1. INTRODUCTION

Notes: *Need to start with practical motivation, and perhaps even put in some figures illustrating implications for real data.*

Then, later in the introduction, we might include discussion like the following; but I think it needs a lot of revision:

Let $\Theta_1, \ldots, \Theta_p$ denote independent and identically distributed random variables, and write

$$\Theta_{(1)} \leq \ldots \leq \Theta_{(p)} \tag{1.1}$$

for their ordered values. There exists a permutation $R = (R_1, \ldots, R_p)$ of $(1, \ldots, p)$ such that $\Theta_{(j)} = \Theta_{R_j}$. If the common distribution of the Θ_js is continuous then the inequalities in (1.1) are all strict and the permutation is unique.

In practice we typically do not observe the Θ_js directly, only in terms of noisy approximations which can be modelled as follows. Let $Q_i = (Q_{i1}, \ldots, Q_{ip})$ denote independent and identically distributed random p-vectors with finite variance and zero mean, independent also of $\Theta = (\Theta_1, \ldots, \Theta_p)$. We observe

$$X_i = (X_{i1}, \ldots, X_{ip}) = Q_i + \Theta \tag{1.2}$$

for $1 \leq i \leq n$. The mean vector,

$$\bar{X} = (\bar{X}_1, \ldots, \bar{X}_p) = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{Q} + \Theta \,, \tag{1.3}$$

is an empirical approximation to Θ. (Here $\bar{Q} = n^{-1} \sum_i Q_i$ equals the mean of the p-vectors Q_i.) The components of $\bar{X}$ can also be ranked, as

$$\bar{X}_{(1)} \leq \ldots \leq \bar{X}_{(p)} \,, \tag{1.4}$$

and there is a permutation $\widehat{R}_1, \ldots, \widehat{R}_p$ of $1, \ldots, p$ such that $\bar{X}_{(j)} = X_{\widehat{R}_j}$ for each j. If the common distribution of the Θ_js is continuous then, regardless of the distribution of the components of Q_i, the inequalities in (1.4) are strict with probability 1.

The permutation $\widehat{R} = (\widehat{R}_1, \ldots, \widehat{R}_p)$ serves as an approximation to R, and we wish to determine the accuracy of that approximation. In particular, for what values of $j_0 = j_0(n, p)$, and for what relationships between n and p, is it true that

$$P\big(\widehat{R}_j = R_j \;\text{ for }\; 1 \leq j \leq j_0\big) \to 1 \tag{1.5}$$

as n and p diverge? That is, how deeply into the ranking can we go before the connection between the true ranking and its empirical form is seriously degraded by noise?

The answer to this question depends to some extent on the extent of dependence among the components of each Q_i. To elucidate this point, let us consider the case where all the components of Q_i are identical; this is an extreme case of strong dependence. Then the components of $\bar{Q}$ are also also identical. Clearly, in this case $\widehat{R}_j = R_j$ for each j, and so (1.5) holds in a trivial and degenerate fashion. Other strongly dependent cases, although not as clear-cut as this one, can also be shown to be ones where $\widehat{R}_j = R_j$ with high probability for many values of j.

The case which is most difficult, i.e. where the strongest conditions are needed to ensure that (1.5) holds, occurs when the components of Q_i are independent. To emphasise this point we give sufficient conditions for (1.5), and show that when the components of each Q_i are independent, those conditions are also necessary. Our arguments can be modified to show that the conditions continue to be necessary under sufficiently weak dependence, for example if the components are m-dependent where $m = m(n)$ diverges sufficiently slowly as n increases.

The assumptions under which (1.5) holds are determined mainly by the lower tail of the common distribution of the Θ_js. If that distribution has an exponentially light left-hand tail, for example if the tail is like that of a normal distribution, then a sufficient condition for (1.5) is that j_0 should increase at a strictly slower rate than $n^{1/4}\,(\log n)^c$, where the constant c, which can be either positive or negative, depends on the rate of decay of the exponential lower tail of the distribution of Θ. For example, $c = 0$ if the distribution is double exponential, and $c = -\frac{1}{4}$ if it is normal. As indicated in the previous paragraph, the condition $j_0 = o\{n^{1/4}\,(\log n)^c\}$ is also necessary for (1.5) if the components of the Q_is are independent.

These results have several interesting aspects, including: (a) The exponent $\frac{1}{4}$ in the condition $j_0 = o\{n^{1/4}\,(\log n)^c\}$ does not change among different types of distribution with exponential tails; (b) the exponent is quite small, implying that the empirical rankings $\widehat{R}_j$ quite quickly become unreliable as predictors of the true rankings R_j; and (c) the critical condition $j_0 = o\{n^{1/4}\,(\log n)^c\}$ does not depend on the value of p. (We assume that p diverges at no faster than a polynomial rate in n, but we impose no upper bound on the degree of that polynomial.)

The condition $j_0 = o\{n^{1/4} (\log n)^c\}$ changes in important ways if the lower tail of the distribution of the Θ_js decays relatively slowly, for example at the polynomial rate $x^{-\alpha}$ as $x \to \infty$. Examples of this type include Pareto distributions. Here a sufficient condition for (1.5) to hold is $j_0 = o(n^{\alpha/2} p^{1/(2\alpha+1)})$, and this assumption is necessary if the components of the Q_is are independent. In this setting, unlike the exponential case, the value of dimension, p, plays a major role in addition to the sample size, n.

The exponential case can be regarded as the limit, as $\alpha \to \infty$, of the polynomial case. More generally, note that as the left-hand tail of the common distribution of the Θ_js becomes heavier, the value of j_0 can be larger before (1.5) fails. That is, if the distribution of the Θ_js has a heavier left-hand tail then the empirical rankings $\widehat{R}_j$ approximate the true rankings R_j for a greater number of values of j, before they degenerate into noise.

This analysis has focused on cases where the ranks of the Q_js are estimated by ranking empirical means of noisy observations of those quantities; see (1.4). However, essentially similar results are obtained if we rank other measures of location, for example L_q location estimators for general $q \geq 1$. In particular, if we take the variables Q_{ij} to have zero median, rather than zero mean, and continue to define X_i by (1.2) but replace the ranking in (1.4) by a ranking of medians, then the results above and those in section 2 continue to hold, modulo minor changes to the regularity conditions.

The model suggested by (1.2), where data on Θ arise in the form of p-vectors $X_1, \ldots, X_n$, is attractive in a number of high-dimensional settings, for example genomics. There, the jth component X_{ij} of X_i would typically represent the expression level of the jth gene of the ith individual in a sample. However, in other cases the means $\bar{X}_1, \ldots, \bar{X}_p$ at (1.3), or medians or other location estimators, might be computed from quite different datasets, one for each component index j. Moreover, those datasets might be of different sizes, $n_1, \ldots, n_p$ say, and then the argument that they arise naturally in the form of vectors would be inappropriate. This setting can arise when data are used to rank institutions, for example schools or universities. The conclusions discussed earlier in this section, and the theoretical properties developed in section 2 below, continue to apply in this case provided there is an "average" value, n say, of the n_js which represents all of them, in the sense that

$n = O(\min_j n_j)$ and $\max_j n_j = O(n)$ as n diverges. Additionally, in such cases it is often realistic to make the assumption that the corresponding centred means (or medians, etc) $\bar{Q}_j = n^{-1} \sum_i Q_{ij}$ are stochastically independent of one another, and so the particular results that are valid in this case are immediately available.

2. THEORETICAL PROPERTIES

We shall assume one of two types of lower tail for the common distribution function, F, of the random variables Θ_j: either it decreases exponentially fast, in which case we suppose that $F(-x) \asymp x^\beta \exp(-C_0 x^\alpha)$ as $x \to \infty$, where $\alpha > 0$ and $-\infty < \beta < \infty$; or it decreases polynomially fast, in which case $F(-x) \asymp x^{-\alpha}$ as $x \to \infty$, where $C_0, \alpha > 0$. (The notation $f(x) \asymp g(x)$, for positive functions f and g, will be taken to mean that $f(x)/g(x)$ is bounded away from zero and infinity as $x \to \infty$.) The former case covers distributions such as the normal, exponential and Subbotin; the latter, distributions such as the Pareto, Student's t and non-normal stable laws (e.g. the Cauchy).

It is convenient to impose the shape constraints on the densities, which we assume to exist in the lower tail, rather than on the distribution functions. Therefore we assume that one of the following two conditions holds as $x \to \infty$:

$$(d/dx)\, F(-x) \asymp (d/dx)\, x^\beta \exp\big(-C_0\, x^\alpha\big)\,, \tag{2.1}$$

$$(d/dx)\, F(-x) \asymp (d/dx)\, x^{-\alpha}\,. \tag{2.2}$$

In both (2.1) and (2.2), α must be strictly positive, but β in (2.1) can be any real number. The constant C_0 in (2.1) must be positive. We assume too that:

for fixed constants $C_1, \ldots, C_5 > 0$, where $C_2 > 2\,(C_1 + 1)$ and $C_4 < C_5$, $p = O(n^{C_1})$ as $n \to \infty$, and, for each $j \geq 1$, $E|Q_j|^{C_2} \leq C_3$, and $E(Q_j^2) \in [C_4, C_5]$. (2.3)

Recall from section 1 that we wish to examine the probability that the true ranks R_j, and their estimators $\widehat{R}_j$, are identical over the range $1 \leq j \leq j_0$. We consider both j_0 and p to be functions of n, so that the main dependent variable can be considered to be n. With this interpretation, define

$$\nu_{\rm exp} = \nu_{\rm exp}(n) = n^{1/4}\, (\log n)^{\{(1/\alpha)-1\}/2}\,, \quad \nu_{\rm pol} = \nu_{\rm pol}(n) = n^{\alpha/2}\, p^{1/(2\alpha+1)}\,, \tag{2.4}$$

where the subscripts denote "exponential" and "polynomial," respectively, and refer to the respective cases represented by (2.1) and (2.2). In the theorem below we impose the additional condition that, for some $\epsilon > 0$, $n = O(p^{4+\epsilon})$ when (2.1) holds, or $n = O(p^{\{4/(2\alpha+1)\}+\epsilon})$ in the context of (2.2), thereby ensuring that in these respective instances, $\nu_{\rm exp} = O(p^{1-\delta})$ and $\nu_{\rm pol} = O(p^{1-\delta})$ for some $\delta > 0$. This is to be expected, since we anticipate that we shall not be able to estimate correctly as many as $O(p)$ rankings.

Theorem 1. *Assume* (2.3)*, and that either (a)* (2.1) *holds and* $n = O(p^{4+\epsilon})$ *for some* $\epsilon > 0$*, or (b)* (2.2) *holds and* $n = O(p^{\{4/(2\alpha+1)\}+\epsilon})$ *for some* $\epsilon > 0$*. In case (a), if*

$$j_0 = o(\nu_{\rm exp}) \tag{2.5}$$

as $n \to \infty$ *then* (1.5) *holds. Conversely, when the components of the vectors* Q_i *are independent,* (2.5) *is necessary for* (1.5)*. In case (b), if*

$$j_0 = o(\nu_{\rm pol})\,, \tag{2.6}$$

then (1.5) *obtains. Conversely, when the components of the vectors* Q_i *are independent,* (2.6) *is necessary for* (1.5)*.*

It is of interest to consider cases where the common distribution, F, of the Θ_js is bounded to the left, for example where $F(x) \asymp x^\alpha$ as $x \downarrow 0$. However, it can be shown that in this context, unless p is constrained to be a sufficiently low degree polynomial function of n, very few of the estimated ranks $\widehat{R}_j$ will agree with the correct values R_j.

To indicate why, we first recall the model introduced in section 1, where the estimated ranks $\widehat{R}_j$ are derived by ordering the values of $\bar{Q}_j + \Theta_j$. Here $\bar{Q}_j = n^{-1} \sum_{1 \le i \le n} Q_{ij}$ is the average value of n independent and identically distributed random variables with zero mean. Therefore the means, $\bar{Q}_j$, are of order $n^{-1/2}$. By way of contrast, if we take $\alpha = 1$ in the formula $F(x) \asymp x^\alpha$ as $x \downarrow 0$, for example if F is the uniform distribution on $[0,1]$, then the spacings of the order statistics $\Theta_{(1)} \le \ldots \le \Theta_{(p)}$ are approximately of size p^{-1}. (More concisely, they are of size Z/p where Z has an exponential distribution.) Therefore, if p is of larger order than $n^{1/2}$ then the errors of the "estimators" $\bar{Q}_j + \Theta_j$ of Θ_j, for $1 \le j \le p$, are an order of magnitude larger than the spacings among the Θ_js. This can make it very difficult to estimate the ranks of the Θ_js from the ranks of values of $\bar{Q}_j + \Theta_j$. Indeed, it can be shown that, in the difficult case where the components of the Q_is are independent, and even for fixed j_0, if $\alpha = 1$ and p is of larger order than n^2 then

$$P\big(\widehat{R}_j = R_j \ \text{ for } \ 1 \le j \le j_0\big) \to 0\,. \tag{2.7}$$

Compare (1.5).

This explains why, when $F(x) \asymp x^\alpha$, it can be quite rare for the estimated ranks $\widehat{R}_j$ to match their true values. Indeed, no matter what the value of α and no matter what the value of j_0, property (1.5) will typically fail to hold unless p is no greater than a sufficiently small power of n, in particular unless $p = o(n^{\alpha/2})$, as the next result indicates.

Theorem 2. *Assume that $(d/dx)\,F(x) \asymp x^{\alpha-1}$ as $x \downarrow 0$, where $\alpha > 0$, and that (2.3) holds. Part (a): Instances where (1.5) holds and $p^2/n^\alpha \to 0$. Under the latter condition, (i) if $\alpha < \frac{1}{2}$ then (1.5) holds even for $j_0 = n$; (ii) if $\alpha = \frac{1}{2}$ then (1.5) holds provided that*

$$(\log j_0)^{2\alpha}\,(p^2/n^\alpha) \to 0\,; \tag{2.8}$$

and (iii) if $\alpha > \frac{1}{2}$ then (1.5) holds provided that

$$j_0 = o\big\{(n^{\alpha/2}/p)^{1/(2\alpha-1)}\big\}\,. \tag{2.9}$$

Part (b): Converses to (a)(ii) and (a)(iii). If $p^2/n^\alpha \to 0$ and the components of the vectors Q_i are independent then, if (1.5) holds, so too does (2.8) (if $\alpha = \frac{1}{2}$) or (2.9) (if $\alpha > \frac{1}{2}$). Part (c): Instances where (2.7) holds. If $\alpha > 0$ and $p^2/n^\alpha \to \infty$, and if the components of the vectors Q_i are independent, then (2.7) holds even for $j_0 = 1$.

The proof of Theorem 2 is similar to that of Theorem 1, and so is omitted. Theorem 1 is derived in section P.

P. PROOF OF THEOREM 1

Step 1: Bounds for probability of rank agreements. Given an integer $j_0 \in [1, p-1]$, let $\mathcal{E}(j_0)$ denote the event that $\bar{Q}_{R_j} + \Theta_{R_j} > \bar{Q}_{R_{j_0}} + \Theta_{R_{j_0}}$ for $j_0 + 1 \le j \le p$. Note that

$$\begin{aligned}
&\big\{\widehat{R}_j = R_j \ \text{ for } \ 1 \le j \le j_0\big\} \\
&\qquad \supseteq \Big\{\big|\bar{Q}_{R_j}\big| \le \tfrac{1}{2}\,\min\big(\Theta_{(j)} - \Theta_{(j-1)}, \Theta_{(j+1)} - \Theta_{(j)}\big) \ \text{ for } \ 1 \le j \le j_0\Big\} \cap \mathcal{E}(j_0)\,,
\end{aligned}$$

where we define $\Theta_{(j-1)} = -\infty$ if $j = 1$. Therefore, writing $\tilde{\mathcal{E}}(j_0)$ for the complement of $\mathcal{E}(j_0)$, and defining $\pi(j_0) = P(\widehat{R}_j = R_j \text{ for } 1 \le j \le j_0)$, we deduce that

$$\pi(j_0) \ge P\Big\{\big|\bar{Q}_{R_j}\big| \le \tfrac{1}{2} \min\big(\Theta_{(j)} - \Theta_{(j-1)}, \Theta_{(j+1)} - \Theta_{(j)}\big) \text{ for } 1 \le j \le j_0\,;\ \mathcal{E}(j_0)\Big\}$$
$$\ge 1 - \sum_{j=1}^{j_0} P\Big\{\big|\bar{Q}_{R_j}\big| > \tfrac{1}{2} \min\big(\Theta_{(j)} - \Theta_{(j-1)}, \Theta_{(j+1)} - \Theta_{(j)}\big)\Big\}$$
$$- P\{\tilde{\mathcal{E}}(j_0)\}\,. \qquad (P.1)$$

Also,

$$\big\{\widehat{R}_j = R_j \text{ for } 1 \le j \le j_0\big\}$$
$$= \Big\{\bar{X}_{R_1} \le \ldots \le \bar{X}_{R_{j_0}} \text{ and } \bar{X}_j > \bar{X}_{R_{j_0}} \text{ for } j \notin \{R_1, \ldots, R_{j_0}\}\Big\}$$
$$= \Big\{\Theta_{(j+1)} - \Theta_{(j)} \ge -\big(\bar{Q}_{R_{j+1}} - \bar{Q}_{R_j}\big) \text{ for } 1 \le j \le j_0$$
$$\text{and } \Theta_j - \Theta_{(j_0)} \ge -\big(\bar{Q}_j - \bar{Q}_{R_{j_0}}\big) \text{ for } j \notin \{R_1, \ldots, R_{j_0}\}\Big\},$$

and so

$$\pi(j_0) \le P\Big\{\Theta_{(j+1)} - \Theta_{(j)} \ge -\big(\bar{Q}_{R_{j+1}} - \bar{Q}_{R_j}\big) \text{ for } 1 \le j \le j_0\Big\}. \qquad (P.2)$$

Write $\mathcal{J} = \mathcal{J}(j_0)$ for the set of even integers j satisfying $1 \le j \le j_0$. Let $\mathcal{E}_j$ be the event that $\Theta_{(j+1)} - \Theta_{(j)} \ge -(\bar{Q}_{R_{j+1}} - \bar{Q}_{R_j})$, and let $\pi_1(j_0)$ denote the probability that $\mathcal{E}_j$ holds for all $j \in \mathcal{J}$. Then by (P.2),

$$\pi(j_0) \le \pi_1(j_0)\,. \qquad (P.3)$$

Write $\tilde{\mathcal{E}}_j$ for the complement of $\mathcal{E}_j$, let $\mathcal{O}$ denote the sigma-field generated by $\Theta_1, \Theta_2, \ldots$, and note that, if the components of each Q_i are independent, then the events $\mathcal{E}_j$, for $j \in \mathcal{J}$, are independent conditional on $\mathcal{O}$. Therefore,

$$\pi_1(j_0) = E\bigg\{P\bigg(\bigcap_{j\in\mathcal{J}} \mathcal{E}_j \,\bigg|\, \mathcal{O}\bigg)\bigg\} = E\bigg[\prod_{j\in\mathcal{J}} \big\{1 - P\big(\mathcal{E}_j \,|\, \mathcal{O}\big)\big\}\bigg]$$
$$\le E\bigg[\exp\bigg\{-\sum_{j\in\mathcal{J}} P\big(\tilde{\mathcal{E}}_j \,|\, \mathcal{O}\big)\bigg\}\bigg]. \qquad (P.4)$$

Step 2: Moderate deviation arguments. Using the arguments of Rubin and Sethuraman (1965) and Amosova (1972) it can be shown that, if the constant C_2 in (2.3) satisfies $C_2 > B^2 + 2$ where $B > 0$, then as n (and hence also p) diverges,

$$P\big\{|\bar{Q}_j| > x\,(\mathrm{var}\,\bar{Q}_j)^{1/2}\big\} = \{1 + o(1)\}\,2\,\{1 - \Phi(x)\}\,, \qquad (P.5)$$

$$P\Big[-\big(\bar{Q}_{j_1} - \bar{Q}_{j_2}\big) \ge x\,\big\{\mathrm{var}\,\big(\bar{Q}_{j_1} - \bar{Q}_{j_2}\big)\big\}^{1/2}\Big] = \{1 + o(1)\}\,\{1 - \Phi(x)\}\,, \quad (P.6)$$

uniformly in $0 < x < B\,(\log p)^{1/2}$ and $j, j_1, j_2 \geq 1$ such that $j_1 \neq j_2$. (To derive (P.6), and also (P.8) and (P.10) below, we also need the assumption that the components of each Q_i are independent; this is not required for (P.5), (P.7) or (P.9). In (P.5) and (P.6), Φ denotes the standard normal distribution function.) Therefore, since $C_2 > 2\,(C_1 + 1)$ in (2.3), we can take $B = (2+\epsilon)^{1/2}$ for some $\epsilon > 0$, and then (P.5) and (P.6) hold uniformly in $0 < x < \{(2+\epsilon)\,\log p\}^{1/2}$. It follows that (P.5) and (P.6) imply that

$$P\big\{|\bar{Q}_j| > x\,(\mathrm{var}\,\bar{Q}_j)^{1/2}\big\} = \{1+o(1)\}\,2\,\{1-\Phi(x)\} + o\big(p^{-1}\big)\,, \quad (P.7)$$

$$P\Big[-\big(\bar{Q}_{j_1} - \bar{Q}_{j_2}\big) \geq x\,\big\{\mathrm{var}\,\big(\bar{Q}_{j_1} - \bar{Q}_{j_2}\big)\big\}^{1/2}\Big] = \{1+o(1)\}\,\{1-\Phi(x)\} + o\big(p^{-1}\big) \qquad (P.8)$$

as $n \to \infty$, uniformly in all $x > 0$ and all $j, j_1, j_2 \geq 1$ such that $j_1 \neq j_2$.

Results (P.7) and (P.8) enable us to write down approximations to the series on the right-hand sides of (P.1) and (P.4):

$$\sum_{j=1}^{j_0} P\Big\{\big|\bar{Q}_{R_j}\big| > \tfrac{1}{2}\,\min\big(\Theta_{(j)} - \Theta_{(j-1)}, \Theta_{(j+1)} - \Theta_{(j)}\big)\Big\} = 2\,\{1+o(1)\}\sum_{j=1}^{j_0} P(|N| > T_{1j}) + o(1)\,, \quad (P.9)$$

$$\sum_{j\in\mathcal{J}} P\big(\tilde{\mathcal{E}}_j\big) = \{1+o(1)\}\sum_{j\in\mathcal{J}} P(N > T_{2j}) + o(1)\,, \qquad (P.10)$$

$$E\bigg[\exp\bigg\{-\sum_{j\in\mathcal{J}} P\big(\tilde{\mathcal{E}}_j \,\big|\, \mathcal{O}\big)\bigg\}\bigg] \leq \{1+o(1)\}\,E\bigg[\exp\bigg\{-(1+\Delta)\sum_{j\in\mathcal{J}} P(N > T_{2j} \,|\, \mathcal{O})\bigg\}\bigg]\,, \quad (P.11)$$

where N denotes a standard normal random variable independent of $\mathcal{O}$,

$$T_{1j} = \frac{\min(\Theta_{(j)} - \Theta_{(j-1)}, \Theta_{(j+1)} - \Theta_{(j)})}{2\,(\mathrm{var}\,\bar{Q}_{R_j})^{1/2}}\,, \quad T_{2j} = \frac{\Theta_{(j+1)} - \Theta_{(j)}}{\{\mathrm{var}\,(\bar{Q}_{R_{j+1}} - \bar{Q}_{R_j})\}^{1/2}}\,,$$

and, for a sequence of positive constants δ_n converging to zero, the random variable Δ satisfies $P(|\Delta| \leq \delta_n) = 1$. Not too that

$$P(|N| > T_{4j}) \leq P(|N| > T_{1j}) \leq P(|N| > T_{3j}) + P(|N| > T_{4j})\,, \qquad (P.12)$$

where, defining $\Theta_{(j-1)} = -\infty$ if $j = 1$, we write

$$T_{3j} = \frac{\Theta_{(j)} - \Theta_{(j-1)}}{2\,(\mathrm{var}\,\bar{Q}_{R_j})^{1/2}}\,, \quad T_{4j} = \frac{\Theta_{(j+1)} - \Theta_{(j)}}{2\,(\mathrm{var}\,\bar{Q}_{R_j})^{1/2}}\,.$$

Step 3: Rényi's representation. If $U_{(1)} \le \ldots \le U_{(p)}$ denote the order statistics of a sample of size p drawn from the uniform distribution on $[0,1]$ then, for each p, we can construct independent random variables $Z_1, \ldots, Z_p$ with the standard negative exponential distribution on $[0,1]$, such that, for $1 \le j \le p$, $U_{(j)} = 1 - \exp(-V_j)$ where

$$V_j = \sum_{k=1}^{j} \frac{Z_k}{p-k+1} = w_j + W_j$$

and, uniformly in $1 \le j \le \frac{1}{2}\,p$ and $2 \le p < \infty$,

$$w_j = \sum_{k=p-j+1}^{p} \frac{1}{k} = \frac{j}{p} + O\big(j^2/p^2\big) = O(j/p)\,, \qquad (P.13)$$

$$W_j = \sum_{k=p-j+1}^{p} k^{-1}\,(Z_{p-k+1} - 1)\,, \quad \sup_{1 \le j \le p/2} j^{-1/2}\,|W_j| \le p^{-1}\,W(p)\,, \qquad (P.14)$$

$$\sup_{1 \le j \le p/2} j^{-3/2} \left| W_j - \frac{1}{p} \sum_{k=p-j+1}^{p} (Z_{p-k+1} - 1) \right| \le p^{-2}\,W(p)\,, \qquad (P.15)$$

where the nonnegative random variable $W(p)$, which without loss of generality we take to be common to (P.14) and (P.15), satisfies $P\{W(p) > p^{\epsilon}\} = O(p^{-C})$ for each $C, \epsilon > 0$.

Step 4: Approximations to $U_{(j+1)} - U_{(j)}$. Using the second identity in (P.13), and (P.14), we deduce that

$$\begin{aligned} U_{(j+1)} - U_{(j)} &= (V_{j+1} - V_j)\Big\{1 - \tfrac{1}{2}\,(V_{j+1} + V_j) \\ &\qquad\qquad + \tfrac{1}{6}\,(V_{j+1}^2 + V_j\,V_{j+1} + V_j^2) - \ldots\Big\} \\ &= \frac{Z_{j+1}}{p-j}\left\{1 + \Psi_{j1}\left(\frac{j}{p} + \frac{S_{j1}}{p^{1/2}}\right)\right\}, \end{aligned} \qquad (P.16)$$

uniformly in $1 \le j \le \frac{1}{2}\,p$, where the random variable Ψ_{j1} satisfies, for $k = 1$,

$$P\bigg(\max_{1 \le j \le p/2} |\Psi_{jk}| \le A\bigg) = 1\,, \qquad (P.17)$$

$A > 0$ is an absolute constant, and for each $C, \epsilon > 0$ the nonnegative random variable S_{j1} satisfies, with $k = 1$,

$$P\bigg(\sup_{1 \le j \le p/2} S_{jk} > p^{\epsilon}\bigg) = O\big(p^{-C}\big)\,. \qquad (P.18)$$

Using the third identity in (P.13), and (P.15), we deduce that

$$0 \le U_{(j)} = w_j + W_j - \tfrac{1}{2}\,(w_j + W_j)^2 + \ldots = \frac{j}{p} + \Psi_{j2}\left(\frac{j^2}{p^2} + \frac{j^{1/2}\,S_{j2}}{p}\right), \quad (P.19)$$

where Ψ_{j2} and $S_{j2} \ge 0$ satisfy (P.17) and (P.18), respectively.

Define $D_j = U_{(j+1)} - U_{(j)}$. If the common distribution function of the Θ_js is F then, by Taylor expansion,

$$\Theta_{(j+1)} - \Theta_{(j)} = F^{-1}(U_{(j)} + D_j) - F^{-1}(U_{(j)}) = D_j\,\big(F^{-1}\big)'\,(U_{(j)} + \omega_j\,D_j)\,, \quad (P.20)$$

where $0 \le \omega_j \le 1$. In the case of exponential tails we take, without loss of generality, $C_0 = 1$ in (2.1), and then:

$$\big(F^{-1}\big)'(1-u) \asymp -(d/du)\,\big(\log u^{-1}\big)^{1/\alpha} = \alpha^{-1}\,u^{-1}\,\big(\log u^{-1}\big)^{(1/\alpha)-1} \qquad (P.21)$$

as $u \downarrow 0$. By (P.20) and (P.21),

$$\Theta_{(j+1)} - \Theta_{(j)} = \Psi_j\,\frac{D_j}{U_{(j)} + \omega_j\,D_j}\,\big\{\log(U_{(j)} + \omega_j\,D_j)\big\}^{(1/\alpha)-1}\,, \qquad (P.22)$$

where Ψ_j denotes a random variable satisfying, for constants B_1, B_2 and B_3 satisfying $0 < B_1 < B_2 < \infty$ and $0 < B_3 < 1$,

$$P\Big(B_1 \le \Psi_j \le B_2 \quad \text{for all } j \text{ such that} \quad U_{(j+1)} < B_3\Big) = 1\,. \qquad (P.23)$$

In the case of polynomial tails, (2.2) implies that

$$\big(F^{-1}\big)'(u) \asymp -(d/du)\,u^{-1/\alpha} = \alpha^{-1}\,u^{-(1/\alpha)-1}$$

in place of (P.21), and therefore, instead of (P.22),

$$\Theta_{(j+1)} - \Theta_{(j)} = \Psi_j\,\frac{D_j}{(U_{(j)} + \omega_j\,D_j)^{(1/\alpha)+1}}\,, \qquad (P.24)$$

where Ψ_j again satisfies (P.23).

Step 5: Bounds to probabilities of events based on $\Theta_{(j+1)} - \Theta_{(j)}$. By assumption in Theorem 1, imposed separately in the exponential and polynomial cases, $n = O(p^C)$, and by (2.3), $p = O(n^{C_1})$, where $C, C_1 > 0$. In view of (P.16), (P.19) and (P.22), in the case of exponential tails there exist $B_4, B_5 > 0$ such that, for any choice of constants c_1, c_2 satisfying $0 < c_1 < c_2 < C^{-1}$, and for all $B_6 > 0$,

$$\inf_{j\in[1,n^{c_1}]} P\Big\{(\Theta_{(j+1)} - \Theta_{(j)})\, Z_{j+1}^{-1}\, (\log n)^{1-(1/\alpha)} \geq B_4\, n^{-c_1}\Big\} = 1 - O\big(n^{-B_6}\big)\,, \qquad (P.25)$$

$$\inf_{j\in[n^{c_1},n^{c_2}]} P\Big\{B_4 \leq j\,(\Theta_{(j+1)} - \Theta_{(j)})\, Z_{j+1}^{-1}\, (\log n)^{1-(1/\alpha)} \leq B_5\Big\} = 1 - O\big(n^{-B_6}\big)\,. \qquad (P.26)$$

(Noting the definitions $C = 4 + \epsilon$ and $C = \{4/(2\alpha+1)\} + \epsilon$ (where $\epsilon > 0$), in cases where the lower tail of Θ_j is exponential or polynomial, respectively, we see that the constraint $c_2 < C^{-1}$ permits n^{c_2} to be of size $\nu_{\mathrm{exp}}\, n^{\epsilon_1}$ or $\nu_{\mathrm{pol}}\, n^{\epsilon_1}$ (where $\epsilon_1 > 0$), in these respective instances.) Likewise, in the case of polynomial tails, and using (P.16), (P.19) and (P.24), there exist $B_4, B_5 > 0$ such that, for $0 < c_1 < c_2 < C^{-1}$ and for all $B_6 > 0$,

$$\inf_{j\in[1,n^{c_1}]} P\Big\{(\Theta_{(j+1)} - \Theta_{(j)})\, Z_{j+1}^{-1}\, p^{-1/\alpha} \geq B_4\, n^{-c_1\,\{(1/\alpha)+1\}}\Big\} = 1 - O\big(n^{-B_6}\big)\,, \qquad (P.27)$$

$$\inf_{j\in[n^{c_1},n^{c_2}]} P\Big\{B_4 \leq j^{(1/\alpha)+1}\,(\Theta_{(j+1)} - \Theta_{(j)})\, Z_{j+1}^{-1}\, p^{-1/\alpha} \leq B_5\Big\} = 1 - O\big(n^{-B_6}\big)\,. \qquad (P.28)$$

Step 6: Bounds to the series in (P.9)–(P.11), and completion of proof, in the exponential case. Define $\ell = (\log n)^{(1/\alpha)-1}$, let N be a standard normal random variable independent of $\mathcal{O}$, and let Z be independent of N and have the standard negative exponential distribution. Let K_1 be a positive constant. If a_n is a sequence of positive numbers and f_n is a sequence of nonnegative functions, write $a_n \doteq f_n(K)$ to mean that, for constants $L_1, L_2 > 1$, either (a) $a_n \leq L_1\, f_n(K)$ whenever $K \geq L_2$ and n is sufficiently large, and $a_n \geq L_1^{-1}\, f_n(K)$ whenever $K \leq L_2^{-1}$ and n is sufficiently large, or (b) $a_n \geq L_1^{-1}\, f_n(K)$ whenever $K \geq L_2$ and n is sufficiently large, and $a_n \leq L_1\, f_n(K)$ whenever $K \leq L_2^{-1}$ and n is sufficiently large. Let $0 < c_1 < c_2 < \frac{1}{2}$ and $c_1 < \frac{1}{4}$, and let j_0 and j_1 denote constants satisfying $|j_1 - n^{c_1}| \leq 1$, $j_1 \leq j_0 \leq n^{c_2}$ and $j_1/j_0 \to 0$.

When (2.1) holds with $C_0 = 1$, properties (P.25) and (P.26) imply that, for each $B_6 > 0$,

$$s(n) \equiv \sum_{j=1}^{j_0} P\{|N| > n^{1/2}\, K_1\, (\Theta_{(j+1)} - \Theta_{(j)})\} \qquad (P.29)$$

$$
\begin{aligned}
&\doteq O\big\{ j_1\, P\big(|N| > K_2\, n^{1/2}\, j_1^{-1}\, Z\, \ell\big) + n^{-B_6} \big\} + \sum_{j_1 < j \le j_0} P\big(|N| > K\, n^{1/2}\, j^{-1}\, Z\, \ell\big) \\
&\doteq O\bigg\{ j_1 \bigg(P\big(Z \le n^{-1/2}\, j_1\, \ell^{-1}\big) \\
&\quad + E\Big[\big(n^{1/2}\, j_1^{-1}\, Z\, \ell\big)^{-1} \exp\big\{ -\tfrac{1}{2}\, (K\, n^{1/2}\, j_1^{-1}\, Z\, \ell)^2 \big\}\, I\big(Z > n^{-1/2}\, j_1\, \ell^{-1}\big) \Big] \bigg) \bigg\} \\
&\quad + \sum_{j_1 < j \le j_0} \bigg(P\big(Z \le n^{-1/2}\, j\, \ell^{-1}\big) \\
&\quad + E\Big[\big(n^{1/2}\, j^{-1}\, Z\, \ell\big)^{-1} \exp\big\{ -\tfrac{1}{2}\, (K\, n^{1/2}\, j^{-1}\, Z\, \ell)^2 \big\}\, I\big(Z > n^{-1/2}\, j\, \ell^{-1}\big) \Big] \bigg) \\
&\doteq O\bigg\{ j_1 \bigg(n^{-1/2}\, j_1\, \ell^{-1} \\
&\quad + E\Big[\big(n^{1/2}\, j_1^{-1}\, Z\, \ell\big)^{-1} \exp\big\{ -\tfrac{1}{2}\, (K\, n^{1/2}\, j_1^{-1}\, Z\, \ell)^2 \big\}\, I\big(Z > n^{-1/2}\, j_1\, \ell^{-1}\big) \Big] \bigg) \bigg\} \\
&\quad + \sum_{j_1 < j \le j_0} \bigg(n^{-1/2}\, j\, \ell^{-1} \\
&\quad + E\Big[\big(n^{1/2}\, j^{-1}\, Z\, \ell\big)^{-1} \exp\big\{ -\tfrac{1}{2}\, (K\, n^{1/2}\, j^{-1}\, Z\, \ell)^2 \big\}\, I\big(Z > n^{-1/2}\, j\, \ell^{-1}\big) \Big] \bigg) .
\end{aligned}
$$

Now,

$$
\begin{aligned}
&E\Big[\big(n^{1/2}\, j^{-1}\, Z\, \ell\big)^{-1} \exp\big\{ -\tfrac{1}{2}\, (K\, n^{1/2}\, j^{-1}\, Z\, \ell)^2 \big\}\, I\big(Z > n^{-1/2}\, j\, \ell^{-1}\big) \Big] \\
&\qquad = \int_{n^{-1/2}\, j\, \ell^{-1}}^{\infty} \big(n^{1/2}\, j^{-1}\, z\, \ell\big)^{-1} \exp\big\{ -\tfrac{1}{2}\, (K\, n^{1/2}\, j^{-1}\, z\, \ell)^2 - z \big\}\, dz \\
&\qquad = n^{-1/2}\, j\, \ell \int_1^{\infty} u^{-1} \exp\big\{ -\tfrac{1}{2}\, (K\, u)^2 - n^{-1/2}\, j\, \ell\, u \big\}\, du \asymp n^{-1/2}\, j\, \ell\, .
\end{aligned}
$$

(Here we have used the fact that $j \le j_0 \le n^{c_2}$ where $c_2 < \frac{1}{2}$.) Therefore,

$$
\begin{aligned}
s(n) &\asymp j_1 \cdot n^{-1/2}\, j_1\, \ell^{-1} + \sum_{j_1 < j \le j_0} n^{-1/2}\, j\, \ell^{-1} \asymp n^{-1/2}\, j_1^2\, \ell^{-1} + n^{-1/2}\, j_0^2\, \ell^{-1} \\
&\asymp n^{-1/2}\, j_0^2\, \ell^{-1}\, . \qquad (P.30)
\end{aligned}
$$

(Here we have used the fact that $j_1/j_0 \to 0$.)

The right-hand side of (P.30) converges to zero if and only if $j_0 = o(n^{1/4}\, \ell^{1/2})$, or equivalently, if and only if (2.5) holds. Moreover, in view of (P.12) and depending on the choice of K_1 in the definition of $s(n)$ at (P.29), $s(n)$ can be either an upper bound or a lower bound to the series $\sum_{j \in \mathcal{J}} P(N > T_{2j})$ on the right-hand side of (P.10). Hence, the series on the left-hand side of (P.10) equals $o(1)$ as $n \to \infty$, if and only if (2.5) holds. It therefore follows from (P.1) that (2.5) is sufficient for (1.5).

Conversely, if (1.5) holds then, in view of (P.3), (P.4) and (P.11),

$$\sum_{j \in \mathcal{J}} P(N > T_{2j} \,|\, \mathcal{O}) \to 0$$

in probability. Therefore, by (P.25) and (P.26), with j_0 and j_1 as in the previous paragraph, there exists $K_1 > 0$ such that

$$\sum_{j_1 < j \le j_0} P\{|N| > n^{1/2}\, K_1\, (\Theta_{(j+1)} - \Theta_{(j)}) \,|\, \mathcal{O}\} \to 0$$

in probability. (We can take the sum over all $j \in [j_1 + 1, j_0]$, rather than just over even j, since (P.11) holds for sums over odd j as well as over even j.) Hence, arguing as in the lines below (P.29), we deduce that for sufficiently large $K_2 > 0$,

$$T(n) \equiv \sum_{j_1 < j \le j_0} f(Z_j/\delta_j) \to 0 \tag{P.31}$$

in probability, where the random variables Z_j are independent and have a common exponential distribution, $\delta_j = n^{-1/2}\, j\, \ell^{-1}$ and $f(z) = z^{-1} \exp(-K_2\, z^2)\, I(z > 1)$. We claim that this implies that the expected value of the left-hand side of (P.31) also converges to 0:

$$\sum_{j_1 < j \le j_0} E\{f(Z_j/\delta_j)\} \to 0\,, \tag{P.32}$$

or equivalently that $\sum_{j_1 < j \le j_0} \delta_j \to 0$, and thence (using the argument leading to (P.30)) that $s(n) \asymp n^{-1/2}\, j_0^2\, \ell^{-1} \to 0$, which is equivalent to (2.5). Therefore, if we establish (P.32) then we shall have proved that (1.5) implies (2.5).

It remains to show that (P.31) implies (P.32). This we do by contradiction. If (P.32) fails then, along a subsequence of values of n, the left-hand side of (P.32) converges to a nonzero number. For notational simplicity we shall make the inessential assumptions that the number is finite and that the subsequence involves all n, and we shall take $K_2 = 1$ in the definition of f. In particular,

$$t(n) \equiv \sum_{j_1 < j \le j_0} E\{f(Z_j/\delta_j)\} \to t(\infty)\,, \tag{P.33}$$

where $0 < t(\infty) < \infty$. Now, $t(n) = \{1 + o(1)\}\, \mu(1)\, \delta(n)$, where $\delta(n) = \sum_{j_1 < j \le j_0} \delta_j$ and, for general $\lambda \ge 1$, $\mu(\lambda) = \int_{z > \lambda} z^{-1} \exp(-z^2)\, dz$. Therefore,

$$\delta(n) \to \delta(\infty) \equiv t(\infty)/\mu(1)\,. \tag{P.34}$$

For each $\lambda > 1$ the left-hand side of (P.31) equals $\Delta_1 + \Delta_2$, where, in view of (P.33),

$$E(\Delta_2) = \sum_{j_1 < j \le j_0} E\{f(Z_j/\delta_j)\, I(Z_j > \lambda\, \delta_j)\} = \{1 + o(1)\}\, \mu(\lambda)\, \delta(n) \qquad (P.35)$$

and

$$\Delta_1 = \sum_{j_1 < j \le j_0} f(Z_j/\delta_j)\, I(Z_j \le \lambda\, \delta_j) = \sum_{j_1 < j \le j_0} f(W_j)\, I_j\,,$$

with $W_j = Z_j/\delta_j$ and $I_j = I(\delta_j \le Z_j \le \lambda\, \delta_j)$. However,

$$\sum_{j_1 < j \le j_0} P(I_j = 1) = \mu_1(\lambda)\, \delta(n) + o(1) = \delta(\infty)\, \mu_1(\lambda) + o(1)\,,$$

where $\mu_1(\lambda) = \int_{1<z<\lambda} z^{-1} \exp(-z^2)\, dz$. Therefore, in the limit as $n \to \infty$, Δ_1 equals a sum, S_λ say, of N independent random variables each having the distribution of $f(W)$, where W is uniformly distributed on $[1, \lambda]$, N has a Poisson distribution with mean $\delta(\infty)\, \mu_1(\lambda)$, and N and the summands are independent. The distribution of S_λ is stochastically monotone increasing, in the sense that $P(S_\lambda > s)$ increases with λ. On the other hand, since $\mu(\lambda) \to 0$ as $\lambda \to \infty$ then, by (P.34) and (P.35),

$$\lim_{\lambda \to \infty} \limsup_{n \to \infty} E(\Delta_2) = 0\,.$$

Combining these results we deduce that $\Delta_1 + \Delta_2$, i.e. $T(n)$ at (P.31), does not converge to zero in probability. This contradicts (P.31) and so establishes that $t(\infty)$ must equal zero; that is, (P.32) holds.

Step 7: Completion of proof in the polynomial case. Here we alter the constraints $0 < c_1 < c_2 < \frac{1}{2}$ and $c_1 < \frac{1}{4}$, imposed in the first paragraph of step 6, to $0 < c_1 < c_2 < C^{-1}$ and $c_1 < \frac{1}{2}\, \alpha$, where $C = \{4/(2\alpha + 1)\} + \epsilon$ is as given below (P.26), and $0 < \epsilon < 2/\{2(2\alpha + 1)\}$. We continue to take $1 \le j_1 \le j_0 \le n^{c_2}$ and $|j_1 - n^{c_1}| \le 1$. If we use (P.27) and (P.28) instead of (P.25) and (P.26), respectively, then, arguing as in the lines from (P.29) down, we obtain:

$$s(n) \doteq O\Big[j_1\, P\big\{N > K_2\, n^{1/2}\, j_1^{-1}\, Z\, (p/j_1)^{1/\alpha}\big\} + n^{-B_6}\Big] \\ + \sum_{j_1 < j \le j_0} P\big\{N > K\, n^{1/2}\, j^{-1}\, Z\, (p/j)^{1/\alpha}\big\}\,.$$

Moreover, since $n^{-1/2}\, j_0^{1+(1/\alpha)}\, p^{-1/\alpha} \to 0$ then, uniformly in $j \le j_0$, we have:

$$P\big\{N > K_2\, n^{1/2}\, j^{-1}\, Z\, (p/j)^{1/\alpha}\big\}$$

$$
\begin{aligned}
&\doteq \big\{n^{1/2}\, j^{-1}\, (p/j)^{1/\alpha}\big\}^{-1} \\
&\qquad + E\bigg(\big\{n^{1/2}\, j^{-1}\, Z\, (p/j)^{1/\alpha}\big\}^{-1} \exp\Big[-\tfrac{1}{2}\,\big\{K\, n^{1/2}\, j^{-1}\, Z\, (p/j)^{1/\alpha}\big\}^2\Big] \\
&\qquad\qquad\qquad\qquad\qquad\qquad \times I\big(Z > n^{-1/2}\, j\, (p/j)^{-1/\alpha}\big)\bigg) \\
&= \big\{n^{1/2}\, j^{-1}\, (p/j)^{1/\alpha}\big\}^{-1} + \int_{n^{-1/2}\, j\, (p/j)^{-1/\alpha}}^{\infty} \big\{n^{1/2}\, j^{-1}\, z\, (p/j)^{1/\alpha}\big\}^{-1} \\
&\qquad\qquad\qquad \times \exp\Big[-\tfrac{1}{2}\,\big\{K\, n^{1/2}\, j^{-1}\, z\, (p/j)^{1/\alpha}\big\}^2 - z\Big]\, dz \\
&= n^{-1/2}\, j\, (p/j)^{-1/\alpha} \int_1^{\infty} u^{-1} \exp\big\{-\tfrac{1}{2}\,(K\, u)^2\big\}\, du \asymp n^{-1/2}\, j\, (p/j)^{-1/\alpha}\,.
\end{aligned}
$$

Therefore,

$$
s(n) \asymp \sum_{j=1}^{j_0} n^{-1/2}\, j\, (p/j)^{-1/\alpha} \asymp n^{-1/2}\, j_0^{2+(1/\alpha)}\, p^{-1/\alpha}\,,
$$

and the right-hand side converges to zero if and only if $j_0 = o(\nu_{\rm pol})$, where $\nu_{\rm pol}$ is as at (2.4). The argument in the paragraph below the one containing (P.30) can therefore be used to prove, from (P.1), that (2.6) is sufficient for (1.5). The converse, when the Q_is have independent components, can also be derived as in step 6.

NOT-FOR-PUBLICATION NOTES about the case $F(x) = x^\alpha$ for $0 < x < 1$, with $\alpha > 0$: Here $(F^{-1})'(u) = \alpha^{-1}\, u^{(1/\alpha)-1}$, whence

$$
\Theta_{(j+1)} - \Theta_{(j)} \approx D_j\, U_{(j)}^{(1/\alpha)-1} \approx (Z/p)\, (j/p)^{(1/\alpha)-1} = Z\, j^{(1/\alpha)-1}\, p^{-1/\alpha}\,,
$$

and so

$$
\begin{aligned}
s(n) &\asymp \sum_{j=1}^{j_0} P\big(N > n^{1/2}\, K\, Z\, j^{(1/\alpha)-1}\, p^{-1/\alpha}\big) \\
&\asymp \sum_{j=1}^{j_0} n^{-1/2}\, j^{1-(1/\alpha)}\, p^{1/\alpha} \asymp n^{-1/2}\, p^{1/\alpha} \begin{cases} 1 & \text{if } \alpha < \frac{1}{2} \\ \log j_0 & \text{if } \alpha = \frac{1}{2} \\ j^{2-(1/\alpha)} & \text{if } \alpha > \frac{1}{2} \end{cases} \\
&\geq n^{-1/2}\, p^{1/\alpha}\,.
\end{aligned}
$$

Therefore, if $p \geq n^{\alpha/2}$ then it is not possible for $s(n)$ to converge to zero.