



Some Bounds for the Expectations of Functions on Order Statistics and Their Applications

Arvydas Astrauskas¹

Received: 24 November 2021 / Revised: 29 March 2022 / Accepted: 27 April 2022 /
Published online: 5 June 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Let $X_{1,N} \geq X_{2,N} \geq \dots \geq X_{N,N}$ be the order statistics of independent identically distributed random variables X_k ($1 \leq k \leq N$). For fixed natural K and a nonnegative bounded deterministic function G_N on \mathbb{R}^N satisfying mild conditions of Lebesgue's measurability, we obtain the following bound for the expectations:

$$\begin{aligned} & \mathbb{E}G_N(X_{1,N}, X_{2,N}, \dots, X_{K,N}, X_{K+1,N}, \dots, X_{N,N}) \\ & \leq T \cdot \mathbb{E}G_N(X_{1,N}^{(1)}, X_{1,N}^{(2)}, \dots, X_{1,N}^{(K)}, X_{K+1,N}, \dots, X_{N,N}) + \vartheta_T \end{aligned}$$

for any $T \geq T_0(K)$ and any $N \geq N_0(T)$ large enough; here constants $\vartheta_T > 0$ tend to zero as T approaches infinity; $X_{1,N}^{(i)}$ ($1 \leq i \leq K$) are mutually independent copies of the maximum $X_{1,N}$; and each $X_{1,N}^{(i)}$ is also independent of the sample $\{X_k\}_{1 \leq k \leq N}$. With G_N as relevant indicator functions and $N \rightarrow \infty$, these bounds are applied to study o- and O-type asymptotic properties of the following functions on order statistics: (Appl-1) the numbers of observations near the K th extremes $X_{K,N}$ and (Appl-2) the sums of negative powers of spacings $X_{K,N} - X_{i,N}$ ($K + 1 \leq i \leq N$).

Keywords Functions on order statistics · Extremes · Expectation · Approximation of multidimensional integrals · Number of near-extremes · Weak law of large numbers

Mathematics Subject Classification 2020 60E15 · 60G70 · 60F05 · 60F99 · 62G30 · 28A25 · 41A45

✉ Arvydas Astrauskas
arvydas.astrauskas@mif.vu.lt

¹ Institute of Data Science and Digital Technologies, Vilnius University,
Akademijos str. 4, LT-08412 Vilnius, Lithuania

1 Introduction: Bounds for Expectations of Functions on Order Statistics and their Motivation

Throughout the paper, let X, X_k ($k \in \mathbb{N} := \{1, 2, \dots\}$) be independent identically distributed (i.i.d.) random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with the common hazard rate function $Q(x) := -\log \mathbb{P}(X > x)$ and the right endpoint $x^* := \sup\{x : Q(x) < \infty\} \leq \infty$. (Sometimes we use the additional notation $\bar{F} := e^{-Q}$ for the tail distribution function). Denote by

$$X_{1,N} \geq X_{2,N} \geq \dots \geq X_{N,N} \quad (1.1)$$

the order statistics (variational series) based on the sample $\{X_k\}_{1 \leq k \leq N}$. For fixed K , the terms $X_{K,N}$ are referred to as the K th extremes.

In the present paper we derive, for any large N , asymptotic bounds for the expectation $\mathbb{E}G_N(X_{1,N}, X_{2,N}, \dots, X_{N,N})$ for an arbitrary nonnegative bounded function G_N on \mathbb{R}^N satisfying certain conditions of (Lebesgue's) measurability, when the largest order statistics essentially contribute to the expectation of the function G_N . Loosely speaking, the aforementioned expectation is bounded by a similar expectation multiplied by a large constant T but with the first K extremes replaced by independent versions of maxima, making asymptotic formulas for such expectations less complicated (see Proposition 1.1 in the present section). In the important case when G_N are indicator functions based on order statistics, the corresponding bounds hold true for distributions, rather than expectations.

In the second part of the paper, we apply the present bounds for expectations and distributions (i.e. Proposition 1.1) to investigate the following functions on order statistics, as the sample size N tends to infinity:

(Appl-1) In Sect. 2, we study o- and O-type asymptotic properties in probability of *the numbers of observations near the K th extremes* (shortly, numbers of near-extremes) defined as the cardinality $\mathfrak{N}_{K:N}(IA_N)$ of the subset of the sample observations X_i ($1 \leq i \leq N$) that fall into the scaled interval $X_{K,N} + IA_N$, where $A_N > 0$ are normalizing constants and $I \subset \mathbb{R}$ is a bounded interval such that $0 \in \bar{I}$, the closure of I ; see Theorem 2.1. In Theorem 2.1, the corresponding o- and O-type conditions on regular variation of the function $\bar{F} = e^{-Q}$ at the right endpoint x^* are given in terms that are closely related to assumptions of the maximum domain of attraction for i.i.d. samples. In the proof, we essentially explore key formula (2.1) for the random variables $\mathfrak{N}_{K:N}(IA_N)$. Our Theorem 2.1 extends the corresponding results of Pakes and Steutel [37], who have considered the numbers of near-maxima when $K = 1$, $A_N \equiv \text{const} > 0$ and $I = (-1, 0]$, with continuous Q . The numbers of near-extremes have been extensively studied in the mathematical and statistical literature because of several applications including:

- (I) actuarial mathematics that deals with models of insurance claims; see, e.g. [25, 26, 32].
- (II) estimation theory that deals with estimating the distributional parameters including the tail index and the right endpoint of the random sample; see, e.g. [25–27, 33].

See also [38] for a physical background and further applications of the numbers of near-extremes, and [16, 34] for a brief review and a comprehensive reference list on the subject including distributional properties and limit theorems for the numbers of near-extremes and related quantities.

(Appl-2) Section 3 presents the weak law of large numbers for the sums $S_{K:N}^{(r)}$ (3.1) of negative powers of spacings $X_{K,N} - X_{i,N}$ ($K + 1 \leq i \leq N$) with fixed K and $N \rightarrow \infty$, assuming a finiteness of certain expectations and regular variation of the function $\bar{F} = e^{-Q}$ at x^* in terms of singular integrals; see Theorem 3.1. These results are crucial to the extreme value theory for eigenvalues of rank-one perturbations of random diagonal matrices (i.e. random mean-field Hamiltonians) and the large-time intermittency theory for the related branching random walks in random potential $\{X_k\}_{k \in \mathbb{N}}$. For a physical background of the random mean field models, see [13, 14]. Asymptotic properties of the sums $S_{K:N}^{(r)}$ (3.1) and their applications to the aforementioned models have been earlier studied in the following papers: [14] dealing with the case of Gaussian i.i.d. X_k ($k \in \mathbb{N}$); [22, 23] treating the case of i.i.d. $X_k = \eta_k$ ($k \in \mathbb{N}$) with the exponential distribution; finally, [4, 6] treating the case of i.i.d. variables X_k ($k \in \mathbb{N}$) with the tail distribution $\bar{F} = e^{-Q}$ satisfying certain conditions on smoothness and (integral-type) regular variation at x^* . Theorem 3.1 of the present paper extends the corresponding results from [4, 6, 14, 22, 23].

It is worth noting that, in the aforementioned articles, the functions on order statistics as in (Appl-1) and (Appl-2) have been introduced to study various aspects of the asymptotic structure of the sample observations that fall into a neighbourhood of the extreme order statistics. See Sects. 2 and 3 for more discussions on applications of the functions on order statistics under consideration, with emphasis on the discrete distributions.

In Remark 2.5 of Sect. 2, we provide the key steps in the proof of the results in (Appl-1) and (Appl-2). The methods of our paper (i.e. application of Proposition 1.1 dealing with asymptotic bounds for expectations of relevant indicator functions based on order statistics; truncation of independent maxima; Lebesgue's bounded convergence theorem, etc.) can be also applied to study the asymptotic behaviour of the following functions on order statistics: spacings and ratios of extreme order statistics, ratios of the sums of positive i.i.d. random variables to their extreme terms, multiplicity of extremes of discrete random variables, etc. These and many others functions on order statistics and their asymptotic properties play an important role in several areas of applications, including mathematical statistics, finance and insurance, meteorology and hydrology [3, 10, 18, 20, 30], as well as statistical physics and statistical mechanics [5, 24, 29, 38], among others.

Throughout the paper, we explore the following *representation of order statistics* (1.1) of the i.i.d. sample $\{X_k\}_{1 \leq k \leq N}$: For the hazard rate function Q and the right endpoint $x^* \leq \infty$ of the random variable X , we define the generalized inverse Q^{\leftarrow} by

$$Q^{\leftarrow}(y) := \inf \{x : Q(x) \geq y\} \quad (y > 0) \quad (1.2)$$

(thus, Q^{\leftarrow} is a finite-valued left-continuous nondecreasing function on $\mathbb{R}_+ := (0, \infty)$ such that $Q^{\leftarrow}(y) \rightarrow x^*$ as $y \rightarrow \infty$). Let η_k ($k \in \mathbb{N}$) be i.i.d. random vari-

ables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with the exponential distribution function $\mathbb{P}(\eta_1 > x) = e^{-x}$, $x > 0$. (Shortly, $\{\eta_k\}_{k \in \mathbb{N}}$ is referred to as i.i.d. $Exp(1)$ sequence). Denote by $\eta_{1,N} > \eta_{2,N} > \dots > \eta_{N,N} > 0$ the order statistics based on the sample $\{\eta_k\}_{1 \leq k \leq N}$; here the inequalities are strict with probability 1, because of the continuity of the exponential distribution. Then, according to Lemma 4.1.9 from [20], we obtain that $X_k \stackrel{d}{=} Q^{\leftarrow}(\eta_k)$ ($k \in \mathbb{N}$) and

$$(X_{1,N}, X_{2,N}, \dots, X_{N,N}) \stackrel{d}{=} (Q^{\leftarrow}(\eta_{1,N}), Q^{\leftarrow}(\eta_{2,N}), \dots, Q^{\leftarrow}(\eta_{N,N})), \tag{1.3}$$

where $\stackrel{d}{=}$ means that the distributions of the random vectors in both sides of (1.3) are identical.

Throughout the paper, we also use the following abbreviations. Let \mathbb{R}^N be the N -dimensional real vector space, and let $\mathbb{R}_+^N \subset \mathbb{R}^N$ be the N -dimensional cone of vectors with positive coordinates. Let \mathbb{I}_U stand for the indicator function of the set U , i.e. $\mathbb{I}_U(x) = 1$ if $x \in U$ and $\mathbb{I}_U(x) = 0$ otherwise. By $T_0 \in \mathbb{R}_+$, $N_0 \in \mathbb{R}_+$, etc., we denote various large numbers, values of which may change from one appearance to the next. Similarly, const , const' , etc., stand for various positive constants. Let sequences $\{X_k\}_{k \in \mathbb{N}}$, $\{X_k^{(i)}\}_{k \in \mathbb{N}}$ ($i = 1, 2, \dots$) be independent copies of the i.i.d. sequence with the common hazard rate function Q , and write $X_{1,N}^{(i)}$ for the maxima of the random variables $X_k^{(i)}$ ($1 \leq k \leq N$). (The same notion is transferred to i.i.d. $Exp(1)$ sequences). Finally, for $N \in \mathbb{N}$ and Q^{\leftarrow} given by (1.2), define the function $q: \mathbb{R}_+^N \rightarrow \mathbb{R}^N$ by $q(\mathbf{y}) := (Q^{\leftarrow}(y_1), Q^{\leftarrow}(y_2), \dots, Q^{\leftarrow}(y_N))$ for any $\mathbf{y} = (y_1, y_2, \dots, y_N) \in \mathbb{R}_+^N$. With these abbreviations, we now state the following proposition.

Proposition 1.1 Fix constants $C \in \mathbb{R}_+$ and $K \in \mathbb{N}$, and assume that $G_N: \mathbb{R}^N \rightarrow [0, C]$ is a deterministic function such that $G_N \circ q(\cdot) := G_N(q(\cdot))$ is measurable on \mathbb{R}_+^N . Then, for any $T \geq T_0(K)$ and any $N \geq N_0(T)$,

$$\begin{aligned} &\mathbb{E}G_N(X_{1,N}, X_{2,N}, \dots, X_{K,N}, X_{K+1,N}, \dots, X_{N,N}) \\ &\leq T \cdot \mathbb{E}G_N(X_{1,N}^{(1)}, X_{1,N}^{(2)}, \dots, X_{1,N}^{(K)}, X_{K+1,N}, \dots, X_{N,N}) + C \cdot \varepsilon_T \end{aligned} \tag{1.4}$$

and

$$\begin{aligned} &\mathbb{E}G_N(X_{1,N}, X_{2,N}, \dots, X_{K,N}, X_{K+1,N}, \dots, X_{N,N}) \\ &\leq TN^K \cdot \mathbb{E}G_N(X_1^{(1)}, X_1^{(2)}, \dots, X_1^{(K)}, X_{K+1,N}, \dots, X_{N,N}) + C \cdot \tilde{\varepsilon}_T; \end{aligned} \tag{1.5}$$

here constants $\varepsilon_T > 0$ and $\tilde{\varepsilon}_T > 0$ depend only on T and K such that $\varepsilon_T \rightarrow 0$ and $\tilde{\varepsilon}_T \rightarrow 0$ as $T \rightarrow \infty$.

Bound (1.4) tells us that, in the expectation of functions on order statistics, the first K largest order statistics can be replaced by the independent versions of maximum at the cost of large multiplier T plus the small error $C\varepsilon_T$. Meanwhile, bound (1.5) means that the first K largest order statistics can be simply replaced by the i.i.d. random variables at the cost of large multiplier TN^K plus the small error $C\tilde{\varepsilon}_T$. Recall that the order

statistics $X_{k,N}$ ($k = 1, 2, \dots$) are strongly correlated random variables. Therefore, one may think of the assertions of Proposition 1.1 as a simplification of distributions or expectations of functions on order statistics in large N limit, especially for the functions mentioned in (Appl-1)–(Appl-2) above and for the functions that depend only on the extreme order statistics. On the other hand, bounds (1.4)–(1.5) are relevant tools for the investigation of o- and O-type asymptotic properties in probability (for instance, the weak laws of large numbers, stochastic boundedness and compactness, etc.) of functions on order statistics, rather than the weak convergence of these quantities to a limit with a nondegenerate distribution.

Proof of Proposition 1.1 is given in Sect. 4 and relies on the following two observations:

(J) Consider the function $G_N(\eta_{1,N}, \dots, \eta_{N,N})$ on the exponential order statistics. For any N , the function G_N is approximated in L^1 by a step function defined as a finite linear combination of indicator functions \mathbb{I}_{I_k} of rectangles $I_k \subset \mathbb{R}^N$ with sides parallel to the axes of \mathbb{R}^N (cf. Lemma 4.1). Therefore, the expectation $\mathbb{E}G_N(\eta_{1,N}, \dots, \eta_{N,N})$ is approximated by a finite linear combination of the expectations $\mathbb{E}\mathbb{I}_{I_k}(\eta_{1,N}, \dots, \eta_{N,N})$;

(JJ) Since \mathbb{I}_{I_k} is the function with separated variables, we repeatedly apply the Markov property of the sequence $\eta_{N-l+1,N}$ ($l = 1, 2, \dots, N$) (cf. Lemma 4.4) and the exact asymptotic bounds for the extremes $\eta_{K,N}$ (cf. Lemma 4.2) to derive the bounds for $\mathbb{E}\mathbb{I}_{I_k}(\eta_{1,N}, \dots, \eta_{N,N})$ like in (1.4) and (1.5) (cf. Lemma 4.3). This together with the approximation results from part (J) yields the assertions of Proposition 1.1.

In Proposition 1.1 with the i.i.d. $Exp(1)$ random variables $X_k = \eta_k$ ($k \in \mathbb{N}$) (so that q is the identity function), it suffices to claim the measurability of G_N itself. The following proposition provides the classes of measurable functions that are useful for applications.

Proposition 1.2 (I) For $M \in \mathbb{N}$, if $g: \mathbb{R}^N \rightarrow \mathbb{R}^M$ is continuous, then $g \circ q: \mathbb{R}_+^N \rightarrow \mathbb{R}^M$ is measurable.

(II) If $h: \mathbb{R}^N \rightarrow \mathbb{R}^M$ is measurable and $B \subset \mathbb{R}^M$ is a Borel set, then $\mathbb{I}_{\{x \in \mathbb{R}^N: h(x) \in B\}}$ is a measurable function. Moreover, the sum and product of two finite-valued measurable functions are both measurable.

Remark 1.3 For fixed constants K and T , and N tending to infinity, the right-hand side of (1.5) is bounded from above by a constant $\text{const}(K, T)$ independent of N . (However, $\text{const}(K, T)$ becomes arbitrarily large as T approaches infinity).

Proof of Remark 1.3 For simplicity of notation, consider the function $G_N(\eta_{1,N}, \dots, \eta_{N,N})$ on the exponential order statistics. We first notice that the support of G_N in (1.5) can be restricted to the cone $\mathbb{R}_{con}^N := \{(x_1, x_2, \dots, x_N) \in \mathbb{R}^N: x_1 > x_2 > \dots > x_N > 0\}$. Consequently, since $0 \leq G_N \leq C$ and the random variables $\eta_1^{(1)}, \eta_1^{(2)}, \dots, \eta_1^{(K)}, \eta_{K+1,N}$ are mutually independent, we obtain that the right-hand side of (1.5) does not exceed

$$CTN^K \mathbb{P}(\eta_1^{(1)} > \dots > \eta_1^{(K)} > \eta_{K+1,N}) + C\tilde{\varepsilon}_T \leq CTN^K \mathbb{E}(e^{-K\eta_{K+1,N}}) + C\tilde{\varepsilon}_T.$$

Here the random variable $e^{-\eta_{K+1,N}}$ is beta-distributed with the density

$$(K + 1) \binom{N}{K + 1} x^K (1 - x)^{N-K-1} \quad (0 < x < 1).$$

Therefore, simple calculations of beta functions show that $\mathbb{E}(e^{-K\eta_{K+1,N}})$ is of the order N^{-K} ; in particular, $\mathbb{E}(e^{-K\eta_{K+1,N}}) \leq (2K)!N^{-K}$ for all $N \geq K + 1$. These estimates imply that, for any N large enough, the right-hand side of (1.5) does not exceed constant $\text{const}(K, T) > 0$ depending only on K and T , as claimed. \square

Remark 1.4 Notice that the values of constants $T_0(K)$, $N_0(T)$, ε_T and $\tilde{\varepsilon}_T$ introduced in Proposition 1.1 can be found in Sect. 4 (the proof of Proposition 1.1). However, a derivation of the optimal values of these constants is beyond the scope of the present paper. \square

The organization of the remaining sections of the paper is as follows: Sect. 2 provides o- and O-type asymptotic results in probability for the numbers of observations near extremes. In Sect. 3, we state the weak law of large numbers for the sums of negative powers of spacings. The results of Sect. 1, i.e. Propositions 1.1 and 1.2, are proved in Sect. 4. The main result of Sect. 2, i.e. Theorem 2.1, is proved in Sect. 5. Finally, Theorem 3.1 of Sect. 3 is derived in Sect. 6.

2 Applications to the Numbers of Near-Extremes

Let X, X_k ($k \in \mathbb{N}$) be the i.i.d. random variables with the common distribution function $1 - e^{-Q}$ and the right endpoint $x^* \leq \infty$. Let $X_{k,N}$ ($1 \leq k \leq N$) be the order statistics defined in (1.1)–(1.3). By \mathcal{I}_0 we denote a collection of bounded intervals $I \subset \mathbb{R}$ (closed, open, semiclosed) such that $0 \in \bar{I}$, the closure of I . The main object of the present section is the number of near-extremes defined by

$$\begin{aligned} \mathfrak{N}_{K:N}(IA) &:= \#\{i \in [1, N] \setminus \{K\} : X_{i,N} \in X_{K,N} + IA\} \\ &= \sum_{i=1, i \neq K}^N \mathbb{I}_{\{X_{i,N} \in X_{K,N} + IA\}}, \end{aligned} \tag{2.1}$$

where $I \in \mathcal{I}_0$, $K \in \mathbb{N}$ and $A = A_N \in \mathbb{R}_+$ are positive constants that may or not may depend on N . Here $x + IA$ stands for the shifted and scaled interval $\{x + yA : y \in I\}$. For $I = \{0\}$, the random variable $\mathfrak{M}_{K:N} := \mathfrak{N}_{K:N}(\{0\})$ denotes the multiplicity of the K th order statistics. For $K = 1$ and $I = (-1, 0]$, Eq. (2.1) appears in the context of insurance mathematics: If $X_{1,N}$ denotes the maximum of N claims X_1, X_2, \dots, X_N , then $\mathfrak{N}_{1:N}((-A, 0])$ gives the number of claims close to $X_{1,N}$; see e.g. [25, 32, 37].

The asymptotic behaviour of sequence (2.1) (as $N \rightarrow \infty$) has been first studied by Pakes and Steutel [37] in the case of $K = 1$ and $I = (-1, 0]$. In particular, they have considered $\mathfrak{N}_{1:N}((-A, 0])$, the numbers of observations near maxima with a fixed interval length A , by assuming that $x^* = \infty$ and Q is the continuous function satisfying the following condition: there exists a constant $0 \leq \rho \leq \infty$ such that

$$\lim_{x \rightarrow \infty} Q(x+c) - Q(x) = \rho c \quad \text{for any } c \in \mathbb{R}_+. \quad (2.2)$$

Here $\rho = 0$ (resp., $\rho = \infty$) refers to the heavy tails (resp., light tails) of distributions. Pakes and Steutel [37] have established that $\mathfrak{N}_{1:N}((-A, 0]) \xrightarrow{\mathbb{P}} 0$ as $N \rightarrow \infty$ when $\rho = 0$, and $\mathfrak{N}_{1:N}((-A, 0]) \xrightarrow{\mathbb{P}} \infty$ as $N \rightarrow \infty$ when $\rho = \infty$; here $\xrightarrow{\mathbb{P}}$ denotes the convergence in probability. Moreover, the sequence $\mathfrak{N}_{1:N}((-A, 0])$ converges weakly (as $N \rightarrow \infty$) to a geometrically distributed random variable when $0 < \rho < \infty$. The latter limit theorem (i.e. the weak convergence to a nondegenerate random variable) has been extended and generalized by many authors to the numbers of observations near extreme order statistics with N -dependent or -independent interval lengths, where the distribution $1 - e^{-Q}$ is typically assumed to be continuous and lay in the maximum domain of attraction (*MDA*); e.g. [8, 28, 34–37]. See also [16, 34] for a review of limit theorems for the numbers of observations near order statistics and references therein. Recall that, by the definition, $1 - e^{-Q} \in MDA$ if and only if there exist constants $A_N \in \mathbb{R}_+$ and $B_N \in \mathbb{R}$ such that the sequence $(X_{1,N} - B_N)/A_N$ converges weakly to a nondegenerate random variable. From Chapter 1 in [18] we know that $1 - e^{-Q} \in MDA$ if and only if there exists a positive auxiliary function $h(\cdot) > 0$ on $(-\infty, x^*)$ such that

$$\lim_{x \uparrow x^*} (Q(x + ch(x)) - Q(x)) = \frac{1}{\gamma} \log(1 + \gamma c) \quad (2.3)$$

for some $\gamma \in \mathbb{R}$ and for any c with $1 + \gamma c > 0$.

Notice that, for $\gamma = 0$, the right-hand side of (2.3) is interpreted as c . Moreover, for $0 < \rho < \infty$, condition (2.2) coincides with (2.3) where $\gamma = 0$ and $h(\cdot) \equiv 1/\rho$. On closer inspection of Sections 1.1.3 and 1.2 in [18], one finds that $1 - e^{-Q} \in MDA$ if and only if (2.3) is fulfilled with $h = a(e^Q)$ for some continuous auxiliary function $a(\cdot) > 0$ on \mathbb{R}_+ such that there exists a finite and positive limit $\lim_{x \rightarrow \infty} a(cx)/a(x)$ for any $c > 0$ (i.e. $a(\cdot)$ is Karamata's regularly varying at infinity; see [12]). In this case, the normalizing constants $A_N > 0$ in limit theorems for the sequences $(X_{1,N} - B_N)/A_N$ and $\mathfrak{N}_{K:N}(IA_N)$ are allowed to be proportional to the quantities $a(N)$ defined above; cf. also [34, 37]. For instance, using the expressions for distributions of the random variables $\mathfrak{N}_{K:N}((-A_N, 0))$ in terms of certain distributions of the normalized spacings $(X_{K,N} - X_{j,N})/A_N$ ($j > K$), Nagaraja et al. [34] have derived limit theorems for $\mathfrak{N}_{K:N}((-A_N, 0))$ from that for the spacings. The inverse statements, including the derivation of limit theorems for the spacings from that for the numbers of near-maxima, have been discussed in [37].

The above considerations serve as a starting point for the O- and o-type asymptotic results of the present section on the numbers of near-extremes. Let us introduce the related abbreviations and notations. For functions $f(\cdot) \geq 0$ and $g(\cdot) > 0$ on $(-\infty, x^0)$, we write $f(x) = O(g(x))$ (resp., $f(x) \asymp g(x)$) as $x \uparrow x^0$, if the ratios $f(x)/g(x)$ (resp., $f(x)/g(x) + g(x)/f(x)$) are bounded from above for all x close to but less than x^0 . A measurable function $f(\cdot) > 0$ on \mathbb{R}_+ is said to be in the class *ORV*, if $f(cx) \asymp f(x)$ as $x \rightarrow \infty$, for any $c \in \mathbb{R}_+$. For a detailed account of alternative definitions and properties of $f \in ORV$, see [1] or Chapter 2 in [12]. For random

variables Y_N ($N \in \mathbb{N}$), we write $Y_N \stackrel{\mathbb{P}}{=} O(1)$ as $N \rightarrow \infty$, if $\limsup_N \mathbb{P}(|Y_N| > T) \rightarrow 0$ as $T \rightarrow \infty$. Moreover, $Y_N \xrightarrow{\mathbb{P}} C$ means the convergence in probability of Y_N to $C \in \mathbb{R} \cup \{\infty\}$. Finally, $Q^-(x) := Q(x-)$ denotes the left-continuous version of Q .

Theorem 2.1 *Let $Q^-(x^*) = \infty$. Fix an arbitrary positive function $a(\cdot) \in ORV$, and write $A_N := a(N)$ ($N \in \mathbb{N}$). Then for $K \in \mathbb{N}$ and intervals $I \in \mathcal{I}_0$ fixed arbitrarily, we have the following limits in probability for the random variables $\mathfrak{N}_{K:N}(IA_N)$ as $N \rightarrow \infty$:*

- (i) *If $Q(x) - Q(x - ca(e^{Q(x)})) = O(1)$ as $x \uparrow x^*$, for any $c \in \mathbb{R}_+$, then $\mathfrak{N}_{K:N}(IA_N) \stackrel{\mathbb{P}}{=} O(1)$.*
- (ii) *If $Q(x) - Q(x - ca(e^{Q(x)})) \rightarrow 0$ as $x \uparrow x^*$, for any $c \in \mathbb{R}_+$, then $\mathfrak{N}_{K:N}(IA_N) \xrightarrow{\mathbb{P}} 0$.*
- (iii) *Let*

$$Q^-(x) - Q^-(x - ca(e^{Q^-(x)})) \rightarrow \infty \text{ as } x \uparrow x^*, \text{ for any } c \in \mathbb{R}_+, \tag{2.4}$$

and assume additionally that either $Q(x) - Q^-(x) = O(1)$ as $x \uparrow x^$, or the function $a(\cdot)$ is nondecreasing on \mathbb{R}_+ ; then*

$$\mathfrak{N}_{K:N}(IA_N) \xrightarrow{\mathbb{P}} \infty,$$

provided $(-I) \cap \mathbb{R}_+$ is an interval of positive length.

Theorem 2.1 extends the corresponding results in [37], mentioned at the beginning of this section.

Example 2.2 (*von Mises distribution*) Consider the function $Q(x) = x + \sin x$ ($x \geq 0$). Notice that the distribution $1 - e^{-Q}$ does not belong to the maximum domain of attraction according to Exercise 1.18 in [18], pp. 36. However, the function Q satisfies the corresponding conditions of Theorem 2.1(i), (ii) and (iii), provided $a(x) = O(1)$, $a(x) \rightarrow 0$ and $a(x) \rightarrow \infty$, respectively. \square

Another important example is the i.i.d. random variables X, X_k ($k \in \mathbb{N}$) with discrete distribution. In this case, we are interested in the multiplicity $\mathfrak{M}_{K:N} := \mathfrak{N}_{K:N}(\{0\})$ of the K th extreme order statistics $X_{K,N}$. Parts (i) and (ii) of Theorem 2.1 immediately imply, respectively, parts (i) and (ii) of the following corollary.

Corollary 2.3 (The multiplicity $\mathfrak{M}_{K:N}$ of the K th extreme order statistics in the case of discrete distributions with heavy and medium tails) *Assume that $\mathbb{P}(X \in \mathbb{N}) = 1$ and $x^* = \infty$. Then for fixed $K \in \mathbb{N}$, we have the following limits in probability for the random variables $\mathfrak{M}_{K:N} := \mathfrak{N}_{K:N}(\{0\})$ as $N \rightarrow \infty$:*

- (i) *If $Q(x) - Q(x - 1) = O(1)$ as $x \rightarrow \infty$, then $\mathfrak{M}_{K:N} \stackrel{\mathbb{P}}{=} O(1)$;*
- (ii) *If $Q(x) - Q(x - 1) \rightarrow 0$ as $x \rightarrow \infty$, then $\mathfrak{M}_{K:N} \xrightarrow{\mathbb{P}} 0$.* \square

For $K = 1$, the assertion in part (ii) of Corollary 2.3 is well-known in mathematical literature, where the necessity of the condition in (ii) is also established; e.g. [9, 15]. Notice that case (ii) refers to the heavy tails e^{-Q} , i.e. $Q(x) = o(x)$ as $x \rightarrow \infty$.

Remark 2.4 (The multiplicity $\mathfrak{M}_{K:N}$ of the K th extreme order statistics in the case of discrete distributions with light tails.)

- (j) Assume that $\mathbb{P}(X \in \mathbb{N}) = 1$ and $x^* < \infty$. Then, for any natural $M \geq 2$, the sequence $\mathbb{P}(\mathfrak{M}_{1:N} \leq M - 2) = \mathbb{P}(X_{1,N} - X_{M,N} \geq 1)$ tends to zero (therefore, $\mathfrak{M}_{K:N} \xrightarrow{\mathbb{P}} \infty$), according to the obvious statement that almost surely $X_{k,N} = x^*$ for any $N \geq N_0(\omega)$ with $k \in \mathbb{N}$ fixed arbitrarily.
- (jj) Assume now that $\mathbb{P}(X \in \mathbb{N}) = 1, x^* = \infty$ and

$$\limsup_{x \rightarrow \infty} (Q(x) - Q(x - 1)) > 0. \tag{2.5}$$

[Class (2.5) includes the functions Q with $Q(x) - Q(x - 1) \rightarrow \infty (x \rightarrow \infty)$ referring to the light tails e^{-Q}]. Then Theorem 1.3 from [7] yields that, for any $m \in \mathbb{N} \cup \{0\}$, the limit $\lim_N \mathbb{P}(\mathfrak{M}_{1:N} = m)$ does not exist. Therefore, $\limsup_N \mathbb{P}(\mathfrak{M}_{1:N} \leq m) > 0$ and $\limsup_N \mathbb{P}(\mathfrak{M}_{1:N} > m) > 0$. These limits imply that the sequence $\mathfrak{M}_{1:N}$ does not approach infinity with probability close to one. In particular, for the discrete case with $Q(x) - Q(x - 1) \rightarrow \infty (x \rightarrow \infty)$, the analogue of Theorem 2.1(iii) does not hold.

For discrete distributions with light tails, clustering properties of the extreme order statistics have been studied in [2, 7, 39], among others. □

Remark 2.5 Proof of Theorem 2.1 is given in Sect. 5. We now present the key steps of the proof of Theorem 2.1(i) in the case of $K = 1$ and $I = (-1, 0]$, i.e. for the number $\mathfrak{N}_{1:N} := \mathfrak{N}_{1:N}((-A_N, 0])$ of observations near maxima:

- (I) (Application of Proposition 1.1 dealing with asymptotic bounds for expectations of relevant indicator functions based on order statistics). Introduce the auxiliary quantities $\tilde{\mathfrak{N}}_{1:N} := \#\{1 \leq i \leq N : \tilde{X}_{1,N} - A_N < X_i \leq \tilde{X}_{1,N}\}$, the number of observations near independent maxima $\tilde{X}_{1,N}$. Using bound (1.4) in Proposition 1.1 with $K = 1$ and $G_N(X_{1,N}, \dots, X_{N,N}) := \mathbb{I}_{\{\mathfrak{N}_{1:N} > T\}}$ for $T > 0$, we obtain the following bound for any small $\varepsilon > 0$:

$$\begin{aligned} \limsup_N \mathbb{P}(\mathfrak{N}_{1:N} > T) &\leq \varepsilon^{-1} \limsup_N \mathbb{P}(\tilde{\mathfrak{N}}_{1:N} > T) + \varepsilon \\ &= \varepsilon^{-1} \limsup_N \mathbb{E}\mathbb{P}(\tilde{\mathfrak{N}}_{1:N} > T \mid \tilde{X}_{1,N}) + \varepsilon. \end{aligned} \tag{2.6}$$

Notice that, conditioning on $\tilde{X}_{1,N}$, the random variable $\tilde{\mathfrak{N}}_{1:N}$ is binomially distributed with the sample size N and success probability $p_N(\tilde{X}_{1,N}) := \mathbb{P}(\tilde{X}_{1,N} - A_N < X \leq \tilde{X}_{1,N} \mid \tilde{X}_{1,N})$.

- (II) (Truncation of independent maxima). Next, we prove that there exist constants $l_{T:N}$ and $L_{T:N}$ ($N \in \mathbb{N}, T \geq T_0$) such that $l_{T:N} \leq \tilde{X}_{1,N} \leq L_{T:N}$ with high probability $1 - o(1)$ as first $N \rightarrow \infty$ and then $T \rightarrow \infty$; see Lemma 5.2. Thus, for T large enough, the right-hand side of (2.6) does not exceed

$$\varepsilon^{-1} \limsup_N \mathbb{E}\left(\mathbb{P}(\tilde{\mathfrak{N}}_{1:N} > T \mid \tilde{X}_{1,N}) \mathbb{I}_{\{l_{T:N} \leq \tilde{X}_{1,N} \leq L_{T:N}\}}\right) + 2\varepsilon. \tag{2.7}$$

(III) (*Lebesgue's bounded convergence theorem and limit theorems for sums of independent random variables*). Consider Eq. (2.7). We first apply the Lebesgue's bounded convergence theorem (i.e. Lemma 5.1) to interchange the operations \limsup_N and \mathbb{E} in (2.7). The last limit is then reduced (by using relevant *limit theorems* for the sequence of binomially distributed random variables) to the condition of Theorem 2.1(i). I.e. as $T \rightarrow \infty$, expression (2.7) tends to the constant $2\varepsilon > 0$, picked arbitrarily small. The latter concludes the proof of Theorem 2.1(i). \square

The same arguments as those in Remark 2.5(I)–(III) are applied to prove the remaining asymptotic results of the present section and Sect. 3 dealing with the sums of negative powers of spacings; cf. Sects. 5 and 6 below. They are thought to be useful in the study of asymptotic behaviour of other functions on order statistics. In this paper, we do not prove the necessity of the conditions on limit theorems for functions on order statistics under consideration. However, the proof of the results of Sect. 2 and 3 suggests that our assumptions on the tail distribution $\bar{F} = e^{-Q}$ are very close to necessary conditions.

3 Applications to Sums of Negative Powers of Spacings

As before, let X, X_1, X_2, \dots be i.i.d. random variables with the common tail distribution function $\bar{F} = e^{-Q}$ and the right endpoint $x^* \leq \infty$. Let $X_{0,N} := x^* \geq X_{1,N} \geq \dots \geq X_{N,N}$ be the “extended” variational series based on the sample X_j ($1 \leq j \leq N$).

In the present section, we provide the weak law of large numbers (WLLN) for the sum $N^{-1}S_{K:N}^{(r)}$ as $N \rightarrow \infty$, where

$$S_{K:N}^{(r)} := \sum_{j=K+1}^N (X_{K,N} - X_{j,N})^{-r} \quad (3.1)$$

for fixed constants $K \in \mathbb{N}$ and $r \in \mathbb{R}_+$. Of course, sum (3.1) is allowed to take the value $+\infty$. (Here and in what follows, $1/0+ = +\infty$ and $1/\infty = 0$, by convention). For continuous \bar{F} , with probability one sum (3.1) is finite for all N , since the inequalities in (1.1) are strict. On the other side, if \bar{F} has an atom at $x^* < \infty$, then with probability one sum (3.1) is infinite for any N large enough, since a few extreme order statistics coincide for such N ; cf. Remark 2.4(j). Notice that, for $K = 0$, Eq. (3.1) is the sum of nonnegative i.i.d. random variables; therefore, $N^{-1}S_{0:N}^{(r)}$ converges in probability to the expectation $\mathbb{E}((x^* - X)^{-r}) \leq \infty$ due to the classical WLLN. We will prove the same WLLN for $N^{-1}S_{K:N}^{(r)}$ ($K \geq 1$) under the additional condition on regular variation of the function \bar{F} at the right endpoint x^* ; see Theorems 3.1 and 3.2. This condition is given in terms of singular integrals of the function \bar{F} and ensures that the sum of the first $o(N)$ terms in (3.1) is negligible in large N limit.

Limit theorems for the functions $N^{-1}S_{K:N}^{(r)}$ ($r = 1$ and $=2$) play a crucial role in the extreme value theory for eigenvalues $e_K(\mathbf{H}_N)$ of $N \times N$ symmetric random matrices $\mathbf{H}_N := \mathbf{X}_{N,\text{diag}} + \mathbf{J}_N$ with $N \rightarrow \infty$, known as *rank one perturbations of*

random diagonal matrices $\mathbf{X}_{N,\text{diag}} := \text{diag} \{X_1, X_2, \dots, X_N\}$, where $\{X_i\}_{i \in \mathbb{N}}$ is the i.i.d. sequence as above and \mathbf{J}_N are $N \times N$ deterministic matrices whose every entry is equal to N^{-1} . Under appropriate conditions on the tail distribution \bar{F} (in particular, continuity conditions of \bar{F}), Astrauskas and Molchanov [6] and Astrauskas [4] have established the following phase transition in the asymptotic behaviour of the extreme eigenvalues $e_K(\mathbf{H}_N)$:

(POI) if the function $N^{-1}S_{K:N}^{(1)}$ is asymptotically small enough (i.e. a low statistical concentration of observations X_i ($1 \leq i \leq N$) close to $X_{K,N}$), then the eigenvalues $e_K(\mathbf{H}_N)$ take on a Poissonian character in the limit $N \rightarrow \infty$;

(CLT) if the function $N^{-1}S_{K:N}^{(1)}$ is asymptotically large enough (i.e. a high statistical concentration of observations X_i ($1 \leq i \leq N$) close to $X_{K,N}$), then the central limit theorem for the maximal eigenvalue $e_1(\mathbf{H}_N)$ holds true.

On the other hand, by applying limit theorems for the two largest eigenvalues $e_K(\mathbf{H}_N)$ ($K = 1, 2$) and, as a by-product, asymptotic properties of the sums $N^{-1}S_{1:N}^{(1)}$ and spacings $X_{1,N} - X_{2,N}$, Bogachev and Molchanov [14], Fleischmann and Molchanov [23], Fleischmann and Greven [22] have studied long-time localization properties of the solutions $0 \leq U(t, \cdot) = \{U(t, k) : k = 1, 2, \dots, N\}$ ($t \in \mathbb{R}_+$) to the random “heat” equations $\partial U(t, \cdot) / \partial t = \mathbf{H}_N U(t, \cdot)$ with certain initial conditions, provided t and N both tend to infinity in a suitable relation. Recall that the solution $U(t, \cdot)$ is represented as a linear combination of N orthonormal eigenvectors of \mathbf{H}_N . It is shown that the main asymptotic contribution to $U(t, \cdot)$ comes from the term associated with the maximal eigenvalue and eigenvector, and the contribution from the other terms associated with lower eigenvalues is asymptotically negligible—therefore, the asymptotic properties of $U(t, \cdot)$ are determined by that of the maximal eigenvalue and eigenvector of matrices \mathbf{H}_N . The model is thought to exhibit the localization–delocalization transition at the boundary between the cases (POI) and (CLT) as above; cf. Section 1 in [4].

To state the main theorems of the section, we write $t \wedge s := \min(t, s)$ and, for $c \in \mathbb{R}_+$, abbreviate

$$J_r(x; c) := \int_{c(\bar{F}(x))^{1/r}}^{(x^*-x) \wedge 1} y^{-r-1} (\bar{F}(x-y) - \bar{F}(x)) \, dy \tag{3.2}$$

for any x smaller than but close to x^* .

Theorem 3.1 (Weak law of large numbers) *Let $x^* \leq \infty$, $r \in \mathbb{R}_+$ and $K \in \mathbb{N}$ be fixed constants. If $m^{(r)} := \mathbb{E}((x^* - X)^{-r}) < \infty$ and*

$$J_r(x; c) \rightarrow 0 \text{ as } x \uparrow x^*, \text{ for any } c \in \mathbb{R}_+, \tag{3.3}$$

then

$$N^{-1}S_{K:N}^{(r)} \xrightarrow{\mathbb{P}} m^{(r)} \text{ as } N \rightarrow \infty.$$

Recall that, for $x^* = \infty$, the expectation $m^{(r)}$ is zero, by convention. Theorem 3.1 is complemented by the following theorem.

Theorem 3.2 (Asymptotic negligibility of the maximal term in sums (3.1)) *Let $x^* \leq \infty$ and $K \in \mathbb{N}$ be fixed constants.*

(I) *If $r \in \mathbb{R}_+ \cup \{\infty\}$ and*

$$\overline{F}\left(x - c(\overline{F}(x))^{1/r}\right) / \overline{F}(x) \rightarrow 1 \text{ as } x \uparrow x^*, \text{ for any } c \in \mathbb{R}_+, \tag{3.4}$$

then

$$N^{1/r} (X_{K,N} - X_{K+1,N}) \xrightarrow{\mathbb{P}} \infty \text{ as } N \rightarrow \infty.$$

(II) *If $r \in \mathbb{R}_+$ and $\overline{F}(x)(x^* - x)^{-r} \rightarrow 0$ as $x \uparrow x^*$, then*

$$N^{1/r} (x^* - X_{1,N}) \xrightarrow{\mathbb{P}} \infty \text{ as } N \rightarrow \infty.$$

Proof of Theorem 3.1 is given in Sect. 6, and it heavily relies on Proposition 1.1 and Lemmas 5.1 and 5.2. Part (I) of Theorem 3.2 follows from Lemma 5.4 with $a(x) \equiv x^{-1/r}$ ($x \in \mathbb{R}_+$). For $x^* < \infty$, part (II) of Theorem 3.2 follows from the obvious bound $\mathbb{P}(N^{1/r}(x^* - X_{1,N}) < T) \leq N\overline{F}(x^* - TN^{-1/r}) \rightarrow 0$ as $N \rightarrow \infty$, for any $T \in \mathbb{R}_+$. □

Remark 3.3 Let us overview *the previous papers on limit theorems for sums* (3.1). For the Gaussian i.i.d. random variables X_k ($k \in \mathbb{N}$), Bogachev and Molchanov [14] derived the strong LLN for $N^{-1}S_{1:N}^{(1)}$ by exploring the explicit almost sure bounds for certain spacings in (3.1). Similarly, in the case of exponentially distributed i.i.d. random variables $X_k = \eta_k$ ($k \in \mathbb{N}$), Fleischmann and Molchanov [23] proved the WLLN for $N^{-1}S_{1:N}^{(1)}$, and later on, Fleischmann and Greven [22] extended this result, in particular, determining the convergence rate in the WLLN (cf. Lemma 2.2 in [22]). Astrauskas and Molchanov [6] established the convergence of $N^{-1}S_{1:N}^{(1)}$ in L^1 norm, provided the density function $p = (-\overline{F})'$ satisfies certain conditions of regular variation in terms of singular integrals of p . Astrauskas [4] gave the sketch proof of the WLLN for $N^{-1}S_{K:N}^{(r)}$ with $r \in \{1, 2\}$ and $K \in \mathbb{N}$, provided \overline{F} is a smooth function satisfying conditions like those in Theorems 3.1 and 3.2. In [4, 6], the authors derived further limit theorems for the (properly normalized) sums $S_{K:N}^{(1)}$ and related quantities in the case of $x^* < \infty$ and $\overline{F}_\alpha(x) = (x^* - x)^\alpha$ for x close to x^* ; here $\alpha > 0$. Notice that, for $\alpha > r$, the function \overline{F}_α satisfies the assumptions of Theorem 3.1; meanwhile, for $\alpha \leq r$, the conditions of Theorem 3.1 do not hold. The various WLLN for $N^{-1}S_{K:N}^{(r)}$ from [4, 6, 14, 22, 23] are extended and generalized in Theorem 3.1 of the present paper, where r is an arbitrary positive constant, $K \in \mathbb{N}$ and \overline{F} is allowed to be a discontinuous function. Recall that limit theorems for sums (3.1) and related functions on order statistics play a major role in studying asymptotic properties of the random matrix models mentioned above. □

Remark 3.4 (Comparison of the conditions of Theorems 3.1 and 3.2)

(I) For any $r \in \mathbb{R}_+$, a finiteness of the expectation

$$m^{(r)} = \mathbb{E}((x^* - X)^{-r}) = r \int_{-\infty}^{x^*} (x^* - x)^{-r-1} \bar{F}(x) dx < \infty$$

implies that $\bar{F}(x)(x^* - x)^{-r} \rightarrow 0$ as $x \uparrow x^*$ (cf. Section V.6 in [21]). For $x^* < \infty$, the latter limit yields that \bar{F} is continuous at x^* . Also, condition (3.4) of Theorem 3.2 implies a continuity of \bar{F} at $x^* < \infty$. On the other hand, if \bar{F} is discontinuous at $x^* < \infty$, then $m^{(r)} = \infty$ and almost surely $N^{-1}S_{K:N}^{(r)} = \infty$ for any N large enough [cf. Remark 2.4(j)].

(II) For any $r \in \mathbb{R}_+$, the conditions of Theorem 3.1 imply assumption (3.4) of Theorem 3.2.

(III) Let $x^* = \infty$, and assume that

$$\bar{F}(x - c) / \bar{F}(x) \rightarrow 1 \text{ as } x \rightarrow \infty, \text{ for any } c \in \mathbb{R}_+, \tag{3.5}$$

i.e. Assumption (3.4) with $r = \infty$ (or equivalently, condition (2.2) with $\rho = 0$ introduced by Pakes and Steutel [37]). Then \bar{F} satisfies the conditions of Theorem 3.1 for any $r \in \mathbb{R}_+$. Condition (3.5) is well-known in the classical theory of regularly varying functions; see Chapter 1 in [12]. In particular, (3.5) is equivalent to the slow variation of the function $\bar{F}(\log(\cdot))$ at infinity; therefore, $-\log(\bar{F}(x)) = o(x)$, i.e. heavy tails.

(IV) Assume now that

$$\mathbb{P}(X \in \mathbb{N}) = 1, \quad x^* = \infty \quad \text{and} \quad \limsup_{x \rightarrow \infty} \frac{\bar{F}(x - 1)}{\bar{F}(x)} > 1,$$

i.e. condition (2.5) of Remark 2.4(jj) holds true. Then, for $r \in \mathbb{R}_+$, we have that $m^{(r)} = 0$; however, condition (3.3) of Theorem 3.1 does not hold. (Indeed, there exists a sequence $x = x_n \in \mathbb{N}$ ($n = 1, 2, \dots$) tending to infinity such that the left-hand side of (3.4) converges to $\text{const} > 1$; therefore, (3.3) fails according to part (II) of the present remark). On the other hand, Remark 2.4(jj) with $m = 0$ implies that

$$\begin{aligned} \limsup_N \mathbb{P}(S_{1:N}^{(r)} = \infty) &= \limsup_N \mathbb{P}(\mathfrak{M}_{1:N} \geq 1) > 0 \quad \text{and} \\ \limsup_N \mathbb{P}(N^{-1}S_{1:N}^{(r)} \leq 1) &\geq \limsup_N \mathbb{P}(\mathfrak{M}_{1:N} = 0) > 0, \end{aligned}$$

so that the sequence $N^{-1}S_{1:N}^{(r)}$ does not converge to a constant (finite or infinite) with probability close to one.

(V) If limit (3.4) is fulfilled for all $c \in \mathbb{R}$ (this is a slightly stronger than (3.4) for $c \in \mathbb{R}_+$), then limit (3.4) holds true uniformly in any compact interval of real c due to the monotonicity of \bar{F} . In this case, the function $(\bar{F}(x))^{1/r}$ is called *self-neglecting* as x approaches x^* . The class of self-neglecting functions at infinity is characterized in Section 2.11 of [12].

Let us prove assertions (II)–(III) of Remark 3.4. As for (II), assuming a finiteness of $\mathbb{E}((x^* - X)^{-r})$ and taking into account the first assertion of (I), we obtain the implication (3.3) \Rightarrow (3.4) from the following simple estimates for any $c \in \mathbb{R}_+$ and any x close to x^* :

$$\begin{aligned} J_r(x; c) &\geq \int_{c(\overline{F}(x))^{1/r}}^{2c(\overline{F}(x))^{1/r}} y^{-r-1}(\overline{F}(x - y) - \overline{F}(x)) \, dy \\ &\geq \int_{c(\overline{F}(x))^{1/r}}^{2c(\overline{F}(x))^{1/r}} y^{-r-1} \, dy \left[\overline{F}\left(x - c(\overline{F}(x))^{1/r}\right) - \overline{F}(x) \right] \\ &= \text{const} \left[\overline{F}\left(x - c(\overline{F}(x))^{1/r}\right) (\overline{F}(x))^{-1} - 1 \right] \geq 0. \end{aligned}$$

In part (III), the implication (3.5) \Rightarrow (3.3) follows from the bounds

$$\begin{aligned} J_r(x; c) &\leq (\overline{F}(x - 1) - \overline{F}(x)) \int_{c(\overline{F}(x))^{1/r}}^1 y^{-r-1} \, dy \\ &\leq \text{const} (\overline{F}(x - 1) - \overline{F}(x)) / \overline{F}(x) \quad \text{for any } x \geq x_0. \end{aligned}$$

□

Remark 3.5 (Alternative conditions for the assertions of Theorems 3.1 and 3.2, provided $F := 1 - \overline{F}$ is smooth enough) Suppose that there exists a density $p := F'$ of the distribution function F . Fix $r \in \mathbb{N}$.

(J) Let $m^{(r)} < \infty$, and assume that p has the r th derivatives $p^{(r)}$ in (x_0, x^*) satisfying the following conditions as $x \uparrow x^*$:

$$\text{for } r = 1, \quad p(x) \log \left(\frac{\overline{F}(x)}{(x^* - x) \wedge 1} \right) \rightarrow 0 \quad \text{and} \quad p'(x) ((x^* - x) \wedge 1) \rightarrow 0; \tag{3.6}$$

$$\begin{aligned} \text{for } r = 2, 3, \dots, \quad &p^{(i)}(x) (\overline{F}(x))^{i+1-r} \rightarrow 0 \quad (0 \leq i \leq r - 2) \quad \text{and} \\ &p^{(r-1)}(x) \log \left(\frac{(\overline{F}(x))^{1/r}}{(x^* - x) \wedge 1} \right) \rightarrow 0, \quad \text{and} \quad p^{(r)}(x) ((x^* - x) \wedge 1) \rightarrow 0; \end{aligned} \tag{3.7}$$

with $p^{(0)} := p$. Then condition (3.3) is fulfilled.

(JJ) If $p(x) (\overline{F}(x))^{(1-r)/r} \rightarrow 0$ as $x \uparrow x^*$, then $(\overline{F})^{1/r}$ is self-neglecting at x^* , so that condition (3.4) is fulfilled.

Let us *prove assertions (J) and (JJ)*, provided $x^* = \infty$. For finite x^* , the proof is similar. In order to prove the convergence of integral (3.2) to zero under conditions (3.6)–(3.7), we apply Taylor's theorem to expand the differences $\bar{F}(x - y) - \bar{F}(x) = F(x) - F(x - y)$ in powers of y up to order $r + 1$. Namely, for all $x \geq x_0$ and $0 < y \leq 1$,

$$\bar{F}(x - y) - \bar{F}(x) = \sum_{i=1}^r \frac{(-1)^{i-1} p^{(i-1)}(x)}{i!} y^i + h_r(v) y^{r+1}, \quad (3.8)$$

where $|h_r(v)| = |p^{(r)}(v)|/(r + 1)!$ for some v such that $x - y \leq v \leq x$. By the last Assumptions in (3.6)–(3.7) with $x^* = \infty$, we have that $\sup_{x-1 \leq v \leq x} |h_r(v)| \rightarrow 0$ as $x \rightarrow \infty$. Taking into account the latter limit and Assumptions (3.6)–(3.7) with $x^* = \infty$, we now substitute (3.8) into (3.2) to find that the right-hand side of (3.2) tends to zero as $x \rightarrow \infty$.

To prove assertion (JJ) for $x^* = \infty$, we consider the density $p_r := r^{-1}(\bar{F})^{(1-r)/r} p$ of the function $1 - (\bar{F})^{1/r}$, so that $p_r(x) \rightarrow 0$ as $x \rightarrow \infty$ by the assumption. Fix $c \in \mathbb{R} \setminus \{0\}$, and write $f(x) := c(\bar{F}(x))^{1/r}$. According to the mean value theorem for the function $1 - (\bar{F})^{1/r}$, we have that for any $x \geq x_0$ there exists a real y such that $x - f(x) \leq y \leq x$ and

$$\left[(\bar{F}(x - f(x)))^{1/r} - (\bar{F}(x))^{1/r} \right] (\bar{F}(x))^{-1/r} = p_r(y) f(x) (\bar{F}(x))^{-1/r} = p_r(y) c.$$

Since $p_r(y) \leq \sup_{x-1 \leq z \leq x} p_r(z) \rightarrow 0$ as $x \rightarrow \infty$, we obtain the assertion (JJ) for $x^* = \infty$. \square

4 Proof of Propositions 1.1 and 1.2

Proof of Proposition 1.1 Denote by $\mathcal{I}^N := \{I\}$ the set of all rectangles $I \subseteq \mathbb{R}^N$ (open, closed, semi-open) with sides parallel to the axes of \mathbb{R}^N . Let μ_N be the probability measure on \mathbb{R}^N generated by the random vector $(\eta_{1,N}, \eta_{2,N}, \dots, \eta_{N,N})$. Similarly, let $\mu_{K:N}^{(1)}$ and $\mu_{K:N}^{(2)}$ be the probability measures on \mathbb{R}^N generated by the random vectors, respectively, $(\eta_{1,N}^{(1)}, \dots, \eta_{1,N}^{(K)}, \eta_{K+1,N}, \dots, \eta_{N,N})$ and $(\eta_{1,N}^{(1)}, \dots, \eta_{1,N}^{(K)}, \eta_{K+1,N}, \dots, \eta_{N,N})$ defined above Proposition 1.1. (Throughout we use the abbreviation $\int g \, d\mu := \int_{\mathbb{R}^N} g(\mathbf{x}) \, d\mu(\mathbf{x})$ for the integral with respect to the measure μ on \mathbb{R}^N ; thus, $\mu(D) := \int \mathbb{I}_D \, d\mu$ for a measurable $D \subseteq \mathbb{R}^N$).

The first lemma tells us that the function G_N is approximated in $L^1(\mu_N)$ and $L^1(\mu_{K:N}^{(i)})$ by a step function, i.e. a finite linear combination of indicator functions of rectangles in \mathbb{R}^N . \square

Lemma 4.1 Fix $K \in \mathbb{N}$ and measurable functions $G_N: \mathbb{R}^N \rightarrow [0, C]$. For any $N > K$ and any $\varepsilon > 0$, there exists a step function $S_N := \sum_{j=1}^L \alpha_j \mathbb{I}_{I_j}$ for some constants $L = L(N, \varepsilon) \in \mathbb{N}$, $\alpha_j = \alpha_j(N, \varepsilon) \in (0, C]$ and rectangles $I_j = I_j(N, \varepsilon) \in \mathcal{I}^N$ such that

$$\int |G_N - S_N| d\mu_N < \varepsilon \quad \text{and} \quad \int |G_N - S_N| d\mu_{K:N}^{(i)} < \varepsilon \quad \text{for } i = 1, 2. \tag{4.1}$$

Lemma 4.1 is proved below by using the standard arguments of the Lebesgue’s theory of measures and integration; see, e.g. Theorem 2.4 (Chapter 2) in [40] or Theorem 1.3.20 in [41] and their proof. However, we need to be careful with N which is arbitrarily large, and $\varepsilon > 0$ which is arbitrarily small. Moreover, we work with three different measures μ_N and $\mu_{K:N}^{(i)}$ ($i = 1, 2$) on \mathbb{R}^N . Thus, we give a sketch of the proof, with emphasis on the basic ideas of the Lebesgue’s theory.

For $K \in \mathbb{N}$, $R \in \mathbb{R}_+$ and $N \geq N_0(K, R)$, define the subset

$$\Omega_{K,R,N} := \left\{ (x_1, x_2, \dots, x_N) \in \mathbb{R}^N : \max_{1 \leq i \leq K+1} |x_i - \log N| \leq \log R \right\}. \tag{4.2}$$

The next lemma follows simply from the formulas for distributions of the normalized exponential order statistics $\eta_{K,N} - \log N$.

Lemma 4.2 *For any $K \in \mathbb{N}$, any $R \geq 1$ and any natural $N \geq K + 1$,*

$$\begin{aligned} \mu_N \left(\mathbb{R}^N \setminus \Omega_{K,R,N} \right) &\leq \mathbb{P}(\eta_{K+1,N} < \log(N/R)) + \mathbb{P}(\eta_{1,N} > \log(NR)) \\ &\leq \rho(R) := (K + 1)R^{-1} + 1 - e^{-2/R}. \end{aligned}$$

Proof of Lemma 4.2 For $N \geq K + 1$ and $R \in \mathbb{R}_+$, the Markov’s inequality implies

$$\mathbb{P}(\eta_{K+1,N} < \log(N/R)) \leq \frac{N}{R} \mathbb{E} \left(e^{-\eta_{K+1,N}} \right) = \frac{N}{R} \cdot \frac{K + 1}{N + 1} < \frac{K + 1}{R},$$

where the expression for the last expectation is derived by the same arguments as in the proof of Remark 1.3 above. Using the inequality $\log(1 - x) \geq -2x$ for all $0 \leq x \leq 1/2$, we obtain that

$$\mathbb{P}(\eta_{1,N} > \log(NR)) = 1 - \exp \left\{ N \log \left(1 - \frac{1}{NR} \right) \right\} \leq 1 - e^{-2/R}$$

provided $NR \geq 2$. Lemma 4.2 is proved. □

From Lemma 4.1 we see that the expectation of the function G_N can be replaced (up to a small error) by a linear combination of the functions (indicators) with separated variables. A real-valued function $g(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$, is said to be the function with separated variables if

$$g(\mathbf{x}) \equiv \prod_{k=1}^N f_k(x_k) \quad \text{for some } f_k : \mathbb{R} \rightarrow \mathbb{R} \quad (1 \leq k \leq N). \tag{4.3}$$

For such g instead of G_N , we are able to prove Proposition 1.1 by applying straightforwardly the Markov property of the random sequence $\eta_{N-l+1,N}$ ($l = 1, 2, \dots, N$). I.e. with (4.2) and the abbreviations given above Lemma 4.1, we have the following lemma.

Lemma 4.3 *Fix $K \in \mathbb{N}$. Assume that $g_N : \mathbb{R}^N \rightarrow \mathbb{R}_+$ are functions with separated variables, determined by products (4.3) of N measurable nonnegative bounded functions on \mathbb{R} . Then, for any $R \geq R_0(K)$ and any $N \geq N_0(R)$,*

$$\int g_N \Pi_{\Omega_{K,R,N}} d\mu_N \leq \tau_i(N, R) \int g_N \Pi_{\Omega_{K,R,N}} d\mu_{K:N}^{(i)} \quad \text{for } i = 1, 2;$$

here $\tau_1(N, R) := e^{2RK} R^K K!$ and $\tau_2(N, R) := N^K R^K K!$.

Lemma 4.3 is proved below.

We now derive the claimed assertions of Proposition 1.1 by applying Lemmas 4.1, 4.2 and 4.3. Using the abbreviations given above Lemmas 4.1 and 4.2, we have that, for any $R \geq R_0(K)$ and any $N \geq N_0(R)$,

$$\begin{aligned} \int G_N d\mu_N &= \int G_N \Pi_{\Omega_{K,R,M}} d\mu_N + \int G_N \Pi_{\mathbb{R}^N \setminus \Omega_{K,R,M}} d\mu_N \\ &\leq \int G_N \Pi_{\Omega_{K,R,M}} d\mu_N + C \cdot \varrho(R) \end{aligned} \tag{4.4}$$

by Lemma 4.2.

We first apply Lemma 4.1 [i.e. the first bound in (4.1)] to the last integral in (4.4); afterwards, we apply Lemma 4.3 to the integrals of the corresponding indicator functions; and finally, we again use Lemma 4.1 (i.e. the second bound in (4.1)) to estimate the integral of the corresponding step function with respect to the measures $\mu_{K:N}^{(i)}$ ($i = 1, 2$). More precisely, for any $R \geq R_0(K)$, any $N \geq N_0(R)$ and any $\varepsilon > 0$,

$$\begin{aligned} \int G_N \Pi_{\Omega_{K,R,M}} d\mu_N &\leq \sum_{j=1}^L \alpha_j \int \Pi_{I_j} \Pi_{\Omega_{K,R,N}} d\mu_N + \varepsilon \\ &\leq \tau_i(N, R) \sum_{j=1}^L \alpha_j \int \Pi_{I_j} \Pi_{\Omega_{K,R,N}} d\mu_{K:N}^{(i)} + \varepsilon \\ &\leq \tau_i(N, R) \int G_N d\mu_{K:N}^{(i)} + \varepsilon \cdot \tau_i(N, R) + \varepsilon \end{aligned}$$

for $i = 1, 2$. With $\varepsilon = C\rho(R)(\tau_i(N, R) + 1)^{-1}$, the last estimate and (4.4) imply

$$\int G_N d\mu_N \leq \tau_i(N, R) \int G_N d\mu_{K:N}^{(i)} + C \cdot 2\rho(R) \tag{4.5}$$

for R, N and i as above; here $\rho(R)$ and $\tau_i(N, R)$ are defined in Lemmas 4.2 and 4.3, respectively.

For $i = 1$, we use the following substitutions: $T := \tau_1(N, R) = \tau_1(R) := e^{2RK} R^K K!$; therefore, $R = \tau_1^{-1}(T)$ and $\varepsilon_T := 2\rho(\tau_1^{-1}(T))$. Notice that $\varepsilon_T \rightarrow 0$ as $T \rightarrow \infty$. Thus, substituting these abbreviations into (4.5), we obtain assertion (1.4) of Proposition 1.1.

Similarly, for $i = 2$, we use the following substitutions: $\tau_2(N, R) =: N^K T$ where $T := R^K K!$; therefore, $R = (T/K!)^{1/K}$ and $\tilde{\varepsilon}_T := 2\rho((T/K!)^{1/K})$. Notice that $\tilde{\varepsilon}_T \rightarrow 0$ as $T \rightarrow \infty$. Thus, substituting these abbreviations into (4.5), we obtain assertion (1.5) of Proposition 1.1.

It remains to prove Lemmas 4.1 and 4.3.

Proof of Lemma 4.1 Throughout the proof, fix $N \in \mathbb{N}$, $N > K$, and $\varepsilon > 0$ arbitrarily. It suffices to prove Lemma 4.1 for the “unified” measure on \mathbb{R}^N

$$\tilde{\mu}_N := \mu_N + \mu_{K:N}^{(1)} + \mu_{K:N}^{(2)} \tag{4.6}$$

instead of the measures μ_N and $\mu_{K:N}^{(i)}$ ($i = 1, 2$). Notice that the measure $\tilde{\mu}_N$ is absolutely continuous with respect to the Lebesgue measure m_N on \mathbb{R}^N , and the density of $\tilde{\mu}_N$ is bounded and continuous m_N -a.e. Moreover, $\tilde{\mu}_N(\mathbb{R}^N) = 3$. Because of such regularity properties of $\tilde{\mu}_N$, the theory of the Lebesgue measure and integration remains valid for $\tilde{\mu}_N$ instead of m_N ; see Chapters 1 and 2 of [40].

First, define the simple functions

$$S_N^0 := \sum_j (j\varepsilon/9) \mathbb{I}_{D_j} \tag{4.7}$$

where $D_j = D_j(N, \varepsilon) := \{\mathbf{x} \in \mathbb{R}^N : j\varepsilon/9 \leq G_N(\mathbf{x}) < (j + 1)\varepsilon/9\}$.

Obviously, $S_N^0 : \mathbb{R}^N \rightarrow [0, C]$ is a measurable function, with a finite number $J = J(N, \varepsilon, C)$ of summands, such that

$$\sup_{\mathbf{x}} |G_N(\mathbf{x}) - S_N^0(\mathbf{x})| < \varepsilon/9. \tag{4.8}$$

Because of a regularity of $\tilde{\mu}_N$, there are open subsets $D_j^* = D_j^*(N, \varepsilon) \subset \mathbb{R}^N$ such that $D_j^* \supset D_j$ and $\tilde{\mu}_N(D_j^* \setminus D_j) < \varepsilon 2^{-j} / (3C)$ for all $j \in [1, J]$; see Theorem 3.4(i) (Chapter 1) in [40]. Define the function S_N^* as in (4.7) with D_j replaced by D_j^* for each $j \in [1, J]$. Thus $S_N^* \geq S_N^0$ and, therefore,

$$\begin{aligned} \int |S_N^0 - S_N^*| d\tilde{\mu}_N &= \int (S_N^* - S_N^0) d\tilde{\mu}_N \\ &= \sum_{j=1}^J \frac{j\varepsilon}{9} \tilde{\mu}_N(D_j^* \setminus D_j) < \frac{\varepsilon}{3} \sum_{j \in \mathbb{N}} 2^{-j} = \frac{\varepsilon}{3}. \end{aligned} \tag{4.9}$$

On the other side, for fixed j , the (open) subset D_j^* can be written as a countable union of closed rectangles $I_{m,j} = I_{m,j}(N, \varepsilon) \in \mathcal{I}^N$ with disjoint interiors $I_{m,j}^{int}$, i.e.

$D_j^* = \bigcup_{m \in \mathbb{N}} I_{m,j}$ where $I_{m,j}^{int} \cap I_{n,j}^{int} = \emptyset$ for all $m \neq n$; see Theorem 1.4 (Chapter 1) in [40]. Thus, $D_j^* = \left(\bigcup_{m \in \mathbb{N}} I_{m,j}^{int}\right) \cup B$, where B is the countable union of the boundaries of the rectangles $I_{m,j}$ ($m \in \mathbb{N}$), so that $\tilde{\mu}_N(B) = 0$. Therefore, because of the finiteness of the series

$$\sum_{m \in \mathbb{N}} \tilde{\mu}_N(I_{m,j}) = \sum_{m \in \mathbb{N}} \tilde{\mu}_N(I_{m,j}^{int}) = \tilde{\mu}_N\left(\bigcup_{m \in \mathbb{N}} I_{m,j}^{int}\right) = \tilde{\mu}_N(D_j^*) < \infty,$$

there exists $M = M(N, \varepsilon, j) \in \mathbb{N}$ such that

$$\left| \tilde{\mu}_N(D_j^*) - \sum_{m=1}^M \tilde{\mu}_N(I_{m,j}) \right| \leq \frac{\varepsilon}{3C2^j}. \tag{4.10}$$

Write now

$$S_N := \sum_{j=1}^J \sum_{m=1}^M \frac{j\varepsilon}{9} \mathbb{I}_{I_{m,j}}.$$

Taking into account the definitions of functions S_N^* and S_N , we see that $S_N \leq S_N^*$ $\tilde{\mu}_N$ -a.e. Therefore, we obtain from (4.10) that

$$\begin{aligned} \int |S_N^* - S_N| d\tilde{\mu}_N &= \int S_N^* d\tilde{\mu}_N - \int S_N d\tilde{\mu}_N \\ &= \sum_{j=1}^J \frac{j\varepsilon}{9} \tilde{\mu}_N(D_j^*) - \sum_{j=1}^J \sum_{m=1}^M \frac{j\varepsilon}{9} \tilde{\mu}_N(I_{m,j}) \\ &= \sum_{j=1}^J \frac{j\varepsilon}{9} \left(\tilde{\mu}_N(D_j^*) - \sum_{m=1}^M \tilde{\mu}_N(I_{m,j}) \right) < \frac{\varepsilon}{3} \sum_{j \in \mathbb{N}} 2^{-j} = \frac{\varepsilon}{3}. \end{aligned} \tag{4.11}$$

In view of the definition of measure $\tilde{\mu}_N$ (4.6), from estimates (4.8), (4.9) and (4.11) we obtain the desired assertions of Lemma 4.1. □

Proof of Lemma 4.3 We will use the probabilistic arguments involving Lemma 4.2 and the Markov property of the exponential order statistics. Similarly as in (4.2), we now introduce the event:

$$\Omega_{K,R,N} := \{ \log(N/R) \leq \eta_{K+1,N} < \eta_{1,N} \leq \log(NR) \} \subset \Omega; \tag{4.12}$$

therefore,

$$\mathbb{I}_{\Omega_{K,R,N}} = \prod_{j=1}^{K+1} \Theta_{R,N}(\eta_{j,N}) \quad \text{with} \quad \Theta_{R,N}(\eta) := \mathbb{I}_{[\log(N/R), \log(NR)]}(\eta). \tag{4.13}$$

Throughout the proof, let $h : \mathbb{R} \rightarrow \mathbb{R}_+$ and $f : \mathbb{R}^{N-k} \rightarrow \mathbb{R}_+$ be measurable bounded functions where $1 \leq k \leq K$ and $K \in \mathbb{N}$. Fix integers $R \geq 2$ and $N \geq 2R + K$. With these abbreviations, it suffices to prove that

$$\begin{aligned} & \mathbb{E} \left(h(\eta_{k,N}) f(\eta_{k+1,N}, \dots, \eta_{N,N}) \prod_{j=k}^{K+1} \Theta_{R,N}(\eta_{j,N}) \right) \\ & \leq NRk \mathbb{E} \left(h(\eta_1) \Theta_{R,N}(\eta_1) \right) \mathbb{E} \left(f(\eta_{k+1,N}, \dots, \eta_{N,N}) \prod_{j=k+1}^{K+1} \Theta_{R,N}(\eta_{j,N}) \right) \end{aligned} \tag{4.14}$$

$$\leq e^{2R} Rk \mathbb{E} \left(h(\eta_{1,N}) \Theta_{R,N}(\eta_{1,N}) \right) \mathbb{E} \left(f(\eta_{k+1,N}, \dots, \eta_{N,N}) \prod_{j=k+1}^{K+1} \Theta_{R,N}(\eta_{j,N}) \right). \tag{4.15}$$

Indeed, taking into account the definition of the probability measures μ_N and $\mu_{K:N}^{(i)}$ ($i = 1, 2$) given above Lemma 4.1 and iterating bounds (4.14) and (4.15) over $k \in [1, K] \cap \mathbb{N}$, we obtain the claimed assertions of Lemma 4.3.

To prove (4.14) and (4.15), we need the following property of the exponential order statistics.

Lemma 4.4 *The random sequence $\eta_{N-l+1,N}$ ($l = 1, 2, \dots, N$) is a nonhomogeneous Markov chain with transition probabilities (by inverting “time” $N - l + 1 = k$):*

$$\begin{aligned} \mathbb{P}(\eta_{k,N} > x \mid \eta_{k+1,N}, \dots, \eta_{N,N}) &= \mathbb{P}(\eta_{k,N} > x \mid \eta_{k+1,N}) \\ &= \begin{cases} e^{-k(x-\eta_{k+1,N})} & \text{if } x > \eta_{k+1,N} \\ 1 & \text{if } x \leq \eta_{k+1,N} \end{cases} \end{aligned}$$

almost surely, for all $1 \leq k \leq N - 1$ and $N \geq 3$.

Remark 4.5 For the proof of the Markov property of order statistics of an i.i.d. sample with absolutely continuous distribution, see Theorem 2.4.3 and Section 3.4 in the monograph [3]. In Section 3.4 of [3], the authors also provide a counterexample of i.i.d. sample with discrete distribution, whose order statistics do not obey the Markov property.

Alternatively, Lemma 4.4 can be proved by using the fact that the normalized spacings $\eta_{1,N} - \eta_{2,N}, 2(\eta_{2,N} - \eta_{3,N}), \dots, N(\eta_{N,N} - \eta_{N+1,N})$ are i.i.d. $Exp(1)$ random variables; here $\eta_{N+1,N} := 0$. Cf. also Examples 4.1.5 and 4.1.6 in [20]. □

Now, using Lemma 4.4 and notation (4.12)–(4.13), we obtain that the left-hand side of (4.14) is equal to

$$\begin{aligned}
& \mathbb{E} \left[h(\eta_{k,N}) \Theta_{R,N}(\eta_{k,N}) f(\eta_{k+1,N}, \dots, \eta_{N,N}) \prod_{j=k+1}^{K+1} \Theta_{R,N}(\eta_{j,N}) \right] \\
&= \mathbb{E} \left[\mathbb{E}(h(\eta_{k,N}) \Theta_{R,N}(\eta_{k,N}) | \eta_{k+1,N}) \right. \\
&\quad \left. \times f(\eta_{k+1,N}, \dots, \eta_{N,N}) \prod_{j=k+1}^{K+1} \Theta_{R,N}(\eta_{j,N}) \right]; \tag{4.16}
\end{aligned}$$

here, assuming $\log(N/R) \leq \eta_{k+1,N} \leq \log(NR)$ (or, equivalently, $\Theta_{R,N}(\eta_{k+1,N}) \equiv 1$), we have that

$$\begin{aligned}
& \mathbb{E} \left(h(\eta_{k,N}) \Theta_{R,N}(\eta_{k,N}) | \eta_{k+1,N} \right) \\
&= k e^{k\eta_{k+1,N}} \int_{\eta_{k+1,N}}^{\infty} h(x) \Theta_{R,N}(x) e^{-kx} dx \\
&\leq NRk \int_0^{\infty} h(x) \Theta_{R,N}(x) e^{-x} dx = NRk \mathbb{E}(h(\eta_1) \Theta_{R,N}(\eta_1)) \tag{4.17}
\end{aligned}$$

with probability one. Equations (4.16) and (4.17) imply (4.14).

Further, we notice that the right-hand side of (4.17) is equal to

$$\begin{aligned}
& NRk \int_0^{\infty} h(x) \Theta_{R,N}(x) e^{-x} dx \\
&= Rk \int_0^{\infty} h(x) (1 - e^{-x})^{1-N} \Theta_{R,N}(x) \cdot N(1 - e^{-x})^{N-1} e^{-x} dx \\
&\leq e^{2R} Rk \int_0^{\infty} h(x) \Theta_{R,N}(x) \cdot N(1 - e^{-x})^{N-1} e^{-x} dx \\
&= e^{2R} Rk \mathbb{E}(h(\eta_{1,N}) \Theta_{R,N}(\eta_{1,N})),
\end{aligned}$$

where the last bound follows from the inequality $(1 - e^{-x})^{1-N} \leq e^{2R}$ provided $x \geq \log(N/R)$ and $N \geq 2R \geq 2$. This and (4.16)–(4.17) imply (4.15), as claimed.

Lemma 4.3 is proved, and this concludes the proof of Proposition 1.1. \square

Proof of Proposition 1.2 (I) We first prove that $q: \mathbb{R}_+^N \rightarrow \mathbb{R}^N$ is measurable. Recall that Q is a finite-valued right-continuous nondecreasing function on $(-\infty, x^*)$ such that $Q(x) \rightarrow 0$ as $x \rightarrow -\infty$, and $Q(x) = \infty$ for all $x \geq x^*$. Here and in what follows, we explore the following well-known properties of the left-continuous inverse Q^{\leftarrow} of the function Q defined in (1.2):

Lemma 4.6 (see Section 2 in [19])

- (I) Q^{\leftarrow} is a finite-valued nondecreasing function on \mathbb{R}_+ such that $Q^{\leftarrow}(y)$ tends to x^* as $y \rightarrow \infty$.

(II) For any $y \in \mathbb{R}_+$ and any $x < x^*$,

$$y \leq Q(x) \text{ if and only if } Q^{\leftarrow}(y) \leq x.$$

(III) For any $y \in \mathbb{R}_+$,

$$Q(Q^{\leftarrow}(y)-) \leq y \leq Q(Q^{\leftarrow}(y)).$$

Further on, from Section 4.1 (Chapter 1) of [40] we learn that $q: \mathbb{R}_+^N \rightarrow \mathbb{R}^N$ is measurable, if for any $(x_1, x_2, \dots, x_N) \in \mathbb{R}^N$, the subset

$$\left\{ (y_1, y_2, \dots, y_N) \in \mathbb{R}_+^N : Q^{\leftarrow}(y_1) \leq x_1, Q^{\leftarrow}(y_2) \leq x_2, \dots, Q^{\leftarrow}(y_N) \leq x_N \right\} \tag{4.18}$$

is measurable. Let us consider (4.18). We first notice that, for $x \geq x^*$, the set $\{y \in \mathbb{R}_+ : Q^{\leftarrow}(y) \leq x\}$ is equal to \mathbb{R}_+ . For $x < x^*$, using Lemma 4.6(II) we obtain that $\{y \in \mathbb{R}_+ : Q^{\leftarrow}(y) \leq x\}$ is either the empty set or the interval $(0, Q(x)]$. Summarizing, subset (4.18) is either the empty set or the rectangle $I \in \mathcal{I}^N, I \subseteq \mathbb{R}_+^N$. I.e. we obtain that $q: \mathbb{R}_+^N \rightarrow \mathbb{R}^N$ is a measurable finite-valued function. This and a continuity of $g: \mathbb{R}^N \rightarrow \mathbb{R}^M$ imply (by Property 2 in [40], pp. 29) that $g \circ q: \mathbb{R}_+^N \rightarrow \mathbb{R}^M$ is measurable, as claimed. Part (I) is proved.

Part (II) follows from the definition of measurable functions; cf. Section 4.1 (Chapter 1) in [40]. This concludes the proof of Proposition 1.2. □

5 Proof of Theorem 2.1

(i) By an appropriate linear transformation of constants $A_N = a(N)$, it suffices to prove assertion (i) for the counts $\mathfrak{N}_N^+ := \#\{K < j \leq N : X_{K,N} - A_N < X_{j,N} \leq X_{K,N}\}$ instead of $\mathfrak{N}_{K:N}(IA_N)$ (2.1). Notice that \mathfrak{N}_N^+ is the function on order statistics $X_{K,N}, \dots, X_{N,N}$ satisfying the measurability conditions of Proposition 1.1; cf. Proposition 1.2. We apply bound (1.4) in Proposition 1.1 to replace $X_{K,N}$ by the independent maxima $\tilde{X}_{1,N}$, i.e. $\tilde{X}_{1,N} := \max\{\tilde{X}_1, \dots, \tilde{X}_N\}$ where the i.i.d. sequence $\{\tilde{X}_i\}_{i \in \mathbb{N}}$ is an independent copy of $\{X_i\}_{i \in \mathbb{N}}$. Thus, abbreviating for $\lambda \in \mathbb{R}$,

$$Y_N(\lambda) := \#\{1 \leq j \leq N : \lambda - A_N < X_j \leq \lambda\}, \tag{5.1}$$

we obtain that, for any $M > 0$, any $\varepsilon > 0$ small enough and any $N \geq N_0(\varepsilon)$,

$$\begin{aligned} \mathbb{P}(\mathfrak{N}_N^+ > M) &\leq \varepsilon^{-1} \mathbb{P}(\#\{K < j \leq N : \tilde{X}_{1,N} - A_N < X_{j,N} \leq \tilde{X}_{1,N}\} > M) + \varepsilon \\ &\leq \varepsilon^{-1} \mathbb{P}(Y_N(\tilde{X}_{1,N}) > M) + \varepsilon. \end{aligned} \tag{5.2}$$

Passing in (5.2) to the limits first $N \rightarrow \infty$ and afterwards $M \rightarrow \infty$, and taking into account that $\varepsilon > 0$ is arbitrarily small, we find that the claimed limit $\mathfrak{N}_N^+ \stackrel{\mathbb{P}}{=} O(1)$

follows from $Y_N(\tilde{X}_{1,N}) \stackrel{\mathbb{P}}{=} O(1)$ as $N \rightarrow \infty$. To prove the latter limit we need the following auxiliary lemmas.

Lemma 5.1 (Bounded convergence theorem) *Assume that Z_N ($N \in \mathbb{N}$) are real random variables, and let $l_{T:N} \leq L_{T:N}$ ($N \in \mathbb{N}$, $T \in \mathbb{R}_+$) be real truncation constants such that*

$$\limsup_{N \rightarrow \infty} (\mathbb{P}(Z_N < l_{T:N}) + \mathbb{P}(Z_N > L_{T:N})) \rightarrow 0 \text{ as } T \rightarrow \infty.$$

Assume further that the functions $h_N : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the following conditions:

- (C1) $h_N(Z_N)$ are random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $|h_N(Z_N)| \leq \text{const}$ almost surely for some $\text{const} > 0$ and for any $N \in \mathbb{N}$;
- (C2) for any large $T \in \mathbb{R}_+$ and for any sequence $\{\lambda_N\}_{N \in \mathbb{N}} \subset \mathbb{R}$ such that $l_{T:N} \leq \lambda_N \leq L_{T:N}$ ($N \in \mathbb{N}$), one has $h_N(\lambda_N) \rightarrow 0$ as $N \rightarrow \infty$.

Then $\mathbb{E}|h_N(Z_N)| \rightarrow 0$ as $N \rightarrow \infty$.

Lemma 5.1 is complimented by the following assertion describing truncation constants for extremes of the i.i.d. sample $\{X_k\}_{1 \leq k \leq N}$.

Lemma 5.2 (Truncation of the K th largest order statistics) *Abbreviate $L_{T:N} := Q^{\leftarrow}(\log N + \log T)$ and $l_{T:N} := Q^{\leftarrow}(\log N - \log T)$. For fixed $K \in \mathbb{N}$, any $T \geq T_0(K)$ and any $N \geq N_0(T)$, we have that*

$$\mathbb{P}(X_{K,N} > L_{T:N}) \leq 1/T \text{ and } \mathbb{P}(X_{K,N} < l_{T:N}) \leq T^K e^{-T}.$$

Lemmas 5.1 and 5.2 are proved below.

We now represent the last probability in (5.2) as the integral with respect to the probability measure on \mathbb{R}^{N+1} induced by the random sample $\{X_1, X_2, \dots, X_N, \tilde{X}_{1,N}\}$ of mutually independent variables. Applying Fubini’s theorem to this integral, we observe that the quantities $Z_N := \tilde{X}_{1,N}$, $h_N(Z_N) := \mathbb{P}(Y_N(Z_N) > M \mid Z_N)$ satisfy the condition (C1) of Lemma 5.1. Therefore, by Lemma 5.1 with Z_N and $h_N(Z_N)$ as above and the constants $l_{T:N} \leq L_{T:N}$ as in Lemma 5.2, we obtain the claimed limit

$$\begin{aligned} & \limsup_N \mathbb{P}(Y_N(\tilde{X}_{1,N}) > M) \\ &= \limsup_N \mathbb{E} \mathbb{P}(Y_N(\tilde{X}_{1,N}) > M \mid \tilde{X}_{1,N}) \rightarrow 0 \text{ as } M \rightarrow \infty, \end{aligned}$$

provided the following limit holds true:

$$\limsup_N \mathbb{P}(Y_N(\lambda_N) > M) \rightarrow 0 \text{ as } M \rightarrow \infty, \tag{5.3}$$

for any nondecreasing sequence of constants λ_N such that $l_{T:N} \leq \lambda_N \leq L_{T:N}$ ($N \geq N_0(T)$) with $T > 1$ fixed arbitrarily. Since the random variable $Y_N(\lambda_N)$ is binomially

distributed with sample size N and success probability $p_N(\lambda_N) := \mathbb{P}(\lambda_N - A_N < X \leq \lambda_N)$, desired limit (5.3) is fulfilled if and only if

$$Np_N(\lambda_N) \equiv N e^{-Q(\lambda_N)} (\exp\{Q(\lambda_N) - Q(\lambda_N - A_N)\} - 1) = O(1) \quad (5.4)$$

as $N \rightarrow \infty$; cf. Theorem 2.1.1 and its proof in [31]. To show (5.4), we need the following technical lemma on some properties of the functions Q and a defined in Theorem 2.1. □

Lemma 5.3 (i) *If $Q(x) - Q(x-) = O(1)$ as $x \uparrow x^*$, then for the constants $l_{T:N} \leq \lambda_N \leq L_{T:N}$ defined in Lemma 5.2 with fixed $T > 1$ we have*

$$e^{Q(\lambda_N)} \asymp N \text{ as } N \rightarrow \infty.$$

(ii) *If $a(\cdot) \in ORV$, then $a(cx) \asymp a(x)$ ($x \rightarrow \infty$) uniformly in compact subsets of $c \in \mathbb{R}_+$.*

Proof Part (i) follows from the definition of constants $l_{T:N} \leq \lambda_N \leq L_{T:N}$ and the bounds in Lemma 4.6(III). Part (ii) follows from Theorem 1 in [1]. □

We now notice that the condition of Theorem 2.1(i) implies the assumption of Lemma 5.3(i). Thus, taking into account that $\lambda_N \uparrow x^*$ and using Lemma 5.3, we obtain that, as $N \rightarrow \infty$,

$$\begin{aligned} Np_N(\lambda_N) &\leq \text{const} \exp\{Q(\lambda_N) - Q(\lambda_N - a(N))\} \\ &\leq \text{const} \exp\left\{Q(\lambda_N) - Q\left(\lambda_N - \text{const}'a\left(e^{Q(\lambda_N)}\right)\right)\right\} = O(1) \end{aligned} \quad (5.5)$$

by the condition of Theorem 2.1(i). Thus, Eq. (5.4) is proved, as claimed.

It remains to prove Lemmas 5.1 and 5.2.

Proof of Lemma 5.1 Using the conditions of Lemma 5.1, we have that

$$\begin{aligned} \mathbb{E}|h_N(Z_N)| &\leq \mathbb{E}\left(|h_N(Z_N)| \mathbb{I}_{\{l_{T:N} \leq Z_N \leq L_{T:N}\}}\right) \\ &\quad + \text{const} \left(\mathbb{P}(Z_N < l_{T:N}) + \mathbb{P}(Z_N > L_{T:N})\right). \end{aligned}$$

Therefore, for any $\varepsilon > 0$, there exists $T_0 = T_0(\varepsilon)$ such that for any $T \geq T_0$

$$\limsup_N \mathbb{E}|h_N(Z_N)| \leq \limsup_N \mathbb{E}\left(|h_N(Z_N)| \mathbb{I}_{\{l_{T:N} \leq Z_N \leq L_{T:N}\}}\right) + \varepsilon. \quad (5.6)$$

Fix $T \geq T_0$. Since $|h_N(Z_N)| \leq \text{const}$ with probability 1, it follows from the bounded convergence theorem (see, e.g. Theorem 16.5 in [11]) that the last limit in (5.6) is equal to zero, if $h_N(Z_N) \mathbb{I}_{\{l_{T:N} \leq Z_N \leq L_{T:N}\}}$ converges almost surely to zero, as $N \rightarrow \infty$. The latter convergence holds true, if for any sequence $\{\lambda_N\}_{N \in \mathbb{N}} \subset \mathbb{R}$, one has

$$h_N(\lambda_N) \mathbb{I}_{\{l_{T:N} \leq \lambda_N \leq L_{T:N}\}} \rightarrow 0 \text{ as } N \rightarrow \infty.$$

This in turn follows from condition (C2) of Lemma 5.1, concluding the proof. \square

Proof of Lemma 5.2 Fix $K \in \mathbb{N}$. By the definition of $L_{T:N}$ and Q , we obtain that, for any $N \geq N_0(T)$,

$$\begin{aligned} \mathbb{P}(X_{K,N} > L_{T:N}) &\leq \mathbb{P}(X_{1,N} > L_{T:N}) \\ &\leq N e^{-Q(L_{T:N})} = N \exp\{-Q \circ Q^{\leftarrow}(\log(NT))\} \leq 1/T \end{aligned}$$

according to the right-hand side bound in Lemma 4.6(III) with $y = \log(NT)$.

For $T \geq T_0$ and $N \geq N_0(T)$, consider the probabilities

$$p_N := \mathbb{P}(X \geq l_{T:N}) = \exp\{-Q(l_{T:N}-)\} = \exp\{-Q(Q^{\leftarrow}(\log(N/T))-)\}.$$

Using the properties of the functions Q and Q^{\leftarrow} , in particular, the left-hand side inequality in Lemma 4.6(III) with $y = \log(N/T)$, we obtain that $Np_N \geq T$ and $p_N \rightarrow 0$ as $N \rightarrow \infty$. Therefore, for T and N large enough, the following simple estimates hold true:

$$\begin{aligned} \mathbb{P}(X_{K,N} < l_{T:N}) &= \sum_{j=0}^{K-1} \binom{N}{j} p_N^j (1-p_N)^{N-j} \leq \sum_{j=0}^{K-1} (Np_N)^j (1-p_N)^{N-K} \\ &\leq K(Np_N)^{K-1} (1-p_N)^{N-K} \\ &= K(Np_N)^{K-1} \exp\{(N-K)\log(1-p_N)\} \\ &\leq \text{const} (Np_N)^{K-1} e^{-Np_N} \leq T^K e^{-T}. \end{aligned}$$

Lemma 5.2 is proved, and this concludes the proof of Theorem 2.1(i). \square

(ii) This case is simply reduced to asymptotic properties of spacings associated with the extreme order statistics. Indeed, the claimed assertion of part (ii) is fulfilled, if $\mathfrak{N}_{K:N}((-tA_N, tA_N]) \xrightarrow{\mathbb{P}} 0$ as $N \rightarrow \infty$, for certain $t \in \mathbb{R}_+$. This in turn holds true, if $\min(X_{K-1,N} - X_{K,N}, X_{K,N} - X_{K+1,N})/A_N \xrightarrow{\mathbb{P}} \infty$ as $N \rightarrow \infty$; here $X_{0,N} := +\infty$. Thus, it remains to prove the following lemma that generalizes Theorem 3.2(I) in Sect. 3.

Lemma 5.4 *Under the conditions of Theorem 2.1(ii),*

$$(X_{K,N} - X_{K+1,N})/A_N \xrightarrow{\mathbb{P}} \infty \text{ as } N \rightarrow \infty, \text{ for fixed } K \in \mathbb{N}.$$

Proof Let $\tilde{X}_{1,N}$ be the independent maxima as above, and $F := 1 - e^{-Q}$. Using bound (1.4) in Proposition 1.1 with

$$G_N(X_{K,N}, X_{K+1,N}) = \mathbb{I}_{\{0 \leq X_{K,N} - X_{K+1,N} < MA_N\}}$$

for $M > 1$ and $K \in \mathbb{N}$, we obtain that, for any $\varepsilon > 0$ small enough, $T := \varepsilon^{-2}$ and for any $N \geq N_0(\varepsilon)$,

$$\begin{aligned} \mathbb{P}(X_{K,N} - X_{K+1,N} < MA_N) &\leq \varepsilon^{-1} \mathbb{P}(0 \leq \tilde{X}_{1,N} - X_{1,N} < MA_N) + \varepsilon \\ &= \varepsilon^{-1} \mathbb{P}(\tilde{X}_{1,N} - MA_N < X_{1,N} \leq \tilde{X}_{1,N}) + \varepsilon \\ &\leq \varepsilon^{-1} \mathbb{E} \left[\left(F^N(\tilde{X}_{1,N}) - F^N(\tilde{X}_{1,N} - MA_N) \right) \mathbb{I}_{\{l_{T:N} \leq \tilde{X}_{1,N} \leq L_{T:N}\}} \right] + 3\varepsilon, \end{aligned} \tag{5.7}$$

where the last bound in (5.7) follows from Lemma 5.2 with $T = \varepsilon^{-2}$. By Lemma 5.1, the last expectation in (5.7) tends to 0, provided

$$F^N(\lambda_N) - F^N(\lambda_N - MA_N) \rightarrow 0 \text{ as } N \rightarrow \infty \tag{5.8}$$

for any nondecreasing sequence $\lambda_N \uparrow x^*$ such that $l_{T:N} \leq \lambda_N \leq L_{T:N}$. To estimate the left-hand side of (5.8), we apply the inequality $c^N - b^N \leq N(c - b)$ for $0 \leq b \leq c \leq 1$ and then use the same arguments as in the proof of (5.5). Namely,

$$\begin{aligned} F^N(\lambda_N) - F^N(\lambda_N - MA_N) &\leq N(F(\lambda_N) - F(\lambda_N - MA_N)) \\ &\leq \text{const} \left[\exp \left\{ Q(\lambda_N) - Q(\lambda_N - \text{const}'a(e^{Q(\lambda_N)})) \right\} - 1 \right] \rightarrow 0, \end{aligned}$$

as $N \rightarrow \infty$, by the condition of Theorem 2.1(ii). This concludes the proof of Lemma 5.4, and therefore, part (ii) of Theorem 2.1 is shown.

(iii) The proof of part (iii) is similar to that of part (i). As in part (i), we need to prove that the sequence $\mathfrak{N}_N^- := \#\{K < j \leq N : X_{K,N} - A_N \leq X_{j,N} < X_{K,N}\}$ tends to infinity with probability $1 + o(1)$. As in Eq. (5.2), we obtain, that for any $M > 1$, any $\varepsilon > 0$ and any $N \geq N_0(M, \varepsilon)$

$$\begin{aligned} \mathbb{P}(\mathfrak{N}_N^- \leq M) &\leq \varepsilon^{-1} \mathbb{P}(\#\{K < j \leq N : \tilde{X}_{1,N} - A_N \leq X_{j,N} < \tilde{X}_{1,N}\} \leq M) + \varepsilon \\ &\leq \varepsilon^{-1} \mathbb{P}(Y_N^-(\tilde{X}_{1,N}) \leq M + K) + \varepsilon, \end{aligned}$$

where $Y_N^-(\lambda) := \#\{1 \leq j \leq N : \lambda - A_N \leq X_j < \lambda\}$ for $\lambda \in \mathbb{R}$. Since here $\varepsilon > 0$ is arbitrarily small, it suffices to prove that $Y_N^-(\tilde{X}_{1,N}) \xrightarrow{\mathbb{P}} \infty$ as $N \rightarrow \infty$. Similarly as in part (i), the last limit follows from the limit $Y_N^-(\lambda_N) \xrightarrow{\mathbb{P}} \infty$ for any sequence of constants $\lambda_N \uparrow x^*$ such that $l_{T:N} \leq \lambda_N \leq L_{T:N}$ ($N \geq N_0(T)$), where $T > 0$ is fixed arbitrarily. With $A_N \equiv a(N)$ and the abbreviation $p_N^-(\lambda_N) := \mathbb{P}(\lambda_N - a(N) \leq X < \lambda_N)$, this limit holds true if and only if

$$Np_N^-(\lambda_N) \equiv N e^{-Q^-(\lambda_N)} \left(e^{Q^-(\lambda_N) - Q^-(\lambda_N - a(N))} - 1 \right) \rightarrow \infty \tag{5.9}$$

as $N \rightarrow \infty$; cf. Theorem 2.1.1 and its proof in [31]. To prove the limit in (5.9), we first find from the left-hand side inequality in Lemma 4.6(III) that $N e^{-Q^-(\lambda_N)} \geq 1/T$. Then using the conditions of Theorem 2.1(iii) on the functions $Q(\cdot)$ and $a(\cdot)$ similarly as in part (i), we obtain that $a(N) \geq \text{const} a(e^{Q^-(\lambda_N)})$. Applying these bounds to the expression on the left-hand side of (5.9) and using condition (2.4), we obtain that $Np_N^-(\lambda_N)$ tends to infinity, as claimed.

Theorem 2.1 is proved. □

6 Proof of Theorem 3.1

Since $\overline{F}(x)(x^* - x)^{-r} \rightarrow 0$ as $x \uparrow x^*$ (see Remark 3.4(I)), we notice that $X < x^*$ almost surely. Abbreviate

$$H(x; \lambda) := (\lambda - x)^{-r} - (x^* - x)^{-r} \quad \text{for } x \leq \lambda < x^*. \quad (6.1)$$

Using the weak law of large numbers (WLLN) for the i.i.d. random variables $(x^* - X_j)^{-r}$ ($1 \leq j \leq N$) and taking into account that, for fixed $k \in \mathbb{N}$, the extreme terms $(x^* - X_{k,N})^{-r}$ are of order $o(N)$ in probability (cf. Theorem 3.2(II)), we find that the assertion of Theorem 3.1 follows from the limit

$$N^{-1} \sum_{j=K+1}^N H(X_{j,N}; X_{K,N}) \xrightarrow{\mathbb{P}} 0 \quad \text{as } N \rightarrow \infty. \quad (6.2)$$

Let us prove (6.2). We first notice from Theorem 3.2(I) that, for any $\delta > 0$ and $N \rightarrow \infty$,

$$\begin{aligned} & \mathbb{P} \left(N^{-1} \sum_{j=K+1}^N H(X_{j,N}; X_{K,N}) > \delta \right) \\ &= \mathbb{P} \left(N^{-1} \sum_{j=K+1}^N H(X_{j,N}; X_{K,N}) > \delta, N(X_{K,N} - X_{K+1,N})^r \geq 1 \right) + o(1) \\ &= \mathbb{P} \left(N^{-1} \sum_{j=K+1}^N H(X_{j,N}; X_{K,N}) \mathbb{I}_{\{N(X_{K,N} - X_{K+1,N})^r \geq 1\}} > \delta \right) + o(1) \\ &\leq \mathbb{P} \left(N^{-1} \sum_{j=K+1}^N H(X_{j,N}; X_{K,N}) \mathbb{I}_{\{N(X_{K,N} - X_{j,N})^r \geq 1\}} > \delta \right) + o(1). \end{aligned} \quad (6.3)$$

By using Propositions 1.1 and 1.2, we obtain that, for any $\delta > 0$, any $\varepsilon > 0$ small enough and any $N \geq N_0(\delta, \varepsilon)$, the right-hand side of (6.3) does not exceed

$$\begin{aligned} & \varepsilon^{-1} \mathbb{P} \left(N^{-1} \sum_{j=K+1}^N H(X_{j,N}; \tilde{X}_{1,N}) \mathbb{I}_{\{N(\tilde{X}_{1,N} - X_{j,N})^r \geq 1\}} > \delta \right) + \varepsilon \\ & \leq \varepsilon^{-1} \mathbb{P} \left(N^{-1} \sum_{j=1}^N H(X_j; \tilde{X}_{1,N}) \mathbb{I}_{\{N(\tilde{X}_{1,N} - X_j)^r \geq 1\}} > \delta \right) + \varepsilon, \end{aligned} \quad (6.4)$$

where $\tilde{X}_{1,N}$ is the independent maxima as in the proof of Theorem 2.1(i) in Sect. 5. Applying Lemmas 5.1 and 5.2 to the last probability in (6.4) similarly as in the proof

of Theorem 2.1(i) in Sect. 5, we obtain that the right-hand side of (6.4) tends to ε as $N \rightarrow \infty$, provided the following limit holds true:

$$\lim_N \mathbb{P} \left(N^{-1} \sum_{j=1}^N H(X_j; \lambda_N) \mathbb{I}_{\{N(\lambda_N - X_j)^r \geq 1\}} > \delta \right) = 0 \quad \text{for any } \delta > 0, \quad (6.5)$$

for any sequence of constants $\lambda_N \uparrow x^*$ such that $l_{T:N} \leq \lambda_N \leq L_{T:N}$ for any $N \geq N_0(T)$. Recall that $l_{T:N} \leq L_{T:N}$ are the truncation constants defined in Lemma 5.2 with $T > 1$ fixed arbitrarily. Summarizing the above considerations with $\varepsilon > 0$ arbitrarily small, we see that (6.5) yields desired limit (6.2). To prove (6.5), we first notice that the quantities

$$Y_j^{(N)} := H(X_j; \lambda_N) \mathbb{I}_{\{N(\lambda_N - X_j)^r \geq 1\}} \quad \text{for } 1 \leq j \leq N$$

are nonnegative bounded i.i.d. random variables. Therefore, limit (6.5), i.e. $N^{-1} \sum_{j=1}^N Y_j^{(N)} \xrightarrow{\mathbb{P}} 0$, is fulfilled if and only if $\mathbb{E}Y_1^{(N)} \rightarrow 0$ as $N \rightarrow \infty$; see, e.g. Theorem 1 (Section IX.9) in [21] or Theorem 2.2.11 in [17]. Thus, abbreviating $F := 1 - \bar{F}$, we need to prove that

$$\mathbb{E}Y_1^{(N)} = \int H(x; \lambda_N) \mathbb{I}_{(-\infty, \lambda_N - N^{-1/r}]}(x) dF(x) \rightarrow 0 \quad \text{as } N \rightarrow \infty \quad (6.6)$$

for λ_N as above. Here recall that the function H is defined in (6.1). Also, $N^{-1/r} \asymp (\bar{F}(\lambda_N))^{1/r}$ as $N \rightarrow \infty$ by Lemma 5.3(i) with $Q = -\log \bar{F}$. This implies the inclusion $(-\infty, \lambda_N - N^{-1/r}] \subset (-\infty, \lambda_N - c(\bar{F}(\lambda_N))^{1/r}]$ for some $c > 0$ and any $N \geq N_0$, which in turn is applied to the left-hand side of (6.6) to find that desired limit (6.6) is a consequence of the following condition:

$$\lim_{\lambda \uparrow x^*} \int H(x; \lambda) \mathbb{I}_{(-\infty, \lambda - c(\bar{F}(\lambda))^{1/r}]}(x) dF(x) = 0 \quad \text{for any } c \in \mathbb{R}_+. \quad (6.7)$$

Proof of Theorem 3.1 is concluded by observing that limit (6.7) is a straightforward consequence of the conditions of Theorem 3.1 (cf. also Remark 3.4(I)–(II)) and the following lemma.

Lemma 6.1 *Fix $x^* \leq \infty$ and $r \in \mathbb{R}_+$, and assume that $\bar{F}(x)(x^* - x)^{-r} \rightarrow 0$ as $x \uparrow x^*$. Then, for any $c \in \mathbb{R}_+$,*

$$\begin{aligned} & \int H(x; \lambda) \mathbb{I}_{(-\infty, \lambda - c(\bar{F}(\lambda))^{1/r}]}(x) dF(x) \\ &= r J_r(\lambda; c) + c^{-r} \left(1 - \bar{F}(\lambda - c(\bar{F}(\lambda))^{1/r}) (\bar{F}(\lambda))^{-1} \right) + o(1) \quad \text{as } \lambda \uparrow x^*. \end{aligned} \quad (6.8)$$

It remains to prove Lemma 6.1. For fixed $c > 0$, we write $f(\lambda) := c(\overline{F}(\lambda))^{1/r}$ for the transformed tail distribution, so that

$$f(\lambda) > 0 \text{ for } \lambda < x^*, \text{ and } f(\lambda)/(x^* - \lambda) \rightarrow 0 \text{ as } \lambda \uparrow x^*.$$

Notice that, for fixed $\lambda < x^*$, the function $H(\cdot; \lambda) : (-\infty, \lambda - f(\lambda)] \rightarrow \mathbb{R}_+$ is continuously differentiable nonnegative bounded function with the nonnegative bounded derivative

$$H_\lambda^{(1)}(x) := \frac{dH(x; \lambda)}{dx} = r(\lambda - x)^{-r-1} - r(x^* - x)^{-r-1}. \tag{6.9}$$

For such H and F , we apply the integration by parts formula (e.g. Section V.6 in [21]) and then use the condition of Lemma 6.1, to obtain that

$$\begin{aligned} & \int H(\cdot; \lambda) \Pi_{(-\infty, \lambda - f(\lambda)]} dF \\ &= -H(\lambda - f(\lambda); \lambda) \overline{F}(\lambda - f(\lambda)) + \int_{-\infty}^{\lambda - f(\lambda)} H_\lambda^{(1)} \overline{F} \\ &= -\frac{\overline{F}(\lambda - f(\lambda))}{c^r \overline{F}(\lambda)} + \frac{\overline{F}(\lambda - f(\lambda))}{(x^* - \lambda + f(\lambda))^r} + \int_{-\infty}^{\lambda - f(\lambda)} H_\lambda^{(1)} \overline{F} \\ &= -\frac{\overline{F}(\lambda - f(\lambda))}{c^r \overline{F}(\lambda)} + \int_{-\infty}^{\lambda - f(\lambda)} H_\lambda^{(1)} \overline{F} + o(1) \\ &= -\frac{\overline{F}(\lambda - f(\lambda))}{c^r \overline{F}(\lambda)} + \int_{-\infty}^{\lambda - \tau} H_\lambda^{(1)} \overline{F} + \int_{\lambda - \tau}^{\lambda - f(\lambda)} H_\lambda^{(1)} \overline{F} + o(1) \end{aligned} \tag{6.10}$$

as $\lambda \uparrow x^*$, for $\tau \in \mathbb{R}_+$ specified below.

Let us consider the last two integrals in (6.10). Assume first that $x^* < \infty$. By using the inequality

$$s^{-r-1} - t^{-r-1} \leq (r + 1)(t - s)t^{-1}s^{-r-1} \text{ for all } 0 < s \leq t \text{ and } r > 0,$$

we obtain the following bound for $H_\lambda^{(1)}$ (6.9):

$$H_\lambda^{(1)}(x) \leq r(r + 1)(x^* - \lambda)(x^* - x)^{-1}(\lambda - x)^{-r-1} \text{ for } x < \lambda < x^*. \tag{6.11}$$

It follows from (6.11) that, for any $\tau > 0$,

$$\begin{aligned} \int_{-\infty}^{\lambda - \tau} H_\lambda^{(1)} \overline{F} &\leq \text{const } (x^* - \lambda) \int_{-\infty}^{\lambda - \tau} (\lambda - x)^{-r-2} dx \\ &= \text{const } \tau^{-r-1} (x^* - \lambda) \rightarrow 0 \text{ as } \lambda \uparrow x^*. \end{aligned} \tag{6.12}$$

We now split the last integral in (6.10) into two integrals over the intervals $(\lambda - \tau, 2\lambda - x^*]$ and $(2\lambda - x^*, \lambda - f(\lambda)]$, respectively. Let us estimate the first integral

$\int_{\lambda-\tau}^{2\lambda-x^*}$. For any $\varepsilon > 0$, we have $\overline{F}(x) \leq \varepsilon(x^* - x)^r$ ($\lambda - \tau < x < x^*$), provided $\lambda \uparrow x^*$ and $\tau > 0$ is small enough. For such λ and τ , we apply this bound together with (6.11), and afterwards we use a change of integration variables $y = (x^* - \lambda)(\lambda - x)^{-1} + 1$, to obtain that

$$\begin{aligned} \int_{\lambda-\tau}^{2\lambda-x^*} H_\lambda^{(1)} \overline{F} &\leq \text{const } \varepsilon(x^* - \lambda) \int_{\lambda-\tau}^{2\lambda-x^*} (\lambda - x)^{-r-1} (x^* - x)^{r-1} dx \\ &= \text{const}' \varepsilon \int_{(x^*-\lambda+\tau)/\tau}^2 y^{r-1} dy \leq \text{const}'' \varepsilon. \end{aligned} \tag{6.13}$$

In view of (6.10), (6.12) and (6.13), it remains to estimate the last integral $\int_{2\lambda-x^*}^{\lambda-f(\lambda)}$. Obviously, as $\lambda \uparrow x^*$,

$$\begin{aligned} \int_{2\lambda-x^*}^{\lambda-f(\lambda)} H_\lambda^{(1)}(x) \overline{F}(x) dx &= \int_{2\lambda-x^*}^{\lambda-f(\lambda)} (\lambda - x)^{-r-1} (\overline{F}(x) - \overline{F}(\lambda)) dx \\ &+ r \overline{F}(\lambda) \int_{2\lambda-x^*}^{\lambda-f(\lambda)} (\lambda - x)^{-r-1} dx - r \int_{2\lambda-x^*}^{\lambda-f(\lambda)} (\lambda - x)^{-r-1} \overline{F}(x) dx \end{aligned} \tag{6.14}$$

$$= r J_r(\lambda; c) + c^{-r} - r \int_{2\lambda-x^*}^{\lambda-f(\lambda)} (\lambda - x)^{-r-1} \overline{F}(x) dx + o(1) \tag{6.15}$$

by changing variables $y = \lambda - x$ in integral (6.14) [cf. also (3.2)] and using the condition of Lemma 6.1. Similarly, for any $\varepsilon > 0$, we obtain the bound $\overline{F}(x) \leq \varepsilon(x^* - x)^r$ for $2\lambda - x^* < x < x^*$, provided λ approaches x^* . Consequently, the last integral in (6.15) does not exceed $\varepsilon \int_{2\lambda-x^*}^{\lambda} (x^* - x)^{-1} dx = \varepsilon \log 2$. Taking into account (6.10) with $\lambda \uparrow x^*$ and summarizing the last bound and bounds (6.12)–(6.15) with $\tau > 0$ and $\varepsilon > 0$ chosen arbitrarily small, we finally obtain the assertion of Lemma 6.1 for $x^* < \infty$.

Assume now that $x^* = \infty$. Consider the last two integrals on the right-hand side of (6.10) with $\tau > 1$. The first integral does not exceed $r \int_{-\infty}^{\lambda-\tau} (\lambda - x)^{-r-1} dx = \tau^{-r}$. Meanwhile, the second integral on the right-hand side of (6.10) (by changing variables $y = \lambda - x$ and letting $\lambda \rightarrow \infty$) is equal to

$$r \int_{f(\lambda)}^1 y^{-r-1} (\overline{F}(\lambda - y) - \overline{F}(\lambda)) dy + c^{-r} + o(1).$$

With τ chosen arbitrarily large, these bounds imply that the right-hand side of (6.10) is equal to the right-hand side of (6.8), as claimed.

Lemma 6.1 is proved, and this concludes the proof of Theorem 3.1. □

Acknowledgements The author would like to thank the referee for many suggestions that helped to improve the presentation of the paper.

Data Availability Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The author declares that he has no conflict of interest.

References

1. Aljančić, S., Arandelović, D.: O-regularly varying functions. *Publ. Inst. Math. (Beograd)(N.S.)* **22**(36), 5–22 (1977)
2. Anderson, C.W.: Extreme value theory for a class of discrete distributions with applications to some stochastic processes. *J. Appl. Probab.* **7**(1), 99–113 (1970)
3. Arnold, B.C., Balakrishnan, N., Nagaraja, H.N.: *A First Course in Order Statistics. Classics in Applied Mathematics*, vol. 54. SIAM, Philadelphia (2008)
4. Astrauskas, A.: Limit theorems for the maximal eigenvalues of the mean-field Hamiltonian with random potential. *Lith. Math. J.* **39**(2), 117–133 (1999)
5. Astrauskas, A.: From extreme values of i.i.d. random fields to extreme eigenvalues of finite-volume Anderson Hamiltonian. *Probab. Surv.* **13**, 156–244 (2016)
6. Astrauskas, A., Molchanov, S.A.: The ground state of a random stationary medium in the mean field approximation. In: Sazonov, V.V., Shervashidze, T.L. (eds.) *New Trends in Probability and Statistics*, vol. 1, pp. 668–682. VSP/Mokslas, Utrecht, Vilnius (1991)
7. Athreya, J.S., Sethuraman, S.: On the asymptotics of discrete order statistics. *Stat. Probab. Lett.* **54**(3), 243–249 (2001)
8. Balakrishnan, N., Stepanov, A.: A note on the number of observations near an order statistic. *J. Stat. Plan. Inference* **134**(1), 1–14 (2005)
9. Baryshnikov, Y., Eisenberg, B., Stengle, G.: A necessary and sufficient condition for the existence of the limiting probability of a tie for first place. *Stat. Probab. Lett.* **23**(3), 203–209 (1995)
10. Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.L.: *Statistics of Extremes: Theory and Applications*. John Wiley and Sons, Chichester (2004)
11. Billingsley, P.: *Probability and Measure*, 3rd edn. John Wiley and Sons, New York (1995)
12. Bingham, N.H., Goldie, C.M., Teugels, J.L.: *Regular Variation*. Cambridge University Press, Cambridge (1987)
13. Bogachev, L.V., Molchanov, S.A.: Mean-field models in the theory of random media. I. *Theor. Math. Phys.* **81**(2), 1207–1214 (1989)
14. Bogachev, L.V., Molchanov, S.A.: Mean-field models in the theory of random media. II. *Theor. Math. Phys.* **82**(1), 99–107 (1990)
15. Bruss, F.T., Grübel, R.: On the multiplicity of the maximum in a discrete random sample. *Ann. Appl. Probab.* **13**(4), 1252–1263 (2003)
16. Dembińska, A.: Asymptotic normality of numbers of observations in random regions determined by order statistics. *Statistics* **48**(3), 508–523 (2014)
17. Durrett, R.: *Probability: Theory and Examples*, 5th edn. Cambridge University Press, Cambridge (2019)
18. de Haan, L., Ferreira, A.: *Extreme Value Theory: An Introduction*. Springer, New York (2006)
19. Embrechts, P., Hofert, M.: A note on generalized inverses. *Math. Methods Oper. Res.* **77**(3), 423–432 (2013)
20. Embrechts, P., Klüppelberg, C., Mikosch, T.: *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin (1997)
21. Feller, W.: *An Introduction to Probability Theory and Its Applications*, vol. II, 2nd edn. John Wiley and Sons, New York (1971)
22. Fleischmann, K., Greven, A.: Localization and selection in a mean field branching random walk in a random environment. *Ann. Probab.* **20**(4), 2141–2163 (1992)
23. Fleischmann, K., Molchanov, S.A.: Exact asymptotics in a mean field model with random potential. *Probab. Theory Relat. Fields* **86**(2), 239–251 (1990)
24. Fortin J.-Y., Clusel, M.: Applications of extreme value statistics in physics. *J. Phys. A: Math. Theor.* **48**(18), 183001 (35pp) (2015)
25. Hashorva, E.: On the number of near-maximum insurance claim under dependence. *Insurance Math. Econom.* **32**(1), 37–49 (2003)

26. Hashorva, E.: Bivariate maximum insurance claim and related point processes. *Stat. Probab. Lett.* **69**(2), 117–128 (2004)
27. Hashorva, E., Hüslér, J.: Estimation of tails and related quantities using the number of near-extremes. *Commun. Stat. Theory Methods* **34**(2), 337–349 (2005)
28. Hu, Z., Su, C.: Limit theorems for the number and sum of near-maxima for medium tails. *Stat. Probab. Lett.* **63**(3), 229–237 (2003)
29. König, W.: *The Parabolic Anderson Model: Random Walk in Random Potential.* (Pathways in Mathematics). Birkhäuser/Springer, Basel (2016)
30. Ladoucette, S.A., Teugels, J.L.: Asymptotics for ratios with applications to reinsurance. *Methodol. Comput. Appl. Probab.* **9**(2), 225–242 (2007)
31. Leadbetter, M.R., Lindgren, G., Rootzén, H.: *Extremes and Related Properties of Random Sequences and Processes.* Springer, New York (1983)
32. Li, Y., Pakes, A.G.: On the number of near-maximum insurance claims. *Insurance Math. Econom.* **28**(3), 309–323 (2001)
33. Müller, S.: Tail estimation based on numbers of near m -extremes. *Methodol. Comput. Appl. Probab.* **5**(2), 197–210 (2003)
34. Nagaraja, H.N., Bharath, K., Zhang, F.: Spacings around an order statistic. *Ann. Inst. Stat. Math.* **67**(3), 515–540 (2015)
35. Pakes, A.G.: The number and sum of near-maxima for thin-tailed populations. *Adv. Appl. Probab.* **32**(4), 1100–1116 (2000)
36. Pakes, A.G., Li, Y.: Limit laws for the number of near maxima via the Poisson approximation. *Stat. Probab. Lett.* **40**(4), 395–401 (1998)
37. Pakes, A.G., Steutel, F.W.: On the number of records near the maximum. *Austral. J. Statist.* **39**(2), 179–192 (1997)
38. Sabhapandit, S., Majumdar, S.N.: Density of near-extreme events. *Phys. Rev. Lett.* **98**(14), 140201 (2007)
39. Sethuraman, S.: A clustering law for some discrete order statistics. *J. Appl. Probab.* **40**(1), 226–241 (2003)
40. Stein, E.M., Shakarchi, R.: *Real Analysis: Measure Theory, Integration, and Hilbert Spaces.* Princeton Lectures in Analysis, vol. 3. Princeton University Press, Princeton (2005)
41. Tao, T.: *An Introduction to Measure Theory.* Graduate Studies in Mathematics, vol. 126. Amer. Math. Soc, Providence (2011)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.