J. Appl. Probab. **61,** 198–213 (2024)
*doi:*10.1017/jpr.2023.26

EXTREMA OF A MULTINOMIAL ASSIGNMENT PROCESS

MIKHAIL LIFSHITS ,* *St. Petersburg State University*
GILLES MORDANT ,** *Universität Göttingen*

Abstract

We study the asymptotic behaviour of the expectation of the maxima and minima of a random assignment process generated by a large matrix with multinomial entries. A variety of results is obtained for different sparsity regimes.

Keywords: Expected maxima; minima; multinomial distribution; random assignment process

2020 Mathematics Subject Classification: Primary 60C05
Secondary 05C70; 60K30

1. Introduction and main results

1.1. Random assignment problem

We consider the following *random assignment problem*. Let (X_{ij}) be an $n \times n$ random matrix and let $[1..n]$ denote the set $\{1, 2, \ldots, n\}$. Let $\mathcal{S}_n$ denote the group of permutations $\sigma : [1..n] \mapsto [1..n]$. For every $\sigma \in \mathcal{S}_n$, let $S(\sigma) = \sum_{i=1}^{n} X_{i\sigma(i)}$.

The process $\{S(\sigma),\ \sigma \in \mathcal{S}_n\}$ is called a *random assignment process*. The problem consists in the study of the asymptotic behaviour of its extrema, in particular,

$$\mathbb{E} \max_{\sigma \in \mathcal{S}_n} S(\sigma) \quad \text{and} \quad \mathbb{E} \min_{\sigma \in \mathcal{S}_n} S(\sigma) \qquad \text{as } n \to \infty. \tag{1.1}$$

We refer to [6, 16] for many applications of assignment processes and their extrema in various fields of mathematics.

There are many remarkable results in the area [6, 10, 13, 17], including a famous result [1, 2] proving a conjecture in [11] claiming that $\lim_{n\to\infty} \mathbb{E} \min_{\sigma\in\mathcal{S}_n} S(\sigma) = \pi^2/6$ when the X_{ij} are independent and identically distributed (i.i.d.) standard exponential. Actually, it showed that, when the random variables considered are nonnegative, the distribution of X_{ij} affects the limit in the minimisation problem only through the value of its probability density function at 0.

In the case mentioned, the support of the common distribution is bounded on the left. The situation is very different when dealing with variables having unbounded supports. For obvious reasons, it is more convenient to illustrate this phenomenon for maxima instead of minima. If the common law of the entries is not bounded from above, then the expectation of maxima

Received 8 November 2022; revision received 31 March 2023.

* Postal address: St. Petersburg State University. Russia, 199034, St. Petersburg, University Emb. 7/9. Email: mikhail@lifshits.org

** Postal address: Institut für Mathematische Stochastik, Universität Göttingen. Göttingen, Germany. Email: gilles.mordant@uni-goettingen.de

no longer tends to a finite limit but grows to infinity and the problem consists in evaluation of the corresponding growth order. In this direction, [12] showed that if X_{ij} are i.i.d. standard Gaussian, then $\mathbb{E}\max_{\sigma\in\mathcal{S}_n} S(\sigma) = n\sqrt{2\log n}(1+o(1))$. Some rather general results of this type were recently obtained [5, 9].

Not much is known for the assignment problem in the discrete setting. We mention the case of i.i.d. Poisson random variables studied in [9], and [14], which considered uniform distributions on $[1..n]$, or on $[1..n^2]$, random permutations of $[1..n]$ for each row, and those of $[1..n^2]$ for the whole matrix.

In this article, we study (1.1) for random matrices $X = (X_{ij})_{1\le i,j\le n}$ with a joint *multinomial* distribution of entries. Recall that a multinomial distribution $\mathcal{M}(m, k)$ is the distribution of a random vector $(Y_j)_{j\le k}$ where Y_j records the number of times side j has been rolled on a fair k-sided die independently rolled m times.

Therefore, in our case, the joint law of matrix entries is $\mathcal{M}(m, n^2)$; the entries are integer-valued, negatively dependent random variables with common binomial distribution $\mathcal{B}(m, p)$ with success probability $p = n^{-2}$ and number of trials m.

We allow the dependence $m = m(n)$. As we will see, the presence of this extra parameter m creates space for a variety of asymptotic behaviours for the expectation of the extrema.

1.2. A motivating example

Let us consider an example showing how the problem studied here emerges in information transmission. Let $\mathcal{A} = (a_1, \ldots, a_n)$ be an alphabet of n letters. If u and v are two independent uniformly distributed words of length m, the $n\times n$ matrix X defined by $X_{ij} := \sum_{k=1}^m \mathbf{1}_{\{u_k=a_i, v_k=a_j\}}$, $1\le i,j\le n$, is distributed according to the multinomial law $\mathcal{M}(m, n^2)$. Recall that the Hamming distance between the words is defined by

$$d_{\mathrm{H}}(u, v) := \sum_{k=1}^m \mathbf{1}_{\{u_k\neq v_k\}} = m - \sum_{k=1}^m \mathbf{1}_{\{u_k=v_k\}} = m - \sum_{i=1}^n X_{ii}.$$

Assume that we have received a word v through a noisy channel and we have to decide whether v is just a random word or a word u that passed through an unknown coding $\sigma : \mathcal{A}\mapsto\mathcal{A}$. The answer should clearly depend on the quantity

$$\min_\sigma d_{\mathrm{H}}(\sigma(u), v) = \min_\sigma (m - S(\sigma)) = m - \max_\sigma S(\sigma).$$

1.3. Results

Our setting is an asymptotic one, i.e. we let $n\to\infty$ and allow $m = m_n$ to be a function of n. The results depend heavily on the relation between n and m. Therefore, we consider separately several zones gradually going down from large m to smaller ones. Everywhere we use the notation $p = p_n := n^{-2}$ for the probability, which is naturally related to our basic multinomial law $\mathcal{M}(m, n^2)$.

1.3.1. *Quasi-Gaussian zone* This zone is defined by assumption

$$\frac{mp}{\log n}\to\infty, \tag{1.2}$$

which essentially means that all entries X_{ij} are sufficiently large to be heuristically approximated with Gaussian variables.

Theorem 1.1. *Under assumption* (1.2), $\mathbb{E}\max_\sigma S(\sigma) = (m/n)(1+o(1))$ *and* $\mathbb{E}\min_\sigma S(\sigma) = (m/n)(1+o(1))$.

Proposition 1.1. *Under assumption* (1.2), $(n/m)\max_\sigma S(\sigma) \xrightarrow{\mathbb{P}} 1$ *and* $(n/m)\min_\sigma S(\sigma) \xrightarrow{\mathbb{P}} 1$.

Remark 1.1. The result of Theorem 1.1 can be compared with the fact that the expectation of the sum for a random permutation is m/n. The technique used to prove convergence in probability of the rescaled maximum and minimum does not apply for the cases we consider below. It is an interesting open question whether such concentration results hold in the cases hereafter.

1.3.2. *Critical zone* The critical zone is described by assumption

$$\frac{mp}{\log n} \to c \tag{1.3}$$

with some $c > 0$. Unlike the quasi-Gaussian case, the expectation behaviour of maxima and minima is not the same.

Theorem 1.2. *Under assumption* (1.3) *for all* $c > 0$, $\mathbb{E}\max_\sigma S(\sigma) = cH_* n \log n(1+o(1))$, *where* $H_* = H_*(c)$ *is the unique solution of* $H\log H - (H-1) = 1/c$, $H > 1$, *and, for all* $c > 1$, $\mathbb{E}\min_\sigma S(\sigma) = c\widetilde{H}_* n \log n(1+o(1))$, *where* $\widetilde{H}_* = \widetilde{H}_*(c)$ *is the unique solution of* $H\log H - (H-1) = 1/c$, $0 < H < 1$.

For $c < 1$ the latter equation has no solution and the result for the minimum is completely different, as stated in the next theorem.

Theorem 1.3. *Let* $c < 1$ *and*

$$\limsup \frac{mp}{\log n} \le c. \tag{1.4}$$

Then, $\lim \mathbb{P}(\min_\sigma S(\sigma) = 0) = 1$.

Remark 1.2. The intermediate case $c = 1$ admits a similar treatment, but the result is less attractive. For example, we can replace assumption (1.4) with

$$\frac{mp}{\log n} \le 1 - \frac{\log(b\log n)}{\log n}, \qquad b > 1.$$

1.3.3. *Quasi-Poissonian zone* The quasi-Poissonian zone is described by the assumptions

$$\frac{mp}{\log n} \to 0 \tag{1.5}$$

while, for every $\delta > 0$,

$$mp \gg n^{-\delta}. \tag{1.6}$$

In this zone all entries X_{ij} are well approximated by Poissonian variables with intensity parameter mp. This zone includes moderately growing intensities mp, constant mp, and even a narrow zone of mp slowly decreasing to zero, e.g. with logarithmic speed.

Theorem 1.4. *Under assumptions* (1.5) *and* (1.6),

$$\mathbb{E}\max_\sigma S(\sigma) = \frac{n\log n}{\log((\log n)/mp)}(1+o(1)).$$

Remark 1.3. Note that if $\log(mp) \ll \log\log n$ we obtain the asymptotics $(n\log n)/\log\log n$ as in the Poisson i.i.d. case with constant intensity [9].

https://doi.org/10.1017/jpr.2023.26 Published online by Cambridge University Press

1.3.4. *Rather sparse matrices* In this zone, we go below (1.6) and assume that

$$mp = cn^{-a}(1 + o(1)), \qquad a \in (0, 1). \tag{1.7}$$

Consider first a regular case.

Theorem 1.5. *Assume that* (1.7) *holds and* $a \notin \{1/k, \ k \in \mathbb{N}\}$. *Then there exists a unique positive integer k such that*

$$\frac{1}{k+1} < a < \frac{1}{k} \tag{1.8}$$

and $\mathbb{E}\ \max_\sigma S(\sigma) = kn(1 + o(1))$.

Let us now briefly discuss the irregular case $a = 1/k$ for some integer $k \geq 2$. Since the lower bound $a > 1/(k+1)$ is still true, we can again obtain $\mathbb{E}\ \max_\sigma S(\sigma) \leq kn(1 + o(1))$. However, the opposite bound breaks down and we are only able to prove that $\mathbb{E}\ \max_\sigma S(\sigma) \geq (k-1)n(1 + o(1))$. To summarise, for the assignment process, we have in this case that

$$(k-1)n(1 + o(1)) \leq \mathbb{E}\ \max_\sigma S(\sigma) \leq kn(1 + o(1))$$

and conjecture that $\mathbb{E}\ \max_\sigma S(\sigma) = (k - \kappa)n(1 + o(1))$ for some $\kappa \in [0, 1]$ depending on a and c. Proving this and finding κ is beyond the reach of current techniques.

Very sparse matrices This zone is determined by

$$1 \ll m \ll n. \tag{1.9}$$

Notice that $m \approx n$ is equivalent to $mp \approx n^{-1}$, and thus the current zone is just below the previous one.

Theorem 1.6. *Under assumption* (1.9), $\mathbb{E}\ \max_\sigma S(\sigma) = m(1 + o(1))$.

2. Proofs

Proof of Theorem 1.1. Let X be a $\mathcal{B}(m, p)$-distributed random variable. Then

$$\mathbb{E}\ \exp(\gamma X) = (1 + p(e^\gamma - 1))^m, \qquad \gamma \in \mathbb{R}. \tag{2.1}$$

Now let X_j, $1 \leq j \leq n$, be $\mathcal{B}(m, p)$-distributed random variables. We do not assume any independence. Then, for every $\gamma > 0$,

$$\mathbb{E}\ \exp\left(\gamma \max_{1\leq j\leq n} X_j\right) \leq \mathbb{E} \sum_{j=1}^{n} \exp(\gamma X_j) = n(1 + p(e^\gamma - 1))^m. \tag{2.2}$$

By Jensen's inequality,

$$\exp\left(\gamma \mathbb{E} \max_{1\leq j\leq n} X_j\right) \leq \mathbb{E}\ \exp(\gamma \max_{1\leq j\leq n} X_j) \leq n(1 + p(e^\gamma - 1))^m.$$

It follows that

$$\mathbb{E} \max_{1\leq j\leq n} X_j \leq \gamma^{-1}(\log n + m \log(1 + p(e^\gamma - 1))) \leq \gamma^{-1}(\log n + mp(e^\gamma - 1)).$$

https://doi.org/10.1017/jpr.2023.26 Published online by Cambridge University Press

We choose $\gamma := ((2\log n)/mp)^{1/2}$. By (1.2), $\gamma \to 0$. Using the expansion $\mathrm{e}^{\gamma} - 1 = \gamma + \gamma^2(1+o(1))/2$, we obtain

$$\begin{aligned}\mathbb{E}\max_{1\le j\le n} X_j &\le \gamma^{-1}(\log n + mp[\gamma + \gamma^2(1+o(1))/2]) \\ &= mp + \gamma^{-1}\log n + mp\,\gamma(1+o(1))/2 \\ &= mp + (2mp\log n)^{1/2}(1+o(1)).\end{aligned}$$

Furthermore, by (1.2) the second term is negligible and we obtain $\mathbb{E}\max_{1\le j\le n} X_j \le mp(1+o(1))$.

The same approach applies to the minima. With the same notation we have, for every $\gamma > 0$,

$$\mathbb{E}\exp(-\gamma\min_{1\le j\le n} X_j) \le \mathbb{E}\sum_{j=1}^{n}\exp(-\gamma X_j) = n(1+p(\mathrm{e}^{-\gamma}-1))^m.$$

By Jensen's inequality,

$$\exp\left(-\gamma\mathbb{E}\min_{1\le j\le n} X_j\right) \le \mathbb{E}\exp\left(-\gamma\min_{1\le j\le n} X_j\right) \le n(1+p(\mathrm{e}^{-\gamma}-1))^m.$$

It follows that $\mathbb{E}\min_{1\le j\le n} X_j \ge -\gamma^{-1}(\log n + m\log(1+p(\mathrm{e}^{-\gamma}-1)))$. We still use $\gamma := ((2\log n)/mp)^{1/2} \to 0$. The expansion $\mathrm{e}^{-\gamma} - 1 = -\gamma + \gamma^2(1+o(1))/2$ yields

$$\log(1+p(\mathrm{e}^{-\gamma}-1)) = p(\mathrm{e}^{-\gamma}-1)(1+o(1)) = -p\gamma(1+o(1)) + p\gamma^2(1+o(1))/2.$$

From this we get

$$\begin{aligned}\mathbb{E}\min_{1\le j\le n} X_j &\ge -\gamma^{-1}(\log n + mp[-\gamma(1+o(1)) + \gamma^2(1+o(1))/2]) \\ &= mp(1+o(1)) - \gamma^{-1}\log n - mp\gamma(1+o(1))/2 \\ &= mp(1+o(1)) - (2mp\log n)^{1/2}(1+o(1)).\end{aligned}$$

By (1.2), the second term is negligible and we obtain $\mathbb{E}\min_{1\le j\le n} X_j \ge mp(1+o(1))$.

We now apply these results to the multinomial assignment process. Here, the joint law of the entries X_{ij} is $\mathcal{M}(m, n^2)$ and every X_{ij} follows the binomial law $\mathcal{B}(m, p)$ with $p = n^{-2}$. Our bound for the maxima yields

$$\mathbb{E}\max_{\sigma} S(\sigma) \le \sum_{i=1}^{n}\mathbb{E}\max_{1\le j\le n} X_{ij} = n\cdot\mathbb{E}\max_{1\le j\le n} X_{1j} \le \frac{m}{n}(1+o(1)),$$

while the bound for the minima yields

$$\mathbb{E}\min_{\sigma} S(\sigma) \ge \sum_{i=1}^{n}\mathbb{E}\min_{1\le j\le n} X_{ij} = n\cdot\mathbb{E}\min_{1\le j\le n} X_{1j} \ge \frac{m}{n}(1+o(1)).$$

It follows that $\mathbb{E}\max_{\sigma} S(\sigma) = (m/n)(1+o(1))$ and $\mathbb{E}\min_{\sigma} S(\sigma) = (m/n)(1+o(1))$, as required. □

Proof of Proposition 1.1. The claim will follow by squeezing once we prove that, for all $\varepsilon > 0$, $\mathbb{P}\big(\max_\sigma S(\sigma) \geq (1+\varepsilon)m/n\big) \to 0$ and $\mathbb{P}\big(\min_\sigma S(\sigma) \leq (1-\varepsilon)m/n\big) \to 0$.

Again, let X_j, $1 \leq j \leq n$, be $\mathcal{B}(m, p)$-distributed random variables. Note that, by (2.2), for $\gamma \to 0^+$,

$$\mathbb{E}\ \exp\left\{\gamma \max_{1\leq j\leq n} X_j\right\} \leq \exp\{\log n + mp\gamma(1+o(1))\}.$$

Using Markov's inequality, we get

$$\mathbb{P}\Big(\max_{1\leq j\leq n} X_j \geq (1+\varepsilon)mp\Big) \leq \exp\{\log n + mp\gamma(o(1)-\varepsilon)\},$$

in which we set $\gamma = (\log n/mp)^{1/2}$ to get

$$\mathbb{P}\Big(\max_{1\leq j\leq n} X_j \geq (1+\varepsilon)mp\Big) \leq \exp\{\log n - (\log n\, mp)^{1/2}(o(1)+\varepsilon)\}.$$

By (1.2), there exists some n_0 such that, for all $n \geq n_0$, $mp \geq 10 \log n/\varepsilon^2$, and thus

$$\mathbb{P}\Big(\max_{1\leq j\leq n} X_j \geq (1+\varepsilon)mp\Big) \leq n^{-2}$$

for sufficiently large n. We can then conclude that

$$\begin{aligned}\mathbb{P}\left(\max_\sigma S(\sigma) \geq (1+\varepsilon)\frac{m}{n}\right) &\leq \mathbb{P}\left(\max_{1\leq i\leq n}\max_{1\leq j\leq n} X_{ij} \geq (1+\varepsilon)\frac{m}{n^2} = (1+\varepsilon)mp\right)\\ &\leq \sum_{i=1}^{n} \mathbb{P}\Big(\max_{1\leq j\leq n} X_{ij} \geq (1+\varepsilon)mp\Big) \to 0.\end{aligned}$$

The claim for the minimum follows along the same lines. □

Before turning to the proof of Theorem 1.2, we state and prove the following lemma.

Lemma 2.1. *Let $(X_j)_{j=1}^{\eta}$ be negatively associated random variables following the binomial law $\mathcal{B}(m, p)$. Assume that $\eta \to \infty$ and the parameters $m = m(\eta)$ and $p = p(\eta)$ are such that*

$$\frac{mp}{\log \eta} \to c > 0 \quad \textit{as } \eta \to \infty, \tag{2.3}$$

$$p\ \log \eta \to 0 \quad \textit{as } \eta \to \infty. \tag{2.4}$$

Then

$$\mathbb{E}\max_{1\leq j\leq \eta} X_j = cH_* \log \eta(1+o(1)) \qquad \textit{as } \eta \to \infty. \tag{2.5}$$

Further, for every $c > 1$,

$$\mathbb{E}\min_{1\leq j\leq \eta} X_j = c\widetilde{H}_* \log \eta(1+o(1)) \qquad \textit{as } \eta \to \infty. \tag{2.6}$$

Proof. The proof is split into two blocks dealing with the lower and upper bounds required to establish both claims.

For the upper bound in (2.5) and the lower bound in (2.6), let $H > H_*$. Then

$$H \log H - (H-1) > \frac{1}{c}. \tag{2.7}$$

Let $r := Hp$. Then, by (2.3), $mr = Hmp = cH \log \eta(1+o(1))$.

Applying the exponential Chebyshev inequality for every j and every $v > 0$, we obtain

$$\begin{aligned}\mathbb{P}(X_j \geq cH \log \eta + v) &= \mathbb{P}(X_j \geq mr(1+o(1))^{-1} + v) \\ &\leq \frac{\mathbb{E}\, \mathrm{e}^{\gamma X_j}}{\exp\{\gamma mr/(1+o(1)+v)\}} = \left[\frac{1+p(\mathrm{e}^{\gamma}-1)}{\exp\{\gamma r/(1+o(1))\}}\right]^m \mathrm{e}^{-\gamma v}. \end{aligned} \tag{2.8}$$

By choosing the optimal $\gamma := \log([(1-p)r]/[p(1-r)])$ and setting $r^{\mathrm{o}} := r/(1+o(1))$, we have

$$\begin{aligned}\frac{1+p(\mathrm{e}^{\gamma}-1)}{\mathrm{e}^{\gamma r}} &= \left(\frac{p}{r}\right)^{r/(1+o(1))} \exp((1-r^{\mathrm{o}})\log(1-p) - (1-r^{\mathrm{o}})\log(1-r)) \\ &= H^{-Hp/(1+o(1))} \exp(-p + r + O(p^2)) \\ &= \exp(-((1+o(1))H \log H - (H-1))p + O(p^2)).\end{aligned}$$

Hence,

$$\begin{aligned}\left[\frac{1+p(\mathrm{e}^{\gamma}-1)}{\mathrm{e}^{\gamma r}}\right]^m &= \exp(-(H\log H - (H-1) + o(1))mp) \\ &= \exp(-(H \log H - (H-1))c(\log \eta)(1+o(1))) := \eta^{-\beta+o(1)},\end{aligned}$$

where, by (2.7), $\beta = (H \log H - (H-1))c > 1$.

Substituting the above results in (2.8) we obtain $\mathbb{P}(X_j \geq cH \log \eta + v) \leq \eta^{-\beta+o(1)}\mathrm{e}^{-\gamma v}$. It is now trivial that

$$\mathbb{P}\Big(\max_{1\leq j\leq \eta} X_j \geq cH \log \eta + v\Big) \leq \eta^{-(\beta-1)+o(1)}\mathrm{e}^{-\gamma v}.$$

It follows that

$$\begin{aligned}\mathbb{E} \max_{1\leq j\leq\eta} X_j - cH\log\eta &= \mathbb{E}\left(\max_{1\leq j\leq\eta} X_j - cH\log\eta\right) \\ &\leq \mathbb{E}\left(\max_{1\leq j\leq\eta} X_j - cH\log\eta\right)_+ \\ &= \int_0^\infty \mathbb{P}\Big(\max_{1\leq j\leq\eta} X_j \geq cH\log\eta + v\Big)\,\mathrm{d}v \\ &\leq \eta^{-(\beta-1)+o(1)} \int_0^\infty \mathrm{e}^{-\gamma v}\,\mathrm{d}v \\ &= \eta^{-(\beta-1)+o(1)}\frac{1}{\gamma} = \eta^{-(\beta-1)+o(1)}\frac{1}{\log H}(1+o(1)) \to 0.\end{aligned}$$

Therefore, $\mathbb{E}\, \max_{1\leq j\leq\eta} X_j \leq cH\log\eta + o(1)$. By letting $H \searrow H_*$ we obtain the upper bound in (2.5).

https://doi.org/10.1017/jpr.2023.26 Published online by Cambridge University Press

The lower bound in (2.6) is obtained in exactly the same way through the Chebyshev inequality for the lower tails.

We now consider the converse bounds. The lower bound in (2.5) is reached in four steps: we give a Poissonian approximation of binomial laws, then provide a lower bound for this Poissonian approximation. This bound provides a lower bound for the maximum's expectation of *independent* binomial i.i.d. random variables. Finally, using the negative association argument, we reduce the claim to the independence case.

First, let X be a binomial $\mathcal{B}(m, p)$-distributed random variable. Elementary calculations show that Poissonian approximation

$$\mathbb{P}(X=k) = \mathrm{e}^{-mp}\frac{(mp)^k}{k!}(1+o(1))$$

is valid if $p^2m \to 0$, $pk \to 0$, and $k^2/m \to 0$. Indeed, we have

$$\mathbb{P}(X=k) = \mathrm{e}^{-mp}\frac{m!}{k!(m-k)!}p^k(1-p)^{m-k}$$

and, with the limiting behaviour as above,

$$(1-p)^m = \exp(m\log(1-p)) = \exp(-mp + m\,\mathrm{O}(p^2)) = \exp(-mp)(1+\mathrm{o}(1)) \text{ as } p^2m \to 0;$$

$$(1-p)^k = \exp(k\log(1-p)) = \exp(k\,\mathrm{O}(p)) = 1+\mathrm{o}(1) \text{ as } kp \to 0;$$

$$\frac{m!}{(m-k)!} \le m^k;$$

$$\frac{m!}{(m-k)!} \ge m^k\left(1-\frac{k}{m}\right)^k = m^k\exp\left(k\,\mathrm{O}\left(\frac{k}{m}\right)\right) = m^k(1+\mathrm{o}(1)) \text{ as } k^2/m \to 0.$$

Second, let $c > 0$ and $H > 1$. Let $k = \lfloor cH\log\eta + 1\rfloor$ and $\lambda = c(\log\eta)(1+o(1))$.

Then, Stirling's approximation and basic computations yield $\mathrm{e}^{-\lambda}\lambda^k/k! = \eta^{-\beta+o(1)}$, where

$$\beta := c(H\log H - (H-1)). \tag{2.9}$$

Now we combine the results of the two steps. Note that with (2.3), (2.4), and, for $k = cH(\log\eta)(1+o(1))$, all three assumptions of the first step are verified and, with $\lambda = mp$, we obtain $\mathbb{P}(X \ge cH\log\eta) \ge \mathbb{P}(X=k) = \eta^{-\beta+o(1)}$. If $1 < H < H_*$, then $\beta < 1$.

Third, let $(\widetilde{X}_j)_{1\le j\le\eta}$ be independent copies of X. Then

$$\begin{aligned}\mathbb{P}\Big(\max_{1\le j\le\eta}\widetilde{X}_j \le cH\log\eta\Big) &= \mathbb{P}(X \le cH\log\eta)^\eta \le (1-\eta^{-\beta+o(1)})^\eta \\ &\le \exp(-\eta^{1-\beta+o(1)}) \to 0. \end{aligned} \tag{2.10}$$

It follows that

$$\begin{aligned}\mathbb{E}\,\max\tilde{X}_j &\ge \mathbb{E}\left[(\max\tilde{X}_j)\mathbf{1}_{\{\max\tilde{X}_j \ge cH\log\eta\}}\right] \\ &\ge (cH\log\eta)\mathbb{P}(\max\tilde{X}_j \ge cH\log\eta) = cH(\log\eta)(1-o(1)).\end{aligned} \tag{2.11}$$

Fourth, from the disintegration theorem for negatively associated variables [4] (see also [3, Chapter 2, Theorem 2.6, and Lemma 2.2]), we have

$$\mathbb{E}\max_{1\le j\le\eta}X_j \ge \mathbb{E}\max_{1\le j\le\eta}\widetilde{X}_j. \tag{2.12}$$

https://doi.org/10.1017/jpr.2023.26 Published online by Cambridge University Press

Combining this estimate with the result of the third step, for every $H < H_*$ we obtain

$$\mathbb{E} \max_{1\le j\le \eta} X_j \ge cH \log \eta(1+o(1)).$$

Letting $H \nearrow H_*$, we obtain the lower bound in (2.5).

The upper bound in (2.6) follows in a similar way. Now let $k := [cH \log \eta]$. By using Poissonian approximation and Poissonian asymptotics we obtain

$$\mathbb{P}(X \le cH \log \eta) \ge \mathbb{P}(X = k) = \eta^{-\beta+o(1)}$$

with the same β as in (2.9). If $\widetilde{H}_* < H < 1$, then $\beta < 1$.

As before, for independent variables we obtain

$$\mathbb{P}\Big(\min_{1\le j\le \eta} \widetilde{X}_j \ge cH \log \eta\Big) \le \exp\left(-\eta^{1-\beta+o(1)}\right).$$

It follows that

$$\begin{aligned}
\mathbb{E} \min_{1\le j\le \eta} \widetilde{X}_j &= \mathbb{E}\Big[\min_{1\le j\le \eta} \widetilde{X}_j \mathbf{1}_{\{\min_{1\le j\le \eta} \widetilde{X}_j \le cH \log \eta\}}\Big] + \mathbb{E}\Big[\min_{1\le j\le \eta} \widetilde{X}_j \mathbf{1}_{\{\min_{1\le j\le \eta} \widetilde{X}_j > cH \log \eta\}}\Big] \\
&\le cH \log \eta + \sum_{j=1}^{\eta} \mathbb{E}\big[\widetilde{X}_j \mathbf{1}_{\{\min_{1\le i\le \eta, i\ne j} \widetilde{X}_i > cH \log \eta\}}\big] \\
&= cH \log \eta + \eta \mathbb{E}\, \widetilde{X}_1 \mathbb{P}\Big(\min_{2\le i\le \eta} \widetilde{X}_i > cH \log \eta\Big) \\
&\le cH \log \eta + \eta \cdot c \log \eta(1+o(1)) \exp\left(-\eta^{1-\beta+o(1)}\right) \\
&= cH \log \eta + o(1).
\end{aligned}$$

The final negative association argument reads as follows. Since (X_j) are negatively associated, so are $(-X_j)$. From the disintegration theorem cited above, it follows that

$$\mathbb{E} \max_{1\le j\le \eta} (-X_j) \ge \mathbb{E} \max_{1\le j\le \eta} (-\widetilde{X}_j),$$

which is equivalent to $\mathbb{E} \min_{1\le j\le \eta} X_j \le \mathbb{E} \min_{1\le j\le \eta} \widetilde{X}_j$. By combining the obtained results, we have $\mathbb{E} \min_{1\le j\le \eta} X_j \le cH \log \eta(1+o(1))$. Finally, letting $H \searrow \widetilde{H}_*$ we obtain the upper bound in (2.6). □

Proof of Theorem 1.2. Recall that a multinomial distribution is *negatively associated*, see [8] and [3, Chapter 1, Theorem 1.27]. Furthermore, with $p = n^{-2}$, the assumption (2.4) is also valid.

Therefore, the bounds (2.5) and (2.6) apply to the sums of the entries X_{ij}. They yield, respectively,

$$\mathbb{E} \max_{\sigma} S(\sigma) \le \sum_{i=1}^{n} \mathbb{E} \max_{1\le j\le n} X_{ij} \le cH_* n \log n(1+o(1)),$$

$$\mathbb{E} \min_{\sigma} S(\sigma) \ge \sum_{i=1}^{n} \mathbb{E} \min_{1\le j\le n} X_{ij} \ge c\widetilde{H}_* n \log n(1+o(1)).$$

The opposite bounds follow by the 'greedy method' dating back to [15], for instance, and introduced for Gaussian random variables in [12] (and used in [9]) that we recall now. This method allows the construction of a quasi-optimal permutation σ^* that provides a sufficiently large value or sufficiently small value of the assignment process. Recall that $[1..i] := \{1, 2, \ldots, i\}$. Define $\sigma^*(1) := \arg\max_{j\in[1..n]} X_{1j}$, and let, for all $i = 2, \ldots, n$, $\sigma^*(i) := \arg\max_{j\notin\sigma^*([1..i-1])} X_{ij}$. It is natural to call this strategy 'greedy', because at every step we consider row i, take the maximum of its available elements (without considering the influence of this choice on subsequent steps), and then forget row i and the corresponding column $\sigma^*(i)$. The number of variables used at consequent steps is decreasing from n to 1.

By using the greedy method, we have

$$\mathbb{E}\max_{\sigma} S(\sigma) \geq \mathbb{E}\sum_{i=1}^{n} X_{i\,\sigma^*(i)} = \sum_{i=1}^{n}\mathbb{E}\max_{j\notin\sigma^*([1..i-1])} X_{ij} = \sum_{i=1}^{n}\mathbb{E}\max_{1\leq j\leq n-i+1} X_{ij}. \tag{2.13}$$

The latter equality may seem surprising because the index sets $[n]\backslash\sigma^*([1..i-1])$ are random and depend on the matrix X. However, it is justified by the following lemma.

Lemma 2.2. *Let $N_1, N_2 > 0$ be positive integers, and let a random vector $X := (X_j)_{1\leq j\leq N_1+N_2}$ be distributed according to a multinomial law $\mathcal{M}_{m,N_1+N_2}$. Let $X^{(1)} := (X_j)_{1\leq j\leq N_1}$ and $X^{(2)} := (X_j)_{N_1<j\leq N_1+N_2}$. Let $1 \leq q \leq N_2$, and let $\mathcal{J} \subset (N_1, N_1+N_2]$ be a random set of size q determined by $X^{(1)}$. Then the variables $\max_{j\in\mathcal{J}} X_j$ and $\max_{N_1<j\leq N_1+q} X_j$ are equidistributed.*

By applying the asymptotic expression (2.5) to each term of the sum (2.13), and by Stirling's formula, $\sum \log i = \log n! \sim n\log n$, we obtain the desired lower bound: $\mathbb{E}\max_\sigma S(\sigma) \geq cH_* n\log n(1+o(1))$. Replacing maxima by minima in the greedy method and using (2.6) yields the remaining upper bound: $\mathbb{E}\min_\sigma S(\sigma) \leq c\widetilde{H}_* n\log n(1+o(1))$.

This completes the proof of Theorem 1.2 except for the postponed proof of Lemma 2.2. □

Proof of Lemma 2.2. Let $S = S(X^{(1)}) := \sum_{j=1}^{N_1} X_j$. Recall that the conditional distribution of $X^{(2)}$ with respect to $X^{(1)}$ is $\mathcal{M}_{m-S,N_2}$. This means that, for all $x_1 \in \mathbb{N}^{N_1}$ and $x_2 \in \mathbb{N}^{N_2}$,

$$\mathbb{P}(X^{(2)} = x_2, X^{(1)} = x_1) = \mathbb{P}(X^{(1)} = x_1)\mathcal{M}_{m-S(x_1),N_2}(x_2).$$

For every fixed set $J \subset (N_1, N_1+N_2]$ of size q,

$$\mathbb{P}(X^{(2)} = x_2, \mathcal{J} = J) = \sum_{s=0}^{m}\mathbb{P}(\mathcal{J} = J, S = s)\mathcal{M}_{m-s,N_2}(x_2)$$

by summing over $x_1 \in \mathcal{J}^{-1}(J)$, where $\mathcal{J}^{-1}(J)$ is the preimage of the set J under $\mathcal{J}$. Now, for every nonnegative integer μ, by summing over x_2 such that $\max_{j\in J} x_{2j} = \mu$, we obtain

$$\mathbb{P}\Big(\max_{j\in J} X_j = \mu, \mathcal{J} = J\Big) = \sum_{s=0}^{m}\mathbb{P}(\mathcal{J} = J, S = s)\mathcal{M}_{m-s,N_2}\Big(x_2\colon \max_{j\in J} x_{2j} = \mu\Big).$$

The latter factor does not depend on a particular set J due to the exchangeability property of the multinomial law. We may thus write $\mathcal{M}_{m-s,N_2}(x_2\colon \max_{j\in J} x_{2j} = \mu) =: F(m-s, N_2, q, \mu)$ and obtain

$$\mathbb{P}\Big(\max_{j\in J} X_j = \mu, \mathcal{J} = J\Big) = \sum_{s=0}^{m}\mathbb{P}(\mathcal{J} = J, S = s)F(m-s, N_2, q, \mu).$$

https://doi.org/10.1017/jpr.2023.26 Published online by Cambridge University Press

By summing over all sets J of size q, we see that

$$\mathbb{P}\left(\max_{j\in\mathcal{J}} X_j = \mu\right) = \sum_{s=0}^{m} \mathbb{P}(S=s)F(m-s, N_2, q, \mu)$$

does not depend on the specific choice of $\mathcal{J}$, and the claim of the lemma follows. □

Proof of Theorem 1.3. We are going to use an old result from [7] about the existence of perfect matching in a random bipartite graph. Let G be a uniformly distributed $n+n$ bipartite graph with $m = m(n)$ edges. If

$$\lim_{n\to\infty}\left(\frac{m}{n} - \log n\right) = \infty, \tag{2.14}$$

then with probability tending to 1, as $n \to \infty$, G has a perfect matching.

In matrix form, this result asserts the following. Let $Y = Y(n, m) = \{Y_{ij}\}_{1\le i,j\le n}$ be a uniformly distributed random $n \times n$ matrix with entries taking values in $\{0, 1\}$ and satisfying $\sum_{i,j=1}^{n} Y_{ij} = m$. If (2.14) holds, then

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{\sigma} \sum_{i=1}^{n} Y_{i\sigma(i)} = n\right) = 1. \tag{2.15}$$

Now let $X = (X_{ij})$ be our matrix following the multinomial law $\mathcal{M}(m, n^2)$. Introduce the matrix $\widetilde{Y}$ as

$$\widetilde{Y}_{ij} := \begin{cases} 0, & X_{ij} > 0, \\ 1, & X_{ij} = 0. \end{cases}$$

Note that $\mathbb{P}(\widetilde{Y}_{ij} = 1) = \mathbb{P}(X_{ij} = 0) = (1-p)^m = \exp(-mp(1+o(1)))$. Let $T := \sum_{i,j=1}^{n} \widetilde{Y}_{ij}$ be the number of empty cells in our matrix X. Observe that, conditioned on T, the matrix $\widetilde{Y}$ has the same distribution as $Y(n, T)$. Taking into account that the probability in (2.15) is nondecreasing as a function of m, we have, for every positive integer M,

$$\begin{aligned}
\mathbb{P}\left(\min_{\sigma} S(\sigma) = 0\right) &= \mathbb{P}\left(\max_{\sigma} \sum_{i=1}^{n} \widetilde{Y}_{i\sigma(i)} = n\right) \\
&= \sum_{k\ge 0} \mathbb{P}\left(\max_{\sigma} \sum_{i} \tilde{Y}_{i\sigma(i)} = n \mid T = k\right)\mathbb{P}(T=k) \\
&\ge \mathbb{P}\left(\max_{\sigma} \sum_{i=1}^{n} Y(n, M)_{i\sigma(i)} = n\right) \sum_{k\ge M} \mathbb{P}\left(\max_{\sigma} \sum_{i} \tilde{Y}_{i\sigma(i)} = n \mid T = k\right)\mathbb{P}(T=k) \\
&\ge \mathbb{P}(T \ge M)\mathbb{P}\left(\max_{\sigma} \sum_{i=1}^{n} Y(n, M)_{i\sigma(i)} = n\right).
\end{aligned} \tag{2.16}$$

We choose $M = n^\beta$ with $\beta \in (1, 2-c)$ and show that both probabilities in the latter product tend to 1 as $n \to \infty$.

https://doi.org/10.1017/jpr.2023.26 Published online by Cambridge University Press

For the first one, using (1.4), we have $\mathbb{E}\,T = n^2\mathbb{E}\,\widetilde{Y}_{11} = n^2\exp(-mp(1+o(1))) \geq n^{2-c(1+o(1))}$. Furthermore, since the variables $\widetilde{Y}_{ij}$ are negatively correlated, we have $\mathrm{Var}\,T \leq n^2\mathrm{Var}\,\widetilde{Y}_{11} \leq n^2\mathbb{E}\,\widetilde{Y}_{11} = \mathbb{E}\,T$. Finally, using $\beta < 2-c$, by Chebyshev's inequality,

$$\mathbb{P}(T \leq n^\beta) \leq \mathbb{P}(|T - \mathbb{E}\,T| \geq \mathbb{E}\,T - n^\beta) = \mathbb{P}(|T - \mathbb{E}\,T| \geq \mathbb{E}\,T(1+o(1)))$$
$$\leq \frac{\mathrm{Var}\,T}{(\mathbb{E}\,T)^2(1+o(1))} \leq \frac{\mathbb{E}\,T}{(\mathbb{E}\,T)^2(1+o(1))} \to 0.$$

On the other hand, since $\beta > 1$, the assumption (2.14) with $m := M = n^\beta$ is true. Therefore, the second probability in the product (2.16) tends to 1 by the result in [7]. We obtain from (2.16) that $\lim_{n\to\infty}\mathbb{P}(\min_\sigma S(\sigma) = 0) = 1$, which is the desired claim. □

Proof of Theorem 1.4. The proof goes along the same lines as for Theorem 1.2. Instead of the key relation (2.5), we prove the following claim. Let (X_j) be negatively associated random variables following the binomial law $\mathcal{B}(m, p)$. Then, under assumptions (1.5) and (1.6),

$$\mathbb{E}\max_{1\leq j\leq n} X_j = \frac{\log n}{\log((\log n)/mp)}(1+o(1)) \qquad \text{as } n \to \infty. \tag{2.17}$$

For the upper bound in (2.17) that we shall prove now, no lower bound on mp is needed; we only use (1.5).

Let $\beta > 1$, $y := (\beta\log n)/mp$, $r := y/(\log y)$. Notice that under (1.5) we have $y, r \to \infty$. Next, for a binomial $\mathcal{B}(m, p)$ random variable X and for every $v > 0$,

$$\mathbb{P}\left(X \geq \frac{\beta\log n}{\log((\log n)/mp)} + v\right) \leq \mathbb{P}\left(X \geq \frac{\beta\log n}{\log((\beta\log n)/mp)} + v\right)$$
$$= \mathbb{P}\left(X \geq \frac{(\beta\log n)/mp}{\log((\beta\log n)/mp)}mp + v\right)$$
$$= \mathbb{P}\left(X \geq \frac{y}{\log y}mp + v\right) = \mathbb{P}(X \geq rmp + v).$$

In the next calculation we use the Poisson version of the bound for exponential moment, $\mathbb{E}\,\exp(\gamma X) \leq \exp(mp(\mathrm{e}^\gamma - 1))$, that immediately follows from the exact formula (2.1). By applying Chebyshev's inequality with Poisson-optimal parameter $\gamma = \log r$ we obtain

$$\mathbb{P}(X \geq rmp + v) \leq \mathbb{E}\,\exp(\gamma X)\exp(-\gamma(rmp + v))$$
$$\leq \exp(-mp(\gamma r - \mathrm{e}^\gamma + 1) - \gamma v)$$
$$= \exp(-mp(r\log r - r + 1) - \gamma v).$$

Since $r \to \infty$, $r\log r - r + 1 \sim r\log r \sim y = (\beta\log n)/mp$. It follows that

$$\mathbb{P}(X \geq rmp + v) \leq \exp(-\beta\log n(1+o(1)) - \gamma v) = n^{-\beta(1+o(1))}\exp(-\gamma v),$$

$$\mathbb{P}\left(\max_{1\leq j\leq n} X_j \geq rmp + v\right) \leq n\mathbb{P}(X \geq rmp + v) \leq n^{-(\beta-1)(1+o(1))}\exp(-\gamma v).$$

Hence,

$$\mathbb{E}\max_{1\leq j\leq n} X_j \leq rmp + n^{-(\beta-1)(1+o(1))}\int_0^\infty \exp(-\gamma v)\,\mathrm{d}v = rmp + n^{-(\beta-1)(1+o(1))}\gamma^{-1}.$$

Note that $rmp\gamma = r\log r\, mp \sim ymp = \beta \log n \to \infty$, and hence we conclude that $n^{-(1-\beta)(1+o(1))}\gamma^{-1}$ is negligible compared to rmp; thus,

$$\mathbb{E} \max_{1\le j\le n} X_j \le rmp(1+o(1)) \sim \frac{\beta \log n}{\log((\log n)/mp)}$$

and the required upper bound follows by letting $\beta \searrow 1$.

For the lower bound, let $\beta \in (0, 1)$, $y := (\beta \log n)/mp$, $r := y/(\log y)$, and

$$k := rmp = \frac{y}{\log y} mp = \frac{\beta \log n}{\log y}.$$

Assumption (1.5) yields $y \to \infty$, $k = o(\log n)$, $\mathrm{e}^k = n^{o(1)}$, and $\mathrm{e}^{mp} = n^{o(1)}$.

On the other hand, under assumption (1.6) we have $|\log(mp)| \ll \log n$, which yields $\log y \ll \log n$, hence $k \to \infty$.

Rewriting the binomial probability mass function as

$$\mathbb{P}(X=k) = \mathrm{e}^{-mp}\frac{(mp)^k}{k!}\left[\frac{m(m-1)\cdots(m-k+1)}{m^k} \times \left(1-\frac{pm}{m}\right)^{m-k} \mathrm{e}^{pm}\right],$$

and observing that the product between square brackets converges to 1 for m, p, k as above, we can make a Poissonian approximation. Using the latter, we obtain

$$\begin{aligned}\mathbb{P}(X\ge k) \ge \mathbb{P}(X=k) \sim \mathrm{e}^{-mp}\frac{(mp)^k}{k!} &\sim \mathrm{e}^{-mp}\mathrm{e}^k(2\pi k)^{-1/2}\left(\frac{mp}{k}\right)^k \\ &= n^{o(1)}r^{-k} = n^{o(1)}r^{-rmp} = n^{o(1)}\exp(-r\log r\, mp) \\ &= n^{o(1)}\exp(-y(1+o(1))mp) = n^{-\beta+o(1)}.\end{aligned}$$

By repeating the arguments from (2.10), (2.11), and (2.12), we obtain

$$\mathbb{E} \max_{1\le j\le n} X_j \ge k(1+o(1)) = \frac{y}{\log y} mp(1+o(1)) = \frac{\beta \log n}{\log y}(1+o(1));$$

letting $\beta \nearrow 1$ provides the required lower bound in (2.17).

Once (2.17) is proved, the proof of Theorem 1.4 is completed by the same simple arguments (including the greedy method) as for Theorem 1.2. Indeed, combining (2.17) with (2.13) implies that

$$\mathbb{E} \max_{\sigma} S(\sigma) \ge \sum_{i=1}^n \mathbb{E} \max_{1\le j\le n-i+1} X_{ij} \ge \sum_{i=1}^n \frac{\log(n-i+1)}{\log\,[\log(n-i+1)/mp]}(1+o(1)),$$

from which the claim easily follows. □

Proof of Theorem 1.5. We first establish the upper bound. Let X_j, $1 \le j \le n$, be $\mathcal{B}(m, p)$-binomial random variables. We have

$$\begin{aligned}\mathbb{E} \max_{1\le j\le n} X_j &= \mathbb{E}\left[\max_{1\le j\le n} X_j \mathbf{1}_{\{\max_{1\le j\le n} X_j \le k\}}\right] + \mathbb{E}\left[\max_{1\le j\le n} X_j \mathbf{1}_{\{\max_{1\le j\le n} X_j > k\}}\right] \\ &\le k + \sum_{j=1}^n \mathbb{E}\left[X_j \mathbf{1}_{\{X_j>k\}}\right] = k + n\mathbb{E}\left[X_1 \mathbf{1}_{\{X_1>k\}}\right].\end{aligned}$$

Furthermore, since the law of X_1 is $\mathcal{B}(m, p)$,

$$\mathbb{P}(X_1 = \ell) = \frac{m!}{(m-\ell)!}\frac{p^\ell}{\ell!}(1-p)^{m-\ell} \le \frac{m^\ell p^\ell}{\ell!}, \qquad 0 \le \ell \le m.$$

Hence,

$$\mathbb{E}\left[X_1 \mathbf{1}_{\{X_1>k\}}\right] \le \sum_{\ell=k+1}^{\infty} \frac{(mp)^\ell}{(\ell-1)!} = \sum_{q=0}^{\infty} \frac{(mp)^{k+1+q}}{(k+q)!} \le (mp)^{k+1}\mathrm{e}^{mp} = (mp)^{k+1}(1+o(1)).$$

Therefore, $\mathbb{E}\ \max_{1\le j\le n} X_j \le k + c^{k+1}n^{1-a(k+1)}(1+o(1)) = k + o(1)$, where we used the lower bound in (1.8) at the last step.

It follows immediately that

$$\mathbb{E}\ \max_{\sigma} S(\sigma) \le nk(1+\mathrm{O}(1)). \tag{2.18}$$

Turning to the lower bound, for every positive integer v in the *independent case*, we have

$$\begin{aligned}\mathbb{P}\Big(\max_{1\le j\le v} X_j < k\Big) = \mathbb{P}(X_1 < k)^v &= (1-\mathbb{P}(X_1 \ge k))^v \\ &\le (1-\mathbb{P}(X_1 = k))^v \\ &\le \exp\{-v\mathbb{P}(X_1 = k)\} \\ &= \exp\left\{-v\frac{c^k n^{-ak}}{k!}(1+o(1))\right\},\end{aligned}$$

where the second last inequality follows from the inequality $1 + x \le \exp x$, and the last one follows from

$$\mathbb{P}(X_1 = k) = \frac{(mp)^k}{k!}\left[\frac{m(m-1)\cdots(m-k+1)}{m^k}(1-p)^{m-k}\right] = \frac{(mp)^k}{k!}(1+o(1))$$

applied to this case; recall (1.7). Let us choose some arbitrary $\delta \in (0, 1)$. By letting $v = [\delta n]$ and using the upper bound in (1.8) we obtain $\mathbb{P}\big(\max_{1\le j\le[\delta n]} X_j < k\big) \to 0$. It follows that

$$\mathbb{E}\ \max_{1\le j\le[\delta n]} X_j \ge k\mathbb{P}\Big(\max_{1\le j\le[\delta n]} X_j \ge k\Big) = k(1+o(1)).$$

By using the negative association argument (2.12), we also obtain $\mathbb{E}\ \max_{1\le j\le[\delta n]} X_j \ge k(1 + o(1))$ in the multinomial setting. Combining this with the greedy method (2.13) yields

$$\mathbb{E}\ \max_{\sigma} S(\sigma) \ge \sum_{i=1}^{n} \mathbb{E}\ \max_{1\le j\le n-i+1} X_{ij} \ge \sum_{i=1}^{n-[\delta n]+1} k(1+o(1)) = (1-\delta)nk(1+o(1)).$$

By letting $\delta \searrow 0$ we obtain $\mathbb{E}\ \max_\sigma S(\sigma) \ge nk(1+o(1))$. We then conclude from this and (2.18) that $\mathbb{E}\ \max_\sigma S(\sigma) = kn(1+o(1))$. □

Proof of Theorem 1.6. The upper bound $\max_\sigma S(\sigma) \le m$ is trivial; it remains to prove the lower bound.

https://doi.org/10.1017/jpr.2023.26 Published online by Cambridge University Press

Let us denote by $(u_i, v_i)_{1\le i\le m}$ the coordinates of the particles thrown on the square table. All u_i and all v_i are i.i.d. random variables uniformly distributed on the integers $[1..n]$. Let $U_0 = V_0 = \emptyset$, $U_k := \{u_i, 1 \le i \le k\}$, $V_k := \{v_i, 1 \le i \le k\}$, $1 \le k \le m$, and introduce the events $A_k := \{u_k \notin U_{k-1}, v_k \notin V_{k-1}\}$, $1 \le k \le m$. Note that the sequence of events A_k describes the possibility of constructing a matching of size k iteratively by keeping track of the rows and columns already used in the construction of a permutation matrix and 'locking' them. This understanding underlies the validity of (2.19) below. It is obvious that, for each k, $\mathbb{P}(A_k) \ge 1 - 2m/n$; hence, by $m \ll n$,

$$\mathbb{E}\left(\sum_{k=1}^{m} \mathbf{1}_{\{A_k\}}\right) \ge m\left(1 - \frac{2m}{n}\right) = m(1 + o(1)).$$

On the other hand, we have

$$\max_{\sigma} S(\sigma) \ge \sum_{k=1}^{m} \mathbf{1}_{\{A_k\}}, \tag{2.19}$$

which entails the desired $\mathbb{E}\ \max_\sigma S(\sigma) \ge m(1 + o(1))$. □

Acknowledgements

We wish to thank the Editor and two anonymous reviewers for their comments and suggestions.

Funding information

The work of M. Lifshits was supported by RSF grant 21-11-00047. G. Mordant gratefully acknowledges the support of the DFG within SFB 1456.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ALDOUS, D. (1992) Asymptotics in the random assignment problem. *Prob. Theory Relat. Fields* **93**, 507–534.
[2] ALDOUS, D. J. (2001) The $\zeta(2)$ limit in the random assignment problem. *Random Structures* **18**, 381–418.
[3] BULINSKI, A. V. AND SHASHKIN A. P. (2007) *Limit Theorems for Associated Random Fields and Related Systems* (Adv. Ser. Statist. Sci. Appl. Prob. **10**). World Scientific, Singapore.
[4] CHRISTOFIDES, T. C. AND VAGGELATOU, E. (2004) A connection between supermodular ordering and positive/negative association. *J. Multivar. Anal.* **88**, 138–151.
[5] CHENG, Y., LIU, Y., TKOCZ, T. AND XU A. (2023) Typical values of extremal-weight combinatorial structures with independent symmetric weights. *Electron. J. Combinatorics* **30**, P1.12.
[6] COPPERSMITH, D. AND SORKIN, G. B. (1999) Constructive bounds and exact expectations for the random assignment problem. *Random Structures Algorithms* **15**, 113–144.
[7] ERDÖS, P. AND RÉNYI, A. (1964) On random matrices. *Publ. Math. Inst. Hungar. Acad. Sci.* **8**, 455–461.
[8] JOAG-DEV, K. AND PROSCHAN, F. (1983) Negative association of random variables with applications. *Ann. Statist.* **11**, 286–295.
[9] LIFSHITS, M. AND TADEVOSIAN, A. (2022) On the maximum of random assignment process. *Statist. Prob. Lett.* **187**, 109530.
[10] LINUSSON, S. AND WÄSTLUND, J. (2004) A proof of Parisi's conjecture on the random assignment problem. *Prob. Theory Relat. Fields* **128**, 419–440.
[11] MÉZARD, M. AND PARISI, G. (1987) On the solution of the random link matching problems. *J. Physique* **48**, 1451–1459.

https://doi.org/10.1017/jpr.2023.26 Published online by Cambridge University Press

[12] MORDANT, G. AND SEGERS, J. (2021) Maxima and near-maxima of a Gaussian random assignment field. *Statist. Prob. Lett.* **173**, 109087.
[13] NAIR, C., PRABHAKAR, B. AND SHARMA, M. (2005) Proofs of the Parisi and Coppersmith–Sorkin random assignment conjectures. *Random Structures Algorithms* **27**, 413–444.
[14] PARVIAINEN, R. (2004) Random assignment with integer costs. *Combinatorics Prob. Comput.* **13**, 103–113.
[15] STEELE, J. M. (1990). Probability and statistics in the service of computer science: Illustrations using the assignment problem. *Commun. Statist. Theory Meth.* **19**, 4315–4329.
[16] STEELE, J. M. (1997). *Probability Theory and Combinatorial Optimization* (CBMS-NSF Regional Conf. Ser. Appl. Math. **69**). SIAM, Philadelphia, PA.
[17] WÄSTLUND, J. (2009) An easy proof of the $\zeta(2)$ limit in the random assignment problem. *Electron. Commun. Prob.* **14**, 261–269.