Sharpened Localization of the Trailing Point of the Pareto Record Frontier

James Allen Fill[1]
Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD, USA

Daniel Q. Naiman
Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD, USA

Ao Sun
Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD, USA

Abstract

For $d \geq 2$ and i.i.d. d-dimensional observations $X^{(1)}, X^{(2)}, \dots$ with independent Exponential(1) coordinates, we revisit the study by Fill and Naiman (*Electron. J. Probab.*, 25:Paper No. 92, 24 pp., 2020) of the boundary (relative to the closed positive orthant), or "frontier", F_n of the closed Pareto record-setting (RS) region $\mathrm{RS}_n := \{0 \leq x \in \mathbb{R}^d : x \not\prec X^{(i)} \text{ for all } 1 \leq i \leq n\}$ at time n, where $0 \leq x$ means that $0 \leq x_j$ for $1 \leq j \leq d$ and $x \prec y$ means that $x_j < y_j$ for $1 \leq j \leq d$. With $x_+ := \sum_{j=1}^d x_j = \|x\|_1$, let

$$F_n^- := \min\{x_+ : x \in F_n\} \quad \text{and} \quad F_n^+ := \max\{x_+ : x \in F_n\}.$$

Almost surely, there are for each n unique vectors $\lambda_n \in F_n$ and $\tau_n \in F_n$ such that $F_n^+ = (\lambda_n)_+$ and $F_n^- = (\tau_n)_+$; we refer to λ_n and τ_n as the *leading* and *trailing* points, respectively, of the frontier. Fill and Naiman provided rather sharp information about the typical and almost sure behavior of F^+, but somewhat crude information about F^-, namely, that for any $\varepsilon > 0$ and $c_n \to \infty$ we have

$$\mathbb{P}(F_n^- - \ln n \in (-(2+\varepsilon) \ln \ln \ln n, c_n)) \to 1$$

(describing typical behavior) and almost surely

$$\limsup \frac{F_n^- - \ln n}{\ln \ln n} \leq 0 \text{ and } \liminf \frac{F_n^- - \ln n}{\ln \ln \ln n} \in [-2, -1].$$

In this extended abstract we use the theory of *generators* (minima of F_n) together with the first- and second-moment methods to improve considerably the trailing-point location results to

$$F_n^- - (\ln n - \ln \ln \ln n) \xrightarrow{\mathrm{P}} -\ln(d-1)$$

(describing typical behavior) and, for $d \geq 3$, almost surely

$$\begin{aligned} &\limsup [F_n^- - (\ln n - \ln \ln \ln n)] \leq -\ln(d-2) + \ln 2 \\ \text{and } &\liminf [F_n^- - (\ln n - \ln \ln \ln n)] \geq -\ln d - \ln 2. \end{aligned}$$

2012 ACM Subject Classification Mathematics of computing → Stochastic processes

Keywords and phrases Multivariate records, Pareto records, generators, interior generators, minima, maxima, record-setting region, frontier, current records, boundary-crossing probabilities, first moment method, second moment method, orthants

Digital Object Identifier 10.4230/LIPIcs.AofA.2024.28

[1] Corresponding author

35th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2024).
Editors: Cécile Mailler and Sebastian Wild; Article No. 28; pp. 28:1–28:21
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Related Version *Full Version*: https://arxiv.org/abs/2402.17221

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

Acknowledgements We thank the anonymous reviewers for their helpful comments.

1 Introduction, background, and main results

Notation. Throughout this extended abstract we abbreviate the kth iterate of natural logarithm ln by L_k and L_1 by L, and we write $x_+ := \sum_{j=1}^d x_j$ and $x_\times := \prod_{j=1}^d x_j$ for the sum and product, respectively, of coordinates of the d-dimensional vector $x = (x_1, \dots, x_d)$. When $0 \leq x$ the sum x_+ equals the ℓ^1-norm $\|x\|_1$, but we use the notation x_+ more generally. We denote coordinate-wise maximum and minimum of vectors by $\vee$ and $\wedge$, respectively.

Unless otherwise noted, all results of this extended abstract hold for any dimension $d \geq 2$.

The study of univariate records is well established ([1] is a standard reference), but that of multivariate records remains under vigorous development. Fill and Naiman [6] studied the stochastic process (F_n), where F_n is the boundary, or "frontier", for *Pareto records* (consult Definitions 1.1–1.2) in general dimension d when the observed sequence of points $X^{(1)}, X^{(2)}, \dots$ are assumed (as they are throughout this extended abstract, except where otherwise noted) to be i.i.d. (independent and identically distributed) copies of a d-dimensional random vector X with independent Exponential(1) coordinates X_j. Their main goal was to sharpen (in various senses) the assertion in Bai et al. [2] "that nearly all maxima occur in a thin strip sandwiched between [the] two parallel hyper-planes"

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

They did this largely by studying (separately) the maximum and minimum sums of coordinates for points lying in F_n. The results for the maximum sum were rather sharp; less so for the minimum sum. The main aim of this extended abstract is to use the theory of generators (minima of F_n) and the first- and second-moment methods to improve considerably their results about the minimum sum.

1.1 Pareto records and the record-setting region

For the reader's convenience, and with the permission of the authors and the copyright holder, this short subsection is excerpted largely verbatim from [6, Section 1.1].

We begin with some definitions. For a positive integer n, let $[n] := \{1, \dots, n\}$. Thus $[d]^{[n]}$ denotes the set of all functions from $[n]$ into $[d]$, or simply the set of all n-tuples with each entry in $\{1, \dots, d\}$. For d-dimensional vectors $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$, write $x \prec y$ (respectively, $x \leq y$) to mean that $x_j < y_j$ (resp., $x_j \leq y_j$) for $j \in [d]$. (We caution that, with this convention, $\leq$ is weaker than $\preceq$, the latter meaning "$\prec$ or $=$"; indeed, $(0,0) \leq (0,1)$ but we have neither $(0,0) \prec (0,1)$ nor $(0,0) = (0,1)$. This distinction will be important for some of our later discussion of generators.) The notation $x \succ y$ means $y \prec x$, and $x \geq y$ means $y \leq x$; the notation $x < y$ means $x \leq y$ but $x \neq y$, and $y > x$ means $x < y$.

▶ **Definition 1.1.**

(a) *We say that $X^{(k)}$ is a* (Pareto) record *(or that it* sets a record at time k*) if $X^{(k)} \not\prec X^{(i)}$ for all $1 \leq i < k$.*

(b) *If $1 \leq k \leq n$, we say that $X^{(k)}$ is a* current record *(or* remaining record*, or* maximum*) at time n if $X^{(k)} \not\prec X^{(i)}$ for all $1 \leq i \leq n$.*

For $n \geq 1$ (or $n \geq 0$, with the obvious conventions) let ρ_n ($\equiv \rho_{d,n}$) denote the number of remaining records at time n (when the dimension is d).

▶ **Definition 1.2.**

(a) *The* record-setting region *at time n is the (random) closed set of points*

$$\mathrm{RS}_n := \{x \in \mathbb{R}^d : 0 \leq x \not\prec X^{(i)} \text{ for all } 1 \leq i \leq n\}.$$

(b) *We call the (topological) boundary of RS_n (relative to the closed positive orthant determined by the origin) its* frontier *and denote it by F_n.*

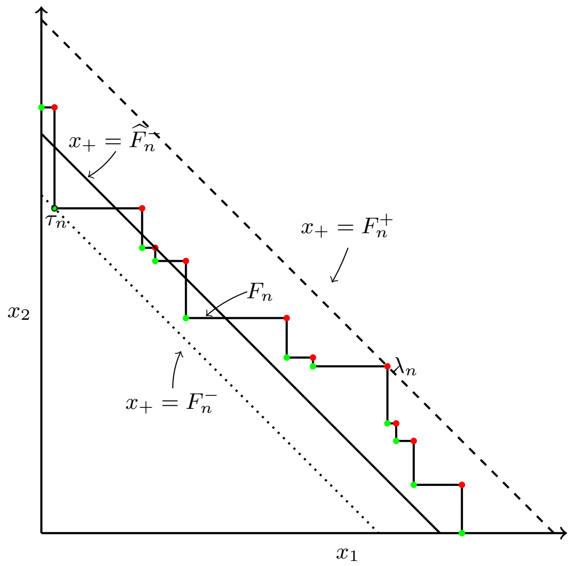

Figure 1 Record frontier F_n based on n observations (for some $n \geq 10$) resulting in 10 current records (shown in red), with the three hyperplanes $x_+ = F_n^+$, $x_+ = F_n^-$, and $x_+ = \widehat{F}_n^-$, the leading point λ_n and the trailing point τ_n. Concerning the three hyperplanes, see Definition 1.4 and (1.2). Generators (see Definition 4.1) are shown in green.

▶ Remark 1.3. The terminology in Definition 1.2(a) is natural since the next observation $X^{(n+1)}$ sets a record if and only if it falls in the record-setting region. Note that

$$\mathrm{RS}_n = \{x \in \mathbb{R}^d : 0 \leq x \not\prec X^{(i)} \text{ for all } 1 \leq i \leq n \\ \text{such that } X^{(i)} \text{ is a current record at time } n\},$$

and that the current records at time n all belong to RS_n but lie on its frontier. Observe also that F_n is a closed subset of RS_n.

This extended abstract primarily concerns the stochastic process (F_n), and specifically the process F^- as defined (along with the process F^+) next (see Figure 1).

▶ **Definition 1.4.** *Recalling that F_n denotes the frontier of RS_n, let*

$$F_n^- := \min\{x_+ : x \in F_n\} \quad \text{and} \quad F_n^+ := \max\{x_+ : x \in F_n\}. \tag{1.1}$$

Almost surely, there are for each n unique vectors $\lambda_n \in F_n$ and $\tau_n \in F_n$ such that $F_n^+ = \lambda_n$ and $F_n^- = \tau_n$; we call λ_n and τ_n the leading *and* trailing *points, respectively, of the frontier.*

Since the sets RS_n decrease (weakly) with n, we have the following trivial consequence.

▶ **Lemma 1.5.** *The process F^- has nondecreasing sample paths.*

1.2 The record-setting frontier; our two main theorems

Fill and Naiman first showed, in a precise sense [6, Theorem 1.4], that the difference between the sum of coordinates (call it Y_n) of a "generic" current record at time n and $\mathrm{L}\, n$ converges in distribution to standard Gumbel. They next translated results from classical extreme value theory due to Kiefer [7] to the setting of multivariate records to produce rather sharp typical-behavior and almost-sure results about the process F^+. For completeness, we repeat their main result [6, Theorem 1.8] for F^+ here, except that we have rather effortlessly extended part (b) of that theorem using Kiefer's "first proof" as described in [6, proof of Theorem 1.8(b)]. We remark that the difference between the top-boundary threshold at about $\mathrm{L}\, n + d\, \mathrm{L}_2\, n$ and bottom-boundary threshold at about $\mathrm{L}\, n + (d-1)\, \mathrm{L}_2\, n$ is a noteworthy feature of F_n^+ discussed further in [6, Section 1.3].

▶ **Theorem 1.6** (Kiefer [7]). *Consider the process F^+ defined at* (1.1).

(a) Typical behavior of F^+: $F_n^+ - [\mathrm{L}\, n + (d-1)\, \mathrm{L}_2\, n - \mathrm{L}((d-1)!)] \xrightarrow{\mathcal{L}} G$.

(b) Top boundaries for F^+: *For any sequence $b_n \to \infty$ that is ultimately monotone increasing,*

$$\mathbb{P}(F_n^+ \geq b_n \text{ i.o.}) = 1 \text{ or } 0 \text{ according as } \textstyle\sum e^{-b_n} b_n^{d-1} \text{ diverges or converges.}$$

In particular, for any $k \geq 2$ we have

$$\mathbb{P}\left(F_n^+ \geq \mathrm{L}\, n + d\, \mathrm{L}_2\, n + \sum_{i=3}^{k} \mathrm{L}_i\, n + c\, \mathrm{L}_{k+1}\, n \text{ i.o.}\right) = \begin{cases} 1 & \text{if } c \leq 1; \\ 0 & \text{if } c > 1. \end{cases}$$

(c) Bottom boundaries for F^+:

$$\mathbb{P}(F_n^+ \leq \mathrm{L}\, n + (d-1)\, \mathrm{L}_2\, n - \mathrm{L}_3\, n - \mathrm{L}((d-1)!) + c \text{ i.o.}) = \begin{cases} 1 & \text{if } c \geq 0; \\ 0 & \text{if } c < 0. \end{cases}$$

◀

From Theorem 1.6 it follows in particular that

$$\frac{F_n^+ - \mathrm{L}\, n}{\mathrm{L}_2\, n} \xrightarrow{\mathrm{P}} d - 1$$

and

$$\liminf \frac{F_n^+ - \mathrm{L}\, n}{\mathrm{L}_2\, n} = d - 1 < d = \limsup \frac{F_n^+ - \mathrm{L}\, n}{\mathrm{L}_2\, n} \quad \text{a.s.}$$

The results derived in [6] for F^- are much less sharp than for F^+. For the reader's convenience, we repeat those results here. Although parts (a) and (c1) were stated with coefficient -3 [rather than $-(2+c)$] for the $\mathrm{L}_3\, n$ term, the improvement we have noted here is pointed out in [6, Remark 3.3].

▶ **Theorem 1.7** ([6], Theorem 1.12). *Consider the process F^- defined at* (1.1).

(a) Typical behavior of F^-:

$$\mathbb{P}(F_n^- \leq \mathrm{L}\, n - (2+c)\, \mathrm{L}_3\, n) \to 0 \text{ if } c > 0, \textit{ and}$$

$$\mathbb{P}(F_n^- \geq \mathrm{L}\, n + c_n) \to 0 \text{ if } c_n \to \infty.$$

(b) Top outer boundaries for F^-: $\mathbb{P}(F_n^- \geq \mathrm{L}\, n + c\, \mathrm{L}_2\, n \text{ i.o.}) = 0$ if $c > 0$.

(c1) Bottom outer boundaries for F^-: $\mathbb{P}(F_n^- \leq \mathrm{L}\, n - (2 + c)\, \mathrm{L}_3\, n \text{ i.o.}) = 0$ if $c > 0$.

(c2) A bottom inner boundary for F^-: $\mathbb{P}(F_n^- \leq \mathrm{L}\, n - \mathrm{L}_3\, n \text{ i.o.}) = 1$. ⌟

Recall that for real-valued random variables Z_n and real numbers a_n, the condition $Z_n = O_{\mathrm{p}}(a_n)$ means that Z_n/a_n is bounded in probability.

The first of two main results of this extended abstract, Theorem 1.8, sharpens Theorem 1.7 considerably. In light of (i) the constant-order variability for a "generic" current record at time n described in the opening paragraph of this subsection and (ii) Theorem 1.6(a), we find it quite surprising that, properly centered but not scaled, F_n^- has a limit in probability.

▶ **Theorem 1.8.** *Consider the process F^- defined at* (1.1)*.*

(a) Typical behavior of F^-:

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

(b) Top outer boundaries for F^-: *If $d \geq 3$, then*

$$\mathbb{P}(F_n^- \geq \mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}(d-2) + \mathrm{L}\, 2 + c \text{ i.o.}) = 0 \text{ if } c > 0.$$

(c) Bottom outer boundaries for F^-:

$$\mathbb{P}(F_n^- \leq \mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}\, d - \mathrm{L}\, 2 - c \text{ i.o.}) = 0 \text{ if } c > 0.$$

Theorem 1.8 gives rise immediately to the following succinct corollary.

▶ **Corollary 1.9.** *Consider the process F^- defined at* (1.1)*.*

(a) Typical behavior of F^-:

$$F_n^- - (\mathrm{L}\, n - \mathrm{L}_3\, n) \xrightarrow{\mathrm{P}} -\mathrm{L}(d-1)$$

and thus

$$\frac{F_n^- - \mathrm{L}\, n}{\mathrm{L}_3\, n} \xrightarrow{\mathrm{P}} -1$$

and, yet more crudely,

$$\frac{F_n^- - \mathrm{L}\, n}{\mathrm{L}_2\, n} \xrightarrow{\mathrm{P}} 0.$$

(b) Almost sure behavior for F^-:

$$\lim \frac{F_n^- - \mathrm{L}\, n}{\mathrm{L}_2\, n} = 0 \text{ a.s.}$$

Further, for fixed $d \geq 3$ we have the refinement

$$F_n^- = \mathrm{L}\, n - \mathrm{L}_3\, n + O(1) \text{ a.s.}$$

▶ Remark 1.10. We do not know how to improve Theorem 1.7(b) when $d = 2$.

Suppose now that instead of F_n^- we consider the somewhat larger quantity

$$\widehat{F}_n^- := \text{(minimum coordinate-sum of any current record at time } n). \tag{1.2}$$

(See Figure 1.) Our second main theorem concerns the process $\widehat{F}^-$; in summary, the same results hold for $\widehat{F}^-$ as for F^- in Theorem 1.8, with a sharper remainder term for $\widehat{F}^-$ in part (a).

▶ **Theorem 1.11.** *Consider the process $\widehat{F}^-$ defined at* (1.2).

(a) Typical behavior of $\widehat{F}^-$:

$$\widehat{F}_n^- = \mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}(d-1) + O_{\mathrm{P}}\left(\frac{1}{\mathrm{L}_2\,n}\right).$$

(b) Top outer boundaries for $\widehat{F}^-$: *If $d \geq 3$, then*

$$\mathbb{P}(\widehat{F}_n^- \geq \mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}(d-2) + \mathrm{L}\,2 + c \text{ i.o.}) = 0 \text{ if } c > 0.$$

(c) Bottom outer boundaries for $\widehat{F}^-$:

$$\mathbb{P}(\widehat{F}_n^- \leq \mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}\,d - \mathrm{L}\,2 - c \text{ i.o.}) = 0 \text{ if } c > 0.$$

As a corollary, the process $\widehat{F}^-$ satisfies the same assertions as for F^- in Corollary 1.9.

▶ Remark 1.12. Combining Theorems 1.8 and 1.11, we find that there is little difference between the two processes, in the sense that

$$\widehat{F}_n^- - F_n^- \xrightarrow{\mathrm{P}} 0,$$

because in fact $0 \leq \widehat{F}_n^- - F_n^- = O_{\mathrm{P}}\left(\frac{\mathrm{L}_3\,n}{\mathrm{L}_2\,n}\right)$.

▶ Remark 1.13. Extending Theorem 1.11, we conjecture that

$$(\mathrm{L}_2\,n)\left(\widehat{F}_n^- - [\mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}(d-1)]\right) \tag{1.3}$$

has a nondegenerate limiting distribution. This is discussed further in Remark 3.3.

1.3 Outline of extended abstract

The proof of Theorem 1.8 relies on Theorem 1.11, so we tackle the latter first. In Sections 2–3 we apply the first moment method and the second moment method, respectively, to the number of remaining records with suitably small coordinate-sum; this leads to the proof of Theorem 1.11 in Appendix B. In Sections 4–5 of this extended abstract we review and extend the theory of generators developed in [5]. In Section 6 we apply the first moment method to the number of generators with suitably small coordinate sum; this, together with the upper bounds on $\widehat{F}^-$ in Theorem 1.11, leads to the proof of Theorem 1.8 in Section 7.

▶ Remark 1.14. Because $F_n^- \leq \widehat{F}_n^-$, Theorem 1.8(b) follows immediately from Theorem 1.11(b), as does Theorem 1.11(c) from Theorem 1.8(c).

More notation. Throughout the extended abstract, the boundaries we consider will without exception have the form

$$b_n := \mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}\,c_n \text{ with } c_n > 0 \text{ and } c_n = \Theta(1). \tag{1.4}$$

Also, we will often use the notation

$$\beta_n := ne^{-b_n} \tag{1.5}$$

The dimension $d \geq 2$ will always remain fixed as $n \to \infty$.

2 Stochastic lower bound on $\widehat{F}_n^-$ via the first moment method

In this section we show how to obtain a suitable stochastic lower bound on $\widehat{F}_n^-$. See Proposition 2.3 for the result. The idea, for a suitably chosen sequence (b_n) is to apply the first moment method (computation of sufficiently small mean, together with application of Markov's inequality) to the count $\rho_n(b_n)$, where

$$\rho_n(b) := \#\{\text{remaining records } r \text{ at epoch } n \text{ with } r_+ \leq b\}. \tag{2.1}$$

Asymptotic determination of the mean is obtained by suitably modifying the asymptotic determination of the mean of $\rho_n = \rho_n(\infty)$ in [2, Section 2].

2.1 Upper (and lower) asymptotic bound(s) on mean

In the next lemma we determine detailed asymptotics for the mean of $\rho_n(b_n)$ when (b_n) is a boundary of interest in establishing Theorems 1.8 and 1.11. The proof is rather elementary, but we defer it to Appendix A. We define

$$J_j(x) := \int_x^\infty (\mathrm{L}\, z)^j e^{-z} \, \mathrm{d}z. \tag{2.2}$$

and note that $J_j(x) \sim (\mathrm{L}\, x)^j e^{-x}$ as $x \to \infty$.

▶ **Lemma 2.1.** *With the notation and assumptions of* (1.4)–(1.5) *and* (2.2)*, as $n \to \infty$ we have*

$$\mathbb{E}\, \rho_n(b_n) = [1 + O(n^{-1}(\mathrm{L}_2\, n)^2)] \frac{1}{(d-1)!} \sum_{j=0}^{d-1} (-1)^j \binom{d-1}{j} (\mathrm{L}\, n)^{d-1-j} J_j(\beta_n), \tag{2.3}$$

or, equivalently,

$$\mathbb{E}\, \rho_n(b_n) = \frac{1}{(d-1)!} \sum_{j=0}^{d-1} (-1)^j \binom{d-1}{j} (\mathrm{L}\, n)^{d-1-j} J_j(\beta_n) + O(n^{-1}(\mathrm{L}\, n)^{d-1-c_n}(\mathrm{L}_2\, n)^2). \tag{2.4}$$

▶ Remark 2.2. We need only lead-order asymptotics for the mean in this section, but (as seen in the proof of Lemma 3.1 found in the full-length paper) we require much more detailed asymptotics for it in the next section – asymptotics with an additive $o(1)$ remainder term, as we have in (2.4).

2.2 Stochastic lower bound on $\widehat{F}_n^-$

We are now in position to apply Markov's inequality to bound the probability of the event $\{\widehat{F}_n^- \leq b_n\} = \{\rho_n(b_n) \geq 1\}$.

▶ **Proposition 2.3** (Stochastic lower bound on $\widehat{F}_n^-$)**.** *With the notation and assumptions of* (1.4)*, as $n \to \infty$ we have*

$$\mathbb{P}(\widehat{F}_n^- \leq b_n) \leq \mathbb{E}\, \rho_n(b_n) = (1 + o(1)) \frac{1}{(d-1)!} (\mathrm{L}\, n)^{d-1-c_n}.$$

3 Stochastic upper bound on $\widehat{F}_n^-$ via second moment method

In this section we show how to obtain a suitable stochastic upper bound on $\widehat{F}_n^-$ (and thus also on F_n^-). See Proposition 3.2 for the result. The idea, for a suitably chosen sequence (b_n), is to apply the second moment method (computation of sufficiently large mean and sufficiently small variance, together with application of Chebyshev's inequality) to the count $\rho_n(b_n)$ [recall the definition (2.1)], which almost surely equals

$$\rho_n^\circ(b_n) := \#\{\text{remaining records } r \text{ at epoch } n \text{ with } r_+ < b_n\}. \tag{3.1}$$

For the mean, we will use Lemma 2.1. The bound on the variance of $\rho_n(b_n)$ is obtained by suitably modifying the already quite technical asymptotic determination of the variance of $\rho_n = \rho_n(\infty)$ in [2, Section 2]; the determination here is quite a bit more technical still.

3.1 Upper bound on variance

We next show that the standard deviation of $\rho_n(b_n)$ is of smaller order of magnitude than the mean – and by enough so that our proof (found in the full-length paper) of Theorem 1.8(b) (for $\widehat{F}^-$, which implies the result for F^-) using the first Borel–Cantelli lemma will succeed. The rather long and rather computationally technical proof of the following result is left for the full-length paper, where the reverse inequality (not needed in this extended abstract) is also established.

▶ **Lemma 3.1.** *With the notation and assumptions of* (1.4)*, as $n \to \infty$ we have*

$$\operatorname{Var} \rho_n(b_n) \le (1 + o(1))\, \mathbb{E}\, \rho_n(b_n). \tag{3.2}$$

3.2 Stochastic upper bound on $\widehat{F}_n^-$

We are now in position to utilize Chebyshev's inequality to provide a bound on $\mathbb{P}(\widehat{F}_n^- \ge b_n) = \mathbb{P}(\rho_n^\circ(b_n) = 0) = \mathbb{P}(\rho_n(b_n) = 0)$.

▶ **Proposition 3.2** (Stochastic upper bound on $\widehat{F}_n^-$)**.** *With the notation and assumptions of* (1.4)*, as $n \to \infty$ we have*

$$\mathbb{P}(F_n^- \ge b_n) \le \mathbb{P}(\widehat{F}_n^- \ge b_n) \le (1 + o(1))\, (d-1)!(\mathrm{L}\, n)^{-(d-1-c_n)} = O((\mathrm{L}\, n)^{-(d-1-c_n)}).$$

Proof. The first asserted inequality follows because $F_n^- \le \widehat{F}_n^-$. Moreover, using Chebyshev's inequality, Lemma 3.1, and Lemma 2.1, we find

$$\begin{aligned}
\mathbb{P}(\widehat{F}_n^- \ge b_n) &\le \mathbb{P}(\rho_n(b_n) = 0) = \mathbb{P}(\rho_n(b_n) - \mathbb{E}\, \rho_n(b_n) \le -\, \mathbb{E}\, \rho_n(b_n)) \\
&\le \frac{\operatorname{Var} \rho_n(b_n)}{[\mathbb{E}\, \rho_n(b_n)]^2} \le (1 + o(1))\, [\mathbb{E}\, \rho_n(b_n)]^{-1} \\
&= (1 + o(1))\, (d-1)!(\mathrm{L}\, n)^{-(d-1-c_n)} \\
&= O((\mathrm{L}\, n)^{-(d-1-c_n)}),
\end{aligned}$$

as desired. ◀

▶ Remark 3.3. Lemma 3.1 and the reverse inequality established in the full-length paper suggest that the law of $\rho_n(b_n)$ might be well approximated by a Poisson distribution with the same mean, but, after attempts using the Stein–Chen method (see, e.g., [4]) or the method of moments, we have been unable to prove such an approximation even in the case that $\mathbb{E}\, \rho_n(b_n)$ has a limit $\lambda \in (0, \infty)$. For fixed $a \in \mathbb{R}$, let $R_n(a)$ denote $\rho_n(b_n)$ when

$$b_n = \mathrm{L}\, n - \mathrm{L}_3\, n - \mathrm{L}(d-1) + \frac{a}{\mathrm{L}_2\, n}, \tag{3.3}$$

i.e., when $c_n = (d-1)e^{-a/\mathrm{L}_2 n}$ in (1.4). Even if a Poisson approximation should fail, we certainly conjecture that $R_n(a)$ converges in distribution to a nondegenerate $R(a)$ as $n \to \infty$ with $\mathbb{P}(R(a) = 0)$ continuous and strictly decreasing in a. In that case, it follows that (1.3) has limiting distribution function $a \mapsto \mathbb{P}(R(a) \geq 1)$.

In particular, if $R(a)$ is Poisson distributed for every a, then (1.3) converges in distribution to $-G^*$, where G^* has a Gumbel distribution with location $-\frac{\mathrm{L}[(d-1)!]}{d-1}$ and scale $\frac{1}{d-1}$.

4 Characterization of generators

The unpublished manuscript [5] by Fill and Naiman developed the concept of generators of multivariate records mainly in connection with an importance-sampling algorithm for generating (simulating) records. We shall find the same concept crucial for our improvement Theorem 1.8(c) to Theorem 1.7(c2), the latter of which was established using a quite different idea, namely, a certain geometric lemma [6, Lemma 3.1]. Accordingly, in this section and the next we review and extend the theory of generators developed in [5]. In this section we provide a characterization of the set of generators that is useful in counting them.

▶ **Definition 4.1.** *Suppose $x \in [0,\infty)^d$.*

(a) *The* closed positive orthant generated *(or* determined*)* by *x is the set*

$$O_x^+ := \{y \in [0,\infty)^d : y \geq x\}.$$

(b) *The minimum points of the frontier F_n are called* generators*. We denote the set of generators at time n by G_n.*

▶ Remark 4.2.

(a) The record-setting region RS_n equals the union $\cup_{g \in G_n} O_g^+$ of closed positive orthants. The elements of G_n are called generators because RS_n is the up-set in $[0,\infty)^d$ generated by G_n with respect to the partial order $\leq$.

(b) The almost surely unique generator with minimum coordinate-sum is the trailing point τ_n, just as the remaining record with maximum coordinate-sum is the leading point λ_n.

There are 11 generators in Figure 1, including the trailing point τ_n at the intersection of F_n and the dotted hyperplane (line) marked with $x_+ = F_n^-$. In terminology we shall establish shortly, 9 of these are interior (i.e., 2-dimensional) generators and 2 of them are 1-dimensional generators.

We now proceed to characterize the set of generators.

Denote the $\rho \equiv \rho_n$ current records at a given time n by $r^{(1)}, \ldots, r^{(\rho)}$ (listed here in arbitrary, but fixed, order). The record-setting region $S \equiv \mathrm{RS}_n$ is then the closed set

$$\begin{aligned} S &= \cap_{i=1}^{\rho} \left[\cup_{k=1}^{d} O^+\left(r_k^{(i)} e^{(k)}\right)\right] = \cup_{k_1=1}^{d} \cdots \cup_{k_\rho=1}^{d} \cap_{i=1}^{\rho} O^+\left(r_{k_i}^{(i)} e^{(k_i)}\right) \\ &= \cup_{k_1=1}^{d} \cdots \cup_{k_\rho=1}^{d} O^+\left(\vee_{i=1}^{\rho} r_{k_i}^{(i)} e^{(k_i)}\right) = \cup_{k \in [d]^{[\rho]}} O^+\left(R_1^{(\Pi_1(k))}, \ldots, R_d^{(\Pi_d(k))}\right), \end{aligned}$$

where $e^{(k)}$ denotes the k^{th} standard basis vector and for $j \in [d]$ and $k \in [d]^{[\rho]}$ we have defined the ordered partition $\Pi(k) = (\Pi_1(k), \ldots, \Pi_d(k))$ of $[\rho]$ by

$$\Pi_j(k) := k^{-1}(\{j\}) = \{i \in [\rho] : k_i = j\},$$

and for $j \in [d]$ and $P \subseteq [\rho]$ we have defined

$$R_j^{(P)} := \vee_{i \in P} r_j^{(i)}.$$

Therefore we have the neat representation

$$S = \bigcup O^{+}\left(R_1^{(\Pi_1)}, \dots, R_d^{(\Pi_d)}\right), \tag{4.1}$$

where the union here is taken over all ordered partitions $\Pi = (\Pi_1, \dots, \Pi_d)$ of $[\rho]$ into d sets; each Π_j is allowed to be empty, in which case $R_j^{(\Pi_j)} := 0$. This shows immediately that every element of $G \equiv G_n$ has in each coordinate either 0 or the value of some record in that coordinate.

To simplify our characterization of generators, we begin by considering only "interior" generators. For any point $x \in O_0^{+}$, let $\nu(x)$ denote the set of non-zero coordinates of x, and observe that x lies in the interior of O_0^{+} if and only if $\nu(x) = [d]$. We call such a point x an *interior* point.

Observe that a point x of the form $\left(R_1^{(\Pi_1)}, \dots, R_d^{(\Pi_d)}\right)$ appearing in (4.1) is interior if and only if all the cells Π of the partition are nonempty. Next, note that $x \in (0,\infty)^d$ is of such a form if and only if there exist d distinct indices $i_1, \dots, i_d$ such that $x_j = r_j^{(i_j)}$ for $j \in [d]$.

We are now in position to state and (in Appendix C) prove a characterization of the set I of interior generators. (Note that $I \subset G \subset S$.)

▶ **Theorem 4.3.** *A point $g \in [0,\infty)^d$ belongs to I if and only if*
(i) *$g \in S$, and*
(ii) *there exist d distinct indices $i_1, \dots, i_d$ such that*

$$g_j = r_j^{(i_j)} = \min\left\{r_j^{(i_\ell)} : \ell \in [d]\right\} \text{ for every } j \in [d]. \tag{4.2}$$

▶ Remark 4.4. Theorem 4.3 gives an injection from the set of interior generators into the set of ordered d-tuples of remaining records.

Now that we have characterized the interior generators, it is straightforward to characterize G in terms of projections of the current records to lower-dimensional coordinate subspaces, but some care must be taken to ensure that the almost sure property of having no coordinate ties remains true after projection. To begin a careful description, given a subset $T = \{j_1, \dots, j_t\}$ of $[d]$ with $|T| = t \in [d]$ and $1 \le j_1 < \cdots < j_t \le d$, define the *projection mapping* $\pi_T : \mathbb{R}^d \to \mathbb{R}^t$ by

$$\pi_T(x_1, \dots, x_d) := (x_{j_1}, \dots, x_{j_t}),$$

and define the *injection mapping* $\iota_T : \mathbb{R}^t \to \mathbb{R}^d$ by

$$\iota_T(x_1, \dots, x_t) := \vee_{k=1}^{t} x_{j_k} e^{(j_k)}.$$

Recall that $\nu(x)$ denotes the set of nonzero coordinates of a point $x \in [0,\infty)^d$. Define the set of *T-generators* to be the set

$$G_T := G \cap \{x : \nu(x) = T\}$$

and observe that G is the disjoint union

$$G = \cup_{T \subseteq [d]} G_T.$$

This observation, together with a characterization of each G_T, thus provides a characterization of G. A characterization of each G_T is obtained by combining the following theorem with Theorem 4.3.

To set up the statement of the theorem, consider the image

$$R_T := \pi_T(R) = \left\{\pi_T(r^{(i)}) : i \in [\rho]\right\} \subset \mathbb{R}^{|T|}$$

under π_T of the set $R := \left\{r^{(i)} : i \in [\rho]\right\}$ of current records, and note that R_T inherits the property of "no ties in any coordinate" from R. Let I_T denote the set of interior generators of R_T, and let $G'_T := \iota_T(I_T)$ denote the injection of I_T into $\mathbb{R}^d$.

▶ **Theorem 4.5.** *For every $T \subseteq [d]$ we have $G_T = G'_T$.*

In light of Theorem 4.5 (which is proved in Appendix C), we call the number of nonzero coordinates of a generator its *dimension*. Figure 2 shows the generators of various dimensions for an example with $d = 3$.

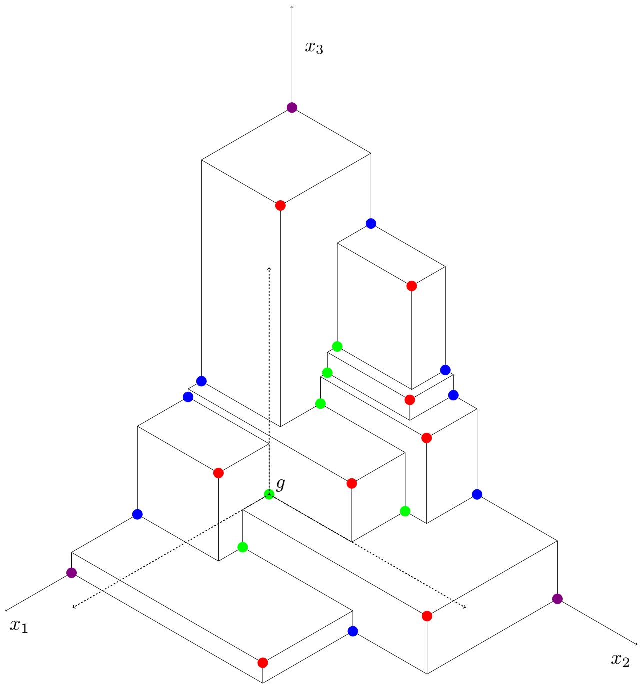

Figure 2 Example of a record frontier in dimension $d = 3$ with $\rho = 8$ remaining records shown in red and the resulting $\gamma = 17$ generators: three one-dimensional generators shown in violet, eight two-dimensional generators shown in blue, and six three-dimensional (interior) generators shown in green. The lower boundary of one of the orthants O_g^+ is shown using dashed lines.

▶ **Example 4.6.** Suppose $d = 4$ and the current records are $(2, 8, 3, 7)$ and $(5, 1, 4, 6)$. Then $|G| = 8$, because $|G_T| = 1$ for precisely eight nonempty subsets T of $[4]$ and $|G_T| = 0$ otherwise. The eight subsets T for which $|G_T| = 1$ are

$$G_{\{1\}} = \{(5, 0, 0, 0)\};\ G_{\{2\}} = \{(0, 8, 0, 0)\};\ G_{\{3\}} = \{(0, 0, 4, 0)\};$$
$$G_{\{4\}} = \{(0, 0, 0, 7)\};\ G_{\{1,2\}} = \{(2, 1, 0, 0)\};\ G_{\{1,4\}} = \{(2, 0, 0, 6)\};$$
$$G_{\{2,3\}} = \{(0, 1, 3, 0)\};\ G_{\{3,4\}} = \{(0, 0, 3, 6)\}.$$

Thus there are four one-dimensional generators, four two-dimensional generators, and no generators with dimension exceeding two.

5 The expected number of generators

The proof of Theorem 1.8(c) requires a tight upper bound on the expected number of generators at time n with suitably small coordinate-sum. In this section we warm up with a result of independent interest, giving an asymptotic approximation for the expected total number of generators at time n. We remark in passing that such an approximation proves useful in the analysis of the importance-sampling record-generating scheme described in [5, Sections 2–4].

5.1 Exact expressions

Let $\gamma_{d,n}$ (respectively, $\iota_{d,n}$) denote the number of generators (resp., interior generators) after a given number n of d-dimensional observations. Our first result relates the expectations of these two quantities.

▶ **Lemma 5.1.** *For integers $d \geq 0$ and $n \geq 0$, we have*

$$\mathbb{E}\,\gamma_{d,n} = \sum_{k=0}^{d} \binom{d}{k} \mathbb{E}\,\iota_{k,n}, \tag{5.1}$$

Proof. This is immediate from Theorem 4.5 and the discussion preceding that theorem. ◀

In Lemma 5.1, note that $\iota_{0,n} = \delta_{0,n}$: There is a single 0-dimensional generator (namely, the origin in $\mathbb{R}^d$) if $n = 0$ and no 0-dimensional generators otherwise. Also note that $\iota_{d,n} = 0$ if $n < d$.

The next result (proved in Appendix C) gives an exact expression for $\mathbb{E}\,\iota_{d,n}$ for $n \geq d \geq 1$. We write $n^{\underline{k}}$ for the falling factorial power

$$n(n-1)\cdots(n-k+1) = k!\tbinom{n}{k}.$$

▶ **Lemma 5.2.** *For integers $n \geq d \geq 1$, we have*

$$\mathbb{E}\,\iota_{d,n} = n^{\underline{d}} \int_{(0,1]^d} x_\times^{d-1} (1 - x_\times)^{n-d}\,\mathrm{d}x. \tag{5.2}$$

▶ Remark 5.3.

(a) The exact expression (5.2) in Lemma 5.2 may be compared to a similar expression for $\mathbb{E}\,\rho_{d,n}$ derived in [2, Section 2]: For $d \geq 1$ and $n \geq 1$ we have

$$\mathbb{E}\,\rho_{d,n} = n \int_{(0,1]^d} (1 - x_\times)^{n-1}\,\mathrm{d}x. \tag{5.3}$$

In fact, by expanding the factor $x_\times^{d-1}$ appearing in the integrand in (5.2) as

$$[1 - (1 - x_\times)]^{d-1} = \sum_{j=0}^{d-1} (-1)^j \binom{d-1}{j} (1 - x_\times)^j,$$

one sees that the expected counts of interior generators and expected counts of remaining records are related by

$$\mathbb{E}\,\iota_{d,n} = n^{\underline{d}} \sum_{j=0}^{d-1} (-1)^j \frac{\binom{d-1}{j}}{n - d + j + 1} \mathbb{E}\,\rho_{d,n-d+j+1}$$

for $n \geq d \geq 1$. But we do not know of any use for this connection.

(b) An alternative expression to (5.3) is

$$\mathbb{E}\,\rho_{d,n} = \sum_{j=1}^{n} (-1)^{j-1} \binom{n}{j} j^{-(d-1)} =: \widehat{H}_n^{(d-1)},$$

a so-called *Roman harmonic number* studied by [8], [9], [10].

5.2 Asymptotics

From here we follow the same outline as for the expected number of remaining records in Bai et al. [2] to obtain an asymptotic expansion for $\mathbb{E}\,\iota_{d,n}$ (see our Theorem 5.7, the main result of Section 5). Accordingly, we begin by considering a Poissonized analogue of $\mathbb{E}\,\iota_{d,n}$, whose proof is rather simple and is included in the full-length paper.

▶ **Lemma 5.4.** *For integers $d \geq 1$ and $n \geq 0$, define*

$$\hat{\iota}_{d,n} := n^d \int_{[0,1)^d} x_\times^{d-1} \exp(-n x_\times) \, \mathrm{d}x.$$

Then, for fixed d, as $n \to \infty$ we have

$$\hat{\iota}_{d,n} = (\mathrm{L}\,n)^{d-1} \sum_{j=0}^{d-1} \frac{(-1)^j \Gamma^{(j)}(d)}{j!(d-1-j)!} (\mathrm{L}\,n)^{-j} + O((n\,\mathrm{L}\,n)^{d-1} e^{-n}).$$

We next bound the difference between $\hat{\iota}_{d,n}$ and

$$\tilde{\iota}_{d,n} := n^d \int_{[0,1)^d} x_\times^{d-1} (1 - x_\times)^n \, \mathrm{d}x. \tag{5.4}$$

▶ **Lemma 5.5.** *For fixed $d \geq 1$, as $n \to \infty$ we have*

$$0 \leq \hat{\iota}_{d,n} - \tilde{\iota}_{d,n} = O(n^{-1} (\mathrm{L}\,n)^{d-1}).$$

Proof. We utilize the elementary inequality

$$e^{-nt}(1 - nt^2) \leq (1-t)^n \leq e^{-nt}$$

for $n \geq 1$ and $0 \leq t \leq 1$ (see [3, Lemma 5]). This yields

$$0 \leq \hat{\iota}_{d,n} - \tilde{\iota}_{d,n} \leq n^{d+1} \int_{[0,1)^d} x_\times^{d+1} \exp(-n x_\times) \, \mathrm{d}x.$$

Proceeding just as in the proof of Lemma 5.4, we find that the last expression here is $O(n^{-1} (\mathrm{L}\,n)^{d-1})$. ◀

▶ **Theorem 5.6.** *For fixed $d \geq 1$, as $n \to \infty$ the expected number of interior generators at time n in dimension d satisfies*

$$\mathbb{E}\,\iota_{d,n} = (\mathrm{L}\,n)^{d-1} \sum_{j=0}^{d-1} \frac{(-1)^j \Gamma^{(j)}(d)}{j!(d-1-j)!} (\mathrm{L}\,n)^{-j} + O(n^{-1} (\mathrm{L}\,n)^{d-1}).$$

Proof. Comparing (5.2) and (5.4) and then invoking Lemma 5.5, we see that

$$\begin{aligned}\mathbb{E}\,\iota_{d,n} &= \frac{n^{\underline{d}}}{(n-d)^d}\tilde{\iota}_{d,n-d} = [1+O(n^{-1})]\,\tilde{\iota}_{d,n-d} \\ &= [1+O(n^{-1})]\ \left[\hat{\iota}_{d,n-d} + O(n^{-1}(\mathrm{L}\,n)^{d-1})\right] \\ &= [1+O(n^{-1})]\,\hat{\iota}_{d,n-d} + O(n^{-1}(\mathrm{L}\,n)^{d-1}).\end{aligned}$$

But, according to Lemma 5.4,

$$\begin{aligned}\hat{\iota}_{d,n-d} &= [\mathrm{L}(n-d)]^{d-1}\sum_{j=0}^{d-1}\frac{(-1)^j\Gamma^{(j)}(d)}{j!(d-1-j)!}[\mathrm{L}(n-d)]^{-j} + O((n\,\mathrm{L}\,n)^{d-1}e^{-n}) \\ &= (\mathrm{L}\,n)^{d-1}\sum_{j=0}^{d-1}\frac{(-1)^j\Gamma^{(j)}(d)}{j!(d-1-j)!}(\mathrm{L}\,n)^{-j} + O(n^{-1}(\mathrm{L}\,n)^{d-2}).\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{E}\,\iota_{d,n} &= [1+O(n^{-1})]\,(\mathrm{L}\,n)^{d-1}\sum_{j=0}^{d-1}\frac{(-1)^j\Gamma^{(j)}(d)}{j!(d-1-j)!}(\mathrm{L}\,n)^{-j} + O(n^{-1}(\mathrm{L}\,n)^{d-1}) \\ &= (\mathrm{L}\,n)^{d-1}\sum_{j=0}^{d-1}\frac{(-1)^j\Gamma^{(j)}(d)}{j!(d-1-j)!}(\mathrm{L}\,n)^{-j} + O(n^{-1}(\mathrm{L}\,n)^{d-1}),\end{aligned}$$

as claimed. ◀

Combining (5.1) and (5.2), we can obtain an exact expression for $\mathbb{E}\,\gamma_{d,n}$. Similarly, combining (5.1) and Theorem 5.6 we obtain the following asymptotic expansion in powers of logarithm for $\mathbb{E}\,\gamma_{d,n}$ after a little rearrangement.

▶ **Theorem 5.7.** *For fixed $d \geq 1$, as $n \to \infty$ the expected number of generators at time n in dimension d satisfies*

$$\mathbb{E}\,\gamma_{d,n} = (\mathrm{L}\,n)^{d-1}\sum_{j=0}^{d-1} a_{d,j}(\mathrm{L}\,n)^{-j} + O(n^{-1}(\mathrm{L}\,n)^{d-1}),$$

where

$$a_{d,j} := \sum_{k=0}^{j}\binom{d}{d-j+k}\frac{(-1)^k\Gamma^{(k)}(d-j+k)}{k!(d-1-j)!}.$$

⌟

▶ Remark 5.8. Concerning Theorem 5.7:

(a) In particular, $a_{d,0} = 1$, so $\mathbb{E}\,\gamma_{d,n}$ has lead-order asymptotics

$$\mathbb{E}\,\gamma_{d,n} = (\mathrm{L}\,n)^{d-1} + O((\mathrm{L}\,n)^{d-2});$$

this is $(d-1)!$ times as large as the lead-order asymptotics for the expected number of remaining records, namely,

$$\mathbb{E}\,\rho_{d,n} = \frac{(\mathrm{L}\,n)^{d-1}}{(d-1)!} + O((\mathrm{L}\,n)^{d-2}).$$

(b) For $d = 2$ and $n \geq 0$, we have

$$\mathbb{E}\,\gamma_{2,n} = H_n + 1 = \mathbb{E}\,\rho_{2,n} + 1,$$

where $H_n := \sum_{k=1}^{n} k^{-1}$ is the nth harmonic number; and in fact it is easy to see that $\gamma_{2,n} = \rho_{2,n} + 1$. For $d = 3$ and $n \geq 0$, we have

$$\mathbb{E}\,\gamma_{3,n} = H_n^2 + H_n^{(2)} + 1 = 2\,\mathbb{E}\,\rho_{3,n} + 1,$$

where $H_n^{(2)} := \sum_{k=1}^{n} k^{-2}$ is the nth second-order harmonic number; and in fact $\gamma_{3,n} = 2\rho_n + 1$, as established in [5, Corollary 6.6]. There is not such a simple relationship between the exact values of $\rho_{d,n}$ and $\gamma_{d,n}$ for $d \geq 4$; confer [5, Remark 6.7].

(c) We hope to extend the work of Section 5.2 by finding at least lead-order asymptotics for the variance, and also a normal approximation or other limit theorem, for the number $\gamma_{d,n}$ of generators after n observations.

6 Stochastic lower bound on F_n^- via the first moment method

In this section we show how to obtain a suitable stochastic lower bound on F_n^-. See Proposition 6.2 for the result. The idea, for a suitably chosen sequence (b_n) is to apply the first moment method (computation of sufficiently small mean, together with application of Markov's inequality) to the count

$$\gamma_n(b) := (\text{number of generators at epoch } n \text{ with coordinate-sum} \leq b).$$

The bound on the mean of $\gamma_n(b_n)$ is obtained by suitably modifying the proof of Theorem 5.7 [compare also the similar treatment of $\rho_n(b_n)$ in Section 3 and the full-length paper].

▶ **Lemma 6.1.** *With the notation and assumptions of* (1.4)–(1.5) *and* (2.2)*, as* $n \to \infty$ *we have*

$$\mathbb{E}\,\gamma_n(b_n) \leq (1 + o(1))\frac{(\mathrm{L}\,n)^{d-1}}{(d-1)!}(c_n \,\mathrm{L}_2\, n)^{d-1}(\mathrm{L}\,n)^{-c_n}. \tag{6.1}$$

Proof. We will be very brief here. Following very closely along the lines of Section 5, one finds that

$$\begin{aligned}
\mathbb{E}\,\gamma_n(b_n) &\sim \frac{1}{(d-1)!}\int_{ne^{-b_n}}^{\mathrm{L}\,n} (\mathrm{L}\,n - \mathrm{L}\,z)^{d-1} z^{d-1} e^{-z}\,\mathrm{d}z \\
&\leq \frac{(\mathrm{L}\,n)^{d-1}}{(d-1)!}\int_{ne^{-b_n}}^{\infty} z^{d-1} e^{-z}\,\mathrm{d}z \\
&\sim \frac{(\mathrm{L}\,n)^{d-1}}{(d-1)!}(ne^{-b_n})^{d-1}\exp(-ne^{-b_n}) \\
&= \frac{(\mathrm{L}\,n)^{d-1}}{(d-1)!}(c_n \,\mathrm{L}_2\, n)^{d-1}(\mathrm{L}\,n)^{-c_n}.
\end{aligned}$$ ◀

We are now in position to utilize Markov's inequality.

▶ **Proposition 6.2** (Stochastic lower bound on F_n^-)**.** *Fix* $d \geq 2$*. If* $1 \leq c_n = O(1)$ *and*

$$b \equiv b_n := \mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}\,c_n,$$

then

$$\mathbb{P}(F_n^- \leq b_n) \leq \mathbb{E}\,\gamma_n(b_n) \leq (1 + o(1))\frac{(\mathrm{L}\,n)^{d-1}}{(d-1)!}(c_n \,\mathrm{L}_2\, n)^{d-1}(\mathrm{L}\,n)^{-c_n}.$$

7 Proof of Theorem 1.8

In this section we prove Theorem 1.8.

Proof of Theorem 1.8.

(a) This follows readily from Propositions 6.2 and 3.2 (or one can invoke Theorem 1.11 instead of Proposition 3.2).

(b) As noted in Remark 1.14, this is immediate from Theorem 1.11(b), already established in Appendix B.

(c) This follows in the same fashion as our given proof of Theorem 1.11(c), now using Proposition 6.2 in place of Proposition 2.3. We leave the routine details to the reader. ◀

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 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`.

3 Zhi-Dong Bai, Hsien-Kuei Hwang, Wen-Qi Liang, and Tsung-Hsi Tsai. Limit theorems for the number of maxima in random samples from planar regions. *Electron. J. Probab.*, 6:no. 3, 41, 2001. `doi:10.1214/EJP.v6-76`.

4 A.D. Barbour, Lars Holst, and Svante Janson. *Poisson Approximation*. Oxford studies in Probability. Oxford University Press, New York, 1992.

5 James Allen Fill and Daniel Q. Naiman. Generating Pareto records, 2019. arXiv:1901.05621. `arXiv:1901.05621`.

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

7 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.

8 Daniel E. Loeb and Gian-Carlo Rota. Formal power series of logarithmic type. *Adv. Math.*, 75(1):1–118, 1989. `doi:10.1016/0001-8708(89)90079-0`.

9 Steven Roman. The logarithmic binomial formula. *Amer. Math. Monthly*, 99(7):641–648, 1992. `doi:10.2307/2324994`.

10 J. Sesma. The Roman harmonic numbers revisited. *J. Number Theory*, 180:544–565, 2017. `doi:10.1016/j.jnt.2017.05.009`.

A Proof of Lemma 2.1

This appendix is devoted to the (elementary) proof of Lemma 2.1.

Proof of Lemma 2.1. We will prove (2.3) by separately considering (a) upper and (b) lower bounds. Before beginning, we note that the mean in question has the exact expression

$$\mathbb{E}\,\rho_n(b_n) = n \int_{x \geq 0:\, x_+ \leq b_n} e^{-x_+}(1-e^{-x_+})^{n-1}\,\mathrm{d}x = \frac{n}{(d-1)!}\int_0^{b_n} y^{d-1}e^{-y}(1-e^{-y})^{n-1}\,\mathrm{d}y. \quad \text{(A.1)}$$

A key technical tool we will use is the pair of elementary inequalities

$$e^{-nt}(1-nt^2) \leq (1-t)^n \leq e^{-nt} \quad \text{(A.2)}$$

for $n \ge 1$ and $0 \le t \le 1$ (see [3, Lemma 5]). Also, note from the definition (2.2) of the function J_j that

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

and that for $1 \le x < y$ we have

$$0 < J_j(x) - J_j(y) \le (\mathrm{L}\,y)^j (e^{-x} - e^{-y}) = (\mathrm{L}\,y)^j e^{-y}(e^{y-x} - 1). \tag{A.4}$$

(a) Utilizing the upper bound in (A.2) immediately we derive

$$\begin{aligned}
&\mathbb{E}\,\rho_{n+1}(b_{n+1}) \\
&= \frac{n+1}{(d-1)!}\int_0^{b_{n+1}} y^{d-1}e^{-y}(1-e^{-y})^n\,\mathrm{d}y \\
&\le \frac{n+1}{(d-1)!}\int_0^{b_{n+1}} y^{d-1}\exp\left(-ne^{-y}-y\right)\,\mathrm{d}y \\
&= \frac{1+n^{-1}}{(d-1)!}\int_{ne^{-b_{n+1}}}^{n} (\mathrm{L}\,n - \mathrm{L}\,z)^{d-1}e^{-z}\,\mathrm{d}z \\
&\le \frac{1+n^{-1}}{(d-1)!}\int_{ne^{-b_{n+1}}}^{n+1} [\mathrm{L}(n+1) - \mathrm{L}\,z]^{d-1}e^{-z}\,\mathrm{d}z \\
&= \frac{1+n^{-1}}{(d-1)!}\sum_{j=0}^{d-1}(-1)^j\binom{d-1}{j}[\mathrm{L}(n+1)]^{d-1-j}\int_{ne^{-b_{n+1}}}^{n+1}(\mathrm{L}\,z)^j e^{-z}\,\mathrm{d}z \\
&= \frac{1+n^{-1}}{(d-1)!}\sum_{j=0}^{d-1}(-1)^j\binom{d-1}{j}[\mathrm{L}(n+1)]^{d-1-j}[J_j(ne^{-b_{n+1}}) - J_j(n+1)].
\end{aligned} \tag{A.5}$$

That is,

$$\mathbb{E}\,\rho_n(b_n) \le \frac{1+(n-1)^{-1}}{(d-1)!}\sum_{j=0}^{d-1}(-1)^j\binom{d-1}{j}(\mathrm{L}\,n)^{d-1-j}[J_j((n-1)e^{-b_n}) - J_j(n)]. \tag{A.6}$$

By the note following (2.2), $J_j(n) \sim (\mathrm{L}\,n)^j e^{-n}$. Moreover, by (A.4) we have

$$\begin{aligned}
&0 < J_j((n-1)e^{-b_n}) - J_j(\beta_n) \\
&\quad \le (\mathrm{L}\,\beta_n)^j e^{-\beta_n}[\exp(e^{-b_n}) - 1] \\
&\quad \sim (\mathrm{L}\,\beta_n)^j e^{-\beta_n} e^{-b_n} = (\mathrm{L}\,\beta_n)^j e^{-\beta_n} n^{-1} c_n \,\mathrm{L}_2\, n = O\big((\mathrm{L}\,\beta_n)^j e^{-\beta_n} n^{-1}\,\mathrm{L}_2\, n\big)\,.
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}\,\rho_n(b_n) &\le \frac{1}{(d-1)!}\sum_{j=0}^{d-1}(-1)^j\binom{d-1}{j}(\mathrm{L}\,n)^{d-1-j}J_j(\beta_n) + O(n^{-1}(\mathrm{L}\,n)^{d-1-c_n}\,\mathrm{L}_2\, n) \\
&= [1 + O(n^{-1}\,\mathrm{L}_2\, n)]\frac{1}{(d-1)!}\sum_{j=0}^{d-1}(-1)^j\binom{d-1}{j}(\mathrm{L}\,n)^{d-1-j}J_j(\beta_n),
\end{aligned}$$

bettering the claim in the upper-bound direction for the mean at (2.3).

(b) Utilizing the lower bound in (A.2), we find from (A.1) that

$$\begin{aligned}\mathbb{E}\,\rho_n(b_n) &= \frac{n}{(d-1)!}\int_0^{b_n} y^{d-1}e^{-y}(1-e^{-y})^{n-1}\,\mathrm{d}y \\ &\geq \frac{n}{(d-1)!}\int_0^{b_n} y^{d-1}e^{-y}(1-e^{-y})^{n}\,\mathrm{d}y \\ &\geq \frac{n}{(d-1)!}\int_0^{b_n} y^{d-1}\exp\left(-ne^{-y}-y\right)\left(1-ne^{-2y}\right)\,\mathrm{d}y. \end{aligned} \tag{A.7}$$

We derive that the added term in (A.7) satisfies

$$\begin{aligned}&\frac{n}{(d-1)!}\int_0^{b_n} y^{d-1}\exp\left(-ne^{-y}-y\right)\,\mathrm{d}y \\ &= \frac{1}{(d-1)!}\int_{\beta_n}^{n}(\mathrm{L}\,n-\mathrm{L}\,z)^{d-1}e^{-z}\,\mathrm{d}z \\ &= \frac{1}{(d-1)!}\sum_{j=0}^{d-1}(-1)^j\binom{d-1}{j}(\mathrm{L}\,n)^{d-1-j}[J_j(\beta_n)-J_j(n)]. \end{aligned}$$

But $J_j(n) \sim e^{-n}(\mathrm{L}\,n)^j$, whence the added term in (A.7) is lower-bounded by

$$\frac{1}{(d-1)!}\sum_{j=0}^{d-1}(-1)^j\binom{d-1}{j}(\mathrm{L}\,n)^{d-1-j}J_j(\beta_n) - O(e^{-n}(\mathrm{L}\,n)^{d-1}).$$

So it remains to show that the subtracted term in (A.7) can be absorbed into the remainder term in (2.4), which we will do in similar (but easier) fashion to upper-bounding $\mathbb{E}\,\rho_n(b_n)$. Indeed, the subtracted term satisfies

$$0 < \frac{n^2}{(d-1)!}\int_0^{b_n} y^{d-1}\exp\left(-ne^{-y}-3y\right)\,\mathrm{d}y = \frac{n^{-1}}{(d-1)!}\int_{\beta_n}^{n} z^2(\mathrm{L}\,n-\mathrm{L}\,z)^{d-1}e^{-z}\,\mathrm{d}z \tag{A.8}$$

$$\begin{aligned}&\leq \frac{n^{-1}(\mathrm{L}\,n)^{d-1}}{(d-1)!}\int_{\beta_n}^{\infty} z^2 e^{-z}\,\mathrm{d}z \\ &\sim \frac{n^{-1}(\mathrm{L}\,n)^{d-1}}{(d-1)!}\beta_n^2 e^{-\beta_n} \\ &= \frac{n^{-1}(\mathrm{L}\,n)^{d-1}}{(d-1)!}c_n^2(\mathrm{L}_2\,n)^2(\mathrm{L}\,n)^{-c_n} = O(n^{-1}(\mathrm{L}_2\,n)^2(\mathrm{L}\,n)^{d-1-c_n}) \\ &= O\left(n^{-1}(\mathrm{L}_2\,n)^2\frac{1}{(d-1)!}\sum_{j=0}^{d-1}(-1)^j\binom{d-1}{j}(\mathrm{L}\,n)^{d-1-j}J_j(\beta_n)\right), \end{aligned}$$

as desired. ◀

B Proof of Theorem 1.11

In this appendix we prove Theorem 1.11.

Proof of Theorem 1.11.

(a) This follows readily from Propositions 2.3 and 3.2. Here are some details. For $a \in \mathbb{R}$, let

$$b_n(a) := \mathrm{L}\,n - \mathrm{L}_3\,n - \mathrm{L}(d-1) + \frac{a}{\mathrm{L}_2\,n}, \tag{B.1}$$

as at (3.3); this is an instance of (1.4) with $c_n = (d-1)e^{-a/\operatorname{L}_2 n}$. By Proposition 2.3,

$$\mathbb{P}(\widehat{F}_n^- \le b_n(a)) \le (1+o(1))\frac{1}{(d-1)!}(\operatorname{L} n)^{d-1-c_n} \to \frac{1}{(d-1)!}e^{(d-1)a};$$

the last expression here tends to 0 as $a \to -\infty$. Similarly, by Proposition 3.2,

$$\mathbb{P}(\widehat{F}_n^- \ge b_n(a)) \le (1+o(1))(d-1)!(\operatorname{L} n)^{-(d-1-c_n)} \to (d-1)!e^{-(d-1)a},$$

and the last expression here tends to 0 as $a \to \infty$. It follows that the sequence of distributions of (1.3) is tight, i.e., that Theorem 1.11(a) holds.

(b) Like F^- (Lemma 1.5), the process $\widehat{F}^-$ has nondecreasing sample paths. From this it follows that if (b_n) is (ultimately) monotone nondecreasing and (n_j) is any strictly increasing sequence of positive integers, then

$$\{\widehat{F}_n^- \ge b_n \text{ i.o.}(n)\} \subseteq \{\widehat{F}_{n_{j+1}}^- \ge b_{n_j} \text{ i.o.}(j)\}.$$

To complete the proof of part (b), we choose $b_n \equiv \operatorname{L} n - \operatorname{L}_3 n - \operatorname{L}(d-2) + \operatorname{L} 2 + c$ with $c > 0$ and $n_j \equiv 2^j$, bound $\mathbb{P}(\widehat{F}_{n_{j+1}}^- \ge b_{n_j})$ using Proposition 3.2, and apply the first Borel–Cantelli lemma.

Here are the details. If n is even, then

$$\begin{aligned} b_{n/2} &= \operatorname{L}(n/2) - \operatorname{L}_3(n/2) - \operatorname{L}(d-2) + \operatorname{L} 2 + c = \operatorname{L} n - \operatorname{L}_3(n/2) - \operatorname{L}(d-2) + c \\ &\ge \operatorname{L} n - \operatorname{L}_3 n - \operatorname{L}(d-2) + c, \end{aligned}$$

the last expression being the one in Proposition 3.2 with $c_n \equiv e^{-c}(d-2)$. Thus, by that proposition,

$$\begin{aligned} \mathbb{P}(\widehat{F}_{n_{j+1}}^- \ge b_{n_j}) &= \mathbb{P}(\widehat{F}_{n_{j+1}}^- \ge b_{n_{j+1}/2}) \\ &\le \mathbb{P}(\widehat{F}_{n_{j+1}}^- \ge \operatorname{L} n_{j+1} - \operatorname{L}_3 n_{j+1} - \operatorname{L}(d-2) + c) \\ &= O((\operatorname{L} n_{j+1})^{-[d-1-e^{-c}(d-2)]}) = O((j+1)^{-[1+(1-e^{-c})(d-2)]}), \end{aligned}$$

which is summable.

(c) To prove part (c) [which, as noted in Remark 1.14, will also follow immediately once we prove Theorem 1.8(c)], we begin with an argument similar to that for part (b). If (b_n) is (ultimately) monotone nondecreasing and (n_j) is any strictly increasing sequence of positive integers, then

$$\{F_n^- \le b_n \text{ i.o.}(n)\} \subseteq \{F_{n_j}^- \le b_{n_{j+1}} \text{ i.o.}(j)\}.$$

To complete the proof of part (c), we choose $b_n \equiv \operatorname{L} n - \operatorname{L}_3 n - \operatorname{L} d - \operatorname{L} 2 - c$ with $c > 0$ and $n_j \equiv 2^j$, bound $\mathbb{P}(F_{n_j}^- \le b_{n_{j+1}})$ using Proposition 2.3, and apply the first Borel–Cantelli lemma.

Here are the details. First note that

$$b_{2n} = \operatorname{L}(2n) - \operatorname{L}_3(2n) - \operatorname{L} d - \operatorname{L} 2 - c \le \operatorname{L} n - \operatorname{L}_3 n - \operatorname{L} d - c,$$

the bounding expression being the one in Proposition 2.3 with $c_n \equiv e^c d$. Thus, by that proposition,

$$\begin{aligned} \mathbb{P}(F_{n_j}^- \le b_{n_{j+1}}) &\le \mathbb{P}(F_{n_j}^- \le \operatorname{L} n_j - \operatorname{L}_3 n_j - \operatorname{L} d - c) \\ &= O((\operatorname{L} n_j)^{-[e^c d-(d-1)]}) = O(j^{-[1+(e^c-1)d]}), \end{aligned}$$

which is summable. ◀

C Proofs of Theorems 4.3 and 4.5 and Lemma 5.2

Proof of Theorem 4.3. First suppose $g \in I$. Then (i) is automatic from the definition of I. Moreover, we know from our earlier discussion of generators that (ii) holds for $g = \left(R_1^{(\Pi_1)}, \dots, R_d^{(\Pi_d)}\right)$ with the possible exception of the second equality in (4.2). But if that equality does not hold, let $j, \ell \in [d]$ with $j \neq \ell$ satisfy

$$r_j^{(i_\ell)} < R_j^{(\Pi_j)}.$$

We then move i_ℓ from the cell Π_ℓ to the cell Π_j in order to form a new partition, call it Π'. Then

$$g > \left(R_1^{(\Pi'_1)}, \dots, R_d^{(\Pi'_d)}\right) \in S,$$

so g is not a generator.

Next we prove the converse. If g has these two properties, then $g \in (0,\infty)^d$ belongs to S, so all that is left to show is that g is a minimum (with respect to $\leq$) of S. Suppose that $x < g$; we will complete the proof by showing that $x \notin S$.

Let j_0 satisfy $x_{j_0} < g_{j_0}$. Then

$$x_{j_0} < g_{j_0} = r_{j_0}^{(i_{j_0})} \tag{C.1}$$

using (4.2) for the equality. Additionally, for $j \neq j_0$ we have

$$x_j \leq g_j < r_j^{(i_{j_0})}, \tag{C.2}$$

where the second inequality holds by (4.2) because

$$g_j = r_j^{(i_j)} = \min\left\{r_j^{(i_\ell)} : \ell \in [d]\right\},$$

which almost surely is strictly smaller than $r_j^{(i_{j_0})}$ because $i_j \neq i_{j_0}$. Combining (C.1) and (C.2), we see that $x \prec r^{(i_{j_0})}$, and so $x \notin S$. ◀

Proof of Theorem 4.5. Let $t = |T|$. There is no loss of generality (and there is some ease in notation) in supposing that $T = [t]$, and thus $x \in G_T$ if and only if $x \in G$ and $x_{t+1} = \cdots = x_d = 0$. Let $x = (x_1, \dots, x_t, 0, \dots, 0)$ satisfy $\nu(x) = t$. We will show that $x \in G_T$ – equivalently, that $x \in G$ – if and only if $\pi_T(x) \in I_T$ – equivalently, that $x \in \iota_T(I_T) = G'_T$.

Indeed, for x to be a generator, there are two requirements: (i) $x \in S$, and (ii) x is a minimum of S. The requirement (i) is that for each i there should exist $j \in [d]$ such that $x_j \geq r_j^{(i)}$. However, since we assume that $r^{(i)} \succ 0$, such j must belong to $[t]$. We have thus argued that x is in $S = \mathrm{RS}(R)$ (the record-setting region determined by the points in R) if and only if $\pi_T(x) \in \mathrm{RS}(R_T)$.

The requirement (ii) is that $y < x$ must imply $y \notin S$. But note that $y < x$ if and only if y is of the form $y = (y_1, \dots, y_t, 0, \dots, 0)$ with $\pi_T(y) < \pi_T(x)$. Thus requirement (ii) can be rephrased thus: If $y = (y_1, \dots, y_t, 0, \dots, 0)$ with $\pi_T(y) < \pi_T(x)$, then $y \notin \mathrm{RS}$ – equivalently, by what we argued in connection with requirement (i), that $\pi_T(y) \notin \mathrm{RS}(R_T)$.

So we have argued that x is a generator if and only if $\pi_T(x) \in I_T$, i.e., if and only if $x \in G'_T$. This is as desired. ◀

Proof of Lemma 5.2. To facilitate the statement and proof of Lemma 5.2, and in order to follow more closely the analogous treatment of remaining records in [2, Section 2], we may and do switch from Exponential(1) observation coordinates to observations uniformly distributed in $[0,1)^d$. Referring to Theorem 4.3(ii), let us say that the d-tuple $(X^{(i_1)}, \dots, X^{(i_d)})$ of observations (where the indices i_j are distinct elements of $[n]$) *generates* an epoch-n interior generator g if those d observations are all remaining records at epoch n and

$$g_j = X_j^{(i_j)} = \min\{X_j^{(i_\ell)} : \ell \in [d]\} \text{ for every } j \in d.$$

Note that every interior generator is generated by precisely one such generating d-tuple. Thus $\mathbb{E}\,\iota_{d,n}$ equals $n^{\underline{d}}$ times the probability that $(X^{(1)}, \dots, X^{(d)})$ generates an interior generator. Condition on the value $\mathbf{y} := (x^{(1)}, \dots, x^{(d)})$ of this d^2-tuple. According to Theorem 4.3, in order for $\mathbf{y}$ to generate an interior generator, two conditions are required. One is that

$$x_j^{(\ell)} \geq x_j^{(j)} \text{ for every } \ell, j \in d \text{ with } \ell \neq j. \tag{C.3}$$

Let $x := \left(x_1^{(1)}, \dots, x_d^{(d)}\right)$. The other condition is that the remaining $n-d$ observations each need to fall outside O_x^+, guaranteeing the condition $x \in S$ required by Theorem 4.3(i).

Therefore,

$$\mathbb{E}\,\iota_{d,n} = n^{\underline{d}} \int_{\mathbf{y}:\text{(C.3) holds}} \left[1 - \prod(1 - x_j^{(j)})\right]^{n-d} \mathrm{d}\mathbf{y},$$

a d^2-dimensional integral which reduces effortlessly to a d-dimensional integral:

$$\begin{aligned}\mathbb{E}\,\iota_{d,n} &= n^{\underline{d}} \int_{[0,1)^d} \left[\prod(1 - x_j)\right]^{d-1} \left[1 - \prod(1 - x_j)\right]^{n-d} \mathrm{d}x \\ &= n^{\underline{d}} \int_{(0,1]^d} x_\times^{d-1} (1 - x_\times)^{n-d} \, \mathrm{d}x,\end{aligned}$$

as desired. ◀