A New Fine-Scale Berry–Esseen-Type Gumbel-Limit Theorem for Multivariate Maxima

James Allen Fill[1]

Department of Applied Mathematics and Statistics, The Johns Hopkins University,
3400 N. Charles Street, Baltimore, MD 21218-2682, USA

Abstract

For $d \geq 2$ and i.i.d. d-dimensional observations $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots$ with independent Exponential(1) coordinates, let φ_n denote the minimum ℓ^1-norm among the maxima of $\{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}\}$. (A *maximum* from this set is an observation $\mathbf{X}^{(k)}$ with $1 \leq k \leq n$ such that $\mathbf{X}^{(k)} \not\prec \mathbf{X}^{(i)}$ for all $1 \leq i \leq n$, where $\mathbf{x} \prec \mathbf{y}$ means that $x_j < y_j$ for $1 \leq j \leq d$.) Key roles in the study of multivariate Pareto records are played by φ_n and by the more easily handled maximum with the maximum ℓ^1-norm. Fill et al. proved [12, Theorem 1.11(a)] that

$$\varphi_n = \ln n - \ln\ln\ln n - \ln(d-1) + O_{\mathrm{p}}\left(\frac{1}{\ln\ln n}\right),$$

where $Z_n = O_{\mathrm{p}}(a_n)$ means that Z_n/a_n is bounded in probability, and conjectured [12, Remarks 1.13 and 3.3] that

$$(\ln\ln n)\left(\varphi_n - [\ln n - \ln\ln\ln n - \ln(d-1)]\right)$$

has a nondegenerate limiting distribution, suggesting that the limiting distribution might be that of $-G$, where G has a Gumbel distribution with location $-\frac{\ln[(d-1)!]}{d-1}$ and scale $\frac{1}{d-1}$. In the present extended abstract we outline a proof of a Berry–Esseen-type theorem for this convergence in distribution, thereby establishing a very sharp result for φ_n.

2012 ACM Subject Classification Mathematics of computing → Stochastic processes

Keywords and phrases Multivariate maxima, Gumbel distributions, Berry–Esseen-type theorem, Poisson approximation, Chen–Stein method, multivariate Pareto records

Digital Object Identifier 10.4230/LIPIcs.AofA.2026.21

Related Version *Full Version*: `https://arxiv.org/abs/2601.18170` [9]

Funding Research supported by the Acheson J. Duncan Fund for the Advancement of Research in Statistics.

Acknowledgements We thank an anonymous reviewer for helpful comments.

1 Introduction, background, assumptions, notation, and statement of main results

Note. All detailed proofs, including some omitted in this extended abstract due to space limitations, are included in the full paper [9]. Note also that the essential reference [12] cited frequently in this extended abstract is the full paper corresponding to the extended abstract [11].

Some notation. We define $\mathbf{1}(A)$ to be 0 or 1 according as A is false or true. We denote by $\mathcal{L}(Y)$ the distribution, or law, of a random variable Y. For a positive integer n, let $[n] := \{1, \dots, n\}$. We find it very useful to abbreviate the kth iterate of natural logarithm

[1] Corresponding author

37th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2026).
Editor: Konstantinos Panagiotou; Article No. 21; pp. 21:1–21:20
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

ln by L_k and L_1 by L. We fix a dimension $d \geq 2$, and $\mathbf{x}$ is then shorthand for the vector $(x_1, \ldots, x_d)$. We write $\mathbf{0}$ for a vector of 0's and $\mathbf{x} \prec \mathbf{y}$ (or $\mathbf{y} \succ \mathbf{x}$) to mean that $x_j < y_j$ for every $j \in [d]$. For simplicity of notation, we will write $\|\mathbf{x}\|$ (rather than $\|\mathbf{x}\|_{\ell^1}$ or $\|\mathbf{x}\|_1$) for the ℓ^1-norm $\sum_{j=1}^d |x_j|$ ($= \sum_{j=1}^d x_j$ if $\mathbf{x} \succ \mathbf{0}$) of $\mathbf{x}$. We take the *probability simplex* $\mathcal{S}_{d-1}$ to be the set of all $\mathbf{x} \succ \mathbf{0}$ with $\|\mathbf{x}\| = 1$. All the results of this extended abstract hold for any fixed dimension $d \geq 2$, and all asymptotic assertions are as $n \to \infty$ unless otherwise specified.

Univariate records have been quite thoroughly studied (see [1] for a book-length treatment), but there is still fundamental knowledge to uncover about multivariate (Pareto) records, and specifically about (partial) *maxima* [see the next paragraph for a definition of maximum; in the standard terminology of [10], a maximum at epoch n is a *current* or *remaining* (Pareto) *record*], even under the simple category of models (call the category by the name S) that observations are i.i.d. (independent and identically distributed) and coordinates of each observation are independent and continuously distributed (with possibly different distributions for the different coordinates).

We assume throughout this extended abstract that $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots$ is a sequence of i.i.d. copies of a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ taking values in $(0, \infty)^d$ and having independent Exponential(1) coordinates. Call this Model E, a member of Category S. We are interested in the sequence, indexed by n, of the set of maxima of $\{\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}\}$ at epoch n; a *maximum* from this set is an observation $\mathbf{X}^{(k)}$ with $1 \leq k \leq n$ such that $\mathbf{X}^{(k)} \not\prec \mathbf{X}^{(i)}$ for all $1 \leq i \leq n$. Results about maxima under Model E are readily transformed to results under any specific model falling in Category S.

It is natural and important to study the ℓ^1-norms of maxima under Model E. One motivation for this study is that [10, Theorem 1.4] establishes the basic result that the conditional distribution of $\|\mathbf{X}^{(k)}\|$ given that $\mathbf{X}^{(k)}$ is a current record at epoch n doesn't depend on $k \in [n]$ (so let us suppose now that $k = n$) and (after centering by $\ln n$) this "generic maximum" has a weak limit that is standard Gumbel (i.e., has distribution function $x \mapsto e^{-e^{-x}}$, $x \in \mathbb{R}$); moreover, as seen from [8, Remark 2.1(b)], the proof of [10, Theorem 1.4] shows that $\|\mathbf{X}^{(n)}\|$ and $\mathbf{U}_n := \|\mathbf{X}^{(n)}\|^{-1}\mathbf{X}^{(n)}$ are independent, with $\mathbf{U}_n$ uniformly distributed on the probability simplex $\mathcal{S}_{d-1}$. Another motivation is that the observation by Bai et al. [3, last paragraph of Section 1 and Section 3.2] "that nearly all maxima occur in a thin strip sandwiched between [the] two parallel hyper-planes"

$$\|\mathbf{x}\| = \mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}[4(d-1)] \quad \text{and} \quad \|\mathbf{x}\| = \mathrm{L}\, n + 4(d-1)\, \mathrm{L}_2\, n. \tag{1.1}$$

is one key to their proof of asymptotic normality for the number of maxima at epoch n, a result that had been indicated by Baryshnikov [5].

Fill and Naiman [10] initiated, and Fill et al. [12] refined, the study of sharp localization of (separately) the maximum and minimum ℓ^1-norms of maxima (and related quantities) at epoch n under Model E. See [10, Theorem 1.8] (or [12, Theorem 1.6]), due essentially to the classical extreme value theory of Kiefer [13], for treatment of the maximum ℓ^1-norm, denoted there as F_n^+. In particular, the theorem establishes that

$$F_n^+ - \left[\mathrm{L}\, n + (d-1)\, \mathrm{L}_2\, n - \mathrm{L}((d-1)!)\right] \tag{1.2}$$

converges in distribution to standard Gumbel.

This extended abstract concerns the minimum ℓ^1-norm, denoted $\widehat{F}_n^-$ in [12] and more simply φ_n here, of any maximum at epoch n. Fill et al. [12] applied the first- and second-moment methods to the count (for arbitrarily chosen $a \in \mathbb{R}$) of the number of maxima with ℓ^1-norm smaller than

$$b_n^* \equiv b_n^*(a) := \mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}(d-1) + \frac{a}{\mathrm{L}_2\, n} \tag{1.3}$$

to prove [12, Theorem 1.11(a)] that

$$\varphi_n = \mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}(d-1) + O_{\mathrm{p}}\left(\frac{1}{\mathrm{L}_2\, n}\right),$$

where $Z_n = O_{\mathrm{p}}(a_n)$ means that Z_n/a_n is bounded in probability, and conjectured [12, Remarks 1.13 and 3.3] that

$$\varphi_n^\circ := (\mathrm{L}_2\, n)\,(\varphi_n - [\mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}(d-1)]) \tag{1.4}$$

has a nondegenerate limiting distribution, suggesting that the limiting distribution might be that of $-G$, where G has a Gumbel distribution with location $-\frac{\ln[(d-1)!]}{d-1}$ and scale $\frac{1}{d-1}$. In the present extended abstract we prove this convergence in distribution, with a Berry–Esseen-type bound, thereby describing the asymptotic behavior of φ_n on a quite fine scale – in particular, finer than that for the much more easily studied F_n^+.

Here then is the main result of this extended abstract. Recall that if G_0 is distributed standard Gumbel (with distribution function $x \mapsto e^{-e^{-x}}$, $x \in \mathbb{R}$) and $m \in \mathbb{R}$ and $s \in (0, \infty)$, then $m + s\, G_0$ has a Gumbel distribution with location m and scale s. Recall also the definition of *Kolmogorov distance* d_{K} between two probability measures μ_1 and μ_2 on $(\mathbb{R}, \text{Borels})$ with respective distribution functions F_1 and F_2:

$$d_{\mathrm{K}}(\mu_1, \mu_2) := \sup_{t \in \mathbb{R}} |F_1(t) - F_2(t)|.$$

▶ **Theorem 1.1** (main theorem: Gumbel limit). *Fix $d \geq 2$. Let $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots, \mathbf{X}^{(n)}$ be independent random d-dimensional vectors each having independent* Exponential(1) *coordinates, let φ_n denote the minimum ℓ^1-norm of any maximum of these vectors, and define φ_n° as at* (1.4)*. Then φ_n° converges in distribution to $-G$, where G has a Gumbel distribution with location $-\frac{\mathrm{L}[(d-1)!]}{d-1}$ and scale $\frac{1}{d-1}$; moreover,*

$$d_{\mathrm{K}}(\mathcal{L}(\varphi_n^\circ), \mathcal{L}(-G)) = O((\mathrm{L}_2\, n)^{-1/2}\, \mathrm{L}_3\, n). \tag{1.5}$$

In other words, (1.5) asserts precisely that

$$\left| \mathbb{P}(\varphi_n^\circ > a) - \exp\left[-\frac{e^{(d-1)a}}{(d-1)!}\right] \right| = O((\mathrm{L}_2\, n)^{-1/2}\, \mathrm{L}_3\, n) \text{ uniformly in } a \in \mathbb{R}, \tag{1.6}$$

or equivalently that

$$\left| \mathbb{P}(\varphi_n > b_n^*(a)) - \exp\left[-\frac{e^{(d-1)a}}{(d-1)!}\right] \right| = O((\mathrm{L}_2\, n)^{-1/2}\, \mathrm{L}_3\, n) \tag{1.7}$$

uniformly in $a \in \mathbb{R}$. In this extended abstract we have tried for a good rate in Theorem 1.1, but we do not know the optimal rate. The plan for proving (1.7) will be first to prove (1.7) uniformly for $|a| > a_n$, where

$$a_n := \frac{1}{2(d-1)}(\mathrm{L}_3\, n - 2\, \mathrm{L}_4\, n); \tag{1.8}$$

then to prove

$$|\mathbb{P}(\varphi_n > b_n^*(a)) - \mathbb{P}(\varphi_n > b_n(a))| = O((\mathrm{L}_2\, n)^{-1/2}\, \mathrm{L}_3\, n) \text{ uniformly for } |a| \leq a_n, \tag{1.9}$$

where $b_n(a)$ is the following (fairly slight) modification of $b_n^*(a)$, employed for computational convenience:

$$b_n \equiv b_n(a) := \mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}(d-1) - \mathrm{L}\left(1 - \frac{a}{\mathrm{L}_2\, n}\right); \tag{1.10}$$

and finally to prove

$$\begin{aligned}&\left|\mathbb{P}(\varphi_n > b_n(a)) - \exp\left[-\frac{e^{(d-1)a}}{(d-1)!}\right]\right|\\ &\quad = O\left((\mathrm{L}_2\, n)^{-\left(d-\frac{3}{2}\right)}(\mathrm{L}_3\, n)^{-1}\right) = o((\mathrm{L}_2\, n)^{-1/2}\,\mathrm{L}_3\, n) \text{ uniformly for } |a| \le a_n. \qquad (1.11)\end{aligned}$$

As in [12, (2.1)], for $b > 0$ define

$$\rho_n(b) \coloneqq \#\{\text{maxima } \mathbf{r} \text{ at epoch } n \text{ with } \|\mathbf{r}\| \le b\}. \qquad (1.12)$$

Recall that the total variation distance between (the distributions of) two integer-valued random variables W_1 and W_2 can be defined as the (achieved) supremum

$$\mathrm{d_{TV}}(W_1, W_2) \equiv \mathrm{d_{TV}}(\mathcal{L}(W_1), \mathcal{L}(W_2)) := \sup_{A \subseteq \mathbb{Z}} |\mathbb{P}(W_1 \in A) - \mathbb{P}(W_2 \in A)|$$

(it is sometimes defined as twice this quantity) and that

$$\begin{aligned}\mathrm{d_{TV}}(W_1, W_2) &= \tfrac{1}{2}\sup_f |\,\mathbb{E}\, f(W_1) - \mathbb{E}\, f(W_2)| = \inf \mathbb{P}(W_1 \ne W_2) \qquad (1.13)\\ &= \sum_k [\mathbb{P}(W_1 = k) - \mathbb{P}(W_2 = k)]^+, \qquad (1.14)\end{aligned}$$

where in (1.13) the (achieved) supremum is taken over all functions $f : \mathbb{Z} \to \mathbb{R}$ with sup-norm bounded by 1 and the (achieved) infimum is taken over all couplings of $\mathcal{L}(W_1)$ and $\mathcal{L}(W_2)$, and where in (1.14) $x^+ = \max\{x, 0\}$ denotes the positive part of x. Recall also that (restricting to integer-valued random variables $W, W_1, W_2, \ldots$) convergence of W_n to W in distribution is equivalent to $\mathrm{d_{TV}}(W_n, W) \to 0$.

We write $\mathrm{Po}(\lambda)$ as shorthand for the Poisson distribution with mean λ (or, sometimes abusing notation, for a random variable with that distribution). By a standard "switching relation" recasting we have

$$\mathbb{P}(\varphi_n > b_n(a)) = \mathbb{P}(\rho_n(b_n(a)) = 0); \qquad (1.15)$$

further, with

$$\lambda \equiv \lambda(a) := \frac{e^{(d-1)a}}{(d-1)!}, \qquad (1.16)$$

the subtracted term in (1.11) is $e^{-\lambda}$, the probability that $\mathrm{Po}(\lambda) = 0$. Thus our key result (1.11) will follow immediately from the following Poisson approximation theorem, a theorem we believe is of interest in its own right.

▶ **Theorem 1.2** (Poisson approximation). *Fix $d \ge 2$ and recall the notation* (1.12)*,* (1.10)*, and* (1.16)*. Then as $n \to \infty$ we have*

$$\mathrm{d_{TV}}(\rho_n(b_n), \mathrm{Po}(\lambda)) = O\left((\mathrm{L}_2\, n)^{-\left(d-\frac{3}{2}\right)}(\mathrm{L}_3\, n)^{-1}\right) \text{ uniformly for } |a| \le a_n. \qquad (1.17)$$

▶ Remark 1.3. Using Theorem 1.2, the rate of convergence in Theorem 1.1 improves to match that in Theorem 1.2 if $\mathcal{L}(\varphi_n^\circ)$ is replaced by the distribution of $(\mathrm{L}_2\, n)[1 - \exp(-\varphi_n^\circ/\mathrm{L}_2\, n)]$.

For fixed $a \in \mathbb{R}$, it follows from [12, Lemma 2.1] that

$$\mathbb{E}\,\rho_n(b_n^*) \to \lambda \tag{1.18}$$

and from [12, Lemma 3.1 and Remark B.1] that

$$\operatorname{Var}\rho_n(b_n^*) \to \lambda \tag{1.19}$$

as well. The theme of [2], as stated by the authors, is that "Convergence to the Poisson distribution, for the number of occurrences of dependent events, can often be established by computing only first and second moments, but not higher ones", and (notwithstanding the slight discrepancy between b_n^* and b_n) that theme, together with (1.18)–(1.19), is the motivation for the proof of Theorem 1.2.

Before carrying out the technical details of [2] to provide rigorous proofs of Theorems 1.1–1.2, which will require a rather substantial amount of modification of $\rho_n(b_n)$ and many additional calculations, let us give a (very) rough sketch as to why $\rho_n(b_n^*)$ should be approximately distributed as $\mathrm{Po}(\lambda)$, when (as we assume here) n is large. For $\underline{b}_n$ defined at (4.2), $\rho_n(\underline{b}_n) = 0$ with high probability. Moreover, for $\mathbf{x} \succ \mathbf{0}$ satisfying $\underline{b}_n < \|\mathbf{x}\| \le b_n^*$ and $G_n(\mathbf{x})$ a small neighborhood of $\mathbf{x}$, the events

$$E_n(\mathbf{x}) := \{\text{there is at least one maximum of } \{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}\} \text{ in } G_n(\mathbf{x})\} \tag{1.20}$$

are "mostly" independent as $\mathbf{x}$ varies (in the sense that there is only suitably local dependence), with probabilities

$$\begin{aligned}
&\mathbb{P}(E_n(\mathbf{x}))\\
&\approx \mathbb{P}(\text{there is exactly one maximum of } \{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}\} \text{ in } G_n(\mathbf{x}))\\
&\approx n\mathbb{P}(\mathbf{X}^{(n)} \in G_n(\mathbf{x}),\ \mathbf{X}^{(n)} \text{ is a maximum of } \{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}\})\\
&\approx \int_{\mathbf{y} \in G_n(\mathbf{x})} n e^{-\|\mathbf{y}\|}(1 - e^{-\|\mathbf{y}\|})^{n-1}\,\mathrm{d}\mathbf{y}\\
&\approx \int_{\mathbf{y} \in G_n(\mathbf{x})} (n-1)e^{-\|\mathbf{y}\|}\exp\Big[-(n-1)e^{-\|\mathbf{y}\|}\Big]\,\mathrm{d}\mathbf{y};
\end{aligned} \tag{1.21}$$

thus, $\rho_n(b_n^*)$ is approximately Poisson distributed with parameter

$$\begin{aligned}
&\int_{\mathbf{x}\succ\mathbf{0}:\,\underline{b}_n<\|\mathbf{x}\|\le b_n^*} (n-1)e^{-\|\mathbf{x}\|}\exp\Big[-(n-1)e^{-\|\mathbf{x}\|}\Big]\,\mathrm{d}\mathbf{x}\\
&= \int_{(n-1)e^{-b_n^*}}^{(n-1)e^{-\underline{b}_n}} \frac{[\mathrm{L}(n-1) - \mathrm{L}\,z]^{d-1}}{(d-1)!} e^{-z}\,\mathrm{d}z\\
&\approx \frac{(\mathrm{L}\,n)^{d-1}}{(d-1)!}\exp\Big[-(n-1)e^{-b_n^*}\Big] \approx \frac{e^{(d-1)a}}{(d-1)!} = \lambda.
\end{aligned} \tag{1.22}$$

Our rigorous proofs of Theorems 1.1–1.2 proceed as follows. In Section 2 we prove (1.7) uniformly for $|a| > a_n$. In Section 3 we prove (1.9). The remaining sections, with the exception of the speculative final section, are devoted to the proof of Theorem 1.2, which will complete the proof of Theorem 1.1.

In Section 4 we introduce a boundary $(\underline{b}_n)$ that is somewhat smaller than (b_n) and prove that

$$\mathrm{d_{TV}}(\rho_n(b_n), \rho_n(b_n) - \rho_n(\underline{b}_n)) = o((\mathrm{L}\,n)^{-(d-1)}) \text{ uniformly for } a \in \mathbb{R}. \tag{1.23}$$

In Section 5 we introduce a maxima-count N_n that is a Poissonized version of $\rho_n(b_n) - \rho_n(\underline{b}_n)$ and prove that

$$\mathrm{d_{TV}}(\rho_n(b_n) - \rho_n(\underline{b}_n), N_n) = o((\mathrm{L}\,n)^{-(d-1)}) \text{ uniformly for } |a| \le a_n. \tag{1.24}$$

In Section 6 we introduce a quite large boundary $(\overline{b}_n)$ and a random variable $\overline{N}_n$ that modifies N_n by conditioning on there being no points of the Poisson process with ℓ^1-norm larger than $\overline{b}_n$ and prove that

$$\mathrm{d_{TV}}\big(N_n, \overline{N}_n\big) = O((\mathrm{L}\,n)^{-(d-1)}) \text{ uniformly for } |a| \le a_n. \tag{1.25}$$

In Section 7, by discretization we obtain approximations $Z_{n,m}$ to $\overline{N}_n$, where each $Z_{n,m}$ (one random variable for each $m \ge 0$) is a sum of indicator random variables $Z_{n,m,\alpha}$, and prove (for each fixed n and fixed $a \in \mathbb{R}$) that

$$\mathrm{d_{TV}}\big(\overline{N}_n, Z_{n,m}\big) \to 0 \text{ as } m \to \infty. \tag{1.26}$$

This discretization and the proof that the dependence of the random variables $Z_{n,m,\alpha}$ is suitably local are similar to what was done in [3] to prove asymptotic normality of the total number $\rho_n(\infty)$ of maxima, unrestricted by location.

The heart of the proof of Theorem 1.2 is to employ the technique of [2] and prove that

$$\limsup_{m\to\infty} \mathrm{d_{TV}}(Z_{n,m}, \mathrm{Po}(\lambda_{n,m})) = O\left((\mathrm{L}_2\, n)^{-\left(d-\frac{3}{2}\right)} (\mathrm{L}_3\, n)^{-1}\right) \text{ uniformly for } |a| \le a_n \tag{1.27}$$

as $n \to \infty$, where $\lambda_{n,m} := \mathbb{E}\, Z_{n,m}$. Our proof of (1.27) will use results from other sections, including Sections 8–9, so we defer it partially to Section 10 and partially to the full paper [9].

In Section 8 we prove (for each fixed n and fixed $a \in \mathbb{R}$) that

$$\mathrm{d_{TV}}\big(\mathrm{Po}(\lambda_{n,m}), \mathrm{Po}(\mathbb{E}\,\overline{N}_n)\big) \to 0 \text{ as } m \to \infty. \tag{1.28}$$

In Section 9 we reference the proof in the full paper [9] that

$$\mathrm{d_{TV}}\big(\mathrm{Po}(\mathbb{E}\,\overline{N}_n), \mathrm{Po}(\lambda)\big) = O((\mathrm{L}\,n)^{-1}(\mathrm{L}_2\, n)^{1/2}) \text{ uniformly for } |a| \le a_n. \tag{1.29}$$

Theorem 1.2 then follows by first applying the triangle inequality to $\mathrm{d_{TV}}(\rho_n(b_n), \mathrm{Po}(\lambda))$ using (1.23)–(1.29) (for fixed n and fixed $|a| \le a_n$), then taking limit superior as $m \to \infty$ on both sides of the inequality, and finally taking the supremum of both sides over $|a| \le a_n$ on both sides of the inequality.

Our proofs use several results that are either taken from [12] or closely related to results from [12]. For the reader's convenience, such results are collected in Appendix A.

In Section 11 we consider the distribution of the maximum at epoch n having smallest ℓ^1-norm, not just (as in all other sections) its ℓ^1-norm.

2 Distribution tails

▶ **Proposition 2.1.** *With a_n defined at (1.8), we have*

$$\left|\mathbb{P}(\varphi_n > b_n^*(a)) - e^{-\lambda(a)}\right| = O((\mathrm{L}_2\, n)^{-1/2}\, \mathrm{L}_3\, n) \tag{2.1}$$

uniformly for $|a| > a_n$.

Proof. First suppose that $a > a_n$. Then, using Lemma A.1,

$$\begin{aligned}
\mathbb{P}(\varphi_n > b_n^*(a)) &\le \mathbb{P}(\varphi_n > b_n^*(a_n)) \\
&\le (1+o(1))\,(d-1)!(\mathrm{L}\,n)^{-(d-1)\left(1-e^{-\frac{a_n}{\mathrm{L}_2\,n}}\right)} \\
&\sim (d-1)!\exp[-(d-1)a_n] \\
&= (d-1)!(\mathrm{L}_2\,n)^{-1/2}\,\mathrm{L}_3\,n;
\end{aligned}$$

and

$$e^{-\lambda(a)} \le e^{-\lambda(a_n)} = \exp\left[-\frac{(\mathrm{L}_2\,n)^{1/2}(\mathrm{L}_3\,n)^{-1}}{(d-1)!}\right] = o\Big((\mathrm{L}_2\,n)^{-1/2}\,\mathrm{L}_3\,n\Big).$$

Now suppose instead that $a < -a_n$. Then, again using Lemma A.1,

$$\begin{aligned}
1-\mathbb{P}(\varphi_n > b_n^*(a)) &= \mathbb{P}(\varphi_n \le b_n^*(a)) \\
&\le \mathbb{P}(\varphi_n \le b_n^*(-a_n)) \\
&\le (1+o(1))\,\frac{1}{(d-1)!}(\mathrm{L}\,n)^{-(d-1)\left(e^{\frac{a_n}{\mathrm{L}_2\,n}}-1\right)} \\
&\sim \frac{\exp[-(d-1)a_n]}{(d-1)!} \\
&= \frac{(\mathrm{L}_2\,n)^{-1/2}\,\mathrm{L}_3\,n}{(d-1)!};
\end{aligned}$$

and

$$\begin{aligned}
1-e^{-\lambda(a)} &\le 1-e^{-\lambda(-a_n)} \\
&= 1-\exp\left[-\frac{(\mathrm{L}_2\,n)^{-1/2}\,\mathrm{L}_3\,n}{(d-1)!}\right] \\
&\sim \frac{(\mathrm{L}_2\,n)^{-1/2}\,\mathrm{L}_3\,n}{(d-1)!}.
\end{aligned}$$

The proof of the proposition is complete. ◀

3 Slight modification (b_n) of the boundary (b_n^*)

For our Poisson approximation theorem, Theorem 1.2, we find it a bit more convenient to use the boundary

$$b_n(a) := \mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}(d-1) - \mathrm{L}\left(1-\frac{a}{\mathrm{L}_2\,n}\right) \tag{3.1}$$

rather than (1.3). In this section we establish (1.9) as the following proposition.

▶ **Proposition 3.1.** *We have*

$$|\mathbb{P}(\varphi_n > b_n^*(a)) - \mathbb{P}(\varphi_n > b_n(a))| = O((\mathrm{L}_2\,n)^{-1/2}\,\mathrm{L}_3\,n) \text{ uniformly for } |a| \le a_n. \tag{3.2}$$

Proof. First note that for every a satisfying $|a| \le a_n$ we have $b_n(a) \ge b_n^*(a)$, and so the left side of (3.2) equals

$$\begin{aligned}
\mathbb{P}(b_n^*(a) < \varphi_n \le b_n(a)) &= \mathbb{P}(\rho_n(b_n(a)) \ge 1,\, \rho_n(b_n^*(a)) = 0) \\
&\le \mathbb{P}(\rho_n(b_n(a)) - \rho_n(b_n^*(a)) \ge 1) \\
&\le \mathbb{E}[\rho_n(b_n(a)) - \rho_n(b_n^*(a))] =: \Delta_n(a) \equiv \Delta_n.
\end{aligned}$$

Now, using the inequality $1 - x \le e^{-x}$ for $x > 0$ on the second line to follow,

$$\begin{aligned}
\Delta_{n+1} &= \frac{(n+1)}{(d-1)!} \int_{b^*_{n+1}}^{b_{n+1}} y^{d-1} e^{-y} (1 - e^{-y})^n \, \mathrm{d}y \\
&\le \frac{(n+1)}{(d-1)!} \int_{b^*_{n+1}}^{b_{n+1}} y^{d-1} \exp(-ne^{-y} - y) \, \mathrm{d}y \\
&= \frac{1 + n^{-1}}{(d-1)!} \int_{ne^{-b_{n+1}}}^{ne^{-b^*_{n+1}}} (\mathrm{L}\, n - \mathrm{L}\, z)^{d-1} e^{-z} \, \mathrm{d}z \\
&\le (1 + n^{-1}) \frac{(\mathrm{L}\, n)^{d-1}}{(d-1)!} \int_{ne^{-b_{n+1}}}^{ne^{-b^*_{n+1}}} e^{-z} \, \mathrm{d}z \\
&\le (1 + n^{-1}) \frac{(\mathrm{L}\, n)^{d-1}}{(d-1)!} \exp(-ne^{-b_{n+1}}) \times n(e^{-b^*_{n+1}} - e^{-b_{n+1}}).
\end{aligned}$$

But

$$\begin{aligned}
e^{-b_{n+1}} &= (n+1)^{-1} [\mathrm{L}_2(n+1)](d-1) \left(1 - \frac{a}{\mathrm{L}_2(n+1)}\right) \\
&\ge (n+1)^{-1} [\mathrm{L}_2(n+1)](d-1) \left(1 - \frac{a_{n+1}}{\mathrm{L}_2(n+1)}\right)
\end{aligned}$$

and

$$\begin{aligned}
&e^{-b^*_{n+1}} - e^{-b_{n+1}} \\
&\quad \le e^{-b^*_{n+1}} (b_{n+1} - b^*_{n+1}) \\
&\quad = (n+1)^{-1} [\mathrm{L}_2(n+1)](d-1) e^{-a/\mathrm{L}_2 n} \left[-\mathrm{L}\left(1 - \frac{a}{\mathrm{L}_2(n+1)}\right) - \frac{a}{\mathrm{L}_2(n+1)} \right] \\
&\quad \le (1 + o(1)) (n+1)^{-1} [\mathrm{L}_2(n+1)](d-1) \left[-\mathrm{L}\left(1 - \frac{a_{n+1}}{\mathrm{L}_2(n+1)}\right) - \frac{a_{n+1}}{\mathrm{L}_2(n+1)} \right] \\
&\quad \sim \frac{1}{2} (d-1)(n+1)^{-1} [\mathrm{L}_2(n+1)]^{-1} a^2_{n+1} \\
&\quad \sim \frac{1}{8(d-1)} (n+1)^{-1} [\mathrm{L}_2(n+1)]^{-1} [\mathrm{L}_3(n+1)]^2,
\end{aligned}$$

so

$$\Delta_{n+1} \le (1 + o(1)) \frac{[\mathrm{L}_2(n+1)]^{-1/2}\, \mathrm{L}_3(n+1)}{8(d-1)(d-1)!}$$

and

$$\Delta_n = O((\mathrm{L}_2\, n)^{-1/2}\, \mathrm{L}_3\, n),$$

as desired. ◀

4 A boundary $(\underline{b}_n)$ that is somewhat smaller than (b_n)

In this section we introduce a boundary $(\underline{b}_n)$ (not depending on a) that is somewhat smaller than (b_n) (when $a \ge -a_n$) and prove that

$$\mathrm{d}_{\mathrm{TV}}(\rho_n(b_n), \rho_n(b_n) - \rho_n(\underline{b}_n)) = o((\mathrm{L}\, n)^{-(d-1)}) \text{ uniformly in } a \in \mathbb{R}, \tag{4.1}$$

which is (1.23). Indeed, let

$$\underline{b}_n := \mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}(d-1) - \mathrm{L}\,3. \tag{4.2}$$

Then

$$\mathrm{d_{TV}}(\rho_n(b_n), \rho_n(b_n) - \rho_n(\underline{b}_n)) \le \mathbb{P}(\rho_n(\underline{b}_n) \ne 0) \le \mathbb{E}\,\rho_n(\underline{b}_n).$$

By Lemma A.2,

$$\mathbb{E}\,\rho_n(\underline{b}_n) = O((\mathrm{L}\,n)^{-2(d-1)}) = o((\mathrm{L}\,n)^{-(d-1)}),$$

completing the proof of (4.1).

▶ Remark 4.1. Though we will not need it (nor give a proof), more accurate asymptotic determination of $\mathbb{E}\,\rho_n(\underline{b}_n)$ is possible, namely,

$$\mathbb{E}\,\rho_n(\underline{b}_n) = \frac{1}{(d-1)!}(\mathrm{L}\,n)^{-2(d-1)} + O((\mathrm{L}\,n)^{-2(d-1)-1}\,\mathrm{L}_3\,n).$$

5 Poissonization

In this section we introduce a maxima-count N_n that is a Poissonized version of $\rho_n(b_n)-\rho_n(\underline{b}_n)$ and prove that

$$\mathrm{d_{TV}}(\rho_n(b_n) - \rho_n(\underline{b}_n), N_n) = o((\mathrm{L}\,n)^{-(d-1)}) \text{ uniformly for } |a| \le a_n, \tag{5.1}$$

which is (1.24).

Indeed, let N_n count the maxima from a Poisson process with intensity measure

$$\mu_n(\mathrm{d}\mathbf{x}) := n e^{-\|\mathbf{x}\|}\mathbf{1}(\mathbf{x} \succ \mathbf{0})\,\mathrm{d}\mathbf{x} \tag{5.2}$$

that belong to the set

$$A_n \equiv A_n(a) := \{\mathbf{x} \succ \mathbf{0} : \underline{b}_n < \|\mathbf{x}\| \le b_n\}, \tag{5.3}$$

a set which is nonempty because $a \ge -a_n$.

Proof of (5.1). Let

$$B_n := \{\mathbf{x} \succ \mathbf{0} : \|\mathbf{x}\| > \underline{b}_n\}, \tag{5.4}$$

and observe that (with no dependence on $a \in \mathbb{R}$)

$$\begin{aligned} p_n &:= n^{-1}\mu_n(B_n) \\ &= \mathbb{P}(\mathrm{Gamma}(d,1) > \underline{b}_n) = \frac{1}{(d-1)!}\int_{\underline{b}_n}^{\infty} x^{d-1}e^{-x}\,\mathrm{d}x \\ &\sim \frac{\underline{b}_n^{d-1}}{(d-1)!}e^{-\underline{b}_n} \sim \frac{1}{(d-1)!}(\mathrm{L}\,n)^{d-1}n^{-1}(\mathrm{L}_2\,n)\,3(d-1) \\ &= o((\mathrm{L}\,n)^{-(d-1)}). \end{aligned} \tag{5.5}$$

To establish (5.1), we first note that observations (whether $\mathbf{X}^{(i)}$'s or Poisson points) with ℓ^1-norm bounded above by $\underline{b}_n$ play no role in determining which observations falling in A_n are maxima. So we can proceed as in the proof of [16, Theorem 4.1] and note (see, e.g., [15, Proposition 3.6] for justification) that, for each nonnegative integer k, the conditional distribution of N_n given that k points of the Poisson process fall in B_n is exactly the conditional distribution of $\rho_n(b_n) - \rho_n(\underline{b}_n)$ given that k of the n observations $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}$ fall in B_n. Thus, by conditioning and (1.13), $\mathrm{d_{TV}}(\rho_n(b_n) - \rho_n(\underline{b}_n), N_n)$ is bounded by twice the total variation distance between the counts of values falling in B_n; these two counts have distributions Binomial(n, p_n) and Poisson(np_n). Therefore, because the total variation distance between these Binomial and Poisson distributions is bounded by p_n (see, e.g., [4, Chapter 1, (1.23)]), we conclude from (5.5) that

$$\mathrm{d_{TV}}(\rho_n(b_n) - \rho_n(\underline{b}_n), N_n) \le 2p_n = o((\mathrm{L}\, n)^{-(d-1)}),$$

with the bound not depending on a, as claimed at (5.1). ◀

We have observed that we may, and henceforth do, take N_n to count the maxima from a Poisson process with intensity measure

$$\underline{\mu}_n(\mathrm{d}\mathbf{x}) := n e^{-\|\mathbf{x}\|} \mathbf{1}(\mathbf{x} \succ \mathbf{0},\, \|\mathbf{x}\| > \underline{b}_n)\, \mathrm{d}\mathbf{x} \tag{5.6}$$

that belong to the set A_n of (5.3).

6 Conditioning

In this section we introduce

$$\overline{b}_n := \mathrm{L}\, n + 2(d-1)\, \mathrm{L}_2\, n \tag{6.1}$$

and $\overline{N}_n$ such that, for each n, the law of the random variable $\overline{N}_n$ is the conditional distribution of N_n given that the Poisson point process with intensity measure (5.6) has no points with ℓ^1-norm larger than $\overline{b}_n$. Put another way, we can (and do) define $\overline{N}_n$ to be the number of maxima of the Poisson point process with intensity measure

$$\overline{\mu}_n(\mathrm{d}\mathbf{x}) := n e^{-\|\mathbf{x}\|} \mathbf{1}(\mathbf{x} \succ \mathbf{0},\, \underline{b}_n < \|\mathbf{x}\| \le \overline{b}_n)\, \mathrm{d}\mathbf{x} \tag{6.2}$$

that belong to the set A_n of (5.3). The following proposition establishes (1.25).

▶ **Proposition 6.1.** *We have*

$$\mathrm{d_{TV}}\big(N_n, \overline{N}_n\big) = O((\mathrm{L}\, n)^{-(d-1)}) \text{ uniformly for } |a| \le a_n. \tag{6.3}$$

Proof. The key is that $\mathcal{L}(N_n \,|\, E_n^c) = \mathcal{L}(\overline{N}_n)$. It is then easy to check using (1.14) that $\mathrm{d_{TV}}\big(N_n, \overline{N}_n\big)$ is bounded by

$$\frac{\mathbb{P}(E_n)}{1 - \mathbb{P}(E_n)}, \tag{6.4}$$

where E_n is the event that the Poisson process with intensity measure (5.2) has a maximum with ℓ^1-norm larger than $\overline{b}_n$; this bound does not depend on a. A hint that (6.4) is small when n is large is provided by [10, Theorem 1.8] [recall (1.2) above], which, however, deals with the non-Poissonized model. Indeed, such a result follows from Lemma A.3, implying a bound of $O((\mathrm{L}\, n)^{-(d-1)})$ on (6.4) and thereby completing the proof. ◀

7 Discretization

In this section, by discretization we obtain approximations $Z_{n,m}$ to $\overline{N}_n$, where each $Z_{n,m}$ (one random variable for each $m \geq 0$) is a sum of indicator random variables $Z_{n,m,\alpha}$, and prove (for fixed n and a) that (1.26) holds:

$$d_{\mathrm{TV}}\left(\overline{N}_n, Z_{n,m}\right) \to 0 \text{ as } m \to \infty. \tag{7.1}$$

Fix n and define

$$\overline{A}_n := \{\mathbf{x} \succ \mathbf{0} : \underline{b}_n < \|\mathbf{x}\| \leq \overline{b}_n\}. \tag{7.2}$$

Enclose $\overline{A}_n$ in a cube of fixed finite integer side-length, for example, the cube $\left[0, \left\lceil \overline{b}_n \right\rceil\right]^d$. For each integer $m \geq 0$, decompose this cube into sub-cubes $G_{n,m,\alpha}$ of side-length 2^{-m}. (To what extent each $G_{n,m,\alpha}$ contains its boundary will be immaterial.) Let $\overline{\Pi}_n$ denote a Poisson process with intensity measure (6.2), and let $\overline{M}_n$ denote the random measure that counts maxima of $\overline{\Pi}_n$ – so that, in particular,

$$\overline{N}_n = \overline{M}_n(A_n). \tag{7.3}$$

With $\wedge$ denoting minimum, let

$$Z_{n,m,\alpha} := 1 \wedge \overline{M}_n(A_n \cap G_{n,m,\alpha}) \tag{7.4}$$

and

$$I_{n,m} := \{\alpha : A_n \cap G_{n,m,\alpha} \neq \varnothing\}. \tag{7.5}$$

All operations over index α henceforth are taken over the index set $I_{n,m}$ unless otherwise specified. Let

$$Z_{n,m} := \sum_\alpha Z_{n,m,\alpha}. \tag{7.6}$$

Proof of (7.1). Using (1.13) for the first inequality, we have

$$\begin{aligned}
1 - {} & d_{\mathrm{TV}}\left(\overline{N}_n, Z_{n,m}\right) \\
& \geq \mathbb{P}(Z_{n,m} = \overline{N}_n) && (7.7) \\
& \geq \mathbb{P}\left(\cap_\alpha \left\{\overline{\Pi}_n\left(\overline{A}_n \cap G_{n,m,\alpha}\right) \leq 1\right\}\right) \\
& = \exp\left[-\sum_\alpha \overline{\mu}_n\left(\overline{A}_n \cap G_{n,m,\alpha}\right)\right] \prod_\alpha \left[1 + \overline{\mu}_n\left(\overline{A}_n \cap G_{n,m,\alpha}\right)\right] \\
& \to 1 \text{ as } m \to \infty, && (7.8)
\end{aligned}$$

as desired, where the last-step convergence follows by routine application of [7, Lemma preceding Theorem 7.1.2]. ◀

8 A sequence of Poisson distributions indexed by m

In this section, recalling the notation $\lambda_{n,m} = \mathbb{E}\, Z_{n,m}$, we prove (for each fixed n and a) that (1.28) holds:

$$d_{\mathrm{TV}}\left(\mathrm{Po}(\lambda_{n,m}), \mathrm{Po}(\mathbb{E}\,\overline{N}_n)\right) \to 0 \text{ as } m \to \infty. \tag{8.1}$$

Proof. From the argument at (7.7)–(7.8) we see convergence in probability:

$$Z_{n,m} \xrightarrow{\mathrm{P}} \overline{N}_n \text{ as } m \to \infty.$$

Noting from (7.4) that $Z_{n,m}$ is monotonically nondecreasing in m, the same convergence holds almost surely (see [6, Problem 20.22(a)]). Therefore, by the monotone convergence theorem, we also have L^1-convergence and hence

$$\lambda_{n,m} \uparrow \mathbb{E}\,\overline{N}_n \text{ as } m \uparrow \infty. \tag{8.2}$$

Thus, utilizing the semigroup property of Poisson distributions under convolution and (1.13) at the initial step,

$$\begin{aligned} \mathrm{d_{TV}}\big(\mathrm{Po}(\lambda_{n,m}), \mathrm{Po}(\mathbb{E}\,\overline{N}_n)\big) &\le \mathbb{P}\left(\mathrm{Po}\big(\mathbb{E}\,\overline{N}_n - \lambda_{n,m}\big) \neq 0\right) &(8.3)\\ &= 1 - \exp\big[-\big(\mathbb{E}\,\overline{N}_n - \lambda_{n,m}\big)\big] &(8.4)\\ &\to 0 \text{ as } m \to \infty. \end{aligned}$$

◀

9 A sequence of Poisson distributions indexed by n

In this section we reference a proof in the full paper [9, Section 9] that (1.29) holds:

$$\mathrm{d_{TV}}\big(\mathrm{Po}(\mathbb{E}\,\overline{N}_n), \mathrm{Po}(\lambda)\big) = O((\mathrm{L}\,n)^{-1}(\mathrm{L}_2\,n)^{1/2}) \text{ uniformly for } |a| \le a_n. \tag{9.1}$$

By the same sort of argument as at (8.3)–(8.4), we need only prove the following lemma.

▶ **Lemma 9.1.** *As $n \to \infty$, we have*

$$\mathbb{E}\,\overline{N}_n = \lambda + O((\mathrm{L}\,n)^{-1}(\mathrm{L}_2\,n)^{1/2}) \text{ uniformly for } |a| \le a_n. \tag{9.2}$$

Proof. See the full paper [9, Section 9]. ◀

10 Poisson approximation for $Z_{n,m}$

In this section, centrally important to the proof of Theorem 1.2, we employ the technique of [2] to outline a proof that (1.27) holds, namely, that

$$\limsup_{m\to\infty} \mathrm{d_{TV}}(Z_{n,m}, \mathrm{Po}(\lambda_{n,m})) = O\left((\mathrm{L}_2\,n)^{-\left(d-\frac{3}{2}\right)}(\mathrm{L}_3\,n)^{-1}\right) \text{ uniformly for } |a| \le a_n \tag{10.1}$$

as $n \to \infty$, where $\lambda_{n,m} := \mathbb{E}\,Z_{n,m}$. We will make a reasonable effort to match the notation used in [2], but, since we have already used b_n for the boundary (1.10), we will use fraktur font to write $\mathfrak{b}_1, \mathfrak{b}_2, \mathfrak{b}_3$ for the three key quantities b_1, b_2, b_3 in [2].

The goal, then, of this section is to establish a Poisson approximation, with error bound, for $Z_{n,m}$ defined at (7.6), with $Z_{n,m,\alpha}$ defined at (7.4). Treating the sub-cube $G_{n,m,\alpha}$ as a closed cube, let $\underline{\mathbf{g}}_{n,m,\alpha}$ denote its unique minimum (i.e., "southwestern-most point"). Note first that if we make, for each $\alpha \in I_{n,m}$ (with $I_{n,m}$ defined at (7.5)), the definition

$$B_{n,m,\alpha} := \{\gamma \in I_{n,m} : \|\underline{\mathbf{g}}_{n,m,\alpha} \vee \underline{\mathbf{g}}_{n,m,\gamma}\| < \overline{b}_n\}, \tag{10.2}$$

where $\vee$ denotes coordinate-wise maximum, then

for every $\alpha \in I_{n,m}$: $Z_{n,m,\alpha}$ is independent of $\sigma\langle Z_{n,m,\gamma} : \gamma \in I_{n,m} - B_{n,m,\alpha}\rangle$, so $\mathfrak{b}_3 = 0$ in the notation of [2]. In [2], see Theorem 1 and the penultimate full paragraph on p. 12 of [2], from which we also find (noting the factor-of-2 difference in their definition of total variation distance) that

$$\mathrm{d_{TV}}(Z_{n,m}, \mathrm{Po}(\lambda_{n,m})) \le \mathfrak{b}_1 + \mathfrak{b}_2 = \mathbb{E}\, Z_{n,m}^2 - \mathbb{E}\, \mathrm{Po}(\lambda_{n,m})^2 + 2\mathfrak{b}_1. \tag{10.3}$$

With

$$p_{n,m,\alpha} := \mathbb{P}(Z_{n,m,\alpha} = 1) \tag{10.4}$$

and

$$p_{n,m,\alpha,\gamma} := \mathbb{E}(Z_{n,m,\alpha} Z_{n,m,\gamma}), \tag{10.5}$$

the notation in (10.3) is that

$$\begin{aligned} \mathfrak{b}_1 \equiv \mathfrak{b}_1(n,m) &= \sum_{\alpha \in I_{n,m}} \sum_{\gamma \in B_{n,m,\alpha}} p_{n,m,\alpha}\, p_{n,m,\gamma}, \\ \mathfrak{b}_2 \equiv \mathfrak{b}_2(n,m) &= \sum_{\alpha \in I_{n,m}} \sum_{\gamma \ne \alpha :\, \alpha \in B_{n,m,\alpha}} p_{n,m,\alpha,\gamma}. \end{aligned} \tag{10.6}$$

Note that if we use the second of the two expressions for the upper bound in (10.3), then we need treat only $\mathfrak{b}_1$, not $\mathfrak{b}_2$.

For the first of the three terms in that bound, we claim that

$$\mathbb{E}\, Z_{n,m}^2 \to \mathbb{E}\, \overline{N}_n^2 \text{ as } m \to \infty. \tag{10.7}$$

Indeed, by the same argument leading to (8.2), $Z_{n,m}$ converges in L^2 to $\overline{N}_n$, and hence (by the second assertion in [7, Theorem 4.5.1] with $r = 2$) (10.7) holds.

The (subtracted) second of the three terms is

$$\mathbb{E}\, \mathrm{Po}(\lambda_{n,m})^2 = \lambda_{n,m} + \lambda_{n,m}^2 \to \mathbb{E}\, \overline{N}_n + \left(\mathbb{E}\, \overline{N}_n\right)^2 \text{ as } m \to \infty, \tag{10.8}$$

where the convergence here follows from (8.2). Therefore, by (10.7)–(10.8),

$$\mathbb{E}\, Z_{n,m}^2 - \mathbb{E}\, \mathrm{Po}(\lambda_{n,m})^2 \to \mathrm{Var}\, \overline{N}_n - \mathbb{E}\, \overline{N}_n \text{ as } m \to \infty. \tag{10.9}$$

For the proof of (10.1), we need also Lemma A.4, which establishes

$$\mathrm{Var}\, \overline{N}_n = \mathbb{E}\, \overline{N}_n + O\left((\mathrm{L}_2\, n)^{-\left(d - \frac{3}{2}\right)} (\mathrm{L}_3\, n)^{-1} \right) \text{ uniformly for } |a| \le a_n, \tag{10.10}$$

and the following lemma.

▶ **Lemma 10.1.** *As $n \to \infty$ we have*

$$\limsup_{m \to \infty} \mathfrak{b}_1(n,m) = O((\mathrm{L}\, n)^{-(d-1)} (\mathrm{L}_2\, n)^d (\mathrm{L}_3\, n)^{-2}) \text{ uniformly for } |a| \le a_n. \tag{10.11}$$

Proof. See the full paper [9, Section 10.1]. ◀

We are now ready for the proof of (10.1).

Proof of (10.1)**.** Simply combine (10.3) and (10.9)–(10.11). ◀

11 Distribution of the smallest maximum

▶ Remark 11.1. (a) As already discussed in Section 1, if $\mathbf{Y}_n$ denotes a "generic maximum" at epoch n, then $\|\mathbf{Y}_n\|$ and $\mathbf{U}_n := \|\mathbf{Y}_n\|^{-1}\mathbf{Y}_n$ are independent, and $\mathbf{U}_n$ is uniformly distributed on $\mathcal{S}_{d-1}$.

(b) Likewise, if the *largest maximum* $\boldsymbol{\lambda}_n$ is defined as the (almost surely unique) maximum at epoch n with the largest ℓ^1-norm, then it easy to check that $\|\boldsymbol{\lambda}_n\|$ [$\equiv F_n^+$ in the notation of [10] and [12] and (1.2)] and $\mathbf{V}_n := \|\boldsymbol{\lambda}_n\|^{-1}\boldsymbol{\lambda}_n$ are independent, and $\mathbf{V}_n$ is uniformly distributed on $\mathcal{S}_{d-1}$.

(c) There is no such simple result (for finite n) concerning the *smallest maximum* $\boldsymbol{\sigma}_n$, defined as the (almost surely unique) maximum at epoch n with the smallest ℓ^1-norm (equal to φ_n). Indeed, using the fact that $\|\mathbf{X}\|$ is distributed Gamma$(d,1)$, it is not hard to see that $\boldsymbol{\sigma}_2$ has exactly the density

$$\begin{aligned}\mathbb{P}(\boldsymbol{\sigma}_2 \in \mathrm{d}\mathbf{s}) &= 2\,\mathbb{P}(\mathbf{X} \in \mathrm{d}\mathbf{s})\left[\mathbb{P}(\mathbf{X} \prec \mathbf{s}) + \mathbb{P}(\|\mathbf{X}\| > \|\mathbf{s}\|) - \mathbb{P}(\mathbf{X} \succ \mathbf{s})\right]\\ &= 2e^{-\|\mathbf{s}\|}\left[\prod_{j=1}^{d}(1-e^{-s_j}) + e^{-\|\mathbf{s}\|}\sum_{j=1}^{d-1}\frac{\|\mathbf{s}\|^j}{j!}\right]\mathrm{d}\mathbf{s}\end{aligned}$$

for $\mathbf{s} \in (0,\infty)^d$. The continuous density here is not a function of $\|\mathbf{s}\|$, and so $\varphi_2 = \|\boldsymbol{\sigma}_2\|$ and $\|\boldsymbol{\sigma}_2\|^{-1}\boldsymbol{\sigma}_2$ are not independent.

Notwithstanding Remark 11.1(c), we make the following conjecture.

▶ **Conjecture 11.2.** *Define $\boldsymbol{\sigma}_n$ to be the (almost surely unique) maximum of $\{\mathbf{X}^{(1)},\dots,\mathbf{X}^{(n)}\}$ with minimum ℓ^1-norm. Then φ_n and $\varphi_n^{-1}\boldsymbol{\sigma}_n$ are asymptotically independent, and $\varphi_n^{-1}\boldsymbol{\sigma}_n$ converges in distribution to the uniform distribution on $\mathcal{S}_{d-1}$.*

In light of the heuristic argument spanning (1.20)–(1.22) in Section 1, it is quite believable that, for each fixed $a \in \mathbb{R}$ and for large n, the point process of vectors $\mathbf{x} \in (0,\infty)^d$ with $\underline{b}_n < \|\mathbf{x}\| \le b_n^*$ that are maxima of $\{\mathbf{X}^{(1)},\dots,\mathbf{X}^{(n)}\}$ at epoch n is, in some suitable sense, well approximated by a Poisson point process, call it $\mathcal{P}_n$, with intensity measure

$$\nu_n(\mathrm{d}\mathbf{x}) := ne^{-\|\mathbf{x}\|}\exp(-ne^{-\|\mathbf{x}\|})\mathbf{1}(\mathbf{x} \succ \mathbf{0},\, \underline{b}_n < \|\mathbf{x}\| \le b_n^*)\,\mathrm{d}\mathbf{x}.$$

Let H_n denote the event that $\mathcal{P}_n((0,\infty)^d) \ge 1$, and let $\mathbf{U}$ be independent of $\mathcal{P}_n$ and uniformly distributed over $\mathcal{S}_{d-1}$. Over the event H_n, let $\widehat{\boldsymbol{\sigma}}_n$ denote the (almost surely unique) point of $\mathcal{P}_n$ with minimum ℓ^1-norm, say $\widehat{\varphi}_n$; and over H_n^c, let $\widehat{\varphi}_n := 1$ and $\widehat{\boldsymbol{\sigma}}_n := \mathbf{U}$. Then, for Borel $A \subseteq \mathcal{S}_{d-1}$, we would expect to have

$$\mathbb{P}(\varphi_n \le b_n^*,\, \varphi_n^{-1}\boldsymbol{\sigma}_n \in A) \approx \mathbb{P}(H_n \cap \{\widehat{\varphi}_n^{-1}\widehat{\boldsymbol{\sigma}}_n \in A\}) \tag{11.1}$$
$$= \mathbb{P}(H_n)\,\mathbb{P}(\mathbf{U} \in A) \tag{11.2}$$

[indeed, we know that the equality of the right side of (11.1) and the expression in (11.2) is true exactly!] and, in particular,

$$\mathbb{P}(\varphi_n \le b_n^*) \approx \mathbb{P}(H_n). \tag{11.3}$$

Combining (11.2)–(11.3) would then give Conjecture 11.2. Perhaps the point-process approximation by $\mathcal{P}_n$ we have surmised here can be established rigorously using the techniques of [4, Chapter 10], which include [2, Theorem 2].

References

1 Barry C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *Records*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1998. A Wiley-Interscience Publication. doi:10.1002/9781118150412.

2 R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: The Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989. URL: http://links.jstor.org/sici?sici=0091-1798(198901)17:1<9:TMSFPA>2.0.CO;2-X&origin=MSN.

3 Zhi-Dong Bai, Luc Devroye, Hsien-Kuei Hwang, and Tsung-Hsi Tsai. Maxima in hypercubes. *Random Structures Algorithms*, 27(3):290–309, 2005. doi:10.1002/rsa.20053.

4 A. D. Barbour, Lars Holst, and Svante Janson. *Poisson Approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York, 1992. Oxford Science Publications.

5 Yu. Baryshnikov. Supporting-points processes and some of their applications. *Probab. Theory Related Fields*, 117(2):163–182, 2000. doi:10.1007/s004400050002.

6 Patrick Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, anniversary edition edition, 2012. With a foreword by Steve Lalley and a brief biography of Billingsley by Steve Koppes.

7 Kai Lai Chung. *A Course in Probability Theory*. Academic Press, Inc., San Diego, CA, third edition, 2001.

8 James Allen Fill. Breaking multivariate records. *Electron. J. Probab.*, 28:1–27, 2023. doi:10.1214/23-EJP968.

9 James Allen Fill. A new fine-scale Berry–Esseen-type Gumbel-limit theorem for multivariate maxima [full-paper version of this extended abstract], 2026. arXiv:2601.18170.

10 James Allen Fill and Daniel Q. Naiman. The Pareto record frontier. *Electron. J. Probab.*, 25:Paper No. 92, 24 pages, 2020. doi:10.1214/20-ejp492.

11 James Allen Fill, Daniel Q. Naiman, and Ao Sun. Sharpened localization of the trailing point of the Pareto record frontier [extended abstract]. In Cécile Mailler and Sebastian Wild, editors, *35th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2024)*, volume 302 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 28:1–28:21, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.AofA.2024.28.

12 James Allen Fill, Daniel Q. Naiman, and Ao Sun. Sharpened localization of the trailing point of the Pareto record frontier. *Theoret. Comput. Sci.*, 1070:Paper No. 115822, 21, 2026. doi:10.1016/j.tcs.2026.115822.

13 J. Kiefer. Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics*, pages 227–244, 1972.

14 Günter Last and Mathew Penrose. *Lectures on the Poisson Process*, volume 7 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2018.

15 Sidney I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original.

16 Ao Sun. *Studies in Multivariate Pareto Records*. PhD thesis, The Johns Hopkins University, 2025.

A Collected results from Fill et al. [12], and closely related results

For the reader's convenience, this appendix collects results from [12], and establishes closely related results, used in the present work.

Our first lemma is used in the proof of Proposition 2.1.

▶ **Lemma A.1** ([12], Propositions 2.3 and 3.2). *If* $\left(\tilde{b}_n\right)$ *is any sequence satisfying*

$$\tilde{b}_n := \mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}\, \tilde{c}_n \text{ with } \tilde{c}_n > 0 \text{ and } \tilde{c}_n = \Theta(1), \tag{A.1}$$

then as $n \to \infty$ *we have*

$$\mathbb{P}\big(\varphi_n \le \tilde{b}_n\big) \le \mathbb{E}\, \rho_n\big(\tilde{b}_n\big) = (1 + o(1))\, \frac{1}{(d-1)!} (\mathrm{L}\, n)^{d-1-\tilde{c}_n}$$

and

$$\mathbb{P}\big(\varphi_n \ge \tilde{b}_n\big) \le (1 + o(1))\, (d-1)! (\mathrm{L}\, n)^{-(d-1-\tilde{c}_n)}. \qquad ◀$$

Define

$$J_j(x) := \int_x^\infty (\mathrm{L}\, z)^j e^{-z}\, \mathrm{d}z \tag{A.2}$$

and note [12, (A.3)] that

$$J_j(x) \sim (\mathrm{L}\, x)^j e^{-x} \text{ as } x \to \infty. \tag{A.3}$$

The next lemma is used in the proof of (4.1).

▶ **Lemma A.2.** *We have*

$$\mathbb{E}\, \rho_n(\underline{b}_n) = O((\mathrm{L}\, n)^{d-1-\omega_n}) = o((\mathrm{L}\, n)^{-(d-1)}).$$

Proof. To bound $\mathbb{E}\, \rho_n(\underline{b}_n)$, follow along the calculations in the proof of Lemma 2.1 in [11, Appendix A]. One finds (compare [11, (A.6)])

$$\mathbb{E}\, \rho_n(\underline{b}_n) \le \frac{1 + (n-1)^{-1}}{(d-1)!} \sum_{j=0}^{d-1} (-1)^j \binom{d-1}{j} (\mathrm{L}\, n)^{d-1-j} [J_j((n-1)e^{-\underline{b}_n}) - J_j(n)], \tag{A.4}$$

with, by (A.3), $J_j(n) \sim (\mathrm{L}\, n)^j e^{-n}$ and

$$\begin{aligned} J_j((n-1)e^{-\underline{b}_n}) &\sim [\mathrm{L}(n-1) - \underline{b}_n]^j \exp[-(n-1)e^{-\underline{b}_n}] \\ &\sim [\mathrm{L}\, n - \underline{b}_n]^j \exp[-(n-1)e^{-\underline{b}_n}] \\ &= [\mathrm{L}_3\, n + \mathrm{L}(d-1) + \mathrm{L}\, 3]^j \exp[-(n-1)n^{-1}(\mathrm{L}_2\, n)\, 3(d-1)] \\ &\sim (\mathrm{L}_3\, n)^j \exp[-(n-1)n^{-1}(\mathrm{L}_2\, n)\, 3(d-1)] \qquad (A.5) \\ &\sim (\mathrm{L}_3\, n)^j (\mathrm{L}\, n)^{-3(d-1)}. \qquad (A.6) \end{aligned}$$

Thus the $j = 0$ term in (A.4) predominates and yields the desired result. ◀

The next lemma is used in the proof of Proposition 6.1.

▶ **Lemma A.3.** *Let* E_n *be the event that a Poisson process with intensity measure*

$$\mu_n(\mathrm{d}\mathbf{x}) := ne^{-\|\mathbf{x}\|} \mathbf{1}(\mathbf{x} \succ \mathbf{0})\, \mathrm{d}\mathbf{x}$$

has a maximum with ℓ^1*-norm larger than* $\overline{b}_n = \mathrm{L}\, n + 2(d-1)\, \mathrm{L}_2\, n$*. Then*

$$\mathbb{P}(E_n) = O((\mathrm{L}\, n)^{-(d-1)}).$$

Proof. Clearly E_n equals the event that there is at least one point of the Poisson process with ℓ^1-norm larger than $\bar{b}_n$, and the number of Poisson points satisfying this "large ℓ^1-norm" condition is distributed Poisson with parameter

$$\int_{\mathbf{x}\succ\mathbf{0}:\|\mathbf{x}\|>\bar{b}_n} \mu_n(\mathrm{d}\mathbf{x}) = n\,\mathbb{P}(\mathrm{Gamma}(d,1) > \bar{b}_n) = n\,\mathbb{P}(\mathrm{Po}(\bar{b}_n) < d)$$
$$\sim n\,\mathbb{P}(\mathrm{Po}(\bar{b}_n) = d-1) \sim \frac{1}{(d-1)!}(\mathrm{L}\,n)^{-(d-1)}.$$

Thus

$$\mathbb{P}(E_n) = 1 - \exp\left[-(1+o(1))\frac{1}{(d-1)!}(\mathrm{L}\,n)^{-(d-1)}\right]$$
$$\sim \frac{1}{(d-1)!}(\mathrm{L}\,n)^{-(d-1)} = O((\mathrm{L}\,n)^{-(d-1)}),$$

as claimed. ◀

The next lemma is used in the proof of (10.1). Given the convergence of $\operatorname{Var}\rho_n(b_n^*) - \mathbb{E}\,\rho_n(b_n^*)$ to 0 established for fixed a by (1.18) and [12, Lemma 3.1 and Remark B.1], it is not surprising that $\operatorname{Var}\overline{N}_n - \mathbb{E}\,\overline{N}_n = o(1)$, but it is nontrivial to prove this (with suitable uniformity in a) and to further quantify $o(1)$.

▶ **Lemma A.4.** *As $n \to \infty$, uniformly for $|a| \le a_n$ we have*

$$\operatorname{Var}\overline{N}_n - \mathbb{E}\,\overline{N}_n = O\left((\mathrm{L}_2\,n)^{-\left(d-\frac{3}{2}\right)}(\mathrm{L}_3\,n)^{-1}\right). \tag{A.7}$$

Proof. To treat $\operatorname{Var}\overline{N}_n$ we follow the approach to $\operatorname{Var}\rho_n(b_n)$ in [12] and start with the second factorial moment of $\overline{N}_n$. Given $\mathbf{x} \succ \mathbf{0}$, let $O^+_{\mathbf{x}} := \{\mathbf{y} : \mathbf{y} \succ \mathbf{x}\}$ denote the positive orthant in $\mathbb{R}^d$, translated by $\mathbf{x}$. Letting $\mathbf{x} \parallel \mathbf{y}$ denote incomparability of $\mathbf{x}$ and $\mathbf{y}$ with respect to $\prec$ (more precisely, that neither $\mathbf{x}$ weakly dominates $\mathbf{y}$ coordinate-by-coordinate nor vice versa), we have, by the Mecke equation [14, Theorem 4.4],

$$\begin{aligned}
&\mathbb{E}\left[\overline{N}_n\left(\overline{N}_n - 1\right)\right] \\
&= n^2 \int_{\mathbf{x},\mathbf{y}\in A_n:\,\mathbf{x}\parallel\mathbf{y}} e^{-(\|\mathbf{x}\|+\|\mathbf{y}\|)}\,\mathbb{P}\left(\overline{\Pi}_n(O^+_{\mathbf{x}}) = 0 \text{ and } \overline{\Pi}_n(O^+_{\mathbf{y}}) = 0\right)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \\
&= \left(\mathbb{E}\,\overline{N}_n\right)^2 - I_{n,0} + I_n,
\end{aligned} \tag{A.8}$$

where

$$\begin{aligned}
I_{n,0} &:= n^2 \int_{\mathbf{x},\mathbf{y}\in A_n:\,\mathbf{x}\not\parallel\mathbf{y}} e^{-(\|\mathbf{x}\|+\|\mathbf{y}\|)}\,\mathbb{P}\left(\overline{\Pi}_n(O^+_{\mathbf{x}}) = 0\right)\mathbb{P}\left(\overline{\Pi}_n(O^+_{\mathbf{y}}) = 0\right)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}, \\
I_n &:= n^2 \int_{\mathbf{x},\mathbf{y}\in A_n:\,\mathbf{x}\parallel\mathbf{y}} e^{-(\|\mathbf{x}\|+\|\mathbf{y}\|)}\,\left[\mathbb{P}\left(\overline{\Pi}_n(O^+_{\mathbf{x}}) = 0 \text{ and } \overline{\Pi}_n(O^+_{\mathbf{y}}) = 0\right)\right. \\
&\qquad\qquad \left. - \,\mathbb{P}\left(\overline{\Pi}_n(O^+_{\mathbf{x}}) = 0\right)\mathbb{P}\left(\overline{\Pi}_n(O^+_{\mathbf{y}}) = 0\right)\right]\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}.
\end{aligned}$$

Clearly $I_{n,0} \ge 0$, and it is easy to see that the integrand factor in square brackets for I_n is nonnegative, so $I_n \ge 0$, too. From (A.8) it follows that

$$\operatorname{Var}\overline{N}_n - \mathbb{E}\,\overline{N}_n = \mathbb{E}\left[\overline{N}_n\left(\overline{N}_n - 1\right)\right] - \left(\mathbb{E}\,\overline{N}_n\right)^2 = -I_{n,0} + I_n.$$

We will complete the proof of Lemma A.4 by proving that

$$I_{n,0} = O\Big((\mathrm{L}\,n)^{-(d-1)}(\mathrm{L}_2\,n)^2(\mathrm{L}_3\,n)^{-2}\Big) \tag{A.9}$$
$$= o\left((\mathrm{L}_2\,n)^{-\left(d-\frac{3}{2}\right)}(\mathrm{L}_3\,n)^{-1}\right), \text{ uniformly for } |a| \le a_n$$

and

$$I_n = O\left((\mathrm{L}_2\,n)^{-\left(d-\frac{3}{2}\right)}(\mathrm{L}_3\,n)^{-1}\right) \text{ uniformly for } |a| \le a_n. \tag{A.10}$$

To prove (A.9), we begin by noting from [9, (9.8)] that for $\mathbf{x} \in A_n$ we have

$$\mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{x}}) = 0\big) = \exp\Big[-ne^{-\|\mathbf{x}\|}(1-u_n(\|\mathbf{x}\|))\Big] \text{ for } \mathbf{x} \in A_n \tag{A.11}$$

with

$$u_n(w) \equiv u_{n,a}(w)$$
$$= O\Big((\mathrm{L}\,n)^{-2(d-1)}(\mathrm{L}_2\,n)^{d-2}\Big) \text{ uniformly for } w \in (\underline{b}_n, b_n] \text{ and } |a| \le a_n,$$

so that

$$\begin{aligned} I_{n,0} &= 2n^2 \int_{\mathbf{x},\mathbf{y}\in A_n:\,\mathbf{x}\prec\mathbf{y}} e^{-(\|\mathbf{x}\|+\|\mathbf{y}\|)} \exp\Big[-ne^{-\|\mathbf{x}\|}(1-u_n(\|\mathbf{x}\|))\Big] \\ &\qquad \times \exp\Big[-ne^{-\|\mathbf{y}\|}(1-u_n(\|\mathbf{y}\|))\Big]\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \\ &= 2n^2 \int_{\mathbf{x}\in A_n}\int_{\mathbf{z}\succ\mathbf{0}:\,\|\mathbf{z}\|\le b_n-\|\mathbf{x}\|} e^{-\|\mathbf{x}\|}e^{-(\|\mathbf{x}\|)+\|\mathbf{z}\|)} \exp\Big[-ne^{-\|\mathbf{x}\|}(1-u_n(\|\mathbf{x}\|))\Big] \\ &\qquad \times \exp\Big[-ne^{-(\|\mathbf{x}\|+\|\mathbf{z}\|)}(1-u_n(\|\mathbf{x}\|+\|\mathbf{z}\|))\Big]\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{x} \\ &= 2n^2 \int_{\underline{b}_n}^{b_n}\int_0^{b_n-x} \frac{x^{d-1}}{(d-1)!}e^{-2x}\frac{\zeta^{d-1}}{(d-1)!}e^{-\zeta}\exp\big[-ne^{-x}(1-u_n(x))\big] \\ &\qquad \times \exp\big[-ne^{-x}e^{-\zeta}(1-u_n(x+\zeta))\big]\,\mathrm{d}\zeta\,\mathrm{d}x. \end{aligned}$$

From here, change variables using $x = \mathrm{L}\,n - \mathrm{L}\,y$ and $\zeta = -\mathrm{L}\,z$ to obtain, with

$$\widehat{u}_n := \sup_{|a|\le a_n,\,y\in(\underline{b}_n,b_n]} |u_{n,a}(y)| = O\Big((\mathrm{L}\,n)^{-2(d-1)}(\mathrm{L}_2\,n)^{d-2}\Big), \tag{A.12}$$

that

$$\begin{aligned} I_{n,0} &= \frac{2}{[(d-1)!]^2}\int_{\beta_n}^{\underline{\beta}_n}\int_{\beta_n/y}^1 (\mathrm{L}\,n-\mathrm{L}\,y)^{d-1}y\,(-\mathrm{L}\,z)^{d-1} \\ &\qquad \times \exp[-y(1-u_n(\mathrm{L}\,n-\mathrm{L}\,y))]\exp[-yz(1-u_n(\mathrm{L}\,n-\mathrm{L}\,y-\mathrm{L}\,z))]\,\mathrm{d}z\,\mathrm{d}y \\ &\le \frac{2(\mathrm{L}\,n)^{d-1}\exp[-(1-\widehat{u}_n)\beta_n]}{[(d-1)!]^2}\int_{\beta_n}^{\underline{\beta}_n} y\exp[-(1-\widehat{u}_n)y]\int_{\beta_n/y}^1(-\mathrm{L}\,z)^{d-1}\,\mathrm{d}z\,\mathrm{d}y. \end{aligned}$$

But for $y \in [\beta_n, \infty)$ we have

$$\int_{\beta_n/y}^1 (-\mathrm{L}\,z)^{d-1}\,\mathrm{d}z = \int_0^{\mathrm{L}\,y-\mathrm{L}\,\beta_n} w^{d-1}e^{-w}\,\mathrm{d}w \le (d-1)!,$$

so, using (A.12) at the asymptotic equivalence to follow, uniformly for $|a| \le a_n$ we have

$$\begin{aligned} I_{n,0} &\le \frac{2(\mathrm{L}\, n)^{d-1} \exp[-(1-\widehat{u}_n)\beta_n]}{(d-1)!} \int_{\beta_n}^{\infty} y \exp[-(1-\widehat{u}_n)y] \, \mathrm{d}y \\ &\sim \frac{2}{(d-1)!} (\mathrm{L}\, n)^{d-1} \beta_n e^{-2\beta_n} = O\Big((\mathrm{L}\, n)^{-(d-1)} (\mathrm{L}_2\, n)^2 (\mathrm{L}_3\, n)^{-2}\Big), \end{aligned}$$

and (A.9) is established.

Now we turn our attention to I_n and the proof of (A.10). First note that for any $\mathbf{x}, \mathbf{y} \in A_n$ we have

$$\begin{aligned} \mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{y}}) = 0 \,|\, \overline{\Pi}_n(O^+_{\mathbf{x}}) = 0\big) &= \mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{y}}) = 0 \,|\, \overline{\Pi}_n(O^+_{\mathbf{x}\vee\mathbf{y}}) = 0\big) \\ &= \frac{\mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{y}}) = 0\big)}{\mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{x}\vee\mathbf{y}}) = 0\big)}; \end{aligned}$$

thus

$$\begin{aligned} I_n = n^2 \int_{\mathbf{x},\mathbf{y} \in A_n :\, \mathbf{x} \| \mathbf{y}} & e^{-(\|\mathbf{x}\|+\|\mathbf{y}\|)} \mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{x}}) = 0\big)\, \mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{y}}) = 0\big) \\ & \times \left[\frac{1}{\mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{x}\vee\mathbf{y}}) = 0\big)} - 1\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}. \end{aligned}$$

For the first two of the three probabilities appearing in this last expression we use (A.11); for the third, we note simply that

$$\mathbb{P}\big(\overline{\Pi}_n(O^+_{\mathbf{x}\vee\mathbf{y}}) = 0\big) \ge \exp(-n e^{-\|\mathbf{x}\vee\mathbf{y}\|}).$$

We then find

$$\begin{aligned} I_n \le n^2 \int_{\mathbf{x},\mathbf{y} \in A_n :\, \mathbf{x} \| \mathbf{y}} & e^{-(\|\mathbf{x}\|+\|\mathbf{y}\|)} \exp\Big[-n e^{-\|\mathbf{x}\|}(1 - u_n(\|\mathbf{x}\|))\Big] \\ & \times \exp\Big[-n e^{-\|\mathbf{y}\|}(1 - u_n(\|\mathbf{y}\|))\Big] \Big[\exp(n e^{-\|\mathbf{x}\vee\mathbf{y}\|}) - 1\Big] \, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}. \end{aligned}$$

Here, the integrand factors $\exp\big[n e^{-\|\mathbf{x}\|} u_n(\|\mathbf{x}\|)\big]$ and $\exp\big[n e^{-\|\mathbf{y}\|} u_n(\|\mathbf{y}\|)\big]$ are each bounded by

$$\exp\Big(\widehat{u}_n \underline{\beta}_n\Big) = \exp\Big[O\Big((\mathrm{L}\, n)^{-2(d-1)} (\mathrm{L}_2\, n)^{d-1}\Big)\Big] = 1 + o(1).$$

Therefore I_n is bounded by $(1 + o(1))$ times

$$\begin{aligned} & n^2 \int_{\mathbf{x},\mathbf{y} \in A_n :\, \mathbf{x} \| \mathbf{y}} e^{-(\|\mathbf{x}\|+\|\mathbf{y}\|)} \exp\Big(-n e^{-\|\mathbf{x}\|}\Big) \exp\Big(-n e^{-\|\mathbf{y}\|}\Big) \Big[\exp(n e^{-\|\mathbf{x}\vee\mathbf{y}\|}) - 1\Big] \, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y} \\ & \quad \le n^2 \int_{\mathbf{x},\mathbf{y} \succ 0 :\, \mathbf{x} \| \mathbf{y},\, \|\mathbf{x}\| \le b_n,\, \|\mathbf{y}\| \le b_n} e^{-(\|\mathbf{x}\|+\|\mathbf{y}\|)} \exp\Big(-n e^{-\|\mathbf{x}\|}\Big) \\ & \qquad \times \exp\Big(-n e^{-\|\mathbf{y}\|}\Big) \Big[\exp(n e^{-\|\mathbf{x}\vee\mathbf{y}\|}) - 1\Big] \, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}. \qquad \text{(A.13)} \end{aligned}$$

After a simple change of variables, (A.13) is seen to agree with the expression I_n appearing at [12, (B.3)], except that the factor $(n+2)(n+1)$ appearing there becomes n^2 at (A.13) and the restrictions $\mathbf{x}_\times \ge e^{-b_{n+2}}$ and $\mathbf{y}_\times \ge e^{-b_{n+2}}$ on the integral at [12, (B.3)] become our $\mathbf{x}_\times \ge e^{-b_n}$ and $\mathbf{y}_\times \ge e^{-b_n}$. (Here $\mathbf{z}_\times := \prod_{j=1}^d z_j$ for any vector $\mathbf{z}$.) We can decompose our I_n as $\sum_{k=1}^{d-1} \binom{d}{k} I_{n,k}$ in the same fashion as done at [12, (B.7)], and we similarly find, uniformly for $|a| \le a_n$, that

$$I_{n,k} = O((\mathrm{L}\, n)^{d-1} \beta_n^{-(d-1)} e^{-\beta_n}) = O\left((\mathrm{L}_2\, n)^{-\left(d - \frac{3}{2}\right)} (\mathrm{L}_3\, n)^{-1} \right).$$

This establishes (A.10) and completes the proof of the lemma. ◀