following lemma to complete the proof of Theorem 2.

Lemma 2. *Assume that the sequence of random variables Y_n converges to the standard normal distribution with a rate*

$$d_1(Y_n,N(0,1))=O(r_n),$$

where $r_n\to 0$ and $N(0,1)$ denotes a standard normal variable. Then

$$\{Y_n\}\in\mathrm{CLT}(\sqrt{r_n}).$$

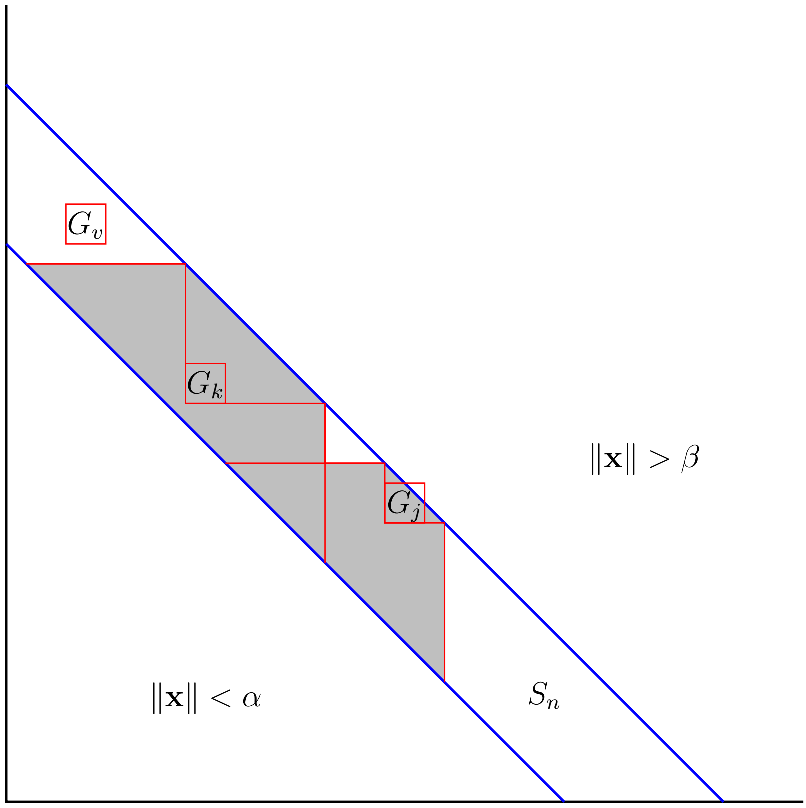

Figure 2: *Possible configurations of* G_v, G_j, *and* G_k.

Proof. Fix n. Let

$$h_z(y) = \begin{cases} 0 & \text{if } y < z, \\ (y-z)/2 & \text{if } z \le y \le z + 2\sqrt{r_n}, \\ \sqrt{r_n} & \text{if } y > z + 2\sqrt{r_n}. \end{cases}$$

Then $\sup_y |h_z(y)| + \sup_y |h_z'(y)| \le \sqrt{r_n} + 1/2$. Without loss of generality, assume $r_n \le 1/4$. Then

$$\begin{aligned} \sqrt{r_n}\, |\mathbb{P}\,(Y_n < z) - \Phi(z)| &= \sqrt{r_n}\, |\mathbb{P}\,(Y_n \ge z) - \mathbb{P}\,(N(0,1) \ge z)| \\ &\le \sup_z |\mathbb{E}[h_z(Y_n)] - \mathbb{E}[h_z(N(0,1))]| \\ &\quad + \sqrt{r_n}\, \mathbb{P}(z \le N(0,1) \le z + 2\sqrt{r_n}) \\ &\le d_1(Y_n, N(0,1)) + r_n \\ &= O(r_n). \end{aligned}$$

Thus $\sup_z |\mathbb{P}\,(Y_n < z) - \Phi(z)| = O(\sqrt{r_n})$. □

Remark. By the same method of proof, we can derive a Berry-Esseen bound for the number of maxima $\hat{K}_n$ in d-simplex of the form

$$\{\hat{K}_n\} \in \mathrm{CLT}\left(n^{-(d-1)/(4d)}(\log n)^d + n^{-1/d}(\log n)^d\right).$$

Acknowledgements

This work was started while the first and the third authors were visiting School of Computer Science, McGill University; they thank the School for its hospitality. The third author also thanks Institut für Stochastik und Mathematische Informatik, J. W. Goethe-Universität (Frankfurt) where part of his work was completed.

Appendix. Various expressions for $\mathbb{E}[K_{n,d}]$.

We collect some expressions for $\mu_{n,d} := \mathbb{E}(K_{n,d})$ in the case of hypercubes. These expressions obviously show the diversity of the nature of the enumeration problem; see also Flajolet et al. (1995), Labelle and Laforest (1995).

Summation formulae.

$$\mu_{n,d} = \sum_{1\le k\le n} \binom{n}{k} (-1)^{k-1} k^{1-d},$$

$$\mu_{n,d} = \sum_{1\le i_1\le\cdots\le i_{d-1}} \frac{1}{i_1\cdots i_{d-1}},$$

$$\mu_{n,d} = \sum_{\substack{i_1+2i_2+\cdots(d-1)i_{d-1}=d-1\\ i_1,\ldots,i_{d-1}\ge 0}} \frac{H_n^{i_1}(H_n^{(2)})^{i_2}\cdots(H_n^{(d-1)})^{i_{d-1}}}{i_1!\cdots i_{d-1}!1^{i_1}\cdots(d-1)^{i_{d-1}}},$$

where $H_n^{(a)} := \sum_{1\le j\le n} 1/j^a$.

Recurrence relations. For $n \ge 1$ and $d \ge 2$

$$\mu_{n,d} = \mu_{n-1,d} + \frac{\mu_{n,d-1}}{n},$$

$$\mu_{n,d} = \sum_{1\le j\le n} \frac{\mu_{j,d-1}}{j},$$

$$\mu_{n,d} = \frac{1}{d-1} \sum_{1\le j\le d-1} H_n^{(d-j)} \mu_{n,j},$$

with $\mu_{n,1} \equiv 1$ for $n \ge 1$.

Integral representations.

$$\mu_{n,d} = n\int_{(0,1)^d} (1 - x_1x_2\cdots x_d)^{n-1}\ \mathrm{d}\mathbf{x},$$
$$\mu_{n,d} = \frac{n}{(d-1)!}\int_0^1 (1-x)^{n-1}(-\log x)^{d-1}\,\mathrm{d}x,$$
$$\mu_{n,d} = \frac{1}{2\pi i}\oint_{|z|=r<1} z^{-d}\prod_{1\le j\le n}\frac{1}{1-z/j}\,\mathrm{d}z,$$
$$\mu_{n,d} = \frac{(-1)^n}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\frac{n!}{s^d(s-1)\cdots(s-n)}\,\mathrm{d}s.$$

Probability expressions.

$$\mu_{n,d} = n\mathbb{E}\left[(1-U_1U_2\cdots U_d)^{n-1}\right],$$
$$\mu_{n,d} = n\mathbb{P}(Y_2+\cdots+Y_n<d),$$

where $U_1, U_2, \ldots, U_d$ are iid uniform $[0,1]$ random variables and the Y_j's are geometric random variables

$$\mathbb{E}[z^{Y_j}] = \frac{1-1/j}{1-z/j} \qquad (2\le j\le n);$$

see Bai et al. (1998).

Also $\mu_{n,d}/n$ is the probability that the first subtree in a random quadtree of n nodes is empty; see Flajolet et al. (1995).

Asymptotic approximations. Let $\rho := (d-1)/\log n$.

(i) If $1\le d\le \log n - M\sqrt{\log n}$, $M>1$ being sufficiently large,

$$\mu_{n,d} = \Gamma(1-\rho)\frac{(\log n)^{d-1}}{(d-1)!}\left(1+O\left(\frac{d}{(\log n-d)^2}\right)\right),$$

uniformly in d;

(ii) if $d=\log n + x\sqrt{\log n}$, where $x=o((\log n)^{1/6})$, then

$$\frac{\mu_{n,d}}{n} = \Phi(x)\left(1+O\left(\frac{1+|x|^3}{\sqrt{\log n}}\right)\right),$$

uniformly in x;

(iii) If $d\ge \log n + M\sqrt{\log n}$, then

$$1-\frac{\mu_{n,d}}{n} = O\left(n^{-\rho\log\rho+\rho-1}(\log n)^{-1/2}\right),$$

uniformly in d; see Hwang (2002) and the references therein.

References

[1] Z.-D. Bai, C.-C. Chao, H.-K. Hwang and W.-Q. Liang (1998). On the variance of the number of maxima in random vectors and its applications. *Annals of Applied Probability*, **8** 886–895.

[2] Z.-D. Bai, H.-K. Hwang, W.-Q. Liang and T.-H. Tsai (2001). Limit theorems for the number of maxima in random samples from planar regions. *Electronic Journal of Probability*, **6** Paper 3, 41 pp.

[3] Z.-D. Bai, H.-K. Hwang and T.-H. Tsai (2003). Berry-Esseen bounds for the number of maxima in planar regions. *Electronic Journal of Probability*, **8** (2003), Paper 9, 26 pp.

[4] A. D. Barbour, L. Holst and S. Janson (1992). *Poisson Approximation*. Oxford Science Publications. Oxford University Press, New York.

[5] A. D. Barbour and A. Xia (2001). The number of two dimensional maxima. *Advances in Applied Probability*, **33** 727–750.

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

[7] Y. Baryshnikov (2000). Supporting-points processes and some of their applications. *Probability Theory and Related Fields*, **117** 163–182.

[8] J. L. Bentley, H. T. Kung, M. Schkolnick and C. D. Thompson (1978). On the average number of maxima in a set of vectors and applications. *Journal of the ACM*, **25** 536–543.

[9] B. A. Berezovskiĭ and S. I. Travkin (1975). Supervision of queues of requests in computer systems. *Automation and Remote Control*, **36** 1719–1725.

[10] C. Blair (1986). Random inequality constraint systems with few variables. *Mathematical Programming*, **35** 133–139.

[11] C. Buchta (1989). On the average number of maxima in a set of vectors. *Information Processing Letters*, **33** 63–65.

[12] H. C. Calpine and A. Golding (1976). Some properties of Pareto-optimal choices in decision problems. *Omega (The International Journal of Management Science)*, **4** 141–147.

[13] A. Carlsund (2003). Notes on the variance of the number of maxima in three dimensions. *Random Structures and Algorithms*, **22** 440–447.

[14] W.-M. Chen, H.-K. Hwang and T.-H. Tsai (2003). Efficient maxima-finding algorithms for random planar samples. *Discrete Mathematics and Theoretical Computer Science (Electronic)*, **6** 107–122.

[15] L. Devroye (1980). A note on finding convex hulls via maximal vectors. *Information Processing Letters*, **11** 53–56.

[16] L. Devroye (1986). *Lecture Notes on Bucket Algorithms*. Birkhäuser Boston, Boston, MA.

[17] L. Devroye (1999). A note on the expected time for finding maxima by list algorithms. *Algorithmica*, **23** 97–108.

[18] M. E. Dyer and J. Walker (1998). Dominance in multi-dimensional multiple-choice knapsack problems. *Asia-Pacific Journal of Operational Research*, **15** 159–168.

[19] P. Flajolet, G. Labelle, L. Laforest and B. Salvy (1995). Hypergeometrics and the cost structure of quadtrees. *Random Structures and Algorithms*, **7** 117–144.

[20] P. Flajolet and B. Salvy (1998). Euler sums and contour integral representations. *Eperimental Mathematics*, **7** 15–35.

[21] P. Flajolet and R. Sedgewick (1995). Mellin transforms and asymptotics: finite differences and Rice's integrals. *Theoretical Computer Science*, **144** 101–124.

[22] M. J. Golin (1994). A provably fast linear-expected-time maxima-finding algorithm. *Algorithmica*, **11** 501–524.

[23] H.-K. Hwang (2002). Phase changes in random recursive structures and algorithms. In *Proceedings of the Workshop on Probability with Applications to Finance and Insurance (Hong Kong, July, 2002)*. Edited by T. L. Lai, H. Yang and S. P. Yung, pp. 82–97, World Scientific, June 2004. (Preprint available at algo.stat.sinica.edu.tw.)

[24] V. M. Ivanin (1975). Estimate of the mathematical expectation of the number of elements of the Pareto set. *Kibernetika* (Kiev), No. 3, 145–146 (in Russian); English translation: *Cybernetics,* **11** 506–507 (1976).

[25] V. M. Ivanin (1976). Calculation of the dispersion of the number of elements of the Pareto set for the choice of independent vectors with independent components. In *Theory of Optimal Decisions* 90–100. Akad. Nauk. Ukrain. SSR Inst. Kibernet., Kiev. (In Russian.)

[26] S. Janson, T. Łuczak and A. Rucinski (2000). *Random Graphs*. John Wiley & Sons, Inc.

[27] A. I. Kuksa and N. Z. Šor (1972). The method of estimating the number of conditionally optimal trajectories of discrete separable dynamic programming. (In Russian). *Kibernetika (Kiev)*, no. 6, 37–44.

[28] G. Labelle and L. Laforest (1995). Combinatorial variations on multidimensional quadtrees. *Journal of Combinatorial Theory, Series A*, **69** 1–16.

[29] B. O'Neill (1980). The number of outcomes in the Pareto-optimal set of discrete bargaining games. *Mathematics of Operations Research,* **6** 571–578.

[30] Yu. V. Prohorov (1953). Asymptotic behavior of the binomial distribution. in *Selected Translations in Mathematical Statistics and Probability*, Vol. 1, pp. 87–95, ISM and AMS, Providence, R.I. (1961); translation from Russian: *Uspehi Matematičeskih Nauk*, **8** (1953), no. 3 (35), 135–142.

[31] C. Vout (1973). Randomly collecting sets of objects or . . . It's another spear-thrower! *Eureka*, No. 36, October, 18–20.