A proof of the Kim-Vu sandwich conjecture

Natalie Behague* Daniel Il'kovič*† Richard Montgomery*

November 3, 2025

Abstract

In 2004, Kim and Vu conjectured that, when $d = \omega(\log n)$, the random d-regular graph $G_d(n)$ can be sandwiched with high probability between two random binomial graphs $G(n,p)$ with edge probabilities asymptotically equal to $\frac{d}{n}$. That is, there should exist $p_* = (1-o(1))\frac{d}{n}$, $p^* = (1+o(1))\frac{d}{n}$ and a coupling (G_*, G, G^*) such that $G_* \sim G(n,p_*)$, $G \sim G_d(n)$, $G^* \sim G(n,p^*)$, and $\mathbb{P}(G_* \subset G \subset G^*) = 1-o(1)$. Known as the *sandwich conjecture*, such a coupling is desirable as it would allow properties of the random regular graph to be inferred from those of the more easily studied binomial random graph. The conjecture was recently shown to be true when $d \gg \log^4 n$ by Gao, Isaev and McKay. In this paper, we prove the sandwich conjecture in full. We do so by analysing a natural coupling procedure introduced in earlier work by Gao, Isaev and McKay, which had only previously been done when $d \gg n/\sqrt{\log n}$.

1 Introduction

Random graphs have been a central object of study in Combinatorics since the foundational work of Erdős and Rényi [10] in 1959. The binomial random graph, or Erdős-Rényi random graph, $G(n,p)$ has vertex set $[n] = \{1,\dots,n\}$ and each potential edge included independently at random with probability p. Along with the closely related uniformly random graph with n vertices and $\lceil p\binom{n}{2}\rceil$ edges, $G(n,p)$ is the most studied random graph model. The next most studied model is probably that of the random regular graph $G_d(n)$, which is chosen uniformly at random from all d-regular graphs with vertex set $[n]$. Throughout this paper, and as is common, we will implicitly assume that dn is even, so that the set of such graphs is non-empty.

The study of random regular graphs began in earnest in the late 1970's, with early work including that by Bender and Canfield [1, 2], Bollobás [3], and Wormald [31, 32]. Compared to $G(n,p)$, where the independence of the edges allows the use of a wide variety of techniques, studying $G_d(n)$ is often much more difficult. For example, from developments [5, 22] of the breakthrough work of Pósa [27] in 1976, it has been long understood when we may expect $G(n,p)$ to be Hamiltonian. On the other hand, it was widely anticipated that, for each $3 \le d \le n-1$, $G_d(n)$ should be Hamiltonian with high probability (i.e., with probability $1-o(1)$), but proving this took the combined work of many authors [4, 6, 7, 23, 28, 29] over the course of 20 years (see the surveys [11, 33] for more details), using a variety of tools in different regimes for d, from the configuration model [4], through switching methods [25], to estimates on the number of regular graphs [26, 30].

*Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK. natalie.behague@warwick.ac.uk, daniel.ilkovic@warwick.ac.uk, and richard.montgomery@warwick.ac.uk.

†Institute of Mathematics, Leipzig University, Augustusplatz 10, 04109 Leipzig, Germany.

NB and RM were supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No. 947978).
DI was supported by the Alexander von Humboldt Foundation in the framework of the Alexander von Humboldt Professorship of Daniel Král' endowed by the Federal Ministry of Education and Research.

This increased difficulty and technicality gives great appeal to finding links from $G_d(n)$ to $G(n,p)$, so that we may hopefully deduce properties of $G_d(n)$ from those known for $G(n,p)$. In 2004, Kim and Vu [20] formalised this desire in their famous 'sandwich conjecture'. They conjectured that, if $d = \omega(\log n)$, then there are $p_*, p^* = (1+o(1))d/n$ and a coupling (G_*, G, G^*) such that $G_* \sim G(n,p_*)$, $G \sim G_d(n)$, $G^* \sim G(n,p^*)$, and $\mathbb{P}(G_* \subset G \subset G^*) = 1-o(1)$. If true, the requirement $d = \omega(\log n)$ cannot be removed; as is well-known, for each $C > 0$ there is some $\varepsilon > 0$ such that if $d = C\log n$ then, with probability $1 - o(n^{-1})$, $G(n,(1+\varepsilon)d/n)$ has minimum degree less than d, and hence contains no d-regular subgraph.

Kim and Vu [20] proved the lower part of their conjecture when $\log n \ll d \ll n^{1/3}/\log^2 n$, and a weakened upper part for the same range of d. More specifically, with high probability their coupling (G_*, G, G^*) satisfied $G_* \subset G$ and $\Delta(G \setminus G^*) \leq (1+o(1))\log n$. In 2017, Dudek, Frieze, Ruciński and Šileikis [9] extended this by showing that the lower part holds when $d = o(n)$, and that it can be generalised to hypergraphs. Subsequently, a major breakthrough was made by Gao, Isaev and McKay [14, 15], who gave the first coupling (G_*, G, G^*) without a weakened upper part, that is, in which $\mathbb{P}(G_* \subset G \subset G^*) = 1 - o(1)$. This allowed them to show that the sandwich conjecture is true if $\min\{d, n-d\} \gg n/\sqrt{\log n}$. Klimošová, Reiher, Ruciński and Šileikis [21] then extended this to show that it holds for $\min\{d, n-d\} \gg (n\log n)^{3/4}$ (while generalising it to biregular graphs). Very significant progress was then made again by Gao, Isaev and McKay [13], who showed that the conjecture is true provided that $d \gg \log^4 n$.

Here, we will prove the sandwich conjecture in full, in the following very slightly stronger form.

Theorem 1.1. *For each $\varepsilon > 0$ there is some $C > 0$ such that the following holds for each $d \geq C\log n$. There is some $(1-\varepsilon)d/n \leq p_*, p^* \leq (1+\varepsilon)d/n$ and a coupling (G_*, G, G^*) of random graphs such that $G_* \sim G(n,p_*)$, $G \sim G_d(n)$, $G^* \sim G(n,p^*)$, and $\mathbb{P}(G_* \subset G \subset G^*) = 1 - o(1)$.*

In their breakthrough initial work, Gao, Isaev and McKay [14, 15] introduced a beautiful natural coupling process, and analysed it in the regime $\min\{d, n-d\} \gg n/\sqrt{\log n}$. In order to prove Theorem 1.1, we analyse this process throughout the whole regime $d \geq C\log n$. As highlighted by Gao, Isaev and McKay [15, Question 2.4], key to this analysis is showing that in certain random graphs F a uniformly chosen d-regular subgraph is likely to include any given edge in F with roughly equal probability, and our success in doing so is of some independent interest (see Theorem 2.4). It allows us to avoid the complications of running two versions of this process, one after another, as was done in the later work of Gao, Isaev and McKay [13]. While this two-process coupling ingeniously avoided the most significant challenge in analysing a coupling generated by the single process, it encounters very significant barriers when d is below $\log^4 n$.

We outline our methods in context with previous work in Section 2, but for now make some brief remarks. As in [13, 14, 15], we work via switching techniques, requiring the counting of certain 'switching paths'. The accuracy we use for our bounds has to be carefully chosen, particularly in the most critical regime, and our analysis is completely separate to previous work. Very broadly, our improvements come in two main areas. Firstly, we work much more directly with random graphs than [13], which used pseudorandom properties from the spectral properties of random graphs. Secondly, our bounds on the number of switching paths do not hold for every d-regular subgraph of our random graph (as they do in [13]), but typically hold for sufficiently many for our proof. This allows us to pivot in a slight but very useful way, considering instead 'switching paths' between two random graphs rather than between a random graph and any one of its d-regular subgraphs.

While, due to [13], the previously open regime for Theorem 1.1 is the upper part of the sandwich for $d = O(\log^4 n)$, our methods here naturally work as long as $d \leq n^\sigma$ for some fixed $\sigma > 0$. With some brief additional work (see Section 8.1), we use our methods to give the upper part of the sandwich for the whole range $d \geq C\log n$. This gives a more straightforward proof of the sandwich

conjecture in the whole regime as, for example, it avoids the complex analytic techniques used by Gao, Isaev and McKay [15, Section 7] when $d = \Omega(n)$. For completion, and as our methods give a slightly alternative approach, we also include a full proof of the lower part of the sandwich (see Section 10). Altogether, then, we give a self-contained proof of Theorem 1.1 which only relies on some quoted results on the concentration of certain random variables (see Section 2.3).

It follows immediately from Theorem 1.1 that, if $d_2 \geq d_1 = \omega(\log n)$ and $d_2 - d_1 = \Omega(d_1)$, and $d_1 n$ and $d_2 n$ are even, then there is a coupling (G_1, G_2) with $G_1 \sim G_{d_1}(n)$, $G_2 \sim G_{d_2}(n)$, and $\mathbb{P}(G_1 \subset G_2) = 1 - o(1)$. (This naturally improves the same conclusion when $d \gg \log^4 n$ by Gao, Isaev and McKay [13].) That this monotonic coupling property for $G(n, d)$ should be possible for $d_1 \leq d_2$ more widely was the subject of some discussion in the community (see [15]), but has appeared only recently in print, in the following form.

Conjecture 1.2 (Gao, Isaev and McKay [15])**.** *Let $d_1, d_2 \in [n-1]$ with $d_1 \leq d_2$ be such that $d_1 n$ and $d_2 n$ are even and $(d_1, d_2) \notin \{(1,2), (n-3, n-2)\}$. Then, there is a coupling (G_1, G_2) of random graphs such that $G_1 \sim G_{d_1}(n)$, $G_2 \sim G_{d_2}(n)$, and, with probability $1 - o(1)$, $G_1 \subset G_2$.*

While the current progress on Conjecture 1.2 can be found outlined in [16], the most interesting cases are perhaps where n is even and $d_2 = d_1 + 1$ (so that good bounds on the probabilities involved may imply the conjecture more widely). Here, the conjecture is known as long as $d_1 = \omega(1)$ and $d_1 \leq n^{1/7}/\log n$, due to work by Gao [12] and then Hollom, Lichev, Mond, Portier and Wang [16].

A version of the sandwich conjecture has also been proposed for graphs chosen uniformly at random from those with vertex set $[n]$ and degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$, by Gao, Isaev and McKay (see [15, Conjecture 1.4]), to generalise the sandwich conjecture beyond the sequences $\mathbf{d} = (d, \ldots, d)$. They conjectured that the corresponding coupling should be possible for sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $d_i = (1 + o(1))d$ for each $i \in [n]$ under the same condition $d = \omega(\log n)$. The results in [13, 14, 15] were shown in this wider generality (sometimes with a stronger bound needed on the distribution of the d_i). Our methods are likely to be able to make further progress here, but to avoid further notational and technical cost we will prove only Theorem 1.1.

As well as this potential generalisation, it is also natural to ask for which functions $\varepsilon = \varepsilon(n)$ and $d = d(n)$ the coupling in Theorem 1.1 exists. Broadly, it seems we should anticipate that the coupling is likely to be possible if, with high probability, $\Delta(G(n, (1-\varepsilon)d/n)) \leq d \leq \delta(G(n, (1+\varepsilon)d/n))$. In this direction, we find the following conjecture appealing, for which we recall that the n-vertex *random graph process* $G_0, G_1, \ldots, G_{\binom{n}{2}}$ begins with the graph G_0 with vertex set $[n]$ and no edges and, for each $i \geq 1$, G_i is formed from G_{i-1} by the addition of an edge selected uniformly and independently at random from $[n]^{(2)} \setminus E(G_{i-1})$.

Conjecture 1.3. *Let $G_0, G_1, \ldots, G_{\binom{n}{2}}$ be the n-vertex random graph process and let $1 \leq d \leq n-1$ be such that dn is even. Then, there is some $G \sim G_d(n)$ for which the following holds with high probability. For each $0 \leq i \leq \binom{n}{2}$, if $\Delta(G_i) \leq d$ then $G_i \subset G$, and if $\delta(G_i) \geq d$ then $G \subset G_i$.*

If true, proving Conjecture 1.3 seems to fundamentally require additional ideas rather than simply a refinement of the techniques presented here. In the next section, we will give some brief details of our notation before discussing in detail our proof and its relation to previous methods. The rest of the paper is then outlined in detail, where the proof of the upper part of Theorem 1.1 appears in Sections 3 to 9 and the proof of the lower part appears in Section 10.

Acknowledgements

The authors would like to thank participants of the workshop 'Combinatorics, Probability and Algorithms' at the Mathematisches Forschungsinstitut Oberwolfach in September 2025 for conversations that led to Conjecture 1.3 and inspired a simplification of the proof of Theorem 1.1.

2 Preliminaries and proof discussion

In this section we will cover some of our notation, before discussing the proof of Theorem 1.1 in detail and giving a brief outline of the paper.

2.1 Notation

A graph G has vertex set $V(G)$ and edge set $E(G)$, and we let $|G| = |V(G)|$ and $e(G) = |E(G)|$. For any two vertex sets U, V of a graph K, we use $e_K(U, V) = |\{(u, v) : u \in U, v \in V, uv \in E(K)\}|$, and the set $N_G(U)$ is the *neighbourhood of* U *in* G, that is, the set of vertices in $V(G) \setminus U$ with at least one neighbouring edge to U. Given a graph G, $v \in V(G)$ and $e \in V(G)^{(2)}$, $G - v$ is the graph $G[V(G) \setminus \{v\}]$, while $G - e$ and $G + e$ are, respectively, the graphs with vertex set $V(G)$ and edge set $E(G) \setminus \{e\}$ and $E(G) \cup \{e\}$. We extend such notation in the usual manner to, for example, $G - A$ and $G + E$ where $A \subset V(G)$ and $E \subset V(G)^{(2)}$. Given graphs G and H, $G \setminus H$ is the graph with vertex set $V(G) \setminus V(H)$ and edge set $E(G) \setminus E(H)$.

All logarithms are natural. We use 'hierarchy' notation to manage the constants we use, where $\alpha \ll \beta$ means that there is some non-negative decreasing function f such that what follows will hold for all $\alpha \leq f(\beta)$. For hierarchies with more than 2 variables these functions are chosen from right to left. For any $a, b, c \in \mathbb{R}$, we say $a = b \pm c$ if $b - c \leq a \leq b + c$.

For each $d, n \in \mathbb{N}$, $\mathcal{K}_d(n)$ is the set of d-regular graphs with vertex set $[n]$, and, for each graph F with $V(F) = [n]$, $\mathcal{K}_d(F) = \{K \in \mathcal{K}_d(n) : K \subset F\}$. For distinct vertices x, y, an x, y-path is a path with endvertices x and y. Though our paths are undirected, we implicitly assume the path starts at x and ends at y. We say the path is (F, K)-*alternating* for graphs F and K if, starting from x, the edges are alternately in F and then K. For each $n \in \mathbb{N}$, we use K_n for the complete graph with vertex set $[n]$.

2.2 Proof discussion

Let $\varepsilon > 0$ be constant and, for convenience in this discussion, take the slightly stronger condition that $d = \omega(\log n)$. To prove Theorem 1.1 we will create the lower part of sandwich (i.e., (G_*, G) with the required properties) using an edge addition process and the upper part of the sandwich (i.e., (G, G^*)) using an edge removal process. These processes are essentially those used by Gao, Isaev and McKay [14, 15] for their first result on the sandwich conjecture, and the main difference here is in how we analyse it. The first process is (roughly) the complement of the other, but the analysis is separate so we discuss them one after another in Section 2.2.1 below, starting with the process for the lower part of the sandwich. Both processes have a key result, Lemma 2.2 and Theorem 2.4 respectively, that shows that the critical property we will need to construct the coupling is likely to hold. These are understandable out of context, using only Definitions 2.1 and 2.3. We will prove these key results using switching methods, as discussed in Section 2.2.2. The rest of the paper is briefly outlined in Section 2.2.3.

2.2.1 Coupling processes

The lower part of the sandwich. Start with F as the graph with vertex set $[n]$ and no edges. Iteratively, as long as F is not a d-regular graph, pick a random edge e from $[n]^{(2)}$ and if it is not in $E(F)$ then add it to F with probability

$$\frac{|\{K \in \mathcal{K}_d(n) : F + e \subset K\}|}{\max_{f \in [n]^2 \setminus E(F)} |\{K \in \mathcal{K}_d(n) : F + f \subset K\}|}, \tag{1}$$

where $\mathcal{K}_d(n)$ is the set of d-regular graphs with vertex set $[n]$. Observe that here an edge e is only added to F if F is not d-regular but there is some d-regular graph containing $F+e$; thus the denominator at (1) is always non-zero. Furthermore, as long as the process runs there is some edge e in $[n]^{(2)} \setminus E(F)$ which is added to F with strictly positive probability if it is chosen, so the process stops with probability 1. Let G be the graph produced at the end of the process, which is then a d-regular graph with vertex set $[n]$. The probability at (1) is chosen so that $G \sim G_d(n)$ (as we will later show in Section 10.3).

While doing this, we will simultaneously build a random graph G_*, starting similarly with the graph with vertex set $[n]$ and no edges. When we consider the random edge e, we also consider whether to add the random edge e to G_* or not; when $e \notin E(G_*)$ we do this with the fixed probability $1-\eta$, where $0 < \eta \ll \varepsilon$. Then, if e is not in F when it is considered and the probability at (1) is at least $1-\eta$, we can couple the random addition of e so that e is always added to F if it is added to G_*. As discussed below, with high probability, for most of this process, the probability at (1) is very close to 1. Thus, if we build such a random graph G_* and stop when it has $\lceil (1-2\eta)dn/2 \rceil$ edges, then with high probability we will have $G_* \subset F$.

Now, observe that by symmetry $\mathbb{P}(G_* = H)$ is the same for any graph H with $V(H) = [n]$ and $e(H) = \lceil (1-2\eta)dn/2 \rceil$. Therefore, G_* has the distribution of a uniformly random graph with vertex set $[n]$ and $\lceil (1-2\eta)dn/2 \rceil$ edges. As $e(G(n,(1-\varepsilon)d/n)) < \lceil (1-2\eta)dn/2 \rceil$ with high probability, we can use a standard coupling of $G(n,(1-\varepsilon)d/n)$ (so that it is within G_* with high probability), to get the exact coupling that we want.

As observed and proved by Gao, Isaev and McKay [13], for each $0 \le i \le dn/2$, conditioned on having i edges, the distribution of F is simple to state. For this, then, for each $0 \le i \le dn/2$ let F_i be the value of F in this process when it has i edges. We use the following definition.

Definition 2.1. Let $d, n, m \in \mathbb{N}$ with $0 \le d \le n-1$, dn even, and $0 \le m \le dn/2$. We say $F \sim F^-(n,d,m)$ if F is a random graph with vertex set $[n]$ which has the same distribution as $H - E$ if $H \sim G_d(n)$ and E is a uniformly random set of m edges from $E(H)$.

Then, for each $0 \le i \le dn/2$, $F_i \sim F^-(n,d,(dn/2)-i)$. We include a proof of this (along the lines of [13]) in Section 10.3. Given this, and the discussion above, to show that the probability at (1) is likely above $1-\eta$ until F has at least $\lfloor (1-\eta)dn/2 \rfloor$ edges (which we can then expect is later than when G_* reaches $\lceil (1-2\eta)dn/2 \rceil$ edges), it will suffice to prove the following result.

Lemma 2.2. *Let $1/n \ll 1/C \ll \eta, \varepsilon \ll 1$. Let $d \ge C \log n$ and $\varepsilon dn \le m \le dn/2$. Let $F \sim F^-(n,d,m)$. Then, with probability $1-o(n^{-2})$, for each $e, f \notin E(F)$, we have*

$$|\{K \in \mathcal{K}_d(n) : F+e \subset K\}| \le (1+\eta)\,|\{K \in \mathcal{K}_d(n) : F+f \subset K\}|. \tag{2}$$

We prove Lemma 2.2 using switching methods (as discussed in Section 2.2.2 and carried out in Section 10.6 using the work more generally in Section 10).

The upper part of the sandwich. The following process is the natural complement of the above process producing F. Start with F as the graph with vertex set $[n]$ and edge set $[n]^{(2)}$. Iteratively, as long as F is not a d-regular graph, pick a random edge e from $[n]^{(2)}$ and if it is in $E(F)$ then remove it from F with probability

$$\frac{|\{K \in \mathcal{K}_d(n) : K \subset F-e\}|}{\max_{f \in E(F)} |\{K \in \mathcal{K}_d(n) : K \subset F-f\}|} = \frac{|\mathcal{K}_d(F-e)|}{\max_{f \in E(F)} |\mathcal{K}_d(F-f)|}, \tag{3}$$

where $\mathcal{K}_d(F-e)$ is the set of d-regular subgraphs of $F-e$ with vertex set $[n]$. Observe that here an edge e is only removed from F if there is some d-regular graph in $F-e$; thus the denominator

at (3) is always non-zero. Furthermore, as before, the process stops with probability 1 with a d-regular graph, G say. The probability at (3) is chosen so that $G \sim G_d(n)$ (as we will later show in Section 3.4).

As before, we will want to argue that the probability at (3) is, with high probability, close to 1 throughout much of the process (which will allow us, similarly, to couple it with a binomial random graph, but one likely to contain F). However, a bound of the form $1-\eta$, with $\eta > 0$ fixed, will not be good enough here as, instead of adding $dn/2$ edges, the process will remove $\binom{n}{2} - dn/2$ edges to get a d-regular graph. This is potentially far many more iterations, which would allow our error terms to accumulate ruinously. We will have to use then a bound in place of $1-\eta$ that depends on the number of edges removed so far.

So much to the negative. In our favour, if we show a natural bound similar to Lemma 2.2 then it gives a better bound when translated into (3). That is, if we can show that, for each $e, f \in E(F)$,

$$|\{K \in \mathcal{K}_d(F) : e \in E(K)\}| \leq (1+\eta)|\{K \in \mathcal{K}_d(F) : f \in E(K)\}|, \tag{4}$$

we can deduce (as we do in Section 3.5.4) that

$$\frac{|\mathcal{K}_d(F-e)|}{\max_{f\in E(F)} |\mathcal{K}_d(F-f)|} = \frac{|\mathcal{K}_d(F)| - |\{K \in \mathcal{K}_d(F) : e \in E(K)\}|}{\max_{f\in E(F)}(|\mathcal{K}_d(F)| - |\{K \in \mathcal{K}_d(F) : f \in E(K)\}|)} \geq 1 - \frac{\eta dn/2}{e(F) - dn/2}. \tag{5}$$

Thus, we will be able to show from a bound like (4) that the probability at (3) will be very close to 1, and only when F has close to $dn/2$ edges (as it nears a d-regular graph) will our bound begin to resemble $1 - O(\eta)$.

This, however, is not good enough. Our 'error term' η in (4) will need to change during the process. For some C_0, μ with $1/n \ll 1/C_0 \ll \mu \ll \varepsilon$, we will set $\eta_m = C_0 \max\left\{\frac{\mu}{\log n}, \left(\frac{n \log n}{m}\right)^{1/8}\right\}$, for each m, and our desired value of η in (4) will be (roughly) $1 - \eta_{e(F)}$, so that we can show that (3) is always at least the right hand side of (5) with (roughly) $\eta_{e(F)}$ replacing η. From (5) (and as proved in Section 3.5.5), as we will use this for each value of $e(F)$ from $\binom{n}{2}$ down to $\left\lfloor \left(1+\frac{\varepsilon}{2}\right)\frac{dn}{2}\right\rfloor$, one key point is that the sum of $\frac{\eta_m dn/2}{m - dn/2}$ over $m = e(F)$ in this range is $O(C_0\mu/\varepsilon)$, so that this will be much smaller than ε.

The changing value of $\eta_{e(F)}$ in the process makes it more difficult to create the corresponding coupling to create G^*. Indeed, starting with G^* as the complete graph with vertex set $[n]$, when we consider an edge e, if $e \in E(G^*)$, then (throughout most of the process) we will wish to remove it with some probability likely to be lower than that at (3), except this will now change depending on the value of $e(F)$. This, however, will be predictable enough that in creating G^* we can use a bound that does not depend on F (see Section 3.5).

The most significant challenge, then, is to prove that bounds like that at (4) are likely to hold for much of the process (with our new, varying, error term in place of η). As we will see (and as discussed below), this is much more challenging than the corresponding bound in the lower part of sandwich. Therefore, here, it is critical for us here that, as observed and proved by Gao, Isaev and McKay [13], conditioned on its number of edges, F has a nice distribution using the following definition.

Definition 2.3. Let $d, n, m \in \mathbb{N}$ with $0 \leq d \leq n-1$, dn even, and $0 \leq m \leq \binom{n}{2} - dn/2$. We say $F \sim F(n,d,m)$ if F is a random graph with vertex set $[n]$ which has the same distribution as $H + E$ if $H \sim G_d(n)$ and E is a uniformly random set of m edges from $[n]^{(2)} \setminus E(H)$.

For each $dn/2 \leq i \leq \binom{n}{2}$, let F_i be the value of F in this process when it has i edges. Then, for each $dn/2 \leq i \leq \binom{n}{2}$, $F_i \sim F(n,d,i-(dn/2))$. We include a proof of this (along the lines of

[13]) in Section 3.4. We arrive now at the main challenge, requiring the key novelties of this paper. From our discussion above (and as proved later in detail in Section 3), it will suffice to prove the following.

Theorem 2.4. *Let $1/n \ll 1/C \ll \mu \ll 1/C_0 \ll \varepsilon \ll 1$. Let $d \geq C \log n$ and $\varepsilon dn \leq m \leq (n-1-d)n/2$. Let $F \sim F(n,d,m)$. Then, with probability $1-o(n^{-2})$, for each $e, f \in E(F)$, we have*

$$|\{K \in \mathcal{K}_d(F) : e \in E(K)\}| \leq \left(1 + C_0 \max\left\{\frac{\mu}{\log n}, \left(\frac{n \log n}{m}\right)^{1/8}\right\}\right) |\{K \in \mathcal{K}_d(F) : f \in E(K)\}|. \tag{6}$$

Similar to Lemma 2.2, we will prove Theorem 2.4 using switching methods, but this will be significantly more difficult. Why this is more challenging is explained below, but the new ideas we use to do this are the heart of our ability to analyse this process as long as $d = \omega(\log n)$, compared to the regime $d \gg n/\sqrt{\log n}$ in which this was done by Gao, Isaev and McKay [15]. (In [13] a modified two-step process is used to neatly avoid some of these difficulties.)

2.2.2 Switching methods to prove Lemma 2.2 and Theorem 2.4

The switching method was introduced by McKay [25], and is a crucial part of our analysis for both the lower and upper part of the sandwich. We will explain these methods first in the context of the lower part of the sandwich as they are much simpler, before discussing the challenges of switching for the upper part and the key new ideas we will introduce. This also allows us to note one key change in the analysis of the lower part of the sandwich which allows a simpler approach when d is large (and is also used for the upper part of the sandwich).

Switching methods for the lower part of the sandwich, i.e., for Lemma 2.2. Let F have $V(F) = [n]$, and let $e, f \in [n]^{(2)} \setminus E(F)$ share no vertices. Suppose, for (2), we wish to compare the number of d-regular subgraphs in $\mathcal{K}_d(n)$ containing $F + e$ with the number containing $F + f$. Let $\mathcal{K}^e = \{K \in \mathcal{K}_d(n) : F + e \subset K, f \notin E(K)\}$ and $\mathcal{K}^f = \{K \in \mathcal{K}_d(n) : F + f \subset K, e \notin E(K)\}$, and note that if we can show that $|\mathcal{K}^e| \leq (1+\eta)|\mathcal{K}^f|$ then (2) will hold.

To use the switching method, we define a notion of 'switching' and consider when we can transform $K \in \mathcal{K}^e$ into $K' \in \mathcal{K}^f$ using a switching. Comparing the number of ways we can do such a transformation on K with the number of ways we could transform K', will allow us to compare $|\mathcal{K}^e|$ to $|\mathcal{K}^f|$. This is best seen via an example.

Suppose for simplicity that $d = o(n)$. Take the auxiliary bipartite graph L with vertex classes $\mathcal{K}^e$ and $\mathcal{K}^f$ where there is an edge between $K \in \mathcal{K}^e$ and $K' \in \mathcal{K}^f$ if $E(K) \triangle E(K')$ is a cycle with length 10 in which e and f appear on opposite sides of the cycle (see Figure 1). Note that such a cycle alternates between $K \setminus K'$ and $K' \setminus K$. We now aim to show that, for some $D > 0$, for each $K \in \mathcal{K}^e$ we have $d_L(K) \geq (1-\eta/3)D$ and for each $K' \in \mathcal{K}^f$ we have $d_L(K') \leq (1+\eta/3)D$, so that, double-counting the edges of L gives

$$|\mathcal{K}^e| \cdot (1-\eta/3)D \leq e(L) \leq |\mathcal{K}^f| \cdot (1+\eta/3)D, \tag{7}$$

and, as $\eta \ll 1$, we thus have $|\mathcal{K}^e| \leq (1+\eta)|\mathcal{K}^f|$.

Letting $K \in \mathcal{K}^e$, for each $K' \in \mathcal{K}^f$ with $KK' \in E(L)$, $E(K) \triangle E(K')$ is the edge set of a length 10 cycle S which alternates between $K \setminus F$ and $K_n \setminus K$ with e and f on opposite sides. Thus, we can bound $d_L(K)$ by counting such cycles. We will have that the vertex degrees of $K \setminus F$ are all $(1 \pm \eta/100)\Delta$ for some Δ (in Lemma 2.2 this will be $2m/n$). Arbitrarily labelling $V(e) = \{v_1, v_6\}$, we will use the notation in Figure 1 to choose such a cycle S vertex by vertex.

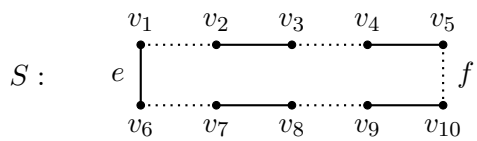

Figure 1: The cycle S alternating between $K \setminus K'$ (solid lines) and $K' \setminus K$ (dotted lines).

After labelling the vertices of f as v_5 and v_{10} (with 2 choices in total), and then choosing v_2 and v_7 (with at most n choices for each), we then have at most $(1+\eta/100)\Delta$ choices for choosing each of v_3, v_4, v_8, v_9, as they are the neighbour in $K \setminus F$ of some previously chosen vertex. Therefore, we have $d_L(K) \leq (1+\eta/100)^4\Delta^4 n^2 \leq (1+\eta/3)\Delta^4 n^2$.

This upper bound however, is actually very close to $d_L(K)$: to ensure for a choice of v_2,v_3,v_4,v_7,v_8, and v_9 that there is some $K' \in \mathcal{K}^f$ with $E(S) = E(K) \triangle E(K')$ we need only make sure that the vertices are all distinct, and that $v_1v_2, v_3v_4, v_6v_7, v_8v_9 \in E(K_n \setminus K)$. When $d = o(n)$, these additional conditions are rather weak as $K_n \setminus K$ is very dense, so we will have at least $(1-\eta/3)\Delta^4 n^2$ such cycles (see Section 10.4). Each such cycle gives rise to an edge in L from K to a different $K' = K - E(S) \cap E(K) + E(S) \setminus E(K)$. Therefore, including the lower bound, we have $d_L(K) = (1 \pm \eta/3)\Delta^4 n^2$. As a similar bound can be shown for f (or using the symmetry between e and f in our set-up), we can do the double-counting at (7) using $D = \Delta^4 n^2$.

For Lemma 2.2, as $F \sim F^-(n,d,m)$, with $m = \omega(n \log n)$ and $d = \omega(\log n)$, a simple concentration inequality (see Section 2.3) and a union bound shows that $K \setminus F$ will be approximately regular, allowing this sketch to be carried out when $d = o(n)$. For larger d it is harder to count these switching cycles as accurately, as no longer is $K_n \setminus K$ so dense. In previous work [13], inequalities like (2) in such a regime were shown to be likely to hold using complex analytic methods. Here, we will use instead two simple ideas to complete the proof when $d = \Omega(n)$. These ideas are very useful for the upper part of the sandwich, so we will introduce them now in the simpler setting of the lower part.

Firstly, we note that we do not need to prove good bounds on the degrees in L that hold for all $K \in \mathcal{K}^e \cup \mathcal{K}^f$. That is, the double-counting at (7) will tolerate good bounds which hold only for most of the graphs $K \in \mathcal{K}^e \cup \mathcal{K}^f$, while for any outliers we can use weaker bounds that are more straightforward to prove. This will allow us to focus on proving good bounds for graphs $K \in \mathcal{K}^e \cup \mathcal{K}^f$ which satisfy some pseudorandom conditions.

Then, in order to show it is likely that most of the graphs $K \in \mathcal{K}_e \cup \mathcal{K}_f$ satisfy some pseudorandom conditions, we exploit the distribution of $F \sim F^-(n,d,m)$. Intuitively, if we form F by taking $K \sim K_d(n)$ and deleting m random edges, then if a property is very likely to hold for K it should be likely to hold for most of the d-regular graphs containing F. (This is proved with the exact parameters we use in Sections 9.1 and 10.5.) This means that instead of studying the properties of the d-regular graphs containing F, we can more simply study a random d-regular graph.

Switching methods for the upper part of the sandwich, i.e., for Theorem 2.4. After using the two ideas above, carrying out a similar switching argument for the upper part (i.e., for Theorem 2.4) will essentially require us to show that if $K \sim K_d(n)$ and F is formed from K by adding m edges uniformly at random from $[n]^{(2)} \setminus E(K)$ (and thus $F \sim F(n,d,m)$) then with high probability we have the following for some suitable ℓ.

A For each distinct $x, y \in [n]$ the number of $(F \setminus K, K)$-alternating x,y-paths of length 2ℓ is well concentrated.

Indeed, if some version of this holds where the paths avoid an arbitrary small set of vertices, then we can put together two such paths between e and f to count degrees in the natural analogue

of the auxiliary graph L discussed above. This is done in detail in Section 4, while here we will focus on **A**.

Compared to similar tasks in the lower part, this is much more difficult in the region $d = o(n)$ as no longer is one of the K or $F \setminus K$ graphs very close to being complete. Our previous discussion would correspond to paths of length $2\ell = 4$, but now we will have to take $\ell = \Omega(\log n/\log\log n)$ when $d = \log^{O(1)} n$, otherwise between some distinct $x, y \in [n]$ no such alternating paths will exist. This pinpoints why the process for the upper part is so much harder to analyse via switchings than the process for the lower part. In their work, Gao, Isaev, and McKay [13] used two processes, one run after the other, so that (very roughly), when such $(F \setminus K, K)$-alternating paths were found, at least K was sparse when compared to $F \setminus K$ (see [13] for more details).

When $F \setminus K$ is sufficiently dense, with $m \geq n^{1+\sigma}$ for any fixed $\sigma > 0$, we use the results of Kim and Vu [19] on the concentration of multivariate polynomials to get sufficiently good bounds on the counts for **A** (see Section 8.2). Thus, we can fix some small $\sigma > 0$, and assume that $d \leq n^{\sigma}$ and $m \leq n^{1+\sigma}$. Take distinct $x, y \in [n]$ and let $\delta = 2m/n$, and note that we may expect $F \setminus K$ to have vertex degrees $(1 + o(1))\delta$. Due to the randomness of $F \setminus K$ and K, if we randomly walk out from x using edges alternately from $F \setminus K$ and then K we should expect this random walk to rapidly mix so that, as long as (say) $\ell \geq 2\log n/\log(\delta d)$, after 2ℓ edges we are roughly no more likely to arrive at y than any other vertex. Heuristically, considering the vertex degrees in $F \setminus K$ and in the d-regular graph K, we should expect there are likely to be $(1 + o(1))\delta^{\ell} d^{\ell}/n$ $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ.

Bounds of this form were proved in certain regions (roughly when $\delta \gg d \log n$) by Gao, Isaev and McKay [13] by working via the spectral properties of binomial random graphs (and using only that K is d-regular), while similar counts on paths within a single graph have been shown by showing certain Markov chains mix rapidly (see, for example, as in [18], building on work in [8]). For Theorem 2.4, we need to work with the two different random graph models, in particular using properties of the more difficult to study random regular graph, and we need to deal with various subtleties due to the accuracy we need and the weaker bound on d that we use. As far as is convenient, then, we rely on pseudorandom conditions of the random graphs involved, in part so that the results we prove may ultimately perhaps prove useful for the study of Conjecture 1.3.

The number of random walks discussed above can be bounded using bounds on the degrees of $F \setminus K$ and the regularity of K. However, when $d = C \log n$ (for our large fixed C), our paths will have length 2ℓ with $\ell = \Omega(\log n/\log\log n)$ and the degrees of $F \setminus K$ are likely to vary by up to $\Omega_C(\delta)$. Thus, we cannot rely alone on the likely concentration of the degrees of $F \setminus K$. However, after revealing the edges of K, for each $v \in [n]$, we will have that $\sum_{u \in N_K(v)} d_{F \setminus K}(u)$ is better concentrated. For any fixed vertex v, even when $d = C \log n$ and $\delta = 2\varepsilon d$, this is likely to be concentrated up to an error term of $O(1/\ell)$. However, to get variables that are this concentrated for *every* vertex v, we ultimately look a step further and consider $\sum_{v' \in N_{F \setminus K}(v)} \sum_{u \in N_K(v')} d_{F \setminus K}(u)$ (see **G2** in Lemma 5.3).

Roughly, then, we count walks alternating between $F \setminus K$ and K using these sums of degrees, and show that most of these walks are actually paths. To do so we rely on pseudorandom properties that later appear as **I3** and **I4**. Then, considering alternating paths from x and y respectively to sets of different endvertices we use lower bounds on the number of edges in $F \setminus K$ between these different sets of endvertices (using **I6** later) to give a lower bound between the number of alternating paths between x and y. A subtlety that is hard to anticipate from our discussion so far is that perhaps the most difficult case for Theorem 2.4 is when $d = C \log n$ and $\delta = \log^{O(1)} n$, but $\delta = \omega(\log^8 n)$. This is because the error term, essentially η, that we wish to use in the counting of our alternating paths is to be smaller than $C_0 \mu/\log n$ (see (6)), but we still have $\ell = \Omega(\log n/\log\log n)$, so that when we build our alternating paths of length ℓ from x, if the next edge from the current endpoint, say v,

is to be in K then v may have a fraction of neighbours in the path so far which exceeds ηd, which here is $O(1)$. This creates some additional small but important complication that, for example, is reflected in the error term we use in **H** in Theorem 6.1 later.

To show that the number of $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ is likely bounded above, we exploit our likely lower bound (applied with a slightly different path length) as well as an inductive upper bound on the number of alternating paths from y. Revealing the edges in $F \setminus K$ not containing x, we deduce that most vertices in $[n] \setminus \{x, y\}$ – those not in Z, say – have at most $(1 + O(\eta))\delta^{\ell-1}d^\ell/n$ $(K, F \setminus K)$-alternating paths of length $2\ell - 1$ to y. Then, revealing the edges in $F \setminus K$ next to x, we get that there are at most $(1+O(\eta))\delta^\ell d^\ell/n$ $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ which do not have their second vertex in $Z \cap N_{F\setminus K}(x)$. As Z was a relatively small set (it will have $O(\eta n)$ vertices), using this and similar properties iteratively will allow us to give an upper bound for all the x, y-paths we are interested in apart from those in an increasingly constrained set of paths. After sufficiently many iterations, a simple bound on the remaining possibilities will be sufficient to give us our desired upper bound. This argument is carried out in Section 7.

2.2.3 Paper outline

In Section 3, we give the details of the edge deletion process we use to create the upper part of the sandwich, reducing the proof of the upper part of the sandwich to Theorem 2.4. In Section 4, using our main switching argument we reduce proving Theorem 4.1 to proving certain pseudorandom conditions concerning switching paths are likely to hold for random graphs F and K. In Sections 5–7 we prove these pseudorandom conditions are likely to hold when $m \le n^{1+\sigma}$ (for any fixed $\sigma > 0$) In Section 5, we give likely upper bounds on the number of $(F \setminus K, K)$-alternating paths of a given length which start from a fixed vertex. In Section 6, we give likely lower bounds on the number of $(F \setminus K, K)$-alternating paths between any two fixed vertices (moreover, on such paths which avoid an arbitrary small set of vertices). In Section 7, we give likely upper bounds on the number of $(F \setminus K, K)$-alternating paths between any two fixed vertices. In Section 8, we give similar bounds to Sections 5–7 when $m \ge n^{1+\sigma}$ (for any fixed small σ). In Section 9, we put this together to prove Theorem 2.4. In combination, this completes the proof for the upper part of the sandwich. In Section 10, we give the lower part of the sandwich and its analysis.

Note that a reader new to the subject may find it more accessible to start with the proof for the lower part of the sandwich in Section 10 before moving on to the proof for the upper part of the sandwich in Sections 3–9.

2.3 Concentration inequalities

We will use the following well-known result of McDiarmid [24, Theorem 2.3(b)].

Theorem 2.5. *Let $X_1, \dots, X_n$ be independent random variables with $0 \le X_i \le 1$ for each $i \in [n]$. Let $S = \sum_{i=1}^n X_i$ and $\mu = \mathbb{E}S$. Then, for any $\varepsilon > 0$,*

$$\mathbb{P}(S \ge (1+\varepsilon)\mu) \le \exp\left(-\frac{\varepsilon^2\mu}{2(1+\varepsilon/3)}\right).$$

We will use Chernoff's bound, in the following standard form (see, for example, [17, Corollary 2.2 and Theorem 2.10]).

Lemma 2.6 (Chernoff's bound)**.** *Let X be a random variable with mean μ which is binomially distributed. Then for any $0 < \gamma < 1$, we have that*

$$\mathbb{P}(|X - \mu| \ge \gamma\mu) \le 2\exp(-\mu\gamma^2/3).$$

In addition, we will use work of Kim and Vu [19] on the concentration of multivariate polynomials, in the form of Theorem 8.6, but quote this in Section 8.2 so that the notation required is close to its application.

3 The upper part of the sandwich

We now give the process producing (G, G^*) for Theorem 1.1 (essentially that of Gao, Isaev and McKay [14]) in Section 3.1, before in the rest of this section showing that the required properties hold (see Claim 3.1), subject only to the proof of Theorem 2.4.

3.1 Edge deletion process

Let $0 < 1/C \ll \varepsilon \ll 1$ and $d \geq C \log n$. Let μ and C_0 be such that $1/C \ll \mu \ll 1/C_0 \ll \varepsilon$. Note that we can assume that $(1+\varepsilon)d/n \leq 1$, for otherwise in Theorem 1.1 we can take $p^* = 1$. For each $0 \leq i < \binom{n}{2} - dn/2$, let

$$\eta_i = C_0 \max\left\{\frac{\mu}{\log n}, \left(\frac{n \log n}{\binom{n}{2} - dn/2 - i}\right)^{1/8}\right\}, \tag{8}$$

and for each $i \geq \binom{n}{2} - dn/2$, let $\eta_i = 0$. Let $e_1, e_2, \ldots$ be edges drawn independently and uniformly at random from $[n]^{(2)}$. Let $x_1, x_2, \ldots$ be independent random variables with $x_i \sim U([0,1])$ for each $i \geq 1$. We will choose the random graphs G and G^* using the sequences $e_1, e_2, \ldots$ and $x_1, x_2, \ldots$.

Produce G: Let F_0 be the complete graph with vertex set $[n]$, and let $\ell(0) = 0$. For each $1 \leq i \leq \binom{n}{2} - dn/2$, let $\ell(i)$ be the least $\ell(i) > \ell(i-1)$ for which $e_{\ell(i)} \in E(F_{i-1})$ and

$$x_{\ell(i)} \leq \frac{|\mathcal{K}_d(F_{i-1} - e_{\ell(i)})|}{\max_{e \in E(F_{i-1})} |\mathcal{K}_d(F_{i-1} - e)|}, \tag{9}$$

and let $F_i = F_{i-1} - e_{\ell(i)}$. Let $G = F_{\binom{n}{2} - dn/2}$.

Produce G^*: Let G_0^* be the complete graph with vertex set $[n]$, and let $m(0) = 0$. Let $R = \lfloor \varepsilon dn/8 \rfloor$. For each $1 \leq i \leq \binom{n}{2}$, let $m(i)$ be the least $m(i) > m(i-1)$ for which $e_{m(i)} \in E(G_{i-1}^*)$ and

$$x_{m(i)} \leq \begin{cases} 1 - \frac{\eta_{i+R-1} dn/2}{\binom{n}{2} - dn/2 - R - (i-1)} & \text{if } i \leq \binom{n}{2} - \frac{dn}{2} - R \\ 1 & \text{otherwise,} \end{cases} \tag{10}$$

and let $G_i^* = G_{i-1}^* - e_{m(i)}$. Let $M \sim \mathrm{Bin}\left((1+\varepsilon)\frac{d}{n}, \binom{n}{2}\right)$, and set $G^* = G_{\binom{n}{2} - M}$.

Note: Observe that, for each $1 \leq i \leq \binom{n}{2} - dn/2$, an edge $e \in E(F_{i-1})$ is only removed from F_{i-1} with probability bigger than 0 if $|\mathcal{K}_d(F_{i-1} - e)| > 0$, and thus F_i always contains at least one d-regular graph. Furthermore, as F_{i-1} is not d-regular, and $\mathcal{K}_d(F_{i-1}) \neq \emptyset$, there is some edge $e \in E(F_{i-1})$ for which $|\mathcal{K}_d(F_{i-1} - e)| > 0$, and thus with probability 1, we will successfully produce $F_1, \ldots, F_{\binom{n}{2} - dn/2}$, and hence G. Similarly, with probability 1, we will successfully produce G^*.

3.2 Analysis overview and deduction of Theorem 1.1

We will show that this process has the following 3 main properties. Recall from Definition 2.3 that $F \sim F(n, d, m)$ is a random graph on n vertices with the same distribution as $H + E$ where $H \sim G_d(n)$ and E is a uniformly random set of m edges from $[n]^{(2)} \setminus E(H)$.

Claim 3.1. a) $G^* \sim G(n, (1+\varepsilon)d/n)$.

b) For each $0 \le i \le \binom{n}{2} - dn/2$, $F_i \sim F(n, d, \binom{n}{2} - dn/2 - i)$.

c) With high probability, $G \subset G^*$.

As $G = F_{\binom{n}{2}-dn/2}$, Claim 3.1 b) implies that $G \sim G_d(n)$. Therefore, Claim 3.1 a), b) and c) directly imply that (G, G^*) satisfies the required properties in Theorem 1.1. Thus, to show the upper part of the sandwich in Theorem 1.1 exists, it is sufficient to prove Claim 3.1. In the rest of this section we will do this, subject only to the proof of Theorem 2.4.

3.3 Distribution of G^*: proof of Claim 3.1 a)

For each $0 \le i \le \binom{n}{2}$, we have that $e(G_i^*) = \binom{n}{2} - i$, and, by symmetry (as the right hand side of (10) depends only on i), the distribution of G_i^* is uniform across all graphs with vertex set $[n]$ and $\binom{n}{2} - i$ edges. Therefore, the distribution of G^* is uniform when conditioned on $e(G^*) = i$, for any $0 \le i \le \binom{n}{2}$. As $G^* = G_{\binom{n}{2}-M}$, for each $0 \le i \le \binom{n}{2}$, we have $\mathbb{P}(e(G^*) = i) = \mathbb{P}(M = i)$. Since $M \sim \mathrm{Bin}\left((1+\varepsilon)\frac{d}{n}, \binom{n}{2}\right)$, we therefore have that $G^* \sim G\left(n, (1+\varepsilon)\frac{d}{n}\right)$.

3.4 Distribution of F_i: proof of Claim 3.1 b)

We now prove Claim 3.1 b), as shown similarly by Gao, Isaev, and McKay for the corresponding result in the lower bound of the sandwich (see [14, Lemma 3]). We will prove by induction on i that, for each $0 \le i \le \binom{n}{2} - dn/2$, $F_i \sim F(n, d, \binom{n}{2} - dn/2 - i)$.

First note that this is trivially true for $i = 0$, and that, if F is any graph with $V(F) = [n]$ and $e(F) = m$, where $dn/2 \le m \le \binom{n}{2}$, then

$$\mathbb{P}(F(n, d, m - dn/2) = F) = \frac{|\mathcal{K}_d(F)|}{|\mathcal{K}_d(n)|} \cdot \binom{\binom{n}{2} - dn/2}{m - dn/2}^{-1}. \tag{11}$$

Assume then that $F_{i-1} \sim F(n, d, \binom{n}{2} - dn/2 - (i-1))$ and let F be an arbitrary graph with $V(F) = [n]$, $e(F) = \binom{n}{2} - i$, and $\mathcal{K}_d(F) \neq \emptyset$. Our aim is to show that

$$\mathbb{P}(F_i = F) = \frac{|\mathcal{K}_d(F)|}{|\mathcal{K}_d(n)|} \cdot \binom{\binom{n}{2} - dn/2}{\binom{n}{2} - dn/2 - i}^{-1}, \tag{12}$$

so that, as F was arbitrary subject to $V(F) = [n]$, $e(F) = \binom{n}{2} - i$, and $\mathcal{K}_d(F) \neq \emptyset$, we have $F_i \sim F(n, d, \binom{n}{2} - dn/2 - i)$.

Let $f \in [n]^{(2)} \setminus E(F)$. Consider $\mathbb{P}(F_i = F | F_{i-1} = F + f)$. If $F_{i-1} = F + f$, the probability that, for the first $j' > \ell(i-1)$ with $e_{j'} \in E(F_{i-1})$, the edge $e_{j'}$ is not immediately removed from $F_{i-1} = F + f$ is, using (9) and that $x_{j'} \sim U([0,1])$,

$$\frac{1}{e(F)+1} \sum_{e' \in E(F+f)} \left(1 - \frac{|\mathcal{K}_d(F + f - e')|}{\max_{e \in E(F+f)} |\mathcal{K}_d(F + f - e)|}\right)$$
$$= 1 - \frac{e(F) + 1 - dn/2}{e(F)+1} \cdot \frac{|\mathcal{K}_d(F+f)|}{\max_{e \in E(F+f)} |\mathcal{K}_d(F + f - e)|},$$

while the probability that $e_{j'} = f$ and $e_{j'}$ is removed from $F_{i-1} = F + f$ is

$$\frac{1}{e(F)+1} \cdot \frac{|\mathcal{K}_d(F)|}{\max_{e \in E(F+f)} |\mathcal{K}_d(F + f - e)|}.$$

Therefore, letting $\hat{p} = \mathbb{P}(F_i = F | F_{i-1} = F + f)$, we have

$$\hat{p} = \frac{1}{e(F)+1} \cdot \frac{|\mathcal{K}_d(F)|}{\max_{e \in E(F+f)} |\mathcal{K}_d(F+f-e)|} + \left(1 - \frac{e(F)+1-dn/2}{e(F)+1} \cdot \frac{|\mathcal{K}_d(F+f)|}{\max_{e \in E(F+f)} |\mathcal{K}_d(F+f-e)|}\right)\hat{p},$$

so that

$$\mathbb{P}(F_i = F | F_{i-1} = F + f) = \hat{p} = \frac{|\mathcal{K}_d(F)|}{(e(F)+1-dn/2)|\mathcal{K}_d(F+f)|}. \tag{13}$$

Then, as by assumption we have that $F_{i-1} \sim F(n, d, \binom{n}{2} - dn/2 - (i-1))$, we see that

$$\begin{aligned}
\mathbb{P}(F_i = F) &= \sum_{f \in [n]^{(2)} \setminus E(F)} \mathbb{P}(F_i = F | F_{i-1} = F + f) \cdot \mathbb{P}(F_{i-1} = F + f) \\
&\overset{(13),(11)}{=} \sum_{f \in [n]^{(2)} \setminus E(F)} \frac{|\mathcal{K}_d(F)|}{(e(F)+1-dn/2)|\mathcal{K}_d(F+f)|} \cdot \frac{|\mathcal{K}_d(F+f)|}{|\mathcal{K}_d(n)|} \cdot \binom{\binom{n}{2} - dn/2}{\binom{n}{2} - dn/2 - (i-1))}^{-1} \\
&= \frac{|\mathcal{K}_d(F)|}{|\mathcal{K}_d(n)|} \cdot \frac{\binom{n}{2} - e(F)}{e(F)+1-dn/2} \cdot \binom{\binom{n}{2} - dn/2}{\binom{n}{2} - dn/2 - (i-1))}^{-1} \\
&= \frac{|\mathcal{K}_d(F)|}{|\mathcal{K}_d(n)|} \cdot \binom{\binom{n}{2} - dn/2}{\binom{n}{2} - dn/2 - i}^{-1},
\end{aligned}$$

and thus (12) holds, as required.

3.5 Containment of G in G^*: proof of Claim 3.1 c)

To prove Claim 3.1 c), it will be convenient to define two more sequences of random graphs. Let H_0 have vertex set $[n]$ and edge set $[n]^{(2)}$. Let $k(0) = 0$. For each $1 \le i \le \binom{n}{2}$ in turn, let $k(i)$ be the least $k > k(i-1)$ such that $e_k \in E(H_{i-1})$, and let $H_i = H_{i-1} - e_{k(i)}$. That is, we form the sequence $H_0, H_1, \ldots$ by removing an edge as soon as it appears in the sequence $e_1, e_2, \ldots$ and $k(i)$ is the number of places we need to look along in this sequence to see i different edges.

Let G_0^+ have vertex set $[n]$ and edge set $[n]^{(2)}$. Recall that $R = \lfloor \varepsilon dn/8 \rfloor$. For each $1 \le i \le \binom{n}{2}$, if

$$x_{k(i)} \le \begin{cases} 1 - \frac{\eta_{i+R-1} dn/2}{\binom{n}{2} - dn/2 - R - i + 1} & \text{if } i \le \binom{n}{2} - dn/2 - R \\ 1 & \text{otherwise} \end{cases} \tag{14}$$

then let $G_i^+ = G_{i-1}^+ - e_{k(i-1)}$ and, otherwise, let $G_i^+ = G_{i-1}^+$. Note that the right hand side of (14) is the same as that in (10). Thus, when an edge e is considered for deletion in the sequences $G_0^*, G_1^*, \ldots$ and $G_0^+, G_1^+, \ldots$, it will be deleted from one only if it is deleted from the other; the difference is that, if it survives this, e will never be deleted in the sequence $G_0^+, G_1^+, \ldots$. It is perhaps also worth noting here that, instead of the choice in Section 3.1, we could instead choose G^* using the sequence $G_0^+, G_1^+, \ldots$. However, this would be a more difficult choice there as we do not have the property that $e(G_i^+) = \binom{n}{2} - i$ for each $0 \le i \le \binom{n}{2}$, and we prefer to keep our processes generating the coupling (G, G^*) as clean as possible.

We will consider the following two properties, both of which we show will occur with high probability, and which together will imply that $G \subset G^*$. For them, we use $N = \binom{n}{2} - dn/2 - 2R$.

B1 For each $1 \le i \le \binom{n}{2} - dn/2 - R$, and each $e \in E(F_{i-1})$,

$$\frac{|\mathcal{K}_d(F_{i-1} - e)|}{\max_{f \in E(F_{i-1})} |\mathcal{K}_d(F_{i-1} - f)|} \ge 1 - \frac{\eta_{i-1} dn/2}{\binom{n}{2} - dn/2 - i + 1}. \tag{15}$$

B2 $e(G_N^+) \le e(H_N) + R$.

Roughly speaking (see also Claim 3.2 below), comparing the right hand side of (15) with that of (9) and (10) we see that, as long as **B1** holds, if an edge e_j in the sequence $e_1, e_2, \dots$ is deleted in the sequence $G_0^*, G_1^*, \dots$ then it will also be deleted from $F_0, F_1, \dots$ as long as the graph $F_{i'}$ from which it is to be deleted does not have more than R more edges than the graph $G_{i''}^+$ from which e_j was deleted. That the difference in the number of edges here is not too large will follow as long as **B2** holds.

To prove Claim 3.1 c), we will prove the following claim (using Theorem 2.4 for Claim 3.2 iv)).

Claim 3.2. i) For each $1 \le i \le \binom{n}{2} - dn/2$, $k(i) \le \ell(i), m(i)$.

ii) If **B2** holds, then $m(i) \le k(i+R)$ for all $0 \le i \le N - R$.

iii) For each $1 \le i \le N$, if **B1** holds and $m(j) \le \ell(j+R)$ for all $1 \le j \le i$, then $F_{i+R} \subset G_i^*$.

iv) **B1** holds with high probability.

v) **B2** holds with high probability.

By Claim 3.2 iv) and v), **B1** and **B2** hold with high probability. When this happens, from Claim 3.2 i) and ii) we have that, for all $1 \le j \le N - R$, $m(j) \le k(j+R) \le \ell(j+R)$. Thus, by Claim 3.2 iii), $G \subset F_N \subset G_{N-R}^*$. As $G^* = G^*_{\binom{n}{2}-M}$, when $M \ge \binom{n}{2} - N + R = dn/2 + 3R$, we have that $G_{N-R}^* \subset G^*$. As $M \sim \mathrm{Bin}\left((1+\varepsilon)\frac{d}{n}, \binom{n}{2}\right)$ and $R = \lfloor \varepsilon dn/8 \rfloor$, by a Chernoff bound (Lemma 2.6), with high probability we have $M \ge (1+\varepsilon)dn/2 - R \ge dn/2 + 3R$. Therefore $G \subset G^*$ with high probability. Thus, it is left only to prove Claim 3.2 i)–v) which we do in Sections 3.5.1–3.5.5 respectively (using Theorem 2.4).

3.5.1 Proof of Claim 3.2 i)

Let $1 \le i \le \binom{n}{2}$. Observe that it follows from the definition of the processes that there are at most $i-1$ unique edges in $e_1, \dots, e_{k(i)-1}$, while the edges $e_{\ell(1)}, \dots, e_{\ell(i)}$ are all unique and the edges $e_{m(1)}, \dots, e_{m(i)}$ are all unique. Thus, $\ell(i), m(i) \ge k(i)$.

3.5.2 Proof of Claim 3.2 ii)

Suppose that **B2** holds. Let $0 \le i \le N - R$ and suppose, for contradiction, that $k(i+R) < m(i)$. Then, there is a set $I \subset [i+R]$ such that $|I| = R+1$ and, for each $j \in I$, $k(j) \notin \{m(i') : i' \in [i]\}$. Note that in particular $|I| \ge 1$. Take $j \in I$. As $k(j) \le k(i+R) < m(i)$, there is some $i' \in [i]$ such that $m(i'-1) < k(j) < m(i')$. From Claim 3.2 i), we have $k(i-1) \le m(i-1)$, so therefore $j \ge i'$.

As $e_{k(j)}$ is the first appearance of that edge in the sequence $e_1, e_2, \dots$, we have $e_{k(j)} \in E(G_{i'-1}^*)$. Thus, as $m(i'-1) < k(j) < m(i')$, we have that (10) does not hold with $(k(j), i')$ replacing $(m(i), i)$, that is,

$$x_{k(j)} > 1 - \frac{\eta_{i'+R-1} dn/2}{\binom{n}{2} - dn/2 - R - i' + 1} \ge 1 - \frac{\eta_{j+R-1} dn/2}{\binom{n}{2} - dn/2 - R - j + 1},$$

where we have used that $j \geq i'$. Therefore, from the condition at (14), we have $e_{k(j)} \in E(G_j^+)$, and hence $e_{k(j)} \in E(G_k^+)$ for each $k \geq j$. In particular, as $j \leq i + R \leq N$, $e_{k(j)} \in E(G_N^+)$. On the other hand, $e_{k(j)} \notin E(H_j)$, and thus $e_{k(j)} \notin E(H_k)$ for any $k \geq j$. In particular, again, $e_{k(j)} \notin E(H_N)$.

Therefore, $\{e_{k(j)} : j \in I\} \subset E(G_N^+) \setminus E(H_N)$, and thus $|E(G_N^+) \setminus E(H_N)| \geq R + 1$. This contradicts **B2**, so we must have had $k(i + R) \geq m(i)$.

3.5.3 Proof of Claim 3.2 iii)

Fix $1 \leq i \leq N$. Suppose that **B1** holds and $m(j) \leq \ell(j+R)$ for all $1 \leq j \leq i$. Let $e \in [n]^{(2)} \setminus E(G_i^*)$. Then, as $e \notin E(G_i^*)$, there is some $j \in [i]$ with $e_{m(j)} = e$ for which (10) holds with j replacing i, so

$$x_{m(j)} \leq 1 - \frac{\eta_{j+R-1} dn/2}{\binom{n}{2} - dn/2 - R - j + 1}.$$

As $m(j) \leq \ell(j + R)$, there is some $j' \leq j + R$ such that $\ell(j' - 1) < m(j) \leq \ell(j')$. Note that

$$x_{m(j)} \leq 1 - \frac{\eta_{j+R-1} dn/2}{\binom{n}{2} - dn/2 - j - R + 1} \leq 1 - \frac{\eta_{j'-1} dn/2}{\binom{n}{2} - dn/2 - j' + 1} \overset{\mathbf{B1}}{\leq} \frac{|\mathcal{K}_d(F_{j'-1} - e_{m(j)})|}{\max_{e \in E(F_{j'-1})} |\mathcal{K}_d(F_{j'-1} - e)|},$$

where we have used that $j' \leq j + R \leq \binom{n}{2} - dn/2 - R$. Then, either $e_{m(j)} \notin E(F_{j'-1})$, or, as $\ell(j' - 1) < m(j) \leq \ell(j')$, condition (9) in the process above implies that $\ell(j') = m(j)$. In either case, as $j' \leq j + R \leq i + R$, we thus have $e \notin E(F_{i+R})$. As $e \in [n]^{(2)} \setminus E(G_i^*)$ was arbitrarily chosen, we thus have $F_{i+R} \subset G_i^*$, as required.

3.5.4 B1 is likely to hold: proof of Claim 3.2 iv)

For each $1 \leq i \leq \binom{n}{2} - dn/2$, we have $F_{i-1} \sim F(n, d, \binom{n}{2} - dn/2 - (i-1))$ by Claim 3.1 b). Therefore, by Theorem 2.4 and a union bound, with high probability we have, for each $1 \leq i \leq \binom{n}{2} - dn/2 - R$, and each $e, f \in E(F_{i-1})$,

$$|\{K \in \mathcal{K}_d(F_{i-1}) : e \in E(K)\}| \leq (1 + \eta_{i-1}) \, |\{K \in \mathcal{K}_d(F_{i-1}) : f \in E(K)\}|. \tag{16}$$

Note also that, by double counting,

$$\begin{aligned} \frac{dn}{2} \cdot |\mathcal{K}_d(F_{i-1})| &= \sum_{f \in E(F_{i-1})} |\{K \in \mathcal{K}_d(F_{i-1}) : f \in E(K)\}| \\ &\geq e(F_{i-1}) \cdot \min_{f \in E(F_{i-1})} |\{K \in \mathcal{K}_d(F_{i-1}) : f \in E(K)\}|. \end{aligned} \tag{17}$$

Therefore, for each $e \in E(F_{i-1})$, we have

$$\begin{aligned} \frac{|\mathcal{K}_d(F_{i-1} - e)|}{\max_{f \in E(F_{i-1})} |\mathcal{K}_d(F_{i-1} - f)|} &= \frac{|\mathcal{K}_d(F_{i-1})| - |\{K \in \mathcal{K}_d(F_{i-1}) : e \in E(K)\}|}{\max_{f \in E(F_{i-1})} (|\mathcal{K}_d(F_{i-1})| - |\{K \in \mathcal{K}_d(F_{i-1}) : f \in E(K)\}|)} \\ &= 1 - \frac{|\{K \in \mathcal{K}_d(F_{i-1}) : e \in E(K)\}| - \min_{f \in E(F_{i-1})} |\{K \in \mathcal{K}_d(F_{i-1}) : f \in E(K)\}|}{|\mathcal{K}_d(F_{i-1})| - \min_{f \in E(F_{i-1})} |\{K \in \mathcal{K}_d(F_{i-1}) : f \in E(K)\}|} \\ &\overset{(16)}{\geq} 1 - \frac{\eta_{i-1} \min_{f \in E(F_{i-1})} |\{K \in \mathcal{K}_d(F_{i-1}) : f \in E(K)\}|}{|\mathcal{K}_d(F_{i-1})| - \min_{f \in E(F_{i-1})} |\{K \in \mathcal{K}_d(F_{i-1}) : f \in E(K)\}|} \\ &\overset{(17)}{\geq} 1 - \frac{\eta_{i-1} \frac{dn}{2e(F_{i-1})} |\mathcal{K}_d(F_{i-1})|}{\left(1 - \frac{dn}{2e(F_{i-1})}\right) |\mathcal{K}_d(F_{i-1})|} \\ &= 1 - \frac{\eta_{i-1} dn/2}{e(F_{i-1}) - dn/2} = 1 - \frac{\eta_{i-1} dn/2}{\binom{n}{2} - dn/2 - i + 1}. \end{aligned}$$

Thus we have that **B1** holds with high probability, as required.

3.5.5 B2 is likely to hold: Proof of Claim 3.2 v)

For each $1 \le i \le \binom{n}{2}$, let $y_i = x_{k(i)}$, and observe that $y_1, \dots, y_{\binom{n}{2}}$ are independent random variables whose distribution is uniform on $[0,1]$. Recalling that $R = \lfloor \varepsilon dn/8 \rfloor$, for each $1 \le i \le \binom{n}{2} - dn/2 - R$, let $Y_i = 0$ if

$$y_i \le 1 - \frac{\eta_{i+R-1} dn/2}{\binom{n}{2} - dn/2 - R - i + 1} \tag{18}$$

and $Y_i = 1$ otherwise. As the condition in (18) depends only on i, we also have that $Y_1, \dots, Y_{\binom{n}{2}}$ are independent random variables. Recalling that $N = \binom{n}{2} - dn/2 - 2R$, let $S = \sum_{i=1}^{N} Y_i$. Observe from our processes producing $H_0, H_1, \dots$ and $G_0^+, G_1^+, \dots$ that $S = e(G_N^+) - e(H_N)$. We have, then,

$$\begin{aligned}
\mathbb{E}S &= \sum_{i=1}^{N} \frac{\eta_{i+R-1} dn/2}{\binom{n}{2} - dn/2 - R - i + 1} \\
&\overset{(8)}{=} \sum_{i=1}^{N} \frac{C_0 dn/2}{\binom{n}{2} - dn/2 - R - i + 1} \max\left\{ \frac{\mu}{\log n}, \left(\frac{n \log n}{\binom{n}{2} - dn/2 - R - i + 1} \right)^{1/8} \right\} \\
&= \frac{C_0 dn}{2} \cdot \sum_{j=R+1}^{\binom{n}{2} - dn/2 - R} \frac{1}{j} \max\left\{ \frac{\mu}{\log n}, \left(\frac{n \log n}{j} \right)^{1/8} \right\} \\
&\le \frac{C_0 dn}{2} \cdot \sum_{j=R+1}^{\binom{n}{2}} \frac{1}{j} \left(\frac{\mu}{\log n} + \left(\frac{n \log n}{j} \right)^{1/8} \right) \\
&\le \frac{C_0 dn}{2} \cdot \frac{2\mu \log(n^2)}{\log n} + \frac{C_0 dn}{2} \int_{j=R}^{\binom{n}{2}} \frac{(n \log n)^{1/8}}{j^{9/8}} dj \\
&\le 2C_0 \mu dn + 4C_0 dn \left(\frac{n \log n}{R} \right)^{1/8} \le 2C_0 \mu dn + 4C_0 dn \left(\frac{10 n \log n}{\varepsilon dn} \right)^{1/8} \le \frac{\lfloor \varepsilon dn/8 \rfloor}{2} = \frac{R}{2},
\end{aligned}$$

where we have used that $d \ge C \log n$ and $1/C \ll \mu \ll 1/C_0 \ll \varepsilon$. Therefore, by Theorem 2.5 applied with $\varepsilon = 1$, we have $\mathbb{P}(S \ge R) \le \exp(-R/8) = o(1)$. Thus, with high probability, $e(G_N^+) - e(H_N) = S \le R$, and hence **B2** holds.

4 Main switching argument

In this section we will show that, if F is roughly regular, and for all but very few graphs $K \in \mathcal{K}_d(F)$ we have good bounds on the number of certain $(F \setminus K, K)$-alternating x, y-paths for each fixed x and y, we can use a switching argument to prove that the edges in F appear in roughly the same number of d-regular subgraphs of F. In certain situations, we want these switching x, y-paths to avoid a set of up to $40 \log n$ vertices Z, and here[1] $|Z|$ can be much larger than ηd as we only use the bound $\eta d \ge C_0$ for some large constant C_0. Thus, if $|(Z \cup \{x\}) \cap N_K(y)|$ is large this can affect significantly the number of x, y-paths leaving y with an edge in K, which explains the error term in **C2** below.

Theorem 4.1. *Let $d \in [n]$, $\delta > 0$ and $0 < 1/C_0, \eta \ll 1$ satisfy $\eta d \ge C_0$ and $\eta^2 \delta \ge C_0 \log n$. Let $4 \le \ell \le 10 \log n$. Let F be a graph with vertex set $[n]$ and $\mathcal{K}_d(F) \ne \emptyset$ for which the following holds.*

[1] For example, using the notation in Theorem 2.4, when $\delta = 2m/n = \omega(\log^8 n)$ and $d = C \log n$, we have $\eta = C_0 \mu / \log n$ and $\eta d = C_0 \mu \cdot C = O(1)$.

C1 *For each $v \in [n]$, $d_F(v) = d + (1 \pm \eta)\delta$.*

Suppose that, for all but at most $n^{-6}|\mathcal{K}_d(F)|$ graphs $K \in \mathcal{K}_d(F)$ the following hold.

C2 *For each distinct $x, y \in [n]$ and each $Z \subset [n]$ with $|Z| \leq 4\ell$, the number of $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ which do not have any internal vertices in Z is at least $\left(1 - \eta - \frac{|(Z \cup \{x\}) \cap N_K(y)|}{d}\right) d^\ell \delta^\ell / n$.*

C3 *For each distinct $x, y \in [n]$, the number of $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ is at most $(1 + \eta)\, d^\ell \delta^\ell / n$.*

Then, for each $e, f \in E(F)$, $|\{K \in \mathcal{K}_d(F) : e \in E(K)\}| \leq (1 + 40\eta)|\{K \in \mathcal{K}_d(F) : f \in E(K)\}|$.

Proof. Let $\mathcal{K}^{\text{bad}}$ be the set of $K \in \mathcal{K}_d(F)$ for which either **C2** or **C3** does not hold, so that, by our assumptions, $|\mathcal{K}^{\text{bad}}| \leq n^{-6}|\mathcal{K}_d(F)|$. We will start by showing that each edge $e \in E(F)$ is contained in at least a small portion of the graphs in $\mathcal{K}_d(F)$, as follows.

Claim 4.2. For each $e \in E(F)$, $|\{K \in \mathcal{K}_d(F) : e \in E(K)\}| \geq d|\mathcal{K}_d(F)|/n^2$.

Proof of Claim 4.2. Let $e \in E(F)$. Suppose, for contradiction, that $|\{K \in \mathcal{K}_d(F) : e \in E(K)\}| < |\mathcal{K}_d(F)|/n^2$. Let $\mathcal{K}_e^+ = \{K \in \mathcal{K}_d(F) : e \in E(K)\}$ and $\mathcal{K}_e^- = \{K \in \mathcal{K}_d(F) : e \notin E(K)\}$. Let L_e be the auxiliary bipartite graph with vertex classes $\mathcal{K}_e^+$ and $\mathcal{K}_e^-$ where there is an edge between $K \in \mathcal{K}_e^+$ and $K' \in \mathcal{K}_e^-$ if $E(K) \triangle E(K')$ is the edge set of a cycle of length $2\ell + 2$.

First, note that, for each $K \in \mathcal{K}_e^+$ and $K' \in \mathcal{K}_e^-$ with $KK' \in E(L_e)$ we have that $E(K) \triangle E(K') - e$ is an $(F \setminus K, K)$-alternating path of length $2\ell + 1$ between the vertices in e. Thus, as $\Delta(F \setminus K) \leq (1 + \eta)\delta$ from **C1**, for each $K \in \mathcal{K}_e^+$ we have $d_{L_e}(K) \leq (1 + \eta)^\ell \delta^\ell d^\ell$.

Now, as $|\mathcal{K}_e^+| = |\{K \in \mathcal{K}_d(F) : e \in E(K)\}| < |\mathcal{K}_d(F)|/n^2$, we have $|\mathcal{K}_e^-| \geq \left(1 - \frac{1}{n^2}\right)|\mathcal{K}_d(F)|$. Thus for at least $|\mathcal{K}_d(F)|/2$ graphs $K' \in \mathcal{K}_e^-$ we have $K' \notin \mathcal{K}^{\text{bad}}$. Let $K' \in \mathcal{K}_e^-$ be such a K' and let u, v be such that $e = uv$, so that $uv \notin E(K')$. Choose $u' \in N_{K'}(u)$, with d possibilities. Then, by **C2** with $Z = \{u\}$ there are at least $\delta^\ell d^\ell / 2n$ $(F \setminus K', K')$-alternating u', v-paths P with length 2ℓ which do not contain u, where we have used that $\eta \ll 1$ and $\eta d \geq C_0$. For each such path P, we have $K' - uu' + uv - (E(P) \cap E(K')) + (E(P) \setminus E(K')) \in \mathcal{K}_e^+$. Therefore, $d_{L_e}(K') \geq \delta^\ell d^{\ell+1}/2n$.

As this holds for at least $|\mathcal{K}_d(F)|/2$ graphs $K' \in \mathcal{K}_e^-$, we have, double-counting the edges of L, that $|\mathcal{K}_e^+| \cdot (1 + \eta)^\ell \delta^\ell d^\ell \geq e(L) \geq \delta^\ell d^{\ell+1} \cdot |\mathcal{K}_d(F)|/4n$. Thus, $|\mathcal{K}_e^+| \geq d|\mathcal{K}_d(F)|/\left((1 + \eta)^\ell \cdot 4n\right) \geq d|\mathcal{K}_d(F)|/n^2$, as required, where we have used that $\ell \leq 10 \log n$ and $\eta \ll 1$. ⊡

Now, let $e, f \in E(F)$ share no vertices, and suppose, for contradiction, that

$$\{K \in \mathcal{K}_d(F) : e \in E(K)\}| > (1 + 18\eta)|\{K \in \mathcal{K}_d(F) : f \in E(K)\}|. \tag{19}$$

Label vertices so that $e = u_1u_2$ and $f = v_1v_2$. Let $\mathcal{K}^e$ be the set of $K \in \mathcal{K}_d(F)$ which contain e but not f, and let $\mathcal{K}^f$ be the set of $K \in \mathcal{K}_d(F)$ which contain f but not e. From (19) and Claim 4.2, we have $|\mathcal{K}^e| > 18\eta|\{K \in \mathcal{K}_d(F) : f \in E(K)\}| > |\mathcal{K}_d(F)|/n^2$, where have used that $\eta d \geq C_0$.

Let $L_{e,f}$ be the bipartite auxiliary graph with vertex classes $\mathcal{K}^e$ and $\mathcal{K}^f$ where there is an edge between $K \in \mathcal{K}^e$ and $K' \in \mathcal{K}^f$ if $(E(K) \triangle E(K')) \setminus \{e, f\}$ is the edge set of two vertex-disjoint paths P_1 and P_2 such that, for each $i \in [2]$, P_i is an $(K' \setminus K, K \setminus K')$-alternating u_i, v_i-path of length $2\ell + 2$. We will show the following claim.

Claim 4.3. For each $K \in \mathcal{K}^e \setminus \mathcal{K}^{\text{bad}}$, $d_{L_{e,f}}(K) \geq (1 - 10\eta)d^{2\ell+2}\delta^{2\ell+2}/n^2$.

Proof of Claim 4.3. Let $K \in \mathcal{K}^e \setminus \mathcal{K}^{\text{bad}}$. Fix $v_1' \in N_K(v_1) \setminus \{u_1, u_2, v_2\}$ and $v_2' \in N_K(v_2) \setminus \{u_1, u_2, v_1, v_1'\}$, for which there are at least $(d-3)(d-4) \geq (1-\eta)d^2$ choices, where we have used that $\eta d \geq C_0$. Let $v_1'' \in N_{F\setminus K}(v_1') \setminus \{u_1, u_2, v_1, v_2, v_2'\}$, with at least $(1-2\eta)\delta$ choices by **C1**. Let P_1 be an $(F \setminus K, K)$-alternating u_1, v_1''-path of length 2ℓ which does not have any internal vertices in $\{u_2, v_1, v_1', v_2, v_2'\}$, noting that from **C2** there are at least $(1-2\eta)\, d^\ell \delta^\ell / n$ choices for P_1, again using that $\eta d \geq C_0$.

Let $Z = V(P_1) \cup \{u_2, v_1, v_1', v_2, v_2'\}$. Let $v_2'' \in N_{F\setminus K}(v_2') \setminus Z$ be such that $|N_K(v_2'') \cap Z| \leq \eta d/2$. Note that the number of edges in K adjacent to Z is at least $\eta d/2 \cdot |\{v : |N_K(v) \cap Z| > \eta d/2\}|$ and at most $d|Z|$. Therefore the number of choices for v_2'' is at least

$$d_{F\setminus K}(v_2') - |Z| - |\{v : |N_K(v) \cap Z| > \eta d/2\}| \geq d_{F\setminus K}(v_2') - |Z| - \frac{d|Z|}{\eta d/2} \geq (1-2\eta)\delta,$$

where we have used that $d_{F\setminus K}(v_2') \geq (1-\eta)\delta$ by **C1**, $\eta^2\delta \geq C_0 \log n$ and $\ell \leq 10 \log n$. Then, let P_2 be an $(F \setminus K, K)$-alternating u_2, v_2''-path of length 2ℓ which does not have any internal vertices in Z, so that, by the choice of v_2'' and by **C2**, we have at least $(1-2\eta)d^\ell\delta^\ell/n$ choices for P_2.

Let $K' = K - E(P_1 \cup P_2) \cap E(K) - \{u_1u_2, v_1v_1', v_2v_2'\} + E(P_1 \cup P_2) \setminus E(K) + \{v_1v_2, v_1'v_1'', v_2'v_2''\}$. Note that $K' \in \mathcal{K}^f$, and that we have made the unique choices for $v_1', v_2', v_1'', P_1, v_2'', P_2$ which produces this choice of K'. Therefore, combining our lower bounds on the choices made, we have

$$d_{L_{e,f}}(K) \geq (1-\eta)d^2 \cdot (1-2\eta)\delta \cdot (1-2\eta)\, \frac{d^\ell\delta^\ell}{n} \cdot (1-2\eta)\delta \cdot (1-2\eta)\frac{d^\ell\delta^\ell}{n} \geq (1-10\eta)\frac{d^{2\ell+2}\delta^{2\ell+2}}{n^2},$$

as required. ⊡

From **C1** and **C3**, we have that, for each $K \in \mathcal{K}_d(F) \setminus \mathcal{K}^{\text{bad}}$ and each distinct $x, y \in [n]$ the number of $(F \setminus K, K)$-alternating x, y-paths of length $2\ell+2$ is at most $(1+\eta)\delta \cdot d \cdot (1+\eta)\, d^\ell\delta^\ell/n$. Thus, applying this to (u_1, v_1) and (u_2, v_2), and using $(1+\eta)^4 \leq (1+5\eta)$, we have the following property.

D1 For each $K' \in \mathcal{K}^f \setminus \mathcal{K}^{\text{bad}}$, we have $d_{L_{e,f}}(K') \leq (1+5\eta)d^{2\ell+2}\delta^{2\ell+2}/n^2$.

Furthermore, using **C1**, we have that for each $K \in \mathcal{K}_d(F)$ and each distinct $x, y \in [n]$ the number of $(F \setminus K, K)$-alternating x, y-paths of length $2\ell+2$ is at most $(1+\eta)^{\ell+1}\delta^{\ell+1}d^{\ell+1} \leq \sqrt{n} \cdot \delta^{\ell+1}d^{\ell+1}$, where we have used that $\ell \leq 10 \log n$ and $\eta \ll 1$. Therefore, we have the following property.

D2 For each $K' \in \mathcal{K}^f$, we have $d_{L_{e,f}}(K') \leq n \cdot \delta^{2\ell+2}d^{2\ell+2}$.

Using that $|\mathcal{K}^{\text{bad}}| \leq n^{-6}|\mathcal{K}_d(F)|$, we have

$$\begin{aligned} e(L_{e,f}) &\overset{\textbf{D1},\textbf{D2}}{\leq} |\mathcal{K}^f| \cdot (1+5\eta)\, \frac{\delta^{2\ell+2}d^{2\ell+2}}{n^2} + \frac{|\mathcal{K}_d(F)|}{n^6} \cdot n \cdot \delta^{2\ell+2}d^{2\ell+2} \\ &\overset{\text{Clm 4.2}}{\leq} |\mathcal{K}^f| \cdot (1+5\eta)\, \frac{\delta^{2\ell+2}d^{2\ell+2}}{n^2} + \frac{|\mathcal{K}^e|}{n} \cdot \frac{\delta^{2\ell+2}d^{2\ell+2}}{n^2}, \end{aligned} \tag{20}$$

where we have used that $|\mathcal{K}^e| > |\mathcal{K}_d(F)|/n^2$.

Furthermore, as $|\mathcal{K}^e| > |\mathcal{K}_d(F)|/n^2$, we have $|\mathcal{K}^e \setminus \mathcal{K}^{\text{bad}}| \geq (1-1/n)|\mathcal{K}^e|$. Thus, by Claim 4.3, we have

$$e(L_{e,f}) \geq (1-1/n)|\mathcal{K}^e| \cdot (1-10\eta)\frac{d^{2\ell+2}\delta^{2\ell+2}}{n^2}.$$

In combination with (20), we have that

$$(1-3/n)\cdot|\mathcal{K}^e|\cdot(1-10\eta)\,\frac{\delta^{2\ell+2}d^{2\ell+2}}{n^2}\le|\mathcal{K}^f|\cdot(1+5\eta)\,\frac{\delta^{2\ell+2}d^{2\ell+2}}{n^2}.$$

As $\eta d\ge C_0$ implies $\eta n\ge C_0$, we have that $|\{K\in\mathcal{K}_d(F):e\in E(K)\}|\le(1+18\eta)\,|\{K\in\mathcal{K}_d(F):f\in E(K)\}|$, a contradiction to (19).

Therefore, for each $e,f\in E(F)$ which share no vertices we have $|\{K\in\mathcal{K}_d(F):e\in E(K)\}|\le(1+18\eta)\,|\{K\in\mathcal{K}_d(F):f\in E(K)\}|$. Let, then, $e,f\in E(F)$. Pick $e'\in E(F)$ with no vertices in $V(e)\cup V(f)$. Then, we have

$$\begin{aligned}|\{K\in\mathcal{K}_d(F):e\in E(F)\}|&\le(1+18\eta)\,|\{K\in\mathcal{K}_d(F):e'\in E(K)\}|\\&\le(1+40\eta)\,|\{K\in\mathcal{K}_d(F):f\in E(K)\}|,\end{aligned}$$

as required. □

5 Upper bounds for alternating paths from a fixed vertex

We now prove the following lemma, which gives a good upper bound on the number of $(F\setminus K,K)$-alternating paths starting from a fixed vertex. (That is, unlike our ultimate goal of **C3**, one of the endpoints of the paths considered is not fixed). As with the next two sections, we consider the regime corresponding to $m\le n^{1+\sigma}$ in Theorem 2.4, for some small fixed $\sigma>0$, but we prove that one of the properties, **G1** in Lemma 5.3, holds more widely so that we can use it in Section 8 and also to conclude that **C1** is likely to hold. As with our results across the next three sections, instead of working in $F(n,d,m)$ here we work with the slightly easier to study random graph F formed by taking $K\sim K_d(n)$ and adding edges independently at random with some probability p chosen so that the expected number of added edges is m. When we apply the main results of these sections, in Section 9, at the cost of a manageable increase in the probability the events considered hold, we will easily be able to translate these results into the model $F(n,d,m)$.

Lemma 5.1. *Let $1/n\ll 1/C\ll\varepsilon\le 1$ and $1/C\ll\sigma,\mu\ll 1$. Let $d\in[n]$ and $m\in\mathbb{N}$ satisfy $d\ge C\log n$ and $\varepsilon dn/2\le m\le n^{1+\sigma}$. Let $\delta=2m/n$ and let*

$$\lambda=\max\left\{\frac{\mu^2}{\log^2 n},\left(\frac{n\log n}{m}\right)^{1/4}\right\}.$$

Let $p=m(\binom{n}{2}-dn/2)^{-1}$. Let $K\in\mathcal{K}_d(n)$, form E by including each element of $[n]^{(2)}\setminus E(K)$ uniformly at random with probability p, and let $F=K+E$.

Then, with probability $1-o(n^{-14})$, we have the following.

E *For each $1\le\ell\le 10\log n$ and $v\in[n]$, the number of $(F\setminus K,K)$-alternating paths with length 2ℓ starting at v is at most $(1+30\lambda)d^\ell\delta^\ell$.*

In Section 5.1 we give pseudorandom conditions (**F1** and **F2**) and show they imply **E**. These conditions concern the degrees in $F\setminus K$, as well as certain sums of degrees, and we prove the likely bounds we need in Section 5.2. Finally, we use this to conclude that Lemma 5.1 holds in Section 5.3.

5.1 Pseudorandom properties for upper bounds for paths from a vertex

To show the number of $(F \setminus K, K)$-alternating paths from a vertex are bounded we will use bounds on the degrees of $F \setminus K$ and, as discussed in Section 2.2.2, a more complicated condition on certain degree sums required for the accuracy we need. Due to this more complicated condition, we prove the lemma by induction by considering not the number of paths, but the sum of the degrees in $F \setminus K$ of the endpoints of the paths we wish to consider.

Lemma 5.2. *Let $0 < \lambda \ll 1$. Let $d, \delta > 0$, let $K \in \mathcal{K}_d(n)$ and let F be a graph with vertex set $[n]$ for which the following properties hold.*

F1 *For each $v \in [n]$, $d_{F\setminus K}(v) \leq (1+\lambda)\,\delta$.*

F2 *For each $v \in [n]$,*

$$\sum_{u \in N_{F\setminus K}(v)} \sum_{u' \in N_K(u)} d_{F\setminus K}(u') \leq \left(1 + \frac{\lambda}{\log n}\right) \delta d \cdot d_{F\setminus K}(v).$$

Then, the following holds.

F3 *For each $1 \leq \ell \leq 10 \log n$ and $v \in [n]$, the number of $(F \setminus K, K)$-alternating paths with length 2ℓ starting at v is at most $(1+30\lambda) d^\ell \delta^\ell$.*

Proof. Let $v' \in [n]$. We will first show by induction that, for each $1 \leq i \leq 10 \log n$,

$$\sum_{u \in V(G)} \#\{(K, F\setminus K)\text{-alternating } v', u\text{-paths length } 2i-1\} \cdot d_{F\setminus K}(u) \leq (1+\lambda) \cdot \left(1 + \frac{\lambda}{\log n}\right)^{i-1} \delta^i d^i. \tag{21}$$

For this, first note that (21) holds for $i = 1$ by **F1**, using that there are d vertices $u \in V(G)$ for which there is a $(K, F \setminus K)$-alternating v', u-path of length 1. Suppose then that $1 < i \leq 10 \log n$ and that (21) holds with i replaced by $i-1$. Then,

$$\begin{aligned}
&\sum_{u \in V(G)} \#\{(K, F\setminus K)\text{-alternating } v', u\text{-paths of length } 2i-1\} \cdot d_{F\setminus K}(u) \\
&\quad \leq \sum_{u' \in V(G)} \#\{(K, F\setminus K)\text{-alternating } v', u'\text{-paths of length } 2i-3\} \cdot \sum_{u'' \in N_{F\setminus K}(u')} \sum_{u \in N_K(u'')} d_{F\setminus K}(u) \\
&\quad \overset{\mathbf{F2}}{\leq} \left(1 + \frac{\lambda}{\log n}\right) \cdot \delta d \cdot \sum_{u' \in V(G)} \#\{(K, F\setminus K)\text{-alternating } v', u'\text{-paths length } 2i-3\} \cdot d_{F\setminus K}(u') \\
&\quad \leq (1+\lambda) \cdot \left(1 + \frac{\lambda}{\log n}\right)^{i-1} \delta^i d^i.
\end{aligned}$$

Therefore, (21) holds for i. Thus, (21) holds for all $1 \leq i \leq 10 \log n$.

For each $1 \leq i \leq 10 \log n$, it therefore follows from (21) that, for each $v' \in [n]$, the number of $(K, F \setminus K)$-alternating paths with length $2i$ starting at v' is at most

$$(1+\lambda) \cdot \left(1 + \frac{\lambda}{\log n}\right)^{i-1} \delta^i d^i \leq \left(1 + 2\lambda + \frac{2(i-1)\lambda}{\log n}\right) \delta^i d^i \leq (1+30\lambda)\,\delta^i d^i,$$

where we have used that $\lambda \ll 1$. Thus, **F3** holds. □

5.2 Likely properties of $F \setminus K$: degree bounds

We now prove a result we mainly use to deduce the pseudorandom conditions used in Lemma 5.2. Note that, as mentioned at the start of the section, the first of these properties is proved for (d, m) more widely so that it can be applied in Section 8 and to show that **C1** is likely to hold.

Lemma 5.3. *Let $1/n \ll 1/C \ll 1/C', \sigma \ll 1$ and $1/C \ll \varepsilon$. Let $d \geq C \log n$ and let m satisfy $\varepsilon dn/2 \leq m \leq \binom{n}{2} - dn/2$. Let $\delta = 2m/n$ and let $p = m\left(\binom{n}{2} - dn/2\right)^{-1}$. Let $K \in \mathcal{K}_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$. Then, the following holds with probability $1 - o(n^{-14})$.*

G1 *For each $v \in [n]$,*

$$\left(1 - C'\sqrt{\frac{\log n}{\delta}}\right)\delta \leq d_{F\setminus K}(v) \leq \left(1 + C'\sqrt{\frac{\log n}{\delta}}\right)\delta.$$

Furthermore, if $d \leq n^{\sigma}$ and $m \leq n^{1+\sigma}$, the following holds with probability $1 - o(n^{-14})$.

G2 *For each $v \in [n]$,*

$$\left(1 - \frac{2}{\delta}\right) d_{F\setminus K}(v) \cdot \delta d \leq \sum_{u \in N_{F\setminus K}(v)} \sum_{u' \in N_K(u)} d_{F\setminus K}(u') \leq \left(1 + \frac{2}{\delta}\right) d_{F\setminus K}(v) \cdot \delta d.$$

Proof. For each $v \in [n]$, by an application of Chernoff's bound, as $d_{F\setminus K}(v)$ is binomially distributed with parameters $n - 1 - d$ and p, and $p(n - 1 - d) = 2m/n = \delta$, we have that

$$\mathbb{P}\left(d_{F\setminus K}(v) \neq \left(1 \pm C'\sqrt{\log n/\delta}\right)\delta\right) \leq 2\exp\left(-\frac{C'^2 \log n}{3\delta} \cdot \delta\right) = o(n^{-15}). \tag{22}$$

Thus, taking a union bound over all $v \in [n]$, with probability $1 - o(n^{-14})$, **G1** holds, as desired.

Let, then, $v \in [n]$ and suppose $d \leq n^{\sigma}$ and $m \leq n^{1+\sigma}$. Reveal the edges of E next to v, and let $A = N_{F\setminus K}(v)$. Note that the probability that there is a vertex $u \in [n]$ with $|N_K(u) \cap A| \geq 100$ is, as $d \leq n^{\sigma}$ and $p \leq 4m/n^2 \leq 4n^{\sigma-1}$, at most

$$n \cdot \binom{d}{100} p^{100} \leq n(pd)^{100} = o(n^{-15}).$$

Thus, with probability $1 - o(n^{-15})$, for each vertex $u \in [n]$, $|N_K(u) \cap A| < 100$. Furthermore, by (22), we have $\delta/2 \leq |A| \leq 2\delta$ with probability $1 - o(n^{-15})$.

Let $\hat{E}_v$ be the set of edges $e \in ([n] \setminus \{v\})^{(2)} \setminus E(K)$ which contain at least one vertex in A. For each $xy \in \hat{E}_v$, let $w_{xy} = |N_K(x) \cap A| + |N_K(y) \cap A|$, so that $w_{xy} \leq 200$ as $|N_K(u) \cap A| < 100$ for each $u \in [n]$. Note that

$$\sum_{xy \in \hat{E}_v} w_{xy} = \sum_{u \in A} \sum_{u' \in N_K(u)} d_{(K_n \setminus K)-v}(u') = \left(1 \pm \frac{1}{4\delta}\right)|A| d(n - 1 - d), \tag{23}$$

where we have used that, for each $u' \in [n]$, we have $d_{(K_n\setminus K)-v}(u') \in \{n-2-d, n-1-d\}$. Therefore, we can take subsets $\hat{E}_{v,i}$, $i \in [200]$, in $\hat{E}_v$ such that $|\hat{E}_{v,i}| \geq |A|d(n-1-d)/400$ and, for each $xy \in \hat{E}_v$, $|\{i \in [200] : xy \in \hat{E}_{v,i}\}| = w_{xy}$. For example, if we randomly assign each $xy \in \hat{E}_v$ to w_{xy} of the $\hat{E}_{v,i}$, $i \in [200]$, for each $xy \in \hat{E}_v$. Then, for each $i \in [200]$, we have $\mathbb{E}|\hat{E}_{v,i}| = \sum_{xy \in \hat{E}_v} \frac{w_{xy}}{200}$, and by

Theorem 2.5 there is a positive probability that $|\hat{E}_{v,i}| \geq 2\mathbb{E}|\hat{E}_{v,i}|/3 \geq |A|d(n-1-d)/400$ for all $i \in [200]$, where we have used (23).

For each $i \in [200]$, let $X_{v,i} = |\hat{E}_{v,i} \cap E|$. Revealing the edges of E not containing v, we have that, for each $i \in [200]$, $X_{v,i}$ is a binomial random variable with mean $\mathbb{E}|X_{v,i}| = p|\hat{E}_{v,i}| \geq |A|dp(n-1-d)/400 \geq \delta^2 d/800$. Therefore, we have, via a Chernoff bound, that, for each $i \in [200]$,

$$\mathbb{P}\left(|X_{v,i} - \mathbb{E}X_{v,i}| > \frac{1}{10\delta} \cdot \mathbb{E}X_{v,i}\right) \leq 2\exp\left(-\frac{\mathbb{E}X_{v,i}}{3 \cdot 100\delta^2}\right) \leq 2\exp\left(-\frac{d}{10^6}\right) = o(n^{-15}), \tag{24}$$

where we have used that $d \geq C\log n$.

Let

$$X_v = \sum_{u \in A} \sum_{u' \in N_K(u)} d_{(F\setminus K)-v}(u') = \sum_{i \in [200]} X_{v,i},$$

Then, by (24), with probability $1 - o(n^{-15})$, we have $|X_{v,i} - \mathbb{E}X_{v,i}| \leq \frac{1}{10\delta} \cdot \mathbb{E}X_{v,i}$ for each $i \in [200]$ and, hence,

$$|X_v - \mathbb{E}X_v| \leq \sum_{i \in [200]} (|X_{v,i} - \mathbb{E}X_{v,i}|) \leq \sum_{i \in [200]} \frac{1}{10\delta} \cdot \mathbb{E}X_{v,i} = \frac{1}{10\delta} \cdot \mathbb{E}X_v. \tag{25}$$

Furthermore, by (23), we have that $\mathbb{E}X_v = \left(1 \pm \frac{1}{4\delta}\right)|A|dp(n-d-1) = \left(1 \pm \frac{1}{4\delta}\right)|A| \cdot d\delta$. Thus, (25) implies that $X_v = \left(1 \pm \frac{1}{2\delta}\right)|A| \cdot \delta d = \left(1 \pm \frac{1}{2\delta}\right) d_{F\setminus K}(v) \cdot \delta d$.

Noting that

$$\left| X_v - \sum_{u \in N_{F\setminus K}(v)} \sum_{u' \in N_K(u)} d_{F\setminus K}(u') \right| \leq \sum_{u \in N_{F\setminus K}(v)} \sum_{u' \in N_K(u)} 1 \leq d \cdot d_{F\setminus K}(v) = \frac{1}{\delta} \cdot d_{F\setminus K}(v) \cdot \delta d,$$

we have that **G2** holds for v with probability $1 - o(n^{-15})$. Taking a union bound over all $v \in [n]$ thus shows that **G2** holds with probability $1 - o(n^{-14})$, as desired. □

5.3 Proof of Lemma 5.1

Combining Lemma 5.2 and Lemma 5.3, it is now easy to prove Lemma 5.1.

Proof of Lemma 5.1. Take the variables as set up in Lemma 5.1, so that $1/n \ll 1/C \ll \varepsilon \leq 1$, $1/C \ll \sigma, \mu \ll 1$, $d \geq C\log n$, $\varepsilon dn/2 \leq m \leq n^{1+\sigma}$, $\delta = 2m/n$ and $p = m\left(\binom{n}{2} - dn/2\right)^{-1}$. Furthermore, we have

$$\lambda = \max\left\{\frac{\mu^2}{\log^2 n}, \left(\frac{n\log n}{m}\right)^{1/4}\right\}. \tag{26}$$

Let $K \sim G_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$. Thus, we want to show that **E** holds with probability $1 - o(n^{-14})$.

Let C' satisfy $1/C \ll 1/C' \ll 1$. As $\delta = 2m/n \geq \varepsilon C\log n$, we have that

$$C'\sqrt{\frac{\log n}{\delta}} \overset{(26)}{\leq} C'\sqrt{\frac{\log n}{\delta}} \cdot \lambda\left(\frac{m}{n\log n}\right)^{1/4} \leq C'\sqrt{\frac{\log n}{\delta}} \cdot \lambda\left(\frac{\delta}{2\log n}\right)^{1/4} = C'\lambda\left(\frac{\log n}{2\delta}\right)^{1/4} \leq \lambda.$$

Also, since $m \geq \varepsilon Cn\log n/2$, we have

$$\frac{2}{\delta} = \frac{n}{m} = \left(\frac{n\log n}{m}\right)^{3/4} \cdot \frac{1}{\log n}\left(\frac{n\log n}{m}\right)^{1/4} \overset{(26)}{\leq} \frac{\lambda}{\log n}.$$

Thus, by Lemma 5.3, we have that **F1** and **F2** hold with probability $1 - o(n^{-14})$. Then, by Lemma 5.2, we have that **E** holds with probability $1 - o(n^{-14})$. □

6 Lower bounds for switching paths

We now prove our likely lower bound on the number of $(F \setminus K, K)$-alternating paths between two fixed vertices, that appears as **C2** in Theorem 4.1, as follows.

Theorem 6.1. *Let $1/n \ll 1/C \ll 1/C_0 \ll \varepsilon \leq 1$ and $1/C \ll \sigma, \mu \ll 1$. Let $d \in [n]$ and $m \in \mathbb{N}$ satisfy $d \geq C \log n$ and $\varepsilon dn/2 \leq m \leq n^{1+\sigma}$. Let $\delta = 2m/n$ and let*

$$\lambda = \max\left\{\frac{\mu^2}{\log^2 n}, \left(\frac{n \log n}{m}\right)^{1/4}\right\}.$$

Let $p = m(\binom{n}{2} - dn/2)^{-1}$. Let $K \sim G_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$. Let ℓ_0 be the smallest integer for which $\delta^{(\ell_0-1)} d^{(\ell_0-1)} \geq n$. Then, with probability $1 - o(n^{-14})$, we have the following.

H *For each distinct $x, y \in [n]$, each $Z \subset [n]$ with $|Z| \leq 40 \log n$, and each ℓ with $10\ell_0 \leq \ell \leq 10 \log n$ the number of $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ which do not have any internal vertices in Z is at least $\left(1 - C_0\lambda - \frac{|(Z \cup \{x\}) \cap N_K(y)|}{d}\right) d^\ell \delta^\ell / n$.*

In Section 6.1, we give pseudorandom conditions (**I1**–**I6**) and show they imply **H**. Of these pseudorandom conditions, **I1** and **I2** will already follow from Lemma 5.3. The conditions **I3**, **I4** and **I5** will be used to show expansion conditions in the graphs K and $F \setminus K$, and we prove **I3** is likely to hold in Section 6.2 and **I4** and **I5** is likely to hold in Section 6.3. The condition **I6** concerns edges in $F \setminus K$ between large sets and is shown to likely hold in Section 6.4. Finally, we put this together in Section 6.5 to prove Theorem 6.1.

6.1 Pseudorandom properties for lower bounds for path counts

To show good bounds on the expansion of certain vertex sets U in a graph H (applied with $H = K$ and $H = F \setminus K$ for Theorem 6.1), we will use that few vertices have more than one neighbour in H in U, so that the size of $N_H(U)$ is close to the sum of the degrees in H of the vertices in U. For this, we will use the following convenient definition.

Definition 6.2. For a vertex set U in a graph F, we define $E_{\geq 2,F}(U)$ to be the set of edges $xy \in E(F)$ with $x \in U$, $y \notin U$ for which there is some $x' \in U \setminus \{x\}$ with $x'y \in E(F)$.

We now give our pseudorandom conditions for bounding below the number of $(F \setminus K, K)$-alternating paths between two fixed vertices, and deduce the bound we need, as follows.

Lemma 6.3. *Let $1/n \ll 1/C \ll 1/C_0 \ll 1$ and $1/C \ll \sigma \ll 1$. Let $d \geq C \log n$ with $d, \delta \leq n^\sigma$. Let $\log^{-3} n \leq \lambda \ll 1$ satisfy $\lambda\delta \geq C_0 \log n$. Let $K \in \mathcal{K}_d(n)$ and let F be a graph with vertex set $[n]$, such that $K \subset F$ and the following properties hold.*

I1 *For each $v \in [n]$, $d_{F \setminus K}(v) = (1 \pm \lambda)\,\delta$.*

I2 *For each $v \in [n]$,*

$$\sum_{u \in N_{F \setminus K}(v)} \sum_{u' \in N_K(u)} d_{F \setminus K}(u') = \left(1 \pm \frac{\lambda}{\log n}\right) d_{F \setminus K}(v) \cdot \delta d.$$

I3 *For each $U \subset [n]$, if there is some $U' \subset [n]$ with $|U'| \leq 4|U|/\delta$ and $U \subset U' \cup N_{F\setminus K}(U')$, then*

$$|E_{\geq 2,K}(U)| + e_K(U) \leq \begin{cases} \frac{\lambda}{\log n} d|U| & \text{if } |U| \leq \frac{\lambda n}{10^6 d \log n} \\ \lambda d|U| & \text{if } |U| \leq \frac{\lambda n}{10^6 d}. \end{cases}$$

I4 *For each $U \subset [n]$ with $|U| \leq \lambda n/100\delta \log n$, if there is some $U' \subset [n]$ with $|U'| \leq 4|U|/d$ and $U \subset U' \cup N_K(U')$, then*

$$|E_{\geq 2,F\setminus K}(U)| + e_{F\setminus K}(U) \leq \frac{\lambda}{\log n}\delta|U|.$$

I5 *For each $U \subset [n]$ with $|U| \leq 100 \log n$, $|\{v \in [n] \setminus U : d_{F\setminus K}(v, U) \geq 10\}| \leq 100 \log n$.*

I6 *For any disjoint sets $U, V \subset [n]$ with $|U|, |V| \geq \frac{\lambda n}{10^8}$, we have*

$$e_{F\setminus K}(U, V) \geq (1 - \lambda)\,\frac{\delta}{n}|U||V|.$$

Let ℓ_0 be the smallest integer for which $\delta^{(\ell_0-1)}d^{(\ell_0-1)} \geq n$. Then, the following holds.

I7 *For each distinct $x, y \in [n]$, each $Z \subset [n]$ with $|Z| \leq 40 \log n$, and each $10\ell_0 \leq \ell \leq 10 \log n$ the number of $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ which do not have any internal vertices in Z is at least $\left(1 - C_0\lambda - \frac{|(Z\cup\{x\})\cap N_K(y)|}{d}\right) d^\ell \delta^\ell/n$.*

Proof. Let ℓ_1 be the smallest integer for which $\delta^{(\ell_1+1)}d^{(\ell_1+1)} \geq \lambda n/(2\cdot 10^7)$, so that $\ell_0 - 3 \leq \ell_1 \leq \ell_0 - 2$ as $\lambda\delta \geq C_0 \log n$. As $d, \delta \leq n^\sigma$, we have that $\ell_1 \geq 1$, and as $d \geq C \log n$ and $\delta \geq C_0 \log n$ we have $\ell_1 \leq \log n/\log\log n$. Let $r = \lceil \lambda n/(8 \cdot 10^7 \cdot \delta^{\ell_1} d^{\ell_1})\rceil \leq \delta d/2$. We will show the following claim.

Claim 6.4. Let $X_0, Y_0 \subset [n]$ be disjoint with $|X_0| = |Y_0| = r$ such that there are disjoint $X_0', Y_0' \subset [n]$ with $X_0 \subset N_K(X_0')$, $Y_0 \subset N_K(Y_0')$ and $|X_0'|, |Y_0'| \leq 4r/d$. Then, for each $W \subset [n]$ with $|W| \leq 3r + \lambda\delta$, there are at least $(1 - C_0\lambda/4)r^2\delta^{2\ell_1+1}d^{2\ell_1}/n$ $(F \setminus K, K)$-alternating paths of length $4\ell_1 + 1$ which start in X_0, end in Y_0, and have no internal vertices in $X_0 \cup Y_0 \cup W$.

Proof of Claim 6.4. Iteratively, for each $1 \leq i \leq \ell_1$, let

$$X_{2i-1} = N_{F\setminus K}(X_{2i-2}) \setminus \left(\left(\cup_{j=0}^{2i-2}(X_j \cup Y_j)\right) \cup N_{F\setminus K}(Y_{2i-2}) \cup W\right),$$

$$Y_{2i-1} = N_{F\setminus K}(Y_{2i-2}) \setminus \left(\left(\cup_{j=0}^{2i-2}(X_j \cup Y_j)\right) \cup N_{F\setminus K}(X_{2i-2}) \cup W\right),$$

$$X_{2i} = N_K(X_{2i-1}) \setminus \left(\left(\cup_{j=0}^{2i-1}(X_j \cup Y_j)\right) \cup N_K(Y_{2i-1}) \cup W\right),$$

and

$$Y_{2i} = N_K(Y_{2i-1}) \setminus \left(\left(\cup_{j=0}^{2i-1}(X_j \cup Y_j)\right) \cup N_K(X_{2i-1}) \cup W\right).$$

Note that, for each edge in $F \setminus K$ between $X_{2\ell_1}$ and $Y_{2\ell_1}$, there is an $(F \setminus K, K)$-alternating path of length $4\ell_1 + 1$ which starts in X_0 and ends in Y_0, and has no internal vertices in $X_0 \cup Y_0 \cup W$, and each such path is different. Thus, to prove the claim, it is sufficient to prove that $e_{F\setminus K}(X_{2\ell_1}, Y_{2\ell_1}) \geq (1 - C_0\lambda/4) \cdot r^2\delta^{2\ell_1+1}d^{2\ell_1}/n$.

We will show by induction that, for each $1 \le i \le \ell_1$,

$$\left(1 - \frac{20\lambda}{\log n} - 20\lambda \cdot \mathbf{1}_{\{i=1\}}\right) \cdot \sum_{v \in X_{2i-2}} d_{F\setminus K}(v) \le |X_{2i-1}| \le \sum_{v \in X_{2i-2}} d_{F\setminus K}(v), \tag{27}$$

$$\left(1 - \frac{20\lambda}{\log n} - 20\lambda \cdot \mathbf{1}_{\{i=1\}}\right) \cdot \sum_{v \in Y_{2i-2}} d_{F\setminus K}(v) \le |Y_{2i-1}| \le \sum_{v \in Y_{2i-2}} d_{F\setminus K}(v), \tag{28}$$

$$\left(1 - \frac{20\lambda}{\log n} - 20\lambda \cdot \mathbf{1}_{\{i\in\{1,\ell_1\}\}}\right) \cdot d\delta \le \frac{\sum_{v \in X_{2i}} d_{F\setminus K}(v)}{\sum_{v \in X_{2i-2}} d_{F\setminus K}(v)} \le \left(1 + \frac{20\lambda}{\log n} + 20\lambda \cdot \mathbf{1}_{\{i\in\{1,\ell_1\}\}}\right) \cdot d\delta, \tag{29}$$

and

$$\left(1 - \frac{20\lambda}{\log n} - 20\lambda \cdot \mathbf{1}_{\{i\in\{1,\ell_1\}\}}\right) \cdot d\delta \le \frac{\sum_{v \in Y_{2i}} d_{F\setminus K}(v)}{\sum_{v \in Y_{2i-2}} d_{F\setminus K}(v)} \le \left(1 + \frac{20\lambda}{\log n} + 20\lambda \cdot \mathbf{1}_{\{i\in\{1,\ell_1\}\}}\right) \cdot d\delta. \tag{30}$$

For this, let $1 \le i \le \ell_1$ and assume that (27)–(30) hold for each $1 \le a < i$.

We start by showing that, for each $1 \le a < i$,

$$\frac{4}{5} d^a \delta^a r \le |X_{2a}|, |Y_{2a}| \le \frac{6}{5} d^a \delta^a r. \tag{31}$$

As $|X_0| = r$, using **I1** we have that $(1-\lambda)\delta r \le \sum_{v \in X_0} d_{F\setminus K}(v) \le (1+\lambda)\delta r$. Thus, by (29) applied for each $1 \le a \le i$, we have, as $i \le \ell_1 = o(\log n)$ and $\lambda \ll 1$, that

$$\sum_{v \in X_{2a}} d_{F\setminus K}(v) \ge \left(1 - \frac{20\lambda}{\log n} - 20\lambda\right) \cdot \left(1 - \frac{20\lambda}{\log n}\right)^{a-1} d^a \delta^a \cdot (1-\lambda)\delta r \ge \frac{9}{10} d^a \delta^{a+1} r \tag{32}$$

and

$$\sum_{v \in X_{2a}} d_{F\setminus K}(v) \le \left(1 + \frac{20\lambda}{\log n} + 20\lambda\right) \cdot \left(1 + \frac{20\lambda}{\log n}\right)^{a-1} d^a \delta^a \cdot (1+\lambda)\delta r \le \frac{11}{10} d^a \delta^{a+1} r. \tag{33}$$

Using **I1**, we therefore have that (31) holds for $|X_{2a}|$. Similarly, (31) holds for $|Y_{2a}|$.

Now, for each $1 \le a < i$, using $|X_0| = |Y_0| = r$ when $a = 1$ and otherwise (32) and (33) for $i = a-1$, by (27) with $i = a$

$$\frac{4}{5} d^{a-1} \delta^a r \le |X_{2a-1}|, |Y_{2a-1}| \le \frac{6}{5} d^{a-1} \delta^a r. \tag{34}$$

If $i = 1$, then set $X_{-1} = X_0'$ and $Y_{-1} = Y_0'$ (using the sets in the claim statement). Whether $i = 1$, or not, we then have that $X_{2i-2} \cup Y_{2i-2} \subset N_K(X_{2i-3} \cup Y_{2i-3})$, and, from the properties of X_0' and Y_0' or by (31) and (34), $|X_{2i-3} \cup Y_{2i-3}| \le 3\delta^{i-1} d^{i-2} r \le 4|X_{2i-2} \cup Y_{2i-2}|/d$. Furthermore, either as $|X_0| = |Y_0| = r$ or by (31) with $a = i-1$, and as $\delta^{\ell_1} d^{\ell_1} < \lambda n/(2 \cdot 10^7)$ by the choice of ℓ_1, we have that

$$|X_{2i-2} \cup Y_{2i-2}| \le 2 \cdot \frac{6}{5} d^{i-1} \delta^{i-1} r < 2 \cdot \frac{\lambda n}{10^7} \cdot d^{i-1-\ell_1} \delta^{i-1-\ell_1} \le \frac{\lambda n}{\delta d} \le \frac{\lambda n}{100 \delta \log n}. \tag{35}$$

Thus, from **I4**, we have,

$$|E_{\geq 2,F\setminus K}(X_{2i-2}\cup Y_{2i-2})| \leq \frac{\lambda}{\log n}\cdot\delta|X_{2i-2}| \leq \frac{\lambda}{\log n}\cdot 2\delta^i d^{i-1}r. \tag{36}$$

Furthermore, we have $\cup_{a=0}^{2i-2}(X_a\cup Y_a)\subset\left(\cup_{a=0}^{2i-3}(X_a\cup Y_a)\right)\cup N_K\left(\cup_{a=0}^{2i-3}(X_a\cup Y_a)\right)$, and, using (31) and (34),

$$|\cup_{a=0}^{2i-3}(X_a\cup Y_a)| \leq 2\cdot\frac{6}{5}\sum_{a=0}^{2i-3}\delta^{\lceil a/2\rceil}d^{\lfloor a/2\rfloor}r \leq 3\delta^{i-1}d^{i-2}r \leq \frac{4}{d}|X_{2i-2}\cup Y_{2i-2}| \leq \frac{4}{d}|\cup_{a=0}^{2i-2}(X_a\cup Y_a)|.$$

Similarly to (35), we have $|\cup_{a=0}^{2i-2}(X_a\cup Y_a)|\leq\lambda n/100\delta\log n$. Thus, from **I4**, we have

$$e_{F\setminus K}\left(\cup_{a=0}^{2i-2}(X_a\cup Y_a)\right) \leq \frac{\lambda}{\log n}\cdot\delta\left|\cup_{a=0}^{2i-2}(X_a\cup Y_a)\right| \leq \frac{\lambda}{\log n}\cdot 3\delta^i d^{i-1}r, \tag{37}$$

where we have used (31) and (34). Therefore, as $\lambda\delta\geq C_0\log n$, this implies that

$$|W| \leq 3r+\lambda\delta \leq \left(\frac{\lambda}{\log n}+\lambda\cdot\mathbf{1}_{\{i=1\}}\right)\cdot\delta^i d^{i-1}r, \tag{38}$$

we have that

$$\begin{aligned}\left(\sum_{v\in X_{2i-2}} d_{F\setminus K}(v)\right)-|X_{2i-1}| &\leq |E_{\geq 2,F\setminus K}(X_{2i-2}\cup Y_{2i-2})| + e_{F\setminus K}\left(\cup_{i'=0}^{2i-2}(X_{i'}\cup Y_{i'})\right)+|W| \\ &\overset{(36),(37),(38)}{\leq} \left(\frac{\lambda}{\log n}+\lambda\cdot\mathbf{1}_{\{i=1\}}\right)\cdot 6\delta^i d^{i-1}r \\ &\overset{(32)}{\leq} \left(\frac{\lambda}{\log n}+\lambda\cdot\mathbf{1}_{\{i=1\}}\right)\cdot 8\cdot\sum_{v\in X_{2i-2}} d_{F\setminus K}(v). \end{aligned}\tag{39}$$

From (39), we get the lower-bound in (27) holds. As $|X_{2i-1}|\leq|N_{F\setminus K}(X_{2i-2})|\leq\sum_{v\in X_{2i-2}}d_{F\setminus K}(v)$, we therefore have that (27) holds entirely. Furthermore, by a similar deduction to those above using **I1**, we have that $\frac{4}{5}d^{i-1}\delta^i r\leq|X_{2i-1}|\leq\frac{6}{5}d^{i-1}\delta^i r$. Similarly, we can deduce that

$$\left(\sum_{v\in Y_{2i-2}} d_{F\setminus K}(v)\right)-|Y_{2i-1}| \leq \left(\frac{\lambda}{\log n}+\lambda\cdot\mathbf{1}_{\{i=1\}}\right)\cdot 8\cdot\sum_{v\in Y_{2i-2}} d_{F\setminus K}(v), \tag{40}$$

that (28) holds, and that $\frac{4}{5}d^{i-1}\delta^i r\leq|Y_{2i-1}|\leq\frac{6}{5}d^{i-1}\delta^i r$.

Now, working similarly to above, we have that $X_{2i-1}\cup Y_{2i-1}\subset N_{F\setminus K}(X_{2i-2}\cup Y_{2i-2})$ and $|X_{2i-2}\cup Y_{2i-2}|\leq 4|X_{2i-1}\cup Y_{2i-1}|/\delta$. Furthermore, we have that

$$|X_{2i-1}\cup Y_{2i-1}| \leq 2\cdot\frac{6}{5}d^{i-1}\delta^i r \leq \frac{2\lambda n}{10^7}\cdot d^{i-1-\ell_1}\delta^{i-\ell_1} \leq \begin{cases}\frac{\lambda n}{10^6 d\log n} & \text{if } i<\ell_1,\\ \frac{\lambda n}{10^6 d} & \text{if } i=\ell_1.\end{cases}$$

Thus, from **I3**, we have

$$\begin{aligned}|E_{\geq 2,K}(X_{2i-1}\cup Y_{2i-1})| &\leq \left(\frac{\lambda}{\log n}+\lambda\cdot\mathbf{1}_{\{i=\ell_1\}}\right)\cdot d|X_{2i-1}\cup Y_{2i-1}| \\ &\leq \left(\frac{\lambda}{\log n}+\lambda\cdot\mathbf{1}_{\{i=\ell_1\}}\right)\cdot 3\delta^i d^i r. \end{aligned}\tag{41}$$

Furthermore, we have that $\cup_{a=0}^{2i-1}(X_a \cup Y_a) \subset \left(\cup_{a=0}^{2i-2}(X_a \cup Y_a)\right) \cup N_{F\setminus K}\left(\cup_{a=0}^{2i-2}(X_a \cup Y_a)\right)$, and $|\cup_{a=0}^{2i-2}(X_a \cup Y_a)| \le 3d^{i-1}\delta^{i-1}r \le 4|\cup_{a=0}^{2i-1}(X_a \cup Y_a)|/\delta$, while

$$\left|\cup_{a=0}^{2i-1}(X_a \cup Y_a)\right| \le 3d^{i-1}\delta^i r \le \frac{3\lambda n}{2\cdot 10^7}\cdot d^{i-1-\ell_1}\delta^{i-\ell_1} \le \begin{cases} \frac{\lambda n}{10^6 d\log n} & \text{if } i<\ell_1, \\ \frac{\lambda n}{10^6 d} & \text{if } i=\ell_1. \end{cases}$$

Thus, from **I3**, we have

$$\begin{aligned} e_K\left(\cup_{a=0}^{2i-1}(X_a \cup Y_a)\right) &\le \left(\frac{\lambda}{\log n} + \lambda\cdot \mathbf{1}_{\{i=\ell_1\}}\right)\cdot d\left|\cup_{a=0}^{2i-1}(X_a \cup Y_a)\right| \\ &\le \left(\frac{\lambda}{\log n} + \lambda\cdot \mathbf{1}_{\{i=\ell_1\}}\right)\cdot 3\delta^i d^i r. \end{aligned} \tag{42}$$

Therefore, we have that

$$\begin{aligned} d|X_{2i-1}| - |X_{2i}| &\le |E_{\ge 2,K}(X_{2i-1}\cup Y_{2i-1})| + e_K\left(\cup_{a=0}^{2i-1}(X_a \cup Y_a)\right) + |W| \\ &\overset{(41),(42),(38)}{\le} \left(\frac{\lambda}{\log n} + \lambda\cdot \mathbf{1}_{\{i\in\{1,\ell_1\}\}}\right)\cdot 7\delta^i d^i r \\ &\le \left(\frac{\lambda}{\log n} + \lambda\cdot \mathbf{1}_{\{i\in\{1,\ell_1\}\}}\right)\cdot 8\cdot d\cdot \sum_{v\in X_{2i-2}} d_{F\setminus K}(v). \end{aligned} \tag{43}$$

Now,

$$\sum_{v\in X_{2i}} d_{F\setminus K}(v) \le \sum_{u\in X_{2i-2}}\sum_{w\in N_{F\setminus K}(u)}\sum_{v\in N_K(w)} d_{F\setminus K}(v) \overset{\mathbf{I2}}{\le} \left(1+\frac{\lambda}{\log n}\right)\cdot \delta d\cdot \sum_{v\in X_{2i-2}} d_{F\setminus K}(v),$$

and

$$\begin{aligned} \sum_{v\in X_{2i}} d_{F\setminus K}(v) &\overset{\mathbf{I1}}{\ge} \left(\sum_{u\in X_{2i-2}}\sum_{w\in N_{F\setminus K}(u)}\sum_{v\in N_K(w)} d_{F\setminus K}(v)\right) - (1+\lambda)\delta d\cdot\left(\sum_{v\in X_{2i-2}} d_{F\setminus K}(v) - |X_{2i-1}|\right) \\ &\qquad - (1+\lambda)\delta(d|X_{2i-1}| - |X_{2i}|) \\ &\overset{\mathbf{I2},(39),(43)}{\ge} \left(1-\frac{\lambda}{\log n}\right)\cdot \delta d\cdot \sum_{v\in X_{2i-2}} d_{F\setminus K}(v) - \delta d\cdot\left(\frac{\lambda}{\log n} + \lambda\cdot \mathbf{1}_{\{i\in\{1,\ell_1\}\}}\right)\cdot 19\cdot \sum_{v\in X_{2i-2}} d_{F\setminus K}(v). \end{aligned}$$

Thus, we have that (29) holds. Similarly (with (40) in place of (39)), (30) holds.

Therefore, by induction, (27)–(30) hold for each $1\le i\le \ell_1$. Applying (29) for each $1\le i\le \ell_1$, and using that $|\sum_{v\in X_0} d_{F\setminus K}(v)| \ge (1-\lambda)\delta r$ by **I1**, we have

$$|X_{2\ell_1}| \ge \frac{\sum_{v\in X_{2\ell_1}} d_{F\setminus K}(v)}{(1+\lambda)\delta} \ge \frac{(1-40\lambda)\cdot(1-\lambda)d^{\ell_1}\delta^{\ell_1+1}r}{(1+\lambda)\delta} \ge (1-45\lambda)d^{\ell_1}\delta^{\ell_1}r \ge \frac{\lambda n}{10^8},$$

where we have used **I1** again and $\ell_1 = o(\log n)$. As, similarly, $|Y_{2\ell_1}| \ge (1-45\lambda)d^{\ell_1}\delta^{\ell_1}r \ge \lambda n/10^8$, we have, by **I6**, that

$$e_{F\setminus K}(X_{2\ell_1}, Y_{2\ell_1}) \ge (1-\lambda)\frac{\delta}{n}|X_{2\ell_1}||Y_{2\ell_1}| \ge (1-\lambda)\frac{\delta}{n}\cdot\Big((1-45\lambda)d^{\ell_1}\delta^{\ell_1}r\Big)^2 \ge \left(1-\frac{C_0\lambda}{4}\right)\cdot\frac{r^2\delta^{2\ell_1+1}d^{2\ell_1}}{n},$$

and thus the claim holds. ⊡

Using this claim, we can now connect vertices with paths with length $4\ell_1+5$, as follows.

Claim 6.5. Let $u,v\in[n]$ be distinct and $W\subset[n]\setminus\{x,y\}$ with $|W|\le\lambda\delta$. Then, there are at least $(1-C_0\lambda/2)\delta^{2\ell_1+3}d^{2\ell_1+2}/n$ $(F\setminus K,K)$-alternating u,v-paths of length $4\ell_1+5$ which have no internal vertices in W.

Proof of Claim 6.5. Let $U_1=N_{F\setminus K}(u)\setminus(W\cup\{v\}\cup N_{F\setminus K}(v))$ and $V_1=N_{F\setminus K}(v)\setminus(W\cup\{u\}\cup N_{F\setminus K}(u))$. Let $U_2=N_K(U_1)\setminus(W\cup\{u,v\}\cup N_K(V_1)\cup V_1)$ and $V_2=N_K(V_1)\setminus(W\cup\{u,v\}\cup N_K(U_1)\cup U_1)$. Using similar arguments to those in the proof of Claim 6.4, we have $(1-2\lambda)\delta d\le|U_2|,|V_2|\le(1+2\lambda)\delta d$.

Let $s=|U_2|$ and $t=|V_2|$, so that $s,t\ge r$, and enumerate $U_2=\{u_1,\dots,u_s\}$ and $V_2=\{v_1,\dots,v_t\}$. For each $i\in[s]$, pick $u_i'\in N_K(u_i)\cap U_1$ and for each $j\in[t]$ pick $v_j'\in N_K(v_j)\cap V_1$. For each $i\in[s]$, let $\tilde U_i=\{u_i',u_{i+1}',\dots,u_{i+r-1}'\}$ and $\hat U_i=\{u_i,u_{i+1},\dots,u_{i+r-1}\}$, and, for each $j\in[t]$, let $\tilde V_j=\{v_j',v_{j+1}',\dots,v_{j+r-1}'\}$ and $\hat V_j=\{v_j,v_{j+1},\dots,v_{j+r-1}\}$, with addition in the indices modulo s and t as appropriate. For each $i\in[s]$ and $j\in[t]$, let $W_{i,j}=W\cup\{u,v\}\cup\tilde U_i\cup\tilde V_j$, so that $|W_{i,j}|\le 3r+\lambda\delta$. Now, by Claim 6.4, there are at least $\left(1-\frac{C_0\lambda}{4}\right)\cdot\frac{r^2\delta^{2\ell_1+1}d^{2\ell_1}}{n}$ $(F\setminus K,K)$-alternating $\hat U_i,\hat V_j$-paths with length $4\ell_1+1$ with no internal vertices in $W_{i,j}$, each of which can be extended to an $(F\setminus K,K)$-alternating u,v-path with length $4\ell_1+5$ with no internal vertices in W. Therefore, as each such path appears for r^2 pairs (i,j), the number of $(F\setminus K,K)$-alternating u,v-paths with length $4\ell_1+5$ with no internal vertices in W is at least

$$\left(1-\frac{C_0\lambda}{4}\right)\cdot\frac{r^2\delta^{2\ell_1+1}d^{2\ell_1}}{n}\cdot st\cdot\frac{1}{r^2}\ge\left(1-\frac{C_0\lambda}{4}\right)(1-2\lambda)^2\cdot\frac{\delta^{2\ell_1+3}d^{2\ell_1+2}}{n}\ge\left(1-\frac{C_0\lambda}{2}\right)\cdot\frac{\delta^{2\ell_1+3}d^{2\ell_1+2}}{n},$$

as required, where we have used that $s,t\ge(1-2\lambda)\delta d$. ⊡

Finally, we can show that **I7** holds. Let $x,y\in[n]$ be distinct and let $Z\subset[n]$ with $|Z|\le 40\log n$. Let $y'\in N_K(y)\setminus(Z\cup\{x\})$, with $d-|(Z\cup\{x\})\cap N_K(y)|$ choices. For each $i\ge 0$, let $\mathcal{P}_i$ be the set of pairs (v,P) such that P is an $(F\setminus K,K)$-alternating y',v-path of length $2i$ with no vertices in $Z\cup\{x\}$ and $d_{F\setminus K}(v,Z\cup V(P)\cup\{x,y\})\le\left(\mathbf{1}_{\{i=0\}}+1/\log n\right)\lambda\delta$. Note that $|\mathcal{P}_0|=1$ as $|Z\cup\{x,y\}|\le\lambda\delta$.

Claim 6.6. For each $1\le i\le 10\log n$,

$$\sum_{(v,P)\in\mathcal{P}_i}d_{F\setminus K}(v)\ge\left(1-4\lambda\cdot\mathbf{1}_{\{i=1\}}-\frac{21\lambda}{\log n}\right)\delta d\cdot\sum_{(v,P)\in\mathcal{P}_{i-1}}d_{F\setminus K}(v).$$

Proof of Claim 6.6. Let $(v,P)\in\mathcal{P}_{i-1}$. Let $Z^+=V(P)\cup Z\cup\{x,y\}$, so that $|Z^+|\le 100\log n$. Let B be the set of $w\in[n]\setminus Z^+$ with $d_{F\setminus K}(w,Z^+)\ge 10$ so that, by **I5**, we have $|B|\le 100\log n$.

For each (u,v') with $u\in N_{F\setminus K}(v)\setminus Z^+$ and $v'\in N_K(u)\setminus(Z^+\cup B)$, there is a path P' such that $P\subset P'$ and $(v',P')\in\mathcal{P}_i$ by definition of B. Therefore,

$$\begin{aligned}\sum_{u\in N_{F\setminus K}(v)}\sum_{v'\in N_K(u)}d_{F\setminus K}(v')\ &-\sum_{(v',P')\in\mathcal{P}_i:P\subset P'}d_{F\setminus K}(v')\\ &\overset{\mathbf{I1}}{\le}\sum_{u\in N_{F\setminus K}(v)\cap Z^+}|N_K(u)|\cdot 2\delta+\sum_{u\in N_{F\setminus K}(v)}\sum_{v'\in N_K(u)\cap(Z^+\cup B)}2\delta\\ &\le 2\delta d\cdot|N_{F\setminus K}(v)\cap Z^+|+(|E_{\ge 2,K}(N_{F\setminus K}(v))|+|Z^+\cup B|)\cdot 2\delta\\ &\overset{\mathbf{I3}}{\le}2d\delta\cdot\left(\mathbf{1}_{\{i=1\}}+1/\log n\right)\lambda\delta+\frac{\lambda d\cdot 2\delta}{\log n}\cdot 2\delta+200\log n\cdot 2\delta\\ &\le\left(2\cdot\mathbf{1}_{\{i=1\}}+10/\log n\right)\cdot\lambda\delta^2 d.\end{aligned}$$

Now, we have $\sum_{u\in N_{F\setminus K}(v)}\sum_{v'\in N_K(u)} d_{F\setminus K}(v') \geq (1-\lambda/\log n)\delta d\cdot d_{F\setminus K}(v)$ by **I2** and $d_{F\setminus K}(v)\geq \delta/2$ by **I1**, so it follows that

$$\sum_{(v',P')\in\mathcal{P}_i:P\subset P'} d_{F\setminus K}(v') \geq \left(1-4\lambda\cdot \mathbf{1}_{\{i=1\}} - 21\lambda\cdot\log^{-1} n\right)\delta d\cdot d_{F\setminus K}(v).$$

Therefore,

$$\begin{aligned}\sum_{(v,P)\in\mathcal{P}_i} d_{F\setminus K}(v) &\geq \sum_{(v,P)\in\mathcal{P}_{i-1}}\sum_{(v',P')\in\mathcal{P}_i:P\subset P'} d_{F\setminus K}(v')\\ &\geq \left(1-4\lambda\cdot \mathbf{1}_{\{i=1\}} - 21\lambda\cdot\log^{-1} n\right)\delta d\cdot \sum_{(v,P)\in\mathcal{P}_{i-1}} d_{F\setminus K}(v).\end{aligned}$$

as claimed. ⊡

Now, let ℓ satisfy $10\ell_0\leq \ell\leq 10\log n$. Let $\ell_2=\ell-(2\ell_1+3)\geq 0$. Let $(v,P)\in\mathcal{P}_{\ell_2}$, so that, by Claim 6.6 applied for each $i\in[\ell_2]$ and **I1**, the number of choices for (v,P) is at least

$$\begin{aligned}\frac{\sum_{(v,P)\in\mathcal{P}_{\ell_2}} d_{F\setminus K}(v)}{(1+\lambda)\delta} &\geq \left(1-4\lambda-\frac{21\lambda}{\log n}\right)\cdot\left(1-\frac{21\lambda}{\log n}\right)^{\ell_2-1}\delta^{\ell_2}d^{\ell_2}\cdot\frac{(1-\lambda)\delta}{(1+\lambda)\delta}\\ &\geq \left(1-\frac{C_0\lambda}{4}\right)\delta^{\ell_2}d^{\ell_2}.\end{aligned}$$

Then, let P' be an $(F\setminus K,K)$-alternating x,v-path of length $4\ell_1+5$ with no internal vertices in $Z\cup V(P)\cup\{y\}$, with at least $(1-C_0\lambda/2)\delta^{2\ell_1+3}d^{2\ell_1+2}/n$ options by Claim 6.5. Then, $(P\cup P')+y'y$ is an $(F\setminus K,K)$-alternating x,y-path with length 2ℓ. Combining our choices for y',P and P', the number of choices is at least

$$\begin{aligned}(d-|(Z\cup\{x\})\cap N_K(y)|)\left(1-\frac{C_0\lambda}{4}\right)\delta^{\ell_2}d^{\ell_2}\cdot\left(1-\frac{C_0\lambda}{2}\right)\frac{\delta^{2\ell_1+3}d^{2\ell_1+2}}{n}&\\ \geq\left(1-C_0\lambda-\frac{|(Z\cup\{x\})\cap N_K(y)|}{d}\right)\frac{\delta^\ell d^\ell}{n},&\end{aligned}$$

as required. Thus, **I7** holds. □

6.2 Likely properties of $G_d(n)$: expansion

The following two lemmas, Lemma 6.7 and Lemma 6.9 will be applied with two different values of λ (λ and $\lambda/\log n$), which motivates why these Lemmas use the bounds $\lambda\geq\frac{1}{64}\log^{-4}n$ and $\lambda\delta d\geq C\log n$. The two lemmas give a likely bound for, respectively, $e_K(U)$ and $|E_{\geq 2,K}(U)|$ with $K\sim G_d(n)$, so that in combination they will be used to prove **I3**. Both of these lemmas are proved using switching techniques. For convenience, after their proofs we will combine them into Corollary 6.11.

Lemma 6.7. *Let $\frac{1}{64}\log^{-4}n\leq\lambda\ll 1$ and $1/C\ll 1$, and let d satisfy $d\geq C\log n$ and $d=o(n)$. Let $\delta\geq C\log n$ satisfy $\lambda\delta d\geq C\log n$, and let $q=2\delta/n$. Let $K\sim G_d(n)$ and $F\sim G(n,q)$. Then, with probability $1-o(n^{-14})$, the following holds.*

- **J** *For each $U\subset[n]$, if $|U|\leq\frac{\lambda n}{10^4}$ and there is some $W\subset[n]$ with $|W|\leq 4|U|/\delta$ and $U\subset W\cup N_F(W)$, then $e_K(U)\leq\frac{\lambda d}{2}|U|$.*

Proof. By Chernoff's bound and a union bound, with probability $1-o(n^{-14})$, we have $\Delta(F)\le 3\delta$. Fixing F with this property, we now show the following holds with probability $1-o(n^{-14})$.

J′ For each $W\subset[n]$, if $|W|\le\frac{\lambda n}{10^3\delta}$, then $e_K(W\cup N_F(W))\le\frac{\lambda d}{8}\cdot\delta|W|$.

This will complete the proof of the lemma as **J′** implies **J**. Indeed, when it holds, if $U\subset[n]$ and $W\subset[n]$ satisfy $|U|\le\frac{\lambda n}{10^4}$, $|W|\le 4|U|/\delta$ and $U\subset W\cup N_F(W)$, then $|W|\le\frac{\lambda n}{10^3\delta}$, so that, from **J′** we have

$$e_K(U)\le e_K(W\cup N_F(W))\le\frac{\lambda d}{8}\cdot\delta|W|\le\frac{\lambda d}{2}|U|,$$

and thus **J** holds.

Let then $W\subset[n]$ with $|W|\le\lambda n/10^3\delta$, and set $m=|W|$. Let $W'=W\cup N_F(W)$, so that $|W'|\le 4\delta m$. For each $i\in\mathbb{N}_0$, let $\mathcal{K}^i$ be the set of $K\in\mathcal{K}_d(n)$ for which $e(K[W'])=i$. We will show the following.

Claim 6.8. For each $i\ge\lambda d\delta m/16$, $|\mathcal{K}^i|\le|\mathcal{K}^{i-1}|/e$.

Proof of Claim 6.8. Let L_i be an auxiliary bipartite graph with vertex classes $\mathcal{K}^i$ and $\mathcal{K}^{i-1}$, and put an edge between $K\in\mathcal{K}^i$ and $K'\in\mathcal{K}^{i-1}$ if $K\triangle K'$ is a cycle with length 6 with exactly two vertices in $W'=W\cup N_F(W)$, which moreover have an edge between them in K.

Firstly, let $K\in\mathcal{K}^i$. Pick v_1,v_2 with $v_1v_2\in E(K[W'])$, with $2i$ possibilities as $K\in\mathcal{K}^i$. Then, pick v_3,v_4 with $v_3v_4\in E(K-W'-N_K(v_2))$, with, as $|W'|\le 4\delta m$, at least $2(dn/2-d\cdot(4\delta m+d))\ge 3dn/4$ possibilities. Then, pick v_5,v_6 with $v_5,v_6\in E(K-W'-\{v_3,v_4\}-N_K(v_4)-N_K(v_1))$, with at least $2(dn/2-d\cdot(4\delta m+2+2d))\ge 3dn/4$ possibilities. Then, noting that $K-\{v_1v_2,v_3v_4,v_5v_6\}+\{v_2v_3,v_4v_5,v_6v_1\}\in\mathcal{K}^{i-1}$ is a neighbour of K in L_i, we have

$$d_{L_i}(K)\ge\frac{1}{2}\cdot 2i\cdot\left(\frac{3dn}{4}\right)^2\ge\frac{1}{2}id^2n^2. \tag{44}$$

Then, let $K'\in\mathcal{K}^{i-1}$. Pick v_2,v_3 with $v_2\in W'$ and $v_2v_3\in E(K')$, with at most $4\delta md$ possibilities. Similarly, pick v_1,v_6 with $v_1\in W'$ and $v_1v_6\in E(K')$, with at most $4\delta md$ possibilities. Then, pick v_4 and v_5 with $v_4v_5\in E(K')$, with at most dn possibilities. Therefore, in total, we have $d_{L_i}(K')\le 16d^3\delta^2m^2n/2$.

Therefore, from this and (44), double-counting the edges of L_i we get

$$\frac{id^2n^2}{2}\cdot|\mathcal{K}^i|\le e(L_i)\le\frac{16d^3\delta^2m^2n}{2}\cdot|\mathcal{K}^{i-1}|,$$

so that, as $i\ge\lambda d\delta m/16$ and $m\le\lambda n/10^3\delta$,

$$\frac{|\mathcal{K}^i|}{|\mathcal{K}^{i-1}|}\le\frac{16d\delta^2m^2}{in}\le\frac{256\delta m}{\lambda n}\le\frac{1}{e},$$

as required. ⊡

Therefore, applying the claim repeatedly, we have, for each $i\ge\lambda d\delta m/8$,

$$|\mathcal{K}^i|\le\exp\left(-\frac{\lambda d\delta m}{16}\right)\cdot|\mathcal{K}_d(n)|\le\exp\left(-\frac{Cm\log n}{16}\right)\cdot|\mathcal{K}_d(n)|,$$

where we have used that $\lambda d\delta\ge C\log n$. In particular,

$$\sum_{i=\lambda d\delta m/8}^{n^2}\frac{|\mathcal{K}^i|}{|\mathcal{K}_d(n)|}\le n^2\exp\left(-\frac{Cm\log n}{16}\right).$$

Now, as there were at most $\binom{n}{m} \leq n^m$ choices for W with $|W| = m$, we have that $\mathbf{J}'$ does not hold with probability at most

$$\sum_{m=1}^{\lambda n/10^3\delta} n^m \cdot n^2 \exp\left(-\frac{Cm\log n}{16}\right) = o(n^{-14}),$$

as required, using that $1/C \ll 1$. □

Similarly to Lemma 6.7, we now prove the following lemma which shows a likely bound on $|E_{\geq 2,K}(U)|$ for many sets U in $K \sim G_d(n)$ using switching techniques.

Lemma 6.9. *Let $\log^{-4} n \leq \lambda \ll 1$ and $1/C \ll 1$. Let $d \geq C\log n$ satisfy $d = o(n)$. Let δ satisfy $\lambda\delta d \geq C\log n$, and let $q = 2\delta/n$. Let $K \sim G_d(n)$ and $F \sim G(n,q)$. Then, with probability $1 - o(n^{-14})$, the following holds.*

K *For each $U \subset [n]$, if $|U| \leq \frac{\lambda n}{10^6 d}$ and there is some $W \subset [n]$ with $|W| \leq 4|U|/\delta$ and $U \subset W \cup N_F(W)$, then $|E_{\geq 2,K}(U)| \leq \frac{\lambda d}{2}|U|$.*

Proof. By Chernoff's bound and a union bound, with probability $1 - o(n^{-14})$, we have $\Delta(F) \leq 3\delta$. Fixing F with this property, we now show the following holds with probability $1 - o(n^{-14})$.

K′ For each $W \subset [n]$, if $|W| \leq \frac{\lambda n}{2\cdot 10^3\delta d}$, then $|E_{\geq 2,K}(W \cup N_F(W))| \leq \frac{\lambda d}{16} \cdot \delta|W|$.

This will easily complete the proof of the lemma as, by Lemma 6.7, we have that, with probability $1 - o(n^{-14})$, **J** holds with $\lambda/64$ in place of λ. Then, if **K′** holds, for each $U \subset [n]$ and $W \subset [n]$ satisfying $|U| \leq \frac{\lambda n}{10^6 d}$, $|W| \leq 4|U|/\delta$ and $U \subset W \cup N_F(W)$, we have $|W \cup N_F(W)| \leq 4\delta|W| \leq 16|U|$ and $|W| \leq \frac{\lambda n}{2\cdot 10^3\delta d}$, so that, from **K′** we have

$$|E_{\geq 2,K}(U)| \leq |E_{\geq 2,K}(W \cup N_F(W))| + e_K(W \cup N_F(W)) \leq \frac{\lambda d}{16} \cdot \delta|W| + \frac{\lambda d}{64} \cdot |W \cup N_F(W)| \leq \frac{\lambda d}{2}|U|,$$

and thus **K** holds.

Let $W \subset [n]$ with $|W| \leq \lambda n/(2 \cdot 10^3\delta d)$, and set $m = |W|$. Let $W' = W \cup N_F(W)$, so that $|W'| \leq 4\delta m$. For each $i \in \mathbb{N}_0$, let $\mathcal{K}^i$ be the set of $K \in \mathcal{K}_d(n)$ for which

$$\sum_{v \in [n]\setminus W'} \max\{d_K(v, W') - 1, 0\} = i. \tag{45}$$

We study this rather than $|E_{\geq 2,K}(W')|$, as removing an edge vv' between $v \in [n] \setminus W'$ and $v' \in W'$ when v has exactly 2 neighbours in W' will reduce $|E_{\geq 2,K}(W')|$ by 2 (and 1 if v has more than 2 neighbours in W'), but the quantity at (45) will always reduce by 1 when v has at least 2 neighbours in W'. We will then use the simple bound that

$$\sum_{v \in [n]\setminus W'} \max\{d_K(v, W') - 1, 0\} \leq |E_{\geq 2,K}(W')| \leq 2 \sum_{v \in [n]\setminus W'} \max\{d_K(v, W') - 1, 0\}. \tag{46}$$

Using switching methods, we now show the following.

Claim 6.10. For each $i \geq \lambda d\delta m/32$, $|\mathcal{K}^i| \leq |\mathcal{K}^{i-1}|/\sqrt{e}$.

Proof of Claim 6.10. Let $i \geq \lambda d\delta m/32$. Let L_i be an auxiliary bipartite graph with vertex classes $\mathcal{K}^i$ and $\mathcal{K}^{i-1}$, and put an edge between $K \in \mathcal{K}^i$ and $K' \in \mathcal{K}^{i-1}$ if $K \triangle K'$ is a cycle with length 6 with exactly 1 vertex in W'.

Firstly, let $K \in \mathcal{K}^i$. Pick v_1, v_2 with $v_1 \in W'$ and $v_1v_2 \in E_{\geq 2,K}(W')$, with at least i possibilities as $K \in \mathcal{K}^i$. Then, pick v_3, v_4 with $v_3v_4 \in E(K - W' - N_K(W') - N_K(v_2))$, with at least $2(dn/2 - d \cdot (4\delta m + 4\delta m \cdot d + d)) \geq 3dn/4$ possibilities, where we have used that $\delta dm \leq \lambda n/(2 \cdot 10^3)$. Then, pick v_5, v_6 with $v_5, v_6 \in E(K - W' - N_K(W') - \{v_3, v_4\} - N_K(v_4))$, with at least $2(dn/2 - d \cdot (4\delta m + 4\delta m \cdot d + 2 + d)) \geq 3dn/4$ possibilities. Note that $K - \{v_1v_2, v_3v_4, v_5v_6\} + \{v_2v_3, v_4v_5, v_6v_1\} \in \mathcal{K}^{i-1}$ is a neighbour of K in L_i, as $v_1v_2 \in E_{\geq 2,K}(W')$ and $v_3, v_4, v_5, v_6 \notin N_K(W')$. Thus we have

$$d_{L_i}(K) \geq i \cdot \left(\frac{3dn}{4}\right)^2 \geq \frac{1}{2} i d^2 n^2. \tag{47}$$

Then, let $K' \in \mathcal{K}^{i-1}$. To switch from K' we need to increase $\sum_{v \in [n] \setminus W'} \max\{d_K(v, W') - 1, 0\}$ by one. If v_1 is the vertex of the cycle in W' and $v_1v_2 \notin E(K')$, it must be the case that $v_2 \in N_{K'}(W')$. This is so that v_2 already has another neighbour from W' in K other than v_1 and so adding v_1v_2 to K' will increase $\sum_{v \in [n] \setminus W'} \max\{d_K(v, W') - 1, 0\}$ by one. Thus we must pick v_1, v_2 with $v_1 \in W'$ and $v_2 \in N_{K'}(W')$, with at most $4\delta m \cdot 4\delta md$ possibilities. Then, pick v_3, v_6 with $v_2v_3, v_1v_6 \in E(K')$, with at most d^2 possibilities. Then, pick v_4 and v_5 with $v_4v_5 \in E(K')$, with at most dn possibilities. Therefore, in total, we have $d_{L_i}(K') \leq 16d^4\delta^2m^2n$.

Therefore, from this and (47), double-counting the edges of L_i we get

$$\frac{id^2n^2}{2} \cdot |\mathcal{K}^i| \leq e(L_i) \leq 16d^4\delta^2m^2n \cdot |\mathcal{K}^{i-1}|,$$

so that, as $i \geq \lambda d\delta m/32$ and $m \leq \lambda n/(2 \cdot 10^3\delta d)$,

$$\frac{|\mathcal{K}^i|}{|\mathcal{K}^{i-1}|} \leq \frac{32d^2\delta^2m^2}{in} \leq \frac{32 \cdot 32d\delta m}{\lambda n} \leq \frac{1}{\sqrt{e}},$$

as required. ⊡

Therefore, applying the claim repeatedly, we have, for each $i \geq \lambda d\delta m/16$,

$$\frac{|\mathcal{K}^i|}{|\mathcal{K}_d(n)|} \leq \exp\left(-\frac{\lambda d\delta m}{64}\right) \leq \exp\left(-\frac{Cm\log n}{64}\right),$$

where we have used that $\lambda d\delta \geq C\log n$. In particular,

$$\sum_{i=\lambda d\delta m/16}^{n^2} \frac{|\mathcal{K}^i|}{|\mathcal{K}_d(n)|} \leq n^2 \exp\left(-\frac{Cm\log n}{64}\right).$$

Thus, using (46), with probability at least $1 - n^2\exp\left(-\frac{Cm\log n}{64}\right)$, we have that $|E_{\geq 2,K}(N_K(W))| \leq 2 \cdot \lambda d\delta m/16$.

Now, as there were at most $\binom{n}{m} \leq n^m$ choices for W with $|W| = m$, we have that $\mathbf{K}'$ does not hold with probability at most

$$\sum_{m=1}^{\lambda n/2 \cdot 10^3\delta d} n^m \cdot n^2 \exp\left(-\frac{Cm\log n}{64}\right) = o(n^{-14}),$$

as required. □

In combination, Lemmas 6.7 and 6.9 imply the following.

Corollary 6.11. *Let $\log^{-4} n \le \lambda \ll 1$ and $1/C \ll 1$. Let $d \ge C \log n$ satisfy $d = o(n)$. Let $m \in \mathbb{N}$ and suppose $\delta = 2m/n$ satisfies $\lambda\delta d \ge C \log n$. Let $p = m(\binom{n}{2} - dn/2)^{-1}$. Let $K \sim G_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$. Then, with probability $1 - o(n^{-14})$, the following holds.*

- **L** *For each $U \subset [n]$, if $|U| \le \frac{\lambda n}{10^6 d}$ and there is some $W \subset [n]$ with $|W| \le 4|U|/\delta$ and $U \subset W \cup N_F(W)$, then $e_K(U) + |E_{\ge 2,K}(U)| \le \lambda d|U|$.*

Proof. Let $q = 2\delta/n$. Independently, select $K \sim G_d(n)$ and $F' \sim G(n,q)$. When $|E(F') \setminus E(K)| \ge m$, then let E be a uniformly random set of m edges from $E(F') \setminus E(K)$, and otherwise let E be a uniformly random set of m edges from $[n]^{(2)} \setminus E(K)$. Then, by symmetry, we have $F \sim F(n,d,m)$.

Now, as $d = o(n)$, we have $\mathbb{E}|E(F') \setminus E(K)| \ge q(\binom{n}{2} - dn/2) \ge 1.1\delta n/2$. By a simple application of a Chernoff bound, then, we have that, with probability $1 - o(n^{-14})$, $|E(F') \setminus E(K)| \ge m$, and when this happens $F \subset F'$. Furthermore, if $F \subset F'$ and **L** holds for F' in place of F, then **L** holds with F. Thus, the corollary follows from Lemmas 6.7 and 6.9. □

6.3 Likely properties of $F \setminus K$: expansion

For **I4**, similarly to our work in Section 6.2, we again split this into Lemma 6.12 and Lemma 6.13, where the first of these holds more easily. We then show a result that implies that **I5** is likely to hold.

Lemma 6.12. *Let $1/n \ll 1/C \ll \varepsilon \le 1$. Let d and λ satisfy $\log^{-3} n \le \lambda \ll 1$, $10^3 d \le \lambda n$, and $d \ge C \log n$. Let $m \in \mathbb{N}$ satisfy $\varepsilon dn/2 \le m \le n^{1+\sigma}$. Let $\delta = 2m/n$ and suppose that $\lambda\delta \ge C \log n$. Let $p = m(\binom{n}{2} - dn/2)^{-1}$. Let $K \sim G_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$.*

Then, with probability $1 - o(n^{-14})$, the following holds.

- **M** *For each $U \subset [n]$ with $|U| \le \lambda n/100\delta \log n$, if there is some $U' \subset [n]$ with $|U'| \le 4|U|/d$ and $U \subset U' \cup N_K(U')$, then*
$$e_{F\setminus K}(U) \le \frac{\lambda}{2\log n} \cdot \delta|U|$$

Proof. Let $U \subset [n]$ with $|U| \le \lambda n/100\delta \log n$ and let $s = |U|$. We will show that, with probability $1 - \exp(-\omega(s))$, U satisfies the property we want, and then take a union bound over all sets U we must consider.

Let $X = e(F \setminus K[U])$ and let $M = \left\lceil \frac{\lambda\delta s}{2\log n} \right\rceil$ so that again, we have

$$\mathbb{P}\left(X \ge \frac{\lambda\delta s}{2\log n}\right) \le \sum_{i=M}^{\binom{s}{2}} \binom{\binom{s}{2}}{i} \cdot p^i \le \sum_{i=M}^{s^2} \left(\frac{es^2 p}{2i}\right)^i \le \sum_{i=M}^{s^2} \left(\frac{es^2 p}{2M}\right)^i$$

As
$$\frac{es^2 p}{2M} \le \frac{esp \cdot \log n}{\lambda\delta} \le \frac{2es \cdot \log n}{\lambda n} \le \frac{2e}{100\delta}$$
and $\lambda\delta \ge C \log n$, we thus obtain that

$$\mathbb{P}\left(X \ge \frac{\lambda\delta s}{2\log n}\right) \le \sum_{i=M}^{s^2} \left(\frac{2e}{100\delta}\right)^i \le 2\left(\frac{2e}{100\delta}\right)^{\frac{\lambda\delta s}{2\log n}} \le 2\left(\frac{2e}{100\delta}\right)^{\frac{Cs}{2}} = \exp(-\omega(s)),$$

where we have used that $\delta \geq C\varepsilon \log n$. Let us now count the number of sets $U \subset [n]$ with $|U| = s$ for which there is some $U' \subset [n]$ with $|U'| \leq 4s/d$ and $U \subset U' \cup N_K(U')$. Note that the condition $|U'| \leq 4|U|/d$ implies that if there is some such U then we must have $1 \leq 4s/d$, which in turn implies that $s \geq d/4 \geq C \log n/4$. Since $|U' \cup N_K(U')| \leq (d+1)|U'|$, the number of possible choices for such a $U \subset U' \cup N_K(U')$ is at most

$$\sum_{i=1}^{4s/d} \binom{n}{i} \cdot 2^{(d+1)i} \leq \sum_{i=1}^{4s/d} \left(n2^{(d+1)}\right)^i \leq 2\left(n2^{(d+1)}\right)^{4s/d} \leq 2^{5s}$$

as $d \geq C \log n$. Thus, the probability that $\mathbf{M}$ does not hold is at most

$$\sum_{s=C\log n/4}^{n} 2^{5s} \cdot \exp\left(-\omega(s)\right) = o(n^{-14}),$$

as required. □

We now give our likely bound on $|E_{\geq 2, F\setminus K}|$, similarly to the bound proved in Lemma 6.12.

Lemma 6.13. *Let $1/n \ll 1/C \ll \varepsilon, \sigma \ll 1$. Let d and λ satisfy $\log^{-3} n \leq \lambda \ll 1$, $10^3 d \leq \lambda n$, and $d \geq C \log n$. Let $m \in \mathbb{N}$ satisfy $\varepsilon dn/2 \leq m \leq n^{1+\sigma}$. Let $\delta = 2m/n$ and suppose that $\lambda\delta \geq C \log n$. Let $p = m(\binom{n}{2} - dn/2)^{-1}$. Let $K \sim G_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$. Then, with probability $1 - o(n^{-14})$, the following holds.*

N *For each $U \subset [n]$ with $|U| \leq \lambda n/100\delta \log n$, if there is some $U' \subset [n]$ with $|U'| \leq 4|U|/d$ and $U \subset U' \cup N_K(U')$, then*

$$|E_{\geq 2, F\setminus K}(U)| \leq \frac{\lambda}{2\log n} \cdot \delta |U|.$$

Proof. Let $U \subset [n]$ with $|U| \leq \lambda n/100\delta \log n$, and set $s = |U|$. We will show that, with probability $1 - \exp\left(-\Omega(s)\right)$, U satisfies the property we want, so that the result follows by a union bound over all such U for which there is some $U' \subset [n]$ with $|U'| \leq 4|U|/d$ and $U \subset U' \cup N_K(U')$, as at the end of Lemma 6.12.

Consider a process in which we reveal edges in E by taking each $v \in [n] \setminus U$ and revealing potential edges in E into U until a neighbour of v in U has been revealed or it has been revealed that v has no neighbours in U. Let V be the set of vertices v after this which have a neighbour in U. Then, reveal the rest of the edges between V and U that have not yet been considered, and let X be the number of edges between V and U revealed here. Note that $|E_{\geq 2, F\setminus K}(U)| \leq 2X$.

Now, let $\mathcal{E}$ be the event that $|V| \leq 3\delta s$. As $V = N_{F\setminus K}(U)$, by a simple application of Chernoff's bound, with probability $1 - \exp(-\omega(s))$, $\mathcal{E}$ holds. Let $Y = X$ if $\mathcal{E}$ holds, and $Y = 0$ otherwise. Thus, with probability $1 - \exp(-\omega(s))$, $|E_{\geq 2, F\setminus K}(U)| \leq 2Y$.

For $M = \left\lceil \frac{\lambda\delta s}{4\log n} \right\rceil = \Omega(s)$, we have, then

$$\mathbb{P}\left(Y \geq \frac{\lambda\delta s}{4\log n}\right) \leq \sum_{i=M}^{|U||V|} \binom{|U||V|}{i} p^i \leq \sum_{i=M}^{3\delta s^2} \binom{s \cdot 3\delta s}{i} p^i \leq \sum_{i=M}^{3\delta s^2} \left(\frac{3e\delta s^2 p}{i}\right)^i \leq \sum_{i=M}^{3\delta s^2} \left(\frac{3e\delta s^2 p}{M}\right)^i$$

As

$$\frac{3e\delta s^2 p}{M} \leq \frac{3esp \cdot 4\log n}{\lambda} \leq \frac{24es\delta \log n}{\lambda n} \leq \frac{24e}{100} \leq e^{-1/3}$$

we see that

$$\mathbb{P}\left(Y \geq \frac{\lambda\delta s}{4\log n}\right) \leq \sum_{i=M}^{3\delta s^2} e^{-i/3} \leq 4e^{-\frac{\lambda\delta s}{12\log n}} \leq \exp\left(-Cs/20\right).$$

Thus $|E_{\geq 2,F'}(U)| \leq 2Y$ with probability $1-\exp\left(-Cs/20\right)$. Similarly to the end of the proof of Lemma 6.12, we know that there are at most 2^{5s} sets U such that $|U| = s$, there is some set U' with $|U'| \leq 4s/d$ and $U \subset U' \cup N_K(U')$; and this is only possible is $s \geq d/4 \geq C\log n/4$. Thus the probability that **N** does not hold is at most

$$\sum_{s=C\log n/4}^{n} 2^{5s} \cdot \exp\left(-Cs/20\right) = o(n^{-14})$$

as required. □

We now prove our result that will imply that **I5** is likely to hold.

Lemma 6.14. *Let $1/n \ll 1/C \ll \varepsilon, \sigma \ll 1$. Let $0 < p \leq n^{\sigma-1}$ and let $H \sim G(n,p)$. Then, with probability $1-o(n^{-14})$, for each $U \subset [n]$ with $|U| \leq 100\log n$, $|\{v \in [n]\setminus U : d_H(v,U) \geq 10\}| \leq 100\log n$.*

Proof. Let $U \subset [n]$ satisfy $|U| \leq 100\log n$. For each $v \in [n]\setminus U$, the probability that $d_H(v,U) \geq 10$ is at most $\binom{|U|}{10}p^{10} \leq |U|^{10}p^{10} \leq n^{-9}$. Then, the probability that $|\{v \in [n]\setminus U : d_H(v) \geq 10\}| > 100\log n$ is, setting $k = 100\log n$, at most $\sum_{i=k}^{n}\binom{n}{i}n^{-9i} \leq \sum_{i=k}^{n} n^i n^{-9i} \leq n \cdot n^{-8k}$. Taking a union bound over all sets $U \subset [n]$ satisfying $|U| \leq 100\log n$, the property in the lemma does not hold with probability at most $2n^k \cdot n^{1-8k} = o(n^{-14})$, as required. □

6.4 Connection properties of $F \setminus K$

We now prove that **I6** is likely to hold, in the following lemma.

Lemma 6.15. *Let $1/n \ll 1/C \ll \varepsilon, \sigma \ll 1$. Let d, λ and μ satisfy $\log^{-3} n \leq \lambda, \mu \ll 1$ and $d \geq C\log n$. Let $m \in \mathbb{N}$ satisfy $\varepsilon dn \leq m \leq n^{1+\sigma}$. Let $\delta = 2m/n$ and suppose that $\lambda^2\mu\delta \geq C\log n$. Let $p = m(\binom{n}{2} - dn/2)^{-1}$. Let $K \sim G_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$. Then, with probability $1-o(n^{-14})$, for each disjoint sets $U, V \subset [n]$ with $|U|, |V| \geq \mu n$, we have $e_{F\setminus K}(U,V) = (1\pm\lambda)\frac{\delta}{n}|U||V|$.*

Proof. First note that, for each disjoint $U, V \subset [n]$ with $|U| \geq |V| \geq \mu n$, we have, as $\lambda\mu n \geq \sqrt{n} \geq 10d$ that

$$\left(1-\frac{\lambda}{5}\right)|U||V| \leq |U||V| - d|U| - d|V| \leq e_{K_n\setminus K}(U,V) \leq |U||V|.$$

Therefore, we have

$$\left(1-\frac{\lambda}{5}\right)p|U||V| \leq \mathbb{E}(e_{F\setminus K}(U,V)) \leq p|U||V|.$$

Note that $p = \frac{\delta}{n-d-1} = \left(1\pm\frac{\lambda}{5}\right)\frac{\delta}{n}$. As $\lambda^2 p|U||V| \geq \lambda^2 \cdot \left(1-\frac{\lambda}{5}\right)\frac{\delta}{n} \cdot |U| \cdot \mu n \geq \lambda^2\mu\delta|U|/2 \geq \frac{C}{2}|U|\log n$, we have by an application of a Chernoff bound (Lemma 2.6) that, with probability $1-\exp(-\frac{C}{100}|U|\log n)$,

$$\left(1-\frac{\lambda}{2}\right)p|U||V| \leq e_{F\setminus K}(U,V) \leq \left(1+\frac{\lambda}{2}\right)p|U||V|.$$

Thus, $e_{F\setminus K}(U,V) \neq (1\pm\lambda)\frac{\delta}{n}|U||V|$ for some $U,V\subset[n]$ with $|U|\geq|V|\geq\mu n$ with probability at most

$$\sum_{m_1=\mu n}^{n}\sum_{m_2=m_1}^{n}\binom{n}{m_1}\binom{n}{m_2}\exp(-\tfrac{C}{100}m_2\log n)=o(n^{-14}). \qquad \square$$

6.5 Proof of Theorem 6.1

Finally, we can combine our work in this section to prove Theorem 6.1.

Proof of Theorem 6.1. Take the variables as set up in Theorem 6.1, so that $1/n \ll 1/C \ll 1/C_0 \ll \varepsilon$, $1/C \ll \mu,\sigma \ll 1$, $d\geq C\log n$, $\varepsilon dn/2\leq m\leq n^{1+\sigma}$, $\delta=2m/n$, and $p=m(\binom{n}{2}-dn/2)^{-1}$. Furthermore, we have

$$\lambda=\max\left\{\frac{\mu^2}{\log^2 n},\left(\frac{n\log n}{m}\right)^{1/4}\right\}, \tag{48}$$

and ℓ_0 is the smallest integer for which $\delta^{(\ell_0-1)}d^{(\ell_0-1)}\geq n$. Let C' satisfy $1/C\ll 1/C'\ll 1$.

Let $K\sim G_d(n)$, form E by including each element of $[n]^{(2)}\setminus E(K)$ uniformly at random with probability p, and let $F=K+E$. By Lemma 6.3, in order to prove Theorem 6.1, it is sufficient to show that **I1**–**I6** hold with probability $1-o(n^{-14})$.

Note that

$$C'\sqrt{\frac{\log n}{\delta}}\overset{(48)}{\leq}C'\sqrt{\frac{\log n}{\delta}}\cdot\lambda\left(\frac{m}{n\log n}\right)^{1/4}=C'\sqrt{\frac{\log n}{\delta}}\cdot\lambda\left(\frac{\delta}{2\log n}\right)^{1/4}=C'\lambda\left(\frac{\log n}{2\delta}\right)^{1/4}\leq\lambda,$$

where we have used that $\delta\geq\varepsilon d\geq\varepsilon C\log n$. Thus, by Lemma 5.3, we have that **I1** holds with probability $1-o(n^{-14})$.

Furthermore, we have

$$\frac{\lambda}{\log n}\overset{(48)}{\geq}\frac{1}{\log n}\left(\frac{n\log n}{m}\right)^{1/4}\geq\frac{1}{\log n}\left(\frac{\log n}{\delta}\right)^{1/4}=\frac{1}{\delta}\cdot\left(\frac{\delta}{\log n}\right)^{3/4}\geq\frac{C'}{\delta},$$

so that, by Lemma 5.3, we have that **I2** holds with probability $1-o(n^{-14})$.

By Corollary 6.11 with $\lambda'=\lambda/\log n$ and C', as $\lambda'\delta d\geq\lambda\delta\geq C'\log n$, we have that, with probability $1-o(n^{-14})$, **I3** holds whenever $|U|\leq\lambda n/10^6 d\log n$. Similarly, Corollary 6.11 applied with λ and C' shows that, with probability $1-o(n^{-14})$, **I3** holds whenever $|U|\leq\lambda n/10^6 d$. Thus, **I3** holds with probability $1-o(n^{-14})$. Furthermore, by Lemmas 6.12 and 6.13 applied with C', we have that, with probability $1-o(n^{-14})$, **I4** holds with probability $1-o(n^{-14})$, while, by Lemma 6.14, **I5** holds with probability $1-o(n^{-14})$.

We have, as $\delta\geq C\varepsilon\log n$, that

$$\lambda^3\delta\overset{(48)}{\geq}\left(\frac{n\log n}{m}\right)^{3/4}\delta\geq\left(\frac{\log n}{\delta}\right)^{3/4}\delta=\log n\cdot\left(\frac{\delta}{\log n}\right)^{1/4}\geq 10^8C'\log n.$$

Thus, by Lemma 6.15 with $\mu=\lambda/10^8$ and C', **I6** holds with probability $1-o(n^{-14})$. This completes the proof that **I1**–**I6** hold with probability $1-o(n^{-14})$, completing the proof of the theorem. $\square$

7 Upper bounds for switching paths

In this section, we are now able to show that our upper bound on the number of $(F\setminus K,K)$-alternating paths between two fixed vertices, that is, **C3**, is likely to hold in the regime $m\leq 1+\sigma$, where σ is a small fixed constant.

Lemma 7.1. *Let $1/n \ll 1/C \ll \mu \ll 1/C_0 \ll \varepsilon \leq 1$ and $1/n \ll \sigma \ll 1$. Let $d \in [n]$ and $m \in \mathbb{N}$ satisfy $d \geq C \log n$ and $\varepsilon dn/2 \leq m \leq n^{1+\sigma}$. Let $\delta = 2m/n$ and let*

$$\eta = C_0 \max\left\{\frac{\mu}{\log n}, \left(\frac{n \log n}{m}\right)^{1/8}\right\}.$$

Let $p = m(\binom{n}{2} - dn/2)^{-1}$. Let $K \sim G_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$. Then, with probability $1 - o(n^{-10})$, we have the following.

O *For each distinct $x, y \in [n]$ and each $\log n \leq \ell \leq 10 \log n$, the number of $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ is at most $(1+\eta)\, d^\ell \delta^\ell / n$.*

Proof. We start by showing that, for each distinct $x, y \in [n]$ and each $\log n \leq \ell \leq 10 \log n$, with some very high probability, a bound on the number of $(F \setminus K, K)$-alternating x, y-paths holds if we restrict these paths to those which do not have their second vertex in some small set of vertices, as follows.

Claim 7.2. For each distinct $x, y \in [n]$ and each $\frac{1}{2} \log n \leq \ell \leq 10 \log n$, with probability $1 - o(n^{-14})$, the following holds.

P1 There is some set $Z \subset [n]$ with $|Z| \leq \eta\delta/5$ for which the number of $(F \setminus K, K)$-alternating x, y-paths of length 2ℓ in which the second vertex is not in Z is at most $\left(1 + \frac{\eta}{5}\right) d^\ell \delta^\ell / n$.

Proof of Claim 7.2. Let $\lambda = (\eta/C_0)^2$. Fix $x, y \in [n]$. Reveal $K \sim G_d(n)$. Reveal the edges of E which do not contain x. By Theorem 6.1 (noting that $\frac{1}{2} \log n > 10\ell_0$ for ℓ_0 as defined in Theorem 6.1) and Lemma 5.1, we have that, with probability $1 - o(n^{-14})$, the following hold.

P2 For each $z \in N_K(y) \setminus \{x\}$ and $w \in [n] \setminus \{x, y, z\}$, the number of $(F \setminus K, K)$-alternating z, w-paths of length $2\ell - 2$ which do not contain $\{x, y\}$ is at least $\left(1 - C_0\lambda - \frac{|N_K(w) \cap \{x,y,z\}|}{d}\right) d^{\ell-1}\delta^{\ell-1}/n$.

P3 For each $z \in N_K(y) \setminus \{x\}$, the number of $(F \setminus K, K)$-alternating paths of length $2\ell - 2$ starting at z is at most $(1 + C_0\lambda)\, d^{\ell-1}\delta^{\ell-1}$.

Let $W = [n] \setminus \big(\{x, y\} \cup N_K(x, y) \cup N_K(N_K(y))\big)$, so that $N_K(y) \cap \big(\{w\} \cup N_K(w)\big) = \emptyset$ for each $w \in W$ and $|W| \geq n - 2 - 2d - d^2 \geq n - 2d^2$. By applying **P2** for each $z \in N_K(y) \setminus \{x\}$, we have the following.

P4 For each $w \in W$, the number of $(K, F \setminus K)$-alternating y, w-paths of length $2\ell - 1$ which do not contain x is at least $(d - |N_K(y) \cap \{x\}|) \cdot (1 - C_0\lambda)\, d^{\ell-1}\delta^{\ell-1}/n$.

Let B be the set of vertices $w \in [n] \setminus \{x, y\}$ for which the number of $(K, F \setminus K)$-alternating y, w-paths of length $2\ell - 1$ not containing x is more than $(1 + \eta/10) d^\ell \delta^{\ell-1}/n$. Then, the number of $(K, F \setminus K)$-alternating paths of length $2\ell - 1$ starting at y which do not contain x is at least

$$|B| \cdot (1 + \eta/10) \frac{d^\ell \delta^{\ell-1}}{n} \;+\; |W \setminus B| \cdot (1 - C_0\lambda)\,(d - |N_K(y) \cap \{x\}|)\, \frac{d^{\ell-1}\delta^{\ell-1}}{n}$$
$$\geq \Big(|B| \cdot \eta/10 + (n - 2d^2)\,(1 - C_0\lambda)\Big)\,(d - |N_K(y) \cap \{x\}|)\, \frac{d^{\ell-1}\delta^{\ell-1}}{n}.$$

From **P3**, this number is at most $(d - |N_K(y) \cap \{x\}|) \cdot (1 + C_0\lambda)\, d^{\ell-1}\delta^{\ell-1}$, so we have

$$|B| \leq \frac{(1 + C_0\lambda) - (n - 2d^2) \cdot (1 - C_0\lambda)\,/n}{\eta/10n} \leq \frac{\left(2C_0\lambda + \frac{2d^2}{n}\right) 10n}{\eta} \leq \frac{\eta n}{20},$$

where we have used that $d \le n^\sigma \le \eta\sqrt{n}/100$ and $\lambda = (\eta/C_0)^2$.

Now, reveal the edges next to x in E, and let $Z = B \cap N_{F\setminus K}(x)$. Note that $\mathbb{E}|Z| \le p\eta n/20 \le \eta\delta/10$. As $\eta\delta \ge C_0\delta(\log n/\delta)^{1/8} \ge C_0 \log n$, by a Chernoff bound (Lemma 2.6) we have that, with probability $1 - o(n^{-14})$, $|Z| \le \eta\delta/5$. Choose C' satisfying $1/c \ll 1/C' \ll 1$ so that $\eta \ge (\delta/\log n)^{3/8}\sqrt{\frac{\log n}{\delta}} \ge (C\varepsilon)^{3/8}\sqrt{\frac{\log n}{\delta}} \ge 20C'\sqrt{\frac{\log n}{\delta}}$. Applying Lemma 5.3, with probability $1 - o(n^{-14})$ we have that $d_{F\setminus K}(v) \le (1+\eta/20)\delta$. Assuming these properties, we have both that $|Z| \le \eta\delta/5$ and that the number of $(F\setminus K, K)$-alternating x,y-paths of length 2ℓ in which the second vertex is not in Z is at most

$$\left(1+\frac{\eta}{20}\right)\delta \cdot \left(1+\frac{\eta}{10}\right)\frac{d^\ell\delta^{\ell-1}}{n} \le \left(1+\frac{\eta}{5}\right)\frac{d^\ell\delta^\ell}{n}.$$

Thus, the claim holds. ⊡

By Claim 7.2, with probability $1 - o(n^{-11})$, **P1** holds for each distinct $x, y \in [n]$ and each $\frac{1}{2}\log n \le \ell \le 10\log n$. By Lemma 5.1, with probability $1 - o(n^{-14})$, the following holds for each $1 \le \ell \le 10\log n$.

P5 For each $z \in [n]$, the number of $(F\setminus K, K)$-alternating paths of length 2ℓ starting at z is at most $(1 + C_0\lambda)\, d^\ell\delta^\ell$.

Assuming **P5**, and that **P1** holds for each distinct $x, y \in [n]$ and each $\frac{1}{2}\log n \le \ell \le 10\log n$, we will now show that **O** holds.

Let $x, y \in [n]$ be distinct, and let $\mathcal{P}_{x,y}$ be the set of $(F\setminus K, K)$-alternating x,y-paths of length 2ℓ. We will now show by induction that, for each $0 \le i \le \ell/2$, there is a set $\mathcal{Q}_i$ of at most $(\eta\delta d/5)^i$ $(F\setminus K, K)$-alternating paths of length $2i$ which start at x such that the number of paths in $\mathcal{P}_{x,y}$ not containing a subpath in $\mathcal{Q}_i$ is at most $\left(1+\frac{\eta}{5}\right)\cdot\frac{d^\ell\delta^\ell}{n}\cdot\sum_{j=0}^{i-1}\left(\frac{\eta}{5}\right)^j$.

This is trivially true for $i = 0$ by letting $\mathcal{Q}_0$ contain only the degenerate path of length 0 with vertex set $\{x\}$. Let, then, $i > 0$, and suppose there is such a set $\mathcal{Q}_{i-1}$. For each $v \in [n]$, using **P1**, let $Z_{i,v}$ be such that $|Z_{i,v}| \le \eta\delta/5$ and the number of $(F\setminus K, K)$-alternating v,y-paths of length $2\ell - 2(i-1)$ in which the second vertex is not in $Z_{i,v}$ is at most $\left(1+\frac{\eta}{5}\right) d^{\ell-(i-1)}\delta^{\ell-(i-1)}/n$.

Let $\mathcal{Q}_i$ be the set of $(F\setminus K, K)$-alternating paths Q of length $2i$ starting at x for which the subpath of Q of length $(2i-2)$ containing x is in $\mathcal{Q}_{i-1}$ and the penultimate vertex of Q is in $Z_{i,v}$ where v is the antepenultimate vertex of Q. Then, the number of paths in $\mathcal{P}_{x,y}$ not containing a subpath in $\mathcal{Q}_i$ is, as $|\mathcal{Q}_{i-1}| \le (\eta\delta d/5)^{i-1}$, at most

$$\left(1+\frac{\eta}{5}\right)\cdot\frac{d^\ell\delta^\ell}{n}\cdot\sum_{j=0}^{i-2}\left(\frac{\eta}{5}\right)^j + |\mathcal{Q}_{i-1}|\cdot\left(1+\frac{\eta}{5}\right)\cdot\frac{d^{\ell-(i-1)}\delta^{\ell-(i-1)}}{n} \le \left(1+\frac{\eta}{5}\right)\cdot\frac{d^\ell\delta^\ell}{n}\cdot\sum_{j=0}^{i-1}\left(\frac{\eta}{5}\right)^j,$$

while $|\mathcal{Q}_i| \le |\mathcal{Q}_{i-1}|\cdot(\eta\delta/5)\cdot d \le (\eta\delta d/5)^i$.

Therefore, by induction, there is a set $\mathcal{Q}_{\lfloor\ell/2\rfloor}$ of at most $(\eta\delta d/5)^{\lfloor\ell/2\rfloor}$ $(F\setminus K, K)$-alternating paths of length $2\lfloor\ell/2\rfloor$ which start at x such that the number of paths in $\mathcal{P}_{x,y}$ not containing a subpath in $\mathcal{Q}_{\lfloor\ell/2\rfloor}$ is at most

$$\left(1+\frac{\eta}{5}\right)\cdot\frac{d^\ell\delta^\ell}{n}\cdot\sum_{j=0}^{\lfloor\ell/2\rfloor-1}\left(\frac{\eta}{5}\right)^j \le \left(1+\frac{\eta}{5}\right)\cdot\frac{d^\ell\delta^\ell}{n}\cdot\frac{1}{1-\frac{\eta}{5}} \le \left(1+\frac{\eta}{2}\right)\cdot\frac{d^\ell\delta^\ell}{n}. \tag{49}$$

Furthermore, using **P5**, the number of paths in $\mathcal{P}_{x,y}$ containing a subpath in $\mathcal{Q}_{\lfloor\ell/2\rfloor}$ is at most

$$|\mathcal{Q}_{\lfloor\ell/2\rfloor}|\cdot(1+C_0\lambda)d^{\ell-\lfloor\ell/2\rfloor}d^{\ell-\lfloor\ell/2\rfloor} \le 2\eta^{\lfloor\ell/2\rfloor}d^\ell d^\ell \le \frac{\eta}{2}\cdot\frac{d^\ell d^\ell}{n},$$

where we have used that $\ell \ge \log n$ and $\eta \ll 1$. In combination with (49), we have that **O** holds for x and y, as required. □

8 Concentration for switching paths in the non-critical range

In this section, we prove corresponding bounds to those in Theorem 6.1 and Lemma 7.1 when $\delta \geq n^\sigma$ for some fixed $\sigma > 0$, using shorter alternating paths, as in the following lemma.

Lemma 8.1. *Let $1/n \ll 1/C \ll 1/k \ll \sigma \leq 1$ and $1/C \ll \varepsilon, \mu, 1/C_0 \leq 1$. Let $d \geq C \log n$ and $\varepsilon dn/2 \leq m \leq \binom{n}{2} - dn/2$. Let $\delta = 2m/n$ and suppose that $\delta \geq n^\sigma$. Let $p = m(\binom{n}{2} - dn/2)^{-1}$ and $\eta = C_0\mu/\log n$. Let $K \sim G_d(n)$, form E by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F = K + E$.*

Then, with probability $1 - o(n^{-10})$, for every distinct $x, y \in [n]$ and every $Z \subset [n]$ with $|Z| \leq 4k$, the number of $(F \setminus K, K)$-alternating x, y-paths with length $2k$ and no interior vertices in Z is $(1 \pm \eta)\delta^k d^k/n$.

As this includes the case where d is large, we will first need to prove a result on the distribution of edges in $G_d(n)$ in this regime, which we do in Section 8.1. Then, in Section 8.2, we quote a result of Kim and Vu [19] and prove Lemma 8.1.

8.1 Edge distribution in $G_d(n)$ for large d.

When d is very large, switching methods in $G_d(n)$ are hard to implement and we need some additional preparation on the likely distribution of edges between large vertex subsets of $G_d(n)$. We will prove the result we need using that the random binomial graph $G(n, q)$ for some $q \approx d/n$ is d-regular with at least some very small probability. Thus, when properties hold with exceptionally high probability in $G(n, q)$, we can infer the property holds with high probability in $G_d(n)$. This was similarly done by Krivelevich, Sudakov, Vu and Wormald [23] using a result of Shamir and Upfal [30]. For completion, we will give a different simple proof of the specific result we need (Claim 8.4), while proving the following lemma.

Lemma 8.2. *Let $1/n \ll 1$. Let $n^{0.99} \leq d < n$. Then, for all but at most $e^{-n}|\mathcal{K}_d(n)|$ graphs $K \in \mathcal{K}_d(n)$, we have the following property.*

Q *For each $U, V \subset [n]$ with $|U|, |V| \geq n^{0.98}$, we have $e_K(U, V) = (1 \pm n^{-0.01})\frac{d}{n}|U||V|$.*

Proof. That the lemma holds when $d \geq n - n^{0.96}$ follows immediately from the following claim.

Claim 8.3. If $d \geq n - n^{0.96}$, then **Q** holds for every $K \in \mathcal{K}_d(n)$.

Proof of Claim 8.3. Let $K \in \mathcal{K}_d(n)$ and $d \geq n - n^{0.96}$. Let $U, V \subset [n]$ with $|U| \geq |V| \geq n^{0.98}$. Then,

$$e_K(U, V) \geq |U||V| - |U|(n - 1 - d) \geq |U||V| - |U| \cdot n^{0.96} \geq (1 - n^{-0.02})|U||V|,$$

so that $e_K(U, V) = (1 \pm n^{-0.02})|U||V| = (1 \pm n^{-0.01})\frac{d}{n}|U||V|$. ⊡

Suppose, then, that $d < n - n^{0.96}$. Let $m = \lfloor n^{0.9} \rfloor$ and

$$q = \frac{d(n - 2m)}{(n - m)(n - m - 1)} = (1 \pm 10n^{-0.1})\frac{d}{n}. \tag{50}$$

Let $G \sim G(n, q)$. We will show that, with at least some small probability, G is d-regular, as follows.

Claim 8.4. With probability at least $\exp(-n^{1.91})$, G is d-regular.

Proof of Claim 8.4. Let $A \subset [n]$ satisfy $|A| = m$ and let $B = [n] \setminus A$. Enumerate $A = \{a_1, \dots, a_m\}$ and $B = \{b_1, \dots, b_{n-m}\}$. For each $i \in [n-m]$, let $d_i = d - d_{G[B]}(b_i)$. Note that $\mathbb{E}(d_{G[B]}(b_i)) = q(n-m-1) \geq d/2 \geq n^{0.99}/2$ for each $i \in [n-m]$. Thus, by a Chernoff bound (Lemma 2.6) and a union bound, with probability at least $1 - 1/2n^2$ we have that, for every $i \in [n-m]$,

$$d_{G[B]}(b_i) = q(n-m-1) \pm \sqrt{qn} \cdot \log^2 n \overset{(50)}{=} \frac{d(n-2m)}{(n-m)} \pm n^{0.6} = d - \frac{dm}{n-m} \pm n^{0.6}.$$

Now,

$$\frac{dm}{n-m} \geq \frac{n^{0.99} \cdot n^{0.9}/2}{n} \geq n^{0.6},$$

and

$$\frac{dm}{n-m} \leq m \cdot \frac{n - n^{0.96}}{n - n^{0.9}} = m - m \cdot \frac{n^{0.96} - n^{0.9}}{n - n^{0.9}} \leq m - n^{0.9} \cdot n^{-0.04}/2 \leq m - n^{0.6}.$$

Therefore, with probability at least $1 - 1/2n^2$, we have that $0 \leq d_i \leq m$ for every $i \in [n-m]$. Furthermore, by the choice of q, note that $\mathbb{P}(e(G-A) = i)$ is maximised at $i = q\binom{n-m}{2} = d(n-2m)/2$ and thus we have $e(G-A) = q\binom{n-m}{2} = d(n-2m)/2$ with probability at least $1/n^2$. When this happens, $\sum_{i \in [n-m]} d_i = dm$. Thus, with probability at least $1/2n^2$, $\sum_{i \in [n-m]} d_i = dm$ and $0 \leq d_i \leq m$ for each $i \in [m]$.

Assuming these properties hold, we can observe there is a bipartite graph H between vertex classes A and B such that $d_H(a_i) = d$ for each $i \in [m]$ and $d_H(b_i) = d_i$ for each $i \in [n-m]$. Indeed, this can be done by connecting b_i to a_k for each $k \in [m]$ for which there is some $k' = k \mod m$ with $\sum_{j<i} d_j < k' \leq \sum_{j \leq i} d_j$. Then, for each $i \in [n-m]$, as $d_i \leq m$, we have that b_i is connected to d_i different vertices in A, and, as $\sum_{j \in [n-m]} d_j = dm$, for each $i \in [m]$, we have that a_i is connected to d different vertices in B.

Futhermore, the probability that $G[A] = \emptyset$ and $G[A, B] = H$ is at least

$$\min\{q, 1-q\}^{m(n-m)+\binom{m}{2}} \overset{(50)}{\geq} \left(\frac{\min\{d, n-d\}}{2n}\right)^{m(n-m)+\binom{m}{2}} \geq n^{-m(n-m)-\binom{m}{2}} \geq 2n^2 \cdot \exp(-n^{1.91}),$$

and when this happens G is d-regular. Thus, we have that $G(n, d/n)$ is d-regular with probability at least $\exp(-n^{1.91})$. ⊡

We now show that $\mathbf{Q}$ is very likely to be true with G in place of K, as follows.

Claim 8.5. The probability that $\mathbf{Q}$ with G in place of K is not satisfied is at most $\exp(-n^{1.92})$.

Proof of Claim 8.5. Let $U, V \subset [n]$ satisfy $|U|, |V| \geq n^{0.98}$. Then, by a Chernoff bound (Lemma 2.6), we have that

$$e_G(U, V) = e_G(U \setminus V, V) + e_G(U \cap V, V \setminus U) + 2e_G(U \cap V) = (1 \pm n^{-0.01}/2)q|U||V|$$

with probability $1 - \exp(-\Omega(n^{-0.02} \cdot n^{1.96} \cdot (d/n))) = 1 - \exp(-\Omega(n^{1.93}))$ using (50) and that $d \geq n^{0.99}$. By a union bound, we therefore have that G satisfies $\mathbf{Q}$ in place of K with probability $1 - 2^{2n} \cdot \exp(-\Omega(n^{1.93})) \geq 1 - \exp(-n^{1.92})$, as required. ⊡

As $\mathbb{P}(G = K')$ is constant on $K' \in \mathcal{K}_d(n)$, if $K \sim K_d(n)$ then the probability that K does not satisfy $\mathbf{Q}$ is, by Claim 8.4 and Claim 8.5, at most

$$\frac{\mathbb{P}(G \text{ does not satisfy } \mathbf{Q})}{\mathbb{P}(G \text{ is } d\text{-regular})} \leq \frac{\exp(-n^{1.92})}{\exp(-n^{1.91})} = \exp(n^{1.91} - n^{1.92}) \leq \exp(-n),$$

and thus the lemma also holds when $d < n - n^{0.96}$. □

8.2 Concentration for switching paths for large d

We will use a result of Kim and Vu [19] on the concentration of positive polynomials with constant degree, whose variables are independent Bernoulli random variables. By the main theorem of [19] (stated without varying probabilities for simplicity), we have the following.

Theorem 8.6. *Let k be a fixed constant. Let $N \in \mathbb{N}$ and $p \in [0,1]$. Let $X_1, \dots, X_N$ be independent random variables with $\mathbb{P}(X_i = 1) = p$ and $\mathbb{P}(X_i = 0) = 1 - p$ for each $i \in [N]$. Let $\mathbf{E}$ be a set of subsets of $[N]$ with size k and let $w(e) \geq 0$ for each $e \in \mathbf{E}$. Let $Y = \sum_{e \in \mathbf{E}} w(e) \cdot \prod_{i \in e} X_i$.*

For each $A \subset [N]$, let $\mathbf{E}_A = \{e \subset [N] \setminus A : e \cup A \in \mathbf{E}\}$ and let $Y_A = \sum_{e \in \mathbf{E}_A} w(e \cup A) \cdot \prod_{i \in e} X_i$. For each $i \geq 0$, let $\mathbb{E}_i Y = \max_{A \subset [N] : |A| = i} \mathbb{E} Y_A$. Let $\mathbb{E}^{\max} Y = \max_{i \geq 0} \mathbb{E}_i Y$ and $\mathbb{E}' Y = \max_{i \geq 1} \mathbb{E}_i Y$.

Then, for each $\alpha > 1$, we have

$$\mathbb{P}\left(|Y - \mathbb{E}Y| > 8^k (k!)^{1/2} (\mathbb{E}^{\max} Y \cdot \mathbb{E}' Y)^{1/2} \cdot \alpha^k\right) = O(\exp(-\alpha + (k-1)\log n)).$$

We can now prove Lemma 8.1 using a fairly direct application of Theorem 8.6.

Proof of Lemma 8.1. Note that as $\varepsilon dn/2 \leq m \leq \binom{n}{2} - dn/2$, we have $(1+\varepsilon)d \leq n - 1$. Thus, $d \leq n/(1+\varepsilon) \leq (1 - \varepsilon/2)n$ and by Lemma 8.2, with probability $1 - o(n^{-10})$, we have the following property when $d \geq n^{0.99}$.

R For each $U, V \subset [n]$ with $|U|, |V| \geq n^{0.98}$, we have $e_K(U,V) = (1 \pm n^{-0.01}) \frac{d}{n} |U||V|$.

We now assume that K has this property if $d \geq n^{0.99}$, and otherwise that $K \in \mathcal{K}_d(n)$, and choose edges from $[n]^{(2)} \setminus E(K)$. For this, let $N = \binom{n}{2} - dn/2$, and enumerate $[n]^{(2)} \setminus E(K) = \{e_1, \dots, e_N\}$. Let $p = m/N = \delta/(n-1-d)$, so that $p(n-1-d) = \delta$. Let X_i, $i \in [N]$, be independent Bernoulli random variables with probability p. Let F' be the graph with vertex set $[n]$ and edge set $\{e_i : i \in [N], X_i = 1\}$.

Fix distinct $x, y \in [n]$ and a set $Z \subset [n]$ with $|Z| \leq 4k$. We will show the property we want holds with probability $1 - o(n^{-12-4k})$ so that the result follows by a union bound. Let $\mathbf{E}$ be the set of $I \subset [N]$ with $|I| = k$ such that if H_I is the graph with vertex set $[n]$ and edge set $\{e_i : i \in I\}$, then there is at least one (H_I, K)-alternating x,y-path of length $2k$ with no interior vertices in Z. For each $I \in \mathbf{E}$, let $w(I)$ be the number of (H_I, K)-alternating x,y-paths of length $2k$ with no interior vertices in Z. Let $Y = \sum_{I \in \mathbf{E}} w(I) \cdot \prod_{i \in I} \mathbf{1}_{\{e_i \in E(F')\}}$. Observe that Y is the number of (F', K)-alternating x,y-paths of length $2k$ with no interior vertices in Z.

The result we want will follow from Theorem 8.6 and the following two claims.

Claim 8.7. $\left|\mathbb{E}Y - \frac{\delta^k d^k}{n}\right| \leq \frac{\eta \delta^k d^k}{2n}$.

Claim 8.8. $\mathbb{E}' Y \leq \left(\frac{8k}{\varepsilon}\right)^k \cdot \frac{d^k \delta^{k-1}}{n}$.

From these claims it follows that $\mathbb{E}^{\max} Y = \max\{\mathbb{E}Y, \mathbb{E}'Y\} \leq 2\delta^k d^k/n$ and that $\mathbb{E}'Y \leq \left(\frac{8k}{\varepsilon}\right)^k \cdot \delta^{k-1} d^k/n$. Let $\alpha = \log^2 n$, and note, for an application of Theorem 8.6, that

$$8^k (k!)^{1/2} (\mathbb{E}^{\max} Y \cdot \mathbb{E}' Y)^{1/2} \cdot \alpha^k \leq \frac{8^k (k!)^{1/2} (4\sqrt{k}/\sqrt{\varepsilon})^k \log^{2k} n \cdot (\delta d)^k}{\delta^{1/2} n} \leq \frac{\eta}{2n} \delta^k d^k,$$

where we have used that $\delta \geq n^{\sigma}$. Therefore, by Claim 8.7 and Theorem 8.6, we have that

$$\mathbb{P}\left(Y \neq (1 \pm \eta) \frac{\delta^k d^k}{n}\right) = O(\exp(-\log^2 n + (k-1)\log n)) = o(n^{-12-4k}),$$

as required. It is left then only to prove the two claims.

Proof of Claim 8.7. We prove this differently according to whether **a)** $d < n^{0.99}$ or **b)** $d \geq n^{0.99}$. Assume first that **a)** $d < n^{0.99}$ and note that in this case, $n/(n-d-1) \leq (1+\eta/10)$. Note too that there are at most $d \cdot (dn)^{k-1}$ paths $v_0v_1 \dots v_{2k}$ such that $v_0 = x$, $v_{2k} = y$, and $v_{2i-1}v_{2i} \in E(K)$ for each $0 \leq i \leq k-1$. Thus we have,

$$\mathbb{E}Y \leq d \cdot (dn)^{k-1} \cdot p^k = \frac{d^k}{n} \cdot \left(\frac{\delta n}{n-1-d}\right)^k \leq \left(1+\frac{\eta}{2}\right)\frac{\delta^k d^k}{n},$$

where we have used that $\eta = C_0\mu/\log n$ and $d < n^{0.99}$.

On the other hand, letting $v_0 = x$ and $v_{2k} = y$, as $\eta d = C_0\mu d/\log n \geq C_0\mu C$, there are at least $(1-\eta/4)d$ choices of $v_{2k-1} \in N_K(v_{2k}) \setminus (Z \cup \{v_0\})$. Then, for each $1 \leq i \leq k-1$ in turn, choose (v_{2i-1}, v_{2i}) so that $v_{2i-1}v_{2i} \in E(K)$ and v_{2i-1} and v_{2i} are distinct from $v_0, v_1, \dots, v_{2i-2}, v_{2k-1}$ and v_{2k}, and not in Z, and $v_{2i-2}v_{2i-1}, v_{2i}v_{2k-1} \notin E(K)$. For each such i, when we go to pick (v_{2i-1}, v_{2i}), let

$$W_i = Z \cup \{v_0, v_1, \dots, v_{2i-2}\} \cup \{v_{2k-1}, v_{2k}\} \cup N_K(v_{2i-2}, v_{2k-1}),$$

and note that $|W_i| \leq 6k+1+2d$. When we choose (v_{2i-1}, v_{2i}), any (v_{2i-1}, v_{2i}) with $v_{2i-1}v_{2i} \in E(K-W_i)$ has the properties we want, and

$$2e(K-W_i) \geq d(n-2|W_i|) \geq \left(1-\frac{\eta}{10k}\right)d(n-d-1),$$

where we have used that $d < n^{0.99}$ and $1/n \ll \mu, 1/k, 1/C_0$ and $\eta = C_0\mu/\log n$. Therefore, we have

$$\mathbb{E}Y \geq \left(1-\frac{\eta}{2k}\right)d\left(\left(1-\frac{\eta}{10k}\right)d(n-d-1)\right)^{k-1}p^k \geq \left(1-\frac{\eta}{4}\right)\frac{d^k\delta^k}{(n-d-1)} \geq \left(1-\frac{\eta}{2}\right)\frac{\delta^k d^k}{n}.$$

Thus, the claim holds when $d < n^{0.99}$.

Therefore, we can assume that **b)** $d \geq n^{0.99}$, so that, in particular, **R** holds. Note, moreover, that, as $d \leq (1-\varepsilon/2)n$, we have $n-d-1 \geq \varepsilon n/3$. Let $v_0 = x$ and $v_{2k} = y$. For each $1 \leq i \leq 2k-3$, pick $v_i \in [n] \setminus (\{v_{2k}, v_0, v_1, \dots, v_{i-1}\} \cup Z)$ such that $v_i \in N_{K_n \setminus K}(v_{i-1})$ if i is odd and $v_i \in N_K(v_{i-1})$ is even. Note that the number of ways to do this is at least $(d-6k)^{k-2}(n-d-1-6k)^{k-1} = (1 \pm \frac{\eta}{10k})^{2k-3}d^{k-2}(n-1-d)^{k-1} = (1 \pm \eta/4)d^{k-2}(n-1-d)^{k-1}$, where we have used that $d \geq n^{0.99}$ and $n-1-d \geq \delta \geq n^{\sigma}$. Then, pick distinct v_{2k-2} and v_{2k-1} in $[n] \setminus (\{v_{2k}, v_0, v_1, \dots, v_{2k-3}\} \cup Z)$ with $v_{2k-2} \in N_K(v_{2k-3})$, $v_{2k-1} \in N_K(v_{2k})$ and $v_{2k-2}v_{2k-1} \notin E(K)$, so that, by **R**, the number of choices for v_{2k-2} and v_{2k-1} is, as $d \geq n^{0.99}$ and $n-1-d \geq \varepsilon n/3$,

$$(1 \pm 2n^{-0.01})d^2 - (1 \pm n^{-0.01})d^2 \cdot \frac{d}{n} = d^2 \cdot \left(\frac{n-d-1}{n} \pm 4n^{-0.01}\right) = \left(1 \pm \frac{\eta}{10}\right)\frac{d^2(n-d-1)}{n}.$$

Therefore, the number of choices for $v_1, \dots, v_{2k-1}$ is $(1 \pm \eta/2)d^k(n-d-1)^k/n$. Thus, we have

$$\mathbb{E}Y = \left(1 \pm \frac{\eta}{2}\right)\frac{d^k(n-d-1)^k p^k}{n} = \left(1 \pm \frac{\eta}{2}\right)\frac{d^k\delta^k}{n},$$

as required, completing the proof of the lemma in the case $d \geq n^{0.99}$. ⊡

Proof of Claim 8.8. Let $A \subset [N]$ with $1 \leq |A| \leq k$. Let $v_0v_1 \dots v_{2k}$ be an arbitrary x, y-path (so that $v_0 = x$ and $v_{2k} = y$) with $v_{2j-1}v_{2j} \in E(K)$ and $v_{2j-2}v_{2j-1} \notin E(K)$ for each $j \in [k]$ and $\{e_i : i \in A\} \subset \{v_{2j-2}v_{2j-1} : j \in [k]\}$. We will bound above the number of possibilities for such a path by counting the choices as we determine each of the vertices in this path. First, for each $i \in A$,

choose which edge $v_{2j-2}v_{2j-1}$ is the edge e_i, and the order of its vertices, so that both v_{2j-2} and v_{2j-1} are determined. Note that there at most $(2k)^{|A|} \leq (2k)^k$ choices for this.

Let $V_A = \{v_0, v_{2k}\} \cup \bigcup_{i\in A} V(e_i)$. Observe the following holds for each $j \in [k]$. If $\{v_{2j-1}, v_{2j}\} \subset V_A$, then (v_{2j-1}, v_{2j}) is already determined. If $|\{v_{2j-1}, v_{2j}\} \cap V_A| = 1$, then either v_{2j-1} or v_{2j} is determined, so there are at most d possibilities for (v_{2j-1}, v_{2j}) as $v_{2j-1}v_{2j} \in E(K)$. If $\{v_{2j-1}, v_{2j}\} \cap V_A = \emptyset$, then there are at most dn choices for (v_{2j-1}, v_{2j}) as $v_{2j-1}v_{2j} \in E(K)$. For each $s \in \{0, 1, 2\}$, let r_s be the number of $j \in [k]$ with $|\{v_{2j-1}, v_{2j}\} \cap V_A| = s$. Then, the number of choices for the vertices $v_0, v_1, \ldots, v_{2k}$ not in V_A is at most $d^{r_1}(dn)^{r_0}$. We have $r_1+2r_2+1 = |V_A|$ and $r_0+r_1+r_2 = k$, so $r_0 = k - |V_A|/2 - r_1/2 + 1/2$. In particular, the number of choices for the vertices is at most $d^{r_1}(dn)^{r_0} = d^{r_1}(nd)^{k-|V_A|/2-r_1/2}$.

Suppose first that $|A| < k$. If $v_0 \in V(e_j)$ for some $j \in A$, then (as $v_{2k} \notin \cup_{i\in A}V(e_i)$) $|V_A| = 2|A|+1$ and we can also observe that $r_1 \geq 2$. Indeed, as $|A| < k$, not every edge $v_{2j}v_{2j+1}$, $0 \leq j < k$, is in $\{e_i : i \in A\}$. Thus, we can choose j_0 to be the least $0 \leq j < k$ with $v_{2j}v_{2j+1} \notin \{e_i : i \in A\}$, and j_1 to be the largest $0 \leq j < k$ with $v_{2j}v_{2j+1} \notin \{e_i : i \in A\}$. Then, $v_{2j_1+1} \notin V_A$ and $v_{2j_1+2} \in V_A$, while, as $v_0v_1 \in \{e_i : i \in A\}$, $j_0 \geq 1$, so that $v_{2j_0-1} \in V_A$ but $v_{2j_0} \notin V_A$. Thus, as $j_0 \leq j_1$, we have $r_1 \geq 2$. Therefore, the number of possibilities for $v_0, v_1 \ldots, v_{2k}$ not in V_A is at most

$$d^{r_1}(nd)^{k-|V_A|/2-r_1/2+1/2} = \left(\frac{d}{n}\right)^{r_1/2}(nd)^{k-|A|} \leq \frac{d}{n}\cdot(nd)^{k-|A|}. \tag{51}$$

Suppose then that $v_0 \notin V(e_j)$ for any $j \in A$, so that, therefore, $|V_A| = 2|A| + 2$. We can also observe that $r_1 \geq 1$, as for the smallest $0 \leq j < k$ such that $v_{2j}v_{2j+1} \in \{e_i : i \in A\}$ we have $|\{v_{2j-1}, v_{2j}\} \cap V_A| = 1$. Then, the number of choices for $v_0, v_1, \ldots, v_{2k}$ not in V_A is at most

$$d^{r_1}(nd)^{k-|V_A|/2-r_1/2+1/2} = d^{r_1}(nd)^{k-|A|-1/2-r_1/2} = \left(\frac{d}{n}\right)^{r_1/2}\frac{1}{n^{1/2}}(nd)^{k-|A|} \leq \frac{\sqrt{d}}{n}\cdot(nd)^{k-|A|}. \tag{52}$$

Recall that $n - 1 - d \geq \varepsilon n/3$. Therefore, including our initial choices, from (51) and (52), we have that

$$\begin{aligned}\mathbb{E}Y_A &\leq (2k)^k \cdot \frac{d}{n}\cdot(nd)^{k-|A|}\cdot p^{k-|A|} = (2k)^k\cdot\frac{d^{k-|A|+1}}{n}\cdot\left(\frac{\delta n}{n-1-d}\right)^{k-|A|}\\ &\leq \left(\frac{8k}{\varepsilon}\right)^k\cdot\frac{d^{k-|A|+1}\delta^{k-|A|}}{n} \leq \left(\frac{8k}{\varepsilon}\right)^k\cdot\frac{d^k\delta^{k-1}}{n},\end{aligned}$$

where we have used that $|A| \geq 1$.

Furthermore, if $|A| = k$, then $V_A = \{v_i : 0 \leq i \leq 2k\}$. In this case the number of choices is $O(1)$, so that $\mathbb{E}Y_A = O(1) = o(\delta^{k-1}d^{k-1}/n)$, where we have used that $\delta \geq n^\sigma$ and $1/k \ll \sigma$. Therefore, we have $\mathbb{E}'Y \leq \left(\frac{8k}{\varepsilon}\right)^k \cdot \frac{d^k\delta^{k-1}}{n}$, as required. ⊡

□

9 Proof of Theorem 2.4

In Section 9.1, we prove a simple result that allows us to convert likely properties of (F, K), where $F \sim F(n, d, m)$ is generated by adding random non-edges to $K \sim G_d(n)$, to properties likely to hold for most of the d-regular subgraphs K of $F \sim F(n, d, m)$. Then, in Section 9.2, using the results in Sections 4–8, we prove Theorem 2.4.

9.1 Translating likely properties of $G_d(n)$ into subgraphs of $F(n,d,m)$

Among all graphs F' with $dn/2+m$ edges, where $F \sim F(n,d,m)$, we have that $\mathbb{P}(F=F')$ is proportional to $|\mathcal{K}_d(F')|$. Thus, any property likely to hold for $G \sim G_d(n)$ is likely to hold for most of the d-regular subgraphs of $F \sim F(n,d,m)$. This is important for our argument, so we prove this precisely with the bounds we will use, as follows.

Lemma 9.1. *Let $d,n,m \in \mathbb{N}$ satisfy $0 \le m \le (n-1-d)n/2$. Let $K \sim G_d(n)$, let E be a uniformly random set of m edges from $[n]^{(2)} \setminus E(K)$, and let $F = K+E$. Let $\mathcal{P}$ be such that $(F,K) \notin \mathcal{P}$ with probability $o(n^{-8})$. Then, with probability $1-o(n^{-2})$, $|\{K' \in \mathcal{K}_d(F) : (F,K') \notin \mathcal{P}\}| \le n^{-6} \cdot |\mathcal{K}_d(F)|$.*

Proof. Let $\mathcal{F}$ be the set of graphs F' with $m+dn/2$ edges, vertex set $[n]$, and $\mathcal{K}_d(F') \neq \emptyset$, and note that $\mathbb{P}(F \in \mathcal{F}) = 1$. Let $\mathcal{F}'$ be the set of $F' \in \mathcal{F}$ for which $|\{K' \in \mathcal{K}_d(F') : (F',K') \notin \mathcal{P}\}| > n^{-6}|\mathcal{K}_d(F')|$. We wish to prove that $\mathbb{P}(F \in \mathcal{F}') = o(n^{-2})$. For this, note that

$$\begin{aligned}
\mathbb{P}((F,K) \notin \mathcal{P}) &= \sum_{F' \in \mathcal{F}} \sum_{K' \in \mathcal{K}_d(F')} \mathbb{P}((F,K) = (F',K')) \cdot \mathbf{1}_{(F',K') \notin \mathcal{P}} \\
&= \frac{1}{|\mathcal{K}_d(n)| \binom{\binom{n}{2}-dn/2}{m}} \sum_{F' \in \mathcal{F}} \sum_{K' \in \mathcal{K}_d(F')} \mathbf{1}_{(F',K') \notin \mathcal{P}} \\
&> \frac{1}{|\mathcal{K}_d(n)| \binom{\binom{n}{2}-dn/2}{m}} \sum_{F' \in \mathcal{F}'} \frac{1}{n^6} \cdot |\mathcal{K}_d(F')| = \frac{1}{n^6} \cdot \mathbb{P}(F \in \mathcal{F}').
\end{aligned}$$

Thus, as $\mathbb{P}((F,K) \notin \mathcal{P}) = o(n^{-8})$, we have $\mathbb{P}(F \in \mathcal{F}') = o(n^{-2})$, as required. □

9.2 Proof of Theorem 2.4

Finally, we can combine our work so far to prove Theorem 2.4.

Proof of Theorem 2.4. Take the variables as set up in Theorem 2.4, so that $1/n \ll 1/C \ll \mu \ll 1/C_0 \ll \varepsilon \ll 1$, $d \ge C \log n$, and $\varepsilon dn/2 \le m \le (n-1-d)n/2$. Let

$$\eta = C_0 \max\left\{\frac{\mu}{\log n}, \left(\frac{n \log n}{m}\right)^{1/8}\right\}. \tag{53}$$

Pick k and σ with $1/n \ll 1/k \ll \sigma \ll 1$. If $d \le n^\sigma$, then let $\ell = \lceil 2 \log n \rceil$, and otherwise let $\ell = k$.

Let $p = m(\binom{n}{2} - dn/2)^{-1}$. Let $K \sim G_d(n)$. Let E be a set of m edges chosen uniformly at random from $[n]^{(2)} \setminus E(K)$, and let $F = K+E$. Furthermore, form E' by including each element of $[n]^{(2)} \setminus E(K)$ uniformly at random with probability p, and let $F' = K+E'$. Note that p is chosen so that $\mathbb{P}(|E'| = i)$ is maximised at $i = m$, and, thus $\mathbb{P}(|E'| = m) \ge 1/n^2$. We will show that, that with probability $1-o(n^{-10})$, **C1**, **C2** and **C3** hold for F' in place of F. From this, it holds that, with probability $1-o(n^{-8})$, **C1**, **C2** and **C3** (defined in Theorem 4.1) hold for F. Then, by Lemma 9.1, with probability $1-o(n^{-2})$ we have that **C1** holds and, for all but at most $n^{-6}|\mathcal{K}_d(F)|$ graphs $K \in \mathcal{K}_d(F)$ we have that **C2** and **C3** hold. Therefore, by Theorem 4.1, with probability $1-o(n^{-2})$, (6) holds for each $e,f \notin E(F)$.

Thus, to complete the proof of Theorem 2.4 it suffices to proof that, with probability $1-o(n^{-10})$ **C1**, **C2**, and **C3** hold for F' in place of F. That **C1** holds with probability $1-o(n^{-10})$ follows from Lemma 5.3. If $d \ge n^\sigma$, then **C2** and **C3** hold with probability $1-o(n^{-10})$ by Lemma 8.1. If $d < n^\sigma$, then **C2** and **C3** hold with probability $1-o(n^{-10})$ by Theorem 6.1 and Lemma 7.1, respectively. □

10 The lower part of the sandwich

In this section, we give the lower part (G_*, G) of the sandwich and prove it satisfies Theorem 1.1. As noted, this part of the theorem previously follows from the combination of work by Kim and Vu [20], Dudek, Frieze, Ruciński and Šileikis [9] and Gao, Isaev and McKay [14, 15]. Moreover, in [13], Gao, Isaev and McKay use an edge addition process that is effectively the same as we use here, and analyse it as long as $d = \omega(\log n)$. We include the lower part of the sandwich for completion, and because, due to the use of our corresponding version of Lemma 9.1, the analysis is different. For the coupling (G_*, G), the process used to create G is the complement of the process used for the upper part, while for ease of analysis we generate G_* via a sequence of graphs $H_0, H_1, \dots$ chosen to be more like the complement of the graphs $G_0^+, G_1^+, \dots$ in Section 3.5. This coupling is given in Section 10.1. We then given an overview of the analysis of the process required in Section 10.2, and outline its proof over the rest of this section.

10.1 Edge addition process

Let $1/n \ll 1/C \ll \varepsilon \ll 1$ and $d \geq C \log n$. Let η satisfy $1/C \ll \eta \ll \varepsilon$. Let $e_1, e_2, \dots$ be a sequence of edges drawn independently and uniformly at random from $[n]^{(2)}$. Let $x_1, x_2, \dots$ be independent random variables with $x_i \sim U([0,1])$ for each $i \geq 1$. Let F_0 be the empty graph with vertex set $[n]$, and let $\ell(0) = 0$. For each $1 \leq i \leq dn/2$, do the following. Let $\ell(i)$ be the least $\ell(i) > \ell(i-1)$ for which $e_{\ell(i)} \notin E(F_{i-1})$ and

$$x_{\ell(i)} \leq \frac{|\{K \in \mathcal{K}_d(n) : F_{i-1} + e_{\ell(i)} \subset K\}|}{\max_{e \notin E(F_{i-1})} |\{K \in \mathcal{K}_d(n) : F_{i-1} + e \subset K\}|},$$

and let $F_i = F_{i-1} + e_{\ell(i)}$. Then, let $G = F_{dn/2}$.

Let H_0 be the empty graph with vertex set $[n]$ and let $k(0) = 0$. For each $1 \leq i \leq \binom{n}{2}$, let $k(i)$ be the least $k > k(i)$ such that $e_k \notin \{e_1, \dots, e_{k-1}\}$ and, if $x_i \leq 1 - \eta$ then let $H_i = H_{i-1} + e_{k(i)}$, and, otherwise, let $H_i = H_{i-1}$. Let $N = \lfloor dn/2 - \eta n \rfloor$. Let $M \sim \mathrm{Bin}\left((1-\varepsilon)\frac{d}{n}, \binom{n}{2}\right)$. If $e(H_N) \geq M$, then let G_* have vertex set $[n]$ and edge set a uniformly random subset of M edges of $E(H_N)$. If $e(H_N) < M$, then let G_* have vertex set $[n]$ and edge set a uniformly random subset of M edges of $[n]^{(2)}$.

10.2 Analysis overview

Similarly to G^* in Section 3, it follows by symmetry that for any two graphs G', G'' with $V(G') = V(G'') = [n]$ and $e(G') = e(G'')$ we have $\mathbb{P}(G_* = G') = \mathbb{P}(G_* = G'')$, and thus, due to the choice of M, $G_* \sim G\left(n, (1-\varepsilon)\frac{d}{n}, \binom{n}{2}\right)$. Furthermore, when $e(H_N) \geq M$ we have $G_* \subset H_N$.

Now, similarly to H_i in Section 3.5, $e(H_N) \sim \mathrm{Bin}(N, 1-\eta)$ so, by a Chernoff bound we have that $e(H_N) \geq (1-2\eta)N \geq (1-3\eta)dn/2$ with high probability. As, by another Chernoff bound, $M \leq (1-\varepsilon+\varepsilon/2)dn/2$ with high probability, we have $e(H_N) \geq M$ with high probability. Thus, to prove $G_* \subset G$ with high probability it is sufficient to prove that $H_N \subset G$ with high probability.

Similarly to Claim 3.1 b), we have the following claim, which we prove in Section 10.3. Recall that we say $F \sim F^-(n, d, m)$ if F is a random graph with vertex set $[n]$ which has the same distribution as $K - E$ if $K \sim G_d(n)$ and E is a uniformly-chosen set of m edges from $E(K)$.

Claim 10.1. For each $0 \leq i \leq dn/2$, $F_i \sim F^-(n, d, dn/2 - i)$.

We will also use Lemma 2.2. After some preliminary work in Sections 10.4 and 10.5, we prove this lemma in Section 10.6. We will now show that it follows from Claim 10.1 and Lemma 2.2

that (G_*, G) satisfies the desired properties in Theorem 1.1. We have already noted that $G_* \sim G\left(n, (1-\varepsilon)\frac{d}{n}, \binom{n}{2}\right)$, and it follows from Claim 10.1 with $i = dn/2$, that $G = F_{dn/2} \sim G_d(n)$. Therefore, from the discussion above, we need only show that $\mathbb{P}(H_N \subset G) = 1 - o(1)$.

By Claim 10.1 and Lemma 2.2 (applied with $\varepsilon = \eta$) with a union bound, with high probability, for each $0 \le i \le N$ and each $e, f \notin E(F_{i-1})$, we have

$$|\{K \in \mathcal{K}_d(n) : F_{i-1} + e \subset E(K)\}| \le (1+\eta)\,|\{K \in \mathcal{K}_d(n) : F_{i-1} + f \subset E(K)\}|$$

and in particular,

$$\frac{|\{K \in \mathcal{K}_d(n) : F_{i-1} + e_{\ell(i)} \subset K\}|}{\max_{e \notin E(F_{i-1})} |\{K \in \mathcal{K}_d(n) : F_{i-1} + e \subset K\}|} \ge \frac{1}{1+\eta} > 1 - \eta. \tag{54}$$

Suppose, then, that $e \in E(H_N)$. Thus, there is some $1 \le i \le N$ with $e_{k(i)} = e$, with e not in $e_1, \dots, e_{k(i)-1}$, and with $x_{k(i)} \le 1 - \eta$. Now, the edges $e_{\ell(1)}, e_{\ell(2)}, \dots, e_{\ell(N)}$ are all distinct, so we must have $\ell(N) \ge k(N)$. Thus, there is some $j \in [N]$ with $\ell(j-1) < k(i) \le \ell(j)$. As e is not in $e_1, \dots, e_{k(i)-1}$, $x_{k(i)} \le 1 - \eta$ and (54) holds, we therefore have $\ell(j) = k(i)$ and $e = e_{\ell(j)} \in E(F_j) \subset E(F_{dn/2}) = E(F)$. Thus, $H_N \subset F$.

It is left then only to prove Claim 10.1 and Lemma 2.2. We prove Claim 10.1 in Section 10.3. As discussed in Section 2.2.2, to prove Lemma 2.2 we will use switching techniques, and in Section 10.4 we give the concentration result for switching paths that we will use for this, that is, Lemma 10.2. In Section 10.5, we note that a corresponding version of Lemma 9.1 holds here.

10.3 Distribution of F_i: proof of Claim 10.1

Claim 10.1 can be shown by induction similarly to Claim 3.1b), or we can observe our edge addition process for d is the complement of the edge deletion process for $n - 1 - d$.

Indeed, for each $0 \le i \le dn/2$, let F_i^c be the complement of F_i. Observe that $\ell(i)$ is the least $\ell(i) > \ell(i-1)$ for which $e_{\ell(i)} \in E(F_{i-1}^c)$ and

$$\begin{aligned} x_{\ell(i)} &\le \frac{|\{K \in \mathcal{K}_d(n) : F_{i-1} + e_{\ell(i)} \subset K\}|}{\max_{e \notin E(F_{i-1})} |\{K \in \mathcal{K}_d(n) : F_{i-1} + e \subset K\}|} = \frac{|\{K \in \mathcal{K}_d(n) : K^c \subset F_{i-1}^c - e_{\ell(i)}\}|}{\max_{e \notin E(F_{i-1})} |\{K \in \mathcal{K}_d(n) : K^c \subset F_{i-1}^c - e\}|} \\ &= \frac{|\mathcal{K}_{n-1-d}(F_{i-1}^c - e_{\ell(i)})|}{\max_{e \in E(F_{i-1}^c)} |\mathcal{K}_{n-1-d}(F_{i-1}^c - e)|} \end{aligned}$$

and $F_i^c = F_{i-1}^c - e_{\ell(i)}$.

Thus, we can form F_i^c using an edge deletion process as in Section 3. Thus (as there we did not use the condition $d \ge C \log n$), we have that $F_i^c \sim F(n, n-d-1, i)$ for each $i \in [dn/2]$, from which it can be easily seen that $F_i \sim F^-(n, d, i)$ for each $i \in [dn/2]$, as required.

10.4 Concentration for switching paths

As discussed in Section 2.2.2, in this case at least one of the graphs used for the switching paths in the proof of Lemma 2.2 is dense, so showing the concentration results that we need for the switching paths is much easier than for the upper part of the sandwich. We do this now, where the main subtlety we need to consider is the case where d is very large (so that the graph $K_n \setminus F$ is very sparse).

Lemma 10.2. *Let $1/n \ll 1/C \ll \eta, \varepsilon$. Let $d \ge C \log n$ and $\varepsilon dn \le m \le dn/2$. Let $K \sim G_d(n)$, let E be a uniformly random set of m edges from $E(K)$, and let $F = K - E$. Then, with probability $1 - o(n^{-8})$, the following holds.*

Let $x, y \in [n]$ be distinct and let $Z \subset [n] \setminus \{x, y\}$ satisfy $|Z| \leq 10$. Suppose, moreover, that if $d \geq n - 10^3/\eta$, then $N_{K_n \setminus K}(x, Z \cup \{y\}) = \emptyset$. Then, the number of $(K_n \setminus K, K \setminus F)$-alternating x, y-paths of length 4 is $(1 \pm \eta)(2m)^2(n-1-d)^2/n^3$.

Proof. Note that the result is trivial if $d = n-1$ as there are no such paths, so we can assume that $d < n-1$. By Lemma 8.2, with probability $1 - o(n^{-8})$, the following property holds.

S1 If $d \geq n^{0.99}$, then, for each $U, V \subset [n]$ with $|U|, |V| \geq n^{0.98}$, we have $e_K(U, V) = (1 \pm n^{-0.01})\frac{d}{n}|U||V|$.

Assume, then that **S1** holds. We will proceed differently according to whether **a)** $d < n^{0.99}$, **b)** $n - 1 - d < n^{0.99}$, or **c)** $d \geq n^{0.99}$ and $n - 1 - d \geq n^{0.99}$.

Suppose first that **a)** $d < n^{0.99}$. Let $p = 2m/dn$. Let $E' \subset K$ be chosen by including each edge in K independently at random with probability p, so that $\mathbb{P}(|E'| = m) \geq 1/n^2$, and let $F' = K - E'$. As $m \geq \varepsilon dn \geq \varepsilon Cn \log n$, by a simple application of a Chernoff bound, and a union bound, we have that, with probability $1 - o(n^{-8})$, $d_{K \setminus F'}(v) = \left(1 \pm \frac{\eta}{20}\right)\frac{2m}{n}$ for each $v \in [n]$. Thus, the following property holds with probability $1 - o(n^{-6})$.

S2 For each $v \in [n]$, $d_{K \setminus F}(v) = \left(1 \pm \frac{\eta}{20}\right)\frac{2m}{n}$.

Assume, then, that **S2** holds. Let $x, y \in [n]$ be distinct and suppose $Z \subset [n] \setminus \{x, y\}$ satisfies $|Z| \leq 10$. Choose $v_3 \in N_{K \setminus F}(y) \setminus (Z \cup \{x\})$, with at most $d_{K \setminus F}(y) = \left(1 \pm \frac{\eta}{20}\right)\frac{2m}{n}$ and at least $d_{K \setminus F}(y) - |Z \cup \{x\}| = \left(1 \pm \frac{\eta}{10}\right)\frac{2m}{n}$ choices by **S2** as $m \geq \varepsilon Cn \log n$. Then, choose v_1, v_2 with $v_1 v_2 \in E(K \setminus F)$, $v_1 \in N_{K_n \setminus K}(x) \setminus (Z \cup \{v_3, y\})$ and $v_2 \in N_{K_n \setminus K}(v_3) \setminus (Z \cup \{x, y\})$. Note that the number of choices for v_1 and v_2 is at most $2e(K \setminus F) = 2m$ and at least

$$2e(K \setminus F) - 2d|N_K(x) \cup N_K(v_3) \cup Z \cup \{x, v_3, y\}| \geq 2m - 2d(2d + 13) \geq \left(1 - \frac{\eta}{10}\right) \cdot 2m,$$

where we used that $m \geq \varepsilon dn$ and $d < n^{0.99}$. Thus, the number of choices for v_1, v_2 is $\left(1 \pm \frac{\eta}{10}\right) \cdot 2m$. Altogether, and using that $n = \left(1 \pm \frac{\eta}{10}\right)(n - 1 - d)$ as $d < n^{0.99}$, the number of $(K_n \setminus K, K \setminus F)$-alternating x, y-paths of length 4 is $(1 \pm \eta)(2m)^2(n-1-d)^2/n^3$.

Suppose then that **b)** $n - 1 - d < n^{0.99}$. Take distinct $x, y \in [n]$ and take $Z \subset [n] \setminus \{x, y\}$ with $|Z| \leq 10$ and, if $d \geq n - 10^3/\eta$, then $N_{K_n \setminus K}(x, Z \cup \{y\}) = \emptyset$. Choose $v_1 \in N_{K_n \setminus K}(x) \setminus (Z \cup \{y\})$, with $\left(1 \pm \frac{\eta}{10}\right)(n-1-d)$ choices (where we use that, if $d < n - 10^3/\eta$, then $\eta(n-d-1)/10 > 11 \geq |Z \cup \{y\}|$). We will show that, with probability $1 - o(n^{-21})$, the number of $(K \setminus F, K_n \setminus K)$-alternating v_1, y-paths with length 3 avoiding $Z \cup \{x\}$ is $\left(1 \pm \frac{\eta}{5}\right)(2m)^2(n-1-d)/n^3$. Thus, by a union bound, with probability at least $1 - o(n^{-20})$, this holds for all choices of v_1 and hence the number of $(K_n \setminus K, K \setminus F)$-alternating x, y-paths with length 4 avoiding Z is

$$\left(1 \pm \frac{\eta}{10}\right)(n-1-d) \cdot \left(1 \pm \frac{\eta}{5}\right)\frac{(2m)^2(n-1-d)}{n^3} = (1 \pm \eta)\frac{(2m)^2(n-1-d)^2}{n^3}.$$

Thus, by a union bound, with probability $1 - o(n^{-8})$, this holds for all distinct $x, y \in [n]$ and $Z \subset [n] \setminus \{x, y\}$ with $|Z| \leq 10$ and, if $d \geq n - 10^3/\eta$, then $N_{K_n \setminus K}(x, Z \cup \{y\}) = \emptyset$.

Fix, then, such an x, y, Z and v_1. Let $\hat{E}$ be the set of pairs (u, v) with $u, v \in [n] \setminus (Z \cup \{x, v_1, y\})$, $uv \in E(K_n \setminus K)$ and $v_1 u, vy \in E(K)$. Then, $|\hat{E}| \leq (n-1-d)n$ and, as $n - 1 - d < n^{0.99}$,

$$|\hat{E}| \geq (n-1-d)n - 2(n-1-d)(13 + 2(n-1-d)) \geq \left(1 - \frac{\eta}{10}\right)(n-1-d)n, \tag{55}$$

so that $|\hat{E}| = \left(1 \pm \frac{\eta}{10}\right)(n-1-d)n$. Let $\hat{G}$ be an auxiliary multigraph with vertex class $[n]$ and an edge uv with multiplicity $|\{(u, v), (v, u)\} \cap \hat{E}|$, so that $\hat{G}$ has maximum degree $n - d - 1$. By Vizing's

Theorem for multigraphs, the edge chromatic number of $\hat{G}$ is at most $(n-d-1)+2$. Therefore, we can partition $\hat{E}$ into $n-d+1$ disjoint sets $\hat{E}_1, \hat{E}_2, \ldots, \hat{E}_{n-d+1}$, where each set $\hat{E}_i$ is a matching – that is, a vertex v appears at most once in a pair in $\hat{E}_i$.

Let $p = 2m/dn$. Let $E' \subset E(K)$ be chosen by including each edge in K independently at random with probability p, so that $\mathbb{P}(|E'| = m) \geq 1/n^2$, and let $F' = K - E'$. For each $i \in [n-d+1]$, let $X_i = |\hat{E}_i \cap E|$. Let $B = \{i \in [n-1-d] : |\hat{E}_i| < C \log n\}|$.

Since each $\hat{E}_i$ is a matching, X_i is a binomial random variable with mean $\mathbb{E}(X_i) = p|\hat{E}_i|$. For $i \notin B$ we have $|\hat{E}_i| > C \log n$. For such i, by a Chernoff bound (Lemma 2.6), with probability $1 - o(n^{-22})$ we have $X_i \geq \left(1 \pm \frac{\eta}{20}\right) p^2 |\hat{E}_i|$, using $p > \varepsilon \gg 1/C$ and $\eta \gg 1/C$. Thus by a union bound, with probability $1 - o(n^{-21})$ we have

$$|\hat{E} \cap E'| = \sum_{i=1}^{n-d+1} X_i = \left(1 \pm \frac{\eta}{20}\right) p^2 |\hat{E}| + \sum_{i \in B} \left(X_i - \left(1 \pm \frac{\eta}{20}\right) p^2 |\hat{E}_i|\right). \tag{56}$$

Let us bound the size of B. We have that $|\hat{E}| = \sum_{i=1}^{n-d+1} |\hat{E}_i| \leq |B| C \log n + (n-d+1-|B|)n$ and so rearranging and using (55) we obtain

$$|B| \leq \frac{2n + \eta/10(n-1-d)n}{n - C \log n} \leq 4 + \eta(n-1-d) < n^{0.99}.$$

Thus, using $p > \varepsilon$ and (55) again, we have

$$\sum_{i \in B} |\hat{E}_i| < |B| C \log n < n^{0.99} \cdot C \log n < \frac{\eta}{50} p^2 |\hat{E}|. \tag{57}$$

Now, since $0 \leq X_i \leq |\hat{E}_i|$ we see that

$$-\left(1 + \frac{\eta}{20}\right) \sum_{i \in B} |\hat{E}_i| \leq \sum_{i \in B} \left(X_i - \left(1 \pm \frac{\eta}{20}\right) p^2 |\hat{E}_i|\right) \leq \sum_{i \in B} |\hat{E}_i|,$$

and in particular, by (57) we have $\sum_{i \in B} \left(X_i - \left(1 \pm \frac{\eta}{20}\right) p^2 |\hat{E}_i|\right) = \pm \frac{\eta}{20} p^2 |\hat{E}|$. Putting this together with (56) we have with probability $1 - o(n^{-21})$ that $|\hat{E} \cap E'| = \left(1 \pm \frac{\eta}{10}\right) p^2 |\hat{E}|$. Therefore, with probability $1 - o(n^{-21})$, the number of $(K \setminus F, K_n \setminus K)$-alternating v_1, y-paths with length 3 avoiding $Z \cup \{x\}$ is, as $d = \left(1 \pm \frac{\eta}{10}\right) n$,

$$\left(1 \pm \frac{\eta}{10}\right) p^2 |\hat{E}| = \left(1 \pm \frac{\eta}{10}\right) \left(\frac{2m}{dn}\right)^2 \cdot \left(1 \pm \frac{\eta}{10}\right) (n-1-d)n = (1 \pm \eta) \frac{(2m)^2 (n-1-d)}{n^3},$$

as required.

Suppose then that **c)** $d \geq n^{0.99}$ and $n-1-d \geq n^{0.99}$. Take distinct $x, y \in [n]$, $Z \subset [n] \setminus \{x, y\}$ with $|Z| \leq 10$, and $v_3 \in [n] \setminus (Z \cup \{x, y\})$. We will show that, with probability $1 - o(n^{-21})$, the number of $(K_n \setminus K, K \setminus F)$-alternating x, v_3-paths with length 3 avoiding $Z \cup \{y\}$ is $\left(1 \pm \frac{4\eta}{5}\right) (2m)(n-1-d)^2/n^3$.

We have $|N_{K_n \setminus K}(x) \setminus (Z \cup \{y, v_3\})| = \left(1 \pm \frac{\eta}{5}\right)(n-1-d) \geq n^{0.98}$ and $|N_{K_n \setminus K}(v_3) \setminus (Z \cup \{x, y\})| = \left(1 \pm \frac{\eta}{5}\right)(n-1-d) \geq n^{0.98}$. Thus, by **S1**, the number of pairs (v_1, v_2) with $v_1 v_2 \in E(K)$, $v_1 \in N_{K_n \setminus K}(x) \setminus \{y, v_3\}$ and $v_2 \in N_{K_n \setminus K}(v_3) \setminus \{x, y\}$ is

$$(1 \pm n^{-0.01}) \left(1 \pm \frac{\eta}{5}\right)^2 \frac{d}{n} (n-1-d)^2 = \left(1 \pm \frac{3\eta}{5}\right) \frac{d}{n} (n-d-1)^2.$$

Let A be the set of such pairs (v_1, v_2). For each $i \in [2]$, let $A_i = \{uv : |\{(u,v), (v,u)\} \cap A| = i\}$. Let $p = 2m/dn$. Let $E' \subset E(K)$ be chosen by including each edge in K independently at random

with probability p, so that $\mathbb{P}(|E'| = m) \geq 1/n^2$, and let $F' = K - E'$, and let $X = 2|A_2 \cap E'| + |A_1 \cap E'|$. Note that X is the number of $(K_n \setminus K, K \setminus F')$-alternating x, v_3-paths with length 3 avoiding $Z \cup \{y\}$, and

$$\mathbb{E}X = \left(1 \pm \frac{3\eta}{5}\right) \frac{pd}{n}(n-d-1)^2 = \left(1 \pm \frac{3\eta}{5}\right) \frac{(2m)}{n^2}(n-d-1)^2.$$

Thus, by a Chernoff bound applied to $|A_2 \cap E'|$ and to $|A_1 \cap E'|$, we have, with probability $1 - o(n^{-23})$, that $X = \left(1 \pm \frac{4}{5}\eta\right)(2m)(n-1-d)^2/n^2$. Thus, with probability $1 - o(n^{21})$, the number of $(K_n \setminus K, K \setminus F)$-alternating x, v_3-paths with length 3 avoiding $Z \cup \{y\}$ is $\left(1 \pm \frac{4\eta}{5}\right) \frac{(2m)}{n^2}(n-d-1)^2$.

Therefore, using a union bound and that there are $\left(1 \pm \frac{\eta}{10}\right) pd = \left(1 \pm \frac{\eta}{10}\right) \frac{2m}{n}$ choices for $v_3 \in N_{K \setminus F}(y) \setminus (Z \cup \{x\})$, with probability $1 - o(n^{-20})$, there are $(1 \pm \eta)(2m)^2(n-1-d)^2/n^3$ $(K_n \setminus K, K \setminus F)$-alternating x, y-paths with length 4 avoiding Z. By a union bound, then, with probability $1 - o(n^{-8})$, this holds for all distinct $x, y \in [n]$, and $Z \subset [n] \setminus \{x, y\}$ with $|Z| \leq 10$. □

10.5 Translating likely properties of $G_d(n)$ into subgraphs of $F^-(n, d, m)$

We now note that a corresponding version of Lemma 9.1 for $F^-(n, d, m)$ follows from taking complements in Lemma 9.1.

Lemma 10.3. *Let $d, n, m \in \mathbb{N}$ satisfy $0 \leq m \leq dn/2$. Let $K \sim G_d(n)$, let E be a uniformly random set of m edges from $E(K)$, and let $F = K - E$. Let $\mathcal{P}$ be a set of pairs (F, K) such that $(F, K) \notin \mathcal{P}$ with probability $o(n^{-8})$. Then, with probability $1 - o(n^{-2})$,*

$$|\{K' \in \mathcal{K}_d(n) : (F, K') \notin \mathcal{P} \wedge (F \subset K')\}| \leq \frac{1}{n^6} \cdot |\{K' \in \mathcal{K}_d(n) : F \subset K'\}|. \tag{58}$$

Proof. Let $\bar{K} = K_n \setminus K$, so that $\bar{K} \sim G_{n-1-d}(n)$, and let $\bar{d} = n - 1 - d$. Note, moreover, that the distribution of E is a uniformly random set of m edges from $[n]^{(2)} \setminus E(\bar{K})$, and $0 \leq m \leq (n-1-\bar{d})n/2$. Let $\bar{F} = \bar{K} + E$, and note that, therefore, $\bar{F} \sim F(n, n-1-d, m)$.

Let $\mathcal{P}'$ be the set of pairs (F', K') such that F' has $n\bar{d}/2 + m$ edges, $V(F') = [n]$, K' is $\bar{d}$-regular, $K' \subset F'$ and $(K_n \setminus F', K_n \setminus K') \in \mathcal{P}$. Then, we have that $(\bar{F}, \bar{K}) \in \mathcal{P}'$ with probability $1 - o(n^{-6})$. Therefore, by Lemma 9.1, we have that, with probability $1 - o(n^{-2})$,

$$|\{K' \in \mathcal{K}_{\bar{d}}(\bar{F}) : (\bar{F}, K') \notin \mathcal{P}'\}| \leq \frac{1}{n^6} \cdot |\mathcal{K}_{\bar{d}}(\bar{F})|.$$

Thus, as $|\{K' \in \mathcal{K}_d(n) : (F, K') \notin \mathcal{P} \wedge (F \subset K')\}| = |\{K' \in \mathcal{K}_{\bar{d}}(\bar{F}) : (\bar{F}, K') \notin \mathcal{P}'\}|$ and $|\mathcal{K}_{\bar{d}}(\bar{F})| = |\{K' \in \mathcal{K}_d(n) : F \subset K'\}|$, (58) holds with probability $1 - o(n^{-2})$. □

10.6 Proof of Lemma 2.2

Finally, then, we use our work so far in this section to prove Lemma 2.2 via a switching argument.

Proof of Lemma 2.2. As set-up in Lemma 2.2, let $1/n \ll 1/C \ll \eta, \varepsilon \ll 1$, let $d \geq C \log n$ and $\varepsilon dn \leq m \leq dn/2$, and let $F \sim F^-(n, d, m)$. Let $\mathcal{K}'(F)$ be the set of $K \in \mathcal{K}_d(n)$ for which $F \subset K$ and the following holds.

T1 For each distinct $x, y \in [n]$ and each $Z \subset [n] \setminus \{x, y\}$ with $|Z| \leq 10$ such that if $d \geq n - 10^3/\eta$, then $N_{K_n \setminus K}(x, Z) = \emptyset$, the following holds. The number of $(K_n \setminus K, K \setminus F)$-alternating x, y-paths of length 4 is $\left(1 \pm \frac{\eta}{100}\right)(2m)^2(n-1-d)^2/n^3$.

By Lemma 10.2 and Lemma 10.3, we have that, with probability $1 - o(n^{-2})$,

$$|\mathcal{K}'(F)| \geq \left(1 - \frac{1}{n^6}\right) |\{K \in \mathcal{K}_d(n) : F \subset K\}|. \tag{59}$$

By a simple application of a Chernoff bound and a union bound (and working via a graph F' as in the proof of Lemma 10.2), with probability $1 - o(n^{-2})$, F has the following property.

T2 For each $v \in [n]$, $d_{K \setminus F}(v) = \left(1 \pm \frac{\eta}{100}\right) 2m/n$.

Assuming that **T2** and (59) hold, we will now prove that the property in the lemma holds. We start by proving the following claim.

Claim 10.4. For each $e \in E(K_n \setminus F)$, we have

$$|\{K \in \mathcal{K}_d(n) : F + e \subset K\}| \geq \left(1 - \frac{\eta}{2}\right) \frac{2m}{2m + n(n-1-d)} |\{K \in \mathcal{K}_d(n) : F \subset K\}|.$$

Proof of Claim 10.4. Let $e \in E(K_n \setminus F)$, and suppose, for contradiction, that

$$|\{K \in \mathcal{K}_d(n) : F + e \subset K\}| < \left(1 - \frac{\eta}{2}\right) \frac{2m}{2m + n(n-1-d)} |\{K \in \mathcal{K}_d(n) : F \subset K\}|. \tag{60}$$

Let $\mathcal{K}_e^+ = \{K \in \mathcal{K}_d(n) : F + e \subset K\}$ and $\mathcal{K}_e^- = \{K \in \mathcal{K}_d(n) : F \subset K, e \notin E(K)\}$. Let L be the auxiliary bipartite graph with vertex classes $\mathcal{K}_e^+$ and $\mathcal{K}_e^-$ where there is an edge between $K \in \mathcal{K}_e^+$ and $K' \in \mathcal{K}_e^-$ if $E(K) \triangle E(K')$ is the edge set of a cycle of length 6.

Let $K \in \mathcal{K}_e^+$. Then, we have $e \in E(K \setminus F)$. Let v_1, v_6 be such that $e = v_1 v_6$. As the number of sequences v_2, v_3, v_4, v_5 with $v_2 v_3 \in E(K)$, $v_4 \in N_{K_n \setminus K}(v_3)$, $v_5 \in N_{K_n \setminus K}(v_6)$ is at most $(n-1-d)^2 dn$, we have

$$d_L(K) \leq (n-1-d)^2 dn \leq \frac{m(n-1-d)^2}{\varepsilon}.$$

Now, suppose $K \in K_e^+$ is such that **T1** holds. Then, by picking a neighbour v_2 of v_1 in $K_n \setminus K$ with at most $(n-1-d)$ options, and then applying **T1**, we have

$$d_L(K) \leq (n-1-d) \cdot \left(1 + \frac{\eta}{10}\right) \frac{(2m)^2 (n-1-d)^2}{n^3}.$$

Note that by (60) we have $|K_e^-| = |\{K \in \mathcal{K}_d : F \subset K\}| - |K_e^+| > \frac{|\{K \in \mathcal{K}_d : F \subset K\}|}{2m}$. Thus we have

$$\begin{aligned} e(L) &\leq |\mathcal{K}_e^+| \cdot \left(1 + \frac{\eta}{10}\right) \frac{(2m)^2 (n-1-d)^3}{n^3} + \frac{|\{K \in \mathcal{K}_d : F \subset K\}|}{n^6} \cdot \frac{m(n-1-d)^2}{\varepsilon} \\ &\leq |\mathcal{K}_e^+| \cdot \left(1 + \frac{\eta}{10}\right) \frac{(2m)^2 (n-1-d)^3}{n^3} + \frac{2|\mathcal{K}_e^-|}{\varepsilon (2m) \cdot 2n^2} \cdot \frac{(2m)^3 (n-1-d)^2}{n^4} \\ &\leq |\mathcal{K}_e^+| \cdot \left(1 + \frac{\eta}{10}\right) \frac{(2m)^2 (n-1-d)^3}{n^3} + |\mathcal{K}_e^-| \cdot \frac{\eta}{10} \cdot \frac{(2m)^3 (n-1-d)^2}{n^4}. \end{aligned} \tag{61}$$

Furthermore, by (59) and (60), for at least $(1 - \frac{\eta}{50})|K_e^-|$ graphs $K \in \mathcal{K}_e^-$ we have that **T1** holds. Let $K \in \mathcal{K}_e^-$ be such a K and let v_1, v_6 be such that $e = v_1 v_6$. Choose a neighbour v_2 of v_1 in $K \setminus F$ such that if $d \geq n - 10^3/\eta$ then $v_2 \notin N_{K_n \setminus K}(\{v_6\})$. If $d \geq n - 10^3/\eta$, then the number of possibilities is, by **T2** and as $m/n \geq \varepsilon d \geq C\varepsilon \log n$ at least

$$\left(1 - \frac{\eta}{100}\right) \frac{2m}{n} - |N_{K_n \setminus K}(\{v_1, v_6\})| \geq \left(1 - \frac{\eta}{100}\right) \frac{2m}{n} - \frac{2 \cdot 10^3}{\eta} \geq \left(1 - \frac{\eta}{50}\right) \frac{2m}{n},$$

and we have that $N_{K_n \setminus K}(v_2, \{v_1, v_6\}) = \emptyset$. When $d < n - 10^3\eta$, the number of choices for v_2 is, by **T2** at least $\left(1 - \frac{\eta}{100}\right) 2m/n$, so in either case there are at least $\left(1 - \frac{\eta}{50}\right) 2m/n$ choices. Then, choose a $(K_n \setminus K, K \setminus F)$-alternating v_2, v_6-path with length 4 avoiding v_1 so that, as **T1** holds for K, there are at least $\left(1 - \frac{\eta}{100}\right)(2m)^2(n-1-d)^2/n^3$ possibilities. Thus,

$$d_L(K) \geq \left(1 - \frac{\eta}{50}\right) \frac{2m}{n} \cdot \left(1 - \frac{\eta}{100}\right) \frac{(2m)^2(n-1-d)^2}{n^3} \geq \left(1 - \frac{\eta}{20}\right) \frac{(2m)^3(n-1-d)^2}{n^4}.$$

As this holds for at least $(1 - \frac{\eta}{50})|K_e^-|$ graphs $K \in \mathcal{K}_e^-$, we have

$$e(L) \geq \left(1 - \frac{\eta}{10}\right) |\mathcal{K}_e^-| \cdot \frac{(2m)^3(n-1-d)^2}{n^4}.$$

In combination with (61), we have

$$\left(1 + \frac{\eta}{5}\right) \cdot |\mathcal{K}_e^+| \cdot \frac{(2m)^2(n-1-d)^3}{n^3} \geq \left(1 - \frac{\eta}{5}\right) |\mathcal{K}_e^-| \cdot \frac{(2m)^3(n-1-d)^2}{n^4}.$$

Therefore,

$$|\mathcal{K}_e^+| \geq \left(1 - \frac{\eta}{2}\right) \frac{2m}{n(n-1-d)} |\mathcal{K}_e^-| = \left(1 - \frac{\eta}{2}\right) \frac{2m}{n(n-1-d)} (|\{K \in \mathcal{K}_d(n) : F \subset K\}| - |\mathcal{K}_e^+|),$$

so that

$$|\mathcal{K}_e^+| \cdot \frac{2m + n(n-1-d)}{n(n-1-d)} \geq \left(1 - \frac{\eta}{2}\right) \frac{2m}{n(n-1-d)} |\{K \in \mathcal{K}_d(n) : F \subset K\}|,$$

and thus

$$|\mathcal{K}_e^+| \geq \left(1 - \frac{\eta}{2}\right) \frac{2m}{2m + n(n-1-d)} |\{K \in \mathcal{K}_d(n) : F \subset K\}|,$$

as required. ⊡

Now, if $n - 1 - d \leq \eta m/10n$, then, from Claim 10.4, we have for each $f \in E(K_n \setminus F)$ that

$$\begin{aligned}(1+\eta)|\{K \in \mathcal{K}_d(n) : F + f \subset K\}| &\geq (1+\eta) \cdot \left(1 - \frac{\eta}{2}\right) \frac{2m}{2m + \eta m/10} |\{K \in \mathcal{K}_d(n) : F \subset K\}| \\ &\geq |\{K \in \mathcal{K}_d(n) : F \subset K\}|,\end{aligned}$$

so that (2) holds for every $e \in E(K_n \setminus F)$. Assume, then, that $n - 1 - d > \eta m/10n$. As $m \geq \varepsilon dn$, this implies that $n - 1 - d > \eta\varepsilon d/10 \geq C\eta\varepsilon \log n/10 \geq \sqrt{C} \log n$. Note that this implies that $d < \frac{n}{1+\sqrt{C}\log n} < n - 10^3/\eta$. Next we prove the following claim.

Claim 10.5. Let $e, f \in E(K_n \setminus F)$ share no vertices and let $n - 1 - d > \eta m/10n$. Then,

$$|\{K \in \mathcal{K}_d(n) : F + e \subset E(K)\}| \leq \left(1 + \frac{\eta}{3}\right) |\{K \in \mathcal{K}_d(n) : F + f \subset E(K)\}|. \tag{62}$$

Proof of Claim 10.5. Fixing such $e, f \in E(K_n \setminus F)$, let $\mathcal{K}^e = \{K \in \mathcal{K}_d(n) : F + e \subset K, f \notin E(K)\}$ and $\mathcal{K}^f = \{K \in \mathcal{K}_d(n) : F + f \subset K, e \notin E(K)\}$. Note that (62) follows if we have that $|\mathcal{K}^e| \leq \frac{\eta}{3}|\{K \in \mathcal{K}_d(n) : F + f \subset K\}|$, so we can assume otherwise. Thus, by Claim 10.4, we can assume that

$$|\mathcal{K}^e| > \frac{\eta}{3}|\{K \in \mathcal{K}_d(n) : F + f \subset K\}| \geq \frac{\eta}{3n}|\{K \in \mathcal{K}_d(n) : F \subset K\}|. \tag{63}$$

Form the auxiliary bipartite graph L with vertex classes $\mathcal{K}^e$ and $\mathcal{K}^f$, where there is an edge between $K \in \mathcal{K}^e$ and $K' \in \mathcal{K}^f$ if $E(K) \triangle E(K')$ is the edge set of a 10-cycle in which e and f are opposing edges of the cycle (see Figure 1).

Suppose that $K \in \mathcal{K}^e$ satisfies **T1**. Let x_1, x_2, y_1, y_2 be such that $e = x_1x_2$ and $f = y_1y_2$. By **T1** applied with $x = x_1$, $y = y_1$ and $Z = \{x_2, y_2\}$, we have that the number of $(K_n \setminus K, K \setminus F)$-alternating x_1, y_1-paths P of length 4 avoiding $\{x_2, y_2\}$ is at least $\left(1 - \frac{\eta}{100}\right) \cdot (2m)^2(n-1-d)^2/n^3$. Pick such a path P. By **T1**, the number of $(K_n \setminus K, K \setminus F)$-alternating x_2, y_2-paths of length 4 avoiding $V(P)$ is at at least $\left(1 - \frac{\eta}{100}\right) \cdot (2m)^2(n-1-d)^2/n^3$ Therefore, we have that

$$d_L(K) \geq \left(\left(1 - \frac{\eta}{100}\right) \cdot \frac{(2m)^2(n-1-d)^2}{n^3}\right)^2 \geq \left(1 - \frac{\eta}{6}\right) \cdot \frac{16m^4(n-1-d)^4}{n^6}.$$

As, by (63), $|\mathcal{K}^e| > \frac{\eta}{3n}|\{K \in \mathcal{K}_d(n) : F \subset K\}|$, this holds for at least $\left(1 - \frac{\eta}{50}\right)|\mathcal{K}^e|$ different $K \in \mathcal{K}^e$ by (59). Therefore, we have

$$e(L) \geq \left(1 - \frac{\eta}{50}\right)|\mathcal{K}^e| \cdot \left(1 - \frac{\eta}{6}\right) \frac{16m^4(n-1-d)^4}{n^6} \geq |\mathcal{K}^e| \cdot \left(1 - \frac{\eta}{5}\right) \frac{16m^4(n-1-d)^4}{n^6}. \tag{64}$$

On the other hand, for each $K' \in \mathcal{K}^f$, we have (working similarly to the way we did in Claim 10.4), $d_L(K') \leq (d^2(n-1-d)^2)^2$. Furthermore, for each $K' \in \mathcal{K}^f$ for which **T1** holds, we have $d_L(K') \leq \left(1 + \frac{3\eta}{100}\right) \cdot \frac{16m^4(n-1-d)^4}{n^6}$. As, by (59) and (63), this latter bound holds for all but $|\{K \in \mathcal{K}_d(n) : F \subset K\}|/n^6 \leq |\mathcal{K}^e|/n^3$ graphs $K' \in \mathcal{K}^f$, we have

$$\begin{aligned}
e(L) &\leq |\mathcal{K}^f| \cdot \left(1 + \frac{3\eta}{100}\right) \cdot \frac{16m^4(n-1-d)^4}{n^6} + \frac{|\mathcal{K}^e|}{n^3} \cdot (d^2(n-1-d)^2)^2 \\
&= |\mathcal{K}^f| \cdot \left(1 + \frac{3\eta}{100}\right) \cdot \frac{16m^4(n-1-d)^4}{n^6} + |\mathcal{K}^e| \cdot \frac{d^4n^3}{16m^4} \cdot \frac{16m^4(n-1-d)^4}{n^6} \\
&\leq |\mathcal{K}^f| \cdot \left(1 + \frac{3\eta}{100}\right) \cdot \frac{16m^4(n-1-d)^4}{n^6} + |\mathcal{K}^e| \cdot \frac{\eta}{100} \cdot \frac{16m^4(n-1-d)^4}{n^6},
\end{aligned}$$

where we have used that $m \geq \varepsilon dn$.

In combination with (64), we have that

$$|\mathcal{K}^e| \cdot \left(1 - \frac{\eta}{5} - \frac{\eta}{100}\right) \frac{16m^4(n-1-d)^4}{n^6} \leq |\mathcal{K}^f| \cdot \left(1 + \frac{3\eta}{100}\right) \frac{16m^4(n-1-d)^4}{n^6},$$

and, hence, $|\mathcal{K}^e| \leq (1 + \eta/3)|\mathcal{K}^f|$, so that

$$\begin{aligned}
|\{K \in \mathcal{K}_d(n) : F + e \subset K\}| &\leq |\mathcal{K}^e| + |\{K \in \mathcal{K}_d(n) : F + e + f \subset K\}| \\
&\leq \left(1 + \frac{\eta}{3}\right)|\mathcal{K}^f| + |\{K \in \mathcal{K}_d(n) : F + e + f \subset K\}| \\
&\leq \left(1 + \frac{\eta}{3}\right)|\{K \in \mathcal{K}_d(n) : F + f \subset K\}|,
\end{aligned}$$

as required. ⊡

Let, then $e, f \in [n]^{(2)} \setminus E(F)$. Pick some $e' \notin E(F)$ which is vertex-disjoint from e and f. Then, by Claim 10.5 applied twice, we have

$$\begin{aligned}
|\{K \in \mathcal{K}_d(n) : F + e \subset E(K)\}| &\leq \left(1 + \frac{\eta}{3}\right)|\{K \in \mathcal{K}_d(n) : F + e' \subset E(K)\}| \\
&\leq \left(1 + \frac{\eta}{3}\right)^2|\{K \in \mathcal{K}_d(n) : F + f \subset E(K)\}|,
\end{aligned}$$

so that, as $\eta \ll 1$, (2) holds, as required. □

References

[1] E. A. Bender. The asymptotic number of non-negative integer matrices with given row and column sums. *Discrete Mathematics*, 10(2):217–223, 1974.

[2] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.

[3] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[4] B. Bollobás. Almost all regular graphs are Hamiltonian. *European Journal of Combinatorics*, 4(2):97–106, 1983.

[5] B. Bollobás. The evolution of sparse graphs. *Graph theory and combinatorics (Cambridge, 1983)*, pages 35–57, 1984.

[6] C. Cooper, A. Frieze, and B. Reed. Random regular graphs of non-constant degree: Connectivity and Hamiltonicity. *Combinatorics, Probability and Computing*, 11(3):249–261, 2002.

[7] C. Cooper, A. Frieze, B. Reed, and O. Riordan. Random regular graphs of non-constant degree: Independence and chromatic number. *Combinatorics, Probability and Computing*, 11(4):323–341, 2002.

[8] B. Cuckler and J. Kahn. Hamiltonian cycles in Dirac graphs. *Combinatorica*, 29(3):299–326, 2009.

[9] A. Dudek, A. Frieze, A. Ruciński, and M. Šileikis. Embedding the Erdős–Rényi hypergraph into the random regular hypergraph and Hamiltonicity. *Journal of Combinatorial Theory, Series B*, 122:719–740, 2017.

[10] P. Erdős and A. Rényi. On random graphs I. *Publ Math Debrecen*, 6:290–297, 1959.

[11] A. Frieze. Hamilton cycles in random graphs: a bibliography. *arXiv preprint arXiv:1901.07139*, 2019.

[12] P. Gao. The number of perfect matchings, and the nesting properties, of random regular graphs. *Random Structures & Algorithms*, 62(4):935–955, 2023.

[13] P. Gao, M. Isaev, and B. McKay. Kim-Vu's sandwich conjecture is true for $d \geq \log^4 n$. *arXiv preprint arXiv:2011.09449*, 2023.

[14] P. Gao, M. Isaev, and B. D. McKay. Sandwiching random regular graphs between binomial random graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 690–701. SIAM, 2020.

[15] P. Gao, M. Isaev, and B. D. McKay. Sandwiching dense random regular graphs between binomial random graphs. *Probability Theory and Related Fields*, 184(1-2):115–158, 2022.

[16] L. Hollom, L. Lichev, A. Mond, J. Portier, and Y. Wang. Monotonicity and decompositions of random regular graphs. *arXiv preprint arXiv:2505.22875*, 2025.

[17] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. John Wiley & Sons, 2011.

[18] F. Joos and M. Kühn. Fractional cycle decompositions in hypergraphs. *Random Structures & Algorithms*, 61(3):425–443, 2022.

[19] J. H. Kim and V. H. Vu. Concentration of multivariate polynomials and its applications. *Combinatorica*, 20(3):417–434, 2000.

[20] J. H. Kim and V. H. Vu. Sandwiching random graphs: universality between random graph models. *Advances in Mathematics*, 188(2):444–469, 2004.

[21] T. Klimošová, C. Reiher, A. Ruciński, and M. Šileikis. Sandwiching biregular random graphs. *Combinatorics, Probability and Computing*, 32(1):1–44, 2023.

[22] J. Komlós and E. Szemerédi. Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete mathematics*, 43(1):55–63, 1983.

[23] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald. Random regular graphs of high degree. *Random Structures & Algorithms*, 18(4):346–363, 2001.

[24] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, pages 195–248. Springer, 1998.

[25] B. D. McKay. Subgraphs of random graphs with specified degrees. *Congressus Numerantium*, 33:213–223, 1981.

[26] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs of high degree. *European Journal of Combinatorics*, 11(6):565–580, 1990.

[27] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[28] R. W. Robinson and N. C. Wormald. Almost all cubic graphs are Hamiltonian. *Random Structures & Algorithms*, 3(2):117–125, 1992.

[29] R. W. Robinson and N. C. Wormald. Almost all regular graphs are Hamiltonian. *Random Structures & Algorithms*, 5(2):363–374, 1994.

[30] E. Shamir and E. Upfal. Large regular factors in random graphs. In *North-Holland Mathematics Studies*, volume 87, pages 271–282. Elsevier, 1984.

[31] N. C. Wormald. The asymptotic connectivity of labelled regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):156–167, 1981.

[32] N. C. Wormald. The asymptotic distribution of short cycles in random regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):168–182, 1981.

[33] N. C. Wormald. Models of random regular graphs. *London Mathematical Society lecture note series*, pages 239–298, 1999.