EXPANSION, LONG CYCLES, AND COMPLETE MINORS IN SUPERCRITICAL RANDOM SUBGRAPHS OF THE HYPERCUBE

JOSHUA ERDE[*], MIHYUN KANG[*], AND MICHAEL KRIVELEVICH[‡]

JUNE 9, 2021

ABSTRACT. Analogous to the case of the binomial random graph $G(d + 1, p)$, it is known that the behaviour of a random subgraph of a d-dimensional hypercube, where we include each edge independently with probability p, which we denote by Q_p^d, undergoes a phase transition around the critical value of $p = \frac{1}{d}$. More precisely, standard arguments show that significantly below this value of p, with probability tending to one as $d \to \infty$ (whp for short) all components of this graph have order $O(d)$, whereas Ajtai, Komlós and Szemerédi [1] showed that significantly above this value, in the *supercritical regime*, whp there is a unique 'giant' component of order $\Theta(2^d)$. In $G(d + 1, p)$ much more is known about the complex structure of the random graph which emerges in this supercritical regime. For example, it is known that in this regime whp $G(d + 1, p)$ contains paths and cycles of length $\Omega(d)$, as well as complete minors of order $\Omega(\sqrt{d})$. In this paper we obtain analogous results in Q_p^d. In particular, we show that for supercritical p, i.e., when $p = \frac{1+\epsilon}{d}$ for a positive constant ϵ, whp Q_p^d contains a cycle of length $\Omega\left(\frac{2^d}{d^3(\log d)^3}\right)$ and a complete minor of order $\Omega\left(\frac{2^{\frac{d}{2}}}{d^3(\log d)^3}\right)$. In order to prove these results, we show that whp the largest component of Q_p^d has good edge-expansion properties, a result of independent interest. We also consider the genus of Q_p^d and show that, in this regime of p, whp the genus is $\Omega(2^d)$.

1. INTRODUCTION

The binomial random graph model $G(n, p)$, introduced by Gilbert [22], is a random variable on the subgraphs of the complete graph K_n whose distribution is given by including each edge in the subgraph independently with probability p. Since its introduction, this model has been extensively studied. A particularly striking feature of this model is the 'phase transition' that it undergoes at $p = \frac{1}{n}$, exhibiting vastly different behaviour when $p = \frac{1-\epsilon}{n}$ to when $p = \frac{1+\epsilon}{n}$ (where ϵ is a positive constant). For more background on the theory of random graphs, see [6, 20, 30].

More recently, the following generalisation of this model has been the object of study: Suppose G is an arbitrary graph with minimum degree $\delta(G)$ at least d and let G_p denote the random subgraph of G obtained by retaining each edge of G independently with probability p. When $G = K_{d+1}$, the complete graph on $d + 1$ vertices, we recover the model $G(d + 1, p)$.

If we take the probability p to be a function of (the lower bound d for) the minimum degree of G, rather than the order (i.e., number of vertices) of G, it has been shown that the model G_p shares many properties with $G(d + 1, p)$. In particular, some of the complex behaviour

which occurs whp[1] in $G(d+1,p)$ once we pass the critical point of $p = \frac{1}{d}$ also occurs whp in G_p in the same regime of p. For example, when $p = \frac{1+\epsilon}{d}$ for $\epsilon > 0$, it has been shown that whp G_p contains a path or cycle of length linear in d, see [16, 37, 38]. Furthermore, for this range of p it has been shown by Frieze and Krivelevich [21] that whp G_p is non-planar and by Erde, Kang and Krivelevich [17] that in fact, whp G_p contains a complete minor of order[2] $\tilde{\Omega}\left(\sqrt{d}\right)$ and hence has genus $\tilde{\Omega}(d)$. Similarly, when $p = \omega\left(\frac{1}{d}\right)$, whp G_p contains a path or cycle of length $(1 - o(1))d$, as shown in [35, 42]. In addition, when $p = (1 + \epsilon)\frac{\log d}{d}$ for $\epsilon > 0$, whp G_p contains a path of length d and in fact even a cycle of length $d + 1$, see [23, 35] respectively. All of these results generalise known results about the binomial model $G(d+1,p)$.

The model $G(d+1,p)$ shows that these results are optimal when G can be an *arbitrary* graph, but for *specific* graphs G, they may be far from the truth. One particular graph for which the model G_p has been studied extensively is the *hypercube Q^d*, the graph with vertex set $V\left(Q^d\right) = \{0,1\}^d$ and in which two vertices are adjacent if they differ in exactly one coordinate. Throughout this paper we will write $n := 2^d$ for the order of the hypercube and we note that $|E\left(Q^d\right)| = \frac{nd}{2}$.

Random subgraphs of the hypercube were first studied by Saposhenko [43] and Burtin [12], who showed that Q_p^d has a threshold for connectivity at $p = \frac{1}{2}$; for a fixed constant $p < \frac{1}{2}$ whp Q_p^d is disconnected and for a fixed constant $p > \frac{1}{2}$ whp Q_p^d is connected, and this result was strengthened by Erdős and Spencer [18] and by Bollobás [4]. Bollobás [5] also showed that $p = \frac{1}{2}$ is the threshold for the existence of a perfect matching in Q_p^d. Very recently, answering a longstanding open problem, Condon, Espuny Díaz, Girao, Kühn and Osthus [14] showed that $p = \frac{1}{2}$ is also the threshold for the existence of a Hamilton cycle in Q_p^d.

It was conjectured by Erdős and Spencer [18] that Q_p^d should undergo a similar phase transition at $p = \frac{1}{d}$ as $G(d+1,p)$ does: It is relatively easy to see, by a coupling with a branching process, that when $p = \frac{1-\epsilon}{d}$ for $\epsilon > 0$, whp all components of Q_p^d have order $O(d)$, but they conjectured that when $p = \frac{1+\epsilon}{d}$, whp Q_p^d contains a unique 'giant' component whose order is linear in $n = 2^d$. This conjecture was confirmed by Ajtai, Komlós and Szemerédi [1].

Theorem 1.1 ([1])**.** Let $\epsilon > 0$ be a constant and let $p = \frac{1+\epsilon}{d}$. Then there exists a $\gamma > 0$ such that whp the largest component of Q_p^d has order at least γn.

Ajtai, Komlós and Szemerédi also indicated that it is possible to show that there is a unique 'giant' component of linear order in n, whose order is $\left(\gamma(\epsilon) + o(1)\right)n$ where $\gamma(\epsilon)$ is the survival probability of the $Po(1 + \epsilon)$ branching process. These results were later extended to a wider range of p, describing more precisely the component structure of Q_p^d when $p = \frac{1+\epsilon}{d}$ with $\epsilon = o(1)$ by Bollobás, Kohayakawa and Łuczak [7], by Borgs, Chayes, van der Hofstad, Slade and Spencer [11], and by Hulshof and Nachmias [29], with the correct order of the critical window in this model being only recently identified by van der Hofstad and Nachmias [27]. For a more detailed background on the phase transition in this model, see the survey of van der Hofstad and Nachmias [26].

If we consider the structure of Q_p^d above this critical threshold of $p = \frac{1}{d}$, then we can conclude from the results in the G_p model mentioned earlier that whp Q_p^d contains a path and

---

[1]Throughout the paper, all asymptotics will be considered as $d \to \infty$ and so, in particular, whp (with high probability) means with probability tending to one as $d \to \infty$

[2]The notation $\tilde{\Omega}(\cdot)$ here is hiding a polylogarithmic factor in d.

cycle of length $\Omega(d)$, has genus $\tilde{\Omega}(d)$ and contains a complete minor of order $\tilde{\Omega}(\sqrt{d})$. However, these results seem far from optimal. Indeed, already above this threshold whp Q_p^d contains a component whose order is linear in $n = 2^d$, and so exponentially large in d. Hence, it is possible that Q_p^d could contain a path and cycle of length linear in n. Similarly, it is known that the genus of Q^d is exponentially large in d. In fact, the genus of Q^d has been shown to be precisely $(d-4)2^{d-3}+1$ for $d \geq 2$, see [3, 41], and the size of the largest complete minor in Q^d was shown to be between $2^{\lfloor \frac{d-1}{2} \rfloor}$ and $\sqrt{d}2^{\frac{d}{2}}+1$ in [13]. Hence, it is possible that whp the genus of Q_p^d in the supercritical regime is linear in n and whp the size of the largest complete minor in Q_p^d is of order $\Omega(\sqrt{n})$.

The main aim of this paper is to show that both of these statements are true, at least up to a polylogarithmic factor in n. More precisely, we prove the following theorems concerning the length of the longest cycle and the size of the largest complete minor in Q_p^d.

Theorem 1.2. Let $\epsilon > 0$ be a constant and let $p = \frac{1+\epsilon}{d}$. Then whp Q_p^d contains a cycle of length $\Omega\left(\frac{2^d}{d^3(\log d)^3}\right)$.

Theorem 1.3. Let $\epsilon > 0$ be a constant and let $p = \frac{1+\epsilon}{d}$. Then whp Q_p^d contains a K_t-minor with $t = \Omega\left(\frac{2^{\frac{d}{2}}}{d^3(\log d)^3}\right)$.

We note that here, and elsewhere in the paper, there may be a dependence on the parameter ϵ in the implicit constants in the bounds.

In fact, both of these theorems follow from a more general statement, of independent interest, which says that whp the largest component of Q_p^d has *good edge-expansion* properties. That is, every large enough subset of the largest component 'expands', in the sense that it has a large edge boundary. The notion of graph expansion has turned out to have fundamental importance in diverse areas of discrete mathematics and computer science: For a comprehensive introduction to expander graphs, see the survey of Hoory, Linial and Widgerson [28]. In particular, notions of expansion have turned out to be a powerful tool in the study of random structures: See for example the survey paper of Krivelevich [33].

In what follows, given a subset $S \subseteq V(Q^d)$, we will write $\partial_v(S)$ and $\partial_e(S)$ for the *vertex* and *edge* boundary of S in Q^d respectively. That is, $\partial_v(S)$ is the set of vertices in $V(Q^d) \setminus S$ which have a neighbour in S and $\partial_e(S)$ is the set of edges between S and $V(Q^d) \setminus S$. Furthermore, we will write $\partial_{v,p}(S)$ and $\partial_{e,p}(S)$ for the vertex and edge boundary of S in Q_p^d respectively. Also, given constants $\alpha, \beta > 0$ and a statement A, we will write 'Let $\alpha \ll \beta$. Then A holds' to indicate that there is some fixed, implicit function f such that A holds for all $\alpha \leq f(\beta)$.

Theorem 1.4. Let $0 < \alpha \ll \epsilon$ be constants, let $p = \frac{1+\epsilon}{d}$ and let $L_1 = L_1\left(Q_p^d\right)$ be the largest component of Q_p^d. Then there exists a $\beta > 0$ such that whp every subset $S \subseteq L_1$ satisfying $\alpha n \leq |S| \leq \frac{|V(L_1)|}{2}$ has an edge boundary satisfying $\left|\partial_{e,p}(S)\right| \geq \beta \frac{n}{d^3(\log d)^2}$.

As an immediate corollary, since Q^d is d-regular, and so every set of m edges in Q^d must span at least $\frac{m}{2d}$ vertices, we can deduce that whp the largest component of Q_p^d also has *good vertex-expansion*. In fact, since in Q_p^d we do not expect significantly fewer than m vertices to span m edges, we can do slightly better.

Corollary 1.5. Let $0 < \alpha \ll \epsilon$ be constants, let $p = \frac{1+\epsilon}{d}$ and let $L_1 = L_1\left(Q_p^d\right)$ be the largest component of Q_p^d. Then there exists a $\beta > 0$ such that whp every subset $S \subseteq L_1$ satisfying $\alpha n \leq |S| \leq \frac{|V(L_1)|}{2}$ has a vertex boundary satisfying $\left|\partial_{v,p}(S)\right| \geq \beta \frac{n}{d^3(\log d)^3}$.

We note that, whilst it is perhaps not immediately clear that we cannot improve Theorem 1.4 to show constant edge-expansion for linear sized subsets of $L_1 = L_1\left(Q_p^d\right)$, it is clear that we cannot hope to show constant vertex-expansion. Indeed, Q^d contains a separator of size $O\left(\frac{n}{\sqrt{d}}\right)$, namely the middle layer, and it is not hard to show that whp this separator splits L_1 roughly in half, and so whp there is a linear sized subset of L_1 whose vertex boundary has size $O\left(\frac{n}{\sqrt{d}}\right)$. However, it then follows from a similar argument as in Corollary 1.5 that whp L_1 contains some linear sized subset whose edge boundary has size $O\left(\frac{n\log d}{\sqrt{d}}\right)$.

In particular, this implies that we cannot deduce optimal versions of Theorems 1.2 or 1.3, that is, ones without the polynomial factors in d, straightforwardly from the expansion properties of the largest component of Q_p^d.

In fact, it is perhaps surprising that the expansion properties of $L_1\left(Q_p^d\right)$ are useful at all, since random subgraphs of the hypercube have been used by Moshkowitz and Shapira [39] to construct graphs which have in some way the worst possible expansion properties.

Whilst Theorem 1.3 already implies that the genus of Q_p^d is $\Omega\left(\frac{n}{d^6(\log d)^6}\right)$, with a more careful argument we are able to give a better bound.

Theorem 1.6. Let $\epsilon > 0$ be a constant and let $p = \frac{1+\epsilon}{d}$. Then whp the genus of Q_p^d is $\Omega(n)$.

Note that, since whp $\left|E\left(Q_p^d\right)\right| = \Theta(n)$, and clearly the genus of a graph is at most the number of edges, Theorem 1.6 is optimal up to the value of the leading constant.

The paper is structured as follows. In Section 2 we collect some lemmas which will be useful in the rest of the paper. Then, in Section 3 we prove our main results on the expansion properties of the largest component in Q_p^d (Theorem 1.4 and Corollary 1.5), and use them to deduce Theorems 1.2 and 1.3. In Section 4 we give a proof of Theorem 1.6, and then in Section 5 we mention some open problems.

2. Preliminaries

For real numbers x, y, z we will write $x = y \pm z$ to mean that $y - z \leq x \leq y + z$.

We will want to use the following simple lemma, which is a slight adaptation of a result in [36], to decompose a tree into roughly equal sized parts.

Lemma 2.1. Let T be a tree such that $\Delta(T) \leq C_1$, all but r vertices of T have degree at most $C_2 \leq C_1$ and $|V(T)| \geq \ell$ for some $C_1, C_2, \ell, r > 0$. Then there exist disjoint vertex sets $A_1, \ldots, A_s \subseteq V(T)$ such that

- $V(T) = \bigcup_{i=1}^s A_i$;
- $T[A_i]$ is connected for each $1 \leq i \leq s$;
- $\ell \leq |A_i| \leq C_1 \ell$ for each $1 \leq i \leq r$; and
- $\ell \leq |A_i| \leq C_2 \ell$ for each $r < i \leq s$.

Proof. We choose an arbitrary root w for T. For a vertex v in a rooted tree S, let us write S_v for the subtree of S rooted at v.

We construct the vertex sets A_i inductively. Let us start by setting $T(0) = T$. Given a tree $T(i)$ rooted at w such that $|V(T(i))| \geq \ell$, let v_i be a vertex of maximal distance from w such that $\left|V\left(T(i)_{v_i}\right)\right| \geq \ell$. We take $A_{i+1} = V\left(T(i)_{v_i}\right)$ and let $T(i+1) = T(i) \setminus T(i)_{v_i}$. We stop when $|V(T(i))| < \ell$, and in that case we add $V(T(i))$ to the final A_i. Finally, let us re-order the sets A_i so that they are non-increasing in size.

We claim that the sets $A_1, A_2, \ldots, A_s$ satisfy the conclusion of the lemma. Indeed, the first two properties are clear by construction. Also, we note that by our choice of v_i, $|T(i)_x| < \ell$ for every child x of v_i. Hence, if $v_i \neq w$, then v_i has $d(v_i) - 1$ children and so, since A_i is the union of the vertices of $T(i)_x$ over all children x of v_i, together with a set of size less than ℓ, it follows that $|A_{i+1}| \leq d(v_i)\ell$. Furthermore, if $v_i = w$, then we note that $A_{i+1} = V(T(i))$ and so $|A_{i+1}| \leq d(w)(\ell - 1) + 1 \leq d(w)\ell$. Therefore, since all but r vertices of T have degree at most C_2, the third and fourth properties also hold. $\qquad\square$

We will also need the following results, which allow us to deduce the existence of a long cycle and a large complete minor from vertex-expansion properties of a graph. For wider context on properties of expanding graphs, see the survey of Krivelevich [33].

Theorem 2.2 ([34, Theorem 1]). Let $k \geq 1, t \geq 2$ be integers. Let G be a graph on more than k vertices satisfying

$$|\partial_v(W)| \geq t, \qquad \text{for every } W \subset V \text{ with } \frac{k}{2} \leq |W| \leq k.$$

Then G contains a cycle of length at least $t + 1$.

The following theorem allows us to deduce the existence of a large complete minor in a graph without any small separators. It is easy to see that graphs with good vertex-expansion do not contain any small separators, and in fact it is known (see [33, Section 5]) that the converse is true, in the sense that graphs without small separators must contain large induced subgraphs with good vertex-expansion.

Theorem 2.3 ([31, Theorem 1.2]). Let G be a graph with N vertices and with no K_t-minor. Then G contains a subset S of order $O(t\sqrt{N})$ such that each connected component of $G \setminus S$ has at most $\frac{2}{3}N$ vertices.

We will use the following Chernoff type bounds on the tail probabilities of the binomial distribution, see e.g. [2, Appendix A].

Lemma 2.4. Let $N \in \mathbb{N}$, let $p \in [0,1]$ and let $X \sim \mathrm{Bin}(N, p)$.

(i) For every positive a with $a \leq \frac{Np}{2}$,

$$\mathbb{P}\left(|X - Np| > a\right) < 2\exp\left(-\frac{a^2}{4Np}\right).$$

(ii) For every positive b,

$$\mathbb{P}\left(X > bNp\right) \leq \left(\frac{e}{b}\right)^{bNp}.$$

In particular, the following consequence of Lemma 2.4 on the edge boundary of large sets in Q_p^d will be useful.

Lemma 2.5. Let $k, c, \epsilon > 0$ be constants and let $p = \frac{1+\epsilon}{d}$. Then whp every subset $X \subseteq V(Q^d)$ of size $|X| \geq cnd^{-k}$ has edge boundary in Q_p^d such that

$$\left|\partial_{e,p}(X)\right| \leq |X|\log d.$$

Proof. Since any subset $W \subseteq V(Q^d)$ satisfies $|\partial_e(W)| \leq d|W|$, it follows from Lemma 2.4 (ii) that

$$\mathbb{P}\left(\left|\partial_{e,p}(W)\right| \geq |W|\log d\right) \leq \mathbb{P}\left(\mathrm{Bin}(d|W|, p) \geq |W|\log d\right)$$
$$\leq \left(\frac{e(1+\epsilon)}{\log d}\right)^{|W|\log d} \leq d^{-(k+1)|W|}.$$

Hence, the probability that there exists a subset $X \subseteq V\left(Q_p^d\right)$ not satisfying the conclusion of the lemma is at most

$$\sum_{i \geq cnd^{-k}} \binom{n}{i} d^{-(k+1)i} \leq \sum_{i \geq cnd^{-k}} \left(\frac{en}{d^{k+1}i}\right)^i \leq \sum_{i \geq cnd^{-k}} \left(\frac{e}{cd}\right)^i = o(1).$$

$\square$

We will also need to use the following correlation inequality, which is a consequence of an inequality of Harris [25] and is itself a special case of the FKG-inequality: See for example [2, Section, 6].

Lemma 2.6. Let $\mathscr{A}$ and $\mathscr{B}$ be families of subgraphs of Q^d which are closed under taking supergraphs. Then

$$\mathbb{P}\left(Q_p^d \in \mathscr{A} \cap \mathscr{B}\right) \geq \mathbb{P}\left(Q_p^d \in \mathscr{A}\right)\mathbb{P}\left(Q_p^d \in \mathscr{B}\right).$$

Finally, we will use the following result on vertex-isoperimetry in the hypercube, which is a useful strengthening of Harper's [24] well-known isoperimetric inequality.

Theorem 2.7 ([9, Corollary 2]). Let $A \subseteq V\left(Q^d\right)$ and let $0 \leq \lambda < 1$ and $k \in \mathbb{N}$ be such that

$$|A| = \sum_{i=0}^{k} \binom{d}{i} + \lambda \binom{d}{k+1} \leq \frac{n}{2}.$$

Then there is a matching from A to its complement A^c of size at least $(1-\lambda)\binom{d}{k} + \lambda\binom{d}{k+1}$.

In particular, for any subset $A \subset V\left(Q^d\right)$ such that both A and A^c have size $\Omega(n)$, there is a matching from A to A^c of size $\Omega\left(\frac{n}{\sqrt{d}}\right)$.

3. Edge-expansion in the largest component

We will need some properties of the largest component of Q_p^d given in [1], where the statement below is claimed, although the details of the proof are not spelled out. For completeness we include a proof in Appendix A.

Theorem 3.1 ([1]). Let $0 < \delta_2 \ll \delta_1$, let $p = \frac{1+\delta_1}{d}$ and let γ_1 be the survival probability of the $\mathrm{Po}(1 + \delta_1)$ branching process. Then there exist constants $\delta_3, \delta_4, \delta_5 > 0$, which can be chosen such that $\delta_3 \ll \delta_2$, such that, if Y is the set of vertices of Q_p^d contained in components of order at least $\delta_3 d^2$, then whp:

(i) there is a unique component L_1 of order at least $\delta_2 n$ in Q_p^d and $|V(L_1)| = (\gamma_1 \pm \delta_2)n$;
(ii) $|V(L_1) \triangle Y| \leq \delta_2 n$;
(iii) the set of vertices

$$X = \left\{x \in V\left(Q^d\right) : |\partial_v(x) \cap Y| \leq \delta_4 d\right\}$$

satisfies $|X| \leq 2^{(1-\delta_5)d}$.

With this theorem in hand, let us briefly sketch the strategy to prove Theorem 1.4 on the edge-expansion properties of the largest component of Q_p^d. We will use a sprinkling argument, viewing Q_p^d as the union of two independent random subgraphs $Q_{q_1}^d$ and $Q_{q_2}^d$ where q_1 is a supercritical probability chosen sufficiently close to p such that, if we denote by L_1' and L_1 the largest component in $Q_{q_1}^d$ and Q_p^d respectively, then whp L_1' contains most of the vertices in L_1, which we can guarantee by Theorem 3.1.

After exposing $Q_{q_1}^d$, it will follow from Property (ii) of Theorem 3.1 that whp most vertices in L_1' are contained in components of $Q_{q_1}^d$ of order $\Omega(d^2)$. Using Lemma 2.1, we can split the subgraph of $Q_{q_1}^d$ consisting of the union of these components into a collection $\mathscr{C}$ of connected pieces which all have polynomial size in d, specifically almost all the pieces will have size around $d^{\frac{3}{2}}$. We then show that whp for any partition $\mathscr{C} = \mathscr{C}_A \cup \mathscr{C}_B$ of $\mathscr{C}$ into two parts, each covering a positive proportion of the vertices in L_1', there are exponentially many edge-disjoint paths between vertices in the two partition classes in $Q_{q_2}^d$. We do so via a union bound, showing that the probability that any partition fails to have this property is much smaller than the total number of possible partitions. For partitions in which the two partition classes have many shared neighbours we will use property (iii) of Theorem 3.1 to find these paths and otherwise we will additionally use Theorem 2.7.

Given this, we note that any large subset $S \subseteq V(L_1)$ must either split apart exponentially many of the pieces in $\mathscr{C}$, in which case for each piece which is split there will be some edge of $Q_{q_1}^d$ present in $\partial_{e,p}(S)$, or there must be a partition $\mathscr{C} = \mathscr{C}_A \cup \mathscr{C}_B$ of $\mathscr{C}$ into two parts, each covering a positive proportion of L_1', such that each piece in $\mathscr{C}_A$ is contained in S and almost every piece in $\mathscr{C}_B$ is disjoint from S. Our claim then implies that there are many edge disjoint paths in $Q_{q_2}^d$ between the two partition classes, almost all of which must contribute at least one edge to $\partial_{e,p}(S)$.

Proof of Theorem 1.4. During the proof we will introduce a series of constants $c_1, c_2, \ldots$, with the understanding that each c_i will be as small as is necessary for the argument in terms of α, ϵ and all c_j with $j < i$.

Let us fix $c_2 \ll c_1$, let γ be the survival probability of a $\mathrm{Po}(1+\epsilon)$ branching process and let us choose $\delta_1 < \epsilon$ such that γ_1, the survival probability of a $\mathrm{Po}(1+\delta_1)$ branching process, is such that $\gamma - \gamma_1 \le c_2$. We take δ_1 and $\delta_2 = c_1$ in Theorem 3.1 to find $c_2 \gg c_3 = \delta_3$, $c_4 = \delta_4$ and $c_5 = \delta_5$. Note that we may assume that $c_5 \ll c_4 \ll c_3$.

Our plan is to expose Q_p^d in two steps, with $q_1 = \frac{1+\delta_1}{d}$ and $q_2 = \frac{p-q_1}{1-q_1}$, where we may assume that $q_2 \ge \frac{c_6}{d}$. Let us write $Q_1 = Q_{q_1}^d$ and $Q_2 = Q_1 \cup Q_{q_2}^d$ where the two random subgraphs are chosen independently, so that Q_2 has the same distribution as Q_p^d.

By Theorem 3.1, whp the largest component L_1' in Q_1 is such that $|V(L_1')| = (\gamma_1 \pm c_1)n = (\gamma \pm 2c_1)n$. Furthermore, if Y is the set of vertices of Q_1 contained in components of order at least $c_3 d^2$ then whp $|V(L_1') \triangle Y| \le c_1 n$ and

$$X = \left\{ x \in V\left(Q^d\right) : |\partial_v(x) \cap Y| \le c_4 d \right\} \tag{1}$$

satisfies $|X| \le 2^{(1-c_5)d}$.

Recall that L_1 is the largest component of Q_p^d. We first show that whp the symmetric difference of $V(L_1)$ and Y is small.

Claim 3.2. Whp $|V(L_1) \triangle Y| \le 4c_1 n$.

Proof of Claim 3.2. We apply Theorem 3.1 with $\delta_1 = \epsilon$ and $\delta_2 = c_1$ to $Q_2 = Q_p^d$. It follows that whp there is a unique component L_1 in Q_2 of order at least $c_1 n$ and that $|V(L_1)| = (\gamma \pm c_1)n$. However, since $Q_1 \subseteq Q_2$ and $|V(L_1')| \ge (\gamma - 2c_1)n \ge c_1 n$, it follows that $L_1' \subseteq L_1$. Hence, $|V(L_1) \setminus V(L_1')| \le 3c_1 n$, and so $|V(L_1) \triangle Y| \le 4c_1 n$. □

Next, we want to use Lemma 2.1 to split each component of Q_1 of order at least $c_3 d^2$ into connected pieces of roughly equal size. In order to do so we need the following bound on the degree sequence of Q_1.

Claim 3.3. Whp Q_1 contains at most nd^{-5} vertices of degree at least $\log d$.

Proof of Claim 3.3. For any fixed vertex $v \in V(Q^d)$, the degree of v in $Q_1 = Q_{q_1}^d$ is distributed as $\text{Bin}(d, q_1)$, and so by Lemma 2.4 (ii) we have that

$$\mathbb{P}\left(d_{Q_1}(v) \geq \log d\right) \leq \left(\frac{e(1+\delta_1)}{\log d}\right)^{\log d} \leq d^{-\frac{\log\log d}{2}}.$$

It follows that the expected number of vertices in Q_1 with degree at least $\log d$ is at most $nd^{-\frac{\log\log d}{2}}$. Hence, by Markov's inequality, whp there at most nd^{-5} vertices with degree at least $\log d$. $\qquad\square$

We note that the above argument is rather unoptimised and with a little more care, the bound on the degree of the exceptional vertices could be improved from $\log d$ to $\frac{C\log d}{\log\log d}$ for some suitably large constant C. However, for ease of presentation we have not attempted to optimise the logarithmic factors in the proof.

Assuming that Claim 3.3 holds, by applying Lemma 2.1 with $\ell = \frac{1}{c_{10}}d^{\frac{3}{2}}$ to a spanning tree of each component of Q_1 of order at least $c_3 d^2$, noting that $\Delta(Q_1) \leq \Delta(Q^d) = d$, we can split the collection of these components into connected pieces such that at most nd^{-5} of the pieces have size between $\ell = \frac{1}{c_{10}}d^{\frac{3}{2}}$ and $\ell d = \frac{1}{c_{10}}d^{\frac{5}{2}}$ and the rest have size between ℓ and $\ell\log d = \frac{1}{c_{10}}d^{\frac{3}{2}}\log d$. Let us call this collection of pieces $\mathscr{C}$, noting that $V(\bigcup\mathscr{C}) = Y$.

We next show that whp for any partition $\mathscr{C} = \mathscr{C}_A \cup \mathscr{C}_B$ of $\mathscr{C}$ into two parts, each covering a positive proportion of the vertices of Y, we expose exponentially many edge-disjoint paths between A and B in the second sprinkling step. This argument will require that the pieces in $\mathscr{C}$ are not significantly smaller than $d^{\frac{3}{2}}$, whereas later bounds will be optimised by taking the pieces in $\mathscr{C}$ as small as possible, which motivates our choice of ℓ in the above.

Claim 3.4. Whp for any partition $\mathscr{C} = \mathscr{C}_A \cup \mathscr{C}_B$ of $\mathscr{C}$ into two parts, with $A = V(\bigcup\mathscr{C}_A)$ and $B = V(\bigcup\mathscr{C}_B)$, such that $\min\{|A|, |B|\} \geq c_7 n$, there are at least $c_{11}nd^{-\frac{3}{2}}$ edge-disjoint paths between A and B in $Q_{q_2}^d$.

Proof of Claim 3.4. The total number of pieces in $\mathscr{C}$ is at most $c_{10}nd^{-\frac{3}{2}}$ and so the total number of possible partitions is at most $2^{c_{10}nd^{-\frac{3}{2}}}$. Hence, if we can show that the claim does not hold for a fixed partition with probability at most $\exp\left(-c_{10}nd^{-\frac{3}{2}}\right)$, then whp the claim holds for all partitions by the union bound.

Given a fixed partition $\mathscr{C}_A \cup \mathscr{C}_B$ with A and B as above, let $N(A) = A \cup \partial_{v,q_1}(A)$ be the inclusive neighbourhood of A in Q_1 and similarly let $N(B)$ be the inclusive neighbourhood of B in Q_1. Let $D = N(A) \cap N(B)$. We split into two cases:

  (i) $|D| \geq c_8 nd^{-\frac{1}{2}}$;
  (ii) $|D| < c_8 nd^{-\frac{1}{2}}$.

Case (i): Let us suppose first that $|D| \geq c_8 nd^{-\frac{1}{2}}$. In this case, let $D' = D \setminus X$, where X is as in (1), so that $|D'| \geq \frac{|D|}{2}$, since $|X| \leq n^{1-c_5}$.

Since each $x \in D'$ is in D, x is either in A or has a neighbour in A and similarly x is either in B or has a neighbour in B. Moreover, since $x \notin X$ and $A \cup B = Y$, it follows that x has at least $\frac{c_4 d}{2}$ many neighbours in one of A or B. For each $x \in D'$, let us fix a set E_x of edges witnessing this, consisting of $\frac{c_4 d}{2}$ edges from x to one of the vertex sets A or B together with a single edge from x to the other vertex set if x is not already an element of that set (see Figure 1).
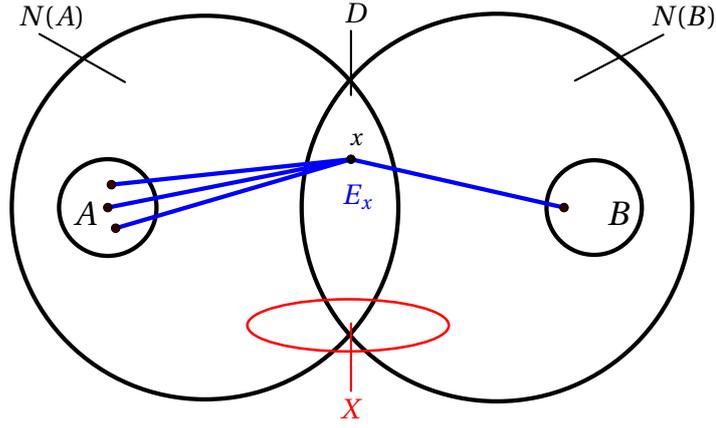
FIGURE 1. A vertex $x \in D' = D \setminus X$ with corresponding edge set E_x consisting of $\frac{c_4 d}{2}$ edges from x to A and an edge from x to B

Let us choose a subset $D'' \subseteq D'$ of size at least $\frac{|D'|}{2}$ such that all $x \in D''$ have the same parity, and hence the edge sets E_x are disjoint for each $x \in D''$.

Then, for each $x \in D''$, there is a path between A and B using edges from E_x in $Q_{q_2}^d$ with probability at least

$$q_2 \left(1 - (1 - q_2)^{\frac{c_4 d}{2}}\right) \geq q_2 \left(1 - \exp\left(-\frac{q_2 c_4 d}{2}\right)\right) \geq \frac{c_6 \left(1 - \exp\left(-\frac{c_4 c_6}{2}\right)\right)}{d} \geq \frac{2 c_9}{d},$$

since $1 - x \leq e^{-x}$ for all $x > 0$ and $q_2 \geq \frac{c_6}{d}$.

Since the sets E_x are disjoint, these events are independent for different $x \in D''$ and so, by Chernoff type bounds (Lemma 2.4 (i)), we have that there are less than $\frac{c_9}{d}|D''| \geq c_{11} n d^{-\frac{3}{2}}$ such paths with probability at most $2 \exp\left(-\frac{c_8 c_9 n d^{-\frac{3}{2}}}{32}\right) \leq \exp\left(-c_{10} n d^{-\frac{3}{2}}\right)$.

Case (ii)**:** Let us now deal with the second case, where $|D| < c_8 n d^{-\frac{1}{2}}$. In this case we note that $|N(A)| \geq |A| \geq c_7 n$ and $|N(A) \cap B| \leq |D| \leq c_8 n d^{-\frac{1}{2}}$, and so $|N(A)^c| \geq |B| - |D| \geq \frac{c_7}{2} n$. Therefore, it follows from Theorem 2.7 that there is a matching F in Q^d of size at least $3 c_8 n d^{-\frac{1}{2}}$ from $N(A)$ to $N(A)^c$.

Since $|X| \ll |D| < c_8 n d^{-\frac{1}{2}}$, if we let F' be those edges in F which do not have an endpoint in either D or X then $|F'| \geq c_8 n d^{-\frac{1}{2}}$. For each edge $uv = e$ in F', where without loss of generality $u \in N(A) \setminus (D \cup X)$, either $u \in A$ or u has at least $c_4 d$ many neighbours in A and similarly either $v \in B$ or v has at least $c_4 d$ many neighbours in B. For each $e \in F'$, let us fix a set of edges E_e witnessing this, consisting of e together with $c_4 d$ edges from u to A if $u \notin A$ and $c_4 d$ edges from v to B if $v \notin B$ (see Figure 2).

Then, for each $uv = e \in F'$, there is a path in $Q_{q_2}^d$ between A and B using edges from E_e with probability at least

$$q_2 \left(1 - (1 - q_2)^{c_4 d}\right)^2 \geq q_2 \left(1 - \exp\left(-q_2 c_4 d\right)\right)^2 \geq \frac{c_6 \left(1 - \exp\left(-c_4 c_6\right)\right)^2}{d} \geq \frac{2 c_9}{d}.$$

Since the sets E_e are disjoint, these events are independent for different $e \in F'$ and so, by Chernoff type bounds (Lemma 2.4 (i)), we have that there are less than $\frac{c_9}{d}|F'| \geq c_{11} n d^{-\frac{3}{2}}$ such paths with probability at most $2 \exp\left(-\frac{c_8 c_9 n d^{-\frac{3}{2}}}{16}\right) \leq \exp\left(-c_{10} n d^{-\frac{3}{2}}\right)$. $\qquad \square$
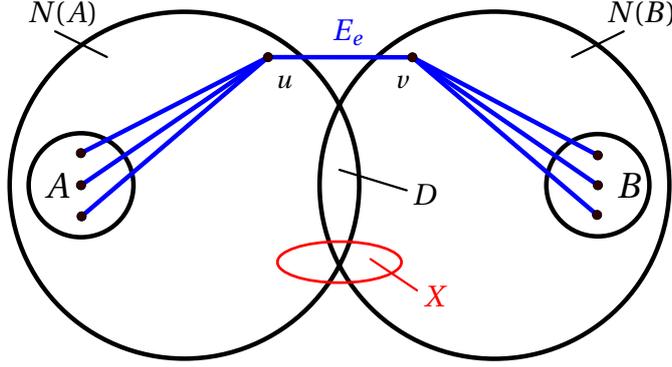
FIGURE 2. An edge $uv = e \in F'$ with corresponding edge set E_e consisting of $c_4 d$ edges from u to A and $c_4 d$ edges from v to B.

Assuming Claims 3.2 and 3.4, our aim is to show that the edge boundary of every subset $S \subseteq L_1$ satisfying $\alpha n \leq |S| \leq \frac{|V(L_1)|}{2}$ is large. Let us fix some such subset S.

Claim 3.5. Whp $\left|\partial_{e,p}(S)\right| \geq \frac{c_{12} n}{d^3 (\log d)^2}$.

Proof of Claim 3.5. We first note that by Claim 3.2 whp $|S \cap Y| \geq (\alpha - 4c_1) n$. Let us write

$$\mathscr{C}_0 = \{C \in \mathscr{C} : S \cap C \neq \emptyset \text{ and } C \setminus S \neq \emptyset\},$$
$$\mathscr{C}_A = \{C \in \mathscr{C} : C \subset S\},$$
$$\mathscr{C}_B = \{C \in \mathscr{C} : S \cap C = \emptyset\},$$

and let $A = V\left(\bigcup \mathscr{C}_A\right)$ and $B = V\left(\bigcup \mathscr{C}_B \cup \bigcup \mathscr{C}_0\right)$.

Since each $C \in \mathscr{C}$ is connected in $Q_1 \subseteq Q_2$, there are at least $|\mathscr{C}_0|$ edges in the edge boundary of S in Q_2. We split into two cases:

(i) $|\mathscr{C}_0| \geq \frac{c_{12} n}{d^3 (\log d)^2}$;
(ii) $|\mathscr{C}_0| < \frac{c_{12} n}{d^3 (\log d)^2}$.

Case (i)**:** Suppose first that $|\mathscr{C}_0| \geq \frac{c_{12} n}{d^3 (\log d)^2}$. Since each $C \in \mathscr{C}_0$ is connected in Q_1 and meets both S and $V(Q^d) \setminus S$, it follows that there are at least $\frac{c_{12} n}{d^3 (\log d)^2}$ many edges in the edge boundary of S in Q_1 and hence, at least this many in the edge boundary of S in Q_2.

Case (ii)**:** Suppose instead that $|\mathscr{C}_0| < \frac{c_{12} n}{d^3 (\log d)^2}$. Since there are at most $n d^{-5}$ pieces in $\mathscr{C}$ of size between $\frac{1}{c_{10}} d^{\frac{3}{2}}$ and $\frac{1}{c_{10}} d^{\frac{5}{2}}$ and the rest have size at most $\frac{1}{c_{10}} d^{\frac{3}{2}} \log d$, it follows that, if we let $Z := V\left(\bigcup \mathscr{C}_0\right)$, then

$$|Z| \leq \frac{2 c_{12} n}{c_{10} d^{\frac{3}{2}} \log d}. \tag{2}$$

In particular, we have that $|A| \geq |S| - |Z| \geq c_7 n$ and $|B| \geq |Y| - |S| \geq c_7 n$, and so, by Claim 3.4, whp there is a family $\mathscr{P}$ of at least $c_{11} n d^{-\frac{3}{2}}$ many edge disjoint paths between A and B in $Q_{q_2}^d$.

Then, by Lemma 2.5 applied with $k = \frac{5}{2}$, whp either $|Z| \leq \frac{c_{11}}{2} n d^{-\frac{5}{2}}$, or Z has edge boundary satisfying

$$\left|\partial_{e,p}(Z)\right| \leq |Z| \log d \leq \frac{2c_{12}}{c_{10}} n d^{-\frac{3}{2}} \leq \frac{c_{11}}{2} n d^{-\frac{3}{2}},$$

where we used (2) to bound $|Z|$ from above. In either case, since every vertex in Z has degree at most d, at most half of the paths in $\mathscr{P}$ meet Z, and every other path in $\mathscr{P}$ must contribute at least one edge to the boundary of S in Q_2, since $A \subseteq S$ and $(B \setminus Z) \cap S = \emptyset$. It follows that the edge boundary of S in Q_2 has size at least $\frac{c_{11}}{2} n d^{-\frac{3}{2}} \geq \frac{c_{12} n}{d^3 (\log d)^2}$. $\qquad \square$

It follows from Claim 3.5 that the theorem holds with $\beta = c_{12}$. $\qquad \square$

Proof of Corollary 1.5. Let β be given by Theorem 1.4. By Lemma 2.5, we have that whp every subset $X \subseteq V(Q^d)$ of size $|X| \geq n d^{-5}$ has edge boundary in Q_p^d such that

$$\left|\partial_{e,p}(X)\right| \leq |X| \log d.$$

So, we may assume that the above and the conclusion of Theorem 1.4 hold. Let $S \subseteq L$ be such that $\alpha n \leq |S| \leq \frac{|V(L_1)|}{2}$ and let $X = \partial_{v,p}(S)$ be the vertex boundary of S in Q_p^d. Then, by our choice of β, we have $\left|\partial_{e,p}(S)\right| \geq \beta \frac{n}{d^3 (\log d)^2}$ and so, since Q^d is d-regular, it follows that $|X| \geq \frac{n}{d^4 (\log d)^2} \geq n d^{-5}$. Hence, $\left|\partial_{e,p}(X)\right| \leq |X| \log d$.

However, since each edge in the edge boundary of S is in the edge boundary of X we obtain

$$|X| \log d \geq \left|\partial_{e,p}(X)\right| \geq \left|\partial_{e,p}(S)\right| \geq \beta \frac{n}{d^3 (\log d)^2}.$$

It follows that $|X| \geq \beta \frac{n}{d^3 (\log d)^3}$, and so the corollary follows $\qquad \square$

As corollaries, we can deduce Theorems 1.2 and 1.3.

Proof of Theorem 1.2. By Corollary 1.5 given $\alpha \ll \epsilon$ there exists a $\beta > 0$ such that whp every subset S of the largest component L_1 of Q_p^d with $\alpha n \leq |S| \leq \frac{|V(L_1)|}{2}$ satisfies $\left|\partial_{v,p}(S)\right| \geq \frac{\beta n}{d^3 (\log d)^3}$. Hence, applying Theorem 2.2 with $k = \frac{|V(L_1)|}{2}$ and $t = \frac{\beta n}{d^3 (\log d)^3}$ we can conclude that L_1 contains a cycle of length $\Omega\left(\frac{n}{d^3 (\log d)^3}\right)$. $\qquad \square$

Proof of Theorem 1.3. If the largest component L_1 of Q_p^d does not contain a K_t-minor, then by Theorem 2.3 there is some constant $C > 0$ such that L_1 contains a subset S of size at most $Ct\sqrt{|V(L_1)|} \leq Ct\sqrt{n}$, such that each component of $G \setminus S$ has order at most $\frac{2|V(L_1)|}{3}$. It follows that we can find some subset of L_1 of size between $\frac{|V(L_1)|}{3}$ and $\frac{|V(L_1)|}{2}$ whose vertex boundary is contained in S, and hence has size at most $Ct\sqrt{n}$.

However, by Corollary 1.5, whp every subset of L_1 of size between $\frac{1}{3}|V(L_1)|$ and $\frac{1}{2}|V(L_1)|$ has a vertex boundary of size at least $\frac{\beta n}{d^3 (\log d)^3}$. It follows that $t \geq \frac{\beta \sqrt{n}}{Cd^3 (\log d)^3} = \Omega\left(\frac{\sqrt{n}}{d^3 (\log d)^3}\right)$. $\qquad \square$

4. Genus of the largest component

Given a graph G let us write $g(G)$ for the *genus* of G, the smallest $k \in \mathbb{N}$ such that G can be embedded on an orientable surface of genus k.

Since the genus of K_t is known to be $\Omega(t^2)$, and the genus of a graph is a *minor monotone function*, that is, if H is a minor of G then $g(H) \leq g(G)$, it follows from Theorem 1.3 that whp $g(Q_p^d) = \Omega\left(\frac{n}{d^3 (\log d)^3}\right)$ in the supercritical regime. However, if we are more careful, we can improve this to a bound of optimal asymptotic order.

Proof of Theorem 1.6. Euler's formula tells us that the genus of a connected graph L satisfies

$$g(L) = \frac{1}{2}\left(e(L) - v(L) - f(L) + 2\right) = \frac{1}{2}\left(\text{excess}(L) - f(L) + 2\right), \tag{3}$$

where $e(L)$ and $v(L)$ are the number of edges and vertices of L respectively, $\text{excess}(L) := e(L) - v(L)$ is the *excess* of L and $f(L)$ is the number of faces of L when embedded on a surface of minimal genus. We will show that whp the largest component L_1 of Q_p^d satisfies

(i) $\text{excess}(L_1) = \Omega(n)$; and

(ii) $f(L_1) = o(n)$,

from which it follows, from (3), that whp $g(L_1)$ is also $\Omega(n)$.

Property (i)**:** To prove that whp $\text{excess}(L_1) = \Omega(n)$, we again argue via a two-round exposure. Let $q_1 = \frac{1+\frac{\epsilon}{2}}{d}$, so that $q_2 = \frac{p - p_1}{1 - p_1} \geq \frac{\epsilon}{2d}$ and let us write as before $Q_1 = Q_{q_1}^d$ and $Q_2 = Q_1 \cup Q_{q_2}^d$ where the two random subgraphs are independent, so that Q_2 has the same distribution as Q_p^d.

By Theorem 1.1 there exists $\alpha > 0$ such that whp there is a component L_1' of Q_1 with $|V(L_1')| \geq \alpha n$. Since L_1' is connected it contains at least $|V(L_1')| - 1$ edges of Q_1.

However, if we let $\beta \ll \alpha$ and let $\mathscr{A}$ be the event that there exists a set of vertices $W \subseteq V\left(Q^d\right)$ of size at least αn such that W spans at most $\beta d|W|$ edges in Q^d and W spans at least $|W| - 1 \geq \frac{|W|}{2}$ edges of Q_1, then it follows, by Lemma 2.4 (ii), that

$$\mathbb{P}(\mathscr{A}) \leq \sum_{i \geq \alpha n} \binom{n}{i} \mathbb{P}\left(\text{Bin}\left(\lfloor \beta d i \rfloor, \frac{2}{d}\right) \geq \frac{i}{2}\right) \leq \sum_{i \geq \alpha n} \left(\frac{en}{i}\right)^i (2\beta e)^{\frac{i}{2}} = \sum_{i \geq \alpha n} \left(\frac{\sqrt{2}e^{\frac{3}{2}}\sqrt{\beta}}{\alpha}\right)^i = o(1).$$

Hence, we can conclude that whp L_1' spans at least $\beta d|V(L_1')| \geq \alpha\beta dn$ many edges of Q^d. Let T be a spanning tree of L_1' in Q_1, then L_1' spans at least $\alpha\beta dn - |V(T)| \geq \frac{\alpha\beta dn}{2}$ edges in $Q^d - T$.

Hence, again by Chernoff type bounds (Lemma 2.4 (i)), we have that the number of edges spanned by L_1' in $Q_{q_2}^d \setminus T$, which we denote by Y, is such that

$$\mathbb{P}\left(Y \leq \frac{\alpha\beta\epsilon n}{8}\right) \leq \mathbb{P}\left(\text{Bin}\left(\left\lceil \frac{\alpha\beta dn}{2}\right\rceil, \frac{\epsilon}{2d}\right) \leq \frac{\alpha\beta\epsilon n}{8}\right) \leq 2\exp\left(-\frac{\alpha\beta\epsilon n}{32}\right) = o(1).$$

So, we can conclude that whp there is some component $L_1 \supseteq L_1'$ in $Q_2 = Q_1 \cup Q_{q_2}^d$ which has excess at least $\frac{\alpha\beta\epsilon n}{8}$. Note that, by Theorem 3.1 whp there is a unique linear sized component in Q_2 and so whp L_1 is in fact the largest component in Q_2. Summing up, whp

$$\text{excess}(L_1) \geq \frac{\alpha\beta\epsilon n}{8}. \tag{4}$$

We note that the typical existence of a component of large excess in Q_p^d can also be deduced in a relatively straightforward manner from Theorem 3.1. Indeed, by Theorem 3.1 almost every vertex in Q^d, and so almost every vertex in the giant component L of Q_1, has linearly many (in d) neighbours in L. It will then follow that whp after sprinkling we add linearly many (in n) edges to the vertex set of L in Q_2.

Property (ii)**:** To prove that whp $f(L_1) = o(n)$, let us consider the expected number of cycles of length s in Q_p^d for a fixed s. We note that we can naively bound the number of cycles of length s in Q^d by $2nd^{s-2}$. Indeed, for any vertex x there are at most d^{s-2} many paths of length $s - 2$ starting at x. However, since the codegree of any pair of vertices in Q^d is at most two, there are at most two ways to complete any path of length $s - 2$ to a cycle of length s. Since there are at most n many ways to choose the initial vertex x, the bound follows.

We mention that this bound is far from tight and indeed, without too much effort, it can be shown that the number of cycles of length s in Q^d is at most $2^d \left(\frac{sd}{2}\right)^{\frac{s}{2}}$ (see [39, Claim 3.2]).

Since each such cycle is in Q_p^d with probability p^s, and $p = \frac{1+\epsilon}{d}$, it follows that the number of cycles of length s, which we denote by X_s, satisfies

$$\mathbb{E}(X_s) \le 2nd^{s-2}p^s = 2nd^{-2}(1+\epsilon)^s.$$

Hence, the expected number of cycles of length at most $\sqrt{\log d}$ in Q_p^d is at most

$$\sum_{s=1}^{\sqrt{\log d}} \mathbb{E}(X_s) \le 2nd^{-2} \sum_{s=1}^{\sqrt{\log d}} (1+\epsilon)^s \le 2nd^{-2}\sqrt{\log d}\, 2^{\sqrt{\log d}} \le nd^{-1}.$$

Hence, by Markov's inequality, whp the number of cycles of length at most $\sqrt{\log d}$ in Q_p^d is $o(n)$. It follows that whp the number of faces of length at most $\sqrt{\log d}$ in any embedding of L_1 is $o(n)$. Furthermore, since in any embedding each edge is in at most two faces, the number of faces of length at least $\sqrt{\log d}$ in any embedding of L_1, which we denote by $Y_{\ge\sqrt{\log d}}$, is at most $\frac{2e(L_1)}{\sqrt{\log d}} \le \frac{2e\left(Q_p^d\right)}{\sqrt{\log d}}$. However, again by Lemma 2.4, we have that whp $e\left(Q_p^d\right) \le dnp = (1+\epsilon)n$ and so whp

$$Y_{\ge\sqrt{\log d}} \le \frac{2e\left(Q_p^d\right)}{\sqrt{\log d}} \le \frac{2(1+\epsilon)n}{\sqrt{\log d}} = o(n).$$

Hence, whp $f(L_1) = o(n)$ and so, by Euler's formula (3) applied to L_1, we have

$$g\left(Q_p^d\right) \ge g(L_1) \ge \frac{1}{2}\left(\text{excess}(L_1) - f(L_1) + 2\right) \overset{(4)}{\ge} \frac{1}{2}\left(\frac{\alpha\beta\epsilon n}{8} - o(n)\right) = \Omega(n).$$

$\square$

5. DISCUSSION

In Theorems 1.2 and 1.3 we have shown that in the supercritical regime when $p = \frac{1+\epsilon}{d}$ for a positive constant ϵ, whp Q_p^d contains a cycle whose length is *almost linear* in $n = 2^d$, that is, up to some polylogarithmic term in n, and similarly a complete minor of order almost $\sqrt{n}$. However, it seems unlikely that these results are best possible, and analogous to the case of $G(d+1, p)$, it is natural to conjecture that for supercritical p whp Q_p^d contains a cycle whose length is in fact *linear* in n, and a complete minor of order $\sqrt{n}$.

Question 5.1. Let $\epsilon > 0$ and $p = \frac{1+\epsilon}{d}$.

(1) Is it the case that whp Q_p^d contains a cycle of length $\Omega(n)$?
(2) Is it the case that whp Q_p^d contains a complete minor of order $\Omega\left(\sqrt{n}\right)$?

Furthermore, in both questions it would also be interesting to know the dependence of the results on ϵ. For example, in $G(d+1, p)$ it is known that for $p = \frac{1+\epsilon}{d}$ the largest component is of order $(2\epsilon + o(\epsilon))d$, the length of the longest cycle is of order $\Theta\left(\epsilon^2\right)d$ (see, for example, [30, Theorem 5.7]) and the order of the largest complete minor is $\Theta\left(\epsilon^{\frac{3}{2}}\right)\sqrt{d}$ (see [19]).

As indicated in the introduction, it seems unlikely that such results can be proven simply by considering the expansion properties of the largest component. However, it is still an interesting question precisely how strong expansion properties of the largest component we can guarantee whp. For example, Krivelevich [32] showed that in the supercritical regime $p = \frac{1+\epsilon}{d}$ in $G(d+1, p)$, whp the largest component contains a subgraph of order $\Theta(d)$ which

13

is a bounded degree α-expander for some $\alpha = \alpha(\epsilon) > 0$, where a graph G is an α-*expander* if for every subset $W \subseteq V(G)$ with $|W| \le \frac{|V(G)|}{2}$ we have that $|\partial_v(W)| \ge \alpha|W|$.

Question 5.2. Let $\epsilon > 0$ and $p = \frac{1+\epsilon}{d}$. For what $\alpha = \alpha(d)$ is it true that whp the largest component of Q_p^d contains a subgraph of order $\Theta(n)$ which is an α-expander?

We note that by a similar argument as in [33, Lemma 2.4], Corollary 1.5 implies that we can take $\alpha(d) = \Omega\left(\frac{1}{d^3(\log d)^3}\right)$, however it seems unlikely that this is optimal.

A careful examination of the proof of Theorem 1.6 shows that whp the genus of the largest component L_1 of Q_p^d is asymptotically equal to half of the excess of L_1. In the case of $G(d+1, p)$, in the supercritical regime, it can be shown that if we delete the largest component, what is left has approximately the distribution of a subcritical random graph, and so whp all other components are unicyclic or trees. From this it follows that the genus of $G(d+1, p)$ is asymptotically equal to the genus of its largest component. Furthermore, by Euler's formula for an arbitrary graph G

$$g(G) = \frac{1}{2}\left(e(G) - v(G) - f(G) + \kappa(G) + 1\right),$$

where $\kappa(G)$ is the number of components of G, and by counting carefully the number of tree-components it is possible to determine asymptotically the genus of $G(d+1, p)$ in this regime, and also the excess of the giant component (see [15, Theorems 1.2 and 1.3]).

However, it is not the case that whp Q_p^d is planar for an arbitrary subcritical p. For example, it is easy to see that whp $Q^4 \subseteq Q_p^d$ when $p = \Theta\left(\frac{1}{d}\right)$, and Q^4 is non-planar. So, even if such a 'symmetry rule' were to hold in this model, it would not be immediate that the main contribution to the genus of Q_p^d is coming from the excess of the largest component. It would be interesting to determine asymptotically the genus of Q_p^d in the supercritical regime.

There are some interesting open questions about the model Q_p^d in the paper of Condon, Espuny Díaz, Girao, Kühn and Osthus [14]. In particular, they used as a crucial part of their proof the fact that whp $Q_{\frac{1}{2}}^d$ contains an 'almost spanning' path, that is, a path containing $(1 - o(1))n$ vertices, and they showed that this property is in fact true for Q_p^d for any constant p.

However, analogous to the case of $G(d+1, p)$, we should perhaps expect such a path to exist for much smaller values of p. In particular, if we expect the sparse random subgraph Q_p^d with $p = \frac{c}{d}$ to contain a path of linear length $f(c)n$ for some function $f(c)$ when $c > 1$, it is natural to conjecture that $f(c) \to 1$ as $c \to \infty$.

Question 5.3. Let $p = \omega\left(\frac{1}{d}\right)$. Is it true that whp Q_p^d contains a path of length $(1 - o(1))n$?

Finally, other notions of random subgraphs of the hypercube Q^d have also been studied. In particular, if we let $Q^d(p)$ denote a random *induced* subgraph of Q^d obtained by retaining each *vertex* independently with probability p, then the typical existence of a giant component in $Q^d(p)$ when $p = \frac{1+\epsilon}{d}$ for a fixed $\epsilon > 0$ was shown by Bollobás, Kohayakawa and Łuczak [8], and this was extended to a broader range of p with $\epsilon = o(1)$ by Reidys [40]. It would be interesting to know if the giant component in $Q^d(p)$ whp also has good expansion properties.

REFERENCES

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

[2] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016.

[3] L. W. Beineke and F. Harary. The genus of the n-cube. *Canadian J. Math.*, 17:494–496, 1965.

[4] B. Bollobás. The evolution of the cube. In *Combinatorial mathematics (Marseille-Luminy, 1981)*, volume 75 of *North-Holland Math. Stud.*, pages 91–97. North-Holland, Amsterdam, 1983.

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

[6] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, second edition, 2001.

[7] B. Bollobás, Y. Kohayakawa, and T. Łuczak. The evolution of random subgraphs of the cube. *Random Structures & Algorithms*, 3(1):55–90, 1992.

[8] B. Bollobás, Y. Kohayakawa, and T. Łuczak. On the evolution of random Boolean functions. In *Extremal problems for finite sets (Visegrád, 1991)*, volume 3 of *Bolyai Soc. Math. Stud.*, pages 137–156. János Bolyai Math. Soc., Budapest, 1994.

[9] B. Bollobás and I. Leader. Matchings and paths in the cube. *Discrete Appl. Math.*, 75(1):1–8, 1997.

[10] B. Bollobás and O. Riordan. A simple branching process approach to the phase transition in $G_{n,p}$. *Electron. J. Combin.*, 19(4):Paper 21, 8, 2012.

[11] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade, and J. Spencer. Random subgraphs of finite graphs. III. The phase transition for the n-cube. *Combinatorica*, 26(4):395–410, 2006.

[12] Ju. D. Burtin. The probability of connectedness of a random subgraph of an n-dimensional cube. *Problemy Peredači Informacii*, 13(2):90–95, 1977.

[13] L. S. Chandran and N. Sivadasan. On the Hadwiger's conjecture for graph products. *Discrete Math.*, 307(2):266–273, 2007.

[14] P. Condon, A. Espuny Díaz, A. Girão, D. Kühn, and D. Osthus. Hamiltonicity of random subgraphs of the hypercube. *arXiv preprint arXiv:2007.02891*, 2020.

[15] C. Dowden, M. Kang, and M. Krivelevich. The genus of the Erdős-Rényi random graph and the fragile genus property. *Random Structures & Algorithms*, 56(1):97–121, 2020.

[16] S. Ehard and F. Joos. Paths and cycles in random subgraphs of graphs with large minimum degree. *Electron. J. Combin.*, 25(2):Paper No. 2.31, 15, 2018.

[17] J. Erde, M. Kang, and M. Krivelevich. Large complete minors in random subgraphs. *Combin. Probab. Comput.*, 2020. https://doi.org/10.1017/S0963548320000607.

[18] P. Erdős and J. Spencer. Evolution of the n-cube. *Comput. Math. Appl.*, 5(1):33–39, 1979.

[19] N. Fountoulakis, D. Kühn, and D. Osthus. Minors in random regular graphs. *Random Structures & Algorithms*, 35(4):444–463, 2009.

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

[21] A. Frieze and M. Krivelevich. On the non-planarity of a random subgraph. *Combin. Probab. Comput.*, 22(5):722–732, 2013.

[22] E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30:1141–1144, 1959.

[23] R. Glebov, H. Naves, and B. Sudakov. The threshold probability for long cycles. *Combin. Probab. Comput.*, 26(2):208–247, 2017.

[24] L. H. Harper. Optimal numberings and isoperimetric problems on graphs. *J. Combinatorial Theory*, 1:385–393, 1966.

[25] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, 56:13–20, 1960.

[26] R. van der Hofstad and A. Nachmias. Unlacing hypercube percolation: a survey. *Metrika*, 77(1):23–50, 2014.

[27] R. van der Hofstad and A. Nachmias. Hypercube percolation. *J. Eur. Math. Soc. (JEMS)*, 19(3):725–814, 2017.

[28] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561, 2006.

[29] T. Hulshof and A. Nachmias. Slightly subcritical hypercube percolation. *Random Structures & Algorithms*, 56(2):557–593, 2020.

[30] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, 2000.

[31] K. Kawarabayashi and B. Reed. A separator theorem in minor-closed classes. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science—FOCS 2010*, pages 153–162. IEEE Computer Soc., Los Alamitos, CA, 2010.

[32] M. Krivelevich. Finding and using expanders in locally sparse graphs. *SIAM J. Discrete Math.*, 32(1):611–623, 2018.

[33] M. Krivelevich. Expanders – how to find them, and what to find in them. In *Surveys in combinatorics 2019*, volume 456 of *London Math. Soc. Lecture Note Ser.*, pages 115–142. Cambridge Univ. Press, Cambridge, 2019.

[34] M. Krivelevich. Long cycles in locally expanding graphs, with applications. *Combinatorica*, 39(1):135–151, 2019.

[35] M. Krivelevich, C. Lee, and B. Sudakov. Long paths and cycles in random subgraphs of graphs with large minimum degree. *Random Structures & Algorithms*, 46(2):320–345, 2015.

[36] M. Krivelevich and A. Nachmias. Coloring complete bipartite graphs from random lists. *Random Structures & Algorithms*, 29(4):436–449, 2006.

[37] M. Krivelevich and W. Samotij. Long paths and cycles in random subgraphs of $\mathcal{H}$-free graphs. *Electron. J. Combin.*, 21(1):Paper 1.30, 18, 2014.

[38] M. Krivelevich and B. Sudakov. The phase transition in random graphs: a simple proof. *Random Structures & Algorithms*, 43(2):131–138, 2013.

[39] G. Moshkovitz and A. Shapira. Decomposing a graph into expanding subgraphs. *Random Structures & Algorithms*, 52(1):158–178, 2018.

[40] C. M. Reidys. Large components in random induced subgraphs of n-cubes. *Discrete Math.*, 309(10):3113–3124, 2009.

[41] G. Ringel. Über drei kombinatorische Probleme am n-dimensionalen Würfel und Würfelgitter. *Abh. Math. Sem. Univ. Hamburg*, 20:10–19, 1955.

[42] O. Riordan. Long cycles in random subgraphs of graphs with large minimum degree. *Random Structures & Algorithms*, 45(4):764–767, 2014.

[43] A. A. Sapoženko. Metric properties of almost all functions of the algebra of logic. *Diskret. Analiz*, (10):91–119, 1967.

Appendix A. Proof of Theorem 3.1

We note that the proof below follows the proof in [1] quite closely. There is a slight error in the original proof, where the vertex-isoperimetric inequality of Harper [24] is used to assert the existence of a large matching between a subset of the hypercube and its complement, however, the claimed size of this matching is too large by a factor of d. This error is easily fixed, and we do so by utilising instead the stronger isoperimetric result given in Theorem 2.7.

Proof of Theorem 3.1. During the proof we will introduce a series of constants $c_1, c_2, \ldots$, with the understanding that each c_i will be as small as is necessary for the argument in terms of δ_1, δ_2 and all c_j with $j < i$. Recall that $p = \frac{1+\delta_1}{d}$ and γ_1 denotes the survival probability of the $\mathrm{Po}(1+\delta_1)$ branching process.

Let us first show that whp $|V(L_1)| \leq (\gamma_1 + \delta_2)n$. To begin with, we note that the number of vertices in the component of Q_p^d containing a fixed vertex v of the hypercube is dominated by a $\mathrm{Bin}(d, p)$ branching process, which tends in distribution to a $\mathrm{Po}(1+\delta_1)$ branching process as $d \to \infty$.

Claim A.1. Whp the number of vertices contained in component of order at most d in Q_p^d is at least $(1 - \gamma_1 - c_1)n$.

Proof of Claim A.1. It follows from standard bounds, see for example [10], that the probability that a $\mathrm{Po}(1+\delta_1)$ branching process dies after growing to size d is $o(1)$, and so the probability that v is contained in a component of order at most d is at least $1 - \gamma_1 - c_2$. If we let Z be the number of vertices contained in components of order at most d in Q_p^d, then it follows that $\mathbb{E}(Z) \geq (1 - \gamma_1 - c_2)n$. However, Z is a $2d$-Lipschitz function of the edges of Q_p^d, that is, by adding or deleting a single edge we can change Z by at most $2d$, and so it follows from a standard application of the Azuma-Hoeffding inequality to the edge-exposure martingale on Q_p^d (see for example [2, Section 7]) that Z is tightly concentrated about its mean, and in particular whp $Z \geq (1 - \gamma_1 - c_1)n$. $\qquad\square$

It follows from Claim A.1 that whp

$$|V(L_1)| \leq n - Z \leq (\gamma_1 + c_1)n \leq (\gamma_1 + \delta_2)n. \tag{5}$$

To continue the argument, we consider a multi-round exposure with $q_1 < p$ chosen such that if $\hat{\gamma}$ is the survival probability of the $\mathrm{Po}(dq_1)$ branching process then $|\gamma_1 - \hat{\gamma}| \le c_3$. Note that we then may assume that we can choose $q_2, q_3 \ge \frac{c_4}{d}$ such that

$$(1 - q_1)(1 - q_2)(1 - q_3) = 1 - p.$$

Let us write $Q_1 = Q_{q_1}^d$, $Q_2 = Q_1 \cup Q_{q_2}^d$ and $Q_3 = Q_2 \cup Q_{q_3}^d$, where the random subgraphs are chosen independently, so that Q_3 has the same distribution as Q_p^d.

By a similar comparison to a branching process as above, we see that the probability that a fixed vertex v is contained in a component of order at least $c_7 d$ in Q_1 is at least $\hat{\gamma} - c_6$. In particular, we expect around $\hat{\gamma}d$ of the neighbours of any fixed vertex to be contained in components of Q_1 of order $\Omega(d)$. Over the next few claims, we will show that we can 'boost' this result to show that almost every vertex has around $\hat{\gamma}d$ neighbours contained in components of Q_2 of order $\Omega(d^2)$.

Claim A.2. Whp all but n^{1-c_8} of the vertices v in Q^d have at least $(\hat{\gamma} - c_5)d$ many neighbours in Q^d which are contained in a component of order at least $c_7 d$ in Q_1. Let us say in this case that v has property (a).

Proof of Claim A.2. Indeed, each neighbour of v in Q^d has probability at least $\hat{\gamma} - c_6$ of being contained in a component of order at least $c_7 d$ in Q_1. Furthermore, since this property is an increasing property of subgraphs of Q^d, by the inequality of Harris (Lemma 2.6), given any subset W of the neighbours of v of size $1 \le i \le d$ the probability that each $w \in W$ is contained in a component of order at least $c_7 d$ in Q_1 is at least $(\hat{\gamma} - c_6)^i$. It follows that the number of neighbours of v which are contained in a component of order at least $c_7 d$ in Q_1 stochastically dominates a $\mathrm{Bin}(d, \hat{\gamma} - c_6)$ random variable. Hence, by Chernoff type bounds (Lemma 2.4 (i)) the probability that v does not have at least $d(\hat{\gamma} - c_5)$ many neighbours in Q^d which are contained in a component of order at least $c_7 d$ in Q_1 is at most n^{-2c_8}, and so v does not have property (a) with probability at most n^{-2c_8}. Hence by Markov's inequality whp the total number of vertices without property (a) is at most n^{1-c_8}. $\square$

Claim A.3. Whp for all but $n^{1-c_{11}}$ of the vertices v in Q^d we can find a collection of at least $c_{10} d$ many disjoint connected subgraphs in Q_1 each of order at least $c_{12} d$ which are adjacent to v in Q^d. Let us say in this case that v has property (b).

Proof of Claim A.3. Indeed, consider without loss of generality the case where v is the origin. We can fix $c_9 d$ disjoint subcubes of dimension $(1 - c_9)d$, each of which contains a vertex adjacent to v, by taking for each $1 \le i \le c_9 d$ the vertex w_i with a one in the ith coordinate and zeros elsewhere, and taking the subcube $Q(i)$ containing w_i, where we vary the last $(1 - c_9)d$ coordinates. Since $c_9 \ll \delta_1$, again by comparison to an appropriate branching process, for each w_i with probability at least $\frac{\hat{\gamma}}{2}$ the component of $Q(i)_{q_1}$ containing w_i has order at least $c_{12} d$. Since the subcubes $Q(i)$ are disjoint, the probability that less than $c_{10} d \le \frac{\hat{\gamma} c_9 d}{4}$ of these components have order at least $c_{12} d$ is at most $n^{-2c_{11}}$ by Lemma 2.4. Therefore, v does not have property (b) with probability at most $n^{-2c_{11}}$. Hence by Markov's inequality whp the total number of vertices without property (b) is at most $n^{1-c_{11}}$. $\square$

Claim A.4. Whp all but $n^{1-c_{13}}$ of the vertices v in Q^d have $(\hat{\gamma} - c_5)d$ many neighbours in Q^d, each of which is contained in some subgraph in Q_1 of order at least $c_{12} d$, where each vertex in this subgraph has property (b). Let us say in this case that v has property (c).

Proof of Claim A.4. By Claim A.2 whp at most n^{1-c_8} vertices of Q^d do not have property (a). Furthermore, by Claim A.3 whp at most $n^{1-c_{11}}$ of the vertices of Q^d do not have property (b),

and so whp the number of vertices at distance at most $c_{12}d$ from a vertex without property (b) is at most

$$n^{1-c_{11}} \sum_{i \le c_{12}d} \binom{d}{i} \le c_{12}dn^{1-c_{11}} \left(\frac{e}{c_{12}}\right)^{c_{12}d} \le n^{1-2c_{13}}$$

since $c_{12} \ll c_{11}$. However, clearly any vertex without property (c) must either not have property (a), or be within distance $c_{12}d$ from a vertex without property (b). Hence, whp the number of vertices without property (c) is at most $n^{1-2c_{13}} + n^{1-c_8} \le n^{1-c_{13}}$. $\qquad \square$

We now use our first sprinkling step to merge most of these disjoint connected subgraphs into ones of quadratic order.

Claim A.5. Whp in $Q_2 = Q_1 \cup Q_{q_2}^d$ all but $n^{1-c_{16}}$ of the vertices v in Q^d are adjacent in Q^d to at least $(\hat{\gamma} - c_5)d$ many vertices lying in components of order at least $c_{15}d^2$ in Q_2. Let us say in this case that v has property (d).

Proof of Claim A.5. After exposing Q_1, by Claim A.4 whp the number of vertices without property (c) is at most $n^{1-c_{13}}$. For each vertex v with property (c), let us fix a family S_v of $(\hat{\gamma} - c_5)d$ subgraphs of Q_1 of order $c_{12}d$ witnessing that v has property (c). Similarly for each vertex v with property (b), let us fix a family T_v of $c_{10}d$ disjoint connected subgraphs of Q_1 of order $c_{12}d$ witnessing that v has property (b).

Consider a fixed subgraph $H \in S_v$ for some v with vertex set $C = V(H)$, and let $C' \subseteq C$ be a set of vertices with the same parity of size $|C|/2$. Let $C' = \{v_1, v_2, \ldots\}$. We will grow a vertex set A containing C which induces a connected subgraph of Q_2 inductively until it reaches an appropriate size. Let us set initially $A(0) = C$.

If $|A(i-1)| \ge c_{15}d^2$ we let $A(i) = A(i-1)$. Otherwise, since v has property (c), each v_i has property (b), and so we can consider the family T_{v_i} of disjoint subgraphs of Q_1 which are adjacent to v_i in Q^d. Since $|A(i-1)| < c_{15}d^2$ we may assume that at most $\frac{c_{10}}{2}d$ of the subgraphs in T_{v_i} meet $A(i-1)$ in more than $\frac{c_{12}}{2}d$ vertices. Let T'_{v_i} be a subfamily of $\frac{c_{10}}{2}d$ subgraphs, each of which meets $A(i-1)$ in at most $\frac{c_{12}}{2}d$ vertices.

We expose the edges between v_i and each subgraph in T'_{v_i} in $Q_{p_2}^d$, and we let $A(i)$ be $A(i-1)$ together with the vertex sets of all the subgraphs in T'_{v_i} which are connected to $A(i-1)$ in this manner. Let A be the final subgraph built in this manner.

We say that stage i is successful if $|A(i-1)| \ge c_{15}d^2$ or at least one of the exposed edges is present in $Q_{p_2}^d$. Note that the probability that a round is successful is at least

$$1 - \left(1 - \frac{c_4}{d}\right)^{\frac{c_{10}}{2}d} \ge c_{14}.$$

Let r be the number of rounds which are successful. Then $|A| \ge \min\{r\frac{c_{12}}{2}d, c_{15}d^2\}$.

Note that, since the parity of each $v_i \in C'$ is the same, and we only expose edges adjacent to v_i in stage i, the outcome of each stage is independent. Hence, r stochastically dominates a $\text{Bin}(|C'|, c_{14})$ random variable, whose expectation is $\frac{c_{12}c_{14}}{2}d$. Hence, it follows again from Lemma 2.4 (i) that the probability that $|A| \le c_{15}d^2$ is at most $n^{-4c_{16}}$.

Hence, we have shown that for any fixed subgraph $C \in S_v$, the probability that C is not contained in a component of Q_2 of order at least $c_{15}d^2$ is at most $n^{-4c_{16}}$. It follows that the expected number of subgraphs C such that C is not contained in a component of Q_2 of order at least $c_{15}d^2$ and $C \in S_v$ for some v is at most $(\hat{\gamma} - c_5)dn^{1-4c_{16}} \le n^{1-3c_{16}}$. As before, it follows from Markov's inequality that whp the number of such subgraphs is at most $n^{1-2c_{16}}$.

In particular, whp the number of vertices with property (c) but not (d) is at most $n^{1-2dc_{16}}$, and so the number of vertices without property (d) is at most $n^{1-c_{16}}$. $\qquad \square$

In particular, since it is an increasing property and $Q_2 \subseteq Q_3$, it follows from Claim A.5 that property (iii) of Theorem 3.1 holds in Q_3, with $\delta_3 = c_{15} \ll \delta_2$, $\delta_4 = (\hat{\gamma} - c_5)$ and $\delta_5 = c_{16}$. Furthermore, since each vertex with property (d) is adjacent to at least $(\hat{\gamma} - c_5)d$ many vertices contained in components of order at least $c_{15}d^2$, it follows by double counting that the number of vertices contained in components of order at least $c_{15}d^2$ in Q_2, and hence in Q_3, is at least

$$\frac{(n - n^{1-c_{16}})(\hat{\gamma} - c_5)d}{d} \geq (\hat{\gamma} - 2c_5)n. \tag{6}$$

Next, we show that whp almost all the vertices in large components in Q_2 belong to the largest component of Q_3. Hence, we will be able to deduce properties (i) and (ii) of Theorem 3.1 from (6), together with (5).

So, to that end, let us denote by X the set of vertices without property (d) and let Y be the set of vertices contained in components of order at least $c_{15}d^2$ in Q_2. Note that, by definition, every vertex not in X is adjacent to at least $(\hat{\gamma} - c_5)d$ many vertices in Y.

Claim A.6. Whp for every partition of $Y = A \cup B$ with $\min\{|A|, |B|\} \geq c_{17}n$ there is an edge between A and B in $Q_3 = Q_2 \cup Q_{q_3}^d$.

Proof of Claim A.6. Indeed, if there is a component C of order at least $c_{15}d^2$ in Q_2 such that $C \cap A$ and $C \cap B$ are both non-empty, then there is a path between A and B in $Q_2 \subseteq Q_3$. Hence, we may restrict ourselves to partitions of Y which partition the components of Q_2 of order at least $c_{15}d^2$. Since there are at most $\frac{n}{c_{15}d^2}$ many such components in Q_2, it follows that the total number of possible partitions to consider is at most $2^{\frac{n}{c_{15}d^2}}$. We will show that the probability that there is no path across such a partition is at most $\exp\left(-\frac{n}{c_{15}d^2}\right)$, and hence by the union bound whp the claim holds.

To this end, let us consider the inclusive neighbourhoods $N(A)$ and $N(B)$ of A and B respectively in Q_2, and let $D = N(A) \cap N(B)$. We split into two cases. Let us suppose first that $|D| \geq \frac{4n}{c_{19}d}$.

In this case, let $D' = D \setminus X$, so that $|D'| \geq \frac{2n}{c_{19}d}$, and let $D'' \subset D'$ be a set of at least $\frac{|D'|}{2} \geq \frac{n}{c_{19}d}$ vertices of the same parity.

For each vertex $x \in D''$ we claim that the probability that there is a path from A to B in $Q_{q_3}^d$ of length at most two which contains x is at least $\frac{c_{18}}{d}$. Indeed, suppose that $x \notin A \cup B$, then since $x \in D$ it has at least one neighbour in both A and B, and since $x \notin X$ it has at least $\frac{(\hat{\gamma} - c_5)}{2}d$ neighbours in one of A or B. Hence, such a path exists with probability at least

$$q_3 \left(1 - (1 - q_3)^{\frac{(\hat{\gamma} - c_5)}{2}d}\right) \geq \frac{c_{18}}{d}.$$

The case where $x \in A$ or B is similar.

However, since these events depend on pairwise disjoint edge sets, and hence are independent for different x, it follows that the probability that no $x \in D''$ joins A to B is at most

$$\left(1 - \frac{c_{18}}{d}\right)^{|D''|} \leq \left(1 - \frac{c_{18}}{d}\right)^{\frac{n}{c_{19}d}} \leq \exp\left(-\frac{n}{c_{15}d^2}\right).$$

In the second case, we may assume that $|D| < \frac{4n}{c_{19}d}$. In this case, we note that $|N(A)| \geq |A| \geq c_{17}n$ and $|N(A)^c| \geq |B| - |D| \geq \frac{c_{17}}{2}n$. Hence, by Theorem 2.7 there is a matching F of size at least $\frac{2c_{20}n}{\sqrt{d}}$ between $N(A)$ and $N(A)^c$. If we let $F' \subseteq F$ be those edges which do not have an endpoint in D or X, then $|F'| \geq \frac{c_{20}n}{\sqrt{d}}$.

For each edge $uv = e$ in F', where without loss of generality $u \in N(A) \setminus (D \cup B)$, either $u \in A$ or u has at least $(\hat{\gamma} - c_5)d$ neighbours in A and similarly either $v \in B$ or v has at least $(\hat{\gamma} - c_5)d$ many neighbours in B.

Hence, we can fix as before events depending on pairwise disjoint edge sets for each $e \in F'$ which imply that there is a path between A and B of length at most 3 containing the edge e. Each of these events happens with probability at least

$$q_3 \left(1 - (1 - q_3)^{(\hat{\gamma} - c_5)d}\right)^2 \geq \frac{c_{18}}{d}$$

and so, as before, the probability that no $e \in F'$ joins A to B is at most

$$\left(1 - \frac{c_{18}}{d}\right)^{|F'|} \leq \left(1 - \frac{c_{18}}{d}\right)^{\frac{c_{20}n}{\sqrt{d}}} \leq \exp\left(-\frac{c_{18}c_{20}n}{d^{\frac{3}{2}}}\right) \leq \exp\left(-\frac{n}{c_{15}d^2}\right).$$

$\square$

Hence, by Claim A.6 whp there is some component L_1 of Q_3 which contains all but at most $c_{17}n$ vertices of Y, that is $|Y \setminus V(L_1)| \leq c_{17}n$, and hence by (6)

$$|V(L_1)| \geq (\hat{\gamma} - 2c_5 - c_{17})n \geq (\gamma_1 - c_3 - 2c_5 - c_{17})n \geq (\gamma_1 - c_1)n.$$

However, by (5) we have $|V(L_1)| \leq (\gamma_1 + c_1 n)$, and so $|V(L_1)| = (\gamma_1 \pm c_1)n = (\gamma_1 \pm \delta_2)n$. Furthermore, by Claim A.1 whp at least $(1 - \gamma_1 - c_1)n$ vertices of Q_p^d are contained in components of order at most d, and hence whp every other component of $Q_3 = Q_p^d$ has order at most $\max\{2c_1 n, d\} \leq \delta_2 n$. Finally, it is clear that $|V(L_1) \triangle Y| \leq 2c_1 n \leq \delta_2 n$ also, and so properties (i) and (ii) of Theorem 3.1 hold. $\square$