Combinatorics, Probability and Computing (2017) 26, 797–825.
doi:10.1017/S0963548317000232

Diameter of the Stochastic Mean-Field Model of Distance

SHANKAR BHAMIDI[1] and REMCO VAN DER HOFSTAD[2]

[1]Department of Statistics, University of North Carolina, Chapel Hill, NC 27599, USA
(e-mail: bhamidi@email.unc.edu)
[2]Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands
(e-mail: rhofstad@win.tue.nl)

Received 1 June 2013; revised 2 April 2017; first published online 7 August 2017

We consider the complete graph $\mathcal{K}_n$ on n vertices with exponential mean n edge lengths. Writing C_{ij} for the weight of the smallest-weight path between vertices $i,j \in [n]$, Janson [18] showed that $\max_{i,j\in[n]} C_{ij}/\log n$ converges in probability to 3. We extend these results by showing that $\max_{i,j\in[n]} C_{ij} - 3\log n$ converges in distribution to some limiting random variable that can be identified via a maximization procedure on a limiting infinite random structure. Interestingly, this limiting random variable has also appeared as the weak limit of the re-centred *graph diameter* of the barely supercritical Erdős–Rényi random graph in [22].

2010 *Mathematics subject classification*: Primary 60C05
Secondary 05C80, 90B15

1. Introduction

We consider the complete graph $\mathcal{K}_n$ on the vertex set $[n] := \{1,2,\ldots,n\}$ and edge set $\mathcal{E}_n := \{\{i,j\} : i < j \in [n]\}$. Each edge $e \in \mathcal{E}_n$, is assigned an exponential mean n edge length E_e, independent across edges. This implies that for any vertex v, the closest neighbour to this vertex is $O_{\mathbb{P}}(1)$ distance away. Define the length of a path π as

$$w(\pi) := \sum_{e\in\pi} E_e. \tag{1.1}$$

This assignment of random edge lengths makes $\mathcal{K}_n$ a (random) metric space often referred to as the *stochastic mean-field model of distance* (see Section 3). By continuity of the distribution of edge lengths, this metric space has unique geodesics. For any two vertices $i,j \in [n]$, let $\pi(i,j)$ denote the shortest path between these two vertices and write C_{ij} for the length of this geodesic; precisely, writing $\mathfrak{P}_{ij}$ for the collection of all paths from i to j, then $C_{ij} = \min_{\mathfrak{P}_{ij}} w(\pi)$. The functional of interest in this paper is the diameter of the metric space:

$$\mathrm{Diam}_w(\mathcal{K}_n) := \max_{i,j\in[n]} C_{ij}. \tag{1.2}$$

We first dive into the statement of the main result, postponing a full discussion to Section 3.

2. Results

The main aim of this paper is to prove that the diameter defined in (1.2) properly re-centred converges to a limiting random variable. We start by constructing this limiting random variable.

2.1. Construction of the limiting random variable

The limiting random variable arises as an optimization problem on an infinite randomly weighted graph $\mathcal{G}_\infty = (\mathcal{V}, \mathcal{E})$. The vertex set of this graph is the set of positive integers $\mathbb{Z}_+ = \{1, 2, \ldots\}$, while the edge set consists of all undirected edges $\mathcal{E} = \{\{i, j\} : i, j \in \mathbb{Z}_+, i \neq j\}$. Let $\mathcal{P}$ be a Poisson process on $\mathbb{R}$ with intensity measure having density

$$\lambda(y) = \mathrm{e}^{-y}, \quad -\infty < y < \infty. \tag{2.1}$$

It is easy to check that $\max\{x : x \in \mathcal{P}\} < \infty$ a.s. Thus we can order the points in $\mathcal{P}$ as $Y_1 > Y_2 > \cdots$. We think of Y_i as the vertex weight at $i \in \mathbb{Z}_+$. The edge weights are easier to describe. Let $(\Lambda_{st})_{s,t \in \mathbb{Z}_+, s<t}$ be a family of independent standard Gumbel random variables, namely Λ_{st} has cumulative distribution function

$$F(x) = \mathrm{e}^{-\mathrm{e}^{-x}}, \quad -\infty < x < \infty. \tag{2.2}$$

The random variable $\Lambda_{s,t}$ gives the weight of an edge $\{s, t\} \in \mathcal{E}$. Now consider the optimization problem

$$\Xi := \max_{s,t \in \mathbb{Z}_+, s<t} (Y_s + Y_t - \Lambda_{st}). \tag{2.3}$$

Although this is not obvious, we shall show that $\Xi < \infty$ a.s. The main result in this paper is as follows. We write $\xrightarrow{\mathrm{w}}$ to denote convergence in distribution.

Theorem 2.1 (diameter asymptotics). *For the diameter of the stochastic mean-field model of distance, as $n \to \infty$,*

$$\max_{i,j \in [n]} C_{ij} - 3 \log n \xrightarrow{\mathrm{w}} \Xi,$$

and

$$\mathbb{E}\Big[\max_{i,j \in [n]} C_{ij}\Big] - 3 \log n \to \mathbb{E}[\Xi], \quad \mathrm{Var}\Big(\max_{i,j \in [n]} C_{ij}\Big) \to \mathrm{Var}(\Xi). \tag{2.4}$$

Remark. Theorem 2.1 solves [18, Problems 1 and 2]. Further, our proof shows that for any fixed $p \geqslant 1$, as $n \to \infty$,

$$\mathbb{E}\Big(\max_{i,j \in [n]} C_{ij} - 3 \log n\Big)^p \to \mathbb{E}(\Xi^p).$$

2.2. Basic notation

Let us briefly describe the notation used in the rest of the paper. We write $\xrightarrow{\mathrm{P}}$ to denote convergence in probability. For a sequence of random variables $(X_n)_{n \geqslant 1}$, we write $X_n = O_{\mathbb{P}}(b_n)$

when $(|X_n|/b_n)_{n\geqslant 1}$ is a tight sequence of random variables as $n \to \infty$, and $X_n = o_{\mathbb{P}}(b_n)$ when $|X_n|/b_n \xrightarrow{\mathrm{P}} 0$ as $n \to \infty$. Further we write $X_n = \Theta_{\mathbb{P}}(b_n)$ if $(X_n/b_n)_{n\geqslant 1}$ and $(b_n/X_n)_{n\geqslant 1}$ are both a tight sequence of random variables. For a non-negative function $n \mapsto g(n)$, we write $f(n) = O(g(n))$ when $|f(n)|/g(n)$ is uniformly bounded, and $f(n) = o(g(n))$ when $\lim_{n\to\infty} f(n)/g(n) = 0$. Furthermore, we write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$. Finally, we write that a sequence of events $(A_n)_{n\geqslant 1}$ occurs *with high probability* (w.h.p.) when $\mathbb{P}(A_n) \to 1$. We use $Y \sim \exp(\lambda)$ for a random variable with exponential distribution with rate λ. At several places we will use the fact that for $Y \sim \exp(1)$ then $\log(1/Y)$ has a standard Gumbel distribution with distribution as in (2.2). To ease notation we will occasionally use symbols representing a real number when an integer is required (*e.g.* $\sum_{j=1}^{\sqrt{n}}$) and suppress the precise $\lfloor \cdot \rfloor$ notation.

3. Background and related results

We now discuss our results and place them in the context of results in the literature.

3.1. Stochastic mean-field model of distance

The stochastic mean-field model of distance has arisen in a number of different contexts in understanding the structure of combinatorial optimization problems in the presence of random data, ranging from shortest path problems [18] to random assignment problems [2, 3], minimal spanning trees [15, 17] and travelling salesman problems [25]; see [6] for a comprehensive survey and related literature. The closest work to this study is the paper by Janson [18]. Recall that C_{ij} denotes the length of the geodesic between two vertices $i, j \in [n]$; by symmetry this has the same distribution for any two vertices in i, j. For any vertex $i \in [n]$, write $\mathrm{Flood}[i] := \max_{j\in[n]} C_{ij}$ for the maximum time started at i to reach all vertices in $\mathcal{K}_n$ (often called the *flooding* time). Then Janson proved that, as $n \to \infty$,

$$\frac{C_{ij}}{\log n} \xrightarrow{\mathrm{P}} 1, \quad \frac{\mathrm{Flood}[i]}{\log n} \xrightarrow{\mathrm{P}} 2, \quad \frac{\mathrm{Diam}_w(\mathcal{K}_n)}{\log n} \xrightarrow{\mathrm{P}} 3, \tag{3.1}$$

and further

$$C_{ij} - \log n \xrightarrow{\mathrm{w}} \Lambda_1 + \Lambda_2 - \Lambda_{12}, \tag{3.2}$$

while

$$\mathrm{Flood}[i] - 2\log n \xrightarrow{\mathrm{w}} \Lambda_1 + \Lambda_2. \tag{3.3}$$

Here $\Lambda_1, \Lambda_2, \Lambda_{12}$ are all independent standard Gumbel random variables as in (2.2). Problems 1 and 2 in [18] then ask whether a result similar to (3.2) and (3.3) holds for the diameter $\mathrm{Diam}_w(\mathcal{K}_n)$ (with (3.1) obviously re-centred by $3\log n$).

The main aim of this paper is to answer this question in the affirmative. We discuss more results about the distribution of Ξ in Section 4.8. In the context of (2.4), for C_{ij} and $\mathrm{Flood}[i]$, Janson also shows convergence of the expectation and variance with explicit limit constants. We have been unable to derive explicit values for the limit constants $\mathbb{E}(\Xi)$ and $\mathrm{Var}(\Xi)$.

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

3.2. Hopcount and extrema

This paper looks at the length of optimal paths (measured in terms of the edge weights). One could also look at the *hopcount* or the number of edges $|\pi(i,j)|$ on the optimal path as well as the largest hopcount $\mathcal{D}^\star = \max_{i,j\in[n]} |\pi(i,j)|$. The entire shortest path tree from a vertex i has the same distribution as a random recursive tree on n vertices (see [24] for a survey). Janson used this in [18] to show that

$$\frac{|\pi(i,j)| - \log n}{\sqrt{\log n}} \xrightarrow{\text{w}} Z,$$

where Z has a standard normal distribution. The maximal hopcount $\mathcal{H}_n(i) = \max_{j\in[n]} |\pi(i,j)|$ from a vertex i has the same distribution as the height of a random recursive tree, which by [12] or [21] satisfies the asymptotics

$$\mathcal{H}_n(i)/\log n \xrightarrow{\text{P}} \mathrm{e} \quad \text{as } n \to \infty.$$

The first-order asymptotics for the maximum hopcount $\mathcal{D}^\star$ were recently proved in [1], showing that

$$\mathcal{D}^\star/\log n \xrightarrow{\text{P}} \alpha^\star,$$

where $\alpha^\star \approx 3.5911$ is the unique solution of the equation $x\log x - x = 1$.

3.3. First passage percolation on random graphs

The last few years have seen progress in the understanding of optimal paths in the presence of edge disorder (usually assumed to have exponential distribution) in the context of various random graph models (see *e.g.* [7, 8, 11] and the references therein). In particular, Proposition 4.4 below with a sketch of proof has appeared in [4, 5, 10].

In the context of our main result, [7] studied the weighted diameter for the random r-regular graphs $\mathcal{G}_{n,r}$ with exponential edge weights, and proved first-order asymptotics. We conjecture that one can adapt the main techniques in this paper to show the second-order asymptotics for $r \geqslant 3$, that is,

$$\mathrm{Diam}_w(\mathcal{G}_{n,r}) - \left(\frac{1}{r-2} + \frac{2}{r}\right)\log n \xrightarrow{\text{w}} \Xi_r, \tag{3.4}$$

for a limit random variable Ξ_r that satisfies that, as $r \to \infty$,

$$r\Xi_r \xrightarrow{\text{w}} \Xi. \tag{3.5}$$

3.4. Diameter of the barely supercritical Erdős–Rényi random graph

Consider the barely supercritical Erdős–Rényi random graph $\mathcal{G}_n(n,(1+\varepsilon)/n)$ where $\varepsilon = \varepsilon_n \to 0$ but $\varepsilon n^3 \to \infty$. It turns out that the random variable Ξ in Theorem 2.1 is closely related to the random variable describing second-order fluctuations for the *graph diameter* $\mathrm{Diam}_g(\mathcal{G}_n(n,(1+\varepsilon)/n))$. Here we use $\mathrm{Diam}_g(\cdot)$ for the graph diameter of a graph, namely the largest *graph* distance between any two vertices in the same component. We now describe this result. Consider the minor modification of the optimization problem defining Ξ in Section 2 where the Poisson process $\mathcal{P}$ generating the vertex weights has intensity measure with density

$$\lambda(y) = \gamma e^{-y}, \quad -\infty < y < \infty.$$

As before, the edge weights Λ_{st} are independent standard Gumbel random variables. Let Ξ_γ denote the random variable corresponding to the optimization problem in (2.3). Let $\lambda = 1+\varepsilon$ and let $\lambda_* < 1$ be the unique value satisfying $\lambda_* e^{-\lambda_*} = \lambda e^{-\lambda}$. After an initial analysis in [13, 14], Riordan and Wormald in [22, Theorem 5.1] showed that there exists a constant $\gamma > 0$ such that

$$\mathrm{Diam}_g(\mathcal{G}_n(n,(1+\varepsilon)/n)) - \frac{\log \varepsilon^3 n}{\log \lambda} - 2\frac{\log \varepsilon^3 n}{\log 1/\lambda_*} \xrightarrow{\mathrm{w}} \Xi_\gamma.$$

We believe that the Poisson cloning technique in [13, 14] coupled with the techniques in this paper may yield an alternative proof of this result, but we defer this to future work.

4. Proofs

We start with the basic ideas behind the main result. We then describe the organization of the rest of the section, which deals with converting this intuitive picture into a proper proof.

4.1. Proof idea

We write $\mathscr{S}_n = (\mathcal{K}_n, (E_e)_{e\in\mathcal{E}_n})$ for the (random) metric space where $(E_e)_{e\in\mathcal{E}_n}$ are i.i.d. mean n exponential random variables. Recall from Janson's result, as described in (3.2), that the distance C_{ij} between typical vertices $i,j \in [n]$ scales like $\log n + O_{\mathbb{P}}(1)$. Intuitively, the extra $2\log n$ in the diameter arises due to the following reason. Consider ranking the vertices according to the distance to their closest neighbour. More precisely, for each vertex $i \in [n]$, write $X_{(i)} = \min_{j\in[n], j\neq i} E_{ij}$, the distance to the closest vertex to i. Arrange these as $X_{(V_1)} > X_{(V_2)} > \cdots > X_{(V_n)}$. We shall show that:

(a) the point process $\mathcal{P}_n = (X_{(V_i)} - \log n)_{i\geqslant 1}$ converges to the Poisson point process $\mathcal{P}$ in Section 2 with intensity measure given by (2.1);

(b) the diameter of $\mathcal{K}_n$ corresponds to the shortest path between a pair of these 'slow' vertices (V_s, V_t);

(c) further, after reaching the closest vertex, the remaining path behaves like a typical optimum path in the original graph $\mathcal{K}_n$ equipped with exponential mean n edge lengths, but now between two disjoint pairs of vertices.

More precisely, part (c) entails that $C_{V_s,V_t} \approx X_{(V_s)} + X_{(V_t)} + d_w(A,B)$, where $A = \{a,b\}$, $B = \{c,d\}$ with a,b,c,d four distinct vertices in $[n]$ and $d_w(A,B)$ is a random variable independent of $X_{(V_t)}, X_{(V_s)}$ having the same distribution as the distance between the sets A,B in the original metric space $\mathscr{S}_n$. The first two terms correspond to the time to get out of these 'slow' vertices, which scale like $\log n + O_{\mathbb{P}}(1)$ by (a) while $d_w(A,B)$ scales like $\log n + O_{\mathbb{P}}(1)$, thus implying that the diameter scales like $3\log n + O_{\mathbb{P}}(1)$. By investigating the fluctuations of $X_{(V_s)}, X_{(V_t)}$ and $d_w(A,B)$, we can also identify the fluctuations of $\max_{i,j\in[n]} C_{ij}$.

Organization of the proof. We start in Section 4.2 by describing the distribution of the shortest path between two disjoint sets of vertices. Section 4.3 proves a weaker version of the Poisson point process limit described in (a) above. Section 4.4 describes the limiting joint distribution of

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

the (properly re-centred) weights of optimal paths between multiple source destination pairs in $\mathscr{S}_n := (\mathcal{K}_n, \{E_e : e \in \mathcal{E}_n\})$. Section 4.5 uses the results in Section 4.3 and 4.4 to study asymptotics for the joint distribution of distances between the slow vertices $(V_s)_{s\in[n]}$. Section 4.6 shows that the diameter of $\mathcal{K}_n$ corresponds to the optimal path between one of the 'first few' slow vertices. The last three sections use these ingredients to show distributional convergence as well as the convergence of the moments of $\mathrm{Diam}_w(\mathcal{K}_n) - 3\log n$ to the limiting random object, thus completing the proof of the main result.

4.2. Explicit distributions for distances between sets of vertices

In this section we explain the proof by Janson of (3.2). We also extend that analysis to the smallest-weight path between disjoint sets of vertices. We remind the reader that the standing assumption henceforth is that each edge has exponential mean n distribution. We start with the following lemma.

Lemma 4.1 (distances between sets of vertices). *Consider two disjoint non-empty sets $A, B \subseteq [n]$. Then,*

$$d_w(A,B) \stackrel{d}{=} \sum_{k=|A|}^{N+|A|-1} \frac{E_k}{k(n-k)}, \tag{4.1}$$

where

(i) *$(E_k)_{k\geqslant 1}$ are i.i.d. mean n exponential random variables;*
(ii) *N is independent of the sequence $(E_k)_{k\geqslant 1}$ with the same distribution as the number of draws required to select the first black ball in an urn containing $|B|$ black balls and $n-|A|-|B|$ white balls, where one is drawing balls without replacement from the urn.*

Proof. We start by exploring the neighbourhood of the set A in a similar way as in [18]. Recall that each edge has an exponential mean n edge length. After having found the ℓth minimal edge and with $k = (|A|+\ell)$, there are $k(n-k)$ edges incident to the found vertices. The minimal edge weight thus has an exponential distribution with mean $n/k(n-k)$. This process stops the first time we find a vertex in B. Since every new vertex added to the cluster of reached vertices is chosen uniformly amongst the set of present unreached vertices, the distribution of the number of steps required to reach a vertex in B has the distribution N asserted in the lemma, independently of the inter-arrival times of new vertices found. Thus the time it takes to find the first element in B is

$$\sum_{\ell=0}^{N-1} \frac{E_k}{(\ell+|A|)(n-\ell-|A|)}. \tag{4.2}$$

Defining $k = \ell + |A|$ proves the claim. □

Now we specialize to a particular case of the above lemma. Fix a vertex, say vertex $v = 1$, and another set $B \subseteq [n] \setminus \{1\}$. For much of what follows, we will be concerned with the optimal path between such a vertex and a set of size $|B| = \Theta(\sqrt{n})$. This is an appropriate time to think about two different but equivalent ways to find such an optimal path.

Process 1. The first way to find the optimal path is the exploration process described in the previous lemma where we start at vertex $v = 1$ and keep adding the closest vertex to the cluster until we hit a vertex in B. Write M_B for the number of vertices other than B that are found in this exploration. Note that the proof of the previous lemma implies that

$$(d_w(\{1\},B),M_B) \stackrel{d}{=} \left(\sum_{k=1}^{N_B} \frac{E_k}{k(n-k)}, N_B\right), \tag{4.3}$$

where N_B is independent of the sequence $(E_k)_{k\geqslant 1}$ and has the same distribution as the number of balls required to get the first black ball when drawing balls without replacement from an urn containing $|B|$ black balls and $n-1-|B|$ white balls.

Process 2. The second way to find the optimal path is the following. We think of water starting at source vertex $v = 1$ at time $t = 0$ percolating through the network at rate one using the edge lengths. Write $\mathsf{SWG}_t^{(1)}$ (an acronym for smallest-weight graph) for the set of vertices reached by time t starting from vertex 1. More precisely,

$$\mathsf{SWG}_t^{(1)} := \{u \in [n]\colon d_w(1,u) \leqslant t\}. \tag{4.4}$$

By convention, vertex $v = 1$ is in $\mathsf{SWG}_t^{(1)}$ for all $t \geqslant 0$. Now note that the size process $(|\mathsf{SWG}_t^{(1)}|)_{t\geqslant 0}$ is a pure-birth Markov process (with respect to the filtration $(\mathcal{F}_t)_{t\geqslant 0} = (\sigma(\mathsf{SWG}_t))_{t\geqslant 0}$) with rate of birth given by $n/k(n-k)$ when the size $|\mathsf{SWG}_t^{(1)}| = k$. Each new vertex added to this cluster is chosen uniformly amongst all available unreached vertices at that time, that is, the vertices $[n]\setminus \mathsf{SWG}_t^{(1)}$. Finally, the distance $d_w(\{1\},B)$ can be recovered as

$$d_w(\{1\},B) := \inf\{t \geqslant 0\colon \mathsf{SWG}_t^{(1)} \cap B \neq \varnothing\}. \tag{4.5}$$

In this section, we use Process 1 to prove the following initial result. We use Process 2 in Section 4.4 below.

Lemma 4.2 (distances between vertex and set of size $b\sqrt{n}$). *Let $B \subseteq [n]$ with $|B| = b\sqrt{n}$. Then as $n \to \infty$,*

$$\left(d_w(\{1\},B) - \frac{1}{2}\log n, M_B/\sqrt{n}\right) \xrightarrow{\text{w}} (\Lambda + \log(\hat{E}/b), \hat{E}/b), \tag{4.6}$$

where $\hat{E}$ is exponential with parameter 1, Λ is Gumbel, and $\hat{E}$ and Λ are independent.

Proof. The above is equivalent to showing

$$(d_w(\{1\},B) - \log M_B, M_B/\sqrt{n}) \xrightarrow{\text{w}} (\Lambda, \hat{E}/b),$$

with $\Lambda, \hat{E}$ independent standard Gumbel and $\exp(1)$ respectively. Fix constants $0 < \alpha < \beta$ and $y \in \mathbb{R}$. Define the event

$$A_n(y,\alpha,\beta) := \{d_w(\{1\},B) - \log M_B \leqslant y\} \cap \{\alpha \leqslant M_B/\sqrt{n} \leqslant \beta\}.$$

Let $(E'_k)_{k\geqslant 1}$ be independent sequence of *mean one* exponential random variables. Equation (4.3) implies

$$\mathbb{P}(A_n(y,\alpha,\beta)) = \sum_{j=\alpha\sqrt{n}}^{\beta\sqrt{n}} \mathbb{P}\left(\sum_{k=1}^{j} \frac{nE'_k}{k(n-k)} - \log j \leqslant y\right) \mathbb{P}(N_B = j). \tag{4.7}$$

Noting that $\sum_{k=1}^{j} 1/j \approx \log j + \gamma$ as $j \to \infty$, where γ is Euler's constant, gives

$$\sum_{k=1}^{j} \frac{nE'_k}{k(n-k)} - \log j \approx \sum_{k=1}^{j} \frac{E'_k - 1}{k} + \gamma + R_n, \tag{4.8}$$

where the error term R_n can be bounded independently of j by

$$|R_n| \leqslant \sum_{k=1}^{\beta\sqrt{n}} \frac{E'_k}{n-k} \xrightarrow{\mathrm{P}} 0, \tag{4.9}$$

as $n \to \infty$. Thus, uniformly for $j \in [\alpha\sqrt{n}, \beta\sqrt{n}]$,

$$\mathbb{P}\left(\sum_{k=1}^{j} \frac{nE'_k}{k(n-k)} - \log j \leqslant y\right) \to \mathbb{P}\left(\sum_{k=1}^{\infty} \frac{E'_k - 1}{k} + \gamma \leqslant y\right).$$

It is easy to check (see *e.g.* [18, Section 3]) that

$$\sum_{k=1}^{\infty} \frac{E'_k - 1}{k} + \gamma \stackrel{d}{=} \Lambda. \tag{4.10}$$

By (4.7), to complete the proof it is enough to show that

$$\mathbb{P}(\alpha \leqslant N_B/\sqrt{n} \leqslant \beta) \to \mathbb{P}(\alpha \leqslant \hat{E}/b \leqslant \beta).$$

This follows easily, since for any $x > 0$

$$\mathbb{P}(N_B > x\sqrt{n}) = \prod_{k=1}^{x\sqrt{n}} \left(1 - \frac{b\sqrt{n}}{n-1-k}\right) \sim \mathrm{e}^{-bx},$$

as $n \to \infty$. □

4.3. Poisson limit for the number of vertices with large minimal edge weights.

The aim of this section is to understand the distribution of edges emanating from the slow vertices, namely the set of vertices for which the closest vertex is at distance $\approx \log n$. For vertex $i \in [n]$, let $X_{(i)} = \min_{j\in[n]} E_{ij}$ denote the minimal edge weight emanating from a given vertex $i \in [n]$. Fix $\alpha \in \mathbb{R}$ and let $N_n(\alpha) = \#\{i \in [n] : X_{(i)} \geqslant \log n - \alpha\}$ denote the number of vertices with minimal outgoing edge weight at least $\log n - \alpha$. We prove the following Poisson limit for $N_n(\alpha)$.

Proposition 4.3 (number of vertices with large minimal edge weight). *As $n \to \infty$,*

$$N_n(\alpha) \xrightarrow{\mathrm{w}} N(\alpha), \tag{4.11}$$

where $N(\alpha)$ is a Poisson random variable with mean e^{α}. More precisely, for n large enough,

$$d_{\mathrm{TV}}(N_n(\alpha), N(\alpha)) \leqslant \frac{2(1+\varepsilon_n)\mathrm{e}^{2\alpha}\log n}{n}, \tag{4.12}$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

where d_{TV} denotes the total variation distance and

$$\varepsilon_n = \exp\left(\frac{\log n - \alpha}{n}\right) - 1.$$

Proof. We use the Stein–Chen method for Poisson approximation. Write

$$N_n(\alpha) = \sum_{i\in[n]} Z_i, \quad Z_i = \mathbb{1}\{X_{(i)} \geqslant \log n - \alpha\}.$$

For fixed $i \in [n]$, note that $X_{(i)}$ has an exponential distribution with mean $n/(n-1)$. Writing $p_n = \mathbb{P}(Z_i = 1)$ so that $\lambda := \mathbb{E}(N_n(\alpha)) = np_n$, it is easy to check that

$$\mathbb{E}(N_n(\alpha)) = (1+\varepsilon_n)\mathrm{e}^{\alpha}. \tag{4.13}$$

Thus, $\lambda \to \mathrm{e}^{\alpha}$ as $n \to \infty$. For each fixed $i \in [n]$, suppose we can couple $N_n(\alpha)$ with a random variable W_i' such that the marginal distribution of W_i' is

$$W_i' + 1 \stackrel{d}{=} N_n(\alpha)\big|_{\{Z_i=1\}}, \tag{4.14}$$

that is, $W_i' + 1$ has the same distribution as $N_n(\alpha)$ conditionally on $\{Z_i = 1\}$. Then Stein–Chen theory [9] implies that in total variation distance

$$d_{\mathrm{TV}}(\mathcal{L}(N_n(\alpha)), \mathrm{Poi}(\lambda)) \leqslant (1\wedge\lambda^{-1}) \sum_{i\in[n]} \mathbb{E}(Z_i)\,\mathbb{E}(|N_n(\alpha) - W_i'|). \tag{4.15}$$

Let us describe W_1', the same construction switching indices works for any i. Let

$$\mathscr{S}_n := \{\mathcal{K}_n, (E_e)_{e\in\mathcal{E}_n}\}$$

be the original edge lengths and let $N_n(\alpha)$ be defined as above for the random metric space $\mathscr{S}_n$. Let us construct the edge lengths of $\mathcal{K}_n$ conditional on the event $\{Z_1 = 1\}$ so that $X_{(1)} - \log n \geqslant -\alpha$. We shall write $\mathscr{S}_n' := \{\mathcal{K}_n, (E_e')_{e\in\mathcal{E}_n}\}$ for $\mathscr{S}_n$ conditioned on this event. Note that this event only affects edges incident to vertex 1 and further, by the lack of memory property of the exponential distribution, every such edge incident to vertex 1 has distribution $\log n - \alpha + E$ where E is an exponential mean n random variable, independently across edges. Thus, we can construct the edge lengths on $\mathscr{S}_n'$ using the edge lengths E_e in $\mathscr{S}_n$ with the following description:

(a) for each edge $e = \{1,i\}$ incident to vertex i, set $E_e' = \log n - \alpha + E_e$,
(b) for any edge not incident to vertex 1, set $E_e' = E_e$.

Define $X_{(i)}'$ analogously to $X_{(i)}$ as the minimal edge length incident to vertex i but in $\mathscr{S}_n'$. Finally, define

$$Z_i' := \mathbb{1}\{X_{(i)}' > \log n - \alpha\}, \quad W_1' = \sum_{v\neq 1} \mathbb{1}\{X_{(v)}' \geqslant \log n - \alpha\}.$$

Then W_1' by construction has the required distribution in (4.14). Note that

$$|N_n(\alpha) - W_1'| \leqslant \mathbb{1}\{X_{(1)} > \log n - \alpha\} + \sum_{i\neq 1} |Z_i - Z_i'|.$$

Taking expectations, by symmetry,

$$\mathbb{E}(|N_n(\alpha) - W_1'|) \leqslant p_n + (n-1)\,\mathbb{E}\,|Z_2 - Z_2'|. \tag{4.16}$$

Now

$$\mathbb{E}|Z_2 - Z_2'| = \mathbb{P}(Z_2 = 1, Z_2' = 0) + \mathbb{P}(Z_2 = 0, Z_2' = 1).$$

Since the edge lengths in $\mathscr{S}_n'$ are at least as large as the edge lengths in $\mathscr{S}_n$, we have $\{Z_2 = 1, Z_2' = 0\} = \varnothing$. For the second term,

$$\{Z_2 = 0, Z_2' = 1\} \equiv \{E_{2,1} < \log n - \alpha, \min_{j \neq 1,2} E_{2,j} \geqslant \log n - \alpha\}.$$

Since $E_{i,j}$ are exponential mean n, we immediately get

$$\mathbb{P}(Z_2 = 0, Z_2' = 1) \leqslant \frac{(1+o(1))\mathrm{e}^{\alpha} \log n}{n^2}.$$

Using this in (4.16), the total variation bound (4.15) completes the proof. □

4.4. Joint convergence of distances between multiple vertices

The aim of this section is to understand the re-centred asymptotic joint distribution of the minimal weight between multiple vertices. To prove this, it turns out that Process 2 using the smallest-weight graph $\mathsf{SWG}_t^{(v)}$ from vertices $v \in [n]$ is more useful than Process 1. Versions of Proposition 4.4 below have appeared in [4, 5, 10]. We give a new proof, both for completeness and since we need a variant of this argument below.

Fix $m \geqslant 2$. Let $(\Lambda_\alpha)_{\alpha \in [m]}$ and $(\Lambda_{\alpha\beta})_{\alpha,\beta \in [m], s<t}$ be independent standard Gumbel random variables. In the following proposition, we identify the limiting distribution of $(d_w(\alpha,\beta) - \log n)_{\alpha,\beta \in [m], \alpha<\beta}$, an extension of the result given in (3.2) proved by Janson [18] for $m = 2$.

Proposition 4.4 (joint distances between many vertices). *As $n \to \infty$,*

$$(d_w(\alpha,\beta) - \log n)_{\alpha,\beta \in [m], \alpha<\beta} \xrightarrow{\mathrm{w}} (\Lambda_\alpha + \Lambda_\beta - \Lambda_{\alpha\beta})_{\alpha,\beta \in [m], \alpha<\beta}. \tag{4.17}$$

Proof. Fix $m \geqslant 2$. Write

$$D(m) := (\Lambda_\alpha + \Lambda_\beta - \Lambda_{\alpha\beta})_{\alpha,\beta \in [m], \alpha<\beta}, \tag{4.18}$$

for the limiting array. The idea of the proof is as follows. We start by sequentially growing the smallest-weight graphs, SWGs, from the m vertices until they meet. This gives us a sequence of collision times $(T_{\alpha\beta})_{\alpha<\beta \in [m]}$. An appropriately chosen linear transformation of these collision times stochastically dominates the array of the lengths of shortest paths. We show that this linear transformation of the collision times converges to the array $D(m)$. A simple limiting argument using the convergence of the marginal distribution of two point distances implies that the joint distribution of the distances themselves converge to $D(m)$ and this completes the proof.

Let us now begin the proof. We write $\mathscr{S}_n$ for the random metric space $(\mathcal{K}_n, (E_e)_{e \in \mathcal{E}_n})$, where once again we remind the reader that E_e are i.i.d. exponential random variables with mean n. Now start the smallest weight cluster $\mathsf{SWG}_t^{(1)}$ from vertex $\alpha = 1$. Write

$$T_1 = \inf\{t \colon |\mathsf{SWG}_t^{(1)}| = \sqrt{n}\} \tag{4.19}$$

for the time for $\mathsf{SWG}_t^{(1)}$ to grow to size $\sqrt{n}$. Now

$$T_1 \stackrel{d}{=} \sum_{k=1}^{\sqrt{n}} nE_k/[k(n-k)],$$

where $\{E_k\}_{k\geqslant 1}$ is a sequence of independent rate one exponential random variables. This implies (see (4.8) and (4.10)) that

$$T_1 - \frac{1}{2}\log n \xrightarrow{\text{w}} \log(1/\hat{E}_1), \tag{4.20}$$

where $\hat{E}_1$ is exponential with mean 1. For every vertex $v \in \mathsf{SWG}_t^{(1)}$, write $B^{(1)}(v) := d_w(1,v)$ for the time when the flow from vertex 1 reaches v. We now work conditionally on the flow cluster $\mathsf{SWG}_{T_1}^{(1)}$. By construction, as $n\to\infty$,

$$\mathbb{P}(2 \notin \mathsf{SWG}_{T_1}^{(1)}) = 1 - \frac{\sqrt{n}}{n} \to 1. \tag{4.21}$$

Further, by the memoryless property of the exponential distribution, conditionally on $\mathsf{SWG}_{T_1}^{(1)}$, for every boundary edge $e=\{u,v\}$ with $u\in \mathsf{SWG}_{T_1}^{(1)}$ and $v \notin \mathsf{SWG}_{T_1}^{(1)}$, the remaining edge length $E_e - (T_1 - B^{(1)}(u))$ has an exponential distribution with mean n, and all these remaining edge lengths are independent.

Freeze the cluster $\mathsf{SWG}_{T_1}^{(1)}$. Start a flow from vertex 2 as the source and write $\mathsf{SWG}_t^{(2)}$ for the smallest-weight graph. Write

$$T_{12} := \inf\{t : \mathsf{SWG}_t^{(2)} \cap \mathsf{SWG}_{T_1}^{(1)} \neq \varnothing\}, \tag{4.22}$$

so that T_{12} is the first time that a vertex in the flow cluster from vertex $\alpha=1$ at time T_1 is hit by the flow cluster from 2. Conditionally on $\mathsf{SWG}_{T_1}^{(1)}$, on the event $\{2\notin \mathsf{SWG}_{T_1}^{(1)}\}$, we have that

(a) the length of the smallest-weight path between 1 and 2 is given by $d_w(1,2) = T_1 + T_{12}$;

(b) the random variable T_{12} has the same distribution as $d_w(\{1\},B)$ in the random (unconditional) metric space $\mathscr{S}_n$ where B is a fixed set of size $\sqrt{n}$.

By Lemma 4.2 with $b=1$ we immediately get

$$\left(T_{12} - \frac{1}{2}\log n, |\mathsf{SWG}_{T_{12}}^{(2)}|/\sqrt{n}\right) \xrightarrow{\text{w}} (\log(1/\hat{E}_2) + \log(\hat{E}_{12}), \hat{E}_{12}), \tag{4.23}$$

where $\hat{E}_2$ and $\hat{E}_{12}$ are independent of each other and independent of $\hat{E}_1$, which arises as the limit in (4.20). Combining (4.20) and (4.23) we get

$$\begin{aligned}(d_w(1,2) - \log n, |\mathsf{SWG}_{T_{12}}^{(2)}|/\sqrt{n}) &= \left(T_{12} - \frac{1}{2}\log n + T_1 - \frac{1}{2}\log n, N/\sqrt{n}\right)\\ &\xrightarrow{\text{w}} (\log(1/\hat{E}_1) + \log(1/\hat{E}_2) + \log(\hat{E}_{12}), \hat{E}_{12}).\end{aligned} \tag{4.24}$$

This proves the claim for $m=2$. We next extend the computation to $m=3$.

For ease of notation, write $\mathcal{B} = \sqrt{n} = |\mathsf{SWG}_{T_1}^{(1)}|$ and $\mathcal{R} = |\mathsf{SWG}_{T_{12}}^{(2)}|$, here $\mathcal{B}$ and $\mathcal{R}$ will be mnemonics for 'black' and 'red' respectively. We now work conditionally on $\mathcal{A} := \mathsf{SWG}_{T_1}^{(1)} \cup \mathsf{SWG}_{T_{12}}^{(2)}$. Since $|\mathcal{A}| = \Theta_{\mathbb{P}}(\sqrt{n})$,

$$\mathbb{P}(3 \notin \mathsf{SWG}_{T_1}^{(1)} \cup \mathsf{SWG}_{T_{12}}^{(2)}) \to 1 \quad \text{as } n\to\infty. \tag{4.25}$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

Freeze the above two flow clusters. Start a flow from vertex $\beta = 3$ and consider the smallest-weight graph $\mathsf{SWG}_t^{(3)}$ emanating from vertex 3. We need to modify this process after the first time it finds a vertex in $\mathcal{A} = \mathsf{SWG}_{T_1}^{(1)} \cup \mathsf{SWG}_{T_{12}}^{(2)}$, namely after time

$$T_3^* = \inf\{t : \mathsf{SWG}_t^{(3)} \cap \mathcal{A} \neq \varnothing\}.$$

Suppose this happens due to $\mathsf{SWG}_{T_3}^{(3)}$ finding a vertex in $\mathsf{SWG}_{T_1}^{(1)}$. Remove all vertices in $\mathsf{SWG}_{T_1}^{(1)}$ and all adjacent edges from $\mathcal{K}_n$ and then continue until the process finds a vertex in $\mathsf{SWG}_{T_{12}}^{(2)}$. Similarly, if this happens due to a vertex in $\mathsf{SWG}_{T_{12}}^{(2)}$ being found, then remove all vertices in $\mathsf{SWG}_{T_{12}}^{(2)}$ and continue. Although this is not quite the smallest-weight graph emanating from vertex 3, to minimize notational overhead, we shall continue to denote this modified process by the same $\{\mathsf{SWG}_t^{(3)}\}_{t\geqslant 0}$. Define the stopping times

$$T_{13} = \inf\{t \geqslant 0 : \mathsf{SWG}_t^{(3)} \cap \mathsf{SWG}_{T_1}^{(1)} \neq \varnothing\},$$

and

$$T_{23} = \inf\{t \geqslant 0 : \mathsf{SWG}_t^{(3)} \cap \mathsf{SWG}_{T_{12}}^{(2)} \neq \varnothing\}.$$

Similarly, define the sizes of the cluster $\mathsf{SWG}_t^{(3)}$ at these stopping times as

$$C_n^{(13)} = |\mathsf{SWG}_{T_{13}}^{(3)}|, \quad C_n^{(23)} = |\mathsf{SWG}_{T_{23}}^{(3)}|. \tag{4.26}$$

Similar to the urn description in (4.3), it is easy to check that conditionally on $\mathcal{A}$ and on the event $\{3 \notin \mathcal{A}\}$, the distribution of the random variables $(T_{13}, T_{23}, C_n^{(13)}, C_n^{(23)})$ can be constructed as follows. Consider an urn with n balls out of which $\mathcal{B} = |\mathsf{SWG}_{T_1}^{(1)}|$ black balls, $\mathcal{R} = |\mathsf{SWG}_{T_{12}}^{(2)}|$ red balls and the remaining $n - \mathcal{B} - \mathcal{R}$ white balls. Also let $(E_k)_{k\geqslant 1}$ be an independent sequence of mean n exponential random variables. Start drawing balls at random without replacement till the first time $\mathcal{N}_1$ we get either a black or a red ball.

(a) Suppose the first ball amongst the black or red balls is a black ball. Remove all black balls so that there are now $(n - \mathcal{N}_1 - \mathcal{B})$ balls in the urn. Continue drawing balls without replacement till we get a red ball. Let $\mathcal{N}_2 > \mathcal{N}_1$ be the time for the first pick of a red ball. Let $C_n^{(13)} = \mathcal{N}_1$, $C_n^{(23)} = \mathcal{N}_2$. Finally, let

$$T_{13} := \sum_{k=1}^{\mathcal{N}_1} \frac{E_k}{k(n-k)}, \quad T_{23} := T_{13} + \sum_{k=\mathcal{N}_1+1}^{\mathcal{N}_2} \frac{E_k}{k(n-k-\mathcal{B})}, \tag{4.27}$$

where, as before, $(E_k)_{k\geqslant 1}$ is an independent sequence of exponential random variables with mean n.

(b) Suppose the first ball amongst black and red balls to be picked is a red ball. Then, in the above formulae, simply interchange the roles of 1 and 2 and $\mathcal{B}$ and $\mathcal{R}$.

Using (4.23) and arguing exactly as in the proof of Lemma 4.2, we see that

$$\left(\frac{C_n^{(13)}}{\sqrt{n}}, \frac{C_n^{(23)}}{\sqrt{n}}, T_{13} - \frac{1}{2}\log n, T_{23} - \frac{1}{2}\log n\right) \tag{4.28}$$
$$\xrightarrow{\text{w}} (\hat{E}_{13}, \hat{E}_{23}/\hat{E}_{12}, \log(1/\hat{E}_3) + \log(\hat{E}_{13}), \log(1/\hat{E}_3) + \log(\hat{E}_{23}/\hat{E}_{12}).$$

Here $\hat{E}_3, \hat{E}_{13}, \hat{E}_{23}$ are independent of $\hat{E}_1, \hat{E}_2, \hat{E}_{12}$ and i.i.d. exponential mean one random variables. Now note that by construction, there is a path of length $D_n(1,3) := T_1 + T_{13}$ between vertices 1

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

and 3 and similarly of length $D_n(2,3) := T_{12} + T_{23}$ between vertices 2 and 3. Thus, by (4.23) and (4.28)

$$d_w(1,3) - \log n \leqslant T_{13} - \frac{1}{2}\log n + T_1 - \frac{1}{2}\log n \xrightarrow{\text{w}} \log(1/\hat{E}_1) + \log(1/\hat{E}_3) + \log(\hat{E}_{13}), \quad (4.29)$$

and

$$\begin{aligned} d_w(2,3) - \log n \leqslant T_{23} - \frac{1}{2}\log n + T_{12} - \frac{1}{2}\log n \\ \xrightarrow{\text{w}} \log(1/\hat{E}_3) + \log(\hat{E}_{23}/\hat{E}_{12}) + \log(1/\hat{E}_2) + \log(\hat{E}_{12}) \\ = \log(1/\hat{E}_2) + \log(1/\hat{E}_3) + \log(\hat{E}_{23}), \end{aligned} \quad (4.30)$$

Thus the limiting array $D(3)$ in (4.18) is a limiting upper bound in the weak sense for the array $d_n(3) := (d_w(\alpha,\beta) - \log n : 1 \leqslant \alpha < \beta \leqslant 3)$. However, we have equality for $m = 2$ by (4.24). Thus the marginals of $d_n(3)$ converge to the marginals of D as $n \to \infty$. This implies $d_n(3) \xrightarrow{\text{w}} D(3)$ as $n \to \infty$.

This entire construction extends inductively for higher values of m. For example, for $m = 4$, we consider the random sets $\mathcal{C}_1 = \mathsf{SWG}^{(1)}_{T_1}, \mathcal{C}_2 = \mathsf{SWG}^{(2)}_{T_{12}}$ and $\mathcal{C}_3 = \mathsf{SWG}^{(3)}_{T_{(12)3}}$, where $T_{(12)3} = \max(T_{12}, T_{13})$. Then we define $\mathcal{A}$ as the union of the above three sets and proceed with the construction as above. This completes the proof. □

Remark. We learned about this reduction from the sums of collision times to lengths of optimal paths via stochastic domination from [23].

The following is an easy corollary of the proof of the above result. Recall that for any two vertices $\alpha, \beta \in [n]$, $\pi(\alpha,\beta)$ denotes the unique shortest path (geodesic) between them.

Corollary 4.5. *Consider the random metric space $\mathscr{S}_n = (\mathcal{K}_n, \{E_e\}_{e\in\mathcal{E}_n})$. Fix $m \geqslant 2$.*

(a) *Let D_n be the event that $\exists \alpha \neq \beta \neq \gamma \in [m]$ such that $\gamma \in \pi(\alpha,\beta)$. Then $\mathbb{P}(D_n) \to 0$ as $n \to \infty$.*
(b) *Fix $1/2 < \vartheta < 1$. Consider the smallest-weight graphs $\{\mathsf{SWG}^{(i)}_{\vartheta\log n}\}_{i\in[m]}$ from these m vertices at time $\vartheta \log n$. Then w.h.p. the shortest paths $\pi(\alpha,\beta)$ are contained in the union of these balls, that is, as $n \to \infty$,*

$$\mathbb{P}\big(\pi(\alpha,\beta) \subseteq \cup_{i=1}^m \mathsf{SWG}^{(i)}_{\vartheta\log n}\ \forall \alpha,\beta \in [m]\big) \to 1.$$

Proof. Part (a) follows from extending (4.21) and (4.25) to general m. Part (b) follows from the above proof, which proves that for any pair of vertices α,β, $\pi(\alpha,\beta)$ can be found in $\mathsf{SWG}^{(\alpha)}_{r_n} \cup \mathsf{SWG}^{(\beta)}_{r_n}$ where $r_n = \frac{1}{2}\log n + O_{\mathbb{P}}(1)$. □

4.5. Distances between vertices with large minimal edge weight

Fix $\alpha \in \mathbb{R}$. Recall that

$$N_n(\alpha) = \sum_{i=1}^{n} \mathbb{1}\{X_{(i)} \geqslant \log n - \alpha\}$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

denotes the number of vertices with minimum outgoing edge length at least $\log n - \alpha$. Fix $m \geqslant 2$ and condition on the event $N_n(\alpha) = m$. Let $V_1, \dots, V_m$ denote the m vertices for which $X_{(V_i)} \geqslant \log n - \alpha$.

Our aim in this section is to understand, conditionally on the event $\{N_n(\alpha) = m\}$, the asymptotic joint distribution of $(d_w(V_i, V_j) : i < j \in [m])$. Recall the array $D(m)$ from (4.18) giving the asymptotic joint distribution of the re-centred (by $\log n$) length of smallest paths between m typical vertices in $\mathscr{S}_n$. The main aim of this section is to prove the following result.

Proposition 4.6 (distances between vertices with large minimal edge weight). *Fix $\alpha \in \mathbb{R}$ and $m \geqslant 2$. Recall* (4.18). *Conditionally on $N_n(\alpha) = m$, as $n \to \infty$,*

$$(d_w(V_i, V_j) - 3\log n + 2\alpha)_{i,j\in[m],i<j} \xrightarrow{\mathrm{w}} (\Lambda_i + \Lambda_j - \Lambda_{ij})_{i,j\in[m],i<j} = D(m). \tag{4.31}$$

We start with two preparatory results, Lemmas 4.7 and 4.8, which we then use to complete the proof of the proposition at the end of this section. Let us start by disentangling exactly what the conditioning event $\{N_n(\alpha) = m\}$ implies about the edge length distribution. To ease notation, we assume without loss of generality that $V_i = i$. Then this conditioning implies that the edge length distributions can be formulated in the following two rules.

(a) *Translation.* Every edge E'_e incident to one of the vertices in $[m]$ is conditioned to be at least $\log n - \alpha$. By the memoryless property of the exponential distribution, we can write $E'_e = \log n - \alpha + E_e$, where (E_e) is an independent family of mean n independent exponential random variables.

(b) *Conditioning.* For every vertex $i \notin [n] \setminus [m]$, the edges $(E'_{i,j})_{j\notin[m]}$ are independent exponential mean n random variables conditioned on

$$X_{(i),[m+1:n]} := \min_{m+1\leqslant j\leqslant n} E'_{i,j} < \log n - \alpha. \tag{4.32}$$

Let us use our original metric space $\mathscr{S}_n$ to sequentially overlay the effect of the above two effects. Recall that we have used $\pi(i, j)$ for the smallest-weight path between i, j in $\mathscr{S}_n$. The following lemma deals with the effect of the simpler translation event (without dealing with the conditioning), and will be the starting point of our analysis.

Lemma 4.7. *Fix $m \geqslant 1$ and consider the metric space $\mathscr{S}_n$. For every edge e incident to one of the vertices in $[m]$, replace the edge E_e by $E_e + \log n - \alpha$. Leave all other edges unchanged. Call this new metric space $\mathscr{S}'_n(tr)$. Write $\pi'(i, j)$ for the smallest-weight path between i, j and write d'_w for the corresponding metric. Then, for all $i, j \in [m]$, and on the event that $E_{ij} + \log n - \alpha > d_w(i, j)$ for all $i, j \in [m]$,*

$$\pi'(i, j) = \pi(i, j), \quad d'_w(i, j) = d_w(i, j) + 2\log n - 2\alpha. \tag{4.33}$$

In particular,

$$(d'_w(i, j) - 3\log n + 2\alpha)_{i,j\in[m],i<j} \xrightarrow{\mathrm{w}} (\Lambda_i + \Lambda_j - \Lambda_{ij})_{i,j\in[m],i<j}. \tag{4.34}$$

Proof. The distributional convergence follows from (4.33) and Proposition 4.4. Equation (4.33) follows since we can construct the smallest-weight path problem for $\mathscr{S}'_n(\mathrm{tr})$ restricted to $[m]$ as follows. To $\mathscr{S}_n$ adjoin m new vertices $\{i' : i' \in [m]\}$. Each new vertex i' has only one edge, namely, to vertex i of length $\log n - \alpha$. Call this new metric space $\mathscr{S}_n^*$ and the corresponding metric d_w^* and smallest-weight path $\pi^*(\cdot,\cdot)$. Then the metric space $\mathscr{S}_n'(\mathrm{tr})$ can be constructed as follows. For $i, j \in [m]$ let $d_w'(i,j) = d_w^*(i',j')$ and $\pi^*(i',j') = \{i' \rightsquigarrow i\} \cup \pi'(i,j) \cup \{j \rightsquigarrow j'\}$. □

By the translation rule (a) and conditionally on $N_n(\alpha) = m$, the distances $(d_w(V_i,V_j))_{i,j\in[m],i<j}$ are identical in distribution (assuming $E_{V_iV_j} + \log n - \alpha > d_w(V_i,V_j)$ for all $i,j \in [m]$ which occurs w.h.p.) to $(d_w'(i,j) - 3\log n + 2\alpha)_{i,j\in[m],i<j}$ *conditionally* on $N_{n,m}(\alpha) = 0$, where $N_{n,m}(\alpha)$ denotes the number of vertices in $[n]\setminus[m]$ for which $X_{(i),[m+1:n]} \geqslant \log n - \alpha$, where $X_{(i),[m+1:n]}$ is defined in (4.32). In order to transfer (4.34) in Lemma 4.7 to the related statement (4.31) in Proposition 4.6, it suffices to prove that the distance matrix $(d_w(i,j) - \log n)_{i,j\in[m],i<j}$ is *asymptotically independent* of $N_{n,m}(\alpha)$. That is the content of the next lemma.

Lemma 4.8. *Fix $\alpha \in \mathbb{R}$ and $m \geqslant 1$. The matrix of distances $(d_w(i,j) - \log n)_{i,j\in[m],i<j}$ is asymptotically independent of $N_{n,m}(\alpha)$. Consequently, conditionally on $N_{n,m}(\alpha) = 0$,*

$$(d_w(i,j) - \log n)_{i,j\in[m],i<j} \xrightarrow{\mathrm{w}} (\Lambda_i + \Lambda_j - \Lambda_{ij})_{i,j\in[m],i<j}. \tag{4.35}$$

Proof. Fix $\ell \geqslant 0$ and a matrix of weights $(w_{ij})_{i,j\in[m],i<j}$. The statement of the lemma is equivalent to the statement that, for every $(w_{ij})_{i,j\in[m],i<j} \in \mathbb{R}^{m(m-1)/2}$ and $m \geqslant 1, \ell \geqslant 0$,

$$\begin{aligned}\mathbb{P}\big((d_w(i,j) - \log n)_{i,j\in[m],i<j} &\leqslant (w_{ij})_{i,j\in[m],i<j}, N_{n,m}(\alpha) \leqslant \ell\big) \\ &= \mathbb{P}\big((d_w(i,j) - \log n)_{i,j\in[m],i<j} \leqslant (w_{ij})_{i,j\in[m],i<j}\big)\,\mathbb{P}(N_{n,m}(\alpha) \leqslant \ell) + o(1).\end{aligned} \tag{4.36}$$

We prove this by proving both inequalities. We note that $(d_w(i,j) - \log n)_{i,j\in[m],i<j}$, as well as $N_{n,m}(\alpha)$, are (sequences of) *increasing* random variables in the edge weights $(E_e)_e$. Thus, by the FKG-inequality for continuous random variables as stated in [19, Section 6], the left-hand side of (4.36) is always *at least* the first term on the right-hand side of (4.36). Note that the condition [19, inequality (6.6)] holds with equality since we deal with independent random variables. This establishes the lower bound (even without the error term).

We next prove the upper bound. We split the edges into two disjoint sets. Let $\mathcal{E}_1 = \{e\colon E_e > \log n - \alpha\}$ and $\mathcal{E}_2 = \{e\colon E_e \leqslant \log n - \alpha\}$. Clearly, $N_{n,m}(\alpha)$ is determined by the edges in $\mathcal{E}_1$, since $N_{n,m}(\alpha)$ is the number of vertices in $[n]\setminus[m]$ for which every edge incident to it is in $\mathcal{E}_1$. Let $\bar{d}_w(i,j)$ be the distance between i and j only using edges in $\mathcal{E}_2$, so that $\bar{d}_w(i,j) \geqslant d_w(i,j)$. We bound from above

$$\begin{aligned}\mathbb{P}\big((d_w(i,j) - \log n)_{i,j\in[m],i<j} &\leqslant (w_{ij})_{i,j\in[m],i<j}, N_{n,m}(\alpha) \leqslant \ell\big) \\ &\leqslant \mathbb{P}\big((\bar{d}_w(i,j) - \log n)_{i,j\in[m],i<j} \leqslant (w_{ij})_{i,j\in[m],i<j}, N_{n,m}(\alpha) \leqslant \ell\big) \\ &\quad + \mathbb{P}\big((d_w(i,j))_{i,j\in[m],i<j} \neq (\bar{d}_w(i,j))_{i,j\in[m],i<j}\big),\end{aligned} \tag{4.37}$$

and bound both terms from above. In the first term, we note that

$$\{(\bar{d}_w(i,j) - \log n)_{i,j\in[m],i<j} \leqslant (w_{ij})_{i,j\in[m],i<j}\}$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

only depends on the edge weights in $\mathcal{E}_2$, while $\{N_{n,m}(\alpha) \leqslant \ell\}$ only depends on the edge weights in $\mathcal{E}_1$, which are disjoint sets. Thus,

$$\begin{aligned} &\mathbb{P}\big((\bar{d}_w(i,j) - \log n)_{i,j\in[m],i<j} \leqslant (w_{ij})_{i,j\in[m],i<j}, N_{n,m}(\alpha) \leqslant \ell\big) \\ &\quad \leqslant \mathbb{P}\big(\{(\bar{d}_w(i,j) - \log n)_{i,j\in[m],i<j} \leqslant (w_{ij})_{i,j\in[m],i<j}\} \circ \{N_{n,m}(\alpha) \leqslant \ell\}\big). \end{aligned} \tag{4.38}$$

By the BK-inequality applied to continuous random variables (see [16, Theorem 1.2]),

$$\begin{aligned} &\mathbb{P}\big((\bar{d}_w(i,j) - \log n)_{i,j\in[m],i<j} \leqslant (w_{ij})_{i,j\in[m],i<j}, N_{n,m}(\alpha) \leqslant \ell\big) \\ &\quad \leqslant \mathbb{P}\big((\bar{d}_w(i,j) - \log n)_{i,j\in[m],i<j} \leqslant (w_{ij})_{i,j\in[m],i<j}\big)\,\mathbb{P}(N_{n,m}(\alpha) \leqslant \ell). \end{aligned} \tag{4.39}$$

By the fact that $\bar{d}_w(i,j) \geqslant d_w(i,j)$, the right-hand side of (4.39) is at most the product of probabilities on the right-hand side of (4.36). It remains to show that the second term in (4.37) is $o(1)$. By symmetry and the union bound,

$$\mathbb{P}\big((d_w(i,j))_{i,j\in[m],i<j} \neq (\bar{d}_w(i,j))_{i,j\in[m],i<j}\big) \leqslant m^2\,\mathbb{P}(d_w(1,2) \neq \bar{d}_w(1,2)). \tag{4.40}$$

By Proposition 4.4, as well as its proof (see Corollary 4.5(b)), $\mathbb{P}(d_w(1,2) \neq \bar{d}_w(1,2)) = o(1)$, which completes the proof of the asymptotic independence. To prove (4.35), it suffices to show that $\liminf_{n\to\infty} \mathbb{P}(N_{n,m}(\alpha) = 0) > 0$, which follows from Proposition 4.3 together with the fact that $N_{n,m}(\alpha) = N_n(\alpha)$ w.h.p. □

Proof of Proposition 4.6. By the translation rule (a) and the conditioning rule (b), the distribution of $(d_w(V_i,V_j))_{i,j\in[m],i<j}$ conditionally on $N_n(\alpha) = m$ is the same as that of $(d'_w(i,j))_{i,j\in[m],i<j}$ in $\mathscr{S}'_n(\mathrm{tr})$ *conditionally on* $N_{n,m}(\alpha) = 0$. The event that $E_{ij} + \log n - \alpha > d_w(i,j)$ for all $i,j \in [m]$ occurs w.h.p., so the identity in (4.33) holds w.h.p. Combining (4.34) in Lemma 4.7 with the asymptotic independence of $(d_w(i,j))_{i,j\in[m],i<j}$ and $N_{n,m}(\alpha)$ in Lemma 4.8 and the fact that $N_{n,m}(\alpha) \xrightarrow{\mathrm{w}} N(\alpha)$ by Proposition 4.3 completes the proof of Proposition 4.6. □

4.6. Reduction to distances between vertices with large minimal edge weights

The previous section analysed distances between the vertices whose minimal outgoing edge is large (*e.g.* $\log n + O_{\mathbb{P}}(1)$). The distances between these vertices are then close to $3\log n + O_{\mathbb{P}}(1)$. The aim of this section is to show that these are the only vertices that matter for the weight diameter. We achieve this by considering distances between vertices whose minimal outgoing edge is 'small' and showing that the distance between such vertices are not large enough to create the diameter and thus can be ignored.

We start with some notation. Fix $\alpha > 0$ and define

$$R_n(\alpha) = \#\{i,j \in [n] \colon X_{(i)} \leqslant \log n - \alpha, X_{(j)} \leqslant \log n + \alpha/2, d_w(i,j) \geqslant 3\log n - \alpha/8\}. \tag{4.41}$$

The random variable $R_n(\alpha)$ counts the number of ordered pairs of vertices $(i,j) \in [n] \times [n]$ that satisfy that the minimal outgoing edge of vertex i is less than $\log n - \alpha$, the minimal outgoing edge of j is less than $\log n+\alpha/2$ and yet the distance between i,j is greater than $3\log n-\alpha/8$. The following result gives an upper bound on the expected value of $R_n(\alpha)$.

Proposition 4.9 (distances from vertices with small minimal weight). *There exists a constant $C>0$ such that, for all $\alpha>0$,*

$$\limsup_{n\to\infty}\mathbb{E}[R_n(\alpha)]\leqslant C\mathrm{e}^{-\alpha/16}. \tag{4.42}$$

Proof. By the union bound,

$$\mathbb{E}[R_n(\alpha)]\leqslant n^2\,\mathbb{P}(d_w(1,2)\geqslant 3\log n-\alpha/8, X_{(1)}\leqslant \log n-\alpha, X_{(2)}\leqslant \log n+\alpha/2). \tag{4.43}$$

Note that

$$(X_{(1)},X_{(2)})\overset{d}{=}\left(\min\left[\frac{n}{n-2}E_1^*,nE_{12}^*\right],\min\left[\frac{n}{n-2}E_2^*,nE_{12}^*\right]\right),$$

where E_1^*,E_2^*,E_{12}^* are independent exponential random variables with mean 1. Here nE_{12}^* represents the weight of the direct edge between vertices 1,2, while for $i\in\{1,2\}$, $nE_i^*/(n-2)$ represents the minimal outgoing edges from vertex i to the remaining vertices $[n]\setminus\{1,2\}$.

On the event $\{d_w(1,2)\geqslant 3\log n-\alpha/8\}$, we have that $nE_{12}^*\geqslant d_w(1,2)\geqslant 3\log n-\alpha/8$. As a result, when $d_w(1,2)\geqslant 3\log n-\alpha/8$, *unless*

$$\max\left(\frac{n}{n-2}E_1^*,\frac{n}{n-2}E_2^*\right)>3\log n-\alpha/8, \tag{4.44}$$

we have that

$$(X_{(1)},X_{(2)})=\left(\frac{n}{n-2}E_1^*,\frac{n}{n-2}E_2^*\right). \tag{4.45}$$

The probability of the event in (4.44) is bounded by $3\mathrm{e}^{\alpha/8}/n^3$ for large n. Since $n^2\mathrm{e}^{\alpha/8}/n^3\to 0$, we can ignore the contribution of this in the proof of Proposition 4.9 and assume (4.45).

Let V_1 be the closest vertex to 1, at distance $X_{(1)}$ (respectively V_2 at distance $X_{(2)}$ from vertex 2). The rest of the smallest-weight path has the same distribution as the smallest-weight path between two sets $A=\{1,V_1\}$ and $B=\{2,V_2\}$ in $\mathscr{S}_n$. Lemma 4.1 thus implies that

$$d_w(i,j)=X_{(1)}+X_{(2)}+\sum_{k=2}^{N-1}\frac{nE_k'}{k(n-k)}, \tag{4.46}$$

where $N=N_1\wedge N_2$ and (N_1,N_2) is a uniform pair of distinct vertices from $[n]\setminus\{1,2\}$ and $(E_k')_{k\geqslant 1}$ are mean one exponential random variables. The distribution of N in Lemma 4.1 was given as the number of draws required to draw the first black ball in an urn containing $N-4$ white and 2 black balls; it is easy to check that this is equivalent to the description above. Writing

$$S_N=\sum_{k=2}^{N-1}\frac{nE_k}{k(n-k)},$$

we get

$$\mathbb{E}[R_n(\alpha)]\leqslant n^2\,\mathbb{P}\big(S_N\geqslant 3\log n-X_{(1)}-X_{(2)}-\alpha/8, X_{(1)}\leqslant \log n-\alpha, X_{(2)}\leqslant \log n+\alpha/2\big). \tag{4.47}$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

Thus,

$$\mathbb{E}[R_n(\alpha)] \leqslant n^2 \int_0^{\log n-\alpha} \int_0^{\log n+\alpha/2} \mathrm{e}^{-(x+y)(n-2)/n}\,\mathbb{P}(S_N \geqslant 3\log n - x - y - \alpha/8)\,\mathrm{d}x\,\mathrm{d}y. \tag{4.48}$$

To complete the proof, we study the tail behaviour of the random variable S_N.

Lemma 4.10 (tail behaviour for random sums). *For any constant $a<2$, there exists a constant $C=C_a<\infty$ such that, for every $x \geqslant 0$,*

$$\mathbb{P}(S_N \geqslant \log n + x) \leqslant C\mathrm{e}^{-ax}. \tag{4.49}$$

Assuming the lemma for the time being, we use this to complete the proof of Proposition 4.9. Using Lemma 4.10 with $a=3/2$ gives

$$\begin{aligned}\mathbb{E}[R_n(\alpha)] &\leqslant Cn^2 \int_0^{\log n-\alpha} \int_0^{\log n+\alpha/2} \mathrm{e}^{-(x+y)}\mathrm{e}^{-a(2\log n - x - y-\alpha/8)}\,\mathrm{d}x\,\mathrm{d}y \\ &= Cn^{2(1-a)} \int_0^{\log n-\alpha} \int_0^{\log n+\alpha/2} \mathrm{e}^{(a-1)(x+y)}\mathrm{e}^{\alpha/8}\,\mathrm{d}x\,\mathrm{d}y \\ &= C\mathrm{e}^{-(a-1)\alpha/2+a\alpha/8} \leqslant C\mathrm{e}^{-\alpha/16}.\end{aligned} \tag{4.50}$$

This completes the proof of Proposition 4.9. □

Proof of Lemma 4.10. We compute the moment generating function of S_N as

$$\begin{aligned}M_{S_N}(a) &= \sum_{j=2}^{n-2} \mathbb{P}(N=j)\,\mathbb{E}[\mathrm{e}^{aS_j}] = \sum_{j=2}^{n-2} \mathbb{P}(N=j) \prod_{k=2}^{j-1} \frac{k(n-k)}{k(n-k)-an} \\ &= \sum_{j=2}^{n-2} \mathbb{P}(N=j)\,\exp\biggl(-\sum_{k=2}^{j-1} \log\biggl(1-\frac{an}{k(n-k)}\biggr)\biggr).\end{aligned} \tag{4.51}$$

Thus,

$$\begin{aligned}\mathbb{P}(S_N \geqslant \log n + x) &\leqslant \mathrm{e}^{-a(\log n + x)} M_{S_N}(a) \\ &\leqslant \mathrm{e}^{-a(\log n + x)} \sum_{j=2}^{n-2} \mathbb{P}(N=j)\,\exp\biggl(-\sum_{k=2}^{j-1} \log\biggl(1-\frac{an}{k(n-k)}\biggr)\biggr).\end{aligned} \tag{4.52}$$

Using $a<2$, we have $an/[k(n-k)]<1$, since $k, n-k \geqslant 2$. Therefore, via Taylor expansion,

$$\log\biggl(1-\frac{an}{k(n-k)}\biggr) \leqslant \frac{an}{k(n-k)} + O\biggl(\frac{n^2}{[k(n-k)]^2}\biggr), \tag{4.53}$$

Using

$$\frac{n}{k(n-k)} = \frac{1}{k} + \frac{1}{n-k},$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

we arrive at

$$\begin{aligned}\mathbb{P}(S_N \geqslant \log n + x) &\leqslant \mathrm{e}{-a(\log n+x)} M_{S_N}(a)\\ &\leqslant C\mathrm{e}^{-a(\log n+x)} \sum_{j=2}^{n-2} \mathbb{P}(N=j)\exp\biggl(a\sum_{k=2}^{j-1}\biggl[\frac{1}{k}+\frac{1}{n-k}\biggr]\biggr)\\ &\leqslant C\mathrm{e}^{-ax}\sum_{j=2}^{n-2}\mathbb{P}(N=j)\mathrm{e}^{a[\log(j/n)-\log(1-j/n)]}\\ &= C\mathrm{e}^{-ax}\,\mathbb{E}\biggl[\biggl(\frac{N/n}{1-N/n}\biggr)^a\biggr].\end{aligned}$$

Note that

$$\mathbb{P}(N=j) = \frac{2(n-j)}{(n-2)(n-3)},$$

so that, by dominated convergence,

$$\mathbb{E}\biggl[\biggl(\frac{N/n}{1-N/n}\biggr)^a\biggr] = \sum_{j=2}^{n-2}\frac{2(n-j)}{(n-2)(n-3)}\biggl(\frac{j/n}{1-j/n}\biggr)^a \to \int_0^1 \frac{u^a}{(1-u)^a} 2(1-u)\,\mathrm{d}u < \infty, \tag{4.54}$$

whenever $a < 2$. □

4.7. The limiting random variable

In this section we prove the finiteness of the random variable $\Xi = \max_{s<t}(Y_s + Y_t - \Lambda_{st})$ in (2.3), which Theorem 2.1 asserts is the limit of the re-centred diameter. In the following lemma we give an alternative expression for its distribution.

Lemma 4.11 (the limiting random variable). *Let $Q = \mathrm{e}^{-\Xi}$. Then*

$$Q = \min_{s<t} \frac{S_s S_t}{E'_{st}}, \tag{4.55}$$

where $S_s = \sum_{i=1}^s E'_i$ and $(E'_i)_{i\geqslant 1}$ and $(E'_{st})_{s<t}$ are i.i.d. exponential random variables with mean 1. In particular, for every $x > 0$,

$$\mathbb{P}(Q > x) = \mathbb{E}\biggl[\prod_{1\leqslant s<t} l(1-\mathrm{e}^{-S_s S_t/x})\biggr], \tag{4.56}$$

and $\mathbb{P}(Q > x) \in (0,1)$ for every $x > 0$. Further, $\lim_{x\downarrow 0}\mathbb{P}(Q > x) = 1$.

Proof. We note that we can write $-\Lambda_{st} = \log(E'_{st})$ and $Y_s = -\log(S_s)$. Indeed, the point process $(\mathrm{e}^{-Y_s})_{s\geqslant 1}$ is a standard Poisson process. Thus,

$$\mathrm{e}^{-\Xi} \stackrel{d}{=} \min_{s<t} \mathrm{e}^{\log(S_s)+\log(S_t)-\log(E'_{st})} = Q. \tag{4.57}$$

Equation (4.56) immediately follows. To prove that $\mathbb{P}(Q > x) \in (0,1)$ for every $x > 0$, we note that $\mathbb{P}(Q > x) < 1$ follows immediately from (4.56) since each of the terms in the product is < 1

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

a.s. To show that $\mathbb{P}(Q > x) > 0$, we first note that

$$\mathbb{P}(Q > x) \geqslant \mathbb{E}\bigg[\prod_{1 \leqslant s < t}(1 - \mathrm{e}^{-S_s S_t/x}) \mathbb{1}_{\{S_1 > 1\}}\bigg]$$
$$= \mathbb{E}\bigg[\prod_{1 \leqslant s < t}(1 - \mathrm{e}^{-S_s S_t/x}) \mid S_1 > 1\bigg] \mathbb{P}(S_1 > 1). \tag{4.58}$$

We compute that $\mathbb{P}(S_1 > 1) = 1/\mathrm{e}$, and observe that by the memoryless property of the exponential random variable S_1, conditionally on $S_1 > 1$, the distribution of $(S_t)_{t \geqslant 1}$ is equal to $(S_t + 1)_{t \geqslant 1}$. Thus,

$$\mathbb{P}(Q > x) \geqslant \mathrm{e}^{-1}\,\mathbb{E}\bigg[\prod_{1 \leqslant s < t}(1 - \mathrm{e}^{-(S_s+1)(S_t+1)/x})\bigg]$$
$$\geqslant \mathrm{e}^{-1} \exp\bigg(\sum_{1 \leqslant s < t} \mathbb{E}\big[\log(1 - \mathrm{e}^{-(S_s+1)(S_t+1)/x})\big]\bigg). \tag{4.59}$$

Next, we compute, using Fubini,

$$\sum_{1 \leqslant s < t} \mathbb{E}\big[\log(1 - \mathrm{e}^{-(S_s+1)(S_t+1)/x})\big] \tag{4.60}$$
$$= \sum_{1 \leqslant s < t} \int_0^\infty \mathrm{d}u \int_0^\infty \mathrm{d}v \frac{u^{s-1}}{(s-1)!} \frac{v^{t-s-1}}{(t-s-1)!} \mathrm{e}^{-(u+v)} \log(1 - \mathrm{e}^{-(u+1)(u+v+1)/x})$$
$$= \int_0^\infty \mathrm{d}u \int_0^\infty \mathrm{d}v \sum_{1 \leqslant s < t} \frac{u^{s-1}}{(s-1)!} \frac{v^{t-s-1}}{(t-s-1)!} \mathrm{e}^{-(u+v)} \log(1 - \mathrm{e}^{-(u+1)(u+v+1)/x})$$
$$= \int_0^\infty \int_0^\infty \log(1 - \mathrm{e}^{-(u+1)(u+v+1)/x})\,\mathrm{d}u\,\mathrm{d}v < \infty.$$

This proves that $\mathbb{P}(Q > x) > 0$ for every $x > 0$. Similarly, for any $\varepsilon > 0$, by conditioning on $S_1 \geqslant \varepsilon$,

$$\mathbb{P}(Q > x) \geqslant \mathrm{e}^{-\varepsilon}\,\mathbb{E}\bigg[\prod_{1 \leqslant s < t}(1 - \mathrm{e}^{-(S_s+\varepsilon)(S_t+\varepsilon)/x})\bigg] \tag{4.61}$$
$$\geqslant \mathrm{e}^{-\varepsilon} \exp\bigg(\sum_{1 \leqslant s < t} \mathbb{E}\big[\log(1 - \mathrm{e}^{-(S_s+\varepsilon)(S_t+\varepsilon)/x})\big]\bigg)$$
$$= \mathrm{e}^{-\varepsilon} \exp\bigg(\int_0^\infty \int_0^\infty \log(1 - \mathrm{e}^{-(u+\varepsilon)(u+v+\varepsilon)/x})\,\mathrm{d}u\,\mathrm{d}v\bigg)$$
$$= \mathrm{e}^{-\varepsilon} \exp\bigg(\varepsilon^{-2} \int_0^\infty \int_0^\infty \log(1 - \mathrm{e}^{-(u'+1)(u'+v'+1)/(x/\varepsilon^2)})\,\mathrm{d}u'\,\mathrm{d}v'\bigg),$$

where the last equality follows by substituting $(u', v') = (\varepsilon u, \varepsilon v)$. For every $\varepsilon > 0$, as $x \searrow 0$, the integrand converges pointwise to zero and thus, by monotone convergence, the integral also converges to 0. Therefore, for every $\varepsilon > 0$,

$$\liminf_{x \downarrow 0} \mathbb{P}(Q > x) \geqslant \mathrm{e}^{-\varepsilon}, \tag{4.62}$$

and letting $\varepsilon \searrow 0$, this completes the proof that $\lim_{x \downarrow 0} \mathbb{P}(Q > x) = 1$. □

4.8. The limiting maximization problem

In this section we combine the various ingredients proved in the previous sections to prove the distributional convergence in Theorem 2.1. We defer the proof of the convergence of moments to the next section. By Proposition 4.3, first note that given any $\varepsilon > 0$, one can choose α large but finite (independent of n) such that for all n large, $\mathbb{P}(N_n(\alpha) \geqslant 2) \geqslant 1-\varepsilon$. Thus by Proposition 4.6, with probability greater than $1-\varepsilon$ for all large n,

$$\mathrm{Diam}_w(K_n) - 3\log n \geqslant d_w(V_1,V_2) - 3\log n \xrightarrow{\mathrm{w}} -2\alpha + \Lambda_1 + \Lambda_2 - \Lambda_{12}. \tag{4.63}$$

As a result the sequence $\{(\mathrm{Diam}_w(K_n) - 3\log n)_- : n \geqslant 1\}$ is tight. Further, using Proposition 4.9, given $\varepsilon > 0$, one can choose $\alpha = \alpha(\varepsilon)$ large enough such that with probability $\geqslant 1-\varepsilon$ for all large n,

$$\mathrm{Diam}_w(K_n) = \max_{s<t\leqslant N_n(\alpha)} d_w(V_s,V_t). \tag{4.64}$$

We note that, again using Propositions 4.6 and 4.3,

$$\max_{s<t\leqslant N_n(\alpha)} d_w(V_s,V_t) - 3\log n \xrightarrow{\mathrm{w}} \max_{s<t\leqslant N(\alpha)} (\Lambda_s + \Lambda_t - \Lambda_{st} - 2\alpha), \tag{4.65}$$

where $N(\alpha)$ is a Poisson random variable with mean e^{α} and the Gumbel variables are independent of $N(\alpha)$. As a result,

$$\mathrm{Diam}_w(\mathcal{K}_n) - 3\log n \xrightarrow{\mathrm{w}} \Xi^*, \tag{4.66}$$

where Ξ^* is the distributional limit as $\alpha \to \infty$ of the right-hand side of (4.65), that is,

$$\max_{s<t\leqslant N(\alpha)} (\Lambda_s + \Lambda_t - \Lambda_{st} - 2\alpha) \xrightarrow{\mathrm{w}} \Xi^*. \tag{4.67}$$

Theorem 4.12 below shows that this weak limit exists and that $\Xi^* = \Xi$ defined in (2.3). This completes the proof of the first assertion of the main theorem, thus showing that

$$\mathrm{Diam}_w(\mathcal{K}_n) - 3\log n \xrightarrow{\mathrm{w}} \Xi.$$

Theorem 4.12 (the limiting variable Ξ). *As $\alpha \to \infty$,*

$$\max_{s<t\leqslant N(\alpha)} (\Lambda_s + \Lambda_t - \Lambda_{st} - 2\alpha) \xrightarrow{\mathrm{w}} \Xi, \tag{4.68}$$

where Ξ is defined in (2.3).

Proof. First note that as $\alpha \to \infty$,

$$\mathrm{e}^{-\alpha} N(\alpha) \xrightarrow{\mathrm{P}} 1. \tag{4.69}$$

For each fixed $\alpha > 0$, define the random variables

$$\hat{\Xi}_\alpha := \max_{s<t\leqslant N(\alpha)} (\Lambda_s + \Lambda_t - \Lambda_{st} - 2\alpha), \quad \Xi_\alpha := \max_{s<t\leqslant \mathrm{e}^{\alpha}} (\Lambda_s + \Lambda_t - \Lambda_{st} - 2\alpha).$$

It then suffices to prove that

$$\Xi_\alpha := \max_{s<t\leqslant \mathrm{e}^{\alpha}} (\Lambda_s + \Lambda_t - \Lambda_{st} - 2\alpha) \xrightarrow{\mathrm{w}} \Xi. \tag{4.70}$$

Then, for each $\varepsilon > 0$, one can sandwich $\hat{\Xi}_\alpha$ between $\Xi_{\alpha-\varepsilon}$ and $\Xi_{\alpha+\varepsilon}$ (with high probability as $\alpha \to \infty$ by (4.69)), and then first let $\alpha \to \infty$ and then $\varepsilon \to 0$ to get the corresponding result for $\hat{\Xi}_\alpha$.

Recall from Section 2, that the Poisson point process $\mathcal{P} = (Y_s)_{s \geqslant 1}$ with intensity measure given by the density function $\lambda(y) = \mathrm{e}^{-y}$. Also recall from (2.3) that we have defined Ξ as

$$\Xi := \max_{s<t}(Y_s + Y_t - \Lambda_{st}).$$

For any fixed $A > 0$, let $\mathcal{P}(A)$ denote $\mathcal{P}$ restricted to the interval $[-A, \infty)$. Write

$$\Xi(A) := \max_{s<t:\, Y_s, Y_t \in \mathcal{P}(A)} (Y_s + Y_t - \Lambda_{st}).$$

Thus, $\Xi(A)$ is the maximum of corresponding pairs (s,t) whose point process values satisfy $Y_s, Y_t \geqslant -A$. Intuitively, one would expect that $\Xi = \Xi(A)$ for large A. We now make his intuition precise. Define

$$\mathcal{R}^{(1)}(A) := \max_{s<t:\, Y_s, Y_t \leqslant -A} (Y_s + Y_t - \Lambda_{st}),$$

and, for $A < B$, let

$$\mathcal{R}^{(2)}(A,B) := \max_{s<t:\, Y_s \geqslant -A, Y_t \leqslant -(A+B)} (Y_s + Y_t - \Lambda_{st}).$$

The random variable $\mathcal{R}^{(1)}(A)$ is the supremum between pairs (s,t) such that $Y_s, Y_t \leqslant -A$ while $\mathcal{R}^{(2)}(A,B)$ corresponds to the supremum between pairs of points (s,t) such that $Y_s > -A$ but $Y_t < -(A+B)$. Note that, for any z,

$$\{\Xi = \Xi(A+B)\} \supseteq \{\Xi(A) > z, \mathcal{R}^{(1)}(A) < z, \mathcal{R}^{(2)}(A,B) < z\}. \tag{4.71}$$

Consider the point process

$$\mathcal{P}^*_\alpha = \sum_{s=1}^{\mathrm{e}^\alpha} \delta\{\Lambda_s - \alpha\}.$$

When arranged in increasing order, write this point process as $Y_1(\alpha) > Y_2(\alpha) > \cdots$. Standard extreme value theory implies that

$$\mathcal{P}^*_\alpha \xrightarrow{\mathrm{w}} \mathcal{P} \quad \text{as } \alpha \to \infty, \tag{4.72}$$

where $\xrightarrow{\mathrm{w}}$ denotes convergence in distribution in the space of point measures on $\mathbb{R}$ equipped with the vague topology. Analogously to $\Xi(A), \mathcal{R}^{(1)}(A), \mathcal{R}^{(2)}(A,B)$, we define the random variables $\Xi_\alpha(A), \mathcal{R}^{(1)}_\alpha(A), \mathcal{R}^{(2)}_\alpha(A,B)$, that is,

$$\Xi_\alpha(A) := \max_{s<t:\, Y_s(\alpha), Y_t(\alpha) \in \mathcal{P}_A(\alpha)} (Y_s(\alpha) + Y_t(\alpha) - \Lambda_{st}),$$

where $\mathcal{P}_\alpha(A)$ is the point process $\mathcal{P}_\alpha$ restricted to the interval $[-A, \infty)$. We define $\mathcal{R}^{(1)}_\alpha(A), \mathcal{R}^{(2)}_\alpha(A)$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

similarly. As before, for any z,

$$\{\Xi_\alpha = \Xi_\alpha(A+B)\} \supseteq \{\Xi_\alpha(A) > z, \mathcal{R}^{(1)}_\alpha(A) < z, \mathcal{R}^{(2)}_\alpha(A,B) < z\}. \tag{4.73}$$

The weak convergence in (4.72) immediately implies that, for any fixed A,

$$\Xi_\alpha(A) \xrightarrow{\text{w}} \Xi(A) \quad \text{as } \alpha \to \infty. \tag{4.74}$$

The following lemma formalizes the notion that for large A, $\Xi = \Xi(A)$ and, similarly, when α is large, $\Xi_\alpha(A) = \Xi_\alpha$. Let us give an intuitive description of how this is achieved. The argument proceeds by showing that for large A, each of the random variables $\mathcal{R}^{(1)}(A), \mathcal{R}^{(1)}_\alpha(A)$, and, for each fixed A, for sufficiently large B, $\mathcal{R}^{(2)}(A,B), \mathcal{R}^{(2)}_\alpha(A,B)$ take large negative values.

Proposition 4.13.

(a) *Fix $x \in \mathbb{R}$. Then*

$$\limsup_{A\to\infty} \mathbb{P}(\mathcal{R}^{(1)}(A) > x) = 0.$$

Further, for each fixed A,

$$\limsup_{B\to\infty} \mathbb{P}(\mathcal{R}^{(2)}(A,B) > x) = 0.$$

(b) *Fix $x \in \mathbb{R}$. Then,*

$$\limsup_{A\to\infty} \limsup_{\alpha\to\infty} \mathbb{P}(\mathcal{R}^{(1)}_\alpha(A) > x) = 0.$$

Further, for each fixed A,

$$\limsup_{B\to\infty} \limsup_{\alpha\to\infty} \mathbb{P}(\mathcal{R}^{(2)}_\alpha(A,B) > x) = 0.$$

We prove the statements in Proposition 4.13 one by one.

Proof of Proposition 4.13(a). We start with $\mathcal{R}^{(1)}(A)$. To simplify notation, we also restrict ourselves to the case $x = 0$. The general x case is identical.

Write

$$\mathcal{N}^{(1)}(A) := \#\{(s,t) : Y_s, Y_t < -A, Y_s + Y_t - \Lambda_{st} \geqslant 0\}.$$

It is enough to show $\limsup_{A\to\infty} \mathbb{E}(\mathcal{N}^{(1)}(A)) = 0$. Conditioning on the point process $\mathcal{P}$, we get

$$\mathbb{E}(\mathcal{N}^{(1)}(A)|\mathcal{P}) = \sum_{(s,t),s<t,Y_s,Y_t<-A} \mathrm{e}^{-\mathrm{e}^{-(Y_s+Y_t)}}.$$

Fix $a > 1$. We use the fact that we can choose A so large such that $\mathrm{e}^{-\mathrm{e}^{C+D}} < \mathrm{e}^{-aC}\mathrm{e}^{-aD}$ for all $C, D > A$. This leads to

$$\mathbb{E}(\mathcal{N}^{(1)}(A)|\mathcal{P}) \leqslant \sum_{(s,t),s<t,Y_s,Y_t<-A} \mathrm{e}^{aY_s}\mathrm{e}^{aY_t}.$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

Since $\{Y_s \in \mathcal{P} : Y_s \leqslant -A\}$ is just a Poisson point process on the interval $(-\infty, -A]$ with density $\lambda(x) = \mathrm{e}^{-x}$, properties of Poisson processes [20, Eqn 3.14] implies that, as $A \to \infty$,

$$\mathbb{E}\Bigg(\sum_{\substack{(s,t),s<t,\\ Y_s,Y_t<-A}} \mathrm{e}^{aY_s}\mathrm{e}^{aY_t}\Bigg) = \frac{1}{2}\bigg(\int_{-\infty}^{-A} \mathrm{e}^{ax}\mathrm{e}^{-x}\,\mathrm{d}x\bigg)^2 = \frac{1}{2}\mathrm{e}^{-2(a-1)A} \to 0.$$

This shows that $\limsup_{A\to\infty} \mathbb{E}(\mathcal{N}^{(1)}(A)) = 0$ and thus completes the proof.

Next fix A and let us deal with $\mathcal{R}^{(2)}(A,B)$. Here we use the fact that $\mathcal{P}(A)$ and $\mathcal{P}^c(A+B) := \mathcal{P} \setminus \mathcal{P}^c(A+B)$ are independent Poisson point processes on the sets $[-A,\infty)$ and $(-\infty, -(A+B))$ with intensity measure with density $\lambda(y) = \mathrm{e}^{-y}$. We condition on $\mathcal{P}(A)$. Fix a point Y_s in $\mathcal{P}(A)$. Then

$$\mathbb{P}\bigg(\sup_{Y_t<-(A+B)} (Y_s + Y_t - \Lambda_{st}) < z|\mathcal{P}(A)\bigg) = \mathbb{E}\bigg(\prod_{t:\, Y_t<-(A+B)} (1-\exp(-e^{-(Y_t-(z-Y_s))}))\bigg). \tag{4.75}$$

Thus to complete the proof, it is enough to show, for any z^* and A,

$$\lim_{B\to\infty} \mathbb{E}\bigg(\prod_{t:Y_t<-(A+B)} (1-\exp(-e^{-(Y_t-(z-Y_s))}))\bigg) \to 1. \tag{4.76}$$

By the dominated convergence theorem, it is enough to show that, as $B \to \infty$,

$$\prod_{t:Y_t<-(A+B)} (1-\exp(-e^{-(Y_t-(z-Y_s))})) \xrightarrow{\mathrm{P}} 1.$$

Taking logarithms, this is equivalent to showing that, as $B \to \infty$,

$$\sum_{t:Y_t<-(A+B)} \log\left(1-\exp(-e^{-(Y_t-(z-Y_s))})\right) \xrightarrow{\mathrm{P}} 0.$$

In turn, this is equivalent to showing that, as $B \to \infty$,

$$\sum_{t:Y_t<-(A+B)} \exp(-e^{-(Y_t-(z-Y_s))}) \xrightarrow{\mathrm{P}} 0.$$

By Campbell's theorem [20],

$$\begin{aligned}\mathbb{E}\bigg(\sum_{t:Y_t<-(A+B)} \exp^{-\mathrm{e}^{-(Y_t-z^*)}}\bigg) &= \int_{-\infty}^{-(A+B)} \exp(-\mathrm{e}^{-(y-z^*)})\mathrm{e}^{-y}\,\mathrm{d}y \\ &= \mathrm{e}^{z^*}\exp(-\mathrm{e}^{A+B+z^*}) \to 0,\end{aligned}$$

as $B \to \infty$. This completes the proof of (4.76) and thus of part (a) of Proposition 4.13. □

Proof of Proposition 4.13(b). This closely follows the proof of part (a). Without giving a full proof, we mainly highlight the differences. We again start with $\mathcal{R}^{(1)}_\alpha(A)$ and again restrict ourselves to the case $x = 0$. The general x case is identical.

Write

$$\mathcal{N}^{(1)}_\alpha(A) := \#\{(s,t) : Y_s(\alpha), Y_t(\alpha) < -A, Y_s(\alpha) + Y_t(\alpha) - \Lambda_{st} \geqslant 0\}.$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

It is enough to show that

$$\limsup_{A\to\infty}\limsup_{\alpha\to\infty}\mathbb{E}(\mathcal{N}_\alpha^{(1)}(A))=0.$$

Conditioning on the point process $\mathcal{P}_\alpha^*$, we now get

$$\begin{aligned}\mathbb{E}(\mathcal{N}_\alpha^{(1)}(A)|\mathcal{P}_\alpha^*) &= \sum_{(s,t),s<t,Y_s(\alpha),Y_t(\alpha)<-A}\exp(-\mathrm{e}^{-(Y_s(\alpha)+Y_t(\alpha))}) \\ &= \sum_{1\leqslant s<t\leqslant \mathrm{e}^\alpha}\mathbb{1}_{\{\Lambda_s,\Lambda_t<-A+\alpha\}}\exp(-\mathrm{e}^{-(\Lambda_s-\alpha)-(\Lambda_t-\alpha)}).\end{aligned} \tag{4.77}$$

Taking expectations and using that Λ_s,Λ_t are independent for $s<t$ leads to

$$\mathbb{E}(\mathcal{N}_\alpha^{(1)}(A))\leqslant\int_{-\infty}^{-A+\alpha}\int_{-\infty}^{-A+\alpha}\mathrm{e}^{-(u-\alpha)}\exp(-\mathrm{e}^{-u})\,\mathrm{e}^{-(v-\alpha)}\exp(-\mathrm{e}^{-v})\exp(-\mathrm{e}^{-(u-\alpha)-(v-\alpha)})\,\mathrm{d}u\,\mathrm{d}v. \tag{4.78}$$

This integral can be bounded by

$$\mathbb{E}(\mathcal{N}_\alpha^{(1)}(A))\leqslant\int_{-\infty}^{-A}\int_{-\infty}^{-A}\mathrm{e}^{-u}\mathrm{e}^{-v}\exp(-\mathrm{e}^{-u-v})\,\mathrm{d}u\,\mathrm{d}v, \tag{4.79}$$

which is independent of α and converges to 0 as $A\to\infty$. The proof for $\mathcal{R}_\alpha^{(2)}(A,B)$ is similar and will be omitted. □

Completing the proof of Theorem 4.12. Proposition 4.13 shows that given any $\varepsilon>0$ one can choose finite $\alpha=\alpha(\varepsilon)$ and $A=A(\varepsilon)$ such that $\mathbb{P}(\Xi=\Xi(A))$ and $\mathbb{P}(\Xi_\alpha(A)=\Xi_\alpha)$ are both at least $1-\varepsilon$. Using (4.74), (4.71) and (4.73) completes the proof of Theorem 4.12. □

4.9. Convergence of moments

Recall that $C_{ij}=d_w(i,j)$. We need to show that

$$\mathbb{E}\Big[\max_{i,j\in[n]}C_{ij}\Big]-3\log n\to\mathbb{E}[\Xi],\qquad \mathrm{Var}\Big(\max_{i,j\in[n]}C_{ij}\Big)\to\mathrm{Var}(\Xi).$$

Since we have already shown convergence in distribution, by uniform integrability for any $p\geqslant1$, to prove that

$$\mathbb{E}\Big[\Big(\max_{i,j\in[n]}C_{ij}-3\log n\Big)^p\Big]\to\mathbb{E}[\Xi^p], \tag{4.80}$$

it suffices to prove that, for some integer q with $q>p/2$,

$$\mathbb{E}\Big[\Big(\max_{i,j\in[n]}C_{ij}-3\log n\Big)^{2q}\Big]=O(1). \tag{4.81}$$

Combined with convergence in distribution, this implies convergence of the moments as well as existence of the moments of the limit random variable Ξ. Note that

$$\mathbb{E}\Big[\Big(\max_{i,j\in[n]}C_{ij}-3\log n\Big)^{2q}\Big]=\mathbb{E}\Big[\Big(\max_{i,j\in[n]}C_{ij}-3\log n\Big)_+^{2q}\Big]+\mathbb{E}\Big[\Big(\max_{i,j\in[n]}C_{ij}-3\log n\Big)_-^{2q}\Big]. \tag{4.82}$$

We start by analysing the first term on the right-hand side of (4.82) by deriving an upper bound on $\max_{i,j\in[n]}C_{ij}-3\log n$, and then prove a lower bound on $\max_{i,j\in[n]}C_{ij}-3\log n$ to obtain a bound on the second term on the right-hand side of (4.82).

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

Upper bound. Let us analyse the first term in (4.82) and show that

$$\mathbb{E}\left[\left(\max_{i,j\in[n]} C_{ij} - 3\log n\right)_+^{2q}\right] = O(1).$$

To prove this assertion, it is enough to show that there exist N,α such that, for all large $n > N$ and $x \geqslant \alpha$, the random variable $\max_{i,j\in[n]} C_{ij} - 3\log n$ has exponential upper tails in the sense that there exist constants $\kappa_1, \kappa_2 > 0$ (independent of x) such that

$$\mathbb{P}\left(\max_{i,j\in[n]} C_{ij} - 3\log n > x\right) \leqslant \kappa_1 \mathrm{e}^{-\kappa_2 x}. \tag{4.83}$$

Now note that

$$\mathbb{1}\left\{\max_{i,j\in[n]} C_{ij} - 3\log n > x\right\} \leqslant \mathbb{1}\left\{\max_{i\in[n]} X_{(i)} > \log n + 4x\right\} + R_n^{(1)}(x) + R_n^{(2)}(x). \tag{4.84}$$

Here $R_n^{(1)}(x) = R_n(8x)$ as in (4.41), that is,

$$R_n^{(1)}(x) = \#\{i,j\in[n]\colon X_{(i)} \leqslant \log n - 8x, X_{(j)} \leqslant \log n + 4x, d_w(i,j) \geqslant 3\log n - x\},$$

while

$$R_n^{(2)}(x) := \#\{(i,j)\colon X_{(i)} > \log n - 8x, X_{(j)} > \log n - 8x, d_w(i,j) > 3\log n + x\}.$$

Recall that, for any $\alpha\in\mathbb{R}$, $N_n(\alpha)$ denotes the number of vertices i with $X_{(i)} \geqslant \log n - \alpha$. For the first term in (4.84), since

$$\mathbb{P}\left(\max_{i\in[n]} X_{(i)} > \log n + 4x\right) = \mathbb{P}(N_n(-4x) \geqslant 1),$$

the Poisson approximation in Proposition 4.3 implies that for x large,

$$\begin{aligned}\mathbb{P}\left(\max_{i\in[n]} X_{(i)} > \log n + 4x\right) &\leqslant \frac{2(1+o(1))\mathrm{e}^{-4x}\log n}{n} + (1-\exp(-\mathrm{e}^{-4x}))\\ &\leqslant (1+o(1))\mathrm{e}^{-4x}.\end{aligned} \tag{4.85}$$

Further, by Proposition 4.9 for n large enough

$$\mathbb{E}(R_n^{(1)}(x)) \leqslant C\mathrm{e}^{-x/2}. \tag{4.86}$$

It remains to analyse $R_n^{(2)}(x)$. Arguing as in the proof of Proposition 4.9,

$$\mathbb{E}(R_n^{(2)}(x)) \leqslant \mathbb{E}(N_n^2(-8x))\,\mathbb{P}(d_w(1,2) > \log n + 17x),$$

where $d_w(1,2)$ is the distance between vertices $1,2$ in $\mathscr{S}_n = \{\mathcal{K}_n, (E_e)_{e\in\mathcal{E}_n}\}$. By Lemma 4.1,

$$d_w(1,2) \stackrel{d}{=} \sum_{k=1}^{N} \frac{E_j}{k(n-k)},$$

where N is uniform on $[n-1]$ independent of $(E_j)_{j\in[n-1]}$ which are mean n exponential random variables. Thus, by Markov's inequality, for any $\alpha > 0$,

$$\mathbb{P}(d_w(1,2) - \log n > 17x) \leqslant \mathrm{e}^{-17\alpha x}\sum_{j=1}^{n-1} \frac{1}{n-1}\exp\left(\alpha\left[\log\frac{j}{n} - \log\left(1-\frac{j}{n}\right)\right]\right).$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

Letting $\beta = 1-\varepsilon$ with $\varepsilon > 0$ small but independent of x, n, we finally get

$$\mathbb{P}(d_w(1,2) - \log n > 17x) \leqslant (1+o(1))\mathrm{e}^{-17\alpha x}\,\mathbb{E}\left(\left[\frac{U}{1-U}\right]^{1-\varepsilon}\right),$$

where $U \sim U[0,1]$. We need to now bound $\mathbb{E}(N_n^2(-8x))$. Write $N_n(-8x) = \sum_{i=1}^n Z_i$ where $Z_i = \mathbb{1}\{X_{(i)} \geqslant \log n + 8x\}$. By Proposition 4.3, $\mathbb{E}(N_n(-8x)) \leqslant 2\mathrm{e}^{8x}$. Further,

$$\mathrm{Var}(N_n(-8x)) \leqslant 2\mathrm{e}^{8x} + n(n-1)\,\mathbb{P}(Z_1 = 1)[\mathbb{P}(Z_2 = 1|Z_1 = 1) - \mathbb{P}(Z_2 = 1)].$$

Given $Z_1 = 1$, the edge weights $(E_{2,i})_{i\neq 2}$ have the same distribution as

$$(\log n - 8x + E_{2,1}, (E_{2,j})_{j\neq 1,2}).$$

Thus,

$$\mathbb{P}(Z_2 = 1|Z_1 = 1) = \mathbb{P}\Big(\min_{j\geqslant 2} E_{2,j} > \log n - 8x\Big) = \exp\left(-\frac{n-2}{n}(\log n - 8x)\right).$$

Combining this, we get that $\mathrm{Var}(N_n(-8x)) \leqslant 4\mathrm{e}^{8x}$ so that $\mathbb{E}([N_n(-8x)]^2) \leqslant 16\mathrm{e}^{16x}$. This results in

$$\mathbb{E}(R_n^{(2)}(x)) \leqslant (1+o(1))16\,\mathbb{E}\left(\left[\frac{U}{1-U}\right]^{1-\varepsilon}\right)\mathrm{e}^{-(1-17\varepsilon)x}. \tag{4.87}$$

Combining (4.85), (4.86) and (4.87) completes the proof of the asserted exponential tail bound in (4.83) and completes the proof of the upper bound. □

Lower bound. Let us now show that

$$\mathbb{E}\left[\Big(\max_{i,j\in[n]} C_{ij} - 3\log n\Big)_-^{2q}\right] = O(1).$$

Recall that V_1, V_2 denote the vertices with the largest and second largest $X_{(i)}$ values. Further

$$\max_{i,j\in[n]} C_{ij} - 3\log n \geqslant_{st} (X_{(V_1)} - \log n)_- + (X_{(V_2)} - \log n)_- + (d_w(1,2) - \log n),$$

where $d_w(1,2)$ is independent of $X_{(V_i)}$ with the same distribution as the length of the optimal path between $1,2$ in $\mathscr{S}_n$ and $\geqslant_{st}$ denotes stochastic domination. By Hölder's inequality,

$$\begin{aligned}\mathbb{E}\left[\Big(\max_{i,j\in[n]} C_{ij} - 3\log n\Big)_-^{2q}\right] \leqslant 3^{2q}\Big(&\mathbb{E}\big([X_{(V_1)} - \log n]_-^{2q}\big) + \mathbb{E}\big([X_{(V_2)} - \log n]_-^{2q}\big)\\ &+ \mathbb{E}\big([d_w(1,2) - \log n]^{2q}\big)\Big).\end{aligned} \tag{4.88}$$

By [18, Proof of Theorem 3.3],

$$\mathbb{E}\big([d_w(1,2) - \log n]^{2q}\big) = O(1).$$

Further,

$$\mathbb{E}\big([X_{(V_1)} - \log n]_-^{2q}\big) \leqslant \mathbb{E}\big([X_{(V_2)} - \log n]_-^{2q}\big).$$

First recall the identity

$$\mathbb{E}(Y^{2q}) = 2q\int_0^\infty y^{2q-1}\,\mathbb{P}(Y > y)\,\mathrm{d}y, \tag{4.89}$$

https://doi.org/10.1017/S0963548317000232 Published online by Cambridge University Press

for any non-negative random variable Y. Second note that

$$[X_{(V_2)} - \log n]_- \equiv 0, \quad \text{if } X_{(V_2)} \geqslant \log n. \tag{4.90}$$

Suppose we show for some $0 < \varepsilon < 1$ small enough and fixed:

$$\mathbb{P}(\log n - X_{(V_2)} \geqslant x) \leqslant 2\exp^{-(1-\varepsilon)e^x} + 2\frac{e^{2x}\log n}{n}, \quad x \leqslant 2^{-1}(1-\varepsilon)\log n. \tag{4.91}$$

Then this bound implies that

$$\mathbb{P}(\log n - X_{(V_2)} \geqslant x) \leqslant \frac{C\log n}{n^{\varepsilon}}, \quad x \in [2^{-1}(1-\varepsilon)\log n, \log n]. \tag{4.92}$$

Using these bounds on the tail probabilities in the moment identity (4.89) with $Y = [X_{V_2} - \log n]_-$ shows that

$$\mathbb{E}\big([X_{(V_2)} - \log n]_-^{2q}\big) = O(1),$$

which completes the proof. Thus we are now left with proving (4.91). However this follows from the Poisson approximation result in Proposition 4.3 since $\mathbb{P}(\log n - X_{(V_2)} \geqslant x) = \mathbb{P}(N_n(x) \leqslant 1)$. □

Acknowledgements

The work of RvdH is supported by the Netherlands Organisation for Scientific Research (NWO) through VICI grant 639.033.806 and the Gravitation NETWORKS grant 024.002.003. The work of SB has been supported in part by NSF-DMS grant 1105581,1310002, 160683, 161307, SES grant 1357622 and in part by an ARO grant. SB thanks the hospitality of EURANDOM and an NWO STAR grant. We thank Júlia Komjáthy for a careful reading of an early version of the paper. We would like to thank the anonymous referee for spotting a subtle conditioning issue in a previous version of the paper which led to substantial improvements in the presentation of the proof. We would like to thank the editor for a number of other suggestions that significantly improved the exposition of the paper.

References

[1] Addario-Berry, L. Broutin, N. and Lugosi, G. (2010) The longest minimum-weight path in a complete graph. *Combin. Probab. Comput.* **19** 1–19.

[2] Aldous, D. (1992) Asymptotics in the random assignment problem. *Probab. Theory Rel. Fields* **93** 507–534.

[3] Aldous, D. J. (2001) The $\zeta(2)$ limit in the random assignment problem. *Random Struct. Alg.* **18** 381–418.

[4] Aldous, D. J. (2010) More uses of exchangeability: Representations of complex random structures. In *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman* (N. H. Bingham and C. M. Goldie, eds), London Mathematical Society Lecture Note Series, Cambridge University Press, pp. 35–63.

[5] Aldous, D. J. and Bhamidi, S. (2010) Edge flows in the complete random-lengths network. *Random Struct. Alg.* **37** 271–311.

[6] Aldous, D. and Steele, J. M. (2004) The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures* (H. Kesten, ed.), Vol. 110 of Encyclopaedia of Mathematical Sciences, Springer, pp. 1–72.

[7] Amini, H. and Lelarge, M. (2015) The diameter of weighted random graphs. *Ann. Appl. Probab.* **25** 1686–1727.

[8] Amini, H. and Peres, Y. (2014) Shortest-weight paths in random regular graphs. *SIAM J. Discrete Math.* **28** 656–672.

[9] Barbour, A. D., Holst, L. and Janson, S. (1992) *Poisson Approximation*, Vol. 2 of Oxford Studies in Probability, The Clarendon Press, Oxford University Press.

[10] Bhamidi, S. (2008) First passage percolation on locally treelike networks I: Dense random graphs. *J. Math. Phys.* **49** 125218.

[11] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2010) First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Probab.* **20** 1907–1965.

[12] Devroye, L. (1987) Branching processes in the analysis of the heights of trees. *Acta Informatica* **24** 277–298.

[13] Ding, J., Kim, J. H., Lubetzky, E. and Peres, Y. (2010) Diameters in supercritical random graphs via first passage percolation. *Combin. Probab. Comput.* **19** 729–751.

[14] Ding, J., Kim, J. H., Lubetzky, E. and Peres, Y. (2011) Anatomy of a young giant component in the random graph. *Random Struct. Alg.* **39** 139–178.

[15] Frieze, A. M. (1985) On the value of a random minimum spanning tree problem. *Discrete Appl. Math.* **10** 47–56.

[16] Goldstein, L. and Rinott, Y. (2007) Functional BKR inequalities, and their duals, with applications. *J. Theoret. Probab.* **20** 275–293.

[17] Janson, S. (1995) The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph. *Random Struct. Alg.* **7** 337–355.

[18] Janson, S. (1999) One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combin. Probab. Comput.* **8** 347–361.

[19] Kemperman, J. H. B. (1977) On the FKG-inequality for measures on a partially ordered space. *Indag. Math.* **39** 313–331.

[20] Kingman, J. F. C. (1993) *Poisson Processes*, Vol. 3 of Oxford Studies in Probability, The Clarendon Press, Oxford University Press.

[21] Pittel, B. (1994) Note on the heights of random recursive trees and random m-ary search trees. *Random Struct. Alg.* **5** 337–347.

[22] Riordan, O. and Wormald, N. (2010) The diameter of sparse random graphs. *Combin. Probab. Comput.* **19** 835–926.

[23] Salez, J. (2013) Joint distribution of distances in large random regular networks. *J. Appl. Probab.* **50** 861–870.

[24] Smythe, R. T. and Mahmoud, H. M. (1994) A survey of recursive trees. *Teor. Ĭmovīr. Mat. Stat.* **51** 1–29.

[25] Wästlund, J. (2010) The mean field traveling salesman and related problems. *Acta Mathematica* **204** 91–150.