Combinatorics, Probability and Computing (2010) 19, 729–751. © Cambridge University Press 2010
doi:10.1017/S0963548310000301

Diameters in Supercritical Random Graphs Via First Passage Percolation

JIAN DING[1], JEONG HAN KIM[2†], EYAL LUBETZKY[3]
and YUVAL PERES[3]

[1]Department of Statistics, UC Berkeley, Berkeley, CA 94720, USA
(e-mail: jding@stat.berkeley.edu)

[2]Department of Mathematics, Yonsei University, Seoul 120-749, Korea
and
National Institute for Mathematical Sciences, Daejeon 305-340, Korea
(e-mail: jehkim@yonsei.ac.kr)

[3]Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA
(e-mail: eyal@microsoft.com, peres@microsoft.com)

Received 27 June 2009; revised 10 August 2010

We study the diameter of $\mathcal{C}_1$, the largest component of the Erdős–Rényi random graph $\mathcal{G}(n,p)$ in the emerging supercritical phase, *i.e.*, for $p=\frac{1+\varepsilon}{n}$ where $\varepsilon^3 n\to\infty$ and $\varepsilon=o(1)$. This parameter was extensively studied for fixed $\varepsilon>0$, yet results for $\varepsilon=o(1)$ outside the critical window were only obtained very recently. Prior to this work, Riordan and Wormald gave precise estimates on the diameter; however, these did not cover the entire supercritical regime (namely, when $\varepsilon^3 n\to\infty$ arbitrarily slowly). Łuczak and Seierstad estimated its order throughout this regime, yet their upper and lower bounds differed by a factor of $\frac{1000}{7}$.

We show that throughout the emerging supercritical phase, *i.e.*, for any $\varepsilon=o(1)$ with $\varepsilon^3 n\to\infty$, the diameter of $\mathcal{C}_1$ is with high probability asymptotic to $D(\varepsilon,n)=(3/\varepsilon)\log(\varepsilon^3 n)$. This constitutes the first proof of the asymptotics of the diameter valid throughout this phase. The proof relies on a recent structure result for the supercritical giant component, which reduces the problem of estimating distances between its vertices to the study of passage times in first-passage percolation. The main advantage of our method is its flexibility. It also implies that in the emerging supercritical phase the diameter of the 2-core of $\mathcal{C}_1$ is w.h.p. asymptotic to $\frac{2}{3}D(\varepsilon,n)$, and the maximal distance in $\mathcal{C}_1$ between any pair of kernel vertices is w.h.p. asymptotic to $\frac{5}{9}D(\varepsilon,n)$.

1. Introduction

The Erdős–Rényi random graph $\mathcal{G}(n,p)$ is perhaps the most fundamental random graph model, and its rich behaviour has been studied in numerous papers since its introduction in 1959 [11]. One of the most famous phenomena exhibited by this model is the *double jump* at the critical $p=1/n$. As discovered by Erdős and Rényi in their celebrated papers from the 1960s, for $p=c/n$

† Research by J. H. Kim was supported by a Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (CRI, no. 2008-0054850).

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

with c fixed, the largest component $\mathcal{C}_1$ has size $O(\log n)$ with high probability (w.h.p.) when $c < 1$, its size is w.h.p. linear in n for $c > 1$, and for $c = 1$ its size has order $n^{2/3}$ (the latter was fully established only decades later by Bollobás [7] and Łuczak [15]). Furthermore, Bollobás found that the critical behaviour extends throughout the regime where $p = (1 \pm \varepsilon)/n$ for $\varepsilon = O(n^{-1/3})$, known as the *critical window* (or scaling window).

Despite the intensive study of this model, one of its key features – the diameter of the largest component – remained unknown in a regime just beyond criticality, namely for $p = (1 + \varepsilon)/n$, where $\varepsilon = o(1)$ and $\varepsilon^3 n \to \infty$ arbitrarily slowly. Our main results determine the asymptotic behaviour of the diameter throughout this regime, as well as the diameter of the 2-core and the maximal distance in it between kernel vertices (see definitions below).

Theorem 1.1. *Let $\mathcal{C}_1$ be the largest component of the random graph $\mathcal{G}(n, p)$ with $p = (1 + \varepsilon)/n$, where $\varepsilon^3 n \to \infty$ and $\varepsilon = o(1)$. Then w.h.p.*

$$\operatorname{diam}(\mathcal{C}_1) = \frac{3 + o(1)}{\varepsilon} \log(\varepsilon^3 n).$$

A key advantage of our method over alternative approaches is its flexibility, demonstrated by the following theorem. Recall that the 2-*core* of $\mathcal{C}_1$, denoted by $\mathcal{C}_1^{(2)}$, is the maximal subgraph of $\mathcal{C}_1$ where every vertex has degree at least degree 2. The *kernel* $\mathcal{K}$ is the multigraph obtained from $\mathcal{C}_1^{(2)}$ by contracting every 2-path (a path where all interior vertices have degree 2) into an edge.

Theorem 1.2. *Let $\mathcal{C}_1^{(2)}$ be the 2-core of the largest component $\mathcal{C}_1$ of $\mathcal{G}(n, p)$ with $p = (1 + \varepsilon)/n$, where $\varepsilon^3 n \to \infty$ and $\varepsilon = o(1)$. Let $\mathcal{K}$ denote the kernel of $\mathcal{C}_1^{(2)}$. Then w.h.p.*

$$\operatorname{diam}(\mathcal{C}_1^{(2)}) = \frac{2 + o(1)}{\varepsilon} \log(\varepsilon^3 n), \tag{1.1}$$

$$\max_{u,v \in \mathcal{K}} \operatorname{dist}_{\mathcal{C}_1}(u, v) = \frac{5/3 + o(1)}{\varepsilon} \log(\varepsilon^3 n). \tag{1.2}$$

(In the above statements, the term w.h.p. denotes a probability tending to 1 as $n \to \infty$, while the $o(1)$-terms denote functions going to 0 as $n \to \infty$.)

Theorem 1.1 completes a long list of studies of the diameter in sparse Erdős–Rényi random graphs. In the subcritical case, $p = (1 - \varepsilon)/n$, where for $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$, Łuczak [16] obtained a precise estimate of $\log_{1/(1-\varepsilon)}(2\varepsilon^3 n) + O(1/\varepsilon)$ for the largest diameter of a component (including the limiting distribution of the additive $O(1/\varepsilon)$-term). In the critical window, it was shown in [19] that the diameter of $\mathcal{C}_1$ has order $n^{1/3}$. See also the recent work [1] studying the limiting distribution within the critical window.

However, analysing the diameter in the supercritical case is considerably more delicate. The asymptotics of this parameter in the fully supercritical regime, $p = (1 + \varepsilon)/n$, where $\varepsilon > 0$ is fixed, were obtained in [12] (see also [9, 8]). Results for the regime $\varepsilon = o(1)$ were only obtained fairly recently. In 2008, Riordan and Wormald [21] proved accurate estimates of the diameter for most of this regime, but did not cover the entire range where the random graph emerges from the critical window (*i.e.*, $\varepsilon^3 n \to \infty$ arbitrarily slowly). Note that, while the gap that remained was extremely small, the authors stated that 'our method does seem to need some concrete lower

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

bound on $\varepsilon^3 n$ tending to infinity as a function of n'. Łuczak and Seierstad [17] gave estimates for the diameter that do apply to the entire supercritical regime, yet their upper and lower bounds differ by a factor of $\frac{1000}{7}$.

Remarks. **(1)** Following the completion of this work, Riordan and Wormald [22] managed to extend their analysis to the entire supercritical regime, thus obtaining a version of Theorem 1.1 with more accurate error-term estimates. Referring to the present work, they stated in [22] that 'Seeing this paper stimulated us to remove the unnecessary restriction on ε.' We emphasize that the proofs in [21, 22] rely on branching process analysis and are quite different from our methods.

(2) By our results and the duality of the supercritical and subcritical regimes, it follows that in the setting of Theorem 1.1, the diameter of $\mathcal{C}_1$ is w.h.p. the largest diameter of a component – larger by an asymptotic factor of 3 compared to the largest diameter of any other component. Compare this to the subcritical case $p = (1 - \varepsilon)/n$, where the largest component $\mathcal{C}_1$ has diameter of order $(1/\varepsilon)\sqrt{\log(\varepsilon^3 n)}$ w.h.p. (since $\mathcal{C}_1$ is w.h.p. a tree whose size has order $\varepsilon^{-2}\log(\varepsilon^3 n)$, and conditional on this it is uniformly distributed on all trees of this size).

In order to establish Theorems 1.1 and 1.2, we apply a recent structure result proved in a companion paper [10]. This result translates the supercritical giant component into a contiguous tractable model constructed in three steps as follows (see Theorems 4.1 and 5.1 for the precise formulation).

(1) Select a multigraph $\mathcal{K}$ uniformly among all graphs with a prescribed size and degree sequence, where almost all degrees are 3.
(2) Replace the edges of $\mathcal{K}$ by paths of i.i.d. geometric lengths.
(3) Attach a Poisson–Galton–Watson tree to each vertex.

Note that step (1) above constructs the kernel, step (2) gives the 2-core and step (3) produces the entire giant component.

Using this tool, the problem of estimating distances in $\mathcal{C}_1$ is reduced to the study of passage times in first-passage percolation (see, *e.g.*, [14] for further information on this thoroughly studied topic). For instance, one may readily deduce from known results [4] on first-passage percolation that the typical distance between kernel vertices in the 2-core is asymptotically $(1/\varepsilon)\log(\varepsilon^3 n)$ (in fact, the limiting distribution of this typical distance is completely determined in [4]); see [10, Corollary 2].

However, in weighted random graphs, maximal distances exhibit a behaviour different from typical distances. In order to prove our main results, we establish sharp large-deviation estimates for these distance variables.

The rest of this paper is organized as follows. Section 2 contains a few preliminary facts required for the proofs. In Section 3 we analyse typical and maximal distances in weighted random regular graphs, as well as in metric graphs (obtained by replacing each weighted edge by a line segment with the corresponding length). Section 4 contains the proofs of the main theorems in the special case of $\varepsilon = o(n^{-1/4})$, where the description of the giant component has a particularly elegant form (Theorem 4.1). We extend these results to the general case of any $\varepsilon = o(1)$ in Section 5.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

2. Preliminaries and notation

A random d-regular $G \sim \mathcal{G}(n,d)$ is a graph uniformly chosen among all graphs on n vertices in which every vertex has degree d. One of the main tools for sampling from this distribution, as well as analysing the behaviour its typical elements, is the *configuration model*, introduced by Bollobás [5] (see [6, 13, 23]).

To construct a graph using this method, associate each of the n vertices with d distinct *half-edges*, and select a uniform perfect matching on these half-edges. The resulting (multi)graph is obtained by contracting each d half-edge into its corresponding vertex (possibly introducing multiple edges or self-loops). Crucially, given that the graph produced is simple, it is uniform over $\mathcal{G}(n,d)$, and for d fixed, the probability of this event is bounded away from 0. Hence, events that hold w.h.p. for the graph obtained via the configuration model also hold w.h.p. for $G \sim \mathcal{G}(n,d)$.

One particularly useful property of the configuration model is that it allows one to construct the graph gradually, exposing the edges of the perfect matching one at a time. This way, each additional edge is uniformly distributed among all possible edges on the remaining (unmatched) half-edges.

The distance between two vertices u,v in an unweighted (undirected) graph G, denoted by $\mathrm{dist}_G(u,v)$, is the number of edges in the shortest path connecting these two vertices. We will use the abbreviation $\mathrm{dist}(u,v)$ when there is no danger of confusion. If G has non-negative weights on its edges, the *length* of a path is replaced by its weight (the sum of weights along its edges), and $\mathrm{dist}(u,v)$ is analogously defined as the weight of the shortest (least heavy) path between u,v.

The diameter of a graph G, denoted by $\mathrm{diam}(G)$, is the maximum of $\mathrm{dist}(u,v)$ over all possible vertices u,v.

It is well known (and easy to show) that $\mathcal{G}(n,d)$ is locally tree-like around a typical vertex. To formalize such statements, we use the following notion. The *tree excess* of a connected set S, denoted by $\mathtt{tx}(S)$, is the maximum number of edges that can be deleted from the induced subgraph on S while still keeping it connected (*i.e.*, the number of extra edges in that induced subgraph beyond $|S|-1$).

3. Random regular graphs with exponential weights

3.1. Diameters of weighted graphs

In this section, we consider a random regular graph for $d \geqslant 3$ fixed, with i.i.d. rate 1 exponential variables on its edges. In fact, here and throughout the paper, we will consider a random d-regular *multigraph* generated via the configuration model (which will prove useful in capturing the geometry of the kernel of the largest component in $\mathcal{G}(n,p)$). Our goal is both to obtain the asymptotic diameter in this graph, and crucially, also establish its decay rate (see (3.3)).

Theorem 3.1. *Fix $d \geqslant 3$ and let $G \sim \mathcal{G}(n,d)$ be a random d-regular multigraph with n vertices and i.i.d. rate* 1 *exponential variables on its edges. Then w.h.p.*

$$\mathrm{diam}(G) = \left(\frac{1}{d-2} + \frac{2}{d}\right)\log n + O(\log\log n).$$

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

To prove the above result, we need to address the exponential decay of the distance between vertices, and introduce the following definition. Let $t > 0$. The *t-radius neighbourhood* of a vertex u, denoted by $B_u(t)$, is

$$B_t(u) \triangleq \{v : \mathrm{dist}(u, v) \leqslant t\}. \tag{3.1}$$

Further define the threshold of $B_t(u)$ reaching a certain size as

$$T_u(s) \triangleq \min\{t : |B_t(u)| \geqslant s\}. \tag{3.2}$$

Given these definitions, we can now formulate the exponential decay of $\mathrm{diam}(G)$ as well as T_u for all $u \in V$.

Theorem 3.2. *Fix $d \geqslant 3$ and let $G \sim \mathcal{G}(n, d)$ be a a random d-regular multigraph with n vertices and i.i.d. rate* 1 *exponential variables on its edges. Then there exists some $c > 0$ so that the following holds. For any $\ell > 0$,*

$$\mathbb{P}\left(\mathrm{diam}(G) \geqslant \left(\frac{1}{d-2} + \frac{2}{d}\right)\log n + 15\log\log n + \ell\right) \leqslant c\mathrm{e}^{-\ell/2} + cn^{-1/2}, \tag{3.3}$$

and for $q = 2\sqrt{dn\log n}$ and a uniformly chosen vertex u,

$$\mathbb{P}\left(T_u(q) \geqslant \frac{\log n}{2(d-2)} + 7\log\log n + \ell\right) \leqslant c\mathrm{e}^{-d\ell} + \frac{c}{n}\mathrm{e}^{-\ell} + 2n^{-3/2}. \tag{3.4}$$

Proof. We first wish to prove (3.4). Fix a vertex u, and consider the following continuous-time exploration process. At time $t = 0$, we have a ball that contains only u. For $t > 0$, our ball contains every v that has $\mathrm{dist}(u, v) \leqslant t$, *i.e.*, it is precisely $B_u(t)$.

In our setting, u is a fixed vertex and both the graph and its weights are random; hence we can expose the edges and their weights as we grow $B_t(u)$. This leads to an equivalent description of the process.

- Start with $B_0(u) = \{u\}$, where u has d (unmatched) half-edges.
- Reveal any matchings (and weights) of these d half-edges connecting them amongst themselves (self-loops at u).
- Repeat the following *exploration step*.
  - Given that there are m half-edges in the current set, denoted by $h_1, \ldots, h_m$, let $\Psi \sim \mathrm{Exp}(m)$ be a rate m exponential variable.
  - Select a uniform half-edge h_i and match it to a uniformly chosen half-edge outside B, thus introducing a new vertex with $d - 1$ new half-edges to the set B.
  - Reveal the matchings (and weights) of any of the half-edges of the newly added vertex whose match is also in B.

To verify the validity of the above process, consider the usual configuration model for generating the neighbourhood of u, a fixed vertex in a random d-regular graph. That process would repeatedly select a uniform unmatched half-edge in the current neighbourhood B and match it to a uniform half-edge in the entire graph. Equivalently, one can grow the neighbourhood by repeatedly revealing every edge within the induced subgraph on B before proceeding to expose

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

edges to new vertices. This method of generating the underlying random regular graph will simultaneously provide us with the weights along the edges. Indeed, we can continuously grow the weights of the half-edges $h_1, \dots, h_m$ in B until one of their rate 1 exponential clocks fires. Since the minimum of m exponentials is exponential with rate m, this is the same as choosing a uniform half-edge h_i after time Ψ (recall that by our conditioning, these m half-edges do not pair within themselves). Note that the final weight of an edge is accumulated between the time of arrival of its first half-edge and the time of its pairing (except edges going back into B whose weights are revealed immediately). Finally, the memoryless property of the exponential distribution guarantees that the process indeed generates a random d-regular graph with i.i.d. rate 1 exponentials.

Let τ_i denote the time of the ith exploration step ($i \geqslant 0$) in the above continuous-time process, and notice that for each i, at time τ_{i+1} we match a uniformly chosen half-edge from the set B_{τ_i} to a uniformly chosen half-edge among all other half-edges (excluding those in B_{τ_i}). Moreover, given $\mathcal{F}_{\tau_i}$, we have that $\tau_{i+1} - \tau_i$ is an exponential variable with rate k, where k is the number of half-edges in $B_{\tau_i}(u)$.

By the uniform choice of the matching, given $\mathcal{F}_{\tau_i}$, the number of half-edges introduced by the new vertex at time τ_{i+1} and connecting back to $B_{\tau_{i+1}}$ is stochastically dominated by a binomial variable

$$\mathrm{Bin}(d-1, \alpha), \quad \text{where } \alpha = \frac{d + (d-2)(i+1)}{dn - 2i} \leqslant \frac{(i+2)d - 2}{dn} \leqslant \frac{i+2}{n},$$

where the first inequality above is valid for, say, every $i \leqslant \frac{n}{2} - 5$. To justify this, observe that after i steps we have some $m \leqslant d + (d-2)i$ half-edges in B_{τ_i} out of some $M \geqslant dn - 2i$ half-edges in total. Moreover, if b half-edges connected back (did not produce a new vertex) then $m = d + (d-2)i - 2b$ whereas $M = dn - 2i - 2b$. Since the probability for each of the new $d-1$ half-edges to connect back into B_{τ_i} is at most

$$\frac{m + d - 2}{M} \leqslant \frac{d + (d-2)i}{dn - 2i},$$

the above statement holds.

Therefore, for every i, the tree-excess of B_{τ_i} is stochastically dominated by a binomial variable $\mathrm{Bin}(id, (i+2)/n)$ (with room to spare). Recalling that $q = 2\sqrt{dn \log n}$ and defining

$$r \triangleq \log^3 n,$$

we have the following for large n:

$$\begin{aligned} \mathbb{P}(\mathtt{tx}(B_{\tau_r}) \geqslant 1) &\leqslant \mathbb{P}\left(\mathrm{Bin}\left(dr, \frac{r+2}{n}\right) \geqslant 1\right) \leqslant O\left(\frac{r^2}{n}\right) = O\left(\frac{\log^6 n}{n}\right), \\ \mathbb{P}(\mathtt{tx}(B_{\tau_r}) \geqslant 2) &\leqslant \mathbb{P}\left(\mathrm{Bin}\left(dr, \frac{r+2}{n}\right) \geqslant 2\right) \leqslant O\left(\frac{r^4}{n^2}\right) = o\left(n^{-3/2}\right). \end{aligned} \tag{3.5}$$

Furthermore, for any k satisfying $r \leqslant k \leqslant 10q$,

$$\mathbb{P}\left(\mathtt{tx}(B_{\tau_k}) \geqslant \frac{k}{\sqrt{r}}\right) \leqslant \mathbb{P}\left(\mathrm{Bin}\left(dk, \frac{k+2}{n}\right) \geqslant \frac{k}{\sqrt{r}}\right) \leqslant \exp\left(-\frac{1}{3}k/\sqrt{r}\right) < n^{-6},$$

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

where the last two inequalities hold for any sufficiently large n by Chernoff's inequality (see, *e.g.*, [2]), noting that $k^2/n = o(k/\sqrt{r})$. Define the event

$$R \triangleq \{\mathtt{tx}(B_{\tau_k}) < k/\sqrt{r} \text{ for all } r \leqslant k \leqslant 10q\}, \tag{3.6}$$

and note that a union bound over all $r \leqslant k \leqslant 10q$ gives that $\mathbb{P}(R) \geqslant 1 - n^{-5}$. (In the rare event that the exploration from u exhausts itself prior to $k = 10q$, the event R applies only to values of k for which τ_k is defined. This occurs with probability $O(1/n^2)$, and hence may be neglected.) At this point, we have two cases concerning $\mathtt{tx}(B_{\tau_r})$.

Case 1: The tree excess of B_{τ_r} is 0 and the event R holds so far. Denote this event by Q_1. In this case, for any $i < r$, conditioning on $\mathtt{tx}(B_{\tau_i}) = 0$ we have that the number of half-edges of B_{τ_i} is $d + di - 2i = (d-2)i + d$, and $\tau_{i+1} - \tau_i \preceq Y_i$, where

$$Y_i \sim \mathrm{Exp}((d-2)i + d) \qquad (i = 0, 1, \ldots, r-1).$$

and all Y_is are independent. (Here and in what follows, $\mu \preceq \nu$ denotes stochastic domination, *i.e.*, $\int f d\mu \leqslant \int f d\nu$ for any increasing function f.) Now, each matching that contributes to the tree excess eliminates two half-edges instead of introducing $d-1$ new ones. Thus, for each $i \geqslant r$, conditioning on $\mathtt{tx}(B_{\tau_i}) < i/\sqrt{r}$ we get that $\tau_{i+1} - \tau_i \preceq Y_i$, where

$$Y_i \sim \mathrm{Exp}((d-2)i + d - 2i/\sqrt{r}) \qquad (i = r, r+1, \ldots) \tag{3.7}$$

Applying the Laplace transform, and noting that for $Y \sim \mathrm{Exp}(\rho)$

$$\mathbb{E}\mathrm{e}^{\lambda Y} = \int_0^\infty \mathrm{e}^{\lambda y} \rho \mathrm{e}^{-\rho y} dy = \frac{\rho}{\rho - \lambda},$$

we obtain that

$$\mathbb{E}\left[\mathrm{e}^{\lambda(\tau_q - \tau_1)} \mid Q_1\right] \leqslant \prod_{i=1}^{r-1} \frac{(d-2)i + d}{(d-2)i + d - \lambda} \prod_{i=r}^{q-1} \frac{(d-2)i + d - 2i/\sqrt{r}}{(d-2)i + d - 2i/\sqrt{r} - \lambda}.$$

Taking $\lambda = d$, we get

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{d(\tau_q - \tau_1)} \mid Q_1\right] &\leqslant \prod_{i=1}^{r-1} \left(1 + \frac{d}{(d-2)i}\right) \prod_{i=r}^{q-1} \left(1 + \frac{d}{(d-2)i - 2i/\sqrt{r}}\right) \\
&\leqslant \exp\left[\frac{d}{d-2} \sum_{i=1}^{r-1} \frac{1}{i} + \frac{d}{d-2} \cdot \frac{1}{1 - \frac{2}{d-2} r^{-1/2}} \sum_{i=r}^{q-1} \frac{1}{i}\right] \\
&\leqslant \exp\left[\frac{d}{d-2}\left(1 + O(r^{-1/2})\right) \log q + 2\right] \leqslant 10 q^{d/(d-2)},
\end{aligned}$$

where the last inequality holds for large n. Hence, in Case 1 we have

$$\mathbb{P}\left(\tau_q - \tau_1 \geqslant \frac{\log n}{2(d-2)} + \frac{\log\log n}{2(d-2)} + \ell \;\middle|\; Q_1\right) \leqslant O(\exp(-d\ell)).$$

As $\tau_1 \sim \mathrm{Exp}(d)$, altogether in this case

$$\mathbb{P}\left(\tau_q \geqslant \frac{\log n}{2(d-2)} + \frac{\log\log n}{2(d-2)} + \ell \;\middle|\; Q_1\right) \leqslant O(\exp(-d\ell)).$$

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

Case 2: The tree excess of B_{τ_r} is 1 and the event R holds so far. Denote this event by Q_2. Here, for any $i < r$, the following holds. Conditioned on $\mathtt{tx}(B_{\tau_i}) \leqslant 1$, the number of half-edges in B_{τ_i} is d for $i = 0$ and at least $(d-2)i$ for $i \geqslant 1$. Hence, $\tau_{i+1} - \tau_i \preceq Y_i$, where the Y_is are independent variables given by

$$Y_i \sim \mathrm{Exp}\big((d-2)(i+1)\big) \quad (i = 0, 1, \ldots, r-1).$$

For $i \geqslant r$, conditioned on $\mathtt{tx}(B_{\tau_i}) < i/\sqrt{r}$ the bounds of Case 1 hold, that is, $\tau_{i+1} - \tau_i \preceq Y_i$ for the variables Y_i as given in (3.7). This yields that

$$\mathbb{E}\big[\mathrm{e}^{\lambda(\tau_q - \tau_1)} \mid Q_2\big] \leqslant \prod_{i=1}^{r-1} \frac{(d-2)(i+1)}{(d-2)(i+1) - \lambda} \prod_{i=r}^{q-1} \frac{(d-2)(i+1) - 2i/\sqrt{r}}{(d-2)(i+1) - 2i/\sqrt{r} - \lambda}.$$

Taking $\lambda = 1$, we get

$$\begin{aligned}
\mathbb{E}\big[\mathrm{e}^{\tau_q - \tau_1} \mid Q_2\big] &\leqslant \prod_{i=1}^{r-1} \left(1 + \frac{1}{(d-2)(i+1)}\right) \prod_{i=r}^{q-1} \left(1 + \frac{1}{(d-2)(i+1) - 2i/\sqrt{r} - 1}\right) \\
&\leqslant \exp\left[\frac{1}{d-2} \sum_{i=2}^{r} \frac{1}{i} + \frac{1}{d-2} \cdot \frac{1}{1 - 2[(d-2)\sqrt{r}]^{-1}} \sum_{i=r}^{q-1} \frac{1}{i}\right] \\
&\leqslant \exp\left[\frac{1}{d-2}\big(1 + O(r^{-1/2})\big) \log q + 2\right] \leqslant 10 q^{1/(d-2)},
\end{aligned}$$

where the last inequality holds for large n. Hence, in Case 2,

$$\mathbb{P}\left(\tau_q - \tau_1 \geqslant \frac{\log n}{2(d-2)} + \frac{\log\log n}{2(d-2)} + \ell \;\middle|\; Q_2\right) \leqslant O(\exp(-\ell)),$$

and again, as $\tau_1 \preceq \mathrm{Exp}(d-2)$, in this case

$$\mathbb{P}\left(\tau_q \geqslant \frac{\log n}{2(d-2)} + \frac{\log\log n}{2(d-2)} + \ell \;\middle|\; Q_2\right) \leqslant O(\exp(-\ell)).$$

Combining the above two cases using (3.5) and that $\mathbb{P}(R) \geqslant 1 - n^{-5}$, we conclude that for some $c > 0$ fixed,

$$\mathbb{P}\left(\tau_q \geqslant \frac{\log n}{2(d-2)} + 7\log\log n + \ell\right) \leqslant c\mathrm{e}^{-d\ell} + \frac{c}{n}\mathrm{e}^{-\ell} + 2n^{-3/2}$$

(here the $7\log\log n$ term eliminated the $\log^6 n$ factor from (3.5)), readily implying (3.4).

To obtain (3.3) we need the next simple lemma.

Lemma 3.3. *With high probability,* $\mathrm{dist}(u, v) \leqslant T_u(q) + T_v(q)$ *for all* u, v.

Proof. Assume that $T_u(q), T_v(q) < \infty$ (*i.e.*, each of the connected components of u and v consists of at least q vertices), otherwise the statement of the lemma holds trivially.

Fix two vertices u, v and consider the aforementioned exploration process. Explore $B_t(u)$ until reaching $t = T_u(q)$, and condition on the event that its tree-excess is $o(q)$, as ensured by the event R defined in (3.6) (this event holds with probability at least $1 - n^{-5}$). Thus, there are $(d - 1 - o(1))q$ half-edges in $B_{T_u(q)}$ except with probability n^{-5}.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

Next, begin exposing $B_t(v)$; each matching adds a uniform half-edge to the neighbourhood of v, and so the probability that $B_{T_v(q)}$ does not intersect $B_{T_u(q)}$ is at most

$$\left(1 - \frac{(d-1-o(1))q}{dn}\right)^q \leqslant \exp(-4(d-1-o(1))\log n) < n^{-7}$$

for any large n. A union bound over u, v now completes the proof. □

To infer (3.3) from (3.4) and the above lemma, argue as follows. Choose $\ell = \frac{1}{d}\log n + k$ in (3.4) to obtain that for some constant $c > 0$, a uniformly chosen vertex u has

$$\mathbb{P}\left(T_u(q) \geqslant \left(\frac{1}{2(d-2)} + \frac{1}{d}\right)\log n + 7\log\log n + k\right) < \frac{c}{n}\mathrm{e}^{-dk} + \frac{c}{n}\mathrm{e}^{-k} + 2n^{-3/2}$$

(with room to spare), and by taking a union bound over u it follows that, for some other $c > 0$, the probability of

$$\left\{T_u(q) \leqslant \left(\frac{1}{2(d-2)} + \frac{1}{d}\right)\log n + 7\log\log n + k\right\} \quad \text{for every } u$$

is at least $1 - c\mathrm{e}^{-k} - 2/\sqrt{n}$. Combining Lemma 3.3 with a choice of $k = \ell/2$ now gives the estimate (3.3) for the decay of the diameter, as required. □

We also need the next simple lemma, which bounds the number of edges in a path achieving $\mathrm{dist}(u, v)$ for any u, v.

Lemma 3.4. *With high probability, for any two vertices u, v, the number of edges in the path achieving* $\mathrm{dist}(u, v)$ *is at most* $4d\mathrm{e}\log n$.

Proof. First consider an arbitrary given d-regular graph H on n vertices with i.i.d. rate 1 exponentials on its edges. Consider a simple path consisting of ℓ edges, and let X_i denote the Exp(1) variable that corresponds to the ith edge. Setting $S_\ell = \sum_{i \leqslant \ell} X_i$,

$$\mathbb{P}(S_\ell \leqslant a\ell) = \mathbb{P}(\mathrm{e}^{-\lambda S_\ell} \geqslant \mathrm{e}^{-\lambda a\ell}) \leqslant \mathrm{e}^{\lambda a\ell}\mathbb{E}\mathrm{e}^{-\lambda S_\ell} = \frac{\mathrm{e}^{\lambda a\ell}}{(1+\lambda)^\ell},$$

where the last equality is due to the fact that $\int_0^\infty \mathrm{e}^{-\lambda x}\mathrm{e}^{-x}dx = 1/(1+\lambda)$. Choosing $\lambda = (1-a)/a$ and subsequently $a = 1/(2d\mathrm{e})$, we get that

$$\mathbb{P}(S_\ell \leqslant a\ell) \leqslant (a\mathrm{e}^{1-a})^\ell \leqslant (2d)^{-\ell}.$$

Fix a starting position u for the path. Summing over all $d(d-1)^{\ell-1}$ possible simple paths of length ℓ originating from u, we have that the probability that one of these paths would have weight smaller than $a\ell$ is at most $2^{-\ell}$. In particular, if A_u denotes the event that, for some integer $\ell \geqslant 2\log_2 n$, there is a simple path originating from u with length ℓ and weight at most $a\ell$, then $\mathbb{P}(A_u) \leqslant 2n^{-2}$. Taking a union bound over the vertices u gives this for any vertex of the graph H w.h.p.

Now take H to be our random d-regular graph. By the estimate of Theorem 3.2 on the diameter, we deduce that, w.h.p., for any u, v either the path achieving $\mathrm{dist}(u, v)$ has at most $2\log_2 n$ edges,

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

or the number of edges it contains is at most

$$\operatorname{diam}(G)/a \leqslant 2de\left(\frac{1}{d-2}+\frac{2}{d}+o(1)\right)\log n < 4de\log n$$

for large n, thus concluding the proof. □

The upper bound on the diameter of the weighted graph G will follow immediately from the results we have established so far. For the lower bound, we need the following lemma.

Lemma 3.5. *Let u, v be two uniformly chosen vertices of the graph G defined in Theorem 3.2. Let $B'_t(x) = \{y : \operatorname{dist}(y, N_G(x)) \leqslant t\}$, where $N_G(x)$ denotes the neighbours of x in G. Then w.h.p. $B'_{t_0}(u) \cap B'_{t_0}(v) = \emptyset$ for $t_0 = \frac{1}{2(d-2)}\log n - 3\log\log n$.*

Proof. Fix a vertex u, consider the exploration process defined in the proof of Theorem 3.2 from the set $N_G(u)$ (w.h.p. comprising $d(d-1)$ half-edges), and again let τ_i be the time of the ith step. As argued before, $\tau_{i+1} - \tau_i \succeq Y_i$, where the Y_is are independent exponential variables given by

$$Y_i \sim \operatorname{Exp}\big((d-2)i + d(d-1)\big)$$

(this follows from the fact that the worst case is when the explored set forms a tree). At this point, setting $b \triangleq d(d-1)-(d-2)$, we get

$$\begin{aligned}\mathbb{P}(\tau_z \leqslant t) &\leqslant \int_{\sum_{i=1}^z x_i \leqslant t} \prod_{i=1}^z \big[(d-2)i+b\big]\mathrm{e}^{-\sum_{i=1}^z((d-2)i+b)x_i}dx_1\dots dx_z \\ &= \int_{0\leqslant y_1\leqslant\dots\leqslant y_z\leqslant t} \prod_{i=1}^z \big[(d-2)i+b\big]\mathrm{e}^{-by_z}\mathrm{e}^{-(d-2)\sum_{i=1}^z y_i}dy_1\dots dy_z,\end{aligned}$$

where $y_k = \sum_{i=0}^{k-1} x_{z-i}$. Letting y play the role of y_z and accounting for all permutations over $y_1,\dots,y_{z-1}$ (giving each such variable the range $[0,y]$),

$$\begin{aligned}\mathbb{P}(\tau_z \leqslant t) &\leqslant \int_0^t \mathrm{e}^{-(d-2+b)y}\frac{\prod_{i=1}^z(i+\frac{b}{d-2})}{(z-1)!} \\ &\quad\times\left(\int_{[0,y]^{z-1}}(d-2)^z\mathrm{e}^{-(d-2)\sum_{i=1}^{z-1}y_i}dy_1\dots dy_{z-1}\right)dy \\ &\leqslant \int_0^t \mathrm{e}^{-(d-2+b)y}\frac{\prod_{i=1}^z(i+\frac{b}{d-2})}{(z-1)!}\cdot\left(\prod_{i=1}^{z-1}\int_0^y(d-2)\mathrm{e}^{-(d-2)y_i}dy_i\right) \\ &\leqslant c(d-2)z^{\frac{b}{d-2}+1}\int_0^t \mathrm{e}^{-d(d-1)y}(1-\mathrm{e}^{-(d-2)y})^{z-1}dy,\end{aligned}$$

where $c>0$ is an absolute constant. Setting $t_0 = \frac{1}{d-2}(\log z - 2\log\log n)$ and $z = \sqrt{n/\log n}$, we obtain that

$$\mathbb{P}(\tau_z \leqslant t_0) \leqslant c(d-2)z^{\frac{b}{d-2}+1}\int_0^{t_0}\mathrm{e}^{-\log^2 n}dy = o(n^{-5}),$$

where we used the fact that $(1-\mathrm{e}^{-(d-2)y})^{z-1} \leqslant \mathrm{e}^{-\log^2 n}$ for all $0 \leqslant y \leqslant t_0$.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

To conclude the proof, observe that the above argument showed that w.h.p. $|B'_{t_0}(u)| \leqslant z$. Choosing another uniform vertex v (which misses this mentioned set with probability $1 - o(1)$) and exposing $B'_{t_0}(v)$, again w.h.p. we obtain a set of size at most z. Crucially, each matching is uniform among the remaining half-edges, and so its probability of hitting the boundary of $|B'_{t_0}(u)|$ is at most z/n. Altogether,

$$\mathbb{P}(B'_{t_0}(u) \cap B'_{t_0}(v) \neq \emptyset) \leqslant \frac{z^2}{n} + o(1) = o(1),$$

as required. □

We are now ready to prove the asymptotic behaviour of the diameter.

Proof of Theorem 3.1. For the upper bound, apply Theorem 3.2 with, say, $\ell = \log\log n$, to obtain that w.h.p.

$$\operatorname{diam}(G) \leqslant \left(\frac{1}{d-2} + \frac{2}{d}\right)\log n + 16\log\log n = \left(\frac{1}{d-2} + \frac{2}{d} + o(1)\right)\log n.$$

It remains to provide a matching asymptotic lower bound.

Set

$$D \triangleq \frac{1}{d}\log n - \frac{1}{d}\log\log n,$$

and call a vertex 'good' if the weight on all the d edges connected to it is larger than D. For a vertex $u \in V$, let A_u denote the event that u is good. Clearly, $\mathbb{P}(A_u) = \frac{\log n}{n}$, and so if $Y = \sum_u \mathbf{1}_{A_u}$ counts the number of good vertices, we have that $\mathbb{E}Y = \log n$. Furthermore,

$$\begin{aligned}\operatorname{Var}(Y) &= \sum_{u,v} \operatorname{Cov}(\mathbf{1}_{A_u}, \mathbf{1}_{A_v}) = \sum_u \operatorname{Var}(\mathbf{1}_{A_u}) + \sum_u \sum_{v:uv\in E(G)} \operatorname{Cov}(\mathbf{1}_{A_u}, \mathbf{1}_{A_v}) \\ &\leqslant \mathbb{E}Y + \sum_u d\mathbb{P}(A_u) = (d+1)\mathbb{E}Y,\end{aligned}$$

and by Chebyshev's inequality we deduce that, for instance, $Y \geqslant \frac{2}{3}\log n$ w.h.p. In particular, the number of pairs of distinct good vertices, denoted by R, satisfies $R \geqslant \frac{1}{4}\log^2 n$ w.h.p.

On the other hand, recalling Lemma 3.5 and taking

$$t_0 = \frac{1}{2(d-2)}\log n - 3\log\log n,$$

we have that two uniform vertices u, v satisfy $B'_{t_0}(u) \cap B'_{t_0}(v) = \emptyset$ w.h.p. where the ball $B'_t(x)$ includes all vertices of distance t from the neighbours of x (excluding x itself). In particular, condition on the events A_u and A_v the probability that $B'_{t_0}(u)$ does not intersect $B'_{t_0}(v)$ remains the same (since the weights on the immediate edges incident to u, v do not play a part). Therefore, for two uniformly chosen vertices u, v we have

$$\mathbb{P}\big(A_u,\, A_v,\, B'_{t_0}(u) \cap B'_{t_0}(v) \neq \emptyset\big) = o(\mathbb{P}(A_u,\, A_v)).$$

Denoting by R' the the number of pairs of good vertices that are of distance at most $2D + \frac{1}{d-2}\log n - 6\log\log n$, we deduce that

$$\mathbb{E}R' = o(\mathbb{E}Y^2) = o((\mathbb{E}Y)^2) = o(\log^2 n).$$

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

By Markov's inequality, $R' \leqslant \frac{1}{8}\log^2 n$ w.h.p., and hence $R - R'$ is w.h.p. non-empty. This implies the existence of two vertices whose distance is at least

$$\left(\frac{1}{d-2} + \frac{2}{d}\right)\log n - 7\log\log n,$$

completing the proof. □

3.2. Diameters of metric graphs

We consider the following continuous analogue of the diameter of a weighted graph (see, *e.g.*, [3] for related information).

Definition. Let $G = (V, E)$ be a graph with non-negative weights on its edges $\{w(e) : e \in E\}$. The corresponding metric graph $\mathcal{X} = \mathcal{X}(G)$ is the graph obtained by replacing every $e \in E$ by a line segment L_e of length $w(e)$, with the (uncountable) vertex set $\cup_e L_e$ and the obvious shortest path metric.

That is, the distance between any $x, y \in \mathcal{X}$, lying on two distinct L_e and $L_{e'}$, respectively, is the minimum of $|x' - x|_{L_e} + |y' - y|_{L_{e'}} + \mathrm{dist}_G(x', y')$ over all 4 possible choices of endpoints x', y' of $L_e, L_{e'}$, respectively, where $|\cdot|_{L_e}$ is the Euclidean distance in the interval $[0, w(e)]$ and we identified x', y' with vertices of G (when $e = e'$ the distance is the minimum of the above and $|x - y|_{L_e}$).

For a fixed integer $d \geqslant 3$, Let $G \sim \mathcal{G}(n, d)$ be a random d-regular graph with i.i.d. rate 1 exponential weights on its edges. At times we will identify the metric graph $\mathcal{X}$ with its points (the union of its line segments). As before, we let $\mathrm{diam}(\mathcal{X}) = \max_{x,y\in\mathcal{X}} \mathrm{dist}(x, y)$.

The next theorem establishes the typical diameter of $\mathcal{X}$, which differs from that of G by a term of $\left(\frac{d-2}{d} + o(1)\right)\log n$.

Theorem 3.6. *Fix $d \geqslant 3$, let $G \sim \mathcal{G}(n, d)$ be a random d-regular multigraph on n vertices with i.i.d.* Exp(1) *variables on its edges, and let $\mathcal{X}(G)$ be its metric graph. Then w.h.p.*

$$\mathrm{diam}(\mathcal{X}) = \left(1 + \frac{1}{d-2}\right)\log n + O(\log\log n).$$

The proof follows the same arguments used to prove Theorem 3.1, but instead of exposing a neighbourhood of an initial vertex, it does so while excluding one of the edges incident to this vertex. We will therefore focus on the modifications required to adapt the original argument to our new setting.

Recall that we defined $T_u(q)$ in the discrete setting as the minimal t such that $|B_t(u)| \geqslant q$. In the continuous case, we will still have $B_t(u)$ count the number of vertices of G, that is, endpoints of line segments in $\mathcal{X}$. Further define $B_t(\bar{e})$ for a directed edge $\bar{e} = (x, y)$ as the set of all vertices of G whose distance from y is at most t, without using the edge $\bar{e}$ (in any direction). That is, $B_t(\bar{e})$ is the set $B_t(y)$ minus vertices that depend on $\bar{e}$ for being included in this set.

A small modification is required in the rare case of a self-loop or multiple-edge incident either to y or to one of its neighbours excluding x. In this case we call the edge $\bar{e} = (x, y)$ 'rare' and

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

take $B_t(\bar{e})$ to be $B_t(\bar{e}')$, where $\bar{e}' = (y, x)$ is the reversed edge. Observe that w.h.p. at most one of these edges is rare, as the probability for each of these independent events is $O(1/n)$.

As before (yet slightly modified in the case of a rare edge), let

$$T_{\bar{e}}(q) = \min\{t : |B_t(\bar{e})| \geqslant q\} + \mathbf{1}_{\{\bar{e} \text{ is rare}\}} \frac{1}{2} w(\bar{e}). \tag{3.8}$$

We can now formulate an exponential decay statement for the $\mathrm{diam}(\mathcal{X})$, analogous to Theorem 3.2.

Theorem 3.7. *Let G and $\mathcal{X}$ be as in Theorem 3.6. Then there exists some $c > 0$ so that the following holds. For any $\ell > 0$,*

$$\mathbb{P}\left(\mathrm{diam}(\mathcal{X}) \geqslant \left(1 + \frac{1}{d-2}\right) \log n + 15 \log\log n + \ell\right) \leqslant c\mathrm{e}^{-\ell} + (2 + c\mathrm{e}^{-\ell/2}) n^{-1/2}, \tag{3.9}$$

and for $q = 2\sqrt{dn \log n}$ and a uniformly chosen directed edge $\bar{e}$,

$$\mathbb{P}\left(T_{\bar{e}}(q) + \frac{1}{2} w(\bar{e}) \geqslant \frac{\log n}{2(d-2)} + 7 \log\log n + \ell\right) \leqslant c\mathrm{e}^{-2\ell} + \frac{c}{n}\mathrm{e}^{-\ell} + 2n^{-3/2}. \tag{3.10}$$

Proof. Consider the exploration process from the endpoint y of the uniformly chosen edge $\bar{e} = (x, y)$, while disregarding $\bar{e}$. This process is identical to the process defined in the previous subsection except that it starts with $d - 1$ half-edges rather than d half-edges.

Letting τ_i be the time of the ith exploration step, we have that there are at most $d - 1 + (d - 2)i$ half-edges in B_{τ_i} out of $dn - 2i$ half-edges in total. As before, the number of half-edges, introduced at time τ_{i+1} and connecting back to $B_{\tau_{i+1}}$, is stochastically dominated by a binomial variable

$$\mathrm{Bin}(d-1, \alpha) \quad \text{where } \alpha = \frac{d - 1 + (d-2)(i+1)}{dn - 2i} \leqslant \frac{i+2}{n}.$$

Thus, we apply the same analysis of the tree excess of B_{τ_i} as in the proof of Theorem 3.2 (stochastically dominating it via a binomial $\mathrm{Bin}(d, (i+2)/n)$) to obtain the exact bounds of (3.5), as well as that $\mathbb{P}(R) \geqslant 1 - n^{-5}$ for R defined in (3.6).

Examining the two cases for $\mathtt{tx}(B_{\tau_r})$, we now have that $\tau_{i+1} - \tau_i \preceq Y_i$ for i.i.d. exponential variables Y_i defined as follows.

Case 1: The tree excess of B_{τ_r} is 0 and the event R holds. We have

$$\begin{aligned} Y_0 &\sim \mathrm{Exp}(d-1), \quad Y_i \sim \mathrm{Exp}(d - 1 + (d-2)i) && (i = 1, \ldots, r-1), \\ Y_i &\sim \mathrm{Exp}(d - 1 + (d-2)i - 2i/\sqrt{r}) && (i = r, r+1, \ldots). \end{aligned}$$

For $d \geqslant 3$ the rates of the above Y_is are all at least 2.

Case 2: The tree excess of B_{τ_r} is 1 and the event R holds. First assume $\bar{e}$ is not rare. In this case there are $d - 1$ half-edges initially: the first step adds $d - 2$ half-edges to the boundary ($d - 1$ new ones at the cost of 1 half-edge lost, as there are no loops or multiple edges), and similarly

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

the second step adds $d-2$ half-edges. Thus,

$$\begin{aligned} &Y_0 \sim \mathrm{Exp}(d-1), \quad Y_1 \sim \mathrm{Exp}(2d-3), \\ &Y_i \sim \mathrm{Exp}(3d-7+(d-2)(i-2)) && (i=2,\dots,r-1), \\ &Y_i \sim \mathrm{Exp}(3d-7+(d-2)(i-2)-2i/\sqrt{r}) && (i=r,r+1,\dots). \end{aligned}$$

Note that for $d \geqslant 3$ the rates of the above Y_is are again all at least 2. In the special case where $\bar{e}$ is rare (occurring with probability $O(1/n)$), the analogous sequence of Y_is will feature a reduced rate of at least 1.

It is then easy to verify that a choice of $\lambda = 2$ for the Laplace transform for a typical edge $\bar{e}$ and $\lambda = 1$ when $\bar{e}$ is rare, plugged into the same calculations as in the proof of Theorem 3.2, gives that

$$\mathbb{P}\left(T_{\bar{e}}(q) \geqslant \frac{\log n}{2(d-2)} + 7\log\log n + \ell \right) \leqslant c\mathrm{e}^{-2\ell} + \frac{c}{n}\mathrm{e}^{-\ell} + 2n^{-3/2}$$

(the middle term corresponds to a rare edge $\bar{e}$; otherwise, Cases 1 and 2 both give a decay rate of 2).

Consider an edge $\bar{e}$ that is not rare. Crucially, $w(\bar{e})$ is independent of $T_{\bar{e}}(q)$ by definition, and as $w(\bar{e}) \sim \mathrm{Exp}(1)$ (and so $\frac{1}{2}w(\bar{e}) \sim \mathrm{Exp}(2)$ has decay rate 2) we immediately obtain (3.10). Alternatively, when $\bar{e}$ is rare, our modified definition of $T_{\bar{e}}(q)$ again verifies (3.10) (this time both $w(\bar{e})$ and the growth of B have rate 1).

Choosing $\ell = \frac{1}{2}\log n + k$ and taking a union bound over $\bar{e}$, we have that

$$\left\{ T_{\bar{e}}(q) + \frac{1}{2}w(\bar{e}) \leqslant \left(\frac{1}{2(d-2)} + \frac{1}{2} \right) \log n + 7\log\log n + k \right\} \quad \text{for every } \bar{e}, \tag{3.11}$$

with probability at least $1 - c\mathrm{e}^{-2k} - (2 + c\mathrm{e}^{-k})n^{-1/2}$.

To conclude (3.9) from (3.11), argue as follows. For two points $x, y \in \mathcal{X}$, write u_x, u_y for their closest points in G, respectively, and let $\bar{e}_x, \bar{e}_y$ denote the directed edges containing x, y and ending at u_x, u_y, respectively. For simplicity, assume first that neither of these edges is rare. Clearly,

$$\mathrm{dist}(x,y) \leqslant \mathrm{dist}(u_x,u_y) + \frac{1}{2}w(\bar{e}_x) + \frac{1}{2}w(\bar{e}_y).$$

Now, by Lemma 3.3, w.h.p. we have $\mathrm{dist}(u,v) \leqslant T_u(q) + T_v(q)$ for all u, v. Furthermore, for any directed edge $\bar{e} = (u,v)$, as $B_t(\bar{e}) \subset B_t(v)$ we get that $T_{\bar{e}}(q) \geqslant T_v(q)$. Altogether, w.h.p.

$$\mathrm{dist}(x,y) \leqslant T_{\bar{e}_x}(q) + \frac{1}{2}w(\bar{e}_x) + T_{\bar{e}_y}(q) + \frac{1}{2}w(\bar{e}_y). \tag{3.12}$$

We claim that the above inequality remains valid if one or more of the edges $\bar{e}_x, \bar{e}_y$ is rare. For instance, if $\bar{e}_x$ is rare we could replace u_x by the other endpoint of $\bar{e}_x$ at a cost of at most $\frac{1}{2}w(\bar{e}_x)$, which is precisely accounted for in the modified definition of $T_{\bar{e}_x}(q)$. The exponential decay of the diameter, as stated in (3.9), now immediately follows from (3.11). □

Proof of Theorem 3.6. The upper bound on $\mathrm{diam}(\mathcal{X})$ follows directly from the exponential decay estimate in (3.9), *e.g.*, by taking $\ell = \log\log n$.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

The proof of the lower bound will follow from essentially the same arguments used to prove Theorem 3.1. Here, instead of defining a 'good' vertex, we call an edge 'good' if its weight is larger than $D = \log n - \log\log n$, and obtain that the probability that a given edge e is good is $\frac{\log n}{n}$. The same second moment argument that showed that R, the number of pairs of distinct good vertices, is at least $\frac{1}{4}\log^2 n$ w.h.p., now yields the same estimate on the number of pairs of distinct good edges.

The argument in Theorem 3.1 then proceeded with an application of Lemma 3.5, which bounds the probability that two uniformly chosen vertices u, v have intersecting neighbourhoods $B'_{t_0}(u)$ and $B'_{t_0}(v)$ for a prescribed t_0 (where $B'_t(x)$ consists of vertices that are of distance at most t from $N_G(x)$, the neighbours of x in G). We now note that the statement of this lemma also holds for any uniformly chosen points $x, y \in \mathcal{X}$, when $N_G(x)$ still stands for the neighbours of x in the graph G (*i.e.*, the endpoints of the incident edges). To see this, simply replace the b by $2(d-1)-(d-2)$ in the proof of that lemma, accounting for the $2(d-1)$ initial half-edges in the exploration process. Since the value of b does not play a role in the proof (as long as it is a constant), the proof holds without any further changes (for the same value of t_0).

Therefore, as before we may deduce that the expected number of pairs of good edges whose distance is at most $2t_0$ is $o(\log^2 n)$, and so w.h.p. (using Markov's inequality) there exists a pair of good edges, e, f, whose distance in $\mathcal{X}$ is at least $2t_0$. The proof is now concluded by choosing the middle points in e, f to obtain a distance of

$$2 \cdot \frac{D}{2} + 2t_0 = \left(1 + \frac{1}{d-2}\right)\log n - 7\log\log n,$$

as required. □

4. The young giant component

In this section, we focus on the giant component in the regime $\varepsilon^3 n \to \infty$ and $\varepsilon = o(n^{-1/4})$. In order to relate the results of the previous section (concerning diameters of weighted random regular graphs) to this setting, we apply the main result of the companion paper [10], which provides a complete and tractable description of $\mathcal{C}_1$.

Theorem 4.1 (Theorem 1 of [10]). *Let $\mathcal{C}_1$ be the largest component of the random graph $\mathcal{G}(n,p)$ for $p = \frac{1+\varepsilon}{n}$, where $\varepsilon^3 n \to \infty$ and $\varepsilon = o(n^{-1/4})$. Then $\mathcal{C}_1$ is one-sided contiguous to the model $\tilde{\mathcal{C}}_1$, constructed in three steps as follows.*

(1) *Let $Z \sim \mathcal{N}\left(\frac{2}{3}\varepsilon^3 n, \varepsilon^3 n\right)$, and select a random 3-regular multigraph $\mathcal{K}$ on $N = 2\lfloor Z \rfloor$ vertices.*
(2) *Replace each edge of $\mathcal{K}$ by a path, where the path lengths are i.i.d.* Geom(ε).
(3) *Attach an independent* Poisson$(1-\varepsilon)$*–Galton–Watson tree to each vertex.*

That is, $\mathbb{P}(\tilde{\mathcal{C}}_1 \in \mathcal{A}) \to 0$ implies $\mathbb{P}(\mathcal{C}_1 \in \mathcal{A}) \to 0$ for any set of graphs $\mathcal{A}$.

In the above, a Poisson(μ)–Galton–Watson tree (or a PGW-tree for short) is the family tree of a Galton–Watson branching process with offspring distribution Poisson(μ), and $\mathcal{N}(\mu, \sigma^2)$ stands for the normal distribution with mean μ and variance σ^2.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

4.1. The 2-core and its kernel

In light of the above theorem, the kernel $\mathcal{K}$ of the giant component in the regime of $p = \frac{1+\varepsilon}{n}$ for $\varepsilon^3 n \to \infty$ and $\varepsilon = o(n^{-1/4})$ can be regarded a random 3-regular graph $G \sim \mathcal{G}(N, 3)$. The 2-core $\mathcal{C}_1^{(2)}$ is then obtained by replacing the edges of this graph by 2-paths (*i.e.*, paths whose interior vertices all have degree 2) of lengths i.i.d. geometric with mean $1/\varepsilon$.

Proof of Theorem 1.2 for the regime $\varepsilon = o(n^{-1/4})$. Let $G \sim \mathcal{G}(N, 3)$ be a random 3-regular graph with i.i.d. Exp(λ) edge weights $\{w(e) : w \in E(G)\}$, where $e^{-\lambda} = 1 - \varepsilon$. Further, let H be the unweighted graph obtained by taking the underlying graph of G, and replacing each of its edges by paths of length i.i.d. geometric variables with mean $1/\varepsilon$. Let K denote the subset of the vertices of H of degree at least 3.

By Theorem 4.1, for N as defined as in step (1) (which in particular satisfies $N = (\frac{4}{3} + o(1))\varepsilon^3 n$ w.h.p.), H corresponds to $\mathcal{C}_1^{(2)}$, the 2-core of the giant component of $\mathcal{G}(n, p)$, and the graph G corresponds to its kernel $\mathcal{K}$.

We can clearly couple G and H such that each 2-path P in H, corresponding to some edge $e \in E(G)$, would satisfy $||P| - w(e)| \leqslant 1$.

By Theorem 3.2, the diameter of the weighted graph G is w.h.p.

$$\operatorname{diam}(G) = \left(\frac{5}{3} + o(1)\right)(1/\lambda)\log N,$$

and by Lemma 3.4 the path achieving it consists of $O(\log N)$ edges. Hence, recalling that $1/\lambda = (1 + o(1))(1/\varepsilon) \to \infty$, the distance between any two vertices $u, v \in K$ in the graph H differs from their distance in G by at most $O(\log N) = o(\operatorname{diam}(G))$, and so w.h.p.

$$\max_{u,v \in K} \operatorname{dist}_H(u, v) = \left(\frac{5}{3} + o(1)\right)(1/\varepsilon)\log(\varepsilon^3 n).$$

Furthermore, by the coupling of G and H, given the metric graph $\mathcal{X}(G)$ we clearly have that the distance between two given points $x, y \in \mathcal{X}$ is, up to a difference of 2, the distance between two other points $u, v \in H$ (simply take the closest points to x, y in the subdivision of the corresponding edges), and *vice versa*. Since we know by Theorem 3.6 that w.h.p.

$$\operatorname{diam}(\mathcal{X}) = (2 + o(1))(1/\lambda)\log N,$$

we can now deduce that w.h.p.

$$\operatorname{diam}(H) = (2 + o(1))(1/\varepsilon)\log(\varepsilon^3 n),$$

as required. □

4.2. The diameter of the giant component

We next wish to prove Theorem 1.1, which establishes the asymptotics of the diameter of the giant component.

The next lemma estimates the diameter of a Poisson–Galton–Watson tree. Throughout the proof, let $0 < \mu < 1$ be some function of n satisfying

$$\mu = 1 - \varepsilon + O(\varepsilon^2).$$

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

Lemma 4.2. *Let T be a* PGW(μ)*-tree for μ as above, and let L_k be the kth level of T. For any $k \geqslant 1/\varepsilon$ we have* $\mathbb{P}(L_k \neq \emptyset) = \Theta\big(\varepsilon \exp\big[-k(\varepsilon + O(\varepsilon^2))\big]\big)$.

Proof. Let T' be a Galton–Watson tree with a binomial offspring distribution $\mathrm{Bin}(b, \mu/b)$. Then T' is precisely the open cluster containing the root after percolating on a b-ary tree T_b, where the percolation probability is μ/b. Let L'_k be the kth level of T'. We will use the next lemma, which gives a sharp estimate for the probability that L'_k is non-empty.

Lemma 4.3 (Theorem 2.1 of [18], restated). *Assign each edge e from level $h-1$ to level h in T_b the edge resistance*

$$r_e = \left(1 - \frac{\mu}{b}\right)\left(\frac{\mu}{b}\right)^{-h}.$$

Let R_k be the effective resistance from the root to level k of T_b. Then,

$$\frac{1}{1+R_k} \leqslant \mathbb{P}(L'_k \neq \emptyset) \leqslant \frac{2}{1+R_k}.$$

In our case, the resistance R_k satisfies (see, *e.g.*, [20, Example 8.3])

$$R_k = \sum_{i=1}^{k} \frac{(1-\frac{\mu}{b})(\frac{\mu}{b})^{-i}}{b(b-1)^{i-1}} = \frac{b-1}{b-\mu} \cdot \frac{\big[(1+\frac{1}{b-1})/\mu\big]^{k+1} - 1}{\big[(1+\frac{1}{b-1})/\mu\big] - 1}.$$

Note that for $k \geqslant 1/\varepsilon$ we have

$$\lim_{b\to\infty} R_k = \frac{(1/\mu)^{k+1}-1}{(1/\mu)-1} = \Theta\big(\varepsilon^{-1}\exp\big[k(\varepsilon + O(\varepsilon^2))\big]\big).$$

Applying Lemma 4.3, we obtain that

$$\frac{1}{1+R_k} \leqslant \mathbb{P}(L'_k \neq \emptyset) \leqslant \frac{2}{1+R_k}.$$

Letting $b \to \infty$ and using the fact that $\mathrm{Bin}(b, \frac{\mu}{b})$ converges to a $\mathrm{Poisson}(\mu)$ distribution, we obtain that, for any $k \geqslant 1/\varepsilon$,

$$\mathbb{P}(L_k \neq \emptyset) = \Theta\big(\varepsilon \exp\big[-k(\varepsilon + O(\varepsilon^2))\big]\big),$$

as required. □

We are now ready to prove the main result.

Proof of Theorem 1.1 for the regime $\varepsilon = o(n^{-1/4})$. For N as defined in Theorem 4.1, let $G \sim \mathcal{G}(N,3)$ be a random 3-regular graph with i.i.d. $\mathrm{Exp}(\lambda)$ edge weights, denoted by $\{w(e) : w \in E(G)\}$, where $\mathrm{e}^{-\lambda} = 1-\varepsilon$. Again, let H be the unweighted graph obtained by taking the underlying graph of G, and replacing each of its edges by paths of length i.i.d. geometric variables with mean $1/\varepsilon$.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

In what follows, we will shift between the shortest distances in the metric graph $\mathcal{X}(G)$ and those in the weighted graph H. Since we have

$$1/\lambda = 1/\log(1/(1-\varepsilon)) = (1/\varepsilon) + O(1),$$

Lemma 3.4, together with the aforementioned coupling between the two models, implies that this shift will only cause an error of $O(\log(\varepsilon^3 n))$, which is easily absorbed in our estimate for the diameter.

Set $\delta > 0$. and consider $m = \Theta(\varepsilon^2 n)$ i.i.d. PGW$(1-\varepsilon)$-trees. By Lemma 4.2, the probability that a given such tree will have height at least

$$h^- = (1-\delta)(1/\varepsilon)\log(\varepsilon^3 n)$$

is, for some $c = c(\delta) > 0$, at least

$$c\varepsilon(\varepsilon^3 n)^{-(1-\delta)(1+O(\varepsilon))} = c\varepsilon(\varepsilon^3 n)^{-1+\delta+o(1)} \triangleq \zeta.$$

Therefore, standard estimates for the binomial variable $\mathrm{Bin}(m,\zeta)$ (mean $m\zeta = (\varepsilon^3 n)^{\delta - o(1)}$) imply that there exist at least two such trees w.h.p.

Let H_u denote the height of the PGW-tree attached to a vertex u in the 2-core. By the above discussion, the two vertices u, v with largest H_u, H_v have $H_u, H_v \geqslant h^-$, and clearly they are uniformly distributed among the vertices of $\mathcal{C}_1^{(2)}$ (by the definition of our model). As such, Lemma 3.5 asserts that their distance in G is w.h.p. at least $2t_0 = (1+o(1))\log N$ (neglecting the distance to their nearest kernel vertices in this lower bound) and re-scaling, we obtain that their distance is at least $(1+o(1))(1/\varepsilon)\log(\varepsilon^3 n)$ in the 2-core. Altogether, the distance between a level-H_u leaf of the PGW-tree of u and a level-H_v leaf of the PGW-tree of v is at least

$$(3 - 2\delta + o(1))(1/\varepsilon)\log(\varepsilon^3 n),$$

and letting $\delta \to 0$ completes the lower bound.

For the upper bound, let $B = \{u \in \mathcal{C}_1^{(2)} : H_u \geqslant 1/\varepsilon\}$. Applying Lemma 4.2 and Markov's inequality (as $\mathbb{E}|B| = O(\varepsilon^3 n)$), we infer that w.h.p.

$$|B| \leqslant \varepsilon^3 n \log\log(\varepsilon^3 n). \tag{4.1}$$

Recall that the attached trees are independent of the 2-core and hence B is independent of the structure of the 2-core. Moreover, for any u,

$$\mathbb{P}\big(H_u \geqslant 1/\varepsilon + \ell \mid u \in B\big) = O\big(\mathrm{e}^{-(1+O(\varepsilon))\varepsilon\ell}\big),$$

since we have $\mathbb{P}(u \in B) = \Theta(\varepsilon)$.

Now, let u' be the closest kernel point to u and let $\bar{e}_u$ be the kernel edge incident to u and ending at u'. Define $T_u = T_{\bar{e}_u}(q) + \frac{1}{2}w(\bar{e}_u)$, where $T_{\bar{e}_u}(q)$ is as given in (3.8), *i.e.*, the minimum time t at which $B_t(\bar{e}_u)$ reaches size $q = 2\sqrt{N\log N}$ (adjusted by $\frac{1}{2}w(\bar{e}_u)$ in the case of rare edges).

For $\delta > 0$, set

$$t_n = \left(\frac{1}{2} + \delta\right)(1/\varepsilon)\log(\varepsilon^3 n).$$

Since H_u and T_u are independent, using (3.10) we deduce that for any u

$$\mathbb{P}(H_u \geqslant \ell_1 \mid u \in B,\ T_u) \leqslant O(\mathrm{e}^{-(1+O(\varepsilon))\varepsilon\ell_1}),$$
$$\mathbb{P}(T_u \geqslant t_n + \ell_2 \mid u \in B,\ H_u) \leqslant O(\mathrm{e}^{-(1+O(\varepsilon))\varepsilon\ell_2}) + O(N^{-3/2}).$$

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

We next wish to bound the upper tail of $H_u + T_u$. To this end, let W be a random variable with $\mathbb{P}(W \geqslant x) = \mathrm{e}^{-(1+O(\varepsilon))\varepsilon x}$ and Z be a random variable on $\{0, n\}$ with $\mathbb{P}(Z = n) = O(N^{-3/2})$. Pick independent variables W_1 and W_2 distributed as W and suppose Z is independent of W_1, W_2. Observe that our previous results give that

$$(H_u \mid u \in B,\ T_u) \preceq W_1 \quad \text{and} \quad (T_u - t_n \mid u \in B,\ H_u) \preceq W_2 + Z. \tag{4.2}$$

By our assumption on W,

$$\mathbb{P}(\mathrm{e}^{(1-\delta)\varepsilon W} \geqslant x) = \mathbb{P}(W \geqslant \log x/((1-\delta)\varepsilon)) \leqslant x^{-\frac{1+O(\varepsilon)}{1-\delta}}$$

and therefore $\mathbb{E}\mathrm{e}^{(1-\delta)\varepsilon W} \leqslant O(1/\delta)$ assuming $\varepsilon = o(\delta)$. We now obtain that $\mathbb{E}\mathrm{e}^{(1-\delta)\varepsilon(W_1+W_2)} \leqslant O(1/\delta^2)$ and thus $\mathbb{P}(W_1 + W_2 \geqslant x) \leqslant O(\mathrm{e}^{-(1-\delta)\varepsilon x}/\delta^2)$. This implies that

$$\mathbb{P}(W_1 + W_2 + Z \geqslant x) \leqslant O(\mathrm{e}^{-(1-\delta)\varepsilon x}/\delta^2) + O(N^{-3/2}),$$

and plugging this into (4.2) we conclude that

$$\mathbb{P}(H_u + T_u \geqslant t_n + \ell \mid u \in B) \leqslant O(\mathrm{e}^{-(1-\delta)\varepsilon\ell}/\delta^2) + O(N^{-3/2}).$$

Taking

$$\ell = (1 + 2\delta)(1/\varepsilon)\log(\varepsilon^3 n)$$

and recalling the bound (4.1) on $|B|$, we conclude that w.h.p. every $u \in B$ satisfies

$$H_u + T_u \leqslant \left(\left(\frac{3}{2} + 3\delta\right)\varepsilon^{-1}\log(\varepsilon^3 n)\right).$$

It remains to treat vertices not in B. Again applying (3.10), w.h.p. we have

$$T_u \leqslant (1 + \delta)(1/\varepsilon)\log(\varepsilon^3 n)$$

for all $u \in \mathcal{C}_1^{(2)}$, and thus for all $u \in \mathcal{C}_1^{(2)} \setminus B$,

$$H_u + T_u \leqslant (1/\varepsilon) + (1 + \delta)(1/\varepsilon)\log(\varepsilon^3 n).$$

Altogether, we conclude that w.h.p. every $u \in \mathcal{C}_1^{(2)}$ satisfies

$$H_u + T_u \leqslant \left(\left(\frac{3}{2} + 3\delta\right)(1/\varepsilon)\log(\varepsilon^3 n)\right),$$

and (3.12) now concludes the proof of the upper bound. □

5. Diameters in the supercritical giant component

The goal of this section is to extend the proofs of Theorems 1.1 and 1.2, provided in the previous section for the special case of $\varepsilon = o(n^{-1/4})$, to any $\varepsilon = o(1)$.

The proofs follow from essentially the same arguments, by replacing Theorem 4.1 with its following more general form.

Theorem 5.1 (Theorem 2 of [10]). *Let $\mathcal{C}_1$ be the largest component of $\mathcal{G}(n,p)$ for $p = \frac{1+\varepsilon}{n}$, where $\varepsilon^3 n \to \infty$ and $\varepsilon \to 0$. Let $\mu < 1$ denote the conjugate of $1 + \varepsilon$, that is, $\mu\mathrm{e}^{-\mu} = (1+\varepsilon)\mathrm{e}^{-(1+\varepsilon)}$. Then $\mathcal{C}_1$ is one-sided contiguous to the following model $\tilde{\mathcal{C}}_1$.*

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

(1) *Let* $\Lambda \sim \mathcal{N}\big(1+\varepsilon-\mu, \frac{1}{\varepsilon n}\big)$ *and assign i.i.d. variables* $D_u \sim \text{Poisson}(\Lambda)$ $(u \in [n])$ *to the vertices, conditioned that* $\sum D_u \mathbf{1}_{D_u \geqslant 3}$ *is even. Let* $N_k = \#\{u : D_u = k\}$ *and* $N = \sum_{k\geqslant 3} N_k$. *Select a random multigraph* $\mathcal{K}$ *on* N *vertices, uniformly among all graphs with* N_k *vertices of degree* k *for* $k \geqslant 3$.

(2) *Replace the edges of* $\mathcal{K}$ *by paths of lengths i.i.d.* $\text{Geom}(1-\mu)$.

(3) *Attach an independent* $\text{Poisson}(\mu)$*–Galton–Watson tree to each vertex.*

That is, $\mathbb{P}(\tilde{\mathcal{C}}_1 \in \mathcal{A}) \to 0$ *implies* $\mathbb{P}(\mathcal{C}_1 \in \mathcal{A}) \to 0$ *for any set of graphs* $\mathcal{A}$.

Indeed, a Taylor-expansion of the above defined parameter μ shows that $\mu = 1-\varepsilon+O(\varepsilon^2)$, hence our treatment of the re-scaling of the weighted graph by paths of length i.i.d. $\text{Geom}(1-\mu)$ will be essentially the same as in the case of i.i.d. $\text{Geom}(\varepsilon)$ variables, and the same applies to the $\text{PGW}(\mu)$-trees (rather than $\text{PGW}(1-\varepsilon)$-trees). Step (1), however, is somewhat different here, as the degree distribution of the kernel is richer. Nevertheless, with only minor modifications the original proofs will hold for this case as well. We next address these adjustments that one needs to make.

5.1. The 2-core and its kernel

To extend the upper bounds from the case $\varepsilon = o(n^{-1/4})$ to the general case, consider the kernel $\mathcal{K}$. While before $\mathcal{K}$ was a random 3-regular graph, now it has a degree distribution of a truncated Poisson, resulting in an order $\varepsilon^k n$ expected number of vertices of degree k for $k \geqslant 3$. As we will now show, this will only assist us in establishing the upper bound on the diameter.

Consider the weighted graph G studied in Section 3. A crucial element in the proofs of all the upper bounds was the decay of the distance of two typical vertices in the 2-core. This was achieved by analysing the exploration process starting from a uniformly chosen vertex. While previously every vertex had 3 half-edges, now the degree distribution is more complicated.

- On one hand, encountering a vertex of large degree in the exploration process would contribute extra half-edges to our boundary, and hence accelerate the exposure speed (by increasing the rate of the exponential waiting times τ_i).
- On the other hand, larger degrees might result in a larger tree-excess, slowing down the growth of the number of exposed vertices.

We now show that the effect of the second item is negligible. First, notice that w.h.p. the largest degree in the kernel is at most $\log N$ (in fact, it is $O\big(\frac{\log N}{\log\log N}\big)$ w.h.p.). Assuming that this holds, (3.5) can be replaced by

$$\mathbb{P}(\mathtt{tx}(B_{\tau_r}) \geqslant 1) \leqslant \mathbb{P}\left(\text{Bin}\left(r\log N, (r+2)\frac{\log N}{N}\right) \geqslant 1\right) = O\left(r^2\frac{\log^2 N}{N}\right) = O\left(\frac{\log^7 N}{N}\right),$$

$$\mathbb{P}(\mathtt{tx}(B_{\tau_r}) \geqslant 2) \leqslant \mathbb{P}\left(\text{Bin}\left(r\log N, (r+2)\frac{\log N}{N}\right) \geqslant 2\right) = O\left(r^4\frac{\log^3 N}{N^2}\right) = o\big(N^{-3/2}\big),$$

which is indeed sufficient. Therefore, the probabilities of the events we condition on to control the tree-excess have the same effect. A straightforward stochastic domination argument for the first item above (where the larger degrees are in our favour) now completes the upper bounds specified in Section 3 for the new degree distribution.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

Establishing the lower bounds is slightly more delicate. Fix $\delta > 0$. By standard large-deviation arguments, there exists some $c = c(\delta) > 1$ such that, with probability at least $1 - \delta$,

$$N_3 = \left(\frac{4}{3} + o(1)\right)\varepsilon^3 n, \qquad N_k \leqslant c\frac{(3\varepsilon)^k}{k!}n \quad \text{for all } k \geqslant 4. \tag{5.1}$$

Set $z = \sqrt{N/\log N}$. Let S_i be the number of unmatched half-edges in our explored set at time τ_i (where τ_i denotes the time of the ith step) and let $X_i = S_i - S_{i-1}$. Recall that after $i \leqslant z$ steps we have at least $N_3 - z$ vertices of degree 3 and at most N_k vertices of degree k for $k \geqslant 4$, and that each matching samples a half-edge uniformly at random. With this in mind, the following variable Y would stochastically dominate the degree of this random half-edge; then

$$\mathbb{P}(Y = 3) = \frac{3(N_3 - z)}{3(N_3 - z) + \sum_{j\geqslant 4} jN_j},$$
$$\mathbb{P}(Y = k) = \frac{kN_k}{3(N_3 - z) + \sum_{j\geqslant 4} jN_j} \quad (k = 4, 5, \ldots).$$

Since each matching to a newly exposed degree-k vertex costs one half-edge and introduces $k - 1$ new ones, it follows that $X_i \preceq Y_i - 2$, where the Y_is are i.i.d. with $Y_i \sim Y$. Setting $\tilde{S}_t = \sum_{i=1}^{t}(Y_i - 2)$ we get that $S_t \preceq S_0 + \tilde{S}_t$ for all t.

By our assumption (5.1), for every $k \geqslant 2$

$$\mathbb{P}(Y_i - 2 = k) \leqslant \frac{c(k+2)(3\varepsilon)^{k+2}n/(k+2)!}{(4 - o(1))\varepsilon^3 n} \leqslant c'\frac{(3\varepsilon)^{k-1}}{(k+1)!}$$

for some constant $c' > 1$. Hence, the Laplace transform for $Y_i - 2$ satisfies

$$\mathbb{E}\mathrm{e}^{\lambda(Y_i-2)} \leqslant \mathrm{e}^{\lambda} + \sum_{k=2}^{\infty}\mathrm{e}^{\lambda k}c'\frac{(3\varepsilon)^{k-1}}{(k+1)!} \leqslant c'\mathrm{e}^{\lambda}\exp(\mathrm{e}^{\lambda}3\varepsilon).$$

Setting $\lambda = \log(1/\varepsilon)$, we arrive at $\mathbb{E}\mathrm{e}^{\lambda X_i} \leqslant c'\mathrm{e}^3/\varepsilon$. Now, an application of Markov's inequality gives that, for large n,

$$\begin{aligned}\mathbb{P}\left(\tilde{S}_t \geqslant \left(1 + \frac{1}{\log\log(1/\varepsilon)}\right)t\right) &\leqslant \left(\frac{c'\mathrm{e}^3}{\varepsilon}\right)^t \exp\left(-\lambda t\left(1 + \frac{1}{\log\log(1/\varepsilon)}\right)\right)\\ &\leqslant \left(c'\mathrm{e}^3\exp\left(-\frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)}\right)\right)^t \leqslant 2^{-t},\end{aligned}$$

where the last inequality holds for any large n. We can now infer that

$$\sum_{t\geqslant\sqrt{1/\varepsilon}}\mathbb{P}\left(\tilde{S}_t \geqslant \left(1 + \frac{1}{\log\log(1/\varepsilon)}\right)t\right) = O(2^{-1/\sqrt{\varepsilon}}) = o(1).$$

On the other hand, as $\mathbb{P}(X_i \geqslant 2) \leqslant \mathbb{P}(Y_i \geqslant 4) = O(\varepsilon)$, it follows that

$$\mathbb{P}\big(X_i = 1 \text{ for all } 1 \leqslant i \leqslant 1/\sqrt{\varepsilon}\big) \geqslant 1 - O(\sqrt{\varepsilon}) = 1 - o(1).$$

Altogether, we conclude that, given (5.1), we have that w.h.p.

$$S_t \leqslant S_0 + t + \frac{t}{\log\log(1/\varepsilon)} = (1 + o(1))t + 3 \quad \text{for all } k = 1, 2, \ldots, z, \tag{5.2}$$

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

where we used the fact that S_0 is the degree of the starting vertex for the exploration and is thus equal to 3 w.h.p.

Letting $\delta \to 0$ (recall its definition above (5.1)), we obtain that (5.2) holds w.h.p. Therefore, we may perform the exploration process according to the argument of Lemma 3.5, and at every step t condition that the event in (5.2) indeed holds up to that point, and so the estimate on τ_z immediately follows in the same manner as before. Finally, recall that the lower bound for the distance between kernel points was obtained from the above ingredients together with the asymptotics of the number of kernel vertices via a simple second moment argument. These asymptotics are the same in our new setting, thus completing the proof of the lower bound for the kernel. An analogous argument gives the corresponding lower bound for the metric graph and the 2-core (here the asymptotic number of edges plays a part in the second moment argument).

Finally, note that, as $\mu = 1 - \varepsilon + O(\varepsilon^2)$, writing the final estimate involving the term $1/\varepsilon$ rather than $1/(1-\mu)$ results in a multiplicative factor of $1 + O(\varepsilon) = 1 + o(1)$, keeping the statement valid.

5.2. The giant component

Having extended the treatment of the 2-core and kernel to the case of $\varepsilon = o(1)$, it remains to address the attached trees. Note that Lemma 4.2 applies directly for the value of μ as defined in Theorem 5.1, and hence our estimates for H_u, the height of the tree attached to a vertex u of the 2-core, remain unchanged. As the rest of the arguments are applications of the results for weighted graphs and metric graphs (already discussed in the previous subsection), they hold without modification.

Acknowledgements

We wish to thank Asaf Nachmias for helpful discussions at an early stage of this project and an anonymous referee for thorough comments.

References

[1] Addario-Berry, L., Broutin, N. and Goldschmidt, C. (2009) The continuum limit of critical random graphs. Available at: `http://arxiv.org/abs/0903.4730`.

[2] Alon, N. and Spencer, J. H. (2008) *The Probabilistic Method*, 3rd edn, Wiley-Interscience Series in Discrete Mathematics and Optimization.

[3] Bandelt, H.-J. and Chepoi, V. (2008) Metric graph theory and geometry: A survey. In *Surveys on Discrete and Computational Geometry*, Vol. 453 of *Contemporary Mathematics*, AMS, Providence, RI, pp. 49–86.

[4] 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.

[5] Bollobás, B. (1980) A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Europ. J. Combin.* **1** 311–316.

[6] Bollobás, B. (2001) *Random Graphs*, 2nd edn, Vol. 73 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press.

[7] Bollobás, B. (1984) The evolution of random graphs. *Trans. Amer. Math. Soc.* **286** 257–274.

[8] Bollobás, B., Janson, S. and Riordan, O. (2007) The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** 3–122.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press

[9] Chung, F. and Lu, L. (2001) The diameter of sparse random graphs. *Adv. Appl. Math.* **26** 257–279.
[10] Ding, J., Kim, J. H., Lubetzky, E. and Peres, Y. Anatomy of a young giant component in the random graph. *Random Struct. Alg.*, in press. Available at: `http://arxiv.org/abs/0906.1839`.
[11] Erdős, P. and Rényi, A. (1959) On random graphs I. *Publ. Math. Debrecen* **6** 290–297.
[12] Fernholz, D. and Ramachandran, V. (2007) The diameter of sparse random graphs. *Random Struct. Alg.* **31** 482–516.
[13] Janson, S., Łuczak, T. and Rucinski, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization.
[14] Kesten, H. (1986) Aspects of first passage percolation. In *École d'Été de Probabilités de Saint-Flour, XIV, 1984*, Vol. 1180 of *Lecture Notes in Mathematics*, Springer, pp. 125–264.
[15] Łuczak, T. (1990) Component behavior near the critical point of the random graph process. *Random Struct. Alg.* **1** 287–310.
[16] Łuczak, T. (1998) Random trees and random graphs. *Random Struct. Alg.* (*Proc. Eighth International Conference 'Random Structures and Algorithms'*, Poznan 1997) **13** 485–500.
[17] Łuczak, T. and Seierstad, T. G. The diameter behavior in the random graph process. Preprint.
[18] Lyons, Russell (1992) Random walks, capacity and percolation on trees. *Ann. Probab.* **20** 2043–2088.
[19] Nachmias, A. and Peres, Y. (2008) Critical random graphs: Diameter and mixing time. *Ann. Probab.* **36** 1267–1286.
[20] Peres, Y. (1997) Probability on trees: An introductory climb. In *Lectures on Probability Theory and Statistics* (Saint-Flour), Vol. 1717 of *Lecture Notes in Mathematics*, Springer, pp. 193–280.
[21] Riordan, O. and Wormald, N. C. (2008) The diameter of sparse random graphs. Preprint, available at: `http://arxiv.org/abs/0808.4067v1`.
[22] Riordan, O. and Wormald, N. C. (2009) The diameter of sparse random graphs. To appear, available at: `http://arxiv.org/abs/0808.4067v2`.
[23] Wormald, N. C. (1999) Models of random regular graphs. In *Surveys in Combinatorics 1999*, Vol. 267 of *London Mathematical Society Lecture Notes*, Cambridge University Press, pp. 239–298.

https://doi.org/10.1017/S0963548310000301 Published online by Cambridge University Press