Electronic Journal of Probability

Vol. 6 (2001) Paper no. 3, pages 1–41.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol6/paper3.abs.html

LIMIT THEOREMS FOR THE NUMBER OF MAXIMA IN RANDOM SAMPLES FROM PLANAR REGIONS

Zhi-Dong Bai
Department of Statistics & Applied Probability, National University of Singapore, Singapore
stabaizd@leonis.nus.edu.sg

Hsien-Kuei Hwang[1], **Wen-Qi Liang and Tsung-Hsi Tsai**
Institute of Statistical Science, Academia Sinica, 115, Taiwan
hkhwang@stat.sinica.edu.tw, wqliang@stat.sinica.edu.tw, chonghi@stat.sinica.edu.tw

Abstract We prove that the number of maximal points in a random sample taken uniformly and independently from a convex polygon is asymptotically normal in the sense of convergence in distribution. Many new results for other planar regions are also derived. In particular, precise Poisson approximation results are given for the number of maxima in regions bounded above by a nondecreasing curve.

Keywords Maximal points, multicriterial optimization, central limit theorems, Poisson approximations, convex polygons.

AMS subject classification Primary. 60D05; Secondary. 60C05

Submitted to EJP on September 29, 2000. Final version accepted on January 22, 2001.

---

[1]Part of the work of this author was done while visiting the Department of Statistics and Applied Probability, National University of Singapore. He thanks the department for their hospitality.

1 Introduction

A point $p_1 = (x_1, y_1)$ is said to *dominate* another point $p_2 = (x_2, y_2)$ if $x_1 \geq x_2$ and $y_1 \geq y_2$. For notational convenience, we write this as $p_1 \succ p_2$. The *maxima* (or *maximal points*) of a sample of points are those points dominated by no other points in the sample. We investigate in this paper distributional properties of the number of maximal points in a random sample taken uniformly and independently from a given planar region.

Such a dominance relation (not restricted to planar points), known more frequently as *Pareto optimality*, is an extremely useful notion in diverse fields ranging from economics to mechanics, from social sciences to algorithmics; see for example Karlin (1959), Bühlmann (1970), Leitman and Marzollo (1975), Preparata and Shamos (1985), Devroye (1986), Steuer (1986), Stadler (1988), Harsanyi (1988), Statnikov and Matusov (1995). It is one of the most natural order relations in multivariate observations because of the lack of intrinsic total order relations. Other terms like "efficiency" in econometrics, "noninferiority" in control, "admissibility" in statistics are all similar notions. One also finds a similar dominance relation used in a card game called "Russian poker." From a probabilistic point of view, we do not distinguish in this paper the difference between $>$ and $\geq$, a key issue, however in other fields. The following quote from the review by J. Stoer and J. Zowe in AMS Mathematical Review (MR: 50 #3928) of Zeleny's book [39] typically describes the general situation encountered in theory and practice:

> *One of the more pertinent criticisms of traditional decision-making theory and practice is directed against the approximation of multiple goal behavior by a single technical criterion. In a realistic model of a technical or economical optimization problem it will be impossible to tie all the given criteria into a single function, which could serve as objective function for an associated mathematical programming problem. It will be more appropriate to handle such problems as problems with a vector-valued objective function. Instead of finding an optimal solution of a single objective function, the problem is now one of locating the set of all nondominated points (also called efficient points or Pareto-optimal points).*

We mention some concrete applications as follows.

1. Finding the maxima of a sample is a prototype problem with many algorithmic and practical applications. Many geometric and graph-theoretic problems can be formulated as maxima-finding problems, including the problem of *computing the minimum independent dominating set* in a permutation graph, the related problem of *finding the shortest maximal increasing subsequence*, the problem of *enumerating restricted empty rectangles*, and the related problem of *computing the largest empty rectangle*. Also the d-dimensional maxima-finding problem is equivalent to the *enclosure problem for planar d-gon* (where the corresponding sides are parallel), the latter problem in turn has several applications in *CAD systems for VLSI circuits*; see [16, 26, 33] for details and references.

2. The notion of dominance was used in Becker et al. (1987) to analyze data from the *Places Rated Almanac*, a collection of nine composite variables constructed for 329 metropolitan areas of the US.

3. The number of maxima is closely related to the performance of maxima-finding algorithms (Devroye (1986) and Preparata and Shamos (1985)); it was also used in probabilistic analysis of simplex algorithms in Blair (1986).

Despite its usefulness in diverse fields, the number of maxima did not receive much attention as far as probabilistic properties are concerned in the literature. This quantity plays, in particular, an important role in the investigation of the stochastic behaviors of maxima-finding algorithms; see Devroye (1986, 1999). The first such study was due to Barndorff-Nielsen and Sobel (1966). Then results appeared sporadically in the control theory and the computer science literatures; see Sholomov (1984) for a survey. We will mention some known results relevant to ours and present new ones.

Let $\mathcal{C}$ be a given measurable region and $M_n(\mathcal{C})$ denote the number of maxima in a random sample of n points taken uniformly and independently from $\mathcal{C}$. Almost all results in the literature are concerned with the mean value of $M_n(\mathcal{C})$. Distributional results are rare, the most studied case being $\mathcal{C} = [0,1]^2$ for which the problem reduces to the number of records in iid (*independent and identically distributed*, here and throughout this paper) sequences of continuous random variables; see Baryshnikov (1987, 2000). From this connection, our study may also be regarded as another line of extensions of the theory of records; see Rényi (1962), Barndorff-Nielsen and Sobel (1966), Arnold et al. (1998). The main contribution of this paper is to provide means of establishing the central limit theorem for $M_n(\mathcal{C})$ when $\mathcal{C}$ is a convex polygon and when $\mathcal{C}$ is some region bounded above by a nondecreasing curve. Indeed, we derive precise Poisson approximation results in the second case (by computing the asymptotics of the total variation and Fortet-Mourier distances). A by-product of our central limit theorems gives the asymptotics of the variance. Many other auxiliary and structural results are also derived.

Results

Known results for $M_n(\mathcal{C})$ when $\mathcal{C} = [0,1]^d$ can be found in Bai et al. (1998), Devroye (1999) and the references therein.

When $\mathcal{C}$ is a convex polygon $\mathcal{P}$, Golin (1993) gave the following "gap theorem":

$$E(M_n(\mathcal{P})) \asymp \begin{cases} n^{1/2}, \\ \log n, \\ 1, \end{cases}$$

namely, $E(M_n(\mathcal{P}))$ is either of order $n^{1/2}$, or of order $\log n$, or bounded, and no other scales are possible. On the other hand, Devroye (1993) showed that if

$$\mathcal{C} = \{(x,y) \,:\, 0 \le x \le 1,\ 0 \le y \le f(x)\}, \tag{1.1}$$

where f is nonincreasing and is either *convex*, or *concave*, or *Lipschitz (of order 1)*, then

$$E(M_n(\mathcal{C})) \sim \pi_0 n^{1/2}, \tag{1.2}$$

where

$$\pi_0 = \left(\frac{\pi}{2}\right)^{1/2} \frac{\int_0^1 |f'(x)|^{1/2}\,\mathrm{d}x}{\left(\int_0^1 f(x)\,\mathrm{d}x\right)^{1/2}}.$$

Note that by Cauchy-Schwarz inequality, it is easily deduced that the right-hand side of π_0 reaches its maximal value when $\mathcal{C}$ is a right triangle with decreasing hypotenuse of the shape ; compare Dwyer (1990).

Our first result shows in a precise way that in the case of a convex polygon $\mathcal{P}$

$$E(M_n(\mathcal{P})) = \pi_1 n^{1/2} + \pi_2 \log n + \pi_3 + O(n^{-1/2}),$$

where π_1 is essentially Devroye's ("discretized") constant, π_2 can assume only values $0, 1/2$ and 1, and π_3 is a complicated constant; see (1.3), (1.6), (3.9) and (3.12) for explicit expressions of these constants. More precisely, let

$$y^* = \sup\{y \,:\, (x,y) \in \mathcal{P}\}, \quad x^* = \sup\{x \,:\, (x,y) \in \mathcal{P}\},$$

and then define

$$I_u = \{(x,y^*) \,:\, (x,y^*) \in \mathcal{P}\}, \quad I_r = \{(x^*,y) : (x^*,y) \in \mathcal{P}\}.$$

In words, if we move a horizontal line from ∞ downwards, then I_u denotes the intersection of this line and $\mathcal{P}$ when they first meet; likewise, I_r denotes the intersection of a vertical line moving from ∞ leftwards and $\mathcal{P}$ when they first meet.

The *upper-right part* of $\mathcal{P}$ is defined as

$$\mathcal{P} \cap \{(x,y) \,:\, x \geq \min I_u,\ y \geq \min I_r\},$$

where $\min I_u := \min\{x \,:\, (x,y^*) \in \mathcal{P}\}$ and $\min I_r := \min\{y,: \, (x^*,y) \in \mathcal{P}\}$. Note that the upper-right part can be either a point, or one or two lines, or a region with nonzero measure. Also note that this definition applies to any other planar regions.

Define $s_1, s_2, \dots, s_\nu$ to be the line segments on the upper-right part of $\mathcal{P}$ which bridge I_u and I_r when $I_u \cap I_r = \emptyset$, where $\emptyset$ denotes the empty set. Also let θ_j be the angle formed by s_j and a horizontal line for $j = 1, \dots, \nu$; see Figure 1 for an illustration.

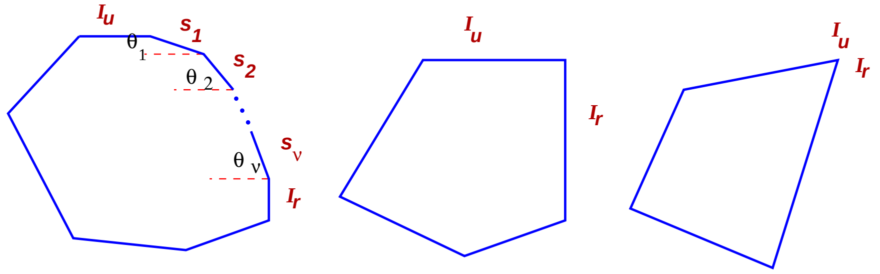

Figure 1: *Possible configurations of convex polygons.*

Let $\mathcal{N}(0,1)$ denote a normal random variable with zero mean and unit variance. Let $|\mathcal{P}|$ and $|s_j|$ denote the area and length, respectively, of $\mathcal{P}$ and s_j.

Theorem 1 (CLT for convex polygons). *The mean of $M_n(\mathcal{P})$ satisfies*

$$E(M_n(\mathcal{P})) = \pi_1 n^{1/2} + \pi_2 \log n + \pi_3 + O\left((\nu+1)n^{-1/2}\right), \tag{1.3}$$

where

$$\pi_1 = \left(\frac{\pi}{2|\mathcal{P}|}\right)^{1/2} \sum_{1 \leq j \leq \nu} |s_j| \left(\cos\theta_j \sin\theta_j\right)^{1/2}. \tag{1.4}$$

$$\pi_2 = \frac{1}{2}\left(1_{\{|I_u|>0\}} + 1_{\{|I_r|>0\}}\right) \in \{0, 1/2, 1\}, \tag{1.5}$$

and π_3 is given in (3.9) and (3.12). If in addition $I_u \cap I_r = \emptyset$ ($\pi_1 > 0$), then

$$\mathrm{Var}(M_n(\mathcal{P})) \sim \sigma_1^2 n^{1/2}, \tag{1.6}$$

and

$$\frac{M_n(\mathcal{P}) - \pi_1 n^{1/2}}{\sigma_1 n^{1/4}} \xrightarrow{D} \mathcal{N}(0,1), \tag{1.7}$$

where $\sigma_1 = (\pi_1(2\log 2 - 1))^{1/2}$ and $\xrightarrow{D}$ denotes convergence in distribution.
If $\pi_1 = 0$ and $\pi_2 > 0$ then

$$\mathrm{Var}(M_n(\mathcal{P})) \sim \pi_2 \log n, \tag{1.8}$$

and

$$\frac{M_n(\mathcal{P}) - \pi_2 \log n}{(\pi_2 \log n)^{1/2}} \xrightarrow{D} \mathcal{N}(0,1). \tag{1.9}$$

The proof of (1.7) consists in first splitting the upper-right part of $\mathcal{P}$ that lies in $\mathcal{P}$ into triangles of the form ◺(except possibly the first and the last). Thus the asymptotic normality of $M_n(\mathcal{P})$ is reduced to that of $M_n(\mathcal{T})$, where $\mathcal{T}$ is the right triangle with corners $(0,0)$, $(0,1)$ and $(1,0)$. It turns out that in this specific case, $M_n(\mathcal{T})$ enjoys many interesting properties; details are examined in the next section. Originally, our proof for the asymptotic normality of $M_n(\mathcal{T})$ proceeded by computing the third and fourth central moments, which was very laborious; see Bai et al. (2000) for details. The proof given here uses the method of moments, and, because of a better manipulation of the recurrence, the result is stronger (convergence of all moments) and the proof is much shorter. The proof of Theorem 1, except (1.8) and (1.9), is then given in Section 3. The case when $I_u \cap I_r \neq \emptyset$ is essentially a special case of Theorem 2; a sketch of proof is given in Section 6. Note that in the special case when $\mathcal{P}$ is the unit square, the result (1.9) can be easily proved by checking Lyapunov's condition.

Our proof actually covers the case when the underlying region $\mathcal{P}$ is not convex, provided that the upper-right part of $\mathcal{P}$ can be split into a finite union of triangles and rectangles; see Section 3.2.

A comparison of (1.3) with the result for the expected number of points on the convex hull of n iid points chosen uniformly from $\mathcal{P}$ by Rényi and Sulanke (1963) shows that there are considerably more maxima than hull points in a random sample when $I_u \cap I_r = \emptyset$ (the former can be used to approximate the latter; see Devroye, 1980). For more information on related results, see Buchta and Reitzner (1997) and the references given there.

The surprising factor $1/2$ in (1.3) and (1.5) suggests several new questions like "Why it is $1/2$?" "Is this factor universal?", and "Since this factor depends (from (1.5)) only on the boundary of the upper-right part of $\mathcal{P}$, what happens for general regions?"

It is for answering these questions that we study the following two problems. First, if we compare (1.2) and (1.3), it is then natural to consider the convergence rate of Devroye's result (1.2). Devroye (1993) observed that the continuous part of f' contributes the $n^{1/2}$ term (as in the expression of π_0), while the discontinuous points contribute $O(\log n)$ maximal points on average. We show that if f is piecewise twice continuously differentiable then the next dominant term in (1.2) is asymptotic to $c \log n$, where c can be explicitly computed in terms of the local behaviors of f near "critical points," that is, points at which $|f'| = 0$, or $|f'| = \infty$ or f' does not exist. This result thus connects the critical points and the $\log n$ term in the asymptotic expression of $E(M_n)$ in a quantitatively precise way. The idea is to break f at all critical points and then to sum over the expected number of maxima (scaled by their area proportions) in each smaller regions with more smooth boundaries (see Figure 8), the hard part being the determination of the coefficient of the log term. This problem is discussed in Section 4.

Second, we consider $E(M_n(\mathcal{C}))$ when $\mathcal{C}$ is bounded by two nondecreasing curves. We show that if the rates that the two functions tend to their values at unity are the same then $E(M_n(\mathcal{C}))$ is asymptotic to a constant; otherwise, it is asymptotic to $c \log n$ for some constant c. In particular, if one curve is linear and the other is either a vertical or horizontal line, then $c = 1/2$. This consideration thus partly interprets the constant $1/2$ in a deeper way.

The last problem we consider is the distribution of $M_n(\mathcal{C})$ when $\mathcal{C}$ is of the shape (1.1), where f is *nondecreasing*. It is well known that when $f(x) \equiv 1$, M_n is well approximated by a Poisson distribution with mean $\log n$ [29]. We derive, under suitable conditions on f, precise asymptotic approximations for the total variation and Fortet-Mourier distances between the distribution of $M_n(\mathcal{C})$ and a suitable Poisson distribution. Actually, our results suggest that Poisson law (with bounded or unbounded mean) is the *universal limit law* of M_n for nondecreasing f; see the discussion in Section 6 for more support of this suggestion.

Recall that the total variation and the Fortet-Mourier distances of two random variables X, Y can be defined, respectively, by

$$\begin{aligned} d_{\mathrm{TV}}(\mathcal{L}(X),\mathcal{L}(Y)) &= \frac{1}{2}\sum_{j\geq 0}\left|P(X=j)-P(Y=j)\right|, \\ d_{\mathrm{FM}}(\mathcal{L}(X),\mathcal{L}(Y)) &= \sum_{j\geq 0}\left|P(X\leq j)-P(Y\leq j)\right|. \end{aligned}$$

Theorem 2 (Poisson approximation). *Assume that f is nondecreasing, $\int_0^1 f(t)\,\mathrm{d}t = 1$ and*

$$\frac{xf(1-x)}{1-F(1-x)} = 1-(1-F(1-x))^{\alpha}\left(a+L(1-F(1-x))\right),$$

where $a,\alpha>0$, $F(x):=\int_0^x f(t)\,\mathrm{d}t$, and

$$L(x) = O\left(|\log x|^{-1}\right), \tag{1.10}$$

as $x\to 0^+$. Then the mean and the variance of M_n satisfy

$$\begin{aligned} E(M_n) &= \frac{\alpha}{\alpha+1}\log n+\frac{\alpha\gamma-\log a}{\alpha+1}+O\left((\log n)^{-1}\right), \\ \mathrm{Var}(M_n) &= \frac{\alpha}{\alpha+1}\log n+\frac{(6\gamma-\pi^2)\alpha^2+(6\gamma-2\pi^2-6\log a)\alpha-6\log a}{6(\alpha+1)^2} \\ &\quad +O\left((\log n)^{-1}\right), \end{aligned} \tag{1.11}$$

where γ is the Euler constant. Furthermore, the distribution of M_n is asymptotically Poisson:

$$\begin{aligned} d_{\mathrm{TV}}(\mathcal{L}(M_n),\mathcal{L}(Y)) &= \frac{\kappa}{\sqrt{2\pi e}\,\lambda}\left(1+O\left((\log n)^{-1/2}\right)\right), \\ d_{\mathrm{FM}}(\mathcal{L}(M_n),\mathcal{L}(Y)) &= \frac{\kappa}{\sqrt{2\pi\lambda}}\left(1+O\left((\log n)^{-1/2}\right)\right), \end{aligned}$$

where Y is a Poisson variate with mean $\lambda=\frac{\alpha}{\alpha+1}\log n+\frac{\alpha\gamma-\log a}{\alpha+1}-1$:

$$P(Y=m)=e^{-\lambda}\frac{\lambda^{m-1}}{(m-1)!}, \qquad (m\geq 1),$$

and

$$\kappa = \left|1-\frac{\pi^2\alpha(\alpha+2)}{6(\alpha+1)^2}\right|.$$

In particular, M_n is asymptotically normally distributed with mean and variance asymptotic to $\frac{\alpha}{\alpha+1}\log n$. We can indeed derive a local limit theorem for M_n.

Note that

$$\kappa = \pm\left(1-\frac{\pi^2\alpha(\alpha+2)}{6(\alpha+1)^2}\right),$$

depending on $\alpha<-1+\pi/\sqrt{\pi^2-6}$ (plus sign) or $\alpha>-1+\pi/\sqrt{\pi^2-6}$ (minus sign). The case when $\alpha=-1+\pi/\sqrt{\pi^2-6}$ is of special interest since $\kappa=0$ and our result reduces to an upper estimate. If we replace (1.10) in this case by $L(x)=O\left(|\log x|^{-3/2}\right)$, then we can show, using the proof techniques for Theorem 2 and the approach in [29], that

$$\begin{aligned} d_{\mathrm{TV}}(\mathcal{L}(M_n),\mathcal{L}(Y)) &= \frac{\kappa_1}{3\sqrt{2\pi}}\left(4e^{-3/2}+1\right)\lambda^{-3/2}\left(1+O\left((\log n)^{-1/2}\right)\right), \\ d_{\mathrm{FM}}(\mathcal{L}(M_n),\mathcal{L}(Y)) &= \frac{2\sqrt{2}\kappa_1}{3\sqrt{\pi e}}\lambda^{-1}\left(1+O\left((\log n)^{-1/2}\right)\right) \end{aligned}$$

where $\kappa_1 = \zeta(3)(\pi^2-6)^{3/2}/\pi^3 - \zeta(3)+1$, $\zeta(3) = \sum_{j\ge1} j^{-3}$.

The proof of Theorem 2 relies on an explicit expression for the moment generating function (in terms of moment generating function of $M_n(\mathcal{C})$ when $\mathcal{C}$ is a rectangle) and a careful analysis of the associated sums and integrals. Details as well as other Poisson approximation results are given in Section 6.

Notation. Throughout this paper, n is the major asymptotic parameter which is taken to be sufficiently large. All limits, whenever unspecified, is taken to be $n\to\infty$. The generic symbols ε, c, and K always represent suitably small, absolute, and large, respectively, positive constants independent of n whose values may vary from one occurrence to another. The symbol $[z^n]F(z)$ denotes the coefficient of z^n in the Taylor expansion of $F(z)$. We write simply M_n when there is no ambiguity of the underlying planar region.

2 Maxima in right triangles

Let $\mathcal{T}$ denote the right triangle with corners $(0,0)$, $(0,1)$ and $(1,0)$. We prove in this section the asymptotic normality of $M_n = M_n(\mathcal{T})$ by the method of moments. The key idea of the proof is to compute the (centralized) moments recursively and then to reduce all major asymptotic estimates to an asymptotic transfer lemma (see Lemma 4).

Theorem 3 (CLT for triangle). *The mean and the variance of M_n satisfy*

$$\begin{aligned} E(M_n) &= \frac{\sqrt{\pi}n!}{\Gamma(n+1/2)} - 1 && (2.1)\\ &= \sqrt{\pi n} - 1 + \frac{\sqrt{\pi}}{8\sqrt{n}} + \frac{\sqrt{\pi}}{128 n^{3/2}} + O\left(n^{-5/2}\right),\\ \mathrm{Var}(M_n) &= \pi n - \frac{\pi\, n!^2}{\Gamma(n+1/2)^2} + \frac{\sqrt{\pi}n!}{\Gamma(n+1/2)}(2\log 2 - 1) + O\left(n^{-K}\right) && (2.2)\\ &= \sigma_2^2\sqrt{n} - \frac{\pi}{4} + \frac{\sigma_2^2}{8\sqrt{n}} - \frac{\pi}{32n} + \frac{\sigma_2^2}{128 n^{3/2}} + \frac{\pi}{128n^2} + O\left(n^{-5/2}\right), \end{aligned}$$

for any $K>0$, where $\sigma_2 = (2\log 2 - 1)^{1/2}\pi^{1/4}$. The distribution of $(M_n - \sqrt{\pi n})/(\sigma_2 n^{1/4})$ is asymptotically normal:

$$\frac{M_n - (\pi n)^{1/2}}{\sigma_2 n^{1/4}} \xrightarrow{D} \mathcal{N}(0,1), \qquad (2.3)$$

where the limit holds with convergence of all moments.

2.1 Moment generating function

Proposition 1. *The moment generating function $f_n(w) := E(e^{M_n w})$ of M_n satisfies the recurrence*

$$f_n(w) = e^w \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) f_j(w) f_k(w), \qquad (2.4)$$

for $n\ge 1$ with the initial condition $f_0(w)=1$, where the sum is extended over all nonnegative integer triples (j,k,ℓ) such that $j+k+\ell = n-1$ and

$$\pi_{j,k,\ell}(n) := \binom{n-1}{j,k,\ell} 2^\ell \int_0^1 x^{2j+\ell}(1-x)^{2k+\ell}\,\mathrm{d}x = \binom{n-1}{j,k,\ell}\frac{(2j+\ell)!(2k+\ell)!2^\ell}{(2n-1)!}. \qquad (2.5)$$

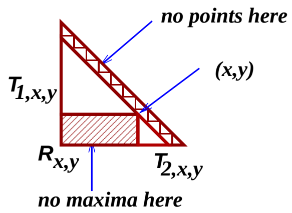

Figure 2: *Division of the original right triangle at the point where $x_j + y_j$ is maximized; the number of maxima is thus computed recursively.*

Proof. Let $X_j = (x_j, y_j)$, $j = 1, \ldots, n$, be n iid points taken uniformly in $\mathcal{T}$. Let $z_j = x_j + y_j$. The idea of the proof is to find the point, say X_1, that maximizes the sum $x_j + y_j$, writing for simplicity $X_1 = (x, y)$; and then to divide the triangle into two smaller triangles, $\mathcal{T}_{1,x,y}$ and $\mathcal{T}_{2,x,y}$, and a rectangle $\mathcal{R}_{x,y}$, as shown in Figure 2. Since the rectangle $\mathcal{R}_{x,y}$ contains no maxima, the number of maxima M_n in $\mathcal{T}$ equals 1 plus those in the two smaller triangles. The probability that there are j points in $\mathcal{T}_{1,x,y}$, k points in $\mathcal{T}_{2,x,y}$ and ℓ points in the rectangle $\mathcal{R}_{x,y}$ is equal to

$$\binom{n-1}{j,k,\ell} x^{2j} y^{2k} (2xy)^\ell \qquad (j + k + \ell = n - 1).$$

Thus we have

$$\begin{aligned}
f_n(w) &= nE\left(e^{M_n w} 1_{\{z_1 = \max\{z_1,\ldots,z_n\}\}}\right) \\
&= \frac{ne^w}{|\mathcal{T}|} \sum_{j+k+\ell=n-1} \binom{n-1}{j,k,\ell} E(e^{M_j w}) E(e^{M_k w}) \int_{\mathcal{T}} x^{2j} y^{2k} (2xy)^\ell \, \mathrm{d}x\, \mathrm{d}y \\
&= 2ne^w \sum_{j+k+\ell=n-1} \binom{n-1}{j,k,\ell} f_j(w) f_k(w) 2^\ell \int_{\mathcal{T}} x^{2j+\ell} y^{2k+\ell} \, \mathrm{d}x\, \mathrm{d}y \\
&= e^w \sum_{j+k+\ell=n-1} \binom{n-1}{j,k,\ell} f_j(w) f_k(w) \frac{(2j+\ell)!(2k+\ell)!2^\ell}{(2n-1)!},
\end{aligned}$$

from which (2.4) follows. $\square$

A "random divorce model?" The above probability distribution has a straightforward probabilistic interpretation: Given $n-1$ couples, we randomly divide them into two groups with sizes t and $2n-2-t$, where $0 \le t \le 2n-2$. Then $\pi_{j,k,\ell}(n)$ is the probability that there are j couples in one group and k couples in the other, the number of "un-coupled" being $\ell = t - 2j = 2n - 2 - t - 2k$ in both groups. From this point of view, the number of maxima in $\mathcal{T}$ can also be interpreted as the total number of steps needed to completely "divorce" $n-1$ couples by repeating the above procedure until no further such divisions are possible, namely, when the sizes of all subproblems (or number of couples) reduce to zero. Our result (2.3) is equivalent to saying that this quantity is asymptotically normally distributed; see Frieze and Pittel (1995) for a similar example.

Let $G(z,w) = e^{-z} \sum_{n\ge0} f_n(w) z^n/n!$. Then G is entire in z satisfying $G(0,w) = 1$ and

$$\frac{\partial}{\partial z} G(z,w) + G(z,w) = e^w \int_0^1 G(x^2 z, w) G((1-x)^2 z, w) \, \mathrm{d}x. \tag{2.6}$$

2.2 Mean and variance

We prove (2.1) and (2.2). Taking derivative on both sides of (2.6) with respective to w and substituting $w = 0$, we obtain $G_1(0) = 0$ and

$$G_1'(z) + G_1(z) = 1 + 2\int_0^1 G_1(x^2 z)\,\mathrm{d}x,$$

where $G_1(z) = (\partial/\partial w)G(z,w)\big|_{w=0}$. Write $G_1(z) = \sum_{n\geq 0} g_1(n)z^n$. Then $g_1(0) = 0$ and by equating coefficient of z^n on both sides

$$(n+1)g_1(n+1) + g_1(n) = \delta_{n0} + \frac{2g_1(n)}{2n+1} \qquad (n \geq 0),$$

where δ_{ab} is the Kronecker symbol. Solving this recurrence, we have

$$g_1(n) = \frac{(-1)^{n-1}}{(2n-1)n!} \qquad (n \geq 1).$$

Thus the mean of M_n is equal to

$$E(M_n) = n![z^n]e^z G_1(z) = \sum_{1\leq j\leq n} \binom{n}{j} \frac{(-1)^{j-1}}{2j-1} = \frac{\sqrt{\pi}\,n!}{\Gamma(n+1/2)} - 1 = \frac{4^n}{\binom{2n}{n}} - 1.$$

This proves (2.1). The asymptotic approximation of $E(M_n)$ follows from Stirling's formula. Note that (2.1) can be proved in a more straightforward way by computing the probability that a point is maximal.

Taking derivative twice with respect to w and substituting $w = 0$ in (2.6), we obtain $G_2(0) = 0$ and

$$G_2'(z) + G_2(z) = 1 + 2\int_0^1 G_2(x^2 z)\,\mathrm{d}x + 4\int_0^1 G_1(x^2 z)\,\mathrm{d}x + 2\int_0^1 G_1(x^2 z)G_1((1-x)^2 z)\,\mathrm{d}x,$$

where $G_2(z) = (\partial^2/\partial w^2)G(z,w)\big|_{w=0}$. Write $G_2(z) = \sum_{n\geq 0} g_2(n)z^n$. We have $g_2(0) = 0$ and

$$(n+1)g_2(n+1) + g_2(n) = \delta_{n0} + \frac{4g_1(n)}{2n+1} + 2\sum_{0\leq j\leq n} g_1(j)g_1(n-j)\frac{(2j)!(2n-2j)!}{(2n+1)!} + \frac{2g_2(n)}{2n+1},$$

for $n \geq 0$. Solving the recurrence, we obtain

$$g_2(n) = \frac{(-1)^{n-1}}{(2n-1)n!} + \sum_{0\leq j\leq n-2} \frac{(-1)^j (n-j)!(2n-2j-1)}{(2n-1)n!} g_{11}(n-j),$$

where

$$g_{11}(n) = \frac{4g_1(n-1)}{n(2n-1)} + \frac{2}{n(2n-1)!} \sum_{0\leq \ell\leq n-1} g_1(\ell)g_1(n-1-\ell)(2\ell)!(2n-2\ell-2)!.$$

Thus the second moment of M_n satisfies for $n \geq 1$

$$\begin{aligned}
E(M_n^2) &= n![z^n]e^z G_2(z) \\
&= E(M_n) + \sum_{2\leq m\leq n} \binom{n}{m} \frac{(-1)^m}{2m-1} \sum_{2\leq j\leq m} (-1)^j j!(2j-1)g_{11}(j) \\
&= E(M_n) + 4\sum_{2\leq m\leq n} \binom{n}{m} \frac{(-1)^m}{2m-1} \sum_{0\leq j\leq m-2} \frac{1}{2j+1} \\
&\quad + 2\sum_{3\leq m\leq n} \binom{n}{m} \frac{(-1)^{m-1}}{2m-1} \sum_{0\leq j\leq m-3} \frac{1}{(2j+1)(2j+3)} \sum_{0\leq \ell\leq j} \frac{\binom{j}{\ell}}{\binom{2j}{2\ell}} \\
&=: E(M_n) + \Sigma_1 + \Sigma_2.
\end{aligned}$$

To derive asymptotics of $E(M_n^2)$, since both sums Σ_1, Σ_2 are of the type $\sum_{n_0 \le j \le n} \binom{n}{j}(-1)^j a_j$ for some sequence a_j, which can essentially be regarded as the n-th difference of a_0, we use the associated integral representation from finite differences to evaluate the sums; see [23].

Lemma 1. *The sums Σ_1 and Σ_2 satisfy the asymptotic expansions*

$$\Sigma_1 \sim \frac{2\sqrt{\pi}\, n!}{\Gamma(n+1/2)} \left(\psi(n+1/2) + \gamma - 2\right) + 4 + \sum_{j\ge 0} \frac{\Gamma(j+1/2)n!}{\Gamma(n+j+3/2)(j+1)}, \tag{2.7}$$

$$\Sigma_2 \sim \pi n - \frac{2\sqrt{\pi} n!}{\Gamma(n+1/2)} \left(\psi(n+1/2) - \log 2 + \gamma\right) - 2 - \sum_{j\ge 0} \frac{\Gamma(j+1/2)n!}{\Gamma(n+j+3/2)(j+1)}, \tag{2.8}$$

where ψ denotes the logarithmic derivative of the Gamma function.

Proof. Observe that the function $\phi_1(s) := \psi(s-1/2)/2 + \gamma/2 + \log 2$ satisfies for $m \ge 2$

$$\phi_1(m) = \sum_{0\le j\le m-2} \frac{1}{2j+1}.$$

Thus we have [23]

$$\Sigma_1 = \frac{4}{2\pi i} \int_{3/2-i\infty}^{3/2+i\infty} \frac{(-1)^{n+1} n!\, \phi_1(s)}{s(s-1)\cdots(s-n)(2s-1)}\, \mathrm{d}s.$$

Since the growth rate of $\psi(s)$ at $\sigma \pm i\infty$ is of logarithmic order, we can evaluate the integral by shifting the line of integration to the left and by taking the residues of the poles encountered into account. There is no pole at $s = 1$ because $\phi_1(1) = 0$. The only singularities are (i) a double pole at $s = 1/2$ and (ii) simple poles at $s = 0$ and $s = -1/2 - j$ for $j \ge 0$. Collecting the residues of these poles, we obtain the asymptotic expansion (2.7).

Similarly, define for $\Re(s) > 1$

$$\phi_2(s) = \sum_{j\ge 0} \left(\frac{1}{2j+3} \int_0^1 \omega(x)^j \,\mathrm{d}x - \frac{1}{2j+2s-1} \int_0^1 \omega(x)^{j+s-2}\, \mathrm{d}x \right),$$

where $\omega(x) := 1 - 2x + 2x^2$. Then

$$\Sigma_2 = \frac{2}{2\pi i} \int_{5/2-i\infty}^{5/2+i\infty} \frac{(-1)^n n!\, \phi_2(s)}{s(s-1)\cdots(s-n)(2s-1)}\, \mathrm{d}s.$$

It is easy to show that $\phi_2(\sigma \pm iT) = O(1)$ for large $T > 1$. Note that $\phi_2(2) = 0$. We are left with the calculation of the residues at the (i) double pole at $s = 1/2$ and (ii) simple poles at $s = 1, 0$ and $s = -j - 1/2$ for $j \ge 0$. We need the following identities

$$\int_0^1 \omega(x)^{-1}\, \mathrm{d}x = \frac{\pi}{2}, \quad \int_0^1 \omega(x)^{-3/2}\, \mathrm{d}x = 2, \quad \int_0^1 \omega(x)^{-2}\, \mathrm{d}x = 1 + \frac{\pi}{2}.$$

¿From these it follows that $\phi_2(1) = -\pi/2$, $\phi_2(0) = 1$, and

$$\phi_2(s) = -\frac{1}{s-1/2} + c_1 + O(|s-1/2|) \qquad (s \sim 1/2),$$

$$\phi_2(s) = -\frac{1}{s+j+1/2} + O(1) \qquad (s \sim -j-1/2,\, j \ge 1),$$

where

$$\begin{aligned}
c_1 &= \sum_{j\ge 0}\frac{1}{2j+3}\int_0^1 \omega(x)^j\,\mathrm{d}x-\frac{1}{2}\int_0^1\frac{\log\omega(x)}{\omega(x)^{3/2}}\,\mathrm{d}x-\sum_{j\ge 1}\frac{1}{2j}\int_0^1\omega(x)^{j-3/2}\,\mathrm{d}x\\
&= -\frac{\pi}{2}+2\int_0^{1/2}\omega(x)^{-3/2}\log\frac{1+\sqrt{\omega(x)}}{\sqrt{\omega(x)}}\,\mathrm{d}x\\
&= 3\log 2-2.
\end{aligned}$$

Collecting residues from the integral representation of Σ_2, we obtain (2.8). □

Noting that there are perfect cancellations of the ψ-terms and the last sums in both expansions (2.7) and (2.8), we obtain

$$E(M_n^2)=E(M_n)+\Sigma_1+\Sigma_2=\pi n+\frac{\sqrt{\pi}n!}{\Gamma(n+1/2)}\,(2\log 2-3)+1+O\left(n^{-K}\right),$$

for any $K>0$. This completes the proof of (2.2) for the variance of M_n. □

2.3 Higher moments

We prove in this section that for $m\ge 1$

$$\begin{aligned}
E\left(M_n-\sqrt{\pi n}\right)^{2m} &\sim h_m n^{m/2}, && (2.9)\\
E\left(M_n-\sqrt{\pi n}\right)^{2m-1} &= o\left(n^{m/2-1/4}\right), && (2.10)
\end{aligned}$$

where $h_m := (2m)!\sigma_2^{2m}/(2^m m!)$. From these the asymptotic normality of $(M_n-\sqrt{\pi n})/(\sigma_2 n^{1/4})$ will follow.

The case $m=1$ holds by (2.1) and (2.2). We prove the remaining cases $m\ge 2$ by induction. A different approach for the asymptotics of higher moments is needed since the preceding one becomes too involved for moments of degree ≥ 3; see Bai et al. [4].

Shifting the mean. Let $P_n(w):=E(e^{(M_n-\sqrt{\pi n})w})=f_n(w)e^{-\sqrt{\pi n}w}$. Then $P_0(w)=1$ and, by (2.4),

$$P_n(w)=\sum_{j+k+\ell=n-1}\pi_{j,k,\ell}(n)P_j(w)P_k(w)e^{\Delta_{n,j,k}w}\qquad (n\ge 1), \tag{2.11}$$

where $\Delta_{n,j,k}:=\sqrt{\pi j}+\sqrt{\pi k}-\sqrt{\pi n}+1$. Define $P_{n,m}:=P_n^{(m)}(0)=E(M_n-\sqrt{\pi n})^m$.

Lemma 2. *The sequences $P_{n,m}$ satisfy the recurrence $P_{0,m}=0$, $P_{1,m}=(1-\sqrt{\pi})^m$ and for $n\ge 2$ and $m\ge 1$*

$$P_{n,m}=\frac{(n-1)!}{\Gamma(n+1/2)}\sum_{0\le j<n}\frac{\Gamma(j+1/2)}{j!}P_{j,m}+R_{n,m}^{(1)}+R_{n,m}^{(2)}, \tag{2.12}$$

where

$$\begin{aligned}
R_{n,m}^{(1)} &:= \sum_{1\le p<m}\binom{m}{p}\sum_{j+k+\ell=n-1}\pi_{j,k,\ell}(n)P_{j,p}P_{k,m-p},\\
R_{n,m}^{(2)} &:= \sum_{\substack{p+q+r=m\\ 0\le p,q<m\\ 1\le r\le m}}\binom{m}{p,q,r}\sum_{j+k+\ell=n-1}\pi_{j,k,\ell}(n)P_{j,p}P_{k,q}\Delta_{n,j,k}^r.
\end{aligned}$$

Proof. By (2.11),

$$\begin{aligned}
P_{n,m} &= \sum_{p+q+r=m} \binom{m}{p,q,r} \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) P_{j,p} P_{k,q} \Delta^r_{n,j,k} \\
&= \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) \left(P_{j,m} + P_{k,m}\right) + R^{(1)}_{n,m} + R^{(2)}_{n,m} \\
&= 2 \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) P_{j,m} + R^{(1)}_{n,m} + R^{(2)}_{n,m}.
\end{aligned}$$

Now, by the integral representation in (2.5), we have

$$\begin{aligned}
&2 \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) P_{j,m} \\
&\quad = 2 \sum_{0\le j<n} \binom{n-1}{j} P_{j,m} \int_0^1 x^{2j}(1-x)^{n-1-j} \sum_{0\le k<n-j} \binom{n-1-j}{k} (2x)^{n-1-j-k}(1-x)^k \,\mathrm{d}x \\
&\quad = 2 \sum_{0\le j<n} \binom{n-1}{j} P_{j,m} \int_0^1 x^{2j}(1-x^2)^{n-1-j} \,\mathrm{d}x.
\end{aligned}$$

The recurrence (2.12) follows. $\square$

Solution and asymptotic transfers of the recurrence. We first study recurrences of the type (2.12).

Lemma 3. *Assume $a_0 = 0$ and*

$$a_n = \frac{(n-1)!}{\Gamma(n+1/2)} \sum_{0\le j<n} \frac{\Gamma(j+1/2)}{j!}\, a_j + b_n \qquad (n \ge 1), \tag{2.13}$$

where b_n is a given sequence. Then for $n \ge 1$

$$a_n = b_n + \frac{n!}{\Gamma(n+1/2)} \sum_{0\le j<n} \frac{\Gamma(j+1/2)}{(j+1)!}\, b_j. \tag{2.14}$$

Proof. Define $\tilde a_n = \Gamma(n+1/2)a_n/n!$ and $\tilde b_n = \Gamma(n+1/2)b_n/n!$. Then $\tilde a_0 = 0$ and

$$\tilde a_n = \frac{1}{n} \sum_{0\le j<n} \tilde a_j + \tilde b_n \qquad (n \ge 1).$$

By considering the difference $n\tilde a_n - (n-1)\tilde a_{n-1}$ and iterating, we obtain

$$\tilde a_n = \tilde b_n + \sum_{1\le j<n} \frac{\tilde b_j}{j+1} \qquad (n \ge 1),$$

which yields (2.14). $\square$

In particular, if $b_n = 1$ for $n \ge 1$, then $a_n = E(X_n) = \sqrt{\pi}\, n!/\Gamma(n+1/2) - 1$.

Lemma 4 (Asymptotic transfer). *Assume that a_n satisfies (2.13) and that $\alpha > 1/2$. If $b_n \sim cn^\alpha$, then*

$$a_n \sim c\frac{2\alpha+1}{2\alpha-1}\, n^\alpha; \tag{2.15}$$

if $b_n = O(n^\alpha)$, or $b_n = o(n^\alpha)$, then $a_n = O(n^\alpha)$, or $a_n = o(n^\alpha)$, respectively.

Proof. By (2.14). □

By (2.12) and the asymptotic transfer lemma, the proofs of (2.9) and (2.10) are reduced, by induction, to estimating the asymptotics of $R^{(1)}_{n,m}$ and $R^{(2)}_{n,m}$.

Asymptotics of $R^{(1)}_{n,m}$. By (2.9), (2.10) and induction, we have for $m \geq 2$

$$\begin{aligned} R^{(1)}_{n,2m-1} &= \sum_{1\leq p\leq 2m-2} \binom{2m-1}{p} \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) P_{j,p} P_{k,2m-1-p} \\ &= o\left(\sum_{1\leq p\leq 2m-2} \binom{2m-1}{p} \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) j^{p/4} k^{(2m-1-p)/4} \right) \\ &= o\left(n^{m/2-1/4}\right). \end{aligned} \tag{2.16}$$

On the other hand,

$$\begin{aligned} R^{(1)}_{n,2m} &= \sum_{1\leq p< m} \binom{2m}{2p} \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) P_{j,p} P_{k,2m-p} + o(n^{m/2}) \\ &\sim \sum_{1\leq p<m} \binom{2m}{2p} h_p h_{m-p} U_{m,p}(n), \end{aligned}$$

where

$$U_{m,p} := \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) j^{p/2} k^{(m-p)/2}.$$

By (2.5)

$$\begin{aligned} U_{m,p}(n) n^{-m/2} &= \int_0^1 \sum_{j+k+\ell=n-1} \binom{n-1}{j,k,\ell} x^{2j}(1-x)^{2k}[2x(1-x)]^\ell (j/n)^{p/2}(k/n)^{(m-p)/2}\, \mathrm{d}x \\ &:= \int_0^1 E\left[(J/n)^{p/2}(K/n)^{(m-p)/2}\right] \mathrm{d}x, \end{aligned}$$

where (J, K, Λ) denotes a trinomial distribution with parameters $(n-1; x^2, (1-x)^2, 2x(1-x))$. By Bernstein's inequality, we have, uniformly in x and for sufficiently large n,

$$\begin{aligned} & P\left(\left|J/n - x^2\right| \geq n^{-1/3}\left[n^{-7/12} \vee x(1-x)\right]\right) \\ &\leq 2\exp\left(-\frac{1}{5}\min\left\{n^{1/3}, n^{2/3}\left[n^{-7/12} \vee x(1-x)\right]\right\}\right) \\ &\leq 2\exp\left(-\frac{1}{5}n^{1/12}\right); \end{aligned}$$

and, similarly,

$$P\left(\left|K/n-(1-x)^2\right|\geq n^{-1/3}\left[n^{-7/12}\vee x(1-x)\right]\right)\leq 2\exp\left(-\frac{1}{5}n^{1/12}\right).$$

When $|J/n-x^2|\leq n^{-1/3}x(1-x)$, $|K/n-(1-x)^2|\leq n^{-1/3}x(1-x)$ and $n^{-1/3}<x<1-n^{-1/3}$, we have

$$\begin{aligned}\left|(J/n)^{p/2}(K/n)^{(m-p)/2}-x^p(1-x)^{(m-p)/2}\right| &\leq p\left|\sqrt{J/n}-x\right|+(m-p)\left|\sqrt{K/n}-(1-x)\right|\\ &= \frac{p|J/n-x^2|}{\sqrt{J/n}+x}+\frac{(m-p)|K/n-(1-x)^2|}{\sqrt{K/n}+(1-x)}\\ &\leq mn^{-1/3}.\end{aligned}$$

Thus,

$$\begin{aligned}U_{m,p}(n)n^{-m/2} &= \int_{n^{-1/3}}^{1-n^{-1/3}} x^p(1-x)^{m-p}\,\mathrm{d}x+O(n^{-1/3})\\ &= \frac{p!(m-p)!}{(m+1)!}+O(n^{-1/3}).\end{aligned}$$

Consequently,

$$R^{(1)}_{n,2m}\sim\frac{n^{m/2}}{m+1}\sum_{1\leq p<m}\frac{\binom{2m}{2p}}{\binom{m}{p}}h_p h_{m-p}=\frac{m-1}{m+1}h_m n^{m/2}. \tag{2.17}$$

Estimate for $R^{(2)}_{n,m}$. We use a slightly different method to estimate $R^{(2)}_{n,m}$ since there are additional cancellations caused by the factor $\Delta^r_{n,j,k}$. Denote again by (J,K,Λ) a trinomial distribution with parameters $(n-1;x^2,(1-x)^2,2x(1-x))$. By induction, (2.9) and (2.10), we have

$$R^{(2)}_{n,m}=\sum_{\substack{p+q+r=m\\ 0\leq p,q<m\\ 1\leq r\leq m}}\binom{m}{p,q,r}O\left(n^{(m+r)/4}\right)V_r(n),$$

where

$$\begin{aligned}V_r(n) &:= n^{-r/2}\sum_{j+k+\ell=n-1}\pi_{j,k,\ell}(n)|\Delta_{j,k}(n)|^r\\ &= \int_0^1 E\left|\sqrt{\pi J/n}+\sqrt{\pi K/n}-\sqrt{\pi}+n^{-1/2}\right|^r\,\mathrm{d}x.\end{aligned}$$

By an argument similar to the estimate of $U_{m,p}(n)$ using Bernstein's inequality, we have

$$\left|\sqrt{\pi J/n}+\sqrt{\pi K/n}-\sqrt{\pi}+n^{-1/2}\right|^r=o(n^{-r/4}),$$

for $|J/n-x^2|\leq n^{-1/3}\left[n^{-7/12}\vee x(1-x)\right]$ and $|K/n-(1-x)^2|\leq n^{-1/3}\left[n^{-7/12}\vee x(1-x)\right]$. Also

$$\begin{aligned}P\left(\left|J/n-x^2\right|\geq n^{-1/3}\left[n^{-7/12}\vee x(1-x)\right]\right) &\leq 2e^{-n^{1/12}/5}\\ P\left(\left|K/n-(1-x)^2\right|\geq n^{-1/3}\left[n^{-7/12}\vee x(1-x)\right]\right) &\leq 2e^{-n^{1/12}/5}.\end{aligned}$$

Thus

$$V_r(n)=o(n^{-r/4})+O\left(e^{-n^{1/12}/5}\right)=o(n^{-r/4}),$$

from which it follows that

$$R^{(2)}_{n,m} = o(n^{m/4}). \tag{2.18}$$

An alternative way of estimating $V_r(n)$ is as follows. Using the inequality $|x| \le e^x + e^{-x}$ for $x \in \mathbb{R}$, we have

$$V_r(n)n^{r/2} \le \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n)\left(e^{r\Delta_{n,j,k}} + e^{-r\Delta_{n,j,k}}\right) =: V_r^+(n) + V_r^-(n).$$

Let $W(z) := \sum_{j\ge 0} e^{r\sqrt{\pi j}} z^j/j!$. Then

$$V_r^+(n) = e^{r(1-\sqrt{\pi n})}(n-1)![z^{n-1}] \int_0^1 W(x^2z)W((1-x)^2z)e^{2x(1-x)z}\,\mathrm{d}x.$$

By the local limit theorem for Poisson distribution (or the saddlepoint method), we have

$$W(y) = O\left(e^{y+r\sqrt{\pi y}}\right) \qquad (y \to \infty).$$

Using Cauchy's integral formula and this estimate, we obtain

$$\begin{aligned}
V_r^+(n) &\le e^{r(1-\sqrt{\pi n})}\left|\frac{(n-1)!}{2\pi i}\oint_{|z|=n} z^{-n}\int_0^1 W(x^2z)W((1-x)^2z)e^{2x(1-x)z}\,\mathrm{d}x\,\mathrm{d}z\right| \\
&\le e^{r(1-\sqrt{\pi n})}\frac{(n-1)!}{2\pi}n^{1-n}\int_{-\pi}^{\pi}\int_0^1 W(x^2n)W((1-x)^2n)e^{2x(1-x)n\cos\theta}\,\mathrm{d}x\,\mathrm{d}\theta \\
&= O\left(n!n^{-n}e^n\int_0^1\int_{-\pi}^{\pi} e^{-2x(1-x)n(1-\cos\theta)}\,\mathrm{d}\theta\,\mathrm{d}x\right).
\end{aligned}$$

By the inequality $1-\cos\theta \ge 2\theta^2/\pi^2$ for $|\theta| \le \pi$, we have

$$V_r^+(n) = O\left(n!n^{-n}e^n n^{-1/2}\int_0^1 x^{-1/2}(1-x)^{-1/2}\,\mathrm{d}x\right) = O(1).$$

Similarly, $V_r^-(n) = O(1)$. Thus $V_r(n) = O(n^{-r/2})$ and (2.18) follows.

Asymptotics of $P_{n,m}$. By the estimates (2.16), (2.17) and (2.18), we obtain

$$\begin{aligned}
R^{(1)}_{n,2m} + R^{(2)}_{n,2m} &\sim \frac{m-1}{m+1} \times \frac{(2m)!}{2^m m!}\sigma^{2m}n^{m/2} \\
R^{(1)}_{n,2m-1} + R^{(2)}_{n,2m-1} &= o\left(n^{m/2-1/4}\right),
\end{aligned}$$

for $m \ge 2$. The results (2.9) and (2.10) follow then from applying the asymptotic transfer lemma. This completes the proof of (2.3) and Theorem 3. □

3 Maxima in convex polygons

We prove Theorem 1 in this section.

3.1 Mean value of $M_n(\mathcal{P})$

We prove (1.3) in this section. Our approach can actually provide an asymptotic expansion but we content ourselves with (1.3) for simplicity.

Let $X_1, \dots, X_n$ be iid random variables uniformly distributed in $\mathcal{P}$. For $\mathcal{A} \subset \mathcal{P}$, denote by $M_n(\mathcal{A})$ the number of maxima of the point set $\{X_1, \dots, X_n\}$ lying in $\mathcal{A}$. Also write $M_n = M_n(\mathcal{P})$.

Observe first that for $\mathcal{A} \subset \mathcal{P}$

$$\begin{aligned} E(M_n(\mathcal{A})) &= nP\{X_1 \in \mathcal{A} \text{ and is a maximal point of } \{X_1, \dots, X_n\}\} \\ &= \frac{n}{|\mathcal{P}|} \int_{\mathcal{A}} \left(1 - \frac{|\mathcal{P} \cap \{s : s \succ t\}|}{|\mathcal{P}|}\right)^{n-1} \mathrm{d}t. \end{aligned} \tag{3.1}$$

The following lemma is useful for estimating Laplace-type integrals encountered in this paper.

Lemma 5. *For $0 \le t \le 1$ and $n \ge 1$, the inequalities*

$$e^{-nt}\left(1 - nt^2\right) \le (1-t)^n \le e^{-nt} \tag{3.2}$$

hold.

Proof. The right inequality is obvious. For the left inequality, we have

$$\begin{aligned} e^{-nt} - (1-t)^n &= e^{-nt}\left(1 - \left(e^t(1-t)\right)^n\right) \\ &\le e^{-nt}\left(1 - (1-t^2)^n\right) \\ &\le e^{-nt} nt^2, \end{aligned}$$

by using the inequalities $e^t \ge 1 + t$ and $(1-t)^n \ge 1 - nt$ (by induction). □

Most Laplace integrals in this paper are of the type

$$\int_0^1 \Big(1 - T(x)\left(1 + O(E(x))\right)\Big)^{n-1} \mathrm{d}x,$$

for some T and E. If we apply the above lemma and the inequality $|e^t - 1| \le te^t$, we obtain

$$\int_0^1 \Big(1 - T(x)\left(1 + O(E(x))\right)\Big)^{n-1} \mathrm{d}x = \int_0^1 e^{-nT(x)}\left(1 + O\left(T(x) + nT^2(x) + nE(x)\right)\right) \mathrm{d}x,$$

which is usually easier to deal with than a "microscopic analysis".

Case 1: $I_u \cap I_r = \emptyset$. Let $\mathcal{P}_j$ be the right triangle with hypotenuse s_j (and under s_j), the angle formed by s_j and the adjacent being θ_j (see Figure 3). Note that part of $\mathcal{P}_j$ may lie outside $\mathcal{P}$ (see Figure 4(b)).

Lemma 6. *If $\mathcal{P}_j \subset \mathcal{P}$, then*

$$E(M_n(\mathcal{P}_j)) = \left(\frac{|\mathcal{P}_j|\,\pi n}{|\mathcal{P}|}\right)^{1/2} - 1 + O\left(n^{-1/2}\right), \tag{3.3}$$

for $j = 1, \dots, \nu$, where $|\mathcal{P}_j| = |s_j|^2 \cos(\theta_j)\sin(\theta_j)/2$.

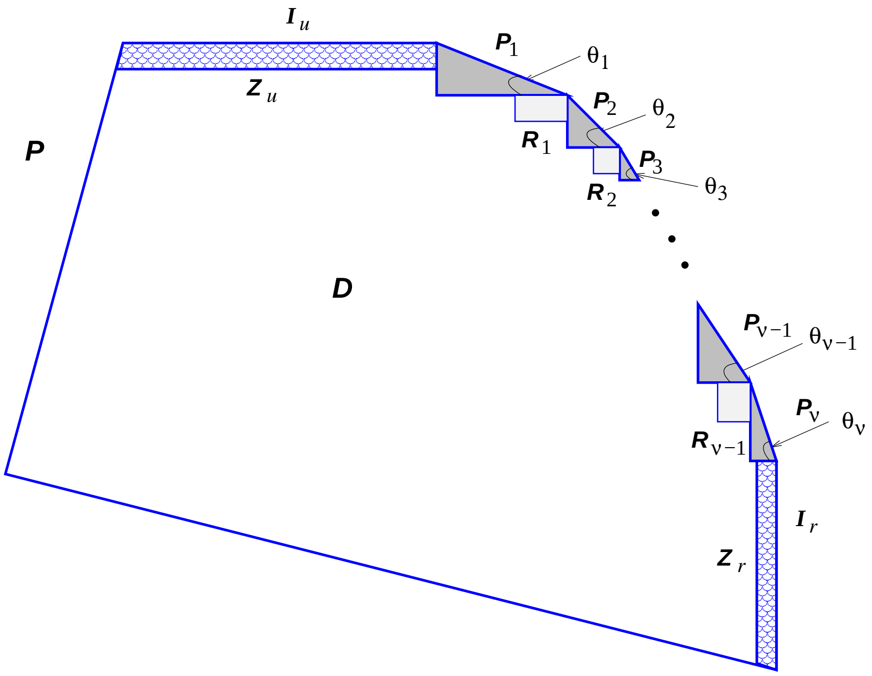

Figure 3: *Dissection of the upper-right part of a convex polygon* $\mathcal{P}$.

Proof. By (3.1) and the integral representation for finite differences [23], we have

$$\begin{aligned} E(M_n(\mathcal{P}_j)) &= \frac{n}{|\mathcal{P}|}\int_0^{|s_j|\cos\theta_j}\int_0^{x\tan\theta_j}\left(1-\frac{\cot\theta_j}{2\,|\mathcal{P}|}\,y^2\right)^{n-1}\mathrm{d}y\,\mathrm{d}x \\ &= \sum_{1\le k\le n}\binom{n}{k}(-1)^{k-1}\frac{(|\mathcal{P}_j|/|\mathcal{P}|)^k}{2k-1} \\ &= \frac{1}{2\pi i}\int_{3/4-i\infty}^{3/4+i\infty}\frac{(-1)^n n!(|\mathcal{P}_j|/|\mathcal{P}|)^s}{s(s-1)\cdots(s-n)(2s-1)}\,\mathrm{d}s \\ &= \frac{(|\mathcal{P}_j|/|\mathcal{P}|)^{1/2}\sqrt{\pi}\,n!}{\Gamma(n+1/2)}-1+O\left(n^{-K}\right), \end{aligned}$$

for any $K>0$. The formula (3.3) follows. □

Let $\mathcal{R}_j$ be a rectangle between $\mathcal{P}_j$ and $\mathcal{P}_{j+1}$ as shown in Figure 3 (the exact position of the lower-left boundary of $\mathcal{R}_j$ being immaterial).

Lemma 7. *For* $j=1,\ldots,\nu-1$,

$$E(M_n(\mathcal{R}_j)) = \frac{\sqrt{\tan\theta_{j+1}}}{2\sqrt{\tan\theta_{j+1}-\tan\theta_j}}\log\frac{\sqrt{\tan\theta_{j+1}}+\sqrt{\tan\theta_{j+1}-\tan\theta_j}}{\sqrt{\tan\theta_{j+1}}-\sqrt{\tan\theta_{j+1}-\tan\theta_j}}+O(n^{-K}), \qquad (3.4)$$

for any $K>0$.

Proof. By (3.1) and a change of variables, we have

$$\begin{aligned}
E(M_n(\mathcal{R}_j)) &= \frac{n}{|\mathcal{P}|}\int_0^\varepsilon\int_0^\varepsilon \left(1-\frac{1}{|\mathcal{P}|}\left(\frac{\tan\theta_j}{2}\,x^2+xy+\frac{\cot\theta_{j+1}}{2}\,y^2\right)\right)^{n-1} \mathrm{d}y\,\mathrm{d}x + O(n^{-K}) \\
&= \frac{n}{|\mathcal{P}|}\int_0^\varepsilon\int_0^{\varepsilon/x} x\left(1-\frac{x^2}{2\,|\mathcal{P}|}\,\zeta(y)\right)^{n-1} \mathrm{d}y\,\mathrm{d}x + O(n^{-K}),
\end{aligned}$$

for any $K>0$, where $\zeta(y)=\tan\theta_j+2y+y^2\cot\theta_{j+1}$. Interchanging the order of integration, the first term on the right-hand side is equal to

$$\begin{aligned}
&\int_0^1 \zeta(y)^{-1}\left(1-\left(1-\frac{\varepsilon^2\zeta(y)}{2|\mathcal{P}|}\right)^n\right)\mathrm{d}y + \int_1^\infty \zeta(y)^{-1}\left(1-\left(1-\frac{\varepsilon^2\zeta(y)}{2|\mathcal{P}|y^2}\right)^n\right)\mathrm{d}y \\
&\qquad = \int_0^\infty \zeta(y)^{-1}\,\mathrm{d}y + O(n^{-K}),
\end{aligned}$$

from which (3.4) follows. □

Note that when $\theta_{j+1}\to\theta_j$, $E(M_n(\mathcal{R}_j))\to 1$, which cancels nicely with the constant in (3.3); this is consistent with the result (2.1) for $\mathcal{T}$.

Let $\mathcal{Z}_u, \mathcal{Z}_r$ be the trapezoids formed by I_u and I_r, respectively, as shown in Figure 3 when $|I_u|>0$ and $|I_r|>0$.

Lemma 8. *If* $|I_u|>0$ *then*

$$E(M_n(\mathcal{Z}_u)) = \frac{1}{2}\log n + \frac{\gamma}{2} + \frac{1}{2}\log\frac{2|I_u|^2\tan\theta_1}{|\mathcal{P}|} + O(n^{-1/2});$$

and if $|I_r|>0$ *then*

$$E(M_n(\mathcal{Z}_r)) = \frac{1}{2}\log n + \frac{\gamma}{2} + \frac{1}{2}\log\frac{2|I_r|^2\cot\theta_\nu}{|\mathcal{P}|} + O(n^{-1/2}).$$

Proof. Again by (3.1), we obtain, using (3.2)

$$\begin{aligned}
E(M_n(\mathcal{Z}_u)) &= \frac{n}{|\mathcal{P}|}\int_0^{|I_u|}\int_0^\varepsilon \left(1-\frac{1}{|\mathcal{P}|}\left(xy+\frac{y^2}{2}\cot\theta_1\right)\right)^{n-1} \mathrm{d}y\,\mathrm{d}x + O(n^{-K}) \\
&= \int_0^{|I_u|} y^{-1}\left(\left(1-\frac{\cot\theta_1}{2|\mathcal{P}|}\,y^2\right)^n - \left(1-\frac{y}{2|\mathcal{P}|}\,(2|I_u|+y\cot\theta_1)\right)^n\right)\mathrm{d}y + O(n^{-K}) \\
&= \int_0^1 y^{-1}\left(\exp\left(-\frac{\cot\theta_1}{2|\mathcal{P}|}\,y^2 n\right) - \exp\left(-\frac{y}{2|\mathcal{P}|}\,(2|I_u|+y\cot\theta_1)\,n\right)\right)\mathrm{d}y + O(n^{-1/2}) \\
&= \frac{1}{2}\log n + \frac{\gamma}{2} + \frac{1}{2}\log\frac{2\,|I_u|^2\tan\theta_1}{|\mathcal{P}|} + O(n^{-1/2}),
\end{aligned}$$

for any $K>0$. The proof for $E(M_n(\mathcal{Z}_r))$ is similar. □

When $|I_u|=0$, there are two further possible cases: $\mathcal{P}_1\subset\mathcal{P}$ and $\mathcal{P}_1\not\subset\mathcal{P}$. Let τ denote the segment on $\mathcal{P}$ having the common intersection point I_u with s_1. In the first case $\mathcal{P}_1\subset\mathcal{P}$, let $\mathcal{T}_u$ be the right triangle with hypotenuse τ and let θ_u be the angle between the hypotenuse and the opposite; see Figure 4(a). In the second case, we denote by $\mathcal{S}_u$ and θ_u the right triangle with hypotenuse τ and the angle formed by τ and the opposite, respectively; see Figure 4(b).

In a parallel manner, when $|I_r|=0$, we distinguish between $\mathcal{P}_\nu\subset\mathcal{P}$ and $\mathcal{P}_\nu\not\subset\mathcal{P}$ and define β, $\mathcal{T}_r$, $\mathcal{S}_r$, and θ_r as shown in Figure 4.

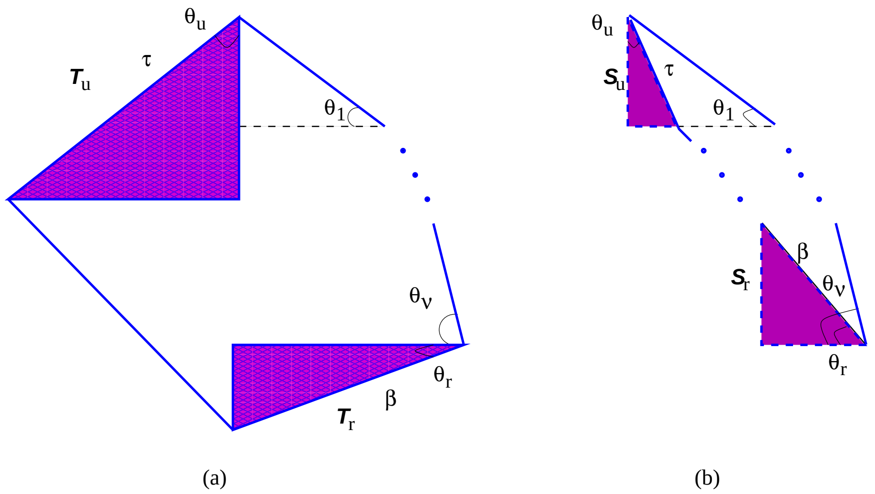

Figure 4: *Illustration of the cases when* $I_u \cap I_r = \emptyset$ *and either* $|I_u| = 0$ *or* $|I_r| = 0$.

Lemma 9. *If* $|I_u| = 0$ *and* $\mathcal{P}_1 \subset \mathcal{P}$ *then*

$$E(M_n(\mathcal{T}_u)) = \frac{1}{\sqrt{\cot\theta_u \cot\theta_1 + 1}} \log \frac{\sqrt{\cot\theta_u \cot\theta_1 + 1} + 1}{\sqrt{\cot\theta_u \cot\theta_1 + 1} - 1} + O(n^{-1/2}); \tag{3.5}$$

if $|I_r| = 0$ *and* $\mathcal{P}_\nu \subset \mathcal{P}$ *then*

$$E(M_n(\mathcal{T}_r)) = \frac{1}{\sqrt{\cot\theta_r \tan\theta_\nu + 1}} \log \frac{\sqrt{\cot\theta_r \tan\theta_\nu + 1} + 1}{\sqrt{\cot\theta_r \tan\theta_\nu + 1} - 1} + O(n^{-1/2}). \tag{3.6}$$

Proof. By (3.1),

$$E(M_n(\mathcal{T}_u)) = \frac{n}{|\mathcal{P}|} \int_0^{\varepsilon} \int_0^{y \tan\theta_u} \left(1 - \frac{1}{|\mathcal{P}|}\left(xy - \frac{\tan\theta_u}{2} x^2 + \frac{\cot\theta_1}{2} y^2\right)\right)^{n-1} \mathrm{d}x\, \mathrm{d}y + O(n^{-K}),$$

for any $K > 0$. The remaining proof of (3.5) is similar to the derivations for (3.4). The result (3.6) is proved in a similar manner. □

Lemma 10. *If* $|I_u| = 0$ *and* $\mathcal{P}_1 \not\subset \mathcal{P}$ *then*

$$E(M_n(\mathcal{S}_u)) = \frac{\tan\theta_u}{\cot\theta_1 - \tan\theta_u} + O(n^{-1/2}); \tag{3.7}$$

if $|I_r| = 0$ *and* $\mathcal{P}_\nu \not\subset \mathcal{P}$ *then*

$$E(M_n(\mathcal{S}_r)) = \frac{\tan\theta_r}{\cot\theta_\nu - \tan\theta_r} + O(n^{-1/2}). \tag{3.8}$$

Proof. These follow from the integral representations

$$\begin{aligned}
E(M_n(\mathcal{S}_u)) &= \frac{n}{|\mathcal{P}|} \int_0^{\varepsilon} \int_0^{y \tan\theta_u} \left(1 - \frac{\tan\theta_1}{2|\mathcal{P}|}(y\cot\theta_1 - x)^2\right)^{n-1} \mathrm{d}x\, \mathrm{d}y + O(n^{-K}), \\
E(M_n(\mathcal{S}_r)) &= \frac{n}{|\mathcal{P}|} \int_0^{\varepsilon} \int_0^{y \tan\theta_r} \left(1 - \frac{\tan\theta_\nu}{2|\mathcal{P}|}(y\cot\theta_\nu - x)^2\right)^{n-1} \mathrm{d}x\, \mathrm{d}y + O(n^{-K}),
\end{aligned}$$

and an analysis similar to the preceding cases. $\square$

When some part of $\mathcal{P}_j$ lies outside $\mathcal{P}$, namely, $\mathcal{P}_j \cap \mathcal{P} \neq \mathcal{P}_j$, we have

$$\begin{aligned} E(M_n(\mathcal{P}_1 \cap \mathcal{P})) &= E(M_n(\mathcal{P}_1) - E(M_n(\mathcal{S}_u)), \\ E(M_n(\mathcal{P}_\nu \cap \mathcal{P})) &= E(M_n(\mathcal{P}_\nu) - E(M_n(\mathcal{S}_r)), \\ E(M_n(\mathcal{P}_j \cap \mathcal{P})) &= E(M_n(\mathcal{P}_j)) + O(n^{-K}) \qquad (j = 2, \ldots, \nu - 1). \end{aligned}$$

Let

$$\mathcal{D} = \mathcal{P} \setminus (\mathcal{P}_1 \cup \cdots \cup \mathcal{P}_\nu \cup \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_{\nu-1} \cup (\mathcal{Z}_u \text{ or } T_u) \cup (\mathcal{Z}_r \text{ or } T_r));$$

(see Figure 3). Then $E(M_n(\mathcal{D})) = O(n^{-K})$, for any $K > 0$, since there exists an $\varepsilon > 0$ such that $|\mathcal{P} \cap \{s : s \succ t\}| > \varepsilon$, for all $t \in \mathcal{D}$.
Thus when $I_u \cap I_r = \emptyset$, we have

$$\begin{aligned} E(M_n) &= \sum_{1 \le j \le \nu} E(M_n(\mathcal{P}_j)) + \sum_{1 \le j \le \nu - 1} E(M_n(\mathcal{R}_j)) + 1_{\{|I_u|>0\}} E(M_n(\mathcal{Z}_u)) + 1_{\{|I_r|>0\}} E(M_n(\mathcal{Z}_r)) \\ &\quad + 1_{\{|I_u|=0\}} 1_{\{\mathcal{P}_1 \subset \mathcal{P}\}} E(M_n(\mathcal{T}_u)) + 1_{\{|I_r|=0\}} 1_{\{\mathcal{P}_\nu \subset \mathcal{P}\}} E(M_n(\mathcal{T}_r)) \\ &\quad - 1_{\{|I_u|=0\}} 1_{\{\mathcal{P}_1 \not\subset \mathcal{P}\}} E(M_n(\mathcal{S}_u)) - 1_{\{|I_r|=0\}} 1_{\{\mathcal{P}_\nu \not\subset \mathcal{P}\}} E(M_n(\mathcal{S}_r)) \\ &\quad + O\left((\nu + 1) n^{-1/2}\right). \end{aligned}$$

This proves (1.3) when $\pi_1 > 0$ with

$$\begin{aligned} \pi_3 &= -\nu + \sum_{1 \le j \le \nu - 1} \frac{\sqrt{\tan\theta_{j+1}}}{2\sqrt{\tan\theta_{j+1} - \tan\theta_j}} \log \frac{\sqrt{\tan\theta_{j+1}} + \sqrt{\tan\theta_{j+1} - \tan\theta_j}}{\sqrt{\tan\theta_{j+1}} - \sqrt{\tan\theta_{j+1} - \tan\theta_j}} \\ &\quad + 1_{\{|I_u|>0\}} \left(\frac{\gamma}{2} + \frac{1}{2} \log \frac{2|I_u|^2 \tan\theta_1}{|\mathcal{P}|}\right) 1 + 1_{\{|I_r|>0\}} \left(\frac{\gamma}{2} + \frac{1}{2} \log \frac{2|I_r|^2 \cot\theta_\nu}{|\mathcal{P}|}\right) \\ &\quad + 1_{\{|I_u|=0\}} 1_{\{\mathcal{P}_1 \subset \mathcal{P}\}} \frac{1}{\sqrt{\cot\theta_u \cot\theta_1 + 1}} \log \frac{\sqrt{\cot\theta_u \cot\theta_1 + 1} + 1}{\sqrt{\cot\theta_u \cot\theta_1 + 1} - 1} \\ &\quad + 1_{\{|I_r|=0\}} 1_{\{\mathcal{P}_\nu \subset \mathcal{P}\}} \frac{1}{\sqrt{\cot\theta_r \tan\theta_\nu + 1}} \log \frac{\sqrt{\cot\theta_r \tan\theta_\nu + 1} + 1}{\sqrt{\cot\theta_r \tan\theta_\nu + 1} - 1} \\ &\quad - 1_{\{|I_u|=0\}} 1_{\{\mathcal{P}_1 \not\subset \mathcal{P}\}} \frac{\tan\theta_u}{\cot\theta_1 - \tan\theta_u} - 1_{\{|I_r|=0\}} 1_{\{\mathcal{P}_\nu \not\subset \mathcal{P}\}} \frac{\tan\theta_r}{\cot\theta_\nu - \tan\theta_r}. \end{aligned} \tag{3.9}$$

Case 2: $I_u \cap I_r \neq \emptyset$**.** We further divide into four cases. First, if $|I_u| > 0$ and $|I_r| > 0$ (see Figure 5(a)), then

$$\begin{aligned} E(M_n) &= \frac{n}{|\mathcal{P}|} \int_0^{|I_u|} \int_0^{|I_r|} \left(1 - \frac{xy}{|\mathcal{P}|}\right)^{n-1} \mathrm{d}y \, \mathrm{d}x + J + O(n^{-K}) \\ &= \log n + \gamma + \log \frac{|I_u|\,|I_r|}{|\mathcal{P}|} + J + O(n^{-1/2}), \end{aligned}$$

where J denotes the contribution of the part in $\mathcal{P}$ outside the rectangle formed by the two segments I_u and I_r. If we are in the case shown in Figure 5(a), then the expected number of maxima contributed from the triangle to the left of I_u (with spanning angle θ_u) is given by

$$\frac{n}{|\mathcal{P}|} \int_0^\varepsilon \int_{x \cot\theta_u}^\varepsilon \left(1 - \frac{1}{|\mathcal{P}|}\left(xy + |I_u| y - \frac{x^2}{2} \cot\theta_u\right)\right)^{n-1} \mathrm{d}y \, \mathrm{d}x + O(n^{-K}) = O(n^{-1}).$$

Similarly, the contribution from other parts of $\mathcal{P}$ and the "overshoot" are all $O(n^{-1})$.

On the other hand, if $|I_u| > 0$ and $|I_r| = 0$ with angle θ_u at the upper-right corner (see Figure 5(b)), then by a similar argument

$$\begin{aligned} E(M_n) &= \frac{n}{|\mathcal{P}|}\int_0^{|I_u|}\int_0^{x\tan\theta_u}\left(1-\frac{1}{|\mathcal{P}|}\left(xy-\frac{y^2}{2}\cot\theta_r\right)\right)^{n-1} \mathrm{d}y\,\mathrm{d}x + O(n^{-1/2}) \\ &= \frac{1}{2}\log n + \frac{\gamma}{2} + \frac{1}{2}\log\frac{2|I_u|^2\tan\theta_u}{|\mathcal{P}|} + O(n^{-1/2}). \end{aligned}$$

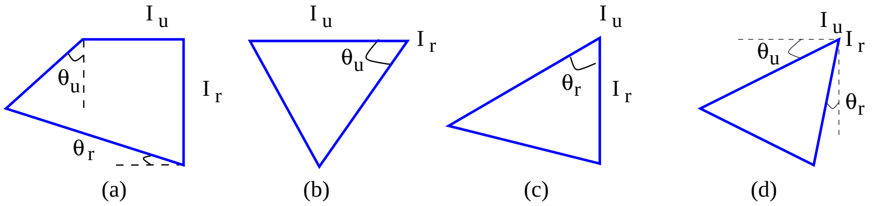

Figure 5: *Illustration of possible cases when* $I_u \cap I_r \neq \emptyset$.

Similarly, if $|I_r| > 0$ and $|I_u| = 0$ with angle θ_r at the upper-right corner (see Figure 5(c)), then

$$E(M_n) = \frac{1}{2}\log n + \frac{\gamma}{2} + \frac{1}{2}\log\frac{2|I_r|^2\tan\theta_r}{|\mathcal{P}|} + O(n^{-1/2}).$$

Finally, if $|I_u| = |I_r| = 0$ with angles θ_u and θ_r indicated as in Figure 5(d), then

$$\begin{aligned} E(M_n) &= \frac{n}{|\mathcal{P}|}\int_0^{\varepsilon}\int_{x\tan\theta_u}^{x\cot\theta_u}\left(1-\frac{1}{|\mathcal{P}|}\left(xy-\frac{\tan\theta_u}{2}x^2-\frac{\tan\theta_r}{2}y^2\right)\right)^{n-1} dy + O(n^{-K}) \\ &= \int_{\tan\theta_u}^{\cot\theta_r}\frac{\mathrm{d}y}{2y-\tan\theta_u-y^2\tan\theta_r} + O(n^{-1/2}) \qquad (3.10) \\ &= \frac{1}{2\sqrt{1+\tan\theta_u\tan\theta_r}}\log\frac{\sqrt{1+\tan\theta_u\tan\theta_r}+1-\tan\theta_u\tan\theta_r}{\sqrt{1+\tan\theta_u\tan\theta_r}-1+\tan\theta_u\tan\theta_r} + O(n^{-1/2}). \qquad (3.11) \end{aligned}$$

Collecting the above results, we have when $I_u \cap I_r \neq \emptyset$

$$\pi_3 = \begin{cases} \gamma + \log\frac{|I_u||I_r|}{|\mathcal{P}|}, & \text{if } |I_u||I_r| > 0; \\ \frac{\gamma}{2} + \frac{1}{2}\log\frac{2|I_u|^2\tan\theta_u}{|\mathcal{P}|}, & \text{if } |I_u| > 0, |I_r| = 0, \\ \frac{\gamma}{2} + \frac{1}{2}\log\frac{2|I_r|^2\tan\theta_r}{|\mathcal{P}|}, & \text{if } |I_u| = 0, |I_r| > 0, \\ \frac{1}{2\sqrt{1+\tan\theta_u\tan\theta_r}}\log\frac{\sqrt{1+\tan\theta_u\tan\theta_r}+1-\tan\theta_u\tan\theta_r}{\sqrt{1+\tan\theta_u\tan\theta_r}-1+\tan\theta_u\tan\theta_r}, & \text{if } |I_u| = |I_r| = 0. \end{cases} \qquad (3.12)$$

This completes the proof of (1.3). □

3.2 Asymptotic normality

We split the upper-right part of $\mathcal{P}$ that lies in $\mathcal{P}$ into ν triangles as shown in Figure 3. Note that all $\mathcal{P}_j$'s except possibly the first and the last are right triangles and that the number of maxima in a right triangle $\mathcal{T}$ is scale invariant. Assume that $\nu \geq 1$. Let $\mathcal{S}_n := \mathcal{P} \setminus \cup_{1\leq j\leq \nu}\mathcal{P}_j$. Then, by our discussions in Section 3.1, $E(M_n(\mathcal{S}_n)) = O(\nu + \log n)$. From this and the inequality

$$\sum_{1\leq j\leq \nu} M_n(\mathcal{P}_j) \leq M_n = \sum_{1\leq j\leq \nu} M_n(\mathcal{P}_j) + M_n(\mathcal{S}_n),$$

it follows that $(M_n - \pi_1 n^{1/2})/(\sigma_1 n^{1/4})$ and $W_n := \sum_{1\leq j\leq \nu}[M_n(\mathcal{P}_j) - (|\mathcal{P}_j|\pi n/|\mathcal{P}|)^{1/2}]/(\sigma_1 n^{1/4})$ have the same limit distribution. Let Φ_n and Φ denote the distribution functions of W_n and $\mathcal{N}(0,1)$, respectively. We show that $|\Phi_n(x) - \Phi(x)| \to 0$ for all $x \in \mathbb{R}$.

To that purpose, assume first that all $\mathcal{P}_j$, $1 \leq j \leq \nu$, are all right triangles. Define Π_n to be the set of all ν-tuples $\rho = (r_1, \ldots, r_\nu)$ of nonnegative integers r_j such that $r_1 + \cdots + r_\nu \leq n$. Let Ω_ρ be the event that there are exactly r_j points lying in $\mathcal{P}_j$, where $\rho \in \Pi_n$. Then $\Phi_n(x) = \sum_{\rho\in\Pi_n} P(\Omega_\rho)\Phi_{n,\rho}(x)$, where $\Phi_{n,\rho}$ is the conditional distribution function of W_n under Ω_ρ. Let $Y_j := M_n(\mathcal{P}_j)$. Observe that, under Ω_ρ, these Y_j's are independent and that the distribution of Y_j under Ω_ρ is the same as that of the number of maxima of r_j iid points taken uniformly at random from $\mathcal{P}_j$. Let $\Psi_{n,\rho}$ be the conditional distribution function (under Ω_ρ) of

$$Z_{n,\rho} := \frac{\sum_{1\leq j\leq \nu}(Y_j - E_\rho(Y_j))}{\left(\sum_{1\leq j\leq \nu} \mathrm{Var}_\rho(Y_j)\right)^{1/2}},$$

where $E_\rho(Y_j) := E(Y_j|\Omega_\rho)$ and $\mathrm{Var}_\rho(Y_j) := \mathrm{Var}(Y_j|\Omega_\rho)$. Define $\mu_j := |\mathcal{P}_j|/|\mathcal{P}|$ and the subset $\Pi_n' \subset \Pi_n$ by

$$\Pi_n' := \left\{\rho = (r_1, \ldots, r_\nu) \in \Pi_n \ : \ |r_j - \mu_j n| \leq (\mu_j n)^{1/2} \log n \text{ for all } j = 1, \ldots, \nu\right\}. \tag{3.13}$$

By (2.3), we have, uniformly for all $\rho \in \Pi_n'$,

$$\sup_{x\in\mathbb{R}} |\Psi_{n,\rho}(x) - \Phi(x)| = o(1).$$

Thus to prove (1.7), it suffices to show that

(i) the probability of the union of Ω_ρ, $\rho \in \Pi_n'$, tends to 1;

(ii) uniformly for all $\rho \in \Pi_n'$, $P(|W_n - Z_{n,\rho}| > \varepsilon \mid \Omega_\rho) \to 0$, for any $\varepsilon > 0$.

¿From these, we obtain, by Slutzky's theorem (suitably modified under our conditioning setting; see [15, p. 254]), that $|\Phi_{n,\rho} - \Phi(x)| \to 0$, uniformly for $\rho \in \Pi_n'$, where $x \in \mathbb{R}$. Thus

$$\begin{aligned} |\Phi_n(x) - \Phi(x)| &\leq P\left(\cup_{\rho\in\Pi_n\setminus\Pi_n'}\Omega_\rho\right) + \sum_{\rho\in\Pi_n'} P(\Omega_\rho)|\Phi_{n,\rho}(x) - \Phi(x)| \\ &\to 0. \end{aligned}$$

The proof of (i) uses $P(X_1 \in \mathcal{P}_j) = \mu_j$ and Bernstein's inequality for binomial distributions (see [15, p. 111]), giving

$$\begin{aligned} P\left(\cup_{\rho\in\Pi_n\setminus\Pi_n'}\Omega_\rho\right) &\leq \sum_{1\leq j\leq \nu} P\left(|Y_j - \mu_j n| > (\mu_j n)^{1/2}\log n\right) \\ &\leq 2\mu_j^{-1/2} \exp\left(-\frac{\log^2 n}{2(1-\mu_j) + 2(\mu_j n)^{-1/2}\log n}\right) \\ &\to 0. \end{aligned}$$

By (3.3), (2.2), and (3.13)), we deduce that

$$\sum_{1\leq j\leq \nu} E_\rho(Y_j) - \pi_1 n^{1/2} = o(n^{1/4}), \qquad \sum_{1\leq j\leq \nu} \mathrm{Var}_\rho(Y_j) - \sigma_1^2 n^{1/2} = o(n^{1/4}),$$

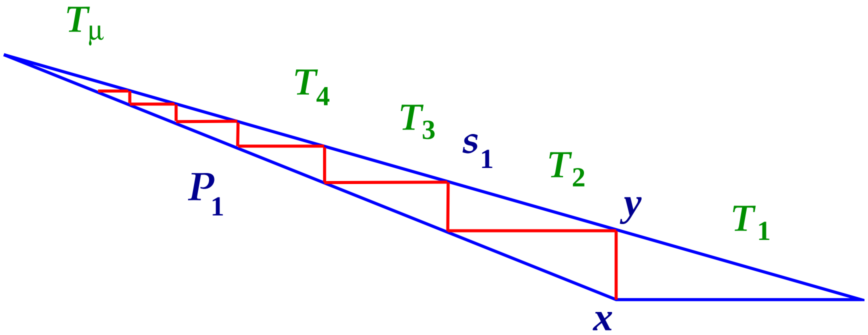

Figure 6: *"Paper-folding" triangulation of* $\mathcal{P}_1$.

for all $\rho \in \Pi_n'$. Thus (ii) follows.

This proves the asymptotic normality of $M_n(\mathcal{P})$ when both $\mathcal{P}_1$ and $\mathcal{P}_\nu$ are right triangles.

Now assume that $\mathcal{P}_1$ is not a right triangle (it is then an obtuse triangle). We split $\mathcal{P}_1$ into 2μ right triangles and an obtuse triangle in the following "paper-folding" way (see Figure 6). Denote by s_1 the hypotenuse of $\mathcal{P}_1$ and call the point opposite to the hypotenuse x. Make a right triangle by connecting a vertical line from x to s_1. Assume that the horizontal line intersects s_1 at y. Draw a vertical line from y to the opposite. This results in two right triangles and an obtuse triangle. Call the right triangle sharing the same hypotenuse with s_1 $\mathcal{T}_1$. Repeating $\mu - 1$ times the same construction in the obtuse triangle yields μ right triangles $\mathcal{T}_i$ along s_1. We take $\mu = \lfloor c \log n \rfloor$, where c is properly chosen so that the expected number of points lying in the obtuse triangle is $\leq n^{1/5}$ and that the expected number of points in each $\mathcal{T}_i$ is $\geq n^{1/5}$. This is achievable since the area of the obtuse triangle contracts exponentially. We then argue similarly as above, noting that the main contribution to $M_n(\mathcal{P}_1)$ comes from $\mathcal{T}_i$. The asymptotic normality of $M_n(\mathcal{P}_1)$ follows as above. The case when $\mathcal{P}_\nu$ is not a right triangle is similar. This completes the proof of (1.7). $\square$

3.3 Variance

Let $M_n^* := (M_n - \pi_1 n^{1/2})/(\sigma_1 n^{1/4})$. For the proof of Theorem 1 when $\pi_1 > 0$, it remains to show (1.6). But this is an easy consequence of the asymptotic normality of M_n and (2.9) with $m = 2$. For,

$$E(M_n - \pi_1 n^{1/2})^4 = O\left(\sum_{1\leq j\leq \nu} E\left(M_n(\mathcal{P}_j) - (|\mathcal{P}_j|\pi n/|\mathcal{P}|)^{1/2}\right)^4\right) = O_\nu(n),$$

implying that $\sup_n E(M_n^{*4}) = O(1)$ and consequently, M_n^{*2} is uniformly integrable. The result (1.6) now follows from dominated convergence theorem and (1.7). $\square$

4 Maxima in regions bounded above by a nonincreasing curve

We consider in this section $\mathcal{C}$ of the form (1.1)

$$\mathcal{C} = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq f(x)\},$$

where $f(x)$ is a nonincreasing function in the unit interval. We may assume for convenience that

$$\int_0^1 f(x)\,\mathrm{d}x = 1.$$

Let $C^2(a,b)$ denote the set of real functions whose second derivatives are continuous in (a,b), where $a<b$. For convenience we also define

$$C_2(a,b) = \{f \in C^2(a,b) \,:\, 0 < |f'(x)| < \infty,\, a < x < b\}.$$

The classification of critical points into $|f'(x)| = 0$, $|f'(x)| = \infty$ or $f'_-(x) \neq f'_+(x)$ leads the study of the expected number of maxima to the following three prototypical cases of decreasing f (see Figure 7):

(i) $f \in C_2(0,1)$;

(ii) $f(x) = \eta > 0$ for $0 \le x \le \xi < 1$ and $f \in C_2(\xi,1)$; or its symmetric (with respect to the line $y = x$) counterpart;

(iii) There is a unique critical point $\xi \in (0,1)$ at which f is continuous (to exclude jumps), $f \in C_2(0,\xi)$ and $f \in C_2(\xi,1)$.

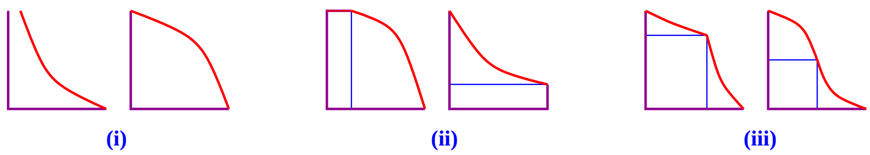

Figure 7: *The three basic prototypes.*

It is obvious that any piecewise $C^2(0,1)$ decreasing functions f can be finitely decomposed into the above prototypes and rectangles (see Figure 8). Our results provide a "transparent" connection between the coefficient of the logarithmic term and the local behavior of f near each critical point. This connection, already observed by Devroye (1993), is made quantitatively more precise here. Note that f may have critical points at the boundary (0 and 1) in the first case. Also if $\xi = 1$ in the second case, then $\mathcal{C}$ is the unit square and the expected number of maxima is asymptotic to $\log n + O(1)$.

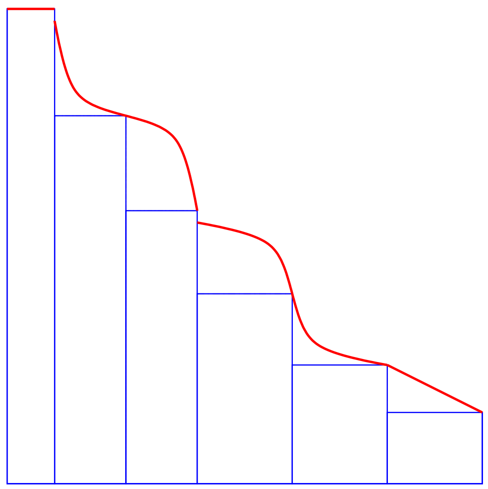

Figure 8: *Decomposition of a general decreasing piecewise C^2-curve at critical points.*

Theorem 4. *Assume that* $f : [0,1] \mapsto [0,\infty)$ *is nonincreasing,* $f(1) = 0$ *and* $\int_0^1 f(x)\,\mathrm{d}x = 1$.
(i) If $f \in C_2(0,1)$ *and*

$$
\begin{aligned}
f'(x) &= -x^{\alpha}(a + o(1)), \qquad (4.1)\\
f'(1-x) &= -x^{\beta}(b + o(1)), \qquad (4.2)
\end{aligned}
$$

as $x \to 0^+$, *where* $a, b > 0$, $\alpha > -2$ *and* $\beta > -1$ *(since* $f(x) \geq 0$*), then*

$$
E(M_n) = \pi_0 n^{1/2} + \frac{1}{3}\left(\frac{\alpha}{\alpha+2} - \frac{\beta}{\beta+2}\right)\log n + o(\log n).
$$

(ii) If $f(x) = \eta > 0$ *for* $0 \leq x \leq \xi < 1$, $f \in C_2(\xi, 1)$ *and*

$$
\begin{aligned}
f'(\xi + x) &= -x^{\alpha}(a + o(1)), \qquad (4.3)\\
f'(1-x) &= -x^{\beta}(b + o(1)),
\end{aligned}
$$

as $x \to 0^+$, *where* $a, b > 0$ *and* $\alpha, \beta > -1$, *then*

$$
E(M_n) = \pi_0 n^{1/2} + \frac{1}{3}\left(\frac{\alpha+3}{\alpha+2} - \frac{\beta}{\beta+2}\right)\log n + o(\log n).
$$

The symmetric version with respect to the line $y = x$ *has the same asymptotic behavior.*
(iii) Assume that ξ *is the unique critical point in* $(0,1)$ *at which* f *is continuous. If* $f \in C_2(0,\xi)$, $f \in C_2(\xi, 1)$, f *satisfies (4.1) and (4.2), and*

$$
\begin{aligned}
f'(\xi - x) &= -x^{\beta'}(b' + o(1)),\\
f'(\xi + x) &= -x^{\alpha'}(a' + o(1)),
\end{aligned}
$$

as $x \to 0^+$, *where* $a', b' > 0$ *and* $\alpha', \beta' > -1$, *then*

$$
E(M_n) = \pi_0 n^{1/2} + \frac{1}{3}\left(\frac{\alpha}{\alpha+2} + \frac{\alpha'}{\alpha'+2} - \frac{\beta}{\beta+2} - \frac{\beta'}{\beta'+2}\right)\log n + o(\log n).
$$

Thus the expected number of maxima, $E(\mathcal{R}_n)$, contributed from the rectangle, satisfies

$$
E(\mathcal{R}_n) = \frac{\log n}{\alpha+2}(1 + o(1)),
$$

in case (ii) and $E(\mathcal{R}_n) = O(1)$ in the last case (see Figure 7). Under stronger conditions, our method of proof can be further refined to make explicit the o-terms; see Bai et al. (2000).

Proof of Theorem 4. (Sketch) The starting point is the integral representation

$$
\begin{aligned}
E(M_n) &= n\int_0^1\int_0^{f(x)}\left(1 - \int_x^{f^{-1}(y)}(f(t) - y)\,\mathrm{d}t\right)^{n-1}\mathrm{d}y\,\mathrm{d}x\\
&= n\int_0^1 |f'(w)|\int_0^w (1 - A(x,w))^{n-1}\,\mathrm{d}x\,\mathrm{d}w,
\end{aligned}
$$

where $|f'(w)| = -f'(w)$ for $0 < w < 1$ and $A(x,w) = \int_{w-x}^{w}(f(t) - f(w))\,\mathrm{d}t$. Take two small quantities $\delta_0 = n^{-1/(\alpha+2)}(\log n)^{2/(\alpha+2)}$ and $\delta_1 = n^{-1/(\beta+2)}(\log n)^{2/(\beta+2)}$ and split the integral into three parts:

$$
\begin{aligned}
E(M_n) &= n\left(\int_0^{\delta_0} + \int_{\delta_0}^{1-\delta_1} + \int_{1-\delta_1}^{1}\right)|f'(w)|\int_0^w (1 - A(x,w))^{n-1}\,\mathrm{d}x\,\mathrm{d}w\\
&=: I_1 + I_2 + I_3.
\end{aligned}
$$

Then prove that

$$\begin{aligned} I_1 &= \frac{(2\pi a)^{1/2}}{\alpha+2} n^{1/2} \delta_0^{(\alpha+2)/2} + \frac{\alpha}{3(\alpha+2)} \log n + \frac{\alpha}{3} \log \delta_0 + o(\log n), \\ I_3 &= \frac{(2\pi b)^{1/2}}{\beta+2} n^{1/2} \delta_1^{(\beta+2)/2} - \frac{\beta}{3(\beta+2)} \log n - \frac{\beta}{3} \log \delta_1 + o(\log n), \end{aligned}$$

and that the main contribution to $E(M_n)$ comes from I_2:

$$I_2 = \pi_0 n^{1/2} - \frac{(2\pi a)^{1/2}}{\alpha+2} n^{1/2} \delta_0^{(\alpha+2)/2} - \frac{(2\pi b)^{1/2}}{\beta+2} n^{1/2} \delta_1^{(\beta+2)/2} + \frac{\beta}{3} \log \delta_1 - \frac{\alpha}{3} \log \delta_0 + o(\log n).$$

For details, see Bai et al. (2000). □

Case (ii)**.** (Sketch) It suffices to consider the contribution from the rectangle $\mathcal{R}_n$ (see Figure 7) since the other part is essentially covered by case (i). By (4.3),

$$\begin{aligned} E(\mathcal{R}_n) &= n \int_0^\xi \int_0^\eta \left(1 - (\xi - x)(\eta - y) - \int_\xi^{f^{-1}(y)} (f(t) - y)\, \mathrm{d}t\right)^{n-1} \mathrm{d}y\, \mathrm{d}x \\ &\sim (\alpha+1) \int_0^{1-\xi} w^{-1} \left(e^{-anw^{\alpha+1}/(\alpha+2)} - e^{-an\xi w^{\alpha+1}/(\alpha+1)}\right) \mathrm{d}w \\ &= \frac{\log n}{\alpha+2} + O(1). \end{aligned}$$

The symmetric version of $E(M_n)$ with respect to the line $y = x$ is similar. □

5 Maxima in regions bounded by two increasing curves

The results in the preceding section give an interpretation of the "magic constant" 1/2 in the expansion (1.3). We give another different "structural interpretation" in this section by considering the region

$$\mathcal{C} = \{(x, y) : 0 \le x \le 1, \min\{f(x), g(x)\} \le y \le \max\{f(x), g(x)\}\},$$

where $f(x)$ and $g(x)$ are two nondecreasing function in the unit interval. Assume that

$$\int_0^1 |f(x) - g(x)|\, \mathrm{d}x = 1 < \infty,$$

and that

$$\left\{ \begin{array}{rcl} f(1-x) &=& f(1) - x^\alpha (a + o(1)), \\ g(1-x) &=& f(1) - x^\beta (b + o(1)), \end{array} \right. \tag{5.1}$$

as $x \to 0^+$, where $a, b, \alpha, \beta > 0$. We may assume that $f(x) \ge g(x)$ in the vicinity of $x = 1$, namely, $\alpha > \beta$ if $\alpha \ne \beta$ and $a < b$ if $\alpha = \beta$.

Theorem 5. *If* $\alpha > \beta \ge 0$ *then*

$$E(M_n) = \left(\frac{1}{\beta+1} - \frac{1}{\alpha+1}\right)(\log n + \gamma) + c_3 + o(1), \tag{5.2}$$

where

$$\begin{aligned} c_3 \quad := \quad & \frac{\log b}{\beta+1} - \frac{\log a}{\alpha+1} + \frac{\beta}{\alpha(\beta+1)} + \frac{\log(\alpha+1)}{\alpha+1} + \frac{\beta(\alpha+1)}{\alpha(\beta+1)}\log(\beta+1) \\ & - \frac{\alpha^2+2\alpha-\beta}{\alpha(\alpha+1)(\beta+1)}\log\alpha + \frac{\alpha-\beta}{\alpha(\beta+1)}\log(\alpha-\beta); \end{aligned}$$

on the other hand, if $\alpha = \beta$ *then*

$$E(M_n) = \frac{1}{\alpha+1}\int_a^b \frac{\mathrm{d}y}{y - \frac{a}{\alpha+1} - \frac{\alpha y^{1+1/\alpha}}{b^{1/\alpha}(\alpha+1)}} + o(1).$$

It is easily shown that

$$\frac{1}{\alpha+1}\int_a^b \frac{\mathrm{d}y}{y - \frac{a}{\alpha+1} - \frac{\alpha y^{1+1/\alpha}}{b^{1/\alpha}(\alpha+1)}} \geq 1,$$

meaning that there is at least one maximal point. Note that the expected number of maxima remains bounded for functions such as $f(x) = \frac{101}{99}x^{1/100}$ and $g(x) = \frac{101}{99}x^{100}$.

Proof of Theorem 5. (Sketch) Take $\delta = \delta_n = n^{-1/(\alpha+1)}(\log n)^{2/(\alpha+1)}$. Then $E(M_n) = I_1 + I_2$, where

$$\begin{aligned} I_1 &= n\int_{1-\delta}^1 \int_{g(x)}^{f(x)} (1 - A(1-x, 1-y))^{n-1}\,\mathrm{d}y\,\mathrm{d}x, \\ I_2 &= n\int_0^{1-\delta} \int_{\min\{f(x),g(x)\}}^{\max\{f(x),g(x)\}} (1 - A(1-x, 1-y))^{n-1}\,\mathrm{d}y\,\mathrm{d}x. \end{aligned}$$

Here $A(1-x, 1-y)$ denotes the area of the region

$$\{(u,v) \,:\, x \leq u \leq 1;\, y \leq v \leq f(1)\,, \min\{f(x), g(x)\} \leq v \leq \max\{f(x), g(x)\}\}\,.$$

By (5.1),

$$A(x,y) = xy - \frac{a+o(1)}{\alpha+1}\,x^{\alpha+1} - \frac{\beta+o(1)}{(\beta+1)b^{1/\beta}}\,y^{1+1/\beta}, \tag{5.3}$$

uniformly for $0 \leq x \leq \delta$ and $\phi(x) \leq y \leq \gamma(x)$.
Since for fixed x the function $A(1-x, 1-y)$ is a nondecreasing function of y,

$$I_2 = O\left(n \max_{\phi(\delta)\leq y\leq\gamma(\delta)} e^{-nA(\delta,y)}\right) = O\left(ne^{-c_4(\log n)^2}\right) = o(1).$$

If $\alpha = \beta$, then by (5.3),

$$\begin{aligned} I_1 &\sim n\int_0^\delta\int_a^b x^\alpha e^{-nA(x,x^\alpha y)}\,\mathrm{d}y\,\mathrm{d}x \\ &\sim n\int_a^b\int_0^\infty x^\alpha \exp\left(-nx^{\alpha+1}\left(y - \frac{a}{\alpha+1} - \frac{\alpha y^{1+1/\alpha}}{(\alpha+1)b^{1/\alpha}}\right)\right)\,\mathrm{d}x\,\mathrm{d}y \\ &\sim \frac{1}{\alpha+1}\int_a^b \frac{\mathrm{d}y}{y - \frac{a}{\alpha+1} - \frac{\alpha y^{1+1/\alpha}}{b^{1/\alpha}(\alpha+1)}}. \end{aligned}$$

On the other hand, if $\alpha > \beta$, then

$$
\begin{aligned}
I_1 &= n \int_0^\delta \int_a^{bx^{\beta-\alpha}} e^{-nA(x,x^\alpha y)}\, \mathrm{d}y\, \mathrm{d}x + o(1) \\
&= n \int_0^{(b/a)^{1/(\alpha-\beta)}} \int_a^{bx^{\beta-\alpha}} x^\alpha \exp\left(-nx^{\alpha+1}\left(y - \frac{a}{\alpha+1} + \frac{\beta b^{-1/\beta}}{\beta+1} x^{-1+\alpha/\beta} y^{1+1/\beta}\right)\right) \mathrm{d}y\, \mathrm{d}x + o(1),
\end{aligned}
$$

¿From this integral representation, the expansion (5.2) is obtained by using Mellin transform techniques. See Bai et al. (2000) for details. □

Remark. If $\beta = 0$ then $E(M_n) \sim \alpha(\log n)/(\alpha+1)$; if, furthermore, f is linear, then $\alpha = 1$ which implies that $E(M_n) \sim (1/2)\log n$. This essentially corresponds to the case $|I_u| = 0$ and $|I_r| > 0$ in Section 3.1. The case when $\alpha = \infty$ and g is linear is similar. When both f and g are linear functions, we have $\alpha = \beta = 1$, and thus

$$
E(M_n(\mathcal{C})) = \int_a^b \frac{\mathrm{d}y}{2y - a - y^2/b} + o(1) = \frac{\sqrt{b}}{2\sqrt{a+b}} \log \frac{\sqrt{b(b+a)} + b - a}{\sqrt{b(b+a)} - b + a} + o(1);
$$

which is essentially the case when $|I_u| = |I_r| = 0$; see (3.11).

6 Maxima in regions bounded above by an increasing curve

We prove Theorem 2 on Poisson approximations in this section.

6.1 Probability generating function

Let

$$
S_m(u) = \frac{u(u+1)\cdots(u+m-1)}{m!} \qquad (m \geq 1)
$$

denote the probability generating function of the number of maxima when the underlying region $\mathcal{C}$ is a rectangle.

Proposition 2. *Let $F(x) = \int_0^x f(t)\,\mathrm{d}t$. Then for any nondecreasing function f*

$$
E(u^{M_n}) = 1 + (u-1) \sum_{1 \leq j \leq n} S_{j-1}(u) \binom{n}{j} \int_0^1 F(x)^{n-j} f(x)^j (1-x)^{j-1}\, \mathrm{d}x \qquad (n \geq 1). \tag{6.1}
$$

Proof. We use a method similar to the proof of (2.4). Let $\mathcal{V} = \{(x_j, y_j) : 1 \leq j \leq n\}$ be n iid points chosen uniformly in $\mathcal{C}$.

We first locate the point $p = (x, y)$ for which $y = \max_{1 \leq j \leq n} y_j$. Then the number of maxima equals one plus those in the (shaded) rectangle (see Figure 9). The probability that there are exactly k points in $\mathcal{V}$ lying in the rectangle formed by (x, y) and $(1, 0)$ is equal to

$$
p_{n,k}(x,y) := \binom{n-1}{k} \left(\int_0^{f^{-1}(y)} f(t)\,\mathrm{d}t + y\left(x - f^{-1}(y)\right)\right)^{n-1-k} (y(1-x))^k .
$$

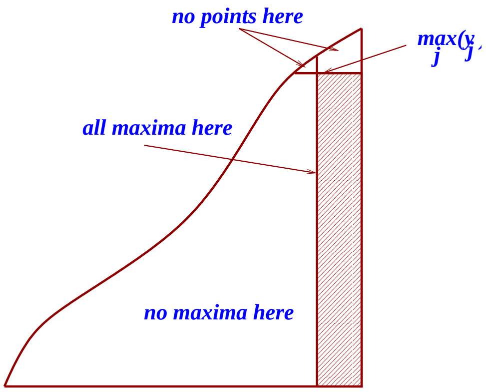

Figure 9: *Division of* $\mathcal{C}$ *at the point with* $\max_{1\le i\le n} y_i$*; all other maxima are inside (and on) the rectangle.*

Thus we have

$$\begin{aligned} E(u^{M_n}) &= nu \sum_{0\le k\le n-1} S_k(u) \int_{\mathcal{C}} p_{n,k}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= u \sum_{0\le k\le n-1} p_{n,k} S_k(u), \end{aligned}$$

where

$$\begin{aligned} p_{n,k} &= \frac{n!}{k!(n-1-k)!} \int_0^1 \int_0^{f(x)} \left(\int_0^{f^{-1}(y)} f(t)\,\mathrm{d}t + y(x - f^{-1}(y)) \right)^{n-1-k} (1-x)^k y^k \,\mathrm{d}y\,\mathrm{d}x \\ &= \frac{n!}{k!(n-1-k)!} \int_0^1 \int_0^x \left(F(v) + f(v)(x-v)\right)^{n-1-k} (1-x)^k f(v)^k f'(v)\,\mathrm{d}v\,\mathrm{d}x, \end{aligned}$$

by the change of variables $v = f^{-1}(y)$. Interchanging the order of integration and making the change of variables $x \mapsto v + (1-v)x$, we obtain

$$p_{n,k} = \frac{n!}{k!(n-1-k)!} \int_0^1 \int_0^1 \left(F(v) + (1-v)f(v)x\right)^{n-1-k} (1-x)^k (1-v)^{k+1} f(v)^k f'(v)\,\mathrm{d}x\,\mathrm{d}v.$$

Expanding the factor $(F(v) + (1-v)f(v)x)^{n-1-k}$ by binomial theorem and then evaluating the inner beta integral, we have

$$\begin{aligned} p_{n,k} &= \sum_{0\le j\le n-1-k} \frac{n!}{(n-1-k-j)!(j+k+1)!} \int_0^1 F(v)^{n-1-k-j} f(v)^{k+j} (1-v)^{k+j+1} f'(v)\,\mathrm{d}v \\ &= \sum_{k+1\le j\le n} \binom{n}{j} \int_0^1 F(v)^{n-j} f(v)^{j-1} (1-v)^j f'(v)\,\mathrm{d}v. \end{aligned}$$

Thus

$$E(u^{M_n}) = u \sum_{1\le j\le n} \binom{n}{j} \left(\int_0^1 F(v)^{n-j} f(v)^{j-1} (1-v)^j f'(v)\,\mathrm{d}v \right) \sum_{0\le k\le j-1} S_k(u).$$

Now

$$u \sum_{0\le k\le j-1} S_k(u) = \frac{u}{2\pi i} \oint z^{-j}(1-z)^{-u-1}\,\mathrm{d}z = \frac{1}{2\pi i} \oint z^{-j}\,\mathrm{d}(1-z)^{-u} = jS_j(u),$$

by an integration by parts. It follows that

$$E(u^{M_n}) = \sum_{1\le j\le n} jS_j(u)\binom{n}{j}\int_0^1 F(v)^{n-j}f(v)^{j-1}(1-v)^j f'(v)\,\mathrm{d}v.$$

We can further remove the factor $f'(v)$ by an integration by parts:

$$\begin{aligned}\int_0^1 F(v)^{n-j}f(v)^{j-1}(1-v)^j f'(v)\,\mathrm{d}v &= \int_0^1 F(v)^{n-j}f(v)^j(1-v)^{j-1}\,\mathrm{d}v\\ &\quad -\frac{n-j}{j}\int_0^1 F(v)^{n-j-1}f(v)^{j+1}(1-v)^j\,\mathrm{d}v.\end{aligned}$$

Thus

$$\begin{aligned}E(u^{M_n}) &= \sum_{1\le j\le n} jS_j(u)\binom{n}{j}\int_0^1 F(v)^{n-j}f(v)^j(1-v)^{j-1}\,\mathrm{d}v\\ &\quad -\sum_{2\le j\le n} S_{j-1}(u)\binom{n}{j-1}(n+1-j)\int_0^1 F(v)^{n-j}f(v)^j(1-v)^{j-1}\,\mathrm{d}v\\ &= n\int_0^1 F(v)^{n-1}f(v)\,\mathrm{d}v + \sum_{1\le j\le n}(jS_j(u)-jS_{j-1}(u))\binom{n}{j}\int_0^1 F(v)^{n-j}f(v)^j(1-v)^{j-1}\,\mathrm{d}v\\ &= 1+(u-1)\sum_{1\le j\le n} S_{j-1}(u)\binom{n}{j}\int_0^1 F(v)^{n-j}f(v)^j(1-v)^{j-1}\,\mathrm{d}v,\end{aligned}$$

since $jS_j(u)-jS_{j-1}(u)=(u-1)S_{j-1}(u)$. This completes the proof. □

Corollary 1. *For any increasing function f,*

$$E(M_n)\le H_n.$$

Proof. This follows from (6.1) and the inequality $(1-x)f(x)\le 1-F(x)$:

$$\begin{aligned}E(M_n) &\le \sum_{1\le j\le n}\binom{n}{j}\int_0^1 F(x)^{n-j}(1-F(x))^{j-1}\,\mathrm{d}F(x)\\ &\le \sum_{1\le j\le n}\binom{n}{j}\int_0^1 y^{n-j}(1-y)^{j-1}\,\mathrm{d}y\\ &= \sum_{1\le j\le n} j^{-1}.\end{aligned}$$

□

6.2 Asymptotics I. Individual terms

Proposition 3. *Under the assumptions of Theorem 2, let $j_0 = n^{\alpha/(\alpha+1)}(\log n)^{2/\alpha}$. If $j\le j_0$ then*

$$\begin{aligned}&\binom{n}{j}\int_0^1 F(x)^{n-j}f(x)^j(1-x)^{j-1}\,\mathrm{d}x\\ &\qquad = j^{-1}\exp\left(-\left(a+O\left(L(j/n)\right)\right)\frac{j^{\alpha+1}}{n^\alpha}\right)\left(1+O\left(\frac{j^\alpha}{n^\alpha}+\frac{j^{2\alpha+1}}{n^{2\alpha}}\right)\right);\end{aligned}\tag{6.2}$$

if $j_0\le j\le n$ then

$$\binom{n}{j}\int_0^1 F(x)^{n-j}f(x)^j(1-x)^{j-1}\,\mathrm{d}x = O\left(j_0^{-1}e^{-(a/2)(\log n)^2}\right).\tag{6.3}$$

Proof. For simplicity, we assume that $L(x) \equiv 0$; the method of proof is easily amended to the desired case. Thus we have

$$\frac{(1-x)f(x)}{1-F(x)} = 1 - a(1-F(x))^\alpha,$$

for $1 - x \leq \varepsilon$, where ε is a small but fixed positive number. First split the integral into two parts:

$$\begin{aligned}\binom{n}{j}\int_0^1 F(x)^{n-j} f(x)^j (1-x)^{j-1}\,\mathrm{d}x &= \binom{n}{j}\left(\int_0^{1-\delta} + \int_{1-\delta}^1\right) F(x)^{n-j} f(x)^j (1-x)^{j-1}\,\mathrm{d}x \\ &=: I_1 + I_2,\end{aligned}$$

where $0 < \delta \leq \varepsilon$. Take $1 - F(1-\delta) = j_0/n = n^{-1/(\alpha+1)}(\log n)^{2/\alpha}$. When $0 \leq x \leq 1-\delta$, we have the inequality

$$(1-x)f(x) \leq \int_x^{1-\delta} f(t)\,\mathrm{d}t + \int_{1-\delta}^1 f(1-\delta)\,\mathrm{d}t = F(1-\delta) + \delta f(1-\delta) - F(x).$$

This yields

$$\begin{aligned}I_1 &\leq \binom{n}{j}\int_0^{F(1-\delta)} y^{n-j}\left(F(1-\delta) + \delta f(1-\delta) - y\right)^{j-1}\,\mathrm{d}y \\ &\leq \binom{n}{j}\int_0^{F(1-\delta)+\delta f(1-\delta)} y^{n-j}\left(F(1-\delta) + \delta f(1-\delta) - y\right)^{j-1}\,\mathrm{d}y \\ &= j^{-1}\left(F(1-\delta) + \delta f(1-\delta)\right)^n \\ &= j^{-1}\left(1 - a(1-F(1-\delta))^\alpha\right)^n \\ &\leq j^{-1} e^{-a(\log n)^{2(\alpha+1)/\alpha}} \\ &\leq j^{-1} e^{-a(\log n)^2}. \end{aligned} \tag{6.4}$$

Estimation of I_2 for large j: $j \geq j_0$. The integral I_2 is equal to

$$I_2 = \binom{n}{j}\int_0^{j_0/n} (1-y)^{n-j} y^{j-1} (1-ay^\alpha)^{j-1}\,\mathrm{d}y.$$

Observe that for $j \geq j_0$, $j/n \geq j_0/n \geq y$. This implies that

$$\frac{\binom{n}{j}(1-y)^{n-j}y^{j-1}(1-ay^\alpha)^{j-1}}{\binom{n}{j+1}(1-y)^{n-j-1}y^j(1-ay^\alpha)^j} = \frac{(j+1)(1-y)}{(n-j)y(1-ay^\alpha)} > 1,$$

for $0 < y \leq j_0/n$. Thus

$$\begin{aligned}I_2 &= I_2(j) \leq I_2(\lfloor j_0 \rfloor) \\ &\leq e^{-(a/2)(\log n)^2}\binom{n}{\lfloor j_0 \rfloor}\int_0^1 (1-y)^{n-\lfloor j_0 \rfloor} y^{\lfloor j_0 \rfloor - 1}\,\mathrm{d}y \\ &= \lfloor j_0 \rfloor^{-1} e^{-(a/2)(\log n)^2} \qquad (j \geq j_0).\end{aligned}$$

This together with (6.4) proves (6.3).

Evaluation of I_2 for small j: $j \leq j_0$. In this case, we have

$$I_2 = \binom{n}{j}\int_0^1 (1-y)^{n-j} y^{j-1} (1-ay^\alpha)^{j-1}\,\mathrm{d}y + I_3,$$

where

$$
\begin{aligned}
I_3 &= -\binom{n}{j}\int_{j_0/n}^{1}(1-y)^{n-j}y^{j-1}(1-ay^\alpha)^{j-1}\,\mathrm{d}y\\
&= O\left(e^{-ajn^{-\alpha/(\alpha+1)}(\log n)^2}\binom{n}{j}\int_0^1(1-y)^{n-j}y^{j-1}\,\mathrm{d}y\right)\\
&= O\left(j^{-1}e^{-a(\log n)^2}\right),
\end{aligned}
$$

for $j\ge j_1 := \lfloor n^{\alpha/(\alpha+1)}\rfloor$. On the other hand, if $j\le j_1-1$, then

$$
y\ge \frac{j_0}{n}\gg\frac{j_1}{n}\ge\frac{j+1}{n}>\frac{j+1}{n+1},
$$

giving

$$
\frac{\binom{n}{j}(1-y)^{n-j}y^{j-1}(1-ay^\alpha)^{j-1}}{\binom{n}{j+1}(1-y)^{n-j-1}y^{j}(1-ay^\alpha)^{j}}=\frac{(j+1)(1-y)}{(n-j)y(1-ay^\alpha)}<1,
$$

for large enough n. Consequently,

$$
I_3=I_3(j)\le I_3(j_1-1)=O\left(j_1^{-1}e^{-a(\log n)^2}\right)=O\left(j^{-1}e^{-a(\log n)^2}\right)\qquad(1\le j\le j_1-1).
$$

Therefore,

$$
I_2=\binom{n}{j}\int_0^1(1-y)^{n-j}y^{j-1}(1-ay^\alpha)^{j-1}\,\mathrm{d}y+O\left(j^{-1}e^{-a(\log n)^2}\right),
$$

uniformly for $1\le j\le j_0$. It remains to evaluate the integral on the right-hand side.
Define $j_2=\lfloor n^{\alpha/(\alpha+2)}\rfloor$. We further divide the interval $1\le j\le j_0$ into two parts: $1\le j<j_2$ and $j_2\le j\le j_0$. The first case is easy since $j=o(n^{\alpha/(\alpha+1)})$ in this range:

$$
\begin{aligned}
I_2 &= \binom{n}{j}\int_0^1(1-y)^{n-j}y^{j-1}\left(1-a(j-1)y^\alpha+O\left(j^2y^{2\alpha}\right)\right)\,\mathrm{d}y+O\left(j^{-1}e^{-a(\log n)^2}\right)\\
&= j^{-1}-\frac{a(j-1)n!\Gamma(j+\alpha)}{j!\Gamma(n+1+\alpha)}+O\left(j^{2\alpha+1}n^{-2\alpha}+j^{-1}e^{-a(\log n)^2}\right)\\
&= j^{-1}e^{-aj^{\alpha+1}/n^\alpha}\left(1+O\left(j^\alpha n^{-\alpha}+j^{2\alpha+1}n^{-2\alpha}\right)\right).
\end{aligned}
$$

This and (6.4) proves (6.2) for $j\le j_2$.
For the remaining case $j_2\le j\le j_0$, we take $r=j/(n+1)$ and $r_0=j^{1/2}n^{-1}\log n$ and decompose the integral into three parts:

$$
\begin{aligned}
I_2 &= \binom{n}{j}\left(\int_0^{r-r_0}+\int_{r-r_0}^{r+r_0}+\int_{r+r_0}^1\right)(1-y)^{n-j}y^{j-1}(1-ay^\alpha)^{j-1}\,\mathrm{d}y+O\left(j^{-1}e^{-a(\log n)^2}\right)\\
&=: I_4+I_5+I_6+O\left(j^{-1}e^{-a(\log n)^2}\right).
\end{aligned}
$$

The integral I_4 is majorized as follows.

$$\begin{aligned}
I_4 &\le \binom{n}{j}\int_0^{j/n-r_0}(1-y)^{n-j}y^{j-1}(1-ay^\alpha)^{j-1}\,\mathrm{d}y\\
&= O\left(\binom{n}{j}\left(1-\frac{j}{n}-r_0\right)^{n-j}\left(\frac{j}{n}-r_0\right)^{j-1}(1-a(j/n-r_0)^\alpha)^{j-1}\right)\\
&= O\left(\binom{n}{j}\left(1-\frac{j}{n}\right)^{n-j}\left(\frac{j}{n}\right)^{j-1}\right.\\
&\qquad\left.\times\ \exp\left(-a\frac{j^{\alpha+1}}{n^\alpha}+(n-j)\log\left(1+\frac{r_0}{1-j/n}\right)+j\log\left(1-\frac{r_0}{j/n}\right)\right)\right)\\
&= O\left(j^{-1}e^{-aj^{\alpha+1}/n^\alpha}T_1\right),
\end{aligned}$$

where, by Stirling's formula,

$$\begin{aligned}
T_1 &= n^{3/2}j^{-1/2}(n-j)^{-1/2}\exp\left((n-j)\log\left(1+\frac{r_0}{1-j/n}\right)+j\log\left(1-\frac{r_0}{j/n}\right)\right)\\
&= O\left(n^{3/2}j^{-1/2}(n-j)^{-1/2}\exp\left(-(n-j)\frac{r_0}{1-j/n}+\frac{jr_0}{j/n}-\frac{r_0^2}{2}\left(\frac{n-j}{(1-j/n)^2}+\frac{j}{(j/n)^2}\right)\right)\right)\\
&= O\left(n^{3/2}j^{-1/2}(n-j)^{-1/2}e^{-(\log n)^2/2}\right)\\
&= O\left(e^{-(\log n)^2/3}\right).
\end{aligned}$$

Thus

$$I_4 = O\left(j^{-1}e^{-aj^{\alpha+1}/n^\alpha-(\log n)^2/3}\right)\qquad (j_2\le j\le j_0).$$

Similarly,

$$I_6 = O\left(j^{-1}e^{-aj^{\alpha+1}/n^\alpha-(\log n)^2/3}\right)\qquad (j_2\le j\le j_0).$$

Let $h(y)=(1-ay^\alpha)^{j-1}$. Then expanding h at $y=r$,

$$I_5=\binom{n}{j}\int_{r-r_0}^{r+r_0}(1-y)^{n-j}y^{j-1}\left(h(r)+h'(r)(y-r)+T_2(y,r)\right)\,\mathrm{d}y,$$

where

$$T_2(y,r)=(y-r)^2\int_0^1(1-t)h''(r+t(y-r))\,\mathrm{d}t=\frac{(y-r)^2}{2}h''(r+\theta(y-r))\qquad (0\le\theta\le 1).$$

Now

$$h''(y)=\left(a^2\alpha^2(j-1)(j-2)y^{2\alpha-2}-a\alpha(\alpha-1)(j-1)y^{\alpha-2}\right)(1-ay^\alpha)^{j-3}.$$

Also for $|y-r|\le r_0$, we have

$$\begin{aligned}
(1-a\left(r+\theta(y-r)\right)^\alpha)^{j-3} &\le (1-ar^\alpha)^{j-3}\exp\left(-a(j-3)\left(r+\theta(y-r)\right)^\alpha-r^\alpha\right)\\
&= O\left((1-ar^\alpha)^{j-3}\,e^{a\alpha j\theta(y-r)r^{\alpha-1}}\right)\\
&= O\left((1-ar^\alpha)^{j-3}\,e^{a\alpha j^{\alpha+1/2}n^{-\alpha}\log n}\right)\\
&= O\left((1-ar^\alpha)^{j-3}\right),
\end{aligned}$$

for $j_2 \le j \le j_0$. Accordingly, for $|y-r| \le r_0$,

$$h''(r+\theta(y-r)) = O\left(h(r)\left(j^2 r^{2\alpha-2} + j r^{\alpha-2}\right)\right).$$

It follows that, by extending the integration limits to the unit interval and by estimating the errors so introduced as in the cases of I_4 and I_6,

$$\begin{aligned} I_5 &= h(r)\binom{n}{j}\int_0^1 (1-y)^{n-j} y^{j-1}\,\mathrm{d}y + h'(r)\binom{n}{j}\int_0^1 (y-r)(1-y)^{n-j}y^{j-1}\,\mathrm{d}y \\ &\quad + O\left(h(r)\left(j^2 r^{2\alpha-2} + j r^{\alpha-2}\right)\binom{n}{j}\int_0^1 (y-r)^2(1-y)^{n-j}y^{j-1}\,\mathrm{d}y + j^{-1}e^{-aj^{\alpha+1}/n^\alpha - (\log n)^2/3}\right) \\ &= h(r)j^{-1} + 0 + O\left(h(r)j^{2\alpha}n^{-2\alpha} + g(r)j^{\alpha-1}n^{-\alpha}\right) \\ &= h(r)j^{-1}\left(1 + O\left(j^{2\alpha+1}n^{-2\alpha} + j^\alpha n^{-\alpha}\right)\right). \end{aligned}$$

This completes the proof of (6.2). □

6.3 Asymptotics II. Probability generating function

Proposition 4. *Under the assumptions of Theorem 2, we have, uniformly for $|u| \le 2$ and $\Re(u) \ge 0$,*

$$E(u^{M_n}) = uW(u)n^{\alpha(u-1)/(\alpha+1)}\left(1 + O\left((\log n)^{-1}\right)\right), \tag{6.5}$$

where

$$W(u) = \frac{\Gamma((\alpha+u)/(\alpha+1))}{\Gamma(u+1)}\, a^{-(u-1)/(\alpha+1)}.$$

If $|u| \le 2$ and $\Re(u) \le 0$ then

$$E(u^{M_n}) = O\left(|u-1|\left(n^{-\alpha} + n^{\alpha(\Re(u)-1)/(\alpha+1)}\right)\right).$$

Proof. By (6.2) and (6.3), we have

$$\begin{aligned} E(u^{M_n}) &= 1 + (u-1)\sum_{1\le j\le j_0} j^{-1}S_{j-1}(u)e^{-[a+O(1/\log(n/j))]j^{\alpha+1}/n^\alpha}\left(1 + O\left(j^{2\alpha+1}n^{-2\alpha} + j^\alpha n^{-\alpha}\right)\right) \\ &\qquad + O\left(|u-1|e^{-(a/2)(\log n)^2}\sum_{j_0\le j\le n} j_0^{-1}|S_{j-1}(u)|\right) \\ &= 1 + (u-1)\sum_{1\le j\le j_0} j^{-1}S_{j-1}(u)e^{-aj^{\alpha+1}/n^\alpha} \\ &\qquad + O\left(|u-1|\sum_{1\le j\le j_0} j^{-1}|S_{j-1}(u)|e^{-aj^{\alpha+1}/(2n^\alpha)}\left((\log n/j)^{-1} + j^{2\alpha+1}n^{-2\alpha} + j^\alpha n^{-\alpha}\right)\right) \\ &\qquad + O\left(|u-1|e^{-(a/2)(\log n)^2}\sum_{j_0\le j\le n} j_0^{-1}|S_{j-1}(u)|\right) \\ &= 1 + (u-1)\sum_{j\ge 1} j^{-1}S_{j-1}(u)e^{-aj^{\alpha+1}/n^\alpha} + T_3, \end{aligned}$$

where, by using the obvious estimate $S_k(u) = O\left(k^{\Re(u)-1}\right)$ for $k \geq 1$ and $|u| \leq 2$,

$$
\begin{aligned}
T_3 &= O\Bigg(|u-1|\int_{j_0}^{\infty} x^{\Re(u)-2} e^{-ax^{\alpha+1}/n^{\alpha}}\,\mathrm{d}x + |u-1|e^{-(a/3)(\log n)^2} \\
&\quad + |u-1|\int_1^{\infty} x^{\Re(u)-2} e^{-ax^{\alpha+1}/(2n^{\alpha})}\left((\log n)^{-1} + x^{2\alpha+1}n^{-2\alpha} + x^{\alpha}n^{-\alpha}\right)\mathrm{d}x\Bigg) \\
&= O\left(|u-1|n^{\alpha(\Re(u)-1)/(\alpha+1)}\left(e^{-a(\log n)^{2(\alpha+1)/\alpha}}(\log n)^{(\Re u-2-\alpha)/(\alpha+1)} + (\log n)^{-1} + n^{-\alpha/(\alpha+1)}\right)\right) \\
&= O\left(|u-1|(\log n)^{-1} n^{\alpha(\Re(u)-1)/(\alpha+1)}\right).
\end{aligned}
$$

We are left with the sum

$$
\begin{aligned}
\Sigma &:= 1 + (u-1)\sum_{j\geq 1} j^{-1} S_{j-1}(u) e^{-aj^{\alpha+1}/n^{\alpha}} \\
&= 1 + \frac{u-1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \Gamma(s) a^{-s} n^{\alpha s} Z(s;u)\,\mathrm{d}s \qquad (c > \max\{0, (\Re(u)-1)/(\alpha+1)\}),
\end{aligned}
$$

where $Z(s;u) := \sum_{j\geq 1} S_{j-1}(u) j^{-1-(\alpha+1)s}$ is absolutely convergent in the half-plane $\Re(s) > (\Re(u)-1)/(\alpha+1)$. By the asymptotic expansion (Flajolet and Odlyzko (1990))

$$
S_j(u) = \frac{j^{u-1}}{\Gamma(u)}\left(1 + \sum_{1\leq l\leq \mu} e_j(u) j^{-l} + O\left(j^{-\mu-1}\right)\right),
$$

uniformly for large j and $|u| \leq 2$, where μ is any fixed positive integer and the $e_j(u)$ are bounded coefficients, we deduce that (i) $Z(s;u)$ is meromorphic in the whole s-plane having simple poles at

$$
s = \frac{u-k}{\alpha+1} \quad \text{with residues } \frac{e_k(u)}{(\alpha+1)\Gamma(u)} \qquad (k = 1, 2, \ldots; u \notin \mathbb{Z});
$$

and (ii)

$$
Z(c+it) = O\left(|t|^{\phi_{\alpha}(c)}\right) \qquad (|t| \to \infty),
$$

where $\phi_{\alpha}(c)$ is a linear function of c; see Proposition 6 of Flajolet et al. (1995).
Thus if $1 \leq \Re(u) \leq 2$, $u \neq 1$,

$$
\begin{aligned}
\Sigma &= 1 + \frac{u-1}{(\alpha+1)\Gamma(u)}\Gamma\left(\frac{u-1}{\alpha+1}\right) a^{-(u-1)/(\alpha+1)} n^{\alpha(u-1)/(\alpha+1)} + (u-1)Z(0;u) \\
&\quad + O\left(|u-1|n^{-\alpha(2-\Re(u))/(\alpha+1)}\right);
\end{aligned}
$$

if $\Re(u) < 1$,

$$
\begin{aligned}
\Sigma &= 1 + (u-1)Z(0;u) + \frac{u-1}{(\alpha+1)\Gamma(u)}\Gamma\left(\frac{u-1}{\alpha+1}\right) a^{-(u-1)/(\alpha+1)} n^{\alpha(u-1)/(\alpha+1)} \\
&\quad + O\left(|u-1|\left(n^{-\alpha} + n^{-\alpha(2-\Re(u))/(\alpha+1)}\right)\right).
\end{aligned}
$$

We complete the proof of the theorem by showing that $1 + (u-1)Z(0;u) = 0$.
To this aim, we substitute the integral representation

$$
n^{-s} = \frac{1}{\Gamma(s)}\int_0^{\infty} e^{-nt} t^{s-1}\,\mathrm{d}t \qquad (n \geq 1; \Re(s) > 0),
$$

in the series definition of Z and interchange the sum and the integral by absolute convergence

$$Z(s;u) = \frac{1}{\Gamma((\alpha+1)s+1)} \int_0^\infty e^{-t}\left(1-e^{-t}\right)^{-u} t^{(\alpha+1)s}\,\mathrm{d}t,$$

for $\Re(s) > \max\{0, (\Re(u)-1)/(\alpha+1)\}$. Now if $\Re(u) < 1$ then

$$1 + (u-1)Z(0;u) = 1 + (u-1)\int_0^\infty e^{-t}\left(1-e^{-t}\right)^{-u}\,\mathrm{d}t = 0;$$

if $1 \le \Re(u) < 2$,

$$1 + (u-1)Z(0;u) = 1 + (u-1)\Gamma(1-u) + (u-1)\int_0^\infty e^{-t}\left(\left(1-e^{-t}\right)^{-u} - t^{-u}\right)\,\mathrm{d}t = 0.$$

The case $\Re(u) \ge 2$ is similar. □

Proof of Theorem 2. (Sketch) The results for the mean and the variance follow from the integral formulae

$$\begin{aligned} E(M_n) &= \frac{1}{2\pi i}\oint_{|w|=\varepsilon} w^{-2}\log E(e^{M_n w})\,\mathrm{d}w, \\ \mathrm{Var}(M_n) &= \frac{2}{2\pi i}\oint_{|w|=\varepsilon} w^{-3}\log E(e^{M_n w})\,\mathrm{d}w, \end{aligned}$$

and Proposition 4 using the approach in [28].

The Poisson approximation formulae are obtained by applying the approach in [29] using Proposition 4, details being omitted here. □

Example. Take $f_0(x) = 1 - a_0(1-x)^\alpha$ and $f(x) := f_0(x)/\int_0^1 f_0(t)\,\mathrm{d}t$, where $0 < a_0 < 1$ and $\alpha > 0$. Then

$$\frac{xf(1-x)}{1-F(1-x)} = 1 - a\,(1-F(1-x))^\alpha\,(1 + O\,(1-F(1-x))^\alpha)\,,$$

where $a = a_0\alpha A^\alpha/(\alpha+1)$. Theorem 2 applies.

Proof of (1.8) and (1.9). If $\mathcal{C}$ is a convex polygon $\mathcal{P}$ with $|I_u| = 0$ and $|I_r| > 0$, then suitable application of Theorem 2 gives

$$\frac{M_n(\mathcal{P}) - \frac{1}{2}\log n}{(\frac{1}{2}\log n)^{1/2}} \xrightarrow{D} \mathcal{N}(0,1),$$

since the region outside the upper-right part can contribute at most $O(1)$ maximal points in probability. Also one easily deduces that (1.6) holds by using (1.11) and results in Section 3.1. The same result holds for the case $|I_u| > 0$ and $|I_r| = 0$ by symmetry. The remaining case $|I_u||I_r| > 0$ is similar.

6.4 Bounded number of maxima.

Theorem 6. *If f is nondecreasing, $\int_0^1 f(t)\,\mathrm{d}t = 1$, and*

$$\frac{xf(1-x)}{1-F(1-x)} = a + o(1) \qquad (0 < a < 1),$$

as $x \to 0^+$, then the number of maxima is bounded:

$$\begin{aligned}
E(M_n) &\sim \frac{1}{a}\log\frac{1}{1-a},\\
\mathrm{Var}(M_n) &\sim \frac{1}{a^2}\left(a\log\frac{1}{1-a} + (a-1)\left(\log(1-a)\right)^2\right),
\end{aligned}$$

and the limiting distribution is still Poisson with mean $\lambda = -\log(1-a)$ (or positive Poisson distribution; see [30, p. 181])

$$E(e^{M_n i\theta}) \to \frac{e^{\lambda e^{i\theta}}-1}{e^{\lambda}-1} = 1 - \frac{1}{a}\left(1-(1-a)^{1-e^{i\theta}}\right). \tag{6.6}$$

Proof. (Sketch) Starting from (6.1), we have, by a much simpler analysis than the proof of Propositions 3 and (4),

$$\begin{aligned}
E(u^{M_n}) &= 1 + (u-1)\sum_{j\geq 1} S_{j-1}(u)\binom{n}{j}\int_0^1 F(x)^{n-j}(1-F(x))^{j-1}\left(\frac{(1-x)f(x)}{(1-F(x))}\right)^{j-1}\mathrm{d}F(x)\\
&\sim 1 + (u-1)\sum_{j\geq 1} S_{j-1}(u)a^{j-1}\binom{n}{j}\int_0^1 F(x)^{n-j}(1-F(x))^{j-1}\,\mathrm{d}F(x)\\
&= 1 + (u-1)\sum_{j\geq 1} S_{j-1}(u)j^{-1}a^{j-1}\\
&= 1 + \frac{1}{a}\left((1-a)^{1-u}-1\right),
\end{aligned}$$

from which (6.6) follows. □

Example. $f(x) = (1-\alpha)(1-x)^{-\alpha}$, where $0 < \alpha < 1$. Then $xf(1-x)/(1-F(1-x)) = 1-\alpha$.

6.5 Number of maxima between $O(1)$ and $O(\log n)$.

Consider

$$\frac{xf(1-x)}{1-F(1-x)} = 1 - L(1-F(1-x))\,(1+o(1)),$$

as $x \to 0^+$, where $L(x) \to 0^+$ is slowly oscillating in the sense that $L(cx)/L(x) \to 1$, as $x \to 0$ for any $c > 0$. Then we have, *heuristically,*

$$\begin{aligned}
&\binom{n}{j}\int_0^1 F(x)^{n-j}(1-F(x))^{j-1}\left(\frac{(1-x)f(x)}{(1-F(x))}\right)^{j-1}\mathrm{d}F(x)\\
&\approx \left(1-L(j/n)(1+o(1))\right)^{j-1}\binom{n}{j}\int_0^1 F(x)^{n-j}(1-F(x))^{j-1}\,\mathrm{d}F(x)\\
&\approx j^{-1}e^{-jL(j/n)(1+o(1))}\\
&\approx j^{-1}e^{-jL(1/n)(1+o(1))}.
\end{aligned}$$

The analytic intuition behind this process is that *the saddlepoint $y = j/n$ of $(1-y)^{n-j}y^j$ is asymptotically not altered by the appearance of the factor $(1-L(y))^{j-1}$, which is more smooth in nature.* This intuition guided most of our analyses in the proof of Theorem 2. Thus, *formally*,

$$\begin{aligned} E(u^{M_n}) &\approx 1+(u-1)\sum_{j\geq 1} j^{-1}S_{j-1}(u)e^{-jL(1/n)} \\ &= 1+\frac{u-1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Gamma(s)L(1/n)^{-s}\left(\sum_{j\geq 1} j^{-1-s}S_{j-1}(u)\right)\mathrm{d}s \\ &\sim \frac{(u-1)\Gamma(u-1)}{\Gamma(u)}L(1/n)^{-(u-1)} \\ &= L(1/n)^{-(u-1)}. \end{aligned}$$

Therefore, we expect again a Poisson distribution in the limit with mean $\log(1/L(1/n))$.

Although we can make this heuristic rigorous by adding suitably technical conditions on L, we only indicate an example to avoid excessive complication.

Example. If $f(x) = e^{(-\log(1-x))^\alpha} / \int_0^1 e^{(-\log x)^\alpha}\,\mathrm{d}x$, where $0<\alpha<1$. Then $L(x) = \alpha(-\log x)^{\alpha-1}$ and we can prove that M_n is asymptotically Poisson with mean $(1-\alpha)\log\log n$.

7 Extensions

We have derived many new quantitative results for the number of maxima in this paper but more new problems arise; we briefly mention some of them.

Convex planar regions. It is stated in Golin (1993) that for convex planar regions $\mathcal{C}$, the mean $E(M_n(\mathcal{C}))$ satisfies either $E(M_n(\mathcal{C})) \asymp n^{1/2}$ or $E(M_n(\mathcal{C})) = O(\log n)$ depending on (in our notation) whether $I_u \cap I_r = \emptyset$, the upper-right part of $\mathcal{C}$ being defined as for convex polygons. By the same method of proving Propositions 1 and 2, it is easy to show that in the case when $I_u \cap I_r \neq \emptyset$, we have

$$E(M_n(\mathcal{C})) \leq H_{n-2} + 2,$$

where H_n denotes the n-th harmonic number. The idea is now to find the two points with maximal x and y coordinates, respectively. If these two points coincide, then there is only one maximal point and the inequality trivially holds. Otherwise, all other maxima lie inside the rectangle formed by the two points for which the expected number of maxima is given by H_{n-2}. One may also derive more precise asymptotics for $E(M_n(\mathcal{C}))$ when $I_u \cap I_r = \emptyset$ using the approaches in Sections 3.1 and 4. A natural question is what happens if we remove convexity?

Convergence rate and large deviations. We naturally expect an error of order $n^{-1/2}$ for our central limit theorems (1.7) and (2.3), but proof is still lacking. What is the "entropy function" for the corresponding large deviation principle? How to refine the asymptotic approximation (1.6) for the variance?

Poisson approximation. Although we have discussed several cases leading to Poisson distribution in the limit, we are still unaware of how "universal" the Poisson limit law is, that is, what is the "minimal condition" besides monotonicity for which $M_n(\mathcal{C})$ is asymptotically Poisson?

Higher dimensions. Most of our approaches either become too tedious or fail for dimension > 2; see Baryshnikov (2000) for a sketch of proof of the asymptotic normality of $M_n([0,1]^d)$.

Average-case analysis of maxima-finding algorithms. Few results are known in this direction when the underlying region is not $[0,1]^d$; see Devroye (1986, 1999). The efficient list algorithm by Bentley et al. (1993) for finding the maxima of a point set performs slightly slower than linear when $\mathcal{C}$ is, say a circle or a right triangle with decreasing hypotenuse, as suggested by simulations; see Golin (1994). But the asymptotic behavior of the expected cost remains open.

Maxima layers. Regarding the set of maximal points as the first layer of maxima, we can consider the second layer maxima of the remaining points after "peeling off" the first layer. Higher layers maxima are defined similarly. The maximum possible layer is called the *depth*. If $\mathcal{C} = [0,1]^2$, maximal layers are closely related to the longest increasing subsequences in random permutations; see Bollobás and Winkler (1988), Aldous and Diaconis (1999), Baik and Rains (1999) for further details. The problem becomes more involved for other planar regions and for higher dimensions. For example, what is the mean value of the second layer maxima when $\mathcal{C} = [0,1]^d$, $d > 2$?

In summary, the study of maxima of random samples is not an isolated subject but instead closely related to many well-studied problems; it also provides natural generalizations of many problems like records, longest increasing subsequences, permutations avoiding certain patterns, partially ordered sets, etc.

Acknowledgements

We thank the referee for useful suggestions.

References

[1] Aldous, D. and Diaconis, P. (1999). Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem. *Bulletin of the American Mathematical Society (New Series)* **36** 413–432.

[2] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records.* John Wiley & Sons, Inc., New York.

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

[4] Bai, Z.-D., Hwang, H.-K., Liang, W.-Q. and Tsai, T.-H. (2000) Limit theorems for the number of maxima in random samples from planar regions. Technical Report C-2000-05, Institute of Statistical Science, Academia Sinica, Taiwan.

[5] Baik, J. and Rains, E. M. (1999). The asymptotics of monotone subsequences of involutions. preprint, 57 pages. `front.math.ucdavis.edu/math.CO/9905084`.

[6] Barndorff-Nielsen, O. and Sobel, M. (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] Baryshnikov, Y. M. (1987) The distribution of the number of nondominating variants. *Soviet Journal of Computer and Systems Sciences* **24** 107–111; translated from *Izvestiya Akademii Nauk SSSR. Tekhnicheskaya Kibernetika*, 47–50 (1986).

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

[9] Becker,R. A., Denby, L., McGill, R. and Wilks, A. R. (1987). Analysis of data from Places Rated Almanac. *American Statistician* **41** 169–186.

[10] Bentley, J. L., Clarkson, L. L. and Levine, D. B. (1993). Fast linear expected-time algorithms for computing maxima and convex hulls. *Algorithmica* **9** 168–183.

[11] Blair, C. (1986). Random linear programs with many variables and few constraints. *Mathematical Programming*, **34** 62–71.

[12] Bollobás, B. and Winkler, P. (1988). The longest chain among random points in Euclidean space. *Proceedings of the American Mathematical Society* **103** 347–353.

[13] Buchta, C. and Reitzner, M. (1997). Equiaffine inner parallel curves of a plane convex body and the convex hulls of randomly chosen points. *Probability Theory and Related Fields* **108** 385–415.

[14] Bühlmann, H. (1970). *Mathematical Methods in Risk Theory.* Die Grundlehren der mathematischen Wissenschaften, Band 172, Springer-Verlag, New York.

[15] Chow, Y. S. and Teicher, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales.* Springer-Verlag, New York.

[16] Datta, A., Maheshwari, A. and Sack, J.-R. (1996). Optimal parallel algorithms for direct dominance problems. *Nordic Journal of Computing* **3** 72–88.

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

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

[19] Devroye, L. (1993). Records, the maximal layer, and uniform distributions in monotone sets. *Computers and Mathematics with Applications* **25** 19–31.

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

[21] Dwyer, R. (1990). Kindler, gentler, average-case analysis for convex hulls and maximal vectors. *SIGACT News* **21** 64–71.

[22] Elster, K.-H. (Ed.) (1993). *Modern Mathematical Methods of Optimization.* Akademie Verlag, Berlin; translated from the Russian by B. Luderer.

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

[24] Frieze, A. and Pittel, B. G. (1995). Probabilistic analysis of an algorithm in the theory of markets in indivisible goods. *Annals of Applied Probability* **5** 768–808.

[25] Golin, M. J. (1993). Maxima in convex regions. in *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, (Austin, TX, 1993), 352–360, ACM, New York.

[26] Güting, R. H., Nurmi, O. and Ottmann, T. (1989). Fast algorithms for direct enclosures and direct dominances. *Journal of Algorithms* **10** 170–186.

[27] Harsanyi, J. C. (1988). *Rational Behavior and Bargaining Equilibrium in Games and Social Situations.* Cambridge University Press, Cambridge; Reprint of the 1986 edition.

[28] Hwang, H.-K. (1998) On convergence rates in the central limit theorems for combinatorial structures. *European Journal of Combinatorics* **19** 329–343.

[29] Hwang, H.-K. (1999). Asymptotics of Poisson approximation to random discrete distributions: an analytic approach. *Advances in Applied Probability* **31** 448–491.

[30] Johnson, N. L., Kotz, S. and Kemp, A. (1992). *Univariate Discrete Distributions.* Second edition, John Wiley & Sons, New York.

[31] Karlin, S. (1959). *Mathematical Methods and Theory in Games, Programming and Economics*, Volume I: Matrix Games, Programming, and Mathematical Economics. Addison-Wesley, Reading, Mass.

[32] Leitmann, G. and Marzollo, A. (Eds.) (1975). *Multicriteria Decision Making.* Springer-Verlag, New York.

[33] Preparata, F. P. and Shamos, M. I. (1985). *Computational Geometry: An Introduction.* Springer-Verlag, New York.

[34] Rényi, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Ann. Fac. Sci. Univ. Clermont-Ferrand No.* **8**, 7–13; also in Selected papers of Alfréd Rényi, Volume III: 1962–1970, Edited by Pál Turán, Akadémiai Kiadó, Budapest (1976).

[35] Rényi, A. and Sulanke, R. (1963). Über die konvexe Hülle von n zufällig gewählten Punkten. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **2** 75–84.

[36] Sholomov, L. A. (1984). Survey of estimational results in choice problems. *Engineering Cybernetics* **21** 51–75.

[37] Stadler, W. (Ed.) (1988). *Multicriteria Optimization in Engineering and in the Sciences.* Plenum Press, New York.

[38] Steuer, R. E. (1986). *Multiple Criteria Optimization: Theory, Computation, and Application.* John Wiley & Sons, New York.

[39] Zeleny, M. (1974). *Linear Multiobjective Programming.* Lecture Notes in Economics and Mathematical Systems, Vol. 95, Springer-Verlag, Berlin-New York.