On random irregular subgraphs

Jacob Fox* Sammy Luo† Huy Tuan Pham‡

Abstract

Let G be a d-regular graph on n vertices. Frieze, Gould, Karoński and Pfender began the study of the following random spanning subgraph model $H = H(G)$. Assign independently to each vertex v of G a uniform random number $x(v) \in [0, 1]$, and an edge (u, v) of G is an edge of H if and only if $x(u) + x(v) \geq 1$. Addressing a problem of Alon and Wei, we prove that if $d = o(n/(\log n)^{12})$, with high probability, for each nonnegative integer $k \leq d$, there are $(1 + o(1))n/(d + 1)$ vertices of degree k in H.

1 Introduction

Given a graph H and a nonnegative integer k, let $m(H, k)$ be the number of vertices in H with degree k, and let $m(H) = \max_k m(H, k)$. It is well-known that in any finite simple graph H, there is some pair of vertices with the same degree, that is, $m(H) \geq 2$. Alon and Wei [1] initiated the study of the following question: Given a d-regular graph G on n vertices, can we find a spanning subgraph H of G such that $m(H)$ is as small as possible? Since such a subgraph H has maximum degree at most d, we clearly have $m(H) \geq \frac{n}{d+1}$. Alon and Wei showed that this lower bound is in fact nearly optimal by proving that every d-regular graph G has a spanning subgraph H satisfying $m(H) \leq (1 + o_n(1))\frac{n}{d+1} + 2$ (in fact, it suffices to assume that G has minimum degree d). They conjectured that for regular graphs, one can guarantee a stronger conclusion, that $m(H, k)$ is close to $\frac{n}{d+1}$ for all $0 \leq k \leq d$.

Conjecture 1.1 ([1], Conjecture 1.1)**.** *Every d-regular graph G on n vertices contains a spanning subgraph H such that*

$$\left|m(H, k) - \frac{n}{d + 1}\right| \leq 2 \qquad \textit{for all } 0 \leq k \leq d.$$

Loosely speaking, we describe a subgraph H as *irregular* when each $m(H, k)$ is close to $\frac{n}{d+1}$, as the degrees of its vertices are close to as uniformly distributed as possible. While the argument of Alon and Wei provides a nearly optimal upper bound to $m(H)$, it does not yield even the following weaker asymptotic version of Conjecture 1.1 that they ask.

Conjecture 1.2 ([1])**.** *Let $d = o(n)$. Then every d-regular graph G on n vertices contains a spanning subgraph H such that*

$$m(H, k) = (1 + o(1))\frac{n}{d + 1} \qquad \textit{for all } 0 \leq k \leq d. \qquad (*)$$

*Department of Mathematics, Stanford University, Stanford, CA 94305. Email: jacobfox@stanford.edu. Research supported by a Packard Fellowship and by NSF Awards DMS-1800053 and DMS-2154169.

†Department of Mathematics, Stanford University, Stanford, CA 94305. Email: sammyluo@stanford.edu. Research supported by NSF GRFP Grant DGE-1656518.

‡Department of Mathematics, Stanford University, Stanford, CA 94305. Email: huypham@stanford.edu. Research supported by a Two Sigma Fellowship.

While both conjectures remain open, significant progress has been made on Conjecture 1.2 in certain ranges of d. In fact, an earlier work of Frieze, Gould, Karoński and Pfender [6] already gave a method for constructing a random spanning subgraph H satisfying ($*$) whenever $d \leq (n/\log n)^{1/4}$. This construction was motivated by their study of the *irregularity strength* $s(G)$ of a graph G, defined as the smallest integer w such that one can assign integer weights in $[w]$ to every edge of G so that the sums of the weights incident to each vertex are distinct. Faudree and Lehel conjectured in [4] that there is some absolute constant C such that $s(G) \leq \frac{n}{d} + C$ whenever G is a d-regular graph on n vertices; see [8, 9, 10] for recent progress on this conjecture.

The natural construction considered in [6] is as follows: For each vertex $v \in V(G)$, pick a weight $x(v) \in [0, 1]$ uniformly and independently at random. Then, for each edge $(u, v) \in E(G)$, we keep it as an edge of H if and only if $x(u) + x(v) \geq 1$. We will henceforth refer to this construction as the *irregular random subgraph* model. It is easy to check that this model satisfies at least the basic properties one would expect from a construction of a very irregular subgraph. Indeed, fix $k \in \{0, 1, \ldots, d\}$ and consider the random variable $X = m(H, k)$. For a vertex v, let $\mathbb{I}_v$ be the indicator random variable for the event that v has degree k in H. Then, $X = \sum_{v \in V(G)} \mathbb{I}_v$. Observe that v has degree k in H if and only if, among the set of $d+1$ numbers consisting of $1 - x(v)$ and the d numbers $x(u)$ with u a neighbor of v in G, $1 - x(v)$ is the $(k+1)^{th}$ largest. It follows that v has degree k in H with probability $1/(d+1)$. By linearity of expectation, $\mathbb{E}[X] = n/(d+1)$. To give an asymptotically tight bound on $m(H, k)$, it then remains to show that the random variable X is well-concentrated around this mean. The authors of [6] do so via an application of the Azuma-Hoeffding inequality. In [1], Alon and Wei show a larger range of values of d over which ($*$) holds with high probability in this construction by using a more careful concentration argument. They start by equitably partitioning the vertices of G into $O(d^2)$ disjoint sets V_i such that within each V_i, each pair of vertices has distance at least 3, and thus the degrees in H of all vertices in V_i are independent. This allows them to apply Chernoff's inequality to show that ($*$) holds with high probability whenever $d = o((n/\log n)^{1/3})$.

In this paper, we take a more direct approach to studying the concentration properties of this irregular random subgraph model, showing the desired asymptotic irregularity property for an even wider range of values of d. First, we take a natural approach of applying the second moment method to study the variance of the random variable X. We show that $\mathrm{Var}(X) \leq \frac{17n}{d+1} = 17\mathbb{E}[X]$, implying by Chebyshev's inequality and the union bound that ($*$) holds with high probability whenever $d = o(n^{1/2})$. This argument is detailed in Section 2. However, one might conjecture much stronger concentration bounds to hold due to exponential-type decay in the distribution of X. By using a much more technical argument employing the martingale Bernstein inequality, we manage to extend the result to all $d = o(n/(\log n)^{12})$, showing the following main theorem.

Theorem 1.3. *Let G be a d-regular graph on n vertices with $d = o(n/(\log n)^{12})$. Then with high probability, the irregular random subgraph $H = H(G)$ satisfies that for $0 \leq k \leq d$, the number of vertices with degree k in H is $(1 + o(1))n/(d + 1)$.*

In particular, this theorem confirms Conjecture 1.2 for all $d = o(n/(\log n)^{12})$. We conjecture that the irregular random subgraph H satisfies Property $*$ with high probability for all $d = o(n/\log n)$. On the other hand, it is likely that for $d = \omega(n/\log n)$, the irregular random subgraph model does not satisfy Property $*$ with high probability, and new ideas would be required to show Conjecture 1.2 in the full range.

The proof of Theorem 1.3 is in Section 3. Since these two arguments are more or less independent, they can be read in any desired order.

2 The second moment argument

In this section, we prove Theorem 1.3 in the case $d = o(n^{1/2})$ through a second moment argument. We establish the following bound on the variance of the random variable $X = m(H, k)$.

Theorem 2.1. *Let G be a d-regular graph on n vertices, and let $H = H(G)$ be the irregular random subgraph. Then for all $0 \leq k \leq d$, the variance of the random variable $X = m(H, k)$ is at most $17\frac{n}{d+1}$.*

Recall that $X = \sum_{v \in V(G)} \mathbb{I}_v$, so

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = n/(d+1) + \sum_{u \neq v} \mathbb{E}[\mathbb{I}_u \mathbb{I}_v] - (n/(d+1))^2. \tag{1}$$

We begin by establishing the following key lemma. It gives an upper bound on the probability that two given vertices each have degree k in the irregular random subgraph model that depends on the codegree of the two vertices in G. Recall that the codegree $d(u, v)$ is the number of common neighbors of vertices u and v in graph G. For a vertex u of G, let $N(u)$ denote the neighborhood of u in G and $N_+(u) = N(u) \cup \{u\}$. We also use $u \sim v$ to denote that vertices u and v are adjacent in graph G.

Lemma 2.2. *For any two vertices u and v,*

$$\mathbb{E}[\mathbb{I}_u \mathbb{I}_v] \leq \frac{1}{(d+1)^2}\left(1 + \frac{16d(u, v)}{d+1}\right).$$

Proof. We have $\mathbb{I}_u = \mathbb{I}_v = 1$ if and only if, for each $w \in \{u, v\}$, $1 - x(w)$ is the $(k+1)^{th}$ largest number among $\{1 - x(w), x(z) : z \sim w\}$. Let $N_u = N(u) \setminus (N(v) \cup \{v\})$, $N_v = N(v) \setminus (N(u) \cup \{u\})$, and $N_{u,v} = N(u) \cap N(v)$. Let $t = d(u, v)$. We consider two cases based on the value of t.
Case 1: $t \geq d/2$.

Condition on an arbitrary realization of $x(z)$ for all $z \in N_u \cup N_v$, and a realization of the set of values $S = \{1 - x(u), 1 - x(v), x(z) : z \in N_{u,v}\}$ (but not the specific value of each $x(z)$ in this set, nor of $x(u)$ or $x(v)$). As $1 - x(u), 1 - x(v), x(z)$ for $z \in N(u) \cup N(v) \setminus \{u, v\}$ are independent and identically distributed, the conditional distribution of $(1 - x(u), 1 - x(v), x(z) : z \in N(u) \cap N(v))$ is given by a uniformly random permutation of the values in S.

Let r_u and r_v be the positions of $1 - x(u), 1 - x(v)$ realized in a random permutation of the values in S (when S is ordered in non-increasing order). When $u \nsim v$, there are at most two choices of r_u and r_v so that $\mathbb{I}_u = \mathbb{I}_v = 1$, depending on whether $1 - x(u)$ and $1 - x(v)$ need to be in the same position or different positions relative to the other $x(z)$. When $u \sim v$, depending on whether $x(u) + x(v) \geq 1$, the positions of $1 - x(u)$ and $1 - x(v)$ among $\{1 - x(u)\} \cup N_u \cup N_{u,v}$ and $\{1 - x(v)\} \cup N_v \cup N_{u,v}$ respectively must either both be k^{th} largest or both be $(k+1)^{th}$ largest. For a fixed set of values S, however, only one of these two cases can yield the correct final ordering of $\{1 - x(u), x(z) : z \sim u\}$ and $\{1 - x(v), x(z) : z \sim v\}$, since going from k^{th} largest to $(k+1)^{th}$ largest means choosing strictly smaller values for $1 - x(u)$ and $1 - x(v)$ among S, which increases $x(u) + x(v)$. Thus, in any case, there are at most two choices of r_u and r_v so that $\mathbb{I}_u = \mathbb{I}_v = 1$. Thus, the conditional probability that $\mathbb{I}_u = \mathbb{I}_v = 1$ is at most $2t!/(t+2)! < 2/t^2$.
Case 2: $t < d/2$.

Condition on an arbitrary realization of $x(z)$ for all $z \in N_{u,v}$, and for each $w \in \{u, v\}$, a realization of the set of values $S_w = \{1 - x(w), x(z) : z \in N_w\}$ (but not the specific value of each $x(z)$ in this set, nor of $x(w)$). As $1-x(u), 1-x(v), x(z)$ for $z \in N(u) \cup N(v) \setminus \{u, v\}$ are independent and identically distributed, the conditional distributions of $(1-x(u), x(z) : z \in N_u)$ and $(1-x(v), x(z) : z \in N_v)$ are uniformly and independently distributed among permutations of S_u and S_v.

Let r_u and r_v be the positions of $1-x(u), 1-x(v)$ realized in random permutations of the values in S_u and S_v respectively. If u and v are not adjacent, there is a unique choice of (r_u, r_v) so that $\mathbb{I}_u = \mathbb{I}_v = 1$, while if u and v are adjacent, as in the previous case we have two possible pairs of positions, but at most one pair can be valid given the set of values in S_u and S_v. Thus, the conditional probability that $\mathbb{I}_u = \mathbb{I}_v = 1$ is at most $1/(d-t)^2$.

Hence, in both cases, we obtain

$$\mathbb{E}[\mathbb{I}_u \mathbb{I}_v] \le \frac{1}{(d+1)^2}\left(1 + \frac{16t}{d+1}\right).$$

□

From Lemma 2.2, the desired bound on $\mathrm{Var}(X)$ follows.

Proof of Theorem 2.1. By Lemma 2.2,

$$\begin{aligned}
\mathbb{E}[X^2] &\le \frac{n}{d+1} + \sum_{u \ne v} \frac{1}{(d+1)^2}\left(1 + \frac{16d(u,v)}{d+1}\right) \\
&= \frac{n}{d+1} + \frac{16}{(d+1)^3}\sum_{u \ne v} d(u,v) + \frac{n(n-1)}{(d+1)^2} \\
&= \frac{n}{d+1} + \frac{16}{(d+1)^3} \cdot nd(d-1) + \frac{n(n-1)}{(d+1)^2} \\
&\le \frac{17n}{d+1} + \left(\frac{n}{d+1}\right)^2,
\end{aligned}$$

where in the second equality, we have used the identity $\sum_{u \ne v} d(u,v) = 2\sum_w \binom{|N(w)|}{2} = nd(d-1)$. Hence, by (1),

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \le \frac{17n}{d+1}. \tag{2}$$

□

This immediately implies the following desired result.

Corollary 2.3. *Theorem 1.3 holds in the case $d = o(n^{1/2})$.*

Proof. By Chebyshev's inequality and the bound (2), the probability that $|X - \mathbb{E}[X]| \ge z$ with $z = \epsilon n/(d+1)$ is at most

$$\frac{\mathrm{Var}(X)}{z^2} \le \frac{17(d+1)}{\epsilon^2 n}.$$

By the union bound over the $d+1$ possible values of k, the probability that there exists k such that $|m(H,k) - n/(d+1)| > \epsilon n/(d+1)$ is at most $\frac{17(d+1)^2}{\epsilon^2 n}$. For $d \le \epsilon^2\sqrt{n}/5 - 1$, we then have that with probability at least $1-\epsilon^2$, $|m(H,k) - n/(d+1)| < \epsilon n/(d+1)$ for all $k \le d$. By taking $\epsilon \to 0$ sufficiently slowly, we obtain that for $d = o(n^{1/2})$, with high probability, $m(H,k) = (1+o(1))n/(d+1)$ for all $k \le d$, as desired. □

In the next section we will see that the range of d can be significantly improved by replacing the use of Chebyshev's inequality with a much stronger concentration result that holds for the random variable X.

3 Exponential concentration via martingale Bernstein inequality

In this section, we use the martingale Bernstein inequality in order to show that the random variable $X = m(H, k)$ defined in the introduction, which is the number of vertices of degree k in the random subgraph $H = H(G)$, is exponentially concentrated on its expected value. This will allow us to prove Theorem 1.3 in full generality. Since Corollary 2.3 handles the case $d = o(n^{1/2})$, to cover the remaining range of d, in this section, we can and will assume that $d = \omega(\log n)$.

We aim to get exponential concentration for X based on the following Bernstein concentration bound for sums of martingale differences. This is essentially due to Freedman [5], but we will use the following version, which follows from combining Corollary 2.3 and Remark 2.1 in the paper by Fan, Grama and Liu [3].

Lemma 3.1. *Let $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n$ be an increasing sequence of σ-algebras. Given random variables $Y_1, \ldots, Y_n$ such that $\mathbb{E}[Y_j|\mathcal{F}_{j-1}] = 0$, let $X_i = \sum_{j=1}^{i} Y_j$, and let $M_i = \sum_{j=1}^{i} \mathbb{E}[Y_j^2 \mid \mathcal{F}_{j-1}]$. Then for any $z \geq 0$, $a > 0$ and $L > 0$,*

$$\mathbb{P}[\max_{i\leq n} |X_i - X_0| \geq z, M_n \leq L] \leq \exp\left(-\frac{1}{2}\frac{(z/a)^2}{L/a^2 + z/a}\right) + \mathbb{P}[\max_{i\leq n} |Y_i| > a].$$

Let π be a permutation of the vertices of the graph chosen uniformly at random. In the following, we identify vertices of the graph with integers in $[n]$ according to the permutation π. Let $\mathcal{F}_i = \sigma(x(u) : u \leq i)$ denote the σ-algebra generated by the random variables $x(u)$ with $u \leq i$. Let $X_i = \mathbb{E}[X \mid \mathcal{F}_i]$, so $\mathbb{E}[X] = X_0$ and $X = X_n$. Let $Y_i = X_i - X_{i-1}$ and let $M_i = \sum_{j=1}^{i} \mathbb{E}[Y_j^2 \mid \mathcal{F}_{j-1}]$. In order to apply Lemma 3.1 in our setting, we need the following two lemmas.

Lemma 3.2. *There exists a constant $C_1 > 0$ such that, with probability at least $1 - n^{-50}$, we have $|Y_i| \leq C_1 \log n$ for all i.*

Lemma 3.3. *Assume that $d = \omega(\log n)$. There exists a constant $C_2 > 0$ such that, with probability at least $1 - n^{-100}$, we have that $M_n \leq C_2(\log n)^{11} n/d$.*

We now give the proof of Theorem 1.3 assuming Lemma 3.2 and Lemma 3.3.

Proof of Theorem 1.3. Let $C_0 = \max(C_1, C_2)$. For $z > 0$, let $\mathcal{E}_1$ be the event that $|X_n - X_0| > z$, $\mathcal{E}_2$ the event that $M_n \leq C_0(\log n)^{11} n/d$, and $\mathcal{E}_3$ the event that $\sup_{j\leq n} |Y_j| \leq C_0(\log n)$. By Lemma 3.1 with $a = C_0(\log n)$ and $L := C_0(\log n)^{11} n/d$, we have for some $c > 0$ that

$$\mathbb{P}[\mathcal{E}_1 \wedge \mathcal{E}_2] \leq \exp\left(-\frac{cz^2}{(\log n)^{11} n/d + (\log n)z}\right) + \mathbb{P}[\overline{\mathcal{E}_3}].$$

Thus,

$$\begin{aligned}\mathbb{P}[\mathcal{E}_1] &\leq \mathbb{P}\left[\mathcal{E}_1 \wedge \mathcal{E}_2\right] + \mathbb{P}[\overline{\mathcal{E}_2}] \\ &\leq \exp\left(-\frac{cz^2}{(\log n)^{11} n/d + (\log n)z}\right) + 2n^{-50},\end{aligned}$$

where we used that $\mathbb{P}[\overline{\mathcal{E}_2}] \le n^{-50}$ by Lemma 3.3 and $\mathbb{P}[\overline{\mathcal{E}_3}] \le n^{-50}$ by Lemma 3.2.

In particular, for any $\epsilon > 0$, if $z = \epsilon n/(d+1)$ with $d < c'\epsilon^2 n/(\log n)^{12}$, then we have

$$\mathbb{P}[|X_n - X_0| > z] < 1/n.$$

Note that $X_0 = \mathbb{E}[X] = n/(d+1)$. By taking $\epsilon \to 0$ sufficiently slowly and using the union bound over the $d+1$ nonnegative integers $k \le d$, we obtain the following conclusion. With high probability, for all $k \le d$, the number of vertices with degree exactly k in H is $(1+o(1))n/(d+1)$. $\square$

It thus remains to prove Lemmas 3.2 and 3.3. This is done in the next two subsections.

3.1 Supremum of the martingale differences

In this subsection, we give the proof of Lemma 3.2, which gives an upper bound on the supremum of the martingale difference Y_i.

Here and throughout, we denote by C a sufficiently large absolute constant, $\kappa = C \log n$, and $k_+ = C \max(k, \kappa)$.

We will use the following three simple claims in the proof, which are standard applications of the Chernoff bound.

Claim 3.4. *Let $h > 0$. Let $I \subseteq \mathbb{R}$ be an interval of positive length. Let $x_1, \ldots, x_m$ be independent and uniformly distributed random variables in I. For a given subinterval of I of length $h|I|/m$, the probability that the number of elements x_i lying in the interval is not contained in $[h \pm \sqrt{\kappa \max(h, \kappa)}]$ is at most $\exp(-\kappa/3)$.*

Proof. The number of variables x_i contained in a fixed interval of length $h|I|/m$ is a sum of m independent $\mathrm{Ber}(h/m)$, which has expectation h. By the multiplicative form of the Chernoff bound, the probability that this deviates from h by more than $\ell := \sqrt{\kappa \max(h, \kappa)}$ is at most

$$\exp\left(-\min\left((\ell/h)^2, \ell/h\right) h/3\right) \le \exp(-\kappa/3). \qquad \square$$

Claim 3.5. *Let $h > 1$. Let $x_1, \ldots, x_m$ be independent and uniformly distributed random variables in an interval I. For each subinterval of length $h|I|/m$, the probability that there are more than $2h$ variables x_i in this interval is at most $\exp(-h/3)$.*

Proof. The number of variables x_i contained in a fixed subinterval of I of length $h|I|/m$ is a sum of m independent $\mathrm{Ber}(h/m)$, which has expectation h. By the multiplicative form of the Chernoff bound, the probability that this is more than $2h$ is at most $\exp(-h/3)$. $\square$

Claim 3.6. *Let $h > 1$. Let $x_1, \ldots, x_m$ be independent and uniformly distributed random variables in an interval I. The probability that there exists an interval of length at least $h|I|/m$ with no element x_i is at most $3m \exp(-h/3)$.*

Proof. Clearly, we may assume $h \le m$. Partition I into $t = \lceil 2m/h \rceil$ equal length intervals. This guarantees that every interval of length $h|I|/m$ contains at least one of these t intervals in the partition. The probability that a given one of these t intervals contains none of $x_1, \ldots, x_m$ is $(1 - 1/t)^m \le e^{-m/t} \le e^{-h/3}$. Taking the union bound over all t intervals, the probability that at least one of the t intervals contains none of $x_1, \ldots, x_m$ is at most $te^{-h/3} \le 3me^{-h/3}$. $\square$

Recall that for a vertex u of G, $N(u)$ denotes the neighborhood of u in G and $N_+(u) = N(u) \cup \{u\}$.

Proof of Lemma 3.2. Observe that

$$\begin{aligned}|Y_u| &= |\mathbb{E}[X|\mathcal{F}_u] - \mathbb{E}[X|\mathcal{F}_{u-1}]| \\ &= \left|\sum_{v\in N_+(u)} (\mathbb{E}[\mathbb{I}_v|\mathcal{F}_u] - \mathbb{E}[\mathbb{I}_v|\mathcal{F}_{u-1}])\right| \\ &\le \mathbb{E}\left[\sum_{v\in N(u)} \mathbb{I}_v \middle| \mathcal{F}_u\right] + \mathbb{E}\left[\sum_{v\in N(u)} \mathbb{I}_v \middle| \mathcal{F}_{u-1}\right] + 1.\end{aligned}$$

Let $S(u) = \sum_{v\in N(u)} \mathbb{I}_v$. Then it suffices to show that $S(u) \le 10\kappa = 10C\log n$ with probability at least $1-3n^{-100}$, as combining with $S(u) \le d \le n$, we have that $\mathbb{E}\left[S(u)|\mathcal{F}_u\right] \le 20\kappa$ with probability at least $1-n^{-99}$. Indeed, if $\mathbb{E}[S(u)|\mathcal{F}_u] > 20\kappa$, then $\mathbb{E}[\mathbb{I}(S(u) > 10\kappa)|\mathcal{F}_u] > 1/n$, and the probability of such an event is at most n^{-99}. Thus, we must have that $\mathbb{P}[\mathbb{E}[S(u)|\mathcal{F}_u] > 10\kappa] < n^{-99}$.

First consider the case $|k/d - 1/2| > \sqrt{\kappa k_+}/d$. Let $\mathcal{G}$ be the σ-algebra generated by $\mathbb{I}(x(v) \notin [k/d \pm \sqrt{\kappa k_+}/d])x(v)$ for all $v \in N(u)$. Condition on $\mathcal{G}$ (i.e., on the set of $v \in N(u)$ with $|x(v) - k/d| > \sqrt{\kappa k_+}/d$ and on the value of $x(v)$ for each such v) satisfying the following properties:

1. All vertices v in $N(u)$ with H-degree k have $|x(v) - k/d| \le \sqrt{\kappa k_+}/d$ with probability at least $1 - n^{-50}$ conditional on $\mathcal{G}$.

2. The number of vertices $v \in N(u)$ such that $|x(v) - k/d| \le \sqrt{\kappa k_+}/d$ is at most $4\sqrt{\kappa k_+}$.

3. The values $x(v)$ for $v \in N(u)$ satisfying $x(v) \notin [k/d \pm \sqrt{\kappa k_+}/d]$ partition $[0, 1] \setminus [k/d \pm \sqrt{\kappa k_+}/d]$ into consecutive intervals of length at most κ/d.

Claim 3.7. *The σ-algebra $\mathcal{G}$ satisfies the above three properties with probability at least $1 - n^{-100}$.*

Proof. By Claim 3.4, with probability at least $1-n^{-100}/4$, for each vertex v, the number of neighbors v' of v with $x(v') \ge 1 - x(v)$ is $dx(v) \pm \sqrt{\kappa \max(\kappa, dx(v))}$. Under such an event, if $v \in N(u)$ has H-degree k, then $k \in [dx(v) \pm \sqrt{\kappa \max(\kappa, dx(v))}]$. This implies that $x(v) \le 2k/d + \kappa/d$, from which we obtain that $|x(v) - k/d| \le 2\frac{\sqrt{\kappa}}{d}\sqrt{\max(\kappa, dx(v))} \le \sqrt{\kappa k_+}/d$.

Claim 3.5 implies that the second property holds with probability at least $1-d\exp(-2\sqrt{\kappa k_+}/3) > 1 - n^{-100}/4$. Claim 3.6 implies that the third property holds with probability at least $1 - 3d\exp(-\kappa/3) > 1 - n^{-100}/4$. $\square$

Let S be the set of vertices $v \in N(u)$ with $|x(v) - k/d| \le \sqrt{\kappa k_+}/d$. By our assumption that $\mathcal{G}$ satisfies the second property above, $|S| \le 4\sqrt{\kappa k_+}$. Partition $[k/d - \sqrt{\kappa k_+}/d, k/d + \sqrt{\kappa k_+}/d]$ into consecutive intervals with endpoints at each value in the set $\{1 - x(w) : w \in N(u), x(w) \in [1 - k/d - \sqrt{\kappa k_+}/d, 1 - k/d + \sqrt{\kappa k_+}/d]\}$, which is $\mathcal{G}$-measurable because $|k/d - 1/2| > \sqrt{\kappa k_+}/d$. Conditioned on $\mathcal{G}$, for each $v \in S$, there is at most one interval I_v in this partition for which v can have H-degree k only if $x(v) \in I_v$. Furthermore, the third property above implies that $|I_v| \le \kappa/d$. Conditional on $\mathcal{G}$, the random variables $x(v)$ for $v \in S$ are independently and uniformly distributed in the interval $[k/d - \sqrt{\kappa k_+/d}, k/d + \sqrt{\kappa k_+/d}]$. Note that $|I_v|/(2\sqrt{\kappa k_+}/d) \le 2\kappa/|S|$. By Claim 3.5, the number of $v \in S$ with $x(v) \in I_v$ is at most 4κ with probability at least $1-\exp(-\kappa/3) \ge 1-n^{-100}$. Hence, with probability at least $1 - 2n^{-100}$, there are at most 4κ neighbors of u with degree k.

Next consider the case $|k/d - 1/2| \le \sqrt{\kappa k_+}/d$. Pick three consecutive intervals J_1, J_2, J_3 whose union contains $[k/d \pm \sqrt{\kappa k_+}/d]$, with the properties that J_2 has length at most κ/d and $\frac{1}{2} \in J_2$, and J_1 and J_3 each has length $\sqrt{\kappa k_+}/d$. It follows that $1 - x \notin J_j$ for $x \in J_j$ and $j = 1, 3$. Let $\mathcal{G}$ be the σ-algebra generated by $\mathbb{I}(x(v) \notin J_1 \cup J_2 \cup J_3)x(v)$ for all $v \in N(u)$, and let $\mathcal{G}_j$ be the σ-algebra generated by $\mathbb{I}(x(v) \notin J_j)x(v)$. Condition on $\mathcal{G}$ and $\mathcal{G}_2, \mathcal{G}_3$ satisfying the following properties:

1. All vertices v in $N(u)$ with H-degree k lie in $J_1 \cup J_2 \cup J_3$ with probability at least $1 - n^{-50}$ conditional on $\mathcal{G}$.

2. There are at most 2κ vertices $v \in N(u)$ with $x(v) \in J_2$, and at most $2\sqrt{\kappa k_+}$ vertices with each of $x(v) \in J_1$ or $x(v) \in J_3$.

3. The values $x(v)$ for $v \in N(u)$ satisfying $x(v) \notin [k/d \pm \sqrt{\kappa k_+}/d]$ or $x(v) \in J_2 \cup J_3$ partition $[0, 1] \setminus J_1$ into consecutive intervals of length at most κ/d.

As in Claim 3.7, we can verify that $\mathcal{G}, \mathcal{G}_2, \mathcal{G}_3$ satisfy these properties with probability at least $1 - n^{-100}$.

Let S be the set of vertices v with $x(v) \in J_1$ (so they are the only remaining vertices with $x(v)$ not measurable in $\mathcal{G}, \mathcal{G}_2, \mathcal{G}_3$). By the second property conditioned on above, $|S| \le 2\sqrt{\kappa k_+}$. By the third property above, the vertices v with $x(v) \notin J_1$ partition $[0, 1] \setminus J_1$ into consecutive intervals each of length at most κ/d. As before, since $1 - x(v) \notin J_1$ for any $v \in S$, for each $v \in S$, conditional on $\mathcal{G}$, $\mathcal{G}_2$, and $\mathcal{G}_3$, there is at most one interval I_v in this partition for which v can have H-degree k only if $x(v) \in I_v$. As before, $|I_v|/|J_1| \le 2\kappa/|S|$, and by Claim 3.5, the number of $v \in S$ with $x(v) \in I_v$ conditional on $\mathcal{G}, \mathcal{G}_2, \mathcal{G}_3$ is at most 4κ with probability at least $1 - n^{-100}$. Thus, the number of vertices with $x(v) \in J_1$ and H-degree k is at most 4κ with probability at least $1 - n^{-100}$. A similar argument shows that the number of vertices with $x(v) \in J_3$ and H-degree k is at most 4κ with probability at least $1 - n^{-100}$. Finally, since there are at most 2κ vertices $v \in N(u)$ with $x(v) \in J_2$, with probability at least $1 - 3n^{-100}$, there are at most 10κ neighbors of u with H-degree k, completing the proof. □

3.2 The variance proxy

In this subsection, we give the proof of Lemma 3.3. Recall that

$$M_n = \sum_{j=1}^{n} \mathbb{E}[Y_j^2 | \mathcal{F}_{j-1}]. \tag{3}$$

We first develop several preliminary estimates that will be useful for the proof of Lemma 3.3. We have that

$$\mathbb{E}[Y_j^2 \mid \mathcal{F}_{j-1}] = \mathbb{E}\left[\left(\sum_{u \in N(j)} (\mathbb{E}[\mathbb{I}_u \mid \mathcal{F}_j] - \mathbb{E}[\mathbb{I}_u \mid \mathcal{F}_{j-1}])\right)^2 \middle| \mathcal{F}_{j-1}\right].$$

For an integer $h \in [0, t]$, define

$$p(x, t, h) = \binom{t}{h} x^h (1 - x)^{t-h}.$$

If $h < 0$ or $h > t$ then by convention we set $p(x,t,h) = 0$. Note that $p(x,t,h)$ is the probability that u has degree k conditioned on the events that $x(u) = x$, and among $d - t$ revealed neighbors v of u in G, there are exactly $k - h$ neighbors with $x(v) \geq 1 - x$.

For each j, let $t(u,j)$ be the number of neighbors of u that arrive after j in the ordering. We then have for u adjacent to j in G that

$$\mathbb{E}[\mathbb{I}_u \mid \mathcal{F}_j] = \mathbb{E}[p(x(u), t(u,j), k - h(u, x(u), j)) \mid \mathcal{F}_j],$$

where $h(u, x(u), j)$ is the number of neighbors v of u arriving by time j with $x(v) \geq 1 - x(u)$. In particular, if $u \leq j$, then

$$\mathbb{E}[\mathbb{I}_u \mid \mathcal{F}_j] = p(x(u), t(u,j), k - h(u, x(u), j)).$$

Note that for a neighbor u of j, u has degree k if $x(j) \geq 1 - x(u)$ and $k - 1 - h(u, x(u), j-1)$ neighbors v' of u coming after j have $x(v') \geq 1 - x(u)$; or $x(j) < 1 - x(u)$ and $k - h(u, x(u), j-1)$ neighbors v' of u coming after j have $x(v') \geq 1 - x(u)$. Thus,

$$\begin{aligned}\mathbb{E}[\mathbb{I}_u \mid \mathcal{F}_{j-1}] = \mathbb{E}[&p(x(u), t(u,j), k - h(u, x(u), j-1)) \cdot (1 - x(u)) \mid \mathcal{F}_{j-1}] \\ &+ \mathbb{E}[p(x(u), t(u,j), k - 1 - h(u, x(u), j-1)) \cdot x(u) \mid \mathcal{F}_{j-1}],\end{aligned}$$

while

$$\begin{aligned}\mathbb{E}[\mathbb{I}_u \mid \mathcal{F}_j] = \mathbb{E}[&p(x(u), t(u,j), k - h(u, x(u), j-1)) \cdot \mathbb{I}(x(j) < 1 - x(u)) \mid \mathcal{F}_{j-1}] \\ &+ \mathbb{E}[p(x(u), t(u,j), k - 1 - h(u, x(u), j-1)) \cdot \mathbb{I}(x(j) \geq 1 - x(u)) \mid \mathcal{F}_{j-1}].\end{aligned}$$

Thus, if j is adjacent to u in G,

$$\begin{aligned}&\mathbb{E}[\mathbb{I}_u \mid \mathcal{F}_j] - \mathbb{E}[\mathbb{I}_u \mid \mathcal{F}_{j-1}] \\ &= \mathbb{E}[(p(x(u), t(u,j), k - 1 - h(u, x(u), j-1)) - p(x(u), t(u,j), k - h(u, x(u), j-1))) \\ &\qquad \cdot (\mathbb{I}(x(j) \geq 1 - x(u)) - x(u)) \mid \mathcal{F}_{j-1}]. \end{aligned} \tag{4}$$

Let $\delta_k(j,u) = p(x(u), t(u,j), k - 1 - h(u, x(u), j-1)) - p(x(u), t(u,j), k - h(u, x(u), j-1))$. Note that $h(u, x(u), j-1)$ is $\sigma(\mathcal{F}_{j-1}, x(u))$-measurable. In particular, if $u \leq j - 1$, then $\delta_k(j,u)$ is $\mathcal{F}_{j-1}$-measurable, and otherwise it is determined by $\mathcal{F}_{j-1}$ and the (independent) choice of $x(u)$, and we have

$$\begin{aligned}\mathbb{E}[Y_j^2 \mid \mathcal{F}_{j-1}] &= \mathbb{E}\left[\left(\sum_{u \in N(j)} \delta_k(j,u)(\mathbb{I}(x(j) \geq 1 - x(u)) - x(u))\right)^2 \middle| \mathcal{F}_{j-1}\right] \\ &\leq 2\mathbb{E}_{x(j)}\left[\left(\sum_{u \in N(j), u \leq j-1} \delta_k(j,u)(\mathbb{I}(x(j) \geq 1 - x(u)) - x(u))\right)^2\right] \\ &\qquad + 2\mathbb{E}_{x(j),x(u)}\left[\left(\sum_{u \in N(j), u > j-1} \delta_k(j,u)(\mathbb{I}(x(j) \geq 1 - x(u)) - x(u))\right)^2\right], \end{aligned} \tag{5}$$

where in the first summand we are averaging over the uniform random variable $x(j)$, and in the second summand we are averaging over independent $x(j)$ and $x(u)$ for all $u > j - 1, u \in N(j)$.

Let

$$A_1(j) := \mathbb{E}_{x(j)}\left[\left(\sum_{u\in N(j), u\le j-1} \delta_k(j,u)(\mathbb{I}(x(j)\ge 1-x(u)) - x(u))\right)^2\right], \tag{6}$$

and let

$$A_2(j) := \mathbb{E}_{x(j),x(u)}\left[\left(\sum_{u\in N(j), u> j-1} \delta_k(j,u)(\mathbb{I}(x(j)\ge 1-x(u)) - x(u))\right)^2\right]. \tag{7}$$

In the next subsection, we will derive some useful bounds on $\delta_k(j,u)$ before turning to the upper bounds for $\sum_j A_1(j)$ and $\sum_j A_2(j)$, which complete the proof of Lemma 3.3.

3.2.1 Bounds on $\delta_k(j,u)$

We will now derive some useful bounds for $\delta_k(j,u)$ for fixed u. For convenience in notation, let $h = k-1-h(u,x(u),j-1)$, $x = x(u)$ and $t = t(u,j-1)$. Note that $m! = \Theta(\sqrt{m}(m/e)^m)$ for all $m \ge 1$ from Stirling's approximation. Letting $\alpha = h/(t-1)$, when $0 < h < t-1$, we have

$$\begin{aligned}
|\delta_k(j,u)| &= |p(x,t,h) - p(x,t,h+1)| \\
&= \frac{t!}{(h+1)!(t-h)!} x^h (1-x)^{t-h-1} \left|(1-x)(h+1) - x(t-h)\right| \\
&\ll \frac{|t(\alpha-x)+1-\alpha-x|}{(\alpha(1-\alpha)t)^{3/2}} \exp\left([\alpha\log(x/\alpha) + (1-\alpha)\log((1-x)/(1-\alpha))](t-1)\right) \\
&\ll \frac{|t(\alpha-x)|+3}{(\alpha(1-\alpha)t)^{3/2}} \exp\left([\alpha\log(x/\alpha) + (1-\alpha)\log((1-x)/(1-\alpha))](t-1)\right).
\end{aligned} \tag{8}$$

We also have the trivial bound $|\delta_k(j,u)| \le 1$, since $\delta_k(j,u)$ is the difference of two probabilities.

Claim 3.8. *There exists an absolute constant $c > 0$ such that the following holds. For each $x \in [0,1]$, the function $f(\alpha) = \alpha\log(x/\alpha) + (1-\alpha)\log((1-x)/(1-\alpha))$ satisfies*

$$f(\alpha) \le -c\min(|x-\alpha|, (1/x + 1/(1-x))(x-\alpha)^2). \tag{9}$$

Proof. Note that

$$f'(\alpha) = \log\frac{x}{\alpha} - \log\frac{1-x}{1-\alpha}, \qquad f''(\alpha) = -\frac{1}{\alpha} - \frac{1}{1-\alpha}.$$

Thus, by Taylor's theorem, noting that $f(x) = f'(x) = 0$,

$$f(\alpha) = f(x) + f'(x)(\alpha - x) + f''(\alpha')(\alpha-x)^2 = -\left(\frac{1}{\alpha'} + \frac{1}{1-\alpha'}\right)(x-\alpha)^2, \tag{10}$$

for some α' between α and x.

By symmetry about $1/2$, we can assume without loss of generality that $x \le 1/2$. If $\alpha \le 2x$, then $\frac{1}{\alpha'} + \frac{1}{1-\alpha'} \ge c\left(\frac{1}{x} + \frac{1}{1-x}\right)$ for any α' between α and x, so (10) implies (9). Otherwise, $\alpha > 2x$, in which case we can verify that for some constants $c_1, c_2 > 0$,

$$f(\alpha) = f(2x) + f(\alpha) - f(2x) \le f(2x) + (\alpha - 2x)f'(2x) \le -c_1 x - c_2(\alpha - 2x) \le -\min(c_1,c_2)|x-\alpha|. \quad \square$$

3.2.2 Bounding $\sum_j A_1(j)$ and $\sum_j A_2(j)$

Recall that

$$A_1(j) = \mathbb{E}_{x(j)}\left[\left(\sum_{u\in N(j), u\le j-1} \delta_k(j,u)(\mathbb{I}(x(j)\ge 1-x(u)) - x(u))\right)^2\right].$$

Claim 3.9. *(a) With probability at least $1-n^2\exp(-c\kappa)$ over the random permutation π, we have that*

$$\left||N(v)\cap[j]| - d\frac{j}{n}\right| \le \max\left(\sqrt{\kappa\frac{d}{n}\min(n-j,j)}, \kappa\right) \tag{11}$$

for all $v\in[n]$ and $j\in[n]$.

(b) With probability at least $1-n^4\exp(-c\kappa)$ (over the randomness of all $x(u)$ with $u\in[n]$), we have that for any interval $I\subseteq[0,1]$ of length at least $2/d$,

$$\big|\,|\{z\in N(v)\cap[j] : x(z)\in I\}| - d|I|j/n\big| \le e_{|I|} \tag{12}$$

for all $v\in[n]$ and $j\in[n]$, where

$$e_r = \kappa + \sqrt{\kappa d r j/n}.$$

Proof. We first prove the bound in (a). Note that under the random permutation π, $|N(v)\cap[j]|$ follows the hypergeometric distribution with parameters (n,d,j). By Hoeffding's Theorem (Theorem 4 of [7], that the hypergeometric distribution concentrates at least as much as the corresponding binomial distribution), we have that

$$\mathbb{P}\left(\left||N(v)\cap[j]| - d\frac{j}{n}\right| > \theta d\frac{j}{n}\right) \le \exp(-\min(\theta/3, \theta^2/2)dj/n).$$

In particular, if $\kappa < \sqrt{dj/n}$, then

$$\mathbb{P}\left(\left||N(v)\cap[j]| - d\frac{j}{n}\right| > \sqrt{\kappa dj/n}\right) \le \exp(-c\kappa).$$

If $\kappa \ge \sqrt{dj/n} \ge 1$, then

$$\mathbb{P}\left(\left||N(v)\cap[j]| - d\frac{j}{n}\right| > \kappa\sqrt{dj/n}\right) \le \exp(-c\kappa\sqrt{dj/n}) \le \exp(-c\kappa).$$

Finally, if $\kappa \ge \sqrt{dj/n}$ and $\sqrt{dj/n} < 1$, then

$$\mathbb{P}\left(\left||N(v)\cap[j]| - d\frac{j}{n}\right| > \kappa\right) \le \exp(-c\kappa).$$

Hence, in all cases,

$$\mathbb{P}\left(\left||N(v)\cap[j]| - d\frac{j}{n}\right| > \max(\sqrt{\kappa dj/n}, \kappa)\right) \le \exp(-c\kappa).$$

Finally, one has $|N(v) \cap [j]| = d - |N(v) \cap [j+1,n]|$, and an identical argument gives that

$$\mathbb{P}\left(\left||N(v) \cap [j]| - d\frac{j}{n}\right| > \max(\sqrt{\kappa d(n-j)/n}, \kappa)\right) \le \exp(-c\kappa).$$

The final bound follows from the union bound over all v and j.

Next, we prove (b). Let $|N(v) \cap [j]| = N_{v,j}$. Conditional on π, the random variable $N_{v,j,\ell,\ell'} := |\{z \in N(v) \cap [j] : x(z) \in [\ell/d, \ell'/d]\}|$ is a binomial random variable with parameters $(N_{v,j}, (\ell'-\ell)/d)$, so

$$\mathbb{E}\left[\exp(\lambda N_{v,j,\ell,\ell'})|\pi\right] = \left(\frac{\ell'-\ell}{d}(e^\lambda - 1) + 1\right)^{|N_{v,j}|}.$$

Next, we note that the function $y \mapsto \left(\frac{\ell'-\ell}{d}(e^\lambda - 1) + 1\right)^y$ is convex, so by Hoeffding's Theorem (Theorem 4 of [7]), we have

$$\begin{aligned}\mathbb{E}\left[\left(\frac{\ell'-\ell}{d}(e^\lambda - 1) + 1\right)^{|N_{v,j}|}\right] &\le \left(\frac{j}{n}\left(\frac{\ell'-\ell}{d}(e^\lambda - 1) + 1\right) + 1 - \frac{j}{n}\right)^d \\ &= \mathbb{E}\left[\exp(\lambda Z)\right],\end{aligned}$$

where Z is binomially distributed with parameters $(d, \frac{j(\ell'-\ell)}{dn})$. This yields Chernoff-type bounds for concentration of $N_{v,j,\ell,\ell'}$, from which similar to the proof of (a), we obtain that with probability at least $1 - \exp(-c\kappa)$,

$$\left||\{z \in N(v) \cap [j] : x(z) \in [\ell/d, \ell'/d]\}| - j(\ell'-\ell)/n\right| \le \frac{1}{2}\max\left(\kappa, \sqrt{\kappa j(\ell'-\ell)/n}\right). \tag{13}$$

By the union bound, we can guarantee this property for all v, j, ℓ, ℓ' with probability at least $1 - n^4\exp(-c\kappa)$.

The final conclusion follows upon noticing that, given any interval I of length at least $2/d$, I is contained in an interval $[\ell/d, \ell'/d]$ of length $(\ell'-\ell)/d \le |I| + 1/d$, and I contains an interval $[\tilde{\ell}/d, \tilde{\ell}'/d]$ of length $(\tilde{\ell}' - \tilde{\ell})/d \ge |I| - 1/d$.

□

Recall that $\kappa = C\log n$ and $k_+ = C\max(k, \kappa)$ for a sufficiently large constant C. In the rest of the section, we will work under the event $\mathcal{E}$ that the events in Claim 3.9 hold. Let $k' = \max(k, \kappa)$.

Lemma 3.10. *Let n be sufficiently large and $j \le n - C\kappa n/d$. We have*

$$\sum \delta_k(j,u) < n^{-10},$$

where the summation ranges over $u \in N(j)$ with $u \le j-1$ such that $|x(u) - k/d| > \sqrt{\kappa k_+}/d$ or $|h/(t-1) - x(u)| > \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)$.

Proof. Under the event $\mathcal{E}$, Claim 3.9(b) tells us that

$$|h(u, x(u), j-1) - x(u)(j-1)d/n| \le e_{x(u)},$$

so

$$|h - (k - 1 - x(j-1)d/n)| \le e_x. \tag{14}$$

By Claim 3.9(a),

$$|t-(d-d(j-1)/n)| \le \max(\kappa, \sqrt{\kappa d \min(n-(j-1), j-1)/n}). \tag{15}$$

From (14) and (15), we have

$$h/(t-1)-x \in \left[\frac{M_1-E_1}{M_2+E_2}, \frac{M_1+E_1}{M_2-E_2}\right], \tag{16}$$

where

$$M_1 = k-1-x(d-1), \qquad M_2 = d-1-d(j-1)/n,$$
$$E_1 = e_x + x(\kappa+\sqrt{\kappa d \min(n-(j-1), j-1)/n}), \qquad E_2 = \kappa+\sqrt{\kappa d\min(n-(j-1), j-1)/n}.$$

Note that

$$E_2 \le \kappa + \sqrt{\kappa d(n-j+1)/n} \le \kappa+\sqrt{\kappa(M_2+1)} \le M_2/2,$$

where we use the fact that $M_2 \ge C\kappa - 1$ for $j \le n - C\kappa n/d$. This, along with (15), implies that

$$t \ge M_2+1-E_2 = 1+\frac{1}{2}M_2 > \frac{1}{2}d(1-(j-1)/n). \tag{17}$$

Furthermore, recall that

$$e_x = \kappa+\sqrt{\kappa dx(j-1)/n}.$$

Let c be the constant in Claim 3.8 and let $C' = cC/6400$. Assume that C is sufficiently large so $C' > 1$. By Claim 3.8, if $|h/(t-1)-x| > \frac{1}{64C'}\kappa/(d(1-j/n))$ and $|h/(t-1)-x|^2/x > \frac{1}{64C'}\kappa/(d(1-j/n))$, then

$$\begin{aligned}
&\exp\left([\alpha\log(x/\alpha)+(1-\alpha)\log((1-x)/(1-\alpha))](t-1)\right)\\
&\le \exp\left(-cd(1-j/n)\min\left(|h/(t-1)-x|, \frac{1}{x}|h/(t-1)-x|^2\right)\right)\\
&< n^{-15},
\end{aligned} \tag{18}$$

upon choosing the constant C in the definition of κ sufficiently large.

We will next show that the conditions of (18) (and hence the conclusion) hold if $|x-k/d| > \sqrt{\kappa k_+}/d$ or $|h/(t-1)-x| > \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)$. Given this, by (8), we have that for $u \in N(j)$ with $u \le j-1$, and either $|x(u)-k/d| \ge \sqrt{\kappa k_+}/d \ge \sqrt{C'\kappa k'}/d$ or $|h/(t-1)-x(u)| > \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)$,

$$\delta_k(j,u) \ll \frac{|t(\alpha-x)|+3}{(\alpha(1-\alpha)t)^{3/2}}n^{-15} < n^{-12}.$$

Summing over at most n such u's, we obtain the desired bound.

We now turn to verify the conditions of (18) when $|x-k/d| > \sqrt{\kappa k_+}/d$ or $|h/(t-1)-x| > \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)$. Let $\eta := |x-k/d|$, and let $\eta_0 := \sqrt{C'\kappa k'}/d$.

Case 1: $|x-k/d| = \eta \ge \eta_0$.

From (16), we first claim that if $\eta \ge \eta_0$ then $|E_1| < |M_1|/2$. Indeed, if $x < 64k'/d$, then

$$|M_1| > \eta d - 2 \ge \sqrt{C'\kappa k'} - 2, \qquad |E_1| < 2(\kappa+\sqrt{\kappa d x}) \le 2(\kappa + 8\sqrt{\kappa k'}) \le 18\sqrt{\kappa k'}.$$

Otherwise, if $x \geq 64k'/d$, then $\eta < x$, and

$$|M_1| > \frac{1}{2}dx, \qquad\qquad |E_1| < 2(\kappa + \sqrt{\kappa dx}).$$

In both cases, recalling that $k' = \max(k, \kappa) \geq \kappa$, we indeed have for a sufficiently large C that $|E_1| < |M_1|/2$. Combining this estimate with the fact that $E_2 \leq M_2/2$ yields

$$4\eta/(1-j/n) > |h/(t-1) - x| > \frac{1}{4}\eta/(1-j/n). \tag{19}$$

In particular, since $k' \geq \kappa$, we have

$$|h/(t-1) - x| > \frac{1}{4}\eta_0/(1-j/n) = \frac{1}{4}\sqrt{C'\kappa k'}/(d(1-j/n)) \geq \frac{1}{64C'}\kappa/(d(1-j/n)). \tag{20}$$

For $\alpha = h/(t-1)$, if $|x - k/d| = \eta \geq \sqrt{C'\kappa k'}/d$, then

$$\frac{1}{x}|h/(t-1) - x|^2 \geq \frac{1}{16}\frac{1}{k/d + \eta}\frac{\eta^2}{(1-j/n)^2}.$$

Thus, if $\eta > (k+\kappa)/d$, then

$$\frac{1}{x}|h/(t-1) - x|^2 \geq \frac{\eta}{32(1-j/n)^2} > \frac{k+\kappa}{32d(1-j/n)^2} > \frac{1}{64C'}\frac{\kappa}{d(1-j/n)},$$

and otherwise if $\eta \leq (k+\kappa)/d$, then

$$\frac{1}{x}|h/(t-1)-x|^2 \geq \frac{1}{16}\frac{d}{2k+\kappa}\frac{\eta_0^2}{(1-j/n)^2} \geq \frac{1}{16}\frac{C'k'}{2k+\kappa}\frac{\kappa}{d(1-j/n)^2} \geq \frac{1}{48}\frac{\kappa}{d(1-j/n)} > \frac{1}{64C'}\frac{\kappa}{d(1-j/n)}.$$

The estimates above and (20) yield that the conditions of (18) hold when $|x-k/d| \geq \sqrt{C'\kappa k'}/d$.
Case 2: $|x-k/d| \leq \eta_0 = \sqrt{C'\kappa k'}/d$ and $|h/(t-1)-x| > \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)$.

Since $|x - k/d| \leq \sqrt{C'\kappa k'}/d$, we have $x \leq 2\max(k/d, \sqrt{C'\kappa k'}/d)$. Then

$$\begin{aligned}\frac{1}{x}|h/(t-1) - x|^2 &\geq \frac{1}{2\max(k/d, \sqrt{C'\kappa k'}/d)} \cdot \frac{\kappa k_+}{d^2(1-j/n)} \geq \frac{1}{2C'(k+\kappa)/d} \cdot \frac{\kappa(\kappa+k)/2}{d^2(1-j/n)} \\ &> \frac{1}{64C'}\frac{\kappa}{d(1-j/n)}.\end{aligned}$$

Furthermore,

$$|h/(t-1) - x| > \frac{\kappa}{d(1-j/n)} > \frac{1}{64C'}\frac{\kappa}{d(1-j/n)}.$$

This yields the conditions of (18) in this case and completes the proof of the lemma. □

We next return to the task of bounding $\sum_j A_1(j)$. Let $\mathcal{U}_j$ be the set of u with $|x(u) - k/d| \leq \sqrt{\kappa k_+}/d$ and $|h/(t-1)-x(u)| \leq \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)$. Note that by symmetry in degrees of H about $d/2$, we can and will assume that $k \leq d/2$. Let C'' be a sufficiently large constant whose choice depends on C. In the following, we first consider values of j such that $j \leq n - C''n\kappa/d$. For these j we have $d\sqrt{1-j/n} \geq \sqrt{C''\kappa d} \geq \sqrt{C\kappa k_+}$, since $d = \omega(\log n)$ and

$d(1-j/n) \geq C''\kappa$. Thus, for C sufficiently large and $k \leq d/2$, we have that $\min(1-\alpha, 1-k/d) \geq 1/4$ (where we recall that $\alpha = h/(t-1)$).

By Claim 3.9(b), the number of $u \in N(j) \cap [j-1]$ with $|x(u) - k/d| \leq \sqrt{\kappa k_+}/d$ is at most $\frac{dj}{n} \cdot 2\frac{\sqrt{\kappa k_+}}{d} + \left(\kappa + \sqrt{2\kappa\sqrt{\kappa k_+} j/n}\right) \leq 4\left(\kappa + \sqrt{\kappa k_+ j/n}\right)$. Furthermore, recall that

$$\begin{aligned} A_1(j) &= \mathbb{E}_{x(j)}\left[\left(\sum_{u \in N(j), u \leq j-1} \delta_k(j,u)(\mathbb{I}(x(j) \geq 1 - x(u)) - x(u))\right)^2\right] \\ &\ll \mathbb{E}_{x(j)}\left[\left(\sum_{u \in N(j), u \leq j-1, u \notin \mathcal{U}_j} \delta_k(j,u)\right)^2\right] \\ &\quad + \mathbb{E}_{x(j)}\left[\left(\sum_{u \in N(j), u \leq j-1, u \in \mathcal{U}_j} \delta_k(j,u)(\mathbb{I}(x(j) \geq 1 - x(u)) - x(u))\right)^2\right]. \end{aligned}$$

If $x(j) < 1 - k/d - \sqrt{\kappa k_+}/d$, then the second term can be bounded by

$$\mathbb{E}_{x(j)}\left[\left(\sum_{u \in N(j), u \leq j-1, u \in \mathcal{U}_j} \delta_k(j,u)(k + \sqrt{\kappa k_+})/d\right)^2\right].$$

Since $x(j) \geq 1 - k/d - \sqrt{\kappa k_+}/d$ occurs with probability at most $k/d + \sqrt{\kappa k_+}/d$, we conclude that

$$\begin{aligned} A_1(j) &\ll \mathbb{E}_{x(j)}\left[\left(\sum_{u \in N(j), u \leq j-1, u \notin \mathcal{U}_j} \delta_k(j,u)\right)^2\right] \\ &\quad + \mathbb{E}_{x(j)}\left[\left(\sum_{u \in N(j), u \leq j-1, u \in \mathcal{U}_j} \delta_k(j,u)\right)^2\right]\left(\frac{(k + \sqrt{\kappa k_+})^2}{d^2} + \frac{k + \sqrt{\kappa k_+}}{d}\right) \\ &\ll n^{-20} + \frac{k + \sqrt{\kappa k_+}}{d}\mathbb{E}_{x(j)}\left[\left(\sum_{u \in N(j), u \leq j-1, u \in \mathcal{U}_j} \delta_k(j,u)\right)^2\right], \end{aligned} \tag{21}$$

where the the bound on the first term follows from Lemma 3.10. In particular, in the following part of the argument, we only need to restrict our attention to the case where $|x - k/d| \leq \sqrt{\kappa k_+}/d$ and $|h/(t-1) - x| \leq \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)$, so $\alpha \in [k/d \pm 2\max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)]$.

Condition on $x(v)$ with $|x(v) - k/d| > \sqrt{\kappa k_+}/d$. Since the event in Claim 3.9(b) holds, the $x(v)$ with $|x(v) - k/d| > \sqrt{\kappa k_+}/d$ partition $[0,1] \backslash [k/d \pm \sqrt{\kappa k_+}/d]$ into consecutive intervals each of length at most κ/d, and the union of any q consecutive intervals has length at most $(2q + 4\kappa)/d$. Indeed, this follows since any interval of length $(2q+4\kappa)/d$ contains at least $(2q+4\kappa) - (\kappa + \sqrt{\kappa(2q+4\kappa)}) > q$ elements $x(v)$.

Since $|x(u) - k/d| \leq \sqrt{\kappa k_+/d}$, if $|h/(t-1) - x(u)| \leq \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)$, then $h/(t-1)$ is contained in an interval centered at k/d of length

$$w := 2\sqrt{\kappa k_+/d} + 2\max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right).$$

Thus, h must be contained in a fixed interval of length at most $(t-1)w$, which implies that u is contained in a (fixed) union of at most $(t-1)w$ consecutive intervals in the partition by $x(v)$ with $|x(v)-k/d| > \sqrt{\kappa k_+}/d$. (Note that each possible value of h determines an interval in the partition that $x(u)$ must lie in.) Hence, $x(u)$ must be contained in a fixed interval of length at most

$$\Big(4\kappa + 2(t-1)w\Big)/d < 12\kappa/d + 16\sqrt{\kappa k_+(1-j/n)}/d,$$

where we have used the fact that $t < 2d(n-j)/n$. The number of such u in $N(j) \cap [j-1]$ is at most $32(\kappa + \sqrt{\kappa k_+(1-j/n)})$.

Letting $\bar{t} = d(n-j)/n$, we have $t = \Theta(\bar{t})$ by Claim 3.9(a), as reasoned in the proof of Lemma 3.10. By (8) and recalling that $\min(1-\alpha, 1-k/d) \geq 1/4$, we have the bound

$$\delta_k(j,u) \ll \frac{3 + t\max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)}{(\max(1,\alpha t))^{3/2}}. \tag{22}$$

Note that if $k < 4\sqrt{\kappa k_+}/\sqrt{1-j/n}$, then $\bar{t}k/d = k(1-j/n) \ll \kappa^2$. Similarly, if $k < 4\kappa/(1-j/n)$, then $\bar{t}k/d = k(1-j/n) \ll \kappa$. On the other hand, if $k \geq \max\left(4\sqrt{\kappa k_+}/\sqrt{1-j/n}, 4\kappa/(1-j/n)\right)$, then

$$\alpha \geq k/d - \max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right) \geq k/d - k/(2d) = k/(2d).$$

Thus, in both cases, $\max(1,\alpha t) \gg \bar{t}\kappa^{-2}k/d$. Hence,

$$\begin{aligned}\delta_k(j,u) &\ll \kappa^3 \frac{3 + \bar{t}\max\left(\sqrt{\kappa k_+}/(d\sqrt{1-j/n}), \kappa/(d(1-j/n))\right)}{(k\bar{t}/d)^{3/2}} \\ &\ll \kappa^3 \frac{\kappa + \sqrt{\kappa k_+(1-j/n)}}{(k(1-j/n))^{3/2}}.\end{aligned} \tag{23}$$

We also have the trivial bound

$$\delta_k(j,u) \leq 1.$$

Hence, we have

$$\begin{aligned}
&\kappa^{-6} \sum_{j=1}^{n-C''\kappa n/d} A_1(j) - n^{-19} \\
&\ll \sum_{j=1}^{n-C''\kappa n/d} \frac{k+\sqrt{\kappa k_+}}{d} \left((\kappa + \sqrt{\kappa k_+(1-j/n)}) \cdot \min\left(1, \frac{\kappa + \sqrt{\kappa k_+(1-j/n)}}{(k(1-j/n))^{3/2}}\right)\right)^2 \\
&\ll \frac{k+\sqrt{\kappa k_+}}{d} \sum_{j\in[1,n-C''\kappa n/d]:k(1-j/n)>1} \left((\kappa + \sqrt{\kappa k(1-j/n)}) \cdot \frac{\kappa + \sqrt{\kappa k_+(1-j/n)}}{(k(1-j/n))^{3/2}}\right)^2 \\
&\qquad + \frac{k+\sqrt{\kappa k_+}}{d} \sum_{j\in[1,n-C''\kappa n/d]:k(1-j/n)\leq 1} \kappa^2 \\
&\ll \frac{k+\sqrt{\kappa k_+}}{d} \Bigg(\kappa^4 \sum_{j<n-n/k} \frac{n^3}{k^3(n-j)^3} + \kappa^3 \sum_{j<n-n/k} \frac{(k+\kappa)n^2}{k^3(n-j)^2}
\end{aligned}$$

$$
\begin{aligned}
&\qquad + \kappa^3 \sum_{j<n-n/k} \frac{n^2}{k^2(n-j)^2} + \sum_{j<n-n/k} \kappa^2 \frac{(k+\kappa)n}{k^2(n-j)} + \kappa^2 \frac{n}{k}\Bigg) \\
&\ll \kappa^4 \frac{k+\kappa}{d} \cdot \frac{n^3}{(n/k)^2 k^3} + \kappa^3 \frac{(k+\kappa)^2 n^2}{k^3 (n/k) d} + \kappa^3 \frac{k+\kappa}{d} \cdot \frac{n^2}{k^2(n/k)} \\
&\qquad + \frac{(\log k)\kappa^2 (k+\kappa)^2 n}{dk^2} + \kappa^2 \frac{n(k+\kappa)}{dk} \\
&\ll \kappa^5 n/d. \qquad (24)
\end{aligned}
$$

Here, we note that $k(1-j/n) > 1$ if and only if $j < n - n/k$. The first inequality follows from applying (21) and (23). The second inequality follows from breaking the summation into the ranges $j < n - n/k$ and $j \geq n - n/k$ and bounding each separately. Specifically, in the range $j < n - n/k$, we use the bound

$$\kappa + \sqrt{\kappa k_+ (1-j/n)} \ll \kappa + \sqrt{\kappa(\kappa+k)(1-j/n)} \ll \kappa + \sqrt{\kappa k (1-j/n)},$$

and in the range $j \geq n - n/k$, we use the bound

$$(\kappa + \sqrt{\kappa k_+ (1-j/n)}) \cdot \min\left(1, \frac{3 + \sqrt{(1-j/n)\kappa k_+}}{(k(1-j/n))^{3/2}}\right) \ll \kappa + \sqrt{\kappa k_+ (1-j/n)} \ll \kappa + \sqrt{\kappa k (1-j/n)} \ll \kappa.$$

The third inequality follows from applications of the inequality $(a+b)^2 \ll a^2 + b^2$. The fourth inequality follows from the elementary bounds $\sum_{j<n-n/k} \frac{1}{(n-j)^s} \ll (n/k)^{-s+1}$ for $s \geq 2$, and $\sum_{j<n-n/k} \frac{1}{n-j} \ll \log k$, as well as the observation $k + \sqrt{\kappa k_+} \ll k + \kappa$. In the last inequality, we use the bound $(k+\kappa)/k \ll \kappa$.

Finally, in the range $j \geq n - C''\kappa n/d$, by Lemma 3.2, we have that $A_1(j) \ll \kappa^2$ and hence

$$\sum_{j=n-C''\kappa n/d}^{n} A_1(j) \ll \kappa^2 \cdot \kappa n/d = \kappa^3 n/d. \qquad (25)$$

Combining (24) and (25), we obtain

$$\sum_{j=1}^{n} A_1(j) \ll (\log n)^{11} n/d. \qquad (26)$$

The control of

$$A_2(j) = \mathbb{E}_{x(j),x(u)}\left[\left(\sum_{u \in N(j), u > j-1} \delta_k(j,u)(\mathbb{I}(x(j) > 1 - x(u)) - x(u))\right)^2\right]$$

is very similar, and we omit the details of the argument. One first considers a realization of the random variables $x(u)$ for $u \in N(j), u > j-1$. The same argument used in bounding A_1 then allows us to prove the same bound on A_2:

$$\sum_{j=1}^{n} A_2(j) \ll (\log n)^{11} n/d. \qquad (27)$$

We are now ready to finish the proof of Lemma 3.3 and hence the proof of Theorem 1.3.

Proof of Lemma 3.3. By (3) and (5), we have

$$M_n \ll \sum_{j=1}^{n}(A_1(j) + A_2(j)) \ll (\log n)^{11} n/d,$$

where in the last inequality we have used (26) and (27). $\square$

References

[1] N. Alon and F. Wei, Irregular subgraphs, preprint, arXiv:2108.02685.

[2] B. Cuckler and F. Lazebnik, Irregularity strength of dense graphs, *J. Graph Theory* **58** (2008), 299–313.

[3] X. Fan, I. Grama, Q. Liu, Hoeffding's inequality for supermartingales, *Stochastic Process. Appl.* **122** (2012), 3545–3559.

[4] R. J. Faudree and J. Lehel, Bound on the irregularity strength of regular graphs, *Colloq. Math. Soc. János Bolyai* **52** (1988), 247–256.

[5] D. A. Freedman, On tail probabilities for martingales, *Ann. Probability* **3** (1975), 100–118.

[6] A. Frieze, R. J. Gould, M. Karoński and F. Pfender, On graph irregularity strength, *J. Graph Theory* **41** (2002), 120–137.

[7] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58** (1963), 13–30.

[8] J. Przybyło, Asymptotic confirmation of the Faudree-Lehel conjecture on irregularity strength for all but extreme degrees, *J. Graph Theory* **100**(1) (2022), 189–204.

[9] J. Przybyło and F. Wei, On the asymptotic confirmation of the Faudree-Lehel Conjecture for general graphs, preprint, arXiv:2109.04317.

[10] J. Przybyło and F. Wei, Short proof on the asymptotic confirmation of the Faudree-Lehel conjecture, preprint, arXiv:2109.13095.