On the Diameter and Radius of Random Subgraphs of the Cube

B. Bollobás[1], Y. Kohayakawa[2], and T. Łuczak[3, 4]

[1] Department of Pure Mathematics and Mathematical Statistics,
University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England

[2] Instituto de Matemática e Estatística, Universidade de São Paulo,
Caixa Postal 20570, 01452–990 São Paulo, SP, Brazil

[3] Mathematical Institute of the Polish Academy of Sciences,
Poznań, Poland

[4] On leave from the Department of Discrete Mathematics,
Adam Mickiewicz University, ul. Matejki 48/49, Poznań, Poland

Abstract. The n-dimensional cube Q^n is the graph whose vertices are the subsets of $\{1, \ldots, n\}$, with two vertices adjacent if and only if their symmetric difference is a singleton. Clearly Q^n has diameter and radius n. Write $M = n2^{n-1} = e(Q^n)$ for the size of Q^n. Let $\widetilde{Q} = (Q_t)_0^M$ be a random Q^n-process. Thus Q_t is a spanning subgraph of Q^n of size t, and Q_t is obtained from Q_{t-1} by the random addition of an edge of Q^n not in Q_{t-1}. Let $t^{(k)} = \tau(\widetilde{Q}; \delta \geq k)$ be the hitting time of the property of having minimal degree at least k. We show that the diameter $d_t = \mathrm{diam}(Q_t)$ of Q_t in almost every $\widetilde{Q}$ behaves as follows: d_t starts infinite and is first finite at time $t^{(1)}$, it equals $n+1$ for $t^{(1)} \leq t < t^{(2)}$, and $d_t = n$ for $t \geq t^{(2)}$. We also show that the radius of Q_t is first finite for $t = t^{(1)}$, when it assumes the value n. These results are deduced from detailed theorems concerning the diameter and radius of the almost surely unique largest component of Q_t for $t = \Omega(M)$.

1. Introduction

Let Q^n be the *n-dimensional cube*, the graph whose vertices are the subsets of $[n] = \{1, \dots, n\}$ and where two such vertices are adjacent if and only if their symmetric difference is a singleton. Note that both the diameter and the radius of Q^n are n. Write $N = 2^n = |Q^n|$ for the order of Q^n and $M = n2^{n-1} = e(Q^n)$ for the size of Q^n. Let $\widetilde{Q} = (Q_t)_0^M$ be a *random Q^n-processes*. This is a Markov chain whose states are spanning subgraphs of Q^n and Q_t $(1 \le t \le M)$ is obtained from Q_{t-1} by the addition of an edge of Q^n not in Q_t, with this edge chosen uniformly at random from all the possibilities. We are interested in the behaviour of $\widetilde{Q}$ for large n; thus we use the terms 'almost surely' and 'almost every' to mean 'with probability tending to 1 as $n \to \infty$'.

If P is a non-trivial monotone increasing property of spanning subgraphs of Q^n we let $\tau_P = \tau(P) = \tau(\widetilde{Q}; P)$ be the hitting time of P in the process $\widetilde{Q} = (Q_t)_0^M$, that is

$$\tau_P = \tau(P) = \tau(\widetilde{Q}; P) = \min\{t : Q_t \text{ has } P\}.$$

The events P we shall consider here are, amongst others, (i) the event $\{\delta \ge k\}$ that the minimal degree should be at least k, (ii) the event $\{\text{diam} < \infty\}$ that the graph should be connected, (iii) the event $\{\text{rad} \le r\}$ that the radius should be at most r, and (iv) the event $\{\text{diam} \le d\}$ that the diameter should be at most d.

One of our main results implies that almost surely

$$\tau(\widetilde{Q}; \delta \ge 1) = \tau(\widetilde{Q}; \text{connectedness}) = \tau(\widetilde{Q}; \text{diam} \le n+1) = \tau(\widetilde{Q}; \text{rad} = n), \tag{1}$$

and

$$\tau(\widetilde{Q}; \text{diam} = n) = \tau(\widetilde{Q}; \delta \ge 2). \tag{2}$$

The fact that the hitting time of connectedness and $\tau(\delta \ge 1)$ almost surely coincide was proved as a remark in [3], where the hitting time of the existence of a perfect matching was shown to equal $\tau(\delta \ge 1)$ almost surely. Other results concerning matchings in random subgraphs of the n-cube are due to Kostochka [13], who investigated in great detail the size of the maximum matching in binomial random subraphs of Q^n. (See also Kostochka [14] for generalisations.)

It follows from the relations above that the diameter of Q_t in a typical $\widetilde{Q} = (Q_t)_0^M$ behaves in the following very simple way. At the beginning of the process the diameter is infinite and when the isolated vertices disappear it becomes finite for the first time, when it assumes value $n+1$. Then it decreases to $n = \mathrm{diam}(Q^n)$, its minimal possible value, exactly when the vertices of degree 1 disappear. As to the radius of Q_t, we see that it becomes finite exactly when the isolated vertices disappear, when it assumes value $n = \mathrm{rad}(Q^n)$.

The results above are obtained as simple corollaries to results concerning the almost surely unique largest component, the 'giant' component (see [17, 18, 19], and for more sophisticated results see [1] and [5]), of Q_t for $t = \Omega(M)$. These results, namely Theorem 10, Corollary 11, and Theorem 13, describe in detail the behaviour of $\mathrm{diam}(L_t)$ and $\mathrm{rad}(L_t)$, where $L_t = L_t(Q_t)$ is the giant component of Q_t, as $t = \Omega(M)$ grows.

Roughly speaking, Corollary 11(i) states that for all fixed k, in a typical process, $\mathrm{diam}(L_t)$ changes from $n+k$ to $n+k-1$ at a sharply defined time, namely $t_k = M(1 - 2^{-1/k}(1-(\log n)/n))$. The behaviour of $\mathrm{diam}(L_t)$ at around these critical values t_k is also obtained. (See Corollary 11(ii).) Note that the results above completely describe $\mathrm{diam}(L_t)$ for any $t = \Omega(M)$. Theorem 13 shows that, perhaps a little surprisingly, the radius $\mathrm{rad}(L_t)$ of L_t is rather stable. Indeed, that result says that almost surely $\mathrm{rad}(L_t) = n-1$ for all $\Omega(M) = t < t^{(1)} = \tau(\widetilde{Q}; \delta \geq 1)$.

Thus, typically, $\mathrm{diam}(L_t)$ decreases steadily as t increases while $\mathrm{rad}(L_t)$ stays constant for a long while at $n-1$, until it *increases* to n at time $t^{(1)}$. This is in sharp contrast with the case of ordinary random graph processes $\widetilde{G} = (G_t)_t$. Indeed, as proved by Burtin (see [8, 9]), if $t = (\omega(\log n)/n)\binom{n}{2}$ and $\omega = \omega(n) \to \infty$ as $n \to \infty$ then almost surely $\mathrm{diam}(G_t) - 1 \leq \mathrm{rad}(G_t) \leq \mathrm{diam}(G_t)$. Thus, in the later stages of an ordinary random graph process, the diameter *as well as* the radius of the evolving graph decrease gradually. Questions concerning the diameter of ordinary random graphs $G_p \in \mathcal{G}(n,p)$ and $G_t \in \mathcal{G}(n,t)$ have been extensively studied, and very precise results are now known for a wide range of p and t. The interested reader is referred to [2, Chapter X], and to Burtin [8, 9].

Finally, we remark that the radius and diameter of random subgraphs of Q^n were also studied in Mahrhold and Weber [12], Kostochka, Sapozhenko, and Weber [16], and Sapozhenko and Weber [18]. However, the model studied by these authors is the one known as the 'mixed' model, where the subgraph of Q^n is chosen by the random deletion of vertices and edges from Q^n. The best results for this model given in the articles above are the ones in [16], where the authors determine the radius of the giant component L up to an additive constant of 2, and the diamater of L up to an additive constant of 8. Let us mention that this model is also studied in Dyer, Frieze, and Foulds [10], where connectivity problems are investigated.

This note is organised as follows. In the next section we give the fundamental lemma, Lemma 2, that is the key to many of our results. In Section 3 we give three preliminary results, which are fully developed in the following two sections into our main theorems concerning the diameter and the radius of the giant component of Q_t. In the last section we mention a problem closely related to the ones we address here.

2. The fundamental lemma

For the study of random Q^n-processes it is often convenient to look first at the binomial model Q_p of a random subgraph of the cube Q^n. As usual, given a graph H and $0 \le p \le 1$, we write $\mathcal{G}(H, p)$ for the space of random spanning subgraphs H_p of H such that every edge of H belongs to H_p independently with probability p. For a set X and $k \ge 0$ we let $X^{(k)}$ denote the set of all k-subsets of X.

Our fundamental lemma, Lemma 2, will be proved in a slightly more general form than needed in this paper. This is done because this version requires only a little more work, and will be used in [7] to study connectivity properties of random subgraphs of Q^n. Before we proceed, recall that Q^n has vertex set $\mathcal{P}([n]) = 2^{[n]}$. To state our result, we shall denote by $Q^{n[-l]}$ a generic graph obtained from Q^n by the deletion of some vertices in such a way that both $\emptyset$ and $[n]$ are left in $Q^{n[-l]}$ and, for every k $(1 \le k \le n-1)$, no more than l vertices from $[n]^{(k)}$ are missing.

Let us start with the following simple observation concerning $Q^{n[-l]}$.

Lemma 1. *The graph $Q^{n[-l]}$ contains at least $n!\,(1 - O(l/n))$ paths of length n joining $\emptyset$ and $[n]$.*

Proof. Let us call a path $P = v_0 v_1 \dots v_s$ in Q^n *proper* if $v_0 = \emptyset$ and $v_s \in [n]^{(s)}$. Let u_i denote the number of proper paths of length i in $Q^{n[-l]}$. Note that for $1 \le i < n$ we have $u_i \ge u_{i-1}(n-i+1) - li!$. Therefore

$$\frac{u_i}{(n)_i} \ge \frac{u_{i-1}}{(n)_{i-1}} - l \binom{n}{i}^{-1},$$

where as usual $(a)_b = a(a-1)\cdots(a-b+1)$. Hence

$$u_n = u_{n-1} \geq n!\left(1 - l\sum_{i=1}^{n-1}\binom{n}{i}^{-1}\right) \geq n!\left(1 - O\left(\frac{l}{n}\right)\right),$$

as required. □

The cornerstone of our paper is Lemma 2 below, which states that the probability that two large sets of vertices should not be connected by a short path in $Q_p^{[-l]} \in \mathcal{G}(Q^{n[-l]}, p)$ is superexponentially small. The proof we give for this result is based on a simple 'splitting' argument for the (large) probability p and the second moment method. Independently, Fill and Pemantle [11] have used the second moment method to prove results of this nature in Q^n; in particular, they determined the oriented first-passage time between two vertices furthest apart in Q^n using an 'enhanced' second moment method. (For improvements of some results in [11], see [4]). Lemmas in the spirit of Lemma 2 below may also be found in Sapozhenko [17], Sapozhenko and Weber [18], and Toman [19]. With Lemma 2 in hand, the proof of the main results will follow a rather natural course, although we shall encounter several technical dificulties on the way.

In the proof below and in the sequel, we only assume our inequalities to be valid for large enough n.

Lemma 2. *Let $l \in \mathbf{N}$ be fixed, and suppose that $0 < \varepsilon = \varepsilon(n) \leq 1$ and that $(\log\log n)/n < p = p(n) < 1$. Then, for all $S \subset [n]^{(1)}$ and $T \subset [n]^{(n-1)}$ with $|S|,\ |T| \geq n^{(1+\varepsilon)/2}$, the probability that in $Q_p^{[-l]} \in \mathcal{G}(Q^{n[-l]}, p)$ there is no S—T path of length $n-2$ is bounded from above by $\exp\{-\varepsilon pn(\log n)/\log\log n\}$.*

Proof. Let $\mathcal{P} = \{P : P \text{ an } S\text{—}T \text{ path in } Q^{n[-l]} \text{ of length } n-2\}$, and let $s = |S|$ and $t = |T|$. For each $v' \in [n]^{(1)}$ and $v'' \in [n]^{(n-1)}$ such that $v' \subset v''$, the graph induced by all vertices $w \in Q^{n[-l]}$ such that $v' \subset w \subset v''$ can be identified with $Q^{(n-2)[-l]}$, so from Lemma 1 we get

$$u = |\mathcal{P}| \geq (st - s)(n-2)!(1 - O(l/n)) \geq \frac{1}{3}st(n-2)! \geq n!/3n^{1-\varepsilon}.$$

Let $\omega = \omega(n)$ with $\log\log n \leq \omega = o(n)$ be fixed. Set $p_0 = \omega/n$. We shall first study the space $\mathcal{G}(Q^{n[-l]}, p_0)$, and, to emphasize this fact, we shall write $\mathbb{E}_{p_0}$ for the expectation and $\mathbb{P}_{p_0}$ for the probability in this space.

Let $X = X(Q_{p_0}^{[-l]})$ be the number of paths in $\mathcal{P}$ that are present in $Q_{p_0}^{[-l]}$. Then

$$\mu = \mathbb{E}_{p_0}(X) = up_0^{n-2} \geq \frac{n!}{3n^{1-\varepsilon}} \left(\frac{\omega}{n}\right)^{n-2} \geq \frac{1}{n^{1-\varepsilon}} \left(\frac{\omega}{\mathrm{e}}\right)^n,$$

and hence μ is rather large. We shall now compute the variance $\sigma^2 = \sigma^2(X) = \mathrm{Var}(X)$ of X and show that $\sigma^2 = O(\mu^2/\omega n^\varepsilon)$. Let us first note that the second factorial moment of X is

$$\begin{aligned} \mathbb{E}_2(X) = \mathbb{E}_{p_0}(X(X-1)) &= \sum_{P\in\mathcal{P}} \sum_{P\neq Q\in\mathcal{P}} \mathbb{P}_{p_0}(P,\, Q \subset Q_{p_0}^{[-l]}) \\ &= \sum\nolimits_1 \mathbb{P}_{p_0}(P,\, Q \subset Q_{p_0}^{[-l]}) + \sum\nolimits_2 \mathbb{P}_{p_0}(P,\, Q \subset Q_{p_0}^{[-l]}), \end{aligned}$$

where $\sum_1$ indicates sum taken over all ordered pairs $(P, Q) \in \mathcal{P} \times \mathcal{P}$ of edge-disjoint paths P and Q, that is such that $E(P) \cap E(Q) = \emptyset$, and $\sum_2$ indicates sum over all the other pairs (P, Q). Clearly

$$\sum\nolimits_1 \mathbb{P}_{p_0}(P,\, Q \subset Q_{p_0}^{[-l]}) = \sum\nolimits_1 \mathbb{P}_{p_0}(P \subset Q_{p_0}^{[-l]})\mathbb{P}_{p_0}(Q \subset Q_{p_0}^{[-l]}) \leq (\mathbb{E}_{p_0}(X))^2.$$

Let us now estimate the sum over ordered pairs of edge-intersecting paths (P, Q). For $k \geq 1$ and $P \in \mathcal{P}$ let $\mathcal{P}_{P,k} = \{Q \in \mathcal{P} : |E(P) \cap E(Q)| = k\}$, and set

$$S_k = \sum_{P\in\mathcal{P}} \sum_{Q\in\mathcal{P}_{P,k}} \mathbb{P}_{p_0}(P,\, Q \subset Q_{p_0}) \geq \sum_{P\in\mathcal{P}} \sum_{Q\in\mathcal{P}_{P,k}} \mathbb{P}_{p_0}(P,\, Q \subset Q_{p_0}^{[-l]}).$$

In order to estimate S_k we shall give bounds on $N_{P,k} = |\mathcal{P}_{P,k}|$ for fixed $P \in \mathcal{P}$.

For any $Q \in \mathcal{P}_{P,k}$ there is an integer vector $\mathbf{i} = \mathbf{i}(Q) = (i_j)_1^k$ with $1 \leq i_1 < \cdots < i_k \leq n-2$ such that the i_jth edge of P and Q coincide for all $1 \leq j \leq k$. Moreover, given $\mathbf{i} = (i_j)_1^k$ as above, there are at most

$$N_{\mathbf{i}} = i_1!(n - i_k - 1)! \prod_{j=1}^{k-1} (i_{j+1} - i_j - 1)!$$

paths $Q \in \mathcal{P}$ such that the i_jth edge of P and Q coincide.

Let $\mathbf{i} = (i_j)_1^k$ as above be given. Define $\mathbf{a} = \mathbf{a}(\mathbf{i}) = (a_j)_0^k$ by setting

$$a_j = \begin{cases} i_1 & \text{if } j = 0 \\ i_{j+1} - i_j - 1 & \text{if } 1 \leq j \leq k-1 \\ n - i_k - 1 & \text{if } j = k. \end{cases}$$

Note that then a_0, $a_k \geq 1$, $a_j \geq 0$ $(1 \leq j \leq k-1)$ and $\sum_0^k a_j = n-k$. Moreover

$$N_{\mathbf{i}} = \prod_0^k a_j = (n-k)! \binom{n-k}{a_0, \ldots, a_k}^{-1} \leq (n-k)!.$$

Thus we have that

$$N_{P,k} \leq \binom{n-2}{k} (n-k)!, \tag{3}$$

and this crude bound will suffice for $k \geq n/10$. Let us now assume that $k < n/10$. We shall estimate $N_{P,k}$ in this range of k more carefully.

Let us write $N_{P,k} = N_{P,k}^{(1)} + N_{P,k}^{(2)}$, where $N_{P,k}^{(1)}$ is the number of paths $Q \in \mathcal{P}_{P,k}$ such that if $\mathbf{i} = \mathbf{i}(Q)$ and $\mathbf{a} = \mathbf{a}(\mathbf{i}) = (a_j)_0^k$ then $a_j \leq n-3k$ for all $0 \leq j \leq k$, and naturally $N_{P,k}^{(2)} = N_{P,k} - N_{P,k}^{(1)}$.

Now, if $\max_{0 \leq j \leq k} a_j \leq n-3k$, then

$$\prod_0^k a_j \leq (n-3k)! \max \left\{ \prod_0^{k-1} b_j : \sum_j b_j = 2k \right\} \leq (n-3k)!(2k)!,$$

and so

$$N_{P,k}^{(1)} \leq \binom{n-2}{k} (n-3k)!(2k)!. \tag{4}$$

To estimate $N_{P,k}^{(2)}$, note that if $Q \in \mathcal{P}_{P,k}$ is counted in $N_{P,k}^{(2)}$ then, for $\mathbf{i} = \mathbf{i}(Q)$, the vector $\mathbf{a} = \mathbf{a}(\mathbf{i}) = (a_j)_0^k$ is such that there is a unique j_0 $(0 \leq j_0 \leq k)$ with $a_{j_0} \geq n-3k+1$, since $\sum_j a_j = n-k$ and $k < n/10$. The number of such vectors $\mathbf{a} = (a_j)_0^k$ with $\sum_0^k a_j = n-k$ is $(k+1)\binom{3k-1}{k}$. Thus

$$N_{P,k}^{(2)} \leq (k+1) \binom{3k-1}{k} (n-k-1)!. \tag{5}$$

Therefore, putting together (3), (4) and (5), we have that for all $k \geq 1$

$$S_k \leq u \binom{n-2}{k} (n-k)! p_0^{2(n-2)-k}$$

and if $k < n/10$ then

$$S_k \leq u \left\{ \binom{n-2}{k} (n-3k)!(2k)! + (k+1) \binom{3k-1}{k} (n-k)! \right\} p_0^{2(n-2)-k}.$$

Hence, recalling that $\mu = \mathbb{E}_{p_0}(X) = up_0^{n-2}$ and $u \geq n!/3n^{1-\varepsilon}$, we see that

$$\begin{aligned}\mu^{-2} \sum_{k \geq n/10} S_k &\leq \frac{3n^{1-\varepsilon}}{n!} \sum_{k \geq n/10} \binom{n-2}{k} (n-k)! p_0^{-k} \\ &\leq 3n^{1-\varepsilon} \sum_{k \geq n/10} \frac{1}{k! p_0^k} \leq 4n^{1-\varepsilon} \left(\frac{10\mathrm{e}}{\omega}\right)^{n/10},\end{aligned}$$

and

$$\begin{aligned}\mu^{-2} \sum_{1 \leq k < n/10} S_k \leq \frac{3n^{1-\varepsilon}}{n!} &\left\{ \sum_{1 \leq k < n/10} \binom{n}{k} (n-3k)!(2k)! p_0^{-k} \right. \\ &\left. \quad + \sum_{1 \leq k < n/10} (k+1) 2^{3k-1} (n-k-1)! p_0^{-k} \right\} \\ \leq 3n^{1-\varepsilon} &\left\{ \sum_{1 \leq k < n/10} \frac{k!}{(n-k)_{2k} p_0^k} \binom{2k}{k} + \sum_{1 \leq k < n/10} \frac{(k+1)2^{3k}}{(n)_{k+1} p_0^k} \right\} \\ \leq 3n^{1-\varepsilon} &\sum_{1 \leq k < n/10} \left\{ \frac{k^k 2^{2k}}{(7n/10)^{2k} (\omega/n)^k} + \frac{k+1}{n} \left(\frac{8}{(9n/10)(\omega/n)} \right)^k \right\} \\ \leq 3n^{1-\varepsilon} &\sum_{1 \leq k < n/10} \left\{ \left(\frac{9k}{n\omega}\right)^k + \frac{k+1}{n} \left(\frac{9}{\omega}\right)^k \right\} \\ \leq 3n^{1-\varepsilon} &\left(\frac{10}{n\omega} + \frac{19}{n\omega} \right) \leq \frac{90}{\omega n^{\varepsilon}}.\end{aligned}$$

Thus $\sigma^2(X)/\mu^2 = O(1/\omega n^{\varepsilon})$ and by Chebyshev's inequality we have $\mathbb{P}_{p_0}(X = 0) = O(1/\omega n^{\varepsilon})$.

We now turn to the original space $\mathcal{G}(Q^{n[-l]}, p)$, where $p = p(n)$ is as given. Write H for a fixed $H = Q^{n[-l]}$. Let us generate a random spanning subgraph G of H by letting $G = \bigcup_1^K H_{p_0}$, where $K = \lfloor pn/\omega \rfloor$, $\omega = \omega(n) = \log\log n$, $p_0 = \omega/n$, and the $H_{p_0} \in \mathcal{G}(H, p_0)$ are chosen independently. Note that $p \geq Kp_0$, and hence it suffices to prove the claimed upper bound for G. By the inequality above, the probability that there is no S—T path of length $n-2$ in G is at most $\{O(1/\omega n^{\varepsilon})\}^K \leq \exp\{-\varepsilon pn (\log n)/\log\log n\}$, completing the proof of our lemma. $\square$

Before we can apply Lemma 2, we need to deal with some technical details concerning the distribution of the neighbours of most vertices in $Q_p \in \mathcal{G}(Q^n, p)$. We start by introducing some further notation and terminology. A *subcube* of Q^n is a subgraph induced in Q^n by a set of the form

$$Q_{S,A} = Q(S,A) = \{x \in Q^n : x \cap S = A\},$$

where $A \subset S \subset [n]$. In the sequel we shall denote the subcube induced by $Q_{S,A}$ by $Q_{S,A} = Q(S,A)$ as well, since this will not cause any confusion. The *dimension* $\dim Q_{S,A}$ of $Q_{S,A}$ is $n-|S|$. Given $x, y \in Q^n$ define the subcube $\langle x,y\rangle = Q\left((x \triangle y)^{\mathrm{c}}, x \cap y\right)$, where $u^{\mathrm{c}} = [n] \setminus u$ and as usual $\triangle$ denotes symmetric difference. In particular, if $x \subset y$ then $\langle x,y\rangle$ equals the 'interval' $[x,y]$ in $Q^n = \mathcal{P}([n])$ regarded as a partial order under inclusion, that is

$$\langle x,y\rangle = [x,y] = \{z \in Q^n : x \subset z \subset y\}.$$

In what follows, given a subcube $Q = \langle x,y\rangle$, we may apply the automorphism $u \in Q^n \mapsto x \triangle u \in Q^n$ of Q^n to translate Q so that it is mapped onto the interval $[\emptyset, x \triangle y]$. In other words, it will suffice to consider subcubes $Q = \langle x,y\rangle$ with $x = \emptyset$. Now let $G \subset Q^n$ be a spanning subgraph of Q^n, and let $u, v \in Q^n$ be given. Let $Q = \langle u,v\rangle$ and let $G' = G[Q]$ be the graph induced by the vertices of $Q = \langle u,v\rangle$ in G. As usual, denote the degree of a vertex $z \in G'$ by $d_{G'}(z)$. Then, we say that $Q = \langle u,v\rangle$ is (u,v)-*good for* G if $d_{G'}(u)$, $d_{G'}(v) \geq m^{2/3}$, where $m = \dim Q$ is the dimension $|u \triangle v|$ of Q.

In the sequel, if J is a graph and $x, y \in J$, we write $d_J(x,y)$ for the distance between x and y in J. In particular, if $x, y \in Q^n$, then $d_{Q^n}(x,y) = |x \triangle y| = \dim\langle x,y\rangle$. Moreover, as usual, if $z \in J$, then $\Gamma(z) = \Gamma_J(z)$ is the set of neighbours of z in J, and if $Z \subset V(J)$, then $\Gamma(Z) = \Gamma_J(Z) = \bigcup_{z\in Z}\Gamma_J(z)$. We may now state a simple technical lemma.

Lemma 3. *Let $x, y \in Q^n$ with $d_{Q^n}(x,y) \geq n/5$ be fixed, and suppose $N_x \subset \Gamma_{Q^n}(x)$, $N_y \subset \Gamma_{Q^n}(y)$ are such that $|N_x|, |N_y| \geq 2n^{2/3}$. Let $1/\log\log n \leq p = p(n) < 1$, and write P_{xy} for the following property concerning $Q_p \in \mathcal{G}(Q^n,p)$.*

(P_{xy}) *There is a vertex $x_0 \in N_x$ and $y_0 \in N_y \cup \Gamma_{Q_p}(N_y)$ such that $d_{Q^n}(x_0,y_0) \leq d_{Q^n}(x,y)+1$ and $\langle x_0,y_0\rangle$ is (x_0,y_0)-good for Q_p.*

Then the probability that P_{xy} holds is at least $1-\exp\{-\Omega(n^{4/3}/\log\log n)\}$.

Proof. We shall analyse two cases. For convenience, considering the automorphism $u \in Q^n \mapsto x \triangle u \in Q^n$ of Q^n, we may and shall assume that $x = \emptyset$. Since $d_{Q^n}(x,y) \geq n/5 \geq 2n^{2/3}$, one of the following two cases must occur.

Case 1. $|[x, y] \cap N_x| \geq n^{2/3}$

Choose $\ell = \lfloor n^{2/3} \rfloor$ vertices $x_1, \ldots, x_\ell$ in $[x, y] \cap N_x$. Clearly there are ℓ vertices $y_1, \ldots, y_\ell \in N_y$ such that $x_i \subset y_i$ for all $1 \leq i \leq \ell$. Note that $d_{Q^n}(x_i, y_i) = |x_i \triangle y_i| = |y_i \setminus x_i| \leq |y| = d_{Q^n}(x, y)$. Let us consider the subcubes $Q_i = [x_i, y_i]$. Then the probability of the event A_i that Q_i should not be (x_i, y_i)-good for $Q_p \in \mathcal{G}(Q^n, p)$ is at most $2\mathbb{P}(X \leq m_i^{2/3}) = \exp\{-\Omega(n/\log\log n)\}$, where X is a binomial random variable with parameters p and $m_i = \dim Q_i$. The events A_i $(1 \leq i \leq \ell)$ are independent and so we have that the probability that P_{xy} fails is $\exp\{-\Omega(n\ell/\log\log n)\} = \exp\{-\Omega(n^{5/3}/\log\log n)\}$.

Case 2. $|N_x \setminus [x, y]| \geq n^{2/3}$

Let $N_x \setminus [x, y] = \{x_1, \ldots, x_\ell\}$ where $\ell = |N_x \setminus [x, y]| \geq n^{2/3}$. We shall analyse two subcases. Suppose $x_j = x \cup \{e_j\} = \{e_j\}$ $(1 \leq j \leq \ell)$. Now let $N_y^+ = N_y \cap \{z \in Q^n : |z| = |y| + 1\}$, and similarly $N_y^- = N_y \cap \{z \in Q^n : |z| = |y| - 1\}$. Suppose $N_y^+ = \{y_1^+, \ldots, y_r^+\}$, where $y_i^+ = y \cup \{f_i\}$ $(1 \leq i \leq r)$. Let $I = \{e_j : 1 \leq j \leq \ell\} \cap \{f_i : 1 \leq i \leq r\}$. Now, if $|I| \geq n^{2/3}$ then Case 2(a) below holds. On the other hand, if $|I| < n^{2/3}$, then $|N_y^- \cup (\{f_i : 1 \leq i \leq r\} \setminus \{e_j : 1 \leq j \leq \ell\})| \geq n^{2/3}$, and Case 2$(b)$ holds.

Case 2(a) There are $\ell' \geq n^{2/3}$ vertices $y_1, \ldots, y_{\ell'} \in N_y^+$ such that for all $1 \leq i \leq \ell'$ there is a $j = j(i)$ $(1 \leq j \leq \ell)$ such that $x_j \subset y_i$.

Note that $j = j(i)$ is uniquely defined for each i, that is the map $i \mapsto j = j(i)$ is injective. Let us consider the cubes $Q_i = [x_{j(i)}, y_i]$ $(1 \leq i \leq \ell')$, and note that as in Case 1 above the probability that P_{xy} fails is $\exp\{-\Omega(n^{5/3}/\log\log n)\}$.

Case 2(b) There are $\ell' \geq n^{2/3}$ vertices $y_1, \ldots, y_{\ell'} \in N_y$ such that $x_j \not\subset y_i$ for any $1 \leq i \leq \ell'$ and $1 \leq j \leq \ell$.

Let $1 \leq i \leq \ell'$. We shall say that the vertices $y_i \cup \{e_j\} \in \Gamma_{Q^n}(y_i)$ $(1 \leq j \leq \ell)$ are *i-suitable*. Note that the number of i-suitable vertices is clearly ℓ. Let $p_0 = p/2$ and pick $Q_{p_0} \in \mathcal{G}(Q^n, p_0)$ randomly. The probability of the event B_i that $\Gamma_{Q_{p_0}}(y_i)$ should not contain at least $n^{1/3}$ i-suitable vertices is at most

$$\mathbb{P}(S_{\ell,p_0} < n^{1/3}) \leq 2\binom{\ell}{\lfloor n^{1/3} \rfloor}(1 - p/2)^{\ell - n^{1/3}} = \exp\{-\Omega(n^{2/3}/\log\log n)\},$$

where S_{ℓ,p_0} is a binomial random variable with parameters ℓ and p_0. The events B_i are independent and hence clearly the probability that no B_i occurs is $\exp\{-\Omega(n^{4/3}/\log\log n)\}$.

Suppose B_{i_0} occurs. Let $y'_1, \ldots, y'_{\ell''}$ be the i_0-suitable vertices adjacent to y_{i_0} in Q_{p_0}. We know that $\ell'' \geq n^{1/3}$. For each $1 \leq k \leq \ell''$ let $1 \leq j = j(k) \leq \ell$ be the unique j such that $x_j \subset y'_k$. Note that then $k \mapsto j = j(k)$ is an injective map. Consider the subcubes $Q_k = [x_{j(k)}, y'_k]$ for $1 \leq k \leq \ell''$ and add edges to Q_{p_0} independently, each with probability $p/2$. It follows that the probability that P_{xy} should fail for $Q_p \in \mathcal{G}(Q^n, p)$ is $\exp\{-\Omega(n^{4/3}/\log\log n)\}$.

Therefore in Case 2 we again have that the probability that the condition P_{xy} fails is as small as required. □

3. Preliminary results

In this section, we apply Lemmas 2 and 3 to show that if p is not too small, then a.e. $Q_p \in \mathcal{G}(Q^n, p)$ is such that any two of its vertices of large degree are connected by a path of length at most n. We also show that this holds for a.e. $Q_t \in \mathcal{G}(Q^n, t)$, where as usual $\mathcal{G}(Q^n, t)$ is the space of all spanning subgraphs of Q^n with t edges, all such graphs being equiprobable.

Let x, $y \in Q^n$ be fixed. For all $0 \leq d \leq n$ and $0 < p = p(n) < 1$, write $\mathcal{G}_{x,y,d}(Q^n, p)$ for the conditional probability space obtained from $\mathcal{G}(Q^n, p)$ by conditioning on the event $\{d_{Q_p}(x),\, d_{Q_p}(y) \geq d\}$. We now use Lemmas 2 and 3 to show the following result concerning $\mathcal{G}_{x,y,d}(Q^n, p)$. We remark that an analogous result for random Boolean functions may be found in Sapozhenko [17].

Lemma 4. *Suppose $3/\log\log n \leq p = p(n) < 1$ and $d = d(n) = 2n^{2/3}$. Let x, $y \in Q^n$ be two fixed vertices in Q^n. Then, with probability $1 - \exp\{-\Omega(n \log\log n)\}$, in the space $\mathcal{G}_{x,y,d}(Q^n, p)$ we have*

$$\begin{aligned} d_{Q_p}(x,y) &= d_{Q^n}(x,y) \quad \text{if} \quad d_{Q^n}(x,y) \geq n - n^{2/3}, \\ d_{Q_p}(x,y) &\leq d_{Q^n}(x,y) + 4 \quad \text{if} \quad d_{Q^n}(x,y) \geq n/5, \\ d_{Q_p}(x,y) &\leq 0.8n \quad \text{if} \quad d_{Q^n}(x,y) \leq n/5. \end{aligned} \tag{6}$$

Proof. We may and shall assume that $x = \emptyset$. Let $S \subset \Gamma_{Q^n}(x)$, $T \subset \Gamma_{Q^n}(y)$ be fixed and suppose $0 < p < 1$. Let $\mathcal{G}_{S,T}(Q^n, p)$ be the space obtained from $\mathcal{G}(Q^n, p)$ by conditioning on the event $\{\Gamma_{Q_p}(x) = S,\, \Gamma_{Q_p}(y) = T\}$. To prove our lemma, it is enough to show that if $|S|$, $|T| \geq 2n^{2/3}$ then, for $Q_p \in \mathcal{G}_{S,T}(Q^n, p)$ we have that $d_{Q_p}(x, y)$ satisfies (6) with probability $1 - \exp\{-\Omega(n \log\log n)\}$. Let us then fix S and T as above, and proceed to prove this assertion.

Let $p_0 = p_1 = p_2 = 1/\log\log n$. Without loss of generality $x = \emptyset$. We analyse three cases.

Case 1. $d_{Q^n}(x, y) = |y| \geq n - n^{2/3}$

Set $S' = S \cap [x, y]$, $T' = T \cap [x, y]$. Then $|S'|, |T'| \geq n^{2/3}$, and hence we may apply Lemma 2 to $[x, y]$. We conclude that in $Q_{p_0} \in \mathcal{G}_{S,T}(Q^n, p_0)$ there is an x–y path of length $d_{Q^n}(x, y)$ with probability $1 - \exp\{-\Omega(n \log\log n)\}$.

Case 2. $n/5 \leq d_{Q^n}(x, y) = |y| < n - n^{2/3}$

Let $N_x = S$, $N_y = T$, and pick $Q_{p_0} \in \mathcal{G}_{S,T}(Q^n, p_0)$ randomly. According to Lemma 3, we know that property P_{xy} holds with probability $1 - \exp\{-\Omega(n^{4/3}/\log\log n)\}$. Hence we may and shall assume that P_{xy} does hold for Q_{p_0}. Let x_0 and y_0 be as in Lemma 3. We know pick $Q_{p_1} \in \mathcal{G}_{S,T}(Q^n, p_1)$ randomly, and apply Lemma 2 to $\langle x_0, y_0 \rangle$. Thus we find that in $Q_{p_0} \cup Q_{p_1}$ there is an x–y path of length at most $d_{Q^n}(x, y) + 4$ with probability $1 - \exp\{-\Omega(n \log\log n)\}$.

Case 3. $d_{Q^n}(x, y) = |y| < n/5$

Pick $Q_{p_0} \in \mathcal{G}_{S,T}(Q^n, p_0)$. Let $Z = \{z \subset [n] \setminus y : |z| = \lceil n/5 \rceil\}$. Then $d_{Q^n}(x, z) = \lceil n/5 \rceil$, and $n/5 \leq d_{Q^n}(y, z) \leq n/2$. Moreover, the random variables $d_{Q_{p_0}}(z)$ $(z \in Z)$ are independent and are such that $d_{Q_{p_0}}(z) \geq 2n^{2/3}$ with probability $1 - o(1)$. Thus, with probability $1 - \exp\{-c^n\}$ for some $c > 1$, there is a vertex $z \in Q^n$ such that $n/5 \leq d_{Q^n}(x, z)$, $d_{Q^n}(y, z) \leq n/2$, and $d_{Q_{p_0}}(z) \geq 2n^{2/3}$. We now pick $Q_{p_i} \in \mathcal{G}_{S,T}(Q^n, p_i)$ $(i = 1,\ 2)$, and argue as in Case 2. $\square$

Lemma 4 has the following immediate corollary.

Corollary 5. *Suppose* $3/\log\log n \leq p = p(n) < 1$. *Let* $\mathcal{A}_0$ *be the event that all* x, $y \in Q_p \in \mathcal{G}(Q^n, p)$ *with* $d_{Q_p}(x)$, $d_{Q_p}(y) \geq 2n^{2/3}$ *are such that (6) in Lemma 4 holds. Then the probability that* $Q_p \in \mathcal{G}(Q^n, p)$ *satisfies* $\mathcal{A}_0$ *is* $1 - \exp\{-\Omega(n \log\log n)\}$.

Proof. For x, $y \in Q^n$, let $\mathcal{A}(x, y)$ be the event that (6) from Lemma 4 holds in Q_p for x and y. Then

$$\mathbb{P}(\mathcal{A}_0 \text{ fails}) \leq \mathbb{P}(\exists x,\ y \in Q^n : \mathcal{A}(x, y) \text{ fails},\ d_{Q_p}(x), d_{Q_p}(y) \geq 2n^{2/3})$$

$$\begin{aligned}&\le 2^{2n}\max_{x,y}\mathbb{P}\big(\mathcal{A}(x,y)\text{ fails, } d_{Q_p}(x), d_{Q_p}(y)\ge 2n^{2/3}\big)\\ &= 2^{2n}\max_{x,y}\mathbb{P}\big(\mathcal{A}(x,y)\text{ fails}\,\big|\, d_{Q_p}(x), d_{Q_p}(y)\ge 2n^{2/3}\big)\\ &\qquad\qquad\qquad\qquad\times\mathbb{P}\big(d_{Q_p}(x), d_{Q_p}(y)\ge 2n^{2/3}\big)\\ &=\exp\{-\Omega(n\log\log n)\}.\end{aligned}$$ □

We now consider Q^n-processes $\widetilde{Q}=(Q_t)_0^M$, and show that vertices of large degree in Q_t are a.s. connected by a short path as long as t is not too small.

Corollary 6. *Almost every random Q^n-process $\widetilde{Q}=(Q_t)_0^M$ is such that, for every $t=t(n)\ge 4M/\log\log n$ and every two vertices x, $y\in Q^n$ such that both of them have at least $3n^{2/3}$ neighbours in Q_t, we have*

$$\begin{aligned} d_{Q_t}(x,y) &= d_{Q^n}(x,y) &&\text{if}\quad d_{Q^n}(x,y)\ge n-n^{2/3},\\ d_{Q_t}(x,y) &\le d_{Q^n}(x,y)+4 &&\text{if}\quad d_{Q^n}(x,y)\ge n/5,\\ d_{Q_t}(x,y) &\le 0.8n &&\text{if}\quad d_{Q^n}(x,y)\le n/5.\end{aligned}\tag{7}$$

Proof. This result is a straightforward consequence of Corollary 5. Let $4M/\log\log n\le t=t(n)\le M$ be fixed and $p=p(n)=\big(1-((\log n)/M)^{1/2}\big)\,t/M$. Then $\mu=pM=\big(1-((\log n)/M)^{1/2}\big)\,t=(1+o(1))t$, and, from standard estimates for binomial random variables,

$$\mathbb{P}\Bigg(|e(Q_p)-\mu|\ge\frac{1}{2}\left(\frac{\log n}{M}\right)^{1/2}\mu\Bigg)\le\exp\left\{-\frac{1}{4}\cdot\frac{\log n}{\log\log n}\right\}=o(1),$$

where $e(Q_p)$ denotes the number of edges in Q_p.

In particular,

$$t-2(M\log n)^{1/2}\le e(Q_p)\le t\tag{8}$$

holds for every $4M/\log\log n\le t\le M$ with probability $1-\exp\{-(\log n)/4\log\log n\}$. Let $\mathcal{G}_c(Q^n,p)$ be the conditional probability space obtained from $\mathcal{G}(Q^n,p)$ conditioning on the event given in (8). Note that if $\mathcal{A}$ is any event concerning $Q_p\in\mathcal{G}(Q^n,p)$, then $\mathbb{P}_c(\mathcal{A})\le(1+o(1))\mathbb{P}(\mathcal{A})$, where $\mathbb{P}_c$ denotes the probability in $\mathcal{G}_c(Q^n,p)$. Let $\mathcal{A}_0$ be the event that (6) is satisfied for all x, $y\in Q_p$ with $d_{Q_p}(x)$, $d_{Q_p}(y)\ge 2n^{2/3}$. Then, by Corollary 5,

$$\mathbb{P}_c(\mathcal{A}_0)=1-(1+o(1))\exp\{-\Omega(n\log\log n)\}=1-\exp\{-\Omega(n\log\log n)\}.$$

Now, we may generate $Q_t \in \mathcal{G}(Q^n, t)$ by first picking $Q_p \in \mathcal{G}_{\mathrm{c}}(Q^n, p)$, and then randomly adding $t' = t - e(Q_p)$ new edges to Q_p. One may check that $Q_{t'} \in \mathcal{G}(Q^n, t')$ is such that its maximal degree $\Delta(Q_{t'})$ satisfies $\Delta(Q_{t'}) \leq n^{2/3}$ with probability $1 - \exp\{-\Omega(n^{4/3}/\log n)\}$. Thus Q_t satisfies the property that (7) of our lemma holds for all $x, y \in Q_t$ with $d_{Q_t}(x)$, $d_{Q_t}(y) \geq 3n^{2/3}$ with probability $1 - \exp\{-\Omega(n \log\log n)\}$. Thus such a property holds for all Q_t in $\widetilde{Q} = (Q_t)_0^M$ with $t \geq 4M/\log\log n$ with probability $1 - M\exp\{-\Omega(n \log\log n)\} = 1 - \exp\{-\Omega(n \log\log n)\}$. $\square$

4. The diameter of a random subgraph of the n-cube

In this section we shall study the behaviour of the diameter of the almost surely unique largest component $L = L(Q_t)$ of $Q_t \in \mathcal{G}(Q^n, t)$ when $t = \Omega(M)$. In fact, our main results will be 'global' ones, in the sense that they will concern random Q^n-processes $\widetilde{Q} = (Q_t)_0^M$, and they will describe the behaviour of $\mathrm{diam}(L)$ as $t = \Omega(M)$ grows.

It follows from Corollary 6 that to estimate the diameter of the largest component L of Q_t from above it is enough to show that each vertex in L of small degree is within a small distance from some vertex of large degree. Our argument below makes this statement precise. To formulate and prove our results, we need to introduce some further definitions. Lemmas 7, 8, and 9 that follow are technical results needed in the proof of one of the main results of this section, Theorem 10.

Let H be a spanning subgraph of Q^n, and let $P_1 = x_0x_1 \ldots x_{\ell_1}$, $P_2 = y_0y_1 \ldots y_{\ell_2}$ be two paths in Q^n such that $y_0 = x_0^{\mathrm{c}} = [n] \setminus x_0$ and $1 \leq k = \ell_1 + \ell_2$. We sometimes refer to the pair (P_1, P_2) with P_1, P_2 as above as a *k-pair*. We say that (P_1, P_2) is a *k-stretching pair* in H if (i) P_1 and P_2 are paths in H and, moreover, (ii) the only edges of H incident to any of $x_1, \ldots, x_{\ell_1}$ are the ones of P_1 and, similarly, the only edges of H incident to any of $y_1, \ldots, y_{\ell_2}$ are the ones of P_2.

The idea here is as follows. Suppose (P_1, P_2) above is a k-stretching pair. Clearly we have $d_H(x_0, y_0) \geq d_{Q^n}(x_0, y_0) = n$, and the two paths P_1, P_2 'stretch out' from the 'core' of H, giving two vertices in H that are apart. Indeed, we certainly have that $d_H(x_{\ell_1}, y_{\ell_2}) \geq n + \ell_1 + \ell_2 = n + k$. Thus the existence of k-stretching pairs is an obstruction for the property of having diameter less than $n + k$. Our main aim is to show that a.s. this is the only such obstruction for any fixed k.

We remark that the very basic idea that stretching pairs are related to the diameter and radius can also be found in Kostochka, Sapozhenko, and Weber [16], although in [16] the exact relationship as given in our Theorem 10 below is not established.

Before we proceed, we need a few more definitions. Let $P_1 = x_0x_1\dots x_{\ell_1}$, $P_2 = y_0y_1\dots y_{\ell_2}$ form a k-pair of paths of Q^n. If P_1 and P_2 satisfy (ii) above, we say that (P_1,P_2) is a *potentially k-stretching pair* in H. If (P_1,P_2) is potentially k-stretching but not k-stretching, that is (i) does not hold, then we say that it is *strictly potentially k-stretching*. Finally, if (P_1,P_2) above is k-stretching in H and furthermore $d_H(x_0)$, $d_H(y_0)\geq n/\log n$, then we say that (P_1,P_2) is a *proper* k-stretching pair.

For $k=1,2,\dots$, let $\mathcal{S}_k^{(\mathrm{s})}$ (respectively $\mathcal{S}_k^{(\mathrm{p})}$) denote the property that a spanning subgraph of Q^n has at least $t\geq M(1-2^{-1/(k+1)})$ edges and contains no k-stretching pair (respectively, no potentially k-stretching pair). Clearly $\mathcal{S}_k^{(\mathrm{p})}$ is an increasing property, whereas $\mathcal{S}_k^{(\mathrm{s})}$ is not. Our next result shows that $\mathcal{S}_k^{(\mathrm{s})}$ is however 'almost surely increasing', and furthermore a threshold function for $\mathcal{S}_k^{(\mathrm{p})}$ is also a threshold function for $\mathcal{S}_k^{(\mathrm{s})}$. For $k\geq 1$ and a Q^n-process $\widetilde{Q}=(Q_t)_0^M$, let $t_k^{(a)}=t_k^{(a)}(\widetilde{Q})$ be the hitting time of property $\mathcal{S}_k^{(a)}$, where $a=\mathrm{s},\mathrm{p}$. Thus

$$t_k^{(a)}=t_k^{(a)}(\widetilde{Q})=\min\{t\geq M(1-2^{-1/(k+1)}) : Q_t \text{ has } \mathcal{S}_k^{(a)}\}.$$

Lemma 7. *Let $k\geq 1$ be fixed and let $\omega=\omega(n)\to\infty$ be such that $\omega\leq\log n$ for all $n\geq 1$. Then almost every random Q^n-process $\widetilde{Q}=(Q_t)_0^M$ satisfies the following.*

(i) We have

$$\begin{aligned} M\left(1-\frac{1}{2^{1/k}}\left(1-\frac{\log n-\omega}{n}\right)\right)& \\ \leq t_k^{(\mathrm{s})}\leq t_k^{(\mathrm{p})}\leq M\left(1-\frac{1}{2^{1/k}}\left(1-\frac{\log n+\omega}{n}\right)\right)&. \qquad (9)\end{aligned}$$

(ii) For $a=\mathrm{s}$ and p, the graph Q_t has property $\mathcal{S}_k^{(a)}$ whenever $t\geq t_k^{(a)}$.

(iii) If $M(1-2^{-1/(k+1)})\leq t<t_k^{(\mathrm{s})}$ then there is a proper k-stretching pair in Q_t.

Proof. (i) Suppose $t=(1-2^{-1/k}(1-(\log n+C)/n))M$, where $C=C(n)=o(\sqrt{n})$, and set $p=p(n)=t/M$. Let $X_k^{(\mathrm{s})}=X_k^{(\mathrm{s})}(Q_t)$ and $X_k^{(\mathrm{p})}=X_k^{(\mathrm{p})}(Q_t)$ denote the number of ordered pairs P_1, P_2 of Q^n-paths such that (P_1,P_2) is k-stretching and, respectively, potentially k-stretching in Q_t. Let us estimate $\mathbb{E}(X_k^{(\mathrm{s})})$. To this end, let us first fix a

k-pair (P_1, P_2) of paths of Q^n. Note that for this pair to be a k-stretching pair in Q_t, we need that certain k edges of Q^n should be in Q_t, and that certain $k_1 = kn + O(1)$ edges of Q^n should not be in Q_t. (Here and in the sequel, the constants implicit in the O-, Ω-, and Θ-notation are allowed to depend on k.) Thus, the probability that (P_1, P_2) should be a k-stretching pair in Q_t is

$$\begin{aligned} \binom{M-k-k_1}{t-k} \Big/ \binom{M}{t} &= \frac{(M-t)_k}{(M)_{k+k_1}} \cdot \frac{t!}{(t-k)!} \\ &= (1+o(1)) \left(1 - \frac{t}{M}\right)^{k_1} \left(\frac{t}{M}\right)^k = (1-p)^{k_1} p^k = \Theta\left((1-p)^{kn}\right). \end{aligned}$$

The number of k-pairs of paths in Q^n is clearly $\Theta(2^n n^k)$, and so we get

$$\mathbb{E}(X_k^{(\mathrm{s})}) = \Theta\left(2^n n^k (1-p)^{kn}\right) = \Theta\left(\mathrm{e}^{-kC}\right), \tag{10}$$

where for the last equality we use that $C = o(\sqrt{n})$. Similar calculations give that $\mathbb{E}(X_k^{(\mathrm{p})}) = \Theta\left(\mathrm{e}^{-kC}\right)$. Now the upper bound in (9) follows from (10) and Markov's inequality.

To see the lower bound for $t_k^{(\mathrm{s})}$, let us suppose that $C = C(n) \to -\infty$ and as above $C = o(\sqrt{n})$. Let us show that the number $X = X(Q_t)$ of k-stretching pairs (P_1, P_2) in $Q_t \in \mathcal{G}(Q^n, t)$ with P_1 a path of length k, and P_2 a trivial path, is concentrated around its expectation. Let $P = x_0 x_1 \ldots x_k$ be a k-path in Q^n. We write $E'(P)$ for the edges of Q^n that have at least one endpoint in $\{x_1, \ldots, x_k\}$. Also, we write $X_P = X_P(Q_t)$ for the indicator function for the event that (P_1, x_0^{c}) should be a k-stretching pair in Q_t. Then $X = \sum_P X_P$, where the sum extends over all paths $P \subset Q^n$ of length k. Note that

$$\mathbb{E}_2(X) = \mathbb{E}(X(X-1)) \leq (\mathbb{E}(X))^2 + S,$$

with $S = \sum_{P,Q} \mathbb{P}(X_P X_Q = 1)$, where the sum is over all pairs (P, Q) of distinct k-paths in Q^n such that $E'(P) \cap E'(Q) \neq \emptyset$. It can be easily checked that if $X_P X_Q = 1$ then at least $k+1$ vertices in $V(P) \cup V(Q)$ have degree at most 2. Thus

$$\mathbb{P}(X_P X_Q = 1) = O\left(n^{2(k+1)} (1-p)^{n(k+1)}\right) = O\left(2^{-n(1+1/k)} n^{k+1} \mathrm{e}^{(k+1)C}\right),$$

and hence

$$S = O\left(2^n n^{2k+1}\right) \max \mathbb{P}(X_P X_Q = 1) \to 0, \tag{11}$$

where the maximum is naturally taken over all pairs (P, Q) of distinct k-paths in Q^n such that $E'(P) \cap E'(Q) \neq \emptyset$.

Thus $\mathrm{Var}(X) = o(\mathbb{E}(X)^2)$ since $C = C(n) \to -\infty$, and so by Chebyshev's inequality we have that $X = (1 + o(1))\mathbb{E}(X)$ almost surely. Therefore a.e. $Q_t \in \mathcal{G}(Q^n, t)$ is such that $X = \Theta(\mathrm{e}^{-kC})$ when $C = C(n) \to -\infty$ and $C = C(n) = o(\sqrt{n})$. Moreover, note that (11) above implies that a.s. any two k-stretching pairs (P, x_0^{c}) and (Q, y_0^{c}) in Q_t, where $P = x_0 x_1 \ldots x_k$ and $Q = y_0 y_1 \ldots y_k$, are such that $E'(P) \cap E'(Q) = \emptyset$. To prove the lower bound for $t_k^{(\mathrm{s})}$, let us consider a Q^n-process $\widetilde{Q} = (Q_t)_0^M$ as a process where random edges of Q^n are successively deleted from the evolving graph Q_t. Let again $t = (1 - 2^{-1/k}(1 - (\log n + C)/n))M$, where $C = C(n) \to -\infty$ and $C = C(n) = o(\sqrt{n})$, and condition on $X = X(Q_t) = \Theta(\mathrm{e}^{-kC}) \to \infty$. We claim that a.s. at least one k-stretching pair (P, x_0^{c}) present in Q_t will still be present in $Q_{t'}$, where $t' = \lceil M(1 - 2^{-1/(k+1)}) \rceil$. Note that $Q_{t'}$ is obtained from Q_t by the random deletion of a certain number of edges. Note also that the probability that no edge in $E(P)$ is deleted in this process, where (P, x_0^{c}) is a fixed stretching pair in Q_t, is bounded away from 0. Since $X \to \infty$, our claim follows, and the lower bound for $t_k^{(\mathrm{s})}$ is proved.

(*ii*) Let us now consider property $\mathcal{S}_k^{(a)}$ for Q_t ($t \geq t_k^{(a)}$). Since $\mathcal{S}_k^{(\mathrm{p})}$ is increasing, trivially Q_t contains no potentially k-stretching pairs for $t \geq t_k^{(\mathrm{p})}$. As remarked earlier, property $\mathcal{S}_k^{(\mathrm{s})}$ is not increasing and hence we need to do a little work to deduce that $\mathcal{S}_k^{(\mathrm{s})}$ a.s. holds for $t \geq t_k^{(\mathrm{s})}$. Set

$$t_0 = M\left(1 - \frac{1}{2^{1/k}}\left(1 - \frac{\log n - \log\log n}{n}\right)\right),$$

and note that $\mathbb{E}(X_k^{(\mathrm{p})}) = \mathbb{E}(X_k^{(\mathrm{p})}(Q_{t_0})) = \Theta\left((\log n)^k\right)$. Also, as shown above, almost surely $t_k^{(\mathrm{s})} \geq t_0$, and hence we may condition on this event.

Since by Markov's inequality almost every $\widetilde{Q} = (Q_t)_0^M$ is such that $X_k^{(\mathrm{p})}(Q_{t_0}) \leq (\log\log n)(\log n)^k$, we may condition on $\widetilde{Q}$ satisfying this property. We claim that almost surely no strictly potentially k-stretching pair (P_1, P_2) of Q^n-paths in Q_{t_0} is k-stretching in Q_t for $t \geq t_0$. Clearly, this proves that a.s. Q_t contains no k-stretching pair for $t \geq t_k^{(\mathrm{s})}$, as required. To check the claim, fix a strictly potentially k-stretching pair (P_1, P_2) in Q_{t_0}. Suppose, as usual, $P_1 = x_0 x_1 \ldots x_{\ell_1}$ and $P_2 = y_0 y_1 \ldots y_{\ell_2}$, where $x_0 = y_0^{\mathrm{c}} = [n] \setminus y_0$. Note that (P_1, P_2) will become k-stretching in some Q_t ($t > t_0$) only if one of the edges of Q^n in $E(P_1) \cup E(P_2)$ not present in Q_{t_0} is added to our evolving graph before any other edge incident to $x_1, \ldots, x_{\ell_1}$, $y_1, \ldots, y_{\ell_2}$. Now, clearly $|E(P_1) \cup E(P_2)| = k$, and the total number of edges of Q^n incident to $x_1, \ldots, x_{\ell_1}$, $y_1, \ldots, y_{\ell_2}$ is $kn + O(1)$. Thus the probability that an edge of $E((P_1) \cup E(P_2)) \setminus E(Q_{t_0})$ should be added to our evolving graph before any other edge incident to the x_i and y_j is is $O(1/n)$. Hence the probability that a strictly potentially k-stretching pair in Q_{t_0} will ever turn into a k-stretching pair is at most $O\big((\log\log n)(\log n)^k/n\big)$. This shows that Q_t in $\widetilde{Q} = (Q_t)_0^M$ has no k-stretching pair if $t \geq t_k^{(\mathrm{s})}$ almost surely, as required.

(*iii*) This follows from second moment calculations analogous to the ones in (*i*). We omit the details. □

The result above tells us that the critical values of t for the existence of stretching and potentially stretching pairs are $t_k = \lfloor M(1 - 2^{-1/(k+1)}(1 - (\log n)/n))\rfloor$. We now look more closely at Q_t when t is near such values.

Lemma 8. *Let $k \geq 1$ be fixed. Suppose $t = M(1 - 2^{-1/k}(1 - (\log n + C)/n))$, where $C = C(n) \to c$ as $n \to \infty$. Let $\lambda = 4 \cdot 2^{-1/k}(1 - 2^{-1/k})^k(1 + (k-1)2^{-1-1/k})\mathrm{e}^{-ck}$. Then*

$$\lim_{n\to\infty} \mathbb{P}(t_k^{(\mathrm{s})} \leq t) = \exp\{-\lambda\}. \tag{12}$$

Proof. Let $X_k^{(\mathrm{s})} = X_k^{(\mathrm{s})}(Q_t)$ count the number of k-stretching pairs in Q_t. Then calculations similar to the ones in the proof of Lemma 7 concerning the random variable $X = X(Q_t)$ show that, for any fixed $r \geq 1$, we have $\lim_{n\to\infty} \mathbb{E}_r(X_k^{(\mathrm{s})}) = \lambda^r$. Thus $X_k^{(\mathrm{s})}$ converges in distribution to a Poisson random variable with mean λ. (See, for instance, Theorem 20 in Chapter I of [2].) Thus (12) follows.

We now give a brief sketch of the calculations. Let $p = t/M$. Below we consider k-pairs (P,Q), where as usual $P = x_0 \dots x_{\ell_1}$, $Q = y_0 \dots y_{\ell_2}$. Let us write $X_{(P,Q)}$ for the 0–1 indicator r.v. of the event that (P,Q) should be a k-stretching pair in Q_t. To avoid double counting of the pair $\{P,Q\}$, in the k-pairs below we assume that, say, $1 \in x_0$. It is easy to see that the total number of such k-pairs is $(1+o(1))2^{n-1}(k+1)n^k$. Among these, $(1+o(1))2^n n^k$ pairs (the ones with ℓ_1 or $\ell_2 = 0$) are such that $\mathbb{E}(X_{(P,Q)}) = (1 + o(1))p^k(1-p)^{nk-2k+1}$, and for the remaining $(1+o(1))2^n(k-1)n^k$ we have $\mathbb{E}(X_{(P,Q)}) = (1+o(1))p^k(1-p)^{nk-2k+2}$. Hence $\mathbb{E}(X_k^{(\mathrm{s})}) = \sum_{(P,Q)} \mathbb{E}(X_{(P,Q)}) = \lambda + o(1)$ as $n \to \infty$.

Now we fix $r \geq 2$. To estimate the rth factorial moment $\mathbb{E}_r(X_k^{(\mathrm{s})})$, we consider r-tuples $U = \langle (P_1,Q_1), \dots, (P_r,Q_r)\rangle$ of k-pairs. In the sequel, we write $\sum^*$ to denote sum over all r-tuples U as above with all the r entries distinct, and $\sum'$ for sum over all such r-tuples with at least one pair of repeated entries. Also, if $P_i = x_0^{(i)} \dots x_{\ell_1^{(i)}}^{(i)}$ and $Q_i =$

$y_0^{(i)} \ldots y_{\ell_2^{(i)}}^{(i)}$ form a k-pair (P_i, Q_i) $(i \in \{1,2\})$, then we say that (P_1, Q_1) and (P_2, Q_2) are *distant* if all the distances between a vertex in $V(P_1) \cup V(Q_1)$ and a vertex in $V(P_2) \cup V(Q_2)$ are at least $\log\log n$. Below we write $\sum_1$ to denote sum over r-tuples U with the (P_i, Q_i) $(1 \le i \le r)$ pairwise distant, and $\sum_2$ to denote sum over r-tuples U with r distinct entries and with at least a pair of entries that are not distant. Then

$$\begin{aligned}
\mathbb{E}_r(X_k^{(\mathrm{s})}) &= \sum\nolimits^* \mathbb{E}(X_{(P_1,Q_1)} \cdots X_{(P_r,Q_r)}) \\
&= \sum\nolimits_1 \mathbb{E}(X_{(P_1,Q_1)} \cdots X_{(P_r,Q_r)}) + \sum\nolimits_2 \mathbb{E}(X_{(P_1,Q_1)} \cdots X_{(P_r,Q_r)}) \\
&= \sum\nolimits_1 \mathbb{E}(X_{(P_1,Q_1)}) \cdots \mathbb{E}(X_{(P_r,Q_r)}) + \sum\nolimits_2 \mathbb{E}(X_{(P_1,Q_1)} \cdots X_{(P_r,Q_r)}).
\end{aligned}$$

Moreover, clearly $\{\mathbb{E}(X_k^{(\mathrm{s})})\}^r = \sum_U \mathbb{E}(X_{(P_1,Q_1)}) \cdots \mathbb{E}(X_{(P_r,Q_r)})$. To show that $\mathbb{E}_r(X_k^{(\mathrm{s})}) = \lambda^r + o(1)$, it suffices to show that

(i) $\sum' \mathbb{E}(X_{(P_1,Q_1)}) \cdots \mathbb{E}(X_{(P_r,Q_r)}) = o(1)$,

(ii) $\sum_2 \mathbb{E}(X_{(P_1,Q_1)}) \cdots \mathbb{E}(X_{(P_r,Q_r)}) = o(1)$,

(iii) $\sum_2 \mathbb{E}(X_{(P_1,Q_1)} \cdots X_{(P_r,Q_r)}) = o(1)$.

Now note that *(i)* and *(ii)* follow easily from the fact that $\sum'$ and $\sum_2$ are sums over a 'small' number of k-pairs U. To finish the proof, note that *(iii)* may be checked by considering the number of vertices of degree at most two in $\bigcup_i V(P_i) \cup \bigcup_j V(Q_j)$, as in the proof of (11). Indeed, for each r-tuple U, we consider a maximal subsequence of k-pairs $\big\langle (P_{i_1}, Q_{i_1}), \ldots, (P_{i_{r'}}, Q_{i_{r'}}) \big\rangle$ of pairwise distant entries $(1 \le i_1 < \cdots < i_{r'} \le r)$, and note that if $X_{(P_1,Q_1)} \cdots X_{(P_r,Q_r)} = 1$ then there are at least $r'k+1$ vertices of degree at most two in $\bigcup_i V(P_i) \cup \bigcup_j V(Q_j)$. However, the number of such U is at most $2^{nr'} n^{O(\log\log n)}$, and hence *(iii)* follows (cf. (11)). $\square$

Before we can prove the main result of this section, we need to introduce one final piece of terminology. Suppose T_1, $T_2 \subset Q^n$ are two subtrees of Q^n, and x_0, $y_0 \in Q^n$ are such that $d_{Q^n}(T_1, x_0) = d_{Q^n}(T_2, y_0) = 1$. Thus $x_0 \notin T_1$ but x_0 is adjacent in Q^n to some vertex of T_1, and similarly for y_0 and T_2. Suppose further that $|T_1| + |T_2| \ge k+1$, and $d_{Q^n}(x_0, y_0) \ge n - k$ for some integer $k \ge 1$. Let us say in this case that $(T_1, T_2; x_0, y_0)$ is a $(k+1)$-*system* in Q^n. Let $H \subset Q^n$ be a spanning subgraph of Q^n. Then we say that the $(k+1)$-system $(T_1, T_2; x_0, y_0)$ is a *weakly* $(k+1)$-*stretching system* in H if all vertices $x \in V(T_1) \cup V(T_2)$ have degree $d_H(x) < 3n^{2/3}$ in H. We then have the following simple lemma.

Lemma 9. *Let $k \geq 1$ be fixed. Suppose $p = 1 - 2^{-1/k}$ and $t = \lfloor pM \rfloor$. Then a.e. $Q_p \in \mathcal{G}(Q^n, p)$ and a.e. $Q_t \in \mathcal{G}(Q^n, t)$ contains no weakly $(k+1)$-stretching system.*

Proof. Fix a $(k+1)$-system $(T_1, T_2; x_0, y_0)$ in Q^n. The probability that this is a weakly $(k+1)$-stretching system in Q_p is clearly at most

$$P_0 = \left\{ \mathbb{P}(S_{n-k,p} < 3n^{2/3}) \right\}^{k+1},$$

where $S_{n-k,p}$ is a binomial random variable with parameters $n-k$ and p. Indeed, we need at least $k+1$ vertices of $V(T_1) \cup V(T_2)$ to have degree less than $3n^{2/3}$, and for any vertex z in $V(T_1) \cup V(T_2)$ there are at least $n-k$ edges of Q^n incident to z that are not incident to any other vertex of $V(T_1) \cup V(T_2)$. Now, if $\ell = \lfloor 3n^{2/3} \rfloor$, then

$$\begin{aligned} \mathbb{P}(S_{n-k,p} < 3n^{2/3}) &\leq 2\binom{n-k}{\ell}(1-p)^{n-k-\ell} \\ &\leq 4(\mathrm{e}n^{1/3})^{3n^{2/3}} 2^{-(n-3n^{2/3})/k} \leq n^{2n^{2/3}} 2^{-n/k}. \end{aligned}$$

Thus, since the number of $(k+1)$-systems in Q^n is at most $O(2^n n^{2(k+1)})$, we have that $Q_p \in \mathcal{G}(Q^n, p)$ contains a weakly $(k+1)$-stretching system with probability

$$O(2^n n^{2(k+1)})P_0 = O\left(2^n n^{2(k+1)(1+n^{2/3})} 2^{-n(1+1/k)}\right) = o(1),$$

as required. The statement for $Q_t \in \mathcal{G}(Q^n, t)$ follows by the monotonicity of the property in question. $\square$

Let us say that a spanning subgraph $H \subset Q^n$ of Q^n satisfies property $\mathcal{D}_k$ if H has $e(H) \geq M(1 - 2^{-1/(k+1)})$ edges, it has a unique component $L = L(H)$ of largest order, and moreover L has diameter $\mathrm{diam}(L) \leq n + k - 1$. For a random Q^n-process $\widetilde{Q} = (Q_t)_0^M$, let

$$t_k^{(\mathrm{d})} = t_k^{(\mathrm{d})}(\widetilde{Q}) = \min\{t \geq M(1 - 2^{-1/(k+1)}) : Q_t \text{ has } \mathcal{D}_k\}.$$

We can now state and prove the main result of this section.

Theorem 10. *Let $k \geq 1$ be fixed. For a.e. random Q^n-process $\widetilde{Q} = (Q_t)_0^M$ we have $t_k^{(\mathrm{d})} = t_k^{(\mathrm{s})}$, and Q_t has property $\mathcal{D}_k$ for all $t \geq t_k^{(\mathrm{d})}$.*

Proof. First recall that a.e. Q^n-process $\widetilde{Q} = (Q_t)_0^M$ is such that Q_t has a unique largest component, the 'giant' component, if, say, $t \geq 2M/n$ (cf. [1, 5]). Moreover, it is easy to check that, quite crudely, almost every $\widetilde{Q}$ is such that if $t \geq M/\log\log\log n$ then Q_t has, apart from the largest component L, only components of order $O(\log\log n)$. Let us condition on our Q^n-process having the properties above.

Now, if $M(1-2^{-1/(k+1)}) \leq t = t(n) < t_k^{(\mathrm{s})}$, and $P_1 = x_0x_1\ldots x_{\ell_1}$, $P_2 = y_0y_1\ldots y_{\ell_2}$ form a proper k-stretching pair, then both x_0 and y_0 belong to the giant component of Q_t, and moreover $d_{Q_t}(x_{\ell_1}, y_{\ell_2}) \geq n+k$. Thus $t_k^{(\mathrm{d})} \geq t_k^{(\mathrm{s})}$ almost surely.

Let us now turn to the reverse inequality. Suppose $t \geq t_k^{(\mathrm{s})} = t_k^{(\mathrm{s})}(\widetilde{Q})$. We may assume that $t_k^{(\mathrm{s})} \geq (1-2^{-1/k})M$, and hence that Q_t does not contain weakly $(k+1)$-stretching systems either. According to Corollary 6, we may and shall assume that (7) of that lemma holds for all x, $y \in Q_t$ with $d_{Q_t}(x)$, $d_{Q_t}(y) \geq 3n^{2/3}$. Let x, $y \in Q_t$ belong to the giant component of Q_t. We claim that $d_{Q_t}(x,y) \leq n+k-1$.

Case 1. $d_{Q_t}(x)$, $d_{Q_t}(y) \geq 3n^{2/3}$

From the assumption that (7) holds for Q_t, in this case we have $d_{Q_t}(x,y) \leq n$.

For the next two cases, let us consider the set W of vertices of Q_t that have degree less than $3n^{2/3}$, and let $H = Q_t[W]$ be the graph induced by W in Q_t.

Case 2. $d_{Q_t}(x) < 3n^{2/3}$, and $d_{Q_t}(y) \geq 3n^{2/3}$

Let $T \subset H = Q_t[W]$ be a spanning tree of the component C_x of x in H. There is a vertex $x_0 \in L$ such that $d_{Q_t}(x_0,T) = 1$, and consequently such that $d_{Q_t}(x_0) \geq 3n^{2/3}$. Thus $d_{Q_t}(x,y) \leq d_{Q_t}(x,x_0) + d_{Q_t}(x_0,y) \leq |T| + n$. So we may assume that $|T| \geq k$, since otherwise we are done. However, as Q_t does not contain weakly $(k+1)$-stretching systems, we must have $|T| = k$, and moreover $T = C_x$ and T must be a path with x as one endpoint and the other endpoint must be adjacent to x_0 in Q_t. We conclude that Q_t contains an induced path $P = x_0x_1\ldots x_k$, where $x_k = x$. We may also assume that $d_{Q_t}(x_0,y) = n$, and hence that $d_{Q^n}(x_0,y) = n$. Since there are no k-stretching pairs in Q_t, there is a vertex $z \in Q^n$, $z \neq x_0,\ldots,x_k$, adjacent in Q_t to some x_i $(1 \leq i \leq k)$. But then $d_{Q_t}(z) \geq 3n^{2/3}$, and $d_{Q_t}(z,y) \leq n-1$. Thus $d_{Q_t}(x,y) \leq n+k-1$, as required.

Case 3. $d_{Q_t}(x)$, $d_{Q_t}(y) < 3n^{2/3}$

Let C_1 and respectively C_2 be the components of x and y in $H = Q_t[W]$. Let $T_i \subset C_i$ be a spanning tree of C_i $(i = 1,\ 2)$. There are vertices x_0, $y_0 \in Q^n$ such that $d_{Q_t}(x_0)$, $d_{Q_t}(y_0) \geq 3n^{2/3}$, and $d_{Q_t}(x_0,T_1) = d_{Q_t}(y_0,T_2) = 1$. Note that $|T_1|$, $|T_2| \leq k$, and hence if $d_{Q^n}(x_0,y_0) \leq n-k-1$ we are done. So we assume that $d_{Q^n}(x_0,y_0) \geq n-k$.

Since $(T_1, T_2; x_0, y_0)$ cannot be a weakly $(k+1)$-stretching system in Q_t, we have $|T_1| + |T_2| \leq k$, and as $d_{Q_t}(x, y) \leq |T_1| + |T_2| + d_{Q_t}(x_0, y_0)$, we assume that $|T_1| + |T_2| = k$. In fact, we may assume that $T_1 = C_1$, and T_1 together with x_0 forms an x_0–x path P_1, and similarly $T_2 = C_2$, and T_2 together with y forms a y_0–y path P_2. Also, we may assume $d_{Q_t}(x_0, y_0) = d_{Q^n}(x_0, y_0) = n$. Since (P_1, P_2) is not a k-stretching pair in Q_t, there is a vertex $z \in Q^n$, $z \notin P_i$, adjacent to some vertex in T_i in Q_t, for $i = 1$, say. However $d_{Q_t}(z) \geq 3n^{2/3}$, and $d_{Q^n}(z, y_0) < n$, and so $d_{Q_t}(x, y) \leq d_{Q_t}(x, z) + d_{Q_t}(z, y) \leq n + k - 1$, as required.

This finishes the proof of our result. □

We can now explicitly describe the behaviour of the diameter of the largest component of Q_t as $t = \Omega(M)$ increases.

Corollary 11. *Let $k \geq 1$ be fixed. (i) If $\omega = \omega(n) \to \infty$ as $n \to \infty$, then a.e. random Q^n-process $\widetilde{Q} = (Q_t)_0^M$ is such that if*

$$M\left(1 - \frac{1}{2^{1/k}}\left(1 - \frac{\log n + \omega}{n}\right)\right) \leq t \leq M\left(1 - \frac{1}{2^{1/(k-1)}}\left(1 - \frac{\log n - \omega}{n}\right)\right)$$

then there is a unique largest component $L = L(Q_t)$ in Q_t and $\operatorname{diam}(L) = n + k - 1$.

(ii) Suppose $C = C(n) \to c$ as $n \to \infty$, and $t = t(n) = M(1-2^{-1/k}(1-(\log n + C)/n))$. As in Lemma 8, let $\lambda = 4 \cdot 2^{-1/k}(1 - 2^{-1/k})^k(1 + (k-1)2^{-1-1/k})\mathrm{e}^{-ck}$. Then a.s. $Q_t \in \mathcal{G}(Q^n, t)$ has a unique largest component $L = L(Q_t)$ and, conditional on this event,

$$\lim_{n\to\infty} \mathbb{P}(\operatorname{diam}(L) = n + k - 1) = \mathrm{e}^{-\lambda}$$

and

$$\lim_{n\to\infty} \mathbb{P}(\operatorname{diam}(L) = n + k) = 1 - \mathrm{e}^{-\lambda}.$$

Proof. Statement (i) follows from Lemma 7 and Theorem 10. Statement (ii) follows from Lemma 8 and Theorem 10. □

An immediate consequence of Corollary 11 is the following. Let $\widetilde{Q} = (Q_t)_0^M$ be a Q^n-process, and let $k \geq 1$ be an integer. Write $t^{(k)} = t^{(k)}(\widetilde{Q})$ for the hitting time of the property $\{\delta \geq k\}$, that is

$$t^{(k)} = t^{(k)}(\widetilde{Q}) = \tau(\widetilde{Q}; \delta \geq k) = \min\{t : \delta(Q_t) \geq k\}.$$

Corollary 12. *In almost every random Q^n-process $\widetilde{Q} = (Q_t)_0^M$, we have*
(i) $t^{(1)} = \tau(\widetilde{Q}; \delta \geq 1) = \tau(\widetilde{Q}; \mathrm{diam} \leq n+1)$,
(ii) $t^{(2)} = \tau(\widetilde{Q}; \delta \geq 2) = \tau(\widetilde{Q}; \mathrm{diam} \leq n)$.

Proof. Let us first recall that, as proved in [3], the graph Q_t in $\widetilde{Q} = (Q_t)_0^M$ almost always becomes connected at time $t = t^{(1)}$. Thus statements concerning the largest component of Q_t from that time onwards are in fact statements about Q_t itself.

Now, almost surely $t^{(1)} \geq (1-\omega/n)M/2$ for any $\omega = \omega(n) \to \infty$, and hence $t_2^{(\mathrm{s})} \leq t^{(1)}$ almost always. Therefore, by Theorem 10, almost surely (i) holds.

To see the statement concerning $\tau(\widetilde{Q}; \mathrm{diam} \leq n)$, note first that a spanning subgraph $H \subset Q^n$ of Q^n contains a 1-stretching pair if and only if it has a vertex of degree 1. Thus almost surely $t_1^{(\mathrm{s})}(\widetilde{Q}) = t^{(2)}$: if $M/\log\log\log n \leq t < t^{(2)}$ then almost surely there is a vertex of degree 1 in Q_t, and hence $t_1^{(\mathrm{s})} \geq t^{(2)}$, and moreover if $t = t^{(2)}$ then there cannot be a 1-stretching pair in Q_t, and hence $t_1^{(\mathrm{s})} \leq t^{(2)}$. Thus we have $t^{(1)} \leq t^{(2)} = t_1^{(\mathrm{s})}$, and hence again by Theorem 10 almost surely (ii) holds. $\square$

5. The radius of a random subgraph of the n-cube

In this section we turn our attention to the radius $\mathrm{rad}(L)$ of the giant component of Q_t $(t = \Omega(M))$ in a typical random Q^n-process $\widetilde{Q} = (Q_t)_0^M$. As mentioned in the introduction, unlike in ordinary random graph processes, the radius of the evolving graph in a typical Q^n-process is rather stable, as the following result shows.

Theorem 13. *Let $0 < \varepsilon < 1/2$ be a constant. Then almost every random Q^n-process $\widetilde{Q} = (Q_t)_0^M$ is such that, for every $t = t(n)$ with $t_0 = \lfloor \varepsilon M \rfloor \leq t < t^{(1)}$, there is a unique largest component $L = L(Q_t)$ in Q_t and $\mathrm{rad}(L) = n-1$.*

Proof. We claim that almost surely, for each vertex x of the giant component L of Q_{t_0}, there exists a vertex y that also belongs L, and such that $d_{Q^n}(x,y) = n-1$. Indeed, notice first that if y has degree at least $\log n$ in Q_{t_0}, then y belongs to the giant component of Q_{t_0}, as the second largest component of Q_{t_0} has order $O(1)$. To see the claim, it now suffices to note that, almost surely, for any $z \in Q^n$, there is a vertex $y \in \Gamma_{Q^n}(z)$ with degree at least $\log n$ in Q_{t_0}. Thus almost surely, for every $t \geq t_0$, the radius $\mathrm{rad}(L_t)$ of the largest component L_t of Q_t is at least $n-1$.

Let us now show that almost surely $\mathrm{rad}(L_t) \leq n-1$ for all $t \geq t_0$. Almost every random Q^n-process $\widetilde{Q} = (Q_t)_0^M$ is such that if $t = t^{(1)} - 1$, then Q_t has a unique isolated vertex $z_0 = z_0(\widetilde{Q})$. In what follows, we condition on $\widetilde{Q}$ having this property. Let $z \in Q^n$ be fixed, and let us further condition on the event that $\{z_0(\widetilde{Q}) = z\}$. It suffices to show that, under these conditions, almost surely $\mathrm{rad}(L_t) \leq n-1$ for $t \geq t_0$, and hence let us now verify this assertion.

Let $\mathcal{G}_z(Q^n, t)$ and respectively $\mathcal{G}_z(Q^n, p)$ be the spaces obtained from $\mathcal{G}(Q^n, t)$ and $\mathcal{G}(Q^n, p)$ by conditioning on the event that z should be an isolated vertex. We now note that Q_{t_0} in $\widetilde{Q} = (Q_t)_0^M$ is simply a random element from $\mathcal{G}_z(Q^n, t_0)$, and therefore we turn to this latter space. However, let us first set $p = p(n) = \big(1 - ((\log n)/M)^{1/2}\big)\,\varepsilon$, and consider $\mathcal{G}_z(Q^n, p)$. Let $\mathcal{A}_0$ be as in Corollary 5. Then

$$\begin{aligned} \mathbb{P}(Q_p \in \mathcal{G}_z(Q^n, p) \text{ fails } \mathcal{A}_0) &\leq \mathbb{P}(Q_p \in \mathcal{G}(Q^n, p) \text{ fails } \mathcal{A}_0)/\mathbb{P}(d_{Q_p}(z) = 0) \\ &= \exp\{-\Omega(n \log\log n)\}(1-p)^{-n} = \exp\{-\Omega(n \log\log n)\}, \end{aligned}$$

and hence a.e. $Q_p \in \mathcal{G}_z(Q^n, p)$ satisfies $\mathcal{A}_0$. It is also easily seen that the vertex $x = z^{\mathrm{c}} = [n] \setminus z$ and all vertices $y \neq z$ of Q^n within Q^n-distance $n/\log n$ from z have degree at least $3n^{2/3}$ in $Q_p \in \mathcal{G}_z(Q^n, p)$.

As in the proof of Corollary 6, we may generate an element from $\mathcal{G}_z(Q^n, t_0)$ by first generating an element Q_p from $\mathcal{G}_z(Q^n, p)$, and then adding a suitable number of random edges to Q_p. Proceeding this way, and using the above fact about property $\mathcal{A}_0$, we may show that a.e. $Q_{t_0} \in \mathcal{G}_z(Q^n, t_0)$ is such that (i) any two vertices x, y of degree at least $3n^{2/3}$ in Q_{t_0} are such that (7) of Corollary 6 holds, and moreover (ii) the vertex $x = z^{\mathrm{c}}$ and all vertices $y \neq z$ within Q^n-distance $n/\log n$ from z have degree at least $3n^{2/3}$. Let $k = \lceil 1/\log_2(1/(1-\varepsilon)) \rceil$. From Lemma 9 we know that a.e. $Q_{t_0} \in \mathcal{G}_z(Q^n, t_0)$ is such that (iii) Q_{t_0} contains no weakly $(k+1)$-stretching systems.

It is now enough to notice that if (i), (ii), and (iii) above hold for Q_{t_0}, then for any graph H with $Q_{t_0} \subset H \subset Q^n$ in which $d_H(z) = 0$ we have that $d_H(x, y) \leq n-1$ for $x = z^{\mathrm{c}}$ and any $y \in H$ distinct from z, as long as x and y belong to the same component in H.

□

One can read out the following corollary from the above result.

Corollary 14. *Almost every random Q^n-process $\widetilde{Q} = (Q_t)_0^M$ is such that*

$$t^{(1)} = \tau(\widetilde{Q}; \delta \geq 1) = \tau(\widetilde{Q}; \mathrm{rad} \leq n). \qquad \square$$

6. Concluding remarks

Let $\widetilde{Q} = (Q_t)_0^M$ be a random Q^n-process. It is known that if t is a little larger than $t_0 = M/n$ almost surely there is a unique largest component in Q_t and moreover this component is of much larger order than the other components (see [1] and [5]). This component is usually referred to as the 'giant' component. One very interesting problem we have not dealt with in this note is that of the determination of the diameter of the giant component of Q_t for t a little larger than t_0. This problem might be harder than the problems we have addressed here: it is likely that finer methods than the ones used in the proof of Lemma 2 might need to be developed. Quite possibly, these methods would be based on the Kruskal–Katona theorem, or more generally on isoperimetric inequalities for the cube, or the martingale method. In particular, it would be interesting to settle the following specific question.

Problem 15. *In a typical random Q^n-process $\widetilde{Q} = (Q_t)_0^M$, is ever the diameter of a component of Q_t superpolynomial?* □

This question concerns Q_t for t close to t_0; there are numerous interesting problems about Q_t for larger t, beyond the hitting time of connectedness. In a sequel [7] we shall study Q_t in this range. In particular, we shall prove a result implying that for almost every cube process the hitting time of k-connectedness is equal to the hitting time of minimal degree at least k. We shall also study the 'mixed model' for random subgraphs of Q^n. Here a random subgraph $Q_{p_{\mathrm{v}},p_{\mathrm{e}}}$ of the cube is chosen by independently deleting vertices and edges from Q^n with probabilities $1-p_{\mathrm{v}}$ and $1-p_{\mathrm{e}}$ respectively. We shall prove some results that are analogous to the ones in this paper, and thereby we shall answer some questions of Kostochka, Sapozhenko and Weber [15]. (See also [16].)

Acknowledgements. This work was started at the LMS Symposium on Probabilistic Methods in Combinatorics, Durham, and it was in part done at ZiF, University of Bielefeld. The first author was partially supported by the NSF under grant DMS-8806097, the second was partially supported by BID/USP, Project 30.01, and the third was partially supported by KBN under grant 2 1087 91 01. The authors would like to express their thanks to Boris Pittel for carefully reading the manuscript and making a number of detailed and insightful comments. They are also indebted to an anonymous referee who drew their attention to [17, 18, 19] and to parts of [16], which appeared after this paper was submitted.

References

[1] Ajtai, M., Komlós, J., Szemerédi, E., Largest random component of a k-cube, *Combinatorica* **2** (1982), 1–7.

[2] Bollobás, B., *Random Graphs*, Academic Press, London 1985, $xvi + 447$pp.

[3] Bollobás, B., Complete matchings in random subgraphs of the cube, *Random Structures and Algorithms* **1** (1990), 95–104.

[4] Bollobás, B., Kohayakawa, Y., On Richardson's model in the cube, to appear.

[5] Bollobás, B., Kohayakawa, Y., Łuczak, T., The evolution of random subgraphs of the cube, *Random Structures and Algorithms* **3** (1992), 55–90.

[6] Bollobás, B., Kohayakawa, Y., Łuczak, T., On the evolution of random Boolean functions, to appear.

[7] Bollobás, B., Kohayakawa, Y., Łuczak, T., Connectivity properties of random subgraphs of the cube, to appear.

[8] Burtin, Yu.D., On extreme metric characteristics of a random graph I. Asymptotic estimates, *Mathematical Notes* **19** (1974), 710–725.

[9] Burtin, Yu.D., On extreme metric characteristics of a random graph II. Limit distributions, *Mathematical Notes* **20** (1975), 83–101.

[10] Dyer, M.E., Frieze, A.M., Foulds, L.R., On the strength of connectivity of random subgraphs of the n-cube, in *Random Graphs '85*, Annals of Discrete Mathematics 33 (Karoński, M., Palka, Z., eds), North-Holland, Amsterdam 1987, pp. 17–40.

[11] Fill, J.A., Pemantle, R., Percolation, first-passage percolation, and covering times for Richardson's model on the n-cube, to appear.

[12] Mahrhold, K., Weber, K., On the radius of random subgraphs of the n-cube, in *Topics in Combinatorics and Graph Theory, Oberwolfach 1990* (Bodendieki, R., Henn, R., eds), Physica, Heidelberg 1990, pp. 505–511.

[13] Kostochka, A.V., Maximum matchings and connected components of random subgraphs of the n-cube (in Russian), *Methods Discrete Anal., Study Boolean Function Graphs* **48** (1989), 42–70.

[14] Kostochka, A.V., Matchings in Random Spanning subgraphs of cubelike graphs, *Random Structures and Algorithms* **1** (1990), 277–285.

[15] Kostochka, A.V., Sapozhenko, A.A., Weber, K., On random cubical graphs, in *Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity* (Nešetřil, J., Fiedler, M., eds), Annals of Discrete Mathematics 51, North–Holland, Amsterdam 1992, pp. 155–160.

[16] Kostochka, A.V., Sapozhenko, A.A., Weber, K., Radius and diameter of random subgraphs of the hypercube, *Random Structures and Algorithms* **4** (1993), 215–229.

[17] Sapozhenko, A.A., Metric properties of almost all Boolean functions (in Russian), *Diskretny Analiz* **10** (1967), 91–119.

[18] Sapozhenko, A.A., Weber, K., On the diameter of random subgraphs of the n-cube, in *Contemporary Methods in Graph Theory*, Wissenschaftsverlag 1990, pp. 507–524.

[19] Toman, E., Geometric structure of Random Boolean functions (in Russian), *Problemy Kibernet.* **35** (1979), 111–132.