Typical Values of Extremal-Weight Combinatorial Structures with Independent Symmetric Weights

Yun Cheng   Yixue Liu   Tomasz Tkocz*   Albert Xu

Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213, U.S.A.

`ttkocz@andrew.cmu.edu`

Abstract

Suppose that the edges of a complete graph are assigned weights independently at random and we ask for the weight of the minimal-weight spanning tree, or perfect matching, or Hamiltonian cycle. For these and several other common optimisation problems, we establish asymptotically tight bounds when the weights are independent copies of a symmetric random variable (satisfying a mild condition on tail probabilities), in particular when the weights are Gaussian.

Mathematics Subject Classifications: 05C80, 90C27, 60C05

1 Introduction

Classical optimisation problems such as the minimum spanning tree, the assignment problem or the shortest path have been extensively studied in the worst case as well as in the average case. For the latter, we usually consider a complete graph on n vertices with each edge having assigned independently at random a nonnegative weight and ask for typical values (in terms of the expectation or high probability bounds) of the weight of the minimum-weight combinatorial structure such as a spanning tree, a perfect matching or a path between two fixed vertices. For example, in the case of the exponential rate 1 weights, Frieze's $\zeta(3)$-result from [18] says that the expected weight of the minimum spanning tree is asymptotic to $\zeta(3)$ as $n \to \infty$, for the assignment problem, resolving Parisi's conjecture from [44], Aldous in [1] showed $\zeta(2)$ to be the asymptotic value (see also [8, 40, 43, 45]), whereas the weight of the shortest path is asymptotic to $\frac{\log n}{n}$, as showed independently by David and Prieditis in [9] and Janson in [31]. Moreover, many limit theorems have

---

also been established (for instance, see [3, 20, 30]), large deviation regimes studied (for instance, see [16, 41]) and various refinements, extensions and constrained versions have been investigated (for instance, see [2, 7, 12, 17, 21, 22, 23, 24, 25, 26, 27, 28, 32, 35, 36, 39]).

In this short note, we consider the case of weights drawn from symmetric distributions. The assignment problem with Gaussian weights has been recently studied in [37, 42] and with general weights in [38]. We identify a mild condition on tails, viz. Chernoff's bound being asymptotically optimal, notably satisfied in the Gaussian case. Under this condition, we are able to find asymptotically tight high probability estimates for the aforementioned optimisation problems (and several others).

2   Results

To motivate our main definition and state our results, we first recall necessary notions and facts. For a random variable X, let $\Lambda\colon \mathbb{R} \to (-\infty, +\infty]$,

$$\Lambda(t) = \log \mathbb{E}e^{tX}, \qquad t \in \mathbb{R},$$

be its log-moment generating function with the Legendre transform (the rate function of X),

$$\Lambda_*(t) = \sup_{s \in \mathbb{R}}\{st - \Lambda(s)\}, \qquad t \in \mathbb{R}.$$

Recall that by Markov's inequality, for every t,

$$\mathbb{P}\left(X > t\right) \leqslant \exp\left(-\sup_{s>0}\{st - \Lambda(s)\}\right).$$

If $\Lambda(t_0) < \infty$ for some $t_0 > 0$, then $\mathbb{E}X \in [-\infty, \infty)$ and for $t > \mathbb{E}X$, in fact we have $\sup_{s>0}\{st - \Lambda(s)\} = \Lambda_*(t)$ (for instance, see Lemma 2.2.5 in [10]). This is then sometimes referred to as Chernoff's inequality,

$$\mathbb{P}\left(X > t\right) \leqslant \exp\left(-\Lambda_*(t)\right), \qquad t > \mathbb{E}X. \tag{1}$$

We say that X has *regular upper tails* if this upper bound is asymptotically tight in the following sense:

$$\mathbb{P}\left(X > t\right) = \exp\left(-(1 + o(1))\Lambda_*(t)\right), \qquad \text{as } t \to +\infty. \tag{2}$$

Note that when X is bounded above, $\mathbb{P}\left(X > t\right) = 0$ and $\Lambda_*(t) = +\infty$ for all $t > \operatorname{ess\,sup}X$, so in this case, condition (2) is vacuously satisfied and X has regular upper tails. Examples with regular upper tails include Gaussian, exponential, gamma, Poisson random variables. On the other hand, it is not difficult to construct random variables without regular upper tails (see Section 4).

For the purposes of this note, we say that a random variable X is *good*, if it is symmetric (meaning $-X$ has the same distribution as X), $\Lambda(t) < \infty$ for all $|t| < \delta$ for some $\delta > 0$ and X has regular upper tails. In particular, all symmetric bounded random variables are

good. If X is good and unbounded, then $\Lambda_* < \infty$ and Λ_* is even, convex strictly increasing on $(0, +\infty)$ with $\Lambda_*(0) = 0$, so that the inverse $\Lambda_*^{-1}\colon [0, \infty) \to [0, \infty)$ of $\Lambda_*|_{[0,\infty)}$ is well-defined. In general, for $t \geqslant 0$, we set $\Lambda_*^{-1}(t) = \inf\{s, \Lambda_*(s) \geqslant t\}$, the usual generalised inverse of Λ_*. For instance, in the degenerate case, when $X = 0$ a.s., $\Lambda \equiv 0$, and $\Lambda_*(0) = 0$ and $\Lambda_* = +\infty$ elsewhere, so $\Lambda_*^{-1} \equiv 0$.

Examples of good random variables include of course standard Gaussian or two-sided exponential.

Our main results provide asymptotically tight high probability bounds on extremal-weight common combinatorial structures (perfect matchings, spanning trees, Hamilton cycles, paths between two fixed vertices, copies of a fixed graph) in complete graphs with edge-weights being i.i.d. copies of a good random variable. Thanks to symmetry, the distribution of the minimum is the same as of negative the maximum and thus we shall only focus on the latter.

Theorem 1. *Let $K_{n,n} = ([n], [n], [n] \times [n])$ be the complete bipartite graph with each edge e assigned an independent copy X_e of a good random variable X with rate function Λ_*. Let $\mathscr{C}_n$ be the set of perfect matchings in $K_{n,n}$ and let W_n be the weight of an optimal matching,*

$$W_n = \max\left\{ \sum_{e \in E(M)} X_e,\ M \in \mathscr{C}_n \right\}.$$

Then,

$$W_n = (1 + o(1)) n \Lambda_*^{-1}(\log n), \qquad w.h.p.[1]$$

Recall that the density of a graph $H = (V, E)$ is $d(H) = |E|/|V|$ and the graph H is called *balanced* if its density is not smaller than the density of any of its subgraphs, that is $\max d(H') = d(H)$ where the maximum is over all subgraphs H' of H.

Theorem 2. *Let $K_n = ([n], \binom{[n]}{2})$ be the complete (undirected) graph with each edge e assigned an independent copy X_e of a good random variable X with rate function Λ_*. Given a set $\mathscr{C}_n$ of subgraphs of K_n, we let W_n be the weight of an optimal one,*

$$W_n = \max\left\{ \sum_{e \in E(H)} X_e,\ H \in \mathscr{C}_n \right\}.$$

In each of the following cases

(a) $\mathscr{C}_n$ is the set of all spanning trees of K_n,

(b) $\mathscr{C}_n$ is the set of all Hamilton cycles in K_n,

(c) $\mathscr{C}_n$ is the set of all paths from vertex 1 to 2 in K_n,

---

[1]with high probability, that is with probability tending to 1 as $n \to \infty$

we have

$$W_n = (1 + o(1))n\Lambda_*^{-1}(\log n), \qquad w.h.p.$$

(d) If $\mathscr{C}_n$ is the set of all copies in K_n of a fixed balanced graph H_0 with ℓ edges and density d, then

$$W_n = (1 + o(1))\ell\Lambda_*^{-1}\left(d^{-1}\log n\right), \qquad w.h.p.$$

In the vast body of works mentioned in passing in the introduction, where the weights are nonnegative, it is natural to think of them as cost and those optimisation problems provide the size of the cheapest structure. Here, when the weights have a symmetric distribution, the paradigm is different and we may think of the random weight X_e assigned to an edge e as a gain from the edge (when positive) or, a loss (when negative, occurring equally likely, by symmetry). It then seems natural to ask for the maximal possible total gain over all structures, so in particular, W_n above is the highest possible gain over all assignments (matchings), spanning trees, etc.

In the case of the optimal matching (the assignment problem), in their independent work [38], Lifshits and Tadevosian have recently obtained the asymptotics of $\mathbb{E}W_n$ for general i.i.d. weights whose quantile function tends to infinity and slowly varies at zero.

In the important case of Gaussian weights, Λ_* is explicit (quadratic). Moreover, the concentration for the supremum of a Gaussian process allows to obtain asymptotic values of the expectation as well. In the case of the optimal matching, the asymptotics of $\mathbb{E}W_n$ was recently found by Mordant and Segers (see Theorem 2.3 in [42]) and, independently, by Lifshits and Tadevosian (see Theorem 1 in [37]), however with a different argument for the lower bound (via a greedy construction, whilst we employ pruning).

Corollary 3. *If in Theorem 1 or 2 we let the distribution of X be standard Gaussian (mean 0, variance 1), then in Theorem 1 as well as Theorem 2 (a), (b), (c), we have*

$$W_n = (1 + o(1))n\sqrt{2\log n}, \quad w.h.p. \quad and \quad \mathbb{E}W_n = (1 + o(1))n\sqrt{2\log n},$$

whilst in Theorem 2 (d),

$$W_n = (1 + o(1))\ell\sqrt{2d^{-1}\log n}, \quad w.h.p. \quad and \quad \mathbb{E}W_n = (1 + o(1))\ell\sqrt{2d^{-1}\log n}.$$

3 Proofs

3.1 Overview

Our upper bounds are based on the union bound. They turn out to be tight and are matched by lower bounds, obtained by means of constructions exploiting threshold probabilities for binomial random graphs.

3.2 Regular tails

We shall need the following simple lemma which establishes asymptotic behaviour of certain sequences showing up in the proofs of our main results. The lemma is based on the tail regularity of good random variables.

Lemma 4. *Let $\alpha > 0$. Let ω_n be a positive sequence such that $\omega_n \to \infty$ with $\omega_n = n^{o(1)}$. Let X be a good random variable and define $x_n = \inf\{t > 0, \ \mathbb{P}(X > t) \leqslant \omega_n n^{-\alpha}\}$. Then, for all n large enough, we have*

$$(1 - o(1))\Lambda_*^{-1}(\alpha \log n) \leqslant x_n \leqslant \Lambda_*^{-1}(\alpha \log n).$$

Proof. Excluding the degenerate situation, when $X = 0$ a.s. (in which case $x_n = 0$ and $\Lambda_*^{-1} \equiv 0$, so the lemma holds trivially), we can assume without loss of generality that $x_n > 0$ for all n (since $\omega_n n^{-\alpha} \to 0$). As $\omega_n \to \infty$, let us also assume that $\omega_n \geqslant 1$ for all n.

By the definition of x_n, for every $\theta \in (0, 1)$, we have $\mathbb{P}(X > \theta x_n) \geqslant \omega_n n^{-\alpha} \geqslant n^{-\alpha}$. On the other hand, by Chernoff's inequality (1), $\mathbb{P}(X > \theta x_n) \leqslant e^{-\Lambda_*(\theta x_n)}$. Combining these two bounds yields $\Lambda_*(\theta x_n) \leqslant \alpha \log n$, so $x_n \leqslant \theta^{-1}\Lambda_*^{-1}(\alpha \log n)$, which proves the upper bound on x_n.

For the lower bound, if X is bounded, say $A = \operatorname{ess\,sup}|X|$, $0 < A < \infty$, then $x_n \to A$ as $n \to \infty$, as well as $\Lambda_*^{-1}(\alpha \log n) \to A$ as $n \to \infty$ (because $\Lambda_*(t) = +\infty$ for every $t > A$). Suppose now that X is not bounded. Then $x_n \to \infty$ as $n \to \infty$. By assumption (2), there is a positive sequence $\varepsilon_n \to 0$ such that $\mathbb{P}(X > x_n) = e^{-(1+\varepsilon_n)\Lambda_*(x_n)}$. By the definition of x_n, $\mathbb{P}(X > x_n) \leqslant \omega_n n^{-\alpha}$, thus $(1 + \varepsilon_n)\Lambda_*(x_n) \geqslant -\log \omega_n + \alpha \log n$, so $\Lambda_*(x_n) \geqslant \theta_n \alpha \log n$ with $\theta_n = (1 + \varepsilon_n)^{-1}(1 - \frac{\log \omega_n}{\alpha \log n})$. By the assumptions on ω_n, we have $\theta_n < 1$ with $\theta_n \to 1$, consequently $\theta_n \alpha \log n \geqslant \Lambda_*(\theta_n \Lambda_*^{-1}(\alpha \log n))$ (the convexity of Λ_* and $\Lambda_*(0) = 0$ imply that $u \mapsto \Lambda_*(u)/u$ is nondecreasing, which we shall also use several times in the sequel). We thus get $x_n \geqslant \theta_n \Lambda_*^{-1}(\alpha \log n) = (1 - o(1))\Lambda_*^{-1}(\alpha \log n)$, as desired. $\qquad\square$

Remark 5. Thanks to Lemma 4, the asymptotic values of W_n from Theorems 1 and 2 can be equivalently stated in terms of the sequences x_n, which may be easier to compute than $\Lambda_*^{-1}(\log n)$ for a given distribution of X.

3.3 Proof of Theorem 1

We begin with a high probability upper bound on W_n. Note that thanks to Chernoff's bound (1), if $X_1, \ldots, X_k$ are i.i.d. copies of X, then for every $t > 0$,

$$\mathbb{P}(X_1 + \cdots + X_k > kt) \leqslant \exp\{-k\Lambda_*(t)\} \tag{3}$$

(by independence, the rate function of $X_1 + \cdots + X_k$ at kt is $k\Lambda_*(t)$). Using first a union bound and then this, for $\delta_n \geqslant 0$, we obtain

$$\mathbb{P}\left(W_n > (1 + \delta_n)n\Lambda_*^{-1}(\log n)\right) \leqslant \sum_{M \in \mathscr{C}_n} \mathbb{P}\left(\sum_{e \in E(M)} X_e > (1 + \delta_n)n\Lambda_*^{-1}(\log n)\right)$$

$$\leqslant \exp\left\{ \log|\mathscr{C}_n| - n\Lambda_* \big((1+\delta_n)\Lambda_*^{-1}(\log n)\big) \right\}$$
$$\leqslant \exp\left\{ \log|\mathscr{C}_n| - (1+\delta_n)n\log n \right\}, \tag{4}$$

where in the last inequality we use the monotonicity of $\Lambda_*(u)/u$. Since $|\mathscr{C}_n| = n! \leqslant \frac{n^{n+1}}{e^n}$, for every $n \geqslant 7$, choosing $\delta_n = n^{-1}$ yields

$$W_n \leqslant (1+n^{-1})n\Lambda_*^{-1}(\log n) \quad \text{with probability at least } 1 - e^{-n}.$$

To establish a matching lower bound, we construct a random subgraph comprising only *large* weights, which contains a perfect matching w.h.p. We set

$$p_n = \frac{2\log n}{n} \qquad \text{and} \qquad x_n = \inf\{t > 0, \ \mathbb{P}(X > t) \leqslant p_n\}. \tag{5}$$

Excluding again the trivial case of $X = 0$ a.s., we have $x_n > 0$ (eventually), and then for every $0 < \delta_n < 1$, we have $p'_n = \mathbb{P}(X > (1-\delta_n)x_n) > p_n$. Since the weights X_e are i.i.d., the random bipartite graph $([n], [n], \{e \in [n] \times [n], \ X_e > (1-\delta_n)x_n\})$ is in fact G_{n,n,p'_n}, so by the classical result of Erdős and Renyi (see [15] or Theorem 6.1 in [19]), w.h.p. it contains a perfect matching which gives

$$W_n \geqslant n(1-\delta_n)x_n, \qquad \text{w.h.p.}$$

We choose δ_n arbitrarily with $\delta_n = o(1)$ as $n \to \infty$ and it remains to show that

$$x_n \geqslant (1-o(1))\Lambda_*^{-1}(\log n).$$

This follows from Lemma 4 applied to $\alpha = 1$ and $\omega_n = 2\log n$. $\qquad\square$

3.4 Proof of Theorem 2

We follow exactly the same strategy as in the proof of Theorem 1.

Case (a), (b), (c): the upper bound. In case (a) and (b) respectively, every graph $G \in \mathscr{C}_n$ has the same number of edges, $n-1$ and n respectively. Repeating verbatim the derivation of (4) yields in each case

$$\mathbb{P}\big(W_n > (1+\delta_n)n\Lambda_*^{-1}(\log n)\big) \leqslant \exp\left\{ \log|\mathscr{C}_n| - (1+\delta_n)n\log n \right\}.$$

The same holds in case (c) because letting $\ell(P)$ be the number of edges on a path P, using (3), we have

$$\mathbb{P}\big(W_n > (1+\delta_n)n\Lambda_*^{-1}(\log n)\big) \leqslant \sum_{P \in \mathscr{C}_n} \mathbb{P}\left(\sum_{e \in E(P)} X_e > (1+\delta_n)n\Lambda_*^{-1}(\log n) \right)$$
$$\leqslant \sum_{P \in \mathscr{C}_n} \exp\left\{ -\ell(P)\Lambda_*\left((1+\delta_n)\frac{n}{\ell(P)}\Lambda_*^{-1}(\log n) \right) \right\}$$

$$\leqslant \sum_{P \in \mathscr{C}_n} \exp\left\{-(1+\delta_n)n \log n\right\},$$

where in the last inequality we use the convexity of Λ_* and $(1+\delta_n)\frac{n}{\ell(P)} \geqslant 1$. In case (a), (b), (c), we have $|\mathscr{C}_n| = n^{n-2}$, $|\mathscr{C}_n| = \frac{1}{2}(n-1)!$, $|\mathscr{C}_n| = \sum_{l=0}^{n-2} \binom{n-2}{l} l! = (n-2)! \sum_{l=0}^{n-2} \frac{1}{(n-2-l)!} \leqslant e(n-2)!$, respectively, so in each case, setting $\delta_n = 0$ suffices to get $\log |\mathscr{C}_n| - (1+\delta_n)n \log n \to -\infty$ and as a result,

$$W_n \leqslant n\Lambda_*^{-1}(\log n), \qquad \text{w.h.p.}$$

Case (a), (b), (c): the lower bound. We define p_n, p_n' and x_n as in the proof of Theorem 1, see (5). Considering the random graph $\mathcal{G} = ([n], \{e \in \binom{[n]}{2}, X_e > (1-\delta_n)x_n\})$ which is distributed as $G_{n,p_n'}$, we get in each case that $W_n \geqslant (n-1)(1-\delta_n)x_n$ w.h.p. This is because $p_n' > \frac{2 \log n}{n}$ guarantees that $\mathcal{G}$ is w.h.p. (a) connected, (b) Hamiltonian, (c) Hamiltonian-connected (in particular $\mathcal{G}$ has a path of length n between vertices 1 and 2), see [13], [33], [4] respectively, or Theorems 4.1, 6.5 and Exercise 6.7.11 in [19]. Then Lemma 4 finishes the argument.

Case (d): the upper bound. Let v and ℓ denote the number of vertices and edges in H_0, respectively. The density of H_0 is $d = \frac{\ell}{v}$. For the cardinality $|\mathscr{C}_n|$, that is the number of copies of H_0 in K_n, we have $|\mathscr{C}_n| \leqslant \binom{n}{v}v! \leqslant n^v$. As in (4), we obtain

$$\begin{aligned}
\mathbb{P}\left(W_n > (1+\delta_n)\ell\Lambda_*^{-1}(d^{-1}\log n)\right) &\leqslant \exp\left\{\log |\mathscr{C}_n| - \ell\Lambda_*\left((1+\delta_n)\Lambda_*^{-1}(d^{-1}\log n)\right)\right\} \\
&\leqslant \exp\left\{v \log n - (1+\delta_n)\ell d^{-1}\log n\right\} \\
&= \exp\left\{-\delta_n v \log n\right\}.
\end{aligned}$$

Choosing, say $\delta_n = (\log n)^{-1/2}$, we conclude

$$W_n \leqslant (1+\delta_n)\ell\Lambda_*^{-1}(d^{-1}\log n), \qquad \text{w.h.p.}$$

Case (d): the lower bound. It is a classical result of Erdős and Rényi from [14] (see also Theorem 5.3 in [19]) that for a balanced graph H_0, when $pn^{1/d} \to \infty$, the random graph $G_{n,p}$ contains a copy of H_0 w.h.p. Therefore, letting $p_n = \omega_n n^{-1/d}$ with $\omega_n \to \infty$, defining x_n as in (5) and considering the random graph $\mathcal{G}$, we obtain as in cases (a), (b), (c), $W_n \geqslant \ell(1-\delta_n)x_n$ w.h.p. It remains to show that $x_n \geqslant (1-o(1))\Lambda_*^{-1}(d^{-1}\log n)$. This follows from Lemma 4 applied to $\alpha = d^{-1}$ as long as we choose $\omega_n \to \infty$, $\omega_n = n^{o(1)}$. $\qquad\square$

Remark 6. Using the common approach based on moment generating functions (for example, see the proof of (A.3) in Appendix A.2 in [5]), we can also obtain the following upper bound on the expectation in Theorems 1 and 2,

$$\mathbb{E}W_n \leqslant l\Lambda_*^{-1}(l^{-1}\log |\mathscr{C}_n|), \tag{6}$$

where $l = \max_{H \in \mathscr{C}_n} |E(H)|$ is the maximal number of edges in the graphs from a given class $\mathscr{C}_n$ (so $l = n, n-1, n, n, \ell$ in Theorems 1 and 2 (a), (b), (c), (d), respectively). Moreover, in each of the cases, the right hand side is asymptotic to the high probability

bound on W_n from Theorems 1 and 2. A proof of (6) can be sketched as follows: using $\max x_k \leqslant \frac{1}{t} \log \sum_k e^{tx_k}$ valid for all $t > 0$ and $x_k \in \mathbb{R}$, the concavity of the log function and independence, we get

$$\mathbb{E}W_n \leqslant \frac{1}{t} \log \sum_{H \in \mathscr{C}_n} \mathbb{E}e^{t \sum_{e \in E(H)} X_e} = \frac{1}{t} \log \sum_{H \in \mathscr{C}_n} e^{|E(H)| \Lambda(t)} \leqslant \frac{\log |\mathscr{C}_n| + l\Lambda(t)}{t}.$$

Taking the infimum over $t > 0$ finishes the argument.

3.5 Proof of Corollary 3

When X is standard Gaussian, $\mathbb{E}e^{tX} = e^{t^2/2}$, so $\Lambda_*(t) = \frac{t^2}{2}$ and $\Lambda_*^{-1}(t) = \sqrt{2t}$. If we let $l = \max_{H \in \mathscr{C}_n} |E(H)|$, then the variance of each Gaussian that W_n takes the maximum over is bounded by l, $\mathrm{Var}(\sum_{e \in E(H)} X_e) = |E(H)| \leqslant l$. From the concentration of the maximum of a Gaussian process around its expectation (see for instance Theorem 7.1 in [34]), we get

$$\mathbb{P}\left(|W_n - \mathbb{E}W_n| \geqslant t\right) \leqslant 2e^{-t^2/(2l)}, \qquad t \geqslant 0. \tag{7}$$

In the case of Theorem 1, $l = n$, so taking, say $t = n$, we get from the above that $|W_n - \mathbb{E}W_n| < n$ w.h.p. Combining this with Theorem 1 which gives $W_n = (1 + o(1))n\sqrt{2 \log n}$ w.h.p., we obtain $\mathbb{E}W_n = (1 + o(1))n\sqrt{2 \log n}$, as desired. We proceed analogously in the case of Theorem 2 (and omit the details). $\qquad \square$

4 Final remarks

4.1 Conclusion

We have determined asymptotically typical values of the weight of the minimal-weight common combinatorial structures in complete graphs with independent identically distributed symmetric weights having regular tails, that is satisfying (2). A natural next step would be to establish limit theorems and the order of fluctuations.

4.2 Fluctuations

Even in the case of Gaussian weights, where many tools exist (e.g. [5, 6, 11]), finding the asymptotic value of the variance seems interesting and challenging. In particular, for optimal matchings, i.e. in the setting of Theorem 1, it follows from (7) that $\mathrm{Var}(W_n) \leqslant 4n$. On the other hand, Mordant and Segers argued in [42] that $\mathrm{Var}(W_n) \geqslant 1$ (exploiting symmetries of the covariance structure). They also noted that by a general phenomenon in superconcentration (Theorem 8.1 of Chatterjee from [5]) and the asymptotics of the mean $\mathbb{E}W_n$, the upper bound in fact improves and we have $\mathrm{Var}(W_n) = o(n)$, however the exact order of fluctuations $\mathrm{Var}(W_n)$ seems elusive.

4.3 Refinements

It is instructive to see the shortcoming of our main result when applied to nonnegative weights. Suppose we consider the minimum spanning tree problem on the complete graph with independent weights Y_e, each uniformly distributed on $[0, 2]$. Since $X_e = Y_e - 1$ is uniform on $[-1, 1]$, applying Theorem 2 (and Remark 5) with $x_n = \inf\{t > 0, \, \mathbb{P}(X > t) \leqslant \frac{2 \log n}{n}\} = 1 - \frac{4 \log n}{n}$, w.h.p., we have

$$-(1 + o(1))n \left(1 - \frac{4 \log n}{n}\right) = W_n = \min_T \sum_{e \in E(T)} X_e = \left(\min_T \sum_{e \in E(T)} Y_e\right) - (n - 1).$$

Without knowing the implicit $o(1)$ term, we cannot infer the asympototic behaviour of $\min_T \sum_{e \in E(T)} Y_e$ (which by Frieze's result from [18], tends to $2\zeta(3)$ in probability). It would be of interest to remedy this and refine the $o(1)$ term for (general) symmetric distributions.

4.4 Measures without regular tails

Recall that a random variable X is said to have regular upper tails if $\lim_{t \to \infty} \frac{-\log \mathbb{P}(X > t)}{\Lambda_*(t)} = 1$, see (2). We show an example of X for which this does not hold. The idea of our example comes from [29]. Fix an increasing sequence of positive numbers $0 = x_0 < x_1 < x_2 < \cdots$ such that $x_n \to \infty$ and let $\log 2 = y_0 < y_1 < y_2 < \cdots$ be an increasing sequence of positive numbers such that $y_n \to \infty$. Define the following nonincreasing right-continuous step function

$$T(t) = \sum_{n=0}^{\infty} e^{-y_n} \mathbf{1}_{[x_n, x_{n+1})}(t), \qquad t \geqslant 0.$$

Let X be a symmetric random variable such that $\mathbb{P}(X > t) = T(t)$, $t \geqslant 0$. In other words, X is discrete taking the values $\pm x_n$ with probabilities $e^{-y_{n-1}} - e^{-y_n}$, $n = 1, 2, \ldots$. Choosing y_n growing *much faster* relative to x_n, it is easily guaranteed that $\Lambda(t) < \infty$ for every $t \in \mathbb{R}$ (it suffices that $\sum e^{tx_n - y_{n-1}} < \infty$ for every $t > 0$). We show that we can choose the y_n such that $\limsup_{t \to \infty} \frac{-\log \mathbb{P}(X > t)}{\Lambda_*(t)} = \infty$ and consequently, X does not have regular upper tails and hence is not *good* (in the sense of (2) of our definition). By Chernoff's inequality, $e^{-y_{n-1}} = \mathbb{P}(X \geqslant x_n) \leqslant e^{-\Lambda_*(x_n)}$, so $\Lambda_*(x_n) \leqslant y_{n-1}$ for every $n = 1, 2, \ldots$. Therefore, by the convexity of Λ_*, for $\alpha \in (0, 1)$,

$$\Lambda_*(\alpha x_n + (1 - \alpha)x_{n+1}) \leqslant \alpha y_{n-1} + (1 - \alpha)y_n$$

and we obtain

$$\begin{aligned}
\frac{-\log \mathbb{P}(X > \alpha x_n + (1 - \alpha)x_{n+1})}{\Lambda_*(\alpha x_n + (1 - \alpha)x_{n+1})} &= \frac{y_n}{\Lambda_*(\alpha x_n + (1 - \alpha)x_{n+1})} \\
&\geqslant \frac{y_n}{\alpha y_{n-1} + (1 - \alpha)y_n}.
\end{aligned}$$

If we choose the sequence (y_n) to grow fast enough, specifically such that $\frac{y_{n-1}}{y_n} \to 0$ (say $y_n = 2^{n^2}$, $n \geqslant 1$), then we get $\limsup_{t\to\infty} \frac{-\log \mathbb{P}(X>t)}{\Lambda_*(t)} \geqslant \frac{1}{1-\alpha}$. Letting $\alpha \to 1$ finishes the argument.

Acknowledgements

References

[1] Aldous, D., The $\zeta(2)$ limit in the random assignment problem. Random Structures Algorithms 18 (2001), no. 4, 381–418.

[2] Benjamini, I., Lubetzky E., Peled, Y., Minimum weight disk triangulations and fillings, Trans. Amer. Math. Soc. 374 (2021), no. 5, 3265–3287.

[3] Bhamidi, S., van der Hofstad, R., Weak disorder asymptotics in the stochastic mean-field model of distance. Ann. Appl. Probab. 22 (2012), no. 1, 29–69.

[4] Bollobás, B., Fenner, T. I., Frieze, A. M., An algorithm for finding Hamilton paths and cycles in random graphs. Combinatorica 7 (1987), no. 4, 327–341.

[5] Chatterjee, S., Superconcentration and related topics. Springer Monographs in Mathematics. Springer, Cham, 2014.

[6] Chatterjee, S., A general method for lower bounds on fluctuations of random variables. Ann. Probab. 47 (2019), no. 4, 2140–2171.

[7] Cooper, C., Frieze, A., Ince, N., Janson, S., Spencer, J., On the length of a random minimum spanning tree. Combin. Probab. Comput. 25 (2016), no. 1, 89–107.

[8] Coppersmith, D., Sorkin, G., Constructive bounds and exact expectations for the random assignment problem. Random Structures Algorithms 15 (1999), no. 2, 113–144.

[9] Davis, R., Prieditis, A., The expected length of a shortest path. Inform. Process. Lett. 46 (1993), no. 3, 135–141.

[10] Dembo, A., Zeitouni, O., Large deviations techniques and applications. Corrected reprint of the second (1998) edition. Stochastic Modelling and Applied Probability, 38. Springer-Verlag, Berlin, 2010.

[11] Ding, J., Eldan, R., Zhai, A., On multiple peaks and moderate deviations for the supremum of a Gaussian field. Ann. Probab. 43 (2015), no. 6, 3468–3493.

[12] Dyer, M., Frieze, A. M., Probabilistic analysis of the generalised assignment problem. Math. Programming 55 (1992), no. 2, Ser. A, 169–181.

[13] Erdős, P.; Rényi, A., On random graphs. I. Publ. Math. Debrecen 6 (1959), 290–297.

[14] Erdős, P.; Rényi, A., On the evolution of random graphs. Magyar Tud. Akad. Mat. Kutató Int. Közl. 5 (1960), 17–61.

[15] Erdős, P., Rényi, A., On random matrices. Magyar Tud. Akad. Mat. Kutató Int. Közl. 8 (1964), 455–461.

[16] Flaxman, A., The lower tail of the random minimum spanning tree. Electron. J. Combin. 14 (2007), no. 1, #N3, 4 pp.

[17] Flaxman, A., Frieze, A. M., Krivelevich, M., On the random 2-stage minimum spanning tree. Random Structures Algorithms 28 (2006), no. 1, 24–36.

[18] Frieze, A. M., On the value of a random minimum spanning tree problem. Discrete Appl. Math. 10 (1985), no. 1, 47–56.

[19] Frieze, A. M., Karoński, M., Introduction to random graphs. Cambridge University Press, Cambridge, 2016.

[20] Frieze, A. M., Pegden, W., Sorkin, G., The distribution of minimum-weight cliques and other subgraphs in graphs with random edge weights. SIAM J. Discrete Math. 32 (2018), no. 3, 2115–2133.

[21] Frieze, A. M., Pegden, W., Sorkin G., Tkocz, T., Minimum-weight combinatorial structures under random cost-constraints, Electron. J. Combin. 28 (2021), no. 1, #P1.20, 18 pp.

[22] Frieze, A. M., Pegden, Tkocz, T., On random multi-dimensional assignment problems. Discrete Appl. Math. 287 (2020), 1–9.

[23] Frieze, A. M., Sorkin, G., The probabilistic relationship between the assignment and asymmetric traveling salesman problems. SIAM J. Comput. 36 (2006/07), no. 5, 1435–1452.

[24] Frieze, A. M., Ruszinkó, M., Thoma, L., A note on random minimum length spanning trees. Electron. J. Combin. 7 (2000), #R41, 5 pp.

[25] Frieze, A. M., Tkocz, T., A randomly weighted minimum spanning tree with a random cost constraint, Electron. J. Combin. 28 (2021), no. 1, #P1.22, 25 pp.

[26] Frieze, A. M., Tkocz, T., Shortest paths with a cost constraint: a probabilistic analysis, Discrete Appl. Math. 302 (2021), 46–53.

[27] Frieze, A. M., Tkocz, T., Probabilistic analysis of algorithms for cost constrained minimum weighted combinatorial objects. Oper. Res. Lett. 49 (2021), no. 3, 400–404.

[28] Frieze, A. M., Tkocz, T., A randomly weighted minimum arborescence with a random cost constraint, Math. Oper. Res. 47 (2022), no. 2, 1664–1680.

[29] Gatzouras, D., Giannopoulos, A., Threshold for the volume spanned by random points with independent coordinates. Israel J. Math. 169 (2009), 125–153.

[30] Janson, S., The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph. Random Structures Algorithms 7 (1995), no. 4, 337–355.

[31] Janson, S., One, two and three times $\log n/n$ for paths in a complete graph with random weights. Random graphs and combinatorial structures (Oberwolfach, 1997). Combin. Probab. Comput. 8 (1999), no. 4, 347–361.

[32] Karp, R. M., A patching algorithm for the nonsymmetric traveling-salesman problem. SIAM J. Comput. 8 (1979), no. 4, 561–573.

[33] Komlós, J. Szemerdi, E., Limit distribution for the existence of Hamiltonian cycles in a random graph. Discrete Math. 43 (1983), no. 1, 55–63.

[34] Ledoux, M., The concentration of measure phenomenon. Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, 2001.

[35] Li, W., Zhang, X., On the difference of expected lengths of minimum spanning trees. Combin. Probab. Comput. 18 (2009), no. 3, 423–434.

[36] Lifshits, M., Mordant, G., Extrema of multinomial assignment process, preprint (2022), `arXiv:2209.05028`.

[37] Lifshits, M., Tadevosian, A., Gaussian Assignment Process, preprint (2021), `arXiv:2107.04883`.

[38] Lifshits, M., Tadevosian, A., On the maximum of random assignment process, Statist. Probab. Lett. 187 (2022), Paper No. 109530, 6 pp.

[39] Lifshits, M., Tadevosian, A., On the maximum of a special random assignment process, preprint (2022), `arXiv:2205.01199`.

[40] Linusson, S., Wästlund, J., A proof of Parisi's conjecture on the random assignment problem. Probab. Theory Related Fields 128 (2004), no. 3, 419–440.

[41] McDiarmid, C., On the method of bounded differences. Surveys in combinatorics, 1989 (Norwich, 1989), 148–188, London Math. Soc. Lecture Note Ser., 141, Cambridge Univ. Press, Cambridge, 1989.

[42] Mordant, G., Segers, J., Maxima and near-maxima of a Gaussian random assignment field, Statist. Probab. Lett. 173 (2021), Paper No. 109087, 8 pp.

[43] Nair, C., Prabhakar, B., Sharma, M., Proofs of the Parisi and Coppersmith-Sorkin random assignment conjectures. Random Structures Algorithms 27 (2005), no. 4, 413–444.

[44] Parisi, G., A Conjecture on random bipartite matching (1998), `arXiv:cond-mat/9801176`.

[45] Wästlund, J., An easy proof of the $\zeta(2)$ limit in the random assignment problem. Electron. Commun. Probab. 14 (2009), 261–269.