Article:

Demyanov, S. and Zhukovskii, M. orcid.org/0000-0001-8763-9533 (2023) Tight concentration of star saturation number in random graphs. Discrete Mathematics, 346 (10). 113572. ISSN: 0012-365X

https://doi.org/10.1016/j.disc.2023.113572

Discrete Mathematics 346 (2023) 113572

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Tight concentration of star saturation number in random graphs

S. Demyanov [a], M. Zhukovskii [b,*]

[a] Moscow Institute of Physics and Technology, Department of Discrete Mathematics, Russian Federation
[b] The University of Sheffield, Department of Computer Science, United Kingdom

ARTICLE INFO

Article history:
Received 25 January 2023
Received in revised form 25 May 2023
Accepted 8 June 2023
Available online 28 June 2023

Keywords:
Random graphs
Saturation number
2 point concentration
Star graphs

ABSTRACT

For given graphs F and G, the minimum number of edges in an inclusion-maximal F-free subgraph of G is called the F-saturation number and denoted $\mathrm{sat}(G, F)$. For the star $F = K_{1,r}$, the asymptotics of $\mathrm{sat}(G(n, p), F)$ is known. We prove a sharper result: whp $\mathrm{sat}(G(n, p), K_{1,r})$ is concentrated in a set of 2 consecutive points.

1. Introduction

The concept of "saturation number" was introduced by Zykov [14] and then independently by Erdős, Hajnal and Moon [1]. They asked about the minimum number of edges in an F-free inclusion-maximal graph on n vertices. In other words, a graph G is *F-saturated* if G is F-free (i.e. does not contain any copy of F), but addition of any edge creates a copy of F. For example, any complete bipartite graph is a K_3-saturated graph. The minimum number of edges in an F-saturated graph is called *the F-saturation number* and is denoted by $\mathrm{sat}(n, F)$.

For example, if $F = K_m$ (i.e. a complete graph on m vertices), then $\mathrm{sat}(n, F)$ is known, this result was obtained by Erdős, Hajnal and Moon [1]: for all $n \geqslant m \geqslant 2$

$$\mathrm{sat}(n, K_m) = (m-2)(n-m+2) + \binom{m-2}{2} = \binom{n}{2} - \binom{n-m+2}{2}.$$

For stars (we denote by $K_{1,r}$ a star with r leaves), the problem was solved by Kászonyi and Tuza [6]:

$$\mathrm{sat}(n, K_{1,r}) = \begin{cases} \binom{r}{2} + \binom{n-r}{2}, & r+1 \leqslant n \leqslant \frac{3r}{2}; \\ \left\lceil \frac{(r-1)n}{2} - \frac{r^2}{8} \right\rceil, & n \geqslant \frac{3r}{2}. \end{cases}$$

The notion of saturation can be generalized to arbitrary host graphs. For a given host graph G, a spanning subgraph H of G is called *F-saturated in G* if H is F-free, but every graph obtained by adding an edge from $E(G) \setminus E(H)$ to H, has at least one copy of F as a subgraph. The minimum number of edges in an F-saturated subgraph of G is denoted by $\mathrm{sat}(G, F)$.

* Corresponding author.
E-mail address: m.zhukovskii@sheffield.ac.uk (M. Zhukovskii).

https://doi.org/10.1016/j.disc.2023.113572

Thus sat(n, F) = sat(K_n, F). In [7] Korándi and Sudakov initiated the study of saturation number of "typical" host graphs. We say that a graph property Q holds with high probability (whp), if $\mathsf{P}(G(n,p)\in Q)\to 1$ as $n\to\infty$. As usual, we denote by $G(n,p)$ the binomial random graph on $[n]:=\{1,\dots,n\}$, i.e. a graph with every edge drawn independently with probability p. Korándi and Sudakov proved that whp $\mathrm{sat}(G(n,p=\mathrm{const}),K_s)=n\log_{\frac{1}{1-p}}n(1+o(1))$. In this paper, we consider only the dense setting, i.e. $p=\mathrm{const}\in(0,1)$.

For stars, the saturation number of the random graph was also studied. In [13] Zito proved that whp

$$\frac{n}{2}-\log_{\frac{1}{1-p}}(np)\leqslant \mathrm{sat}\left(G(n,p),K_{1,2}\right)\leqslant \frac{n}{2}-\log_{\frac{1}{1-p}}\left(\sqrt{n}\right).$$

Note that $\mathrm{sat}(G,K_{1,2})$ is the minimum cardinality of a maximal matching in G.

In [12] Mohammadian and Tayfeh–Rezaie prove that for any fixed $p\in(0,1)$ and any fixed integer $r\geqslant 2$ whp

$$\mathrm{sat}\left(G(n,p),K_{1,r}\right)=\frac{(r-1)n}{2}-(1+o(1))\,(r-1)\log_{\frac{1}{1-p}}n.$$

Here we want to emphasize the fact that, for any maximal matching in G, the deletion of its vertex set leaves only an independent set of G. On the other hand, it is well known that whp any large enough subset of $G(n,p)$ contains a matching (see e.g. [3, Remark 4.3]) and that the independence number of $G(n,p)$ is concentrated in a set of two consecutive points [9–11,2]: for a fixed $0<p<1$ and any $\varepsilon>0$, whp

$$\lfloor\alpha_p(n)-\varepsilon\rfloor\leqslant\alpha(G(n,p))\leqslant\lfloor\alpha_p(n)+\varepsilon\rfloor, \tag{1}$$

where $\alpha_p(n):=2\log_b n-2\log_b\log_b n+2\log_b(e/2)+1$, $b=1/(1-p)$.

Thus, whp $\mathrm{sat}(G(n,p),K_{1,2})$ is equal to a half of the size of the complement to a maximum independent set (if this size is odd, then it is 1/2 less), and so it is concentrated in a set of two consecutive points as well. We further consider $r\geqslant 3$.

In this paper, we show that for all r, whp $\mathrm{sat}(G(n,p),K_{1,r})$ is concentrated in a set of two consecutive points (this is as sharp as possible) and thus significantly improve the result of Mohammadian and Tayfeh–Rezaie [12]. We let

$$\varphi_m(k)=\binom{n}{k}\binom{\binom{k}{2}}{m}p^m(1-p)^{\binom{k}{2}-m}.$$

The main result of our paper is stated below.

Theorem 1. *Let $p\in(0,1)$ be a constant, $r\geqslant 3$. Let $\delta>0$, $0<\varepsilon'\ll\varepsilon\ll\delta$ and $n>n_0(\delta)$ be large enough. Let*

$$x_0=\lfloor\alpha_p(n)+\varepsilon\rfloor\quad\text{and}\quad r'=\left\lceil\frac{r-3}{2}\right\rceil-I(n-x_0,r-1\text{ are odd}).$$

If $\varphi_{r'}(x_0)<\varepsilon'$, then, with probability at least $1-\delta$,

$$\mathrm{sat}(G(n,p),K_{1,r})=\left\lceil\frac{(r-1)(n-x_0)}{2}\right\rceil+\mu,\quad \mu:=r'+1.$$

Otherwise, let $\mu\leqslant r'$ be the smallest non-negative integer such that $\varphi_\mu(x_0)\geqslant\varepsilon'$. Then, with probability at least $1-\delta$,

$$\mathrm{sat}(G(n,p),K_{1,r})\in\left\{\left\lceil\frac{(r-1)(n-x_0)}{2}\right\rceil+\mu,\left\lceil\frac{(r-1)(n-x_0)}{2}\right\rceil+\mu+1\right\}.$$

To prove the theorem, we first show (and this is the trickiest part of the paper) that the almost optimal strategy is

(1) to take a set that induces at most r' edges and has the maximum size, and
(2) to preserve $r-1$ edges adjacent to each of the vertices outside this set,

and, after that, constructively prove the upper bound.

The rest of the paper is organized as follows. Section 2 contains definitions and theorems needed to prove the main result. In Sections 3 and 4 the lower and upper bounds are proved respectively.

2. Preliminaries

2.1. Almost independent sets

Let $\varepsilon>0$ be small enough and let $\xi_m(k)$ be the number of sets of size k that induce exactly m edges in $G(n,p)$. Let α_m be the maximum cardinality of a set of vertices that induces exactly m edges in $G(n,p)$. In particular, $\alpha_0=\alpha(G(n,p))$ is the independence number. Note that $\varphi_m(k)=\mathsf{E}\xi_m(k)$ and that α_m is the maximum k such that $\xi_m(k)\geqslant 1$.

In [5] it was proven that, for any constant $m \in \mathbb{Z}_+$, whp α_m is concentrated in a set of 2 consecutive points. Moreover, these points are the same for different values of m: for every $m \in \mathbb{Z}_+$, whp

$$\lfloor \alpha_p(n) - \varepsilon \rfloor \leqslant \alpha_m \leqslant \lfloor \alpha_p(n) + \varepsilon \rfloor .$$

More precisely, the following is true.

Theorem 2 *([5]). Let $\varepsilon', \varepsilon, \delta, n, x_0$ be defined as in Theorem 1 with the additional requirement that $n \gg m$. We have $\alpha_m \in \{x_0 - 1, x_0\}$ with probability at least $1 - \delta$. If $\varphi_m(x_0) < \varepsilon'$, then $\alpha_m = x_0 - 1$ with probability at least $1 - \delta$. If $\varphi_m(x_0) > 1/\varepsilon'$, then $\alpha_m = x_0$ with probability at least $1 - \delta$.*

Note that, for every m, $\varphi_m(x_0) \ll \varphi_{m+1}(x_0)$. Consider separately two cases distinguished in Theorem 1.

1. If $\varphi_{r'}(x_0) < \varepsilon'$, then $\varphi_m(x_0) < \varepsilon'$ as well for all $m < r'$ implying that $\alpha_m = x_0 - 1$ for all $m \leqslant r'$ with probability at least $1 - \delta$.
2. Otherwise, $\alpha_m = x_0 - 1$ for all $m < \mu$, $\alpha_\mu \in \{x_0 - 1, x_0\}$ and $\alpha_{\mu+1} = x_0$ with probability at least $1 - \delta$.

2.2. Powers of Hamilton cycles

The ℓ-th power of a graph H is obtained by the addition to H of edges between all pairs of vertices that are at distance at most ℓ. A Hamilton ℓ-cycle in a graph G is the ℓ-th power of a Hamilton cycle in G. To prove our main result, we need the following theorem from [4,8].

Theorem 3 *([4,8]). Let $\ell \geqslant 2$ be fixed. Suppose that $pn^{1/\ell} \to \infty$ as $n \to \infty$. Then whp $G(n, p)$ contains a Hamilton ℓ-cycle.*

Note that the ℓ-th power of a Hamilton cycle in a graph on $k(\ell + 1)$ vertices admits a $K_{\ell+1}$-factor, i.e. contains a disjoint union of cliques of size $\ell + 1$ covering all vertices of this graph.

Theorem 3 implies the following.

Lemma 1. *Let $p = \text{const} \in (0, 1)$. Fix a positive integer ℓ. Whp for every set $W \subset [n]$ of size at most $2 \log_b n$ the graph $G(n, p)|_{[n]\setminus W}$ obtained by the deletion of vertices from W contains the ℓ-th power of a Hamilton cycle.*

The proof is standard and is based on a sequential exposure of edges of $G(n, p)$ sufficiently many times and independently with probability slightly bigger than the threshold probability of an appearance of the ℓ-th power of a Hamilton cycle. For the sake of completeness, we give this argument in Appendix.

3. Lower bound

Let G be a graph, and H be $K_{1,r}$-saturated in G with the minimum possible number of edges. Obviously, the maximum degree of H does not exceed $r - 1$. Let us divide the set of vertices of H into two subsets: $V(H) = V_1 \sqcup V_2$, where V_1 comprises all the vertices of H with degrees at most $r - 2$, and all the vertices in V_2 have degree exactly $r - 1$. Let the induced subgraph $H_1 := H[V_1]$ have k vertices and m edges. Then the graph H has at least $\frac{r-1}{2}(n - k) + m$ edges, i.e.

$$\text{sat}(G, K_{1,r}) \geqslant \left\lceil \frac{r-1}{2}(n-k) \right\rceil + m. \tag{2}$$

It is clear that H_1 is also the induced subgraph of G itself. If this were not the case, then we could draw an edge inside H_1 that does not create $K_{1,r}$, a contradiction. So, $|E(G[V_1])| = m$. We will show that if $G = G(n, p)$, then whp it is optimal to take V_1 such that $m \leqslant r' + 1$. For that, let us rewrite (2) as follows:

$$\text{sat}(G, K_{1,r}) \geqslant \frac{r-1}{2}(n-k) + m = \frac{r-1}{2}(n - x_0) + \left(m - \frac{r-1}{2}(k - x_0) \right). \tag{3}$$

The term $\frac{r-1}{2}(n - x_0)$ refers to the number of edges in H, when $G = G(n, p)$, H_1 is an independent set of a maximum size, and there are no edges between V_1 and V_2 in H. We show that the second summand in the right hand side of (3) is whp at least μ, where μ is defined in the statement of Theorem 1. In other words, whp it is impossible to enlarge a maximum set that induces exactly μ edges and get an induced subgraph where the increase in the number of edges is less than the increase in the number of vertices times $\frac{r-1}{2}$.

Lemma 2. *Let $p \in (0, 1)$ be a constant. Then whp for any $k \geqslant x_0 + 1$ there are no induced subgraphs on k vertices in $G(n, p)$ with fewer than $\frac{r-1}{2}(k - x_0) + \mu$ edges.*

Assume that Lemma 2 is true and consider separately three cases.

1. If $k \leqslant x_0 - 1$, then (3) implies $\mathrm{sat}(G, K_{1,r}) \geqslant \lceil \frac{r-1}{2}(n - x_0) + \frac{r-1}{2} \rceil \geqslant \lceil \frac{r-1}{2}(n - x_0) \rceil + \mu$ as needed.
2. If $k = x_0$, then, due to the definition of μ, whp $m \geqslant \mu$, see the discussion in Section 2.1 after the statement of Theorem 2. It readily implies the desired inequality $\mathrm{sat}(G, K_{1,r}) \geqslant \frac{r-1}{2}(n - x_0) + \mu$.
3. Finally, if $k \geqslant x_0 + 1$, then, due to Lemma 2, whp $m - \frac{r-1}{2}(k - x_0) \geqslant \mu$ and therefore $\mathrm{sat}(G, K_{1,r}) \geqslant \frac{r-1}{2}(n - x_0) + \mu$ as well.

This finishes the proof of the lower bound in Theorem 1. In Section 3.1 we give the proof of Lemma 2.

3.1. Proof of Lemma 2

Denote $x_1 := x_0 + 1$. Let X_k be a random variable equal to the number of induced subgraphs on k vertices with the number of edges fewer than $\frac{r-1}{2}(k - x_0) + \mu$, $X = \sum_{k \geqslant x_1} X_k$. Notice that Lemma 2 states that $X = 0$ whp. We will prove it using Markov's inequality. Thus, it is sufficient to show that $\mathbb{E}X \to 0$ as $n \to \infty$.

Due to the linearity of the expectation

$$\mathbb{E}X = \sum_{k \geqslant x_1} \mathbb{E}X_k = \sum_{k \geqslant x_1} \binom{n}{k} \sum_{m < \mu + \frac{r-1}{2}(k - x_0)} \binom{\binom{k}{2}}{m} p^m (1-p)^{\binom{k}{2} - m}. \tag{4}$$

Let us show that the function $f(m) := \binom{\binom{k}{2}}{m} p^m (1-p)^{\binom{k}{2} - m}$ increases in our range. To do this, consider the ratio $\frac{f(m+1)}{f(m)}$ and show that it is greater than 1:

$$\frac{f(m+1)}{f(m)} = \frac{p}{1-p} \times \frac{\binom{k}{2} - m}{m+1} \geqslant \frac{p}{1-p} \times \frac{\binom{k}{2} - \frac{r-1}{2}(k - x_0) - \mu}{\frac{r-1}{2}(k - x_0) + \mu + 1} > 1$$

for n large enough.

Let us rewrite (4) as follows:

$$\mathbb{E}X = \binom{n}{x_1} \sum_{m < \mu + \frac{r-1}{2}} \binom{\binom{x_1}{2}}{m} p^m (1-p)^{\binom{x_1}{2} - m} + \sum_{k \geqslant x_1 + 1} \binom{n}{k} \sum_{m < \mu + \frac{r-1}{2}(k - x_0)} \binom{\binom{k}{2}}{m} p^m (1-p)^{\binom{k}{2} - m}. \tag{5}$$

Consider separately the first and second terms in (5) and show that they tend to zero. We start with the first summand. Let us recall that

$$x_1 = \lfloor 2\log_b n - 2\log_b \log_b n + 2\log_b(e/2) + 1 + \varepsilon \rfloor + 1.$$

Therefore,

$$\frac{x_1 - 1}{2} \geqslant \log_b n - \log_b \log_b n + \log_b(e/2) + \frac{\varepsilon}{2}.$$

We get

$$\begin{aligned}
\binom{n}{x_1}(1-p)^{\binom{x_1}{2}} &\leqslant \left(\frac{en}{x_1}\right)^{x_1} (1-p)^{\binom{x_1}{2}} = \exp\left[x_1 \left(\ln n + 1 - \ln x_1 - \frac{x_1 - 1}{2} \ln b\right)\right] \\
&\leqslant \exp\left[x_1 \left(1 - \ln 2 - \log_b(e/2) \ln b - \frac{\varepsilon}{2} + o(1)\right)\right] \\
&= \exp\left[-x_1 \left(\frac{\varepsilon}{2} + o(1)\right)\right] = \exp\left[-\varepsilon \log_b n(1 + o(1))\right].
\end{aligned}$$

Moreover,

$$\sum_{m < \mu + \frac{r-1}{2}} \binom{\binom{x_1}{2}}{m} \left(\frac{p}{1-p}\right)^m \leqslant \left[\mu + \frac{r-1}{2}\right] \max_m \left(\frac{\binom{x_1}{2} e}{m}\right)^m \left(\frac{p}{1-p}\right)^m = o(x_1^{2\mu + r}).$$

From the above, the first summand in (5) is bounded from above as follows:

$$\binom{n}{x_1}(1-p)^{\binom{x_1}{2}} \sum_{m < \mu + \frac{r-1}{2}} \binom{\binom{x_1}{2}}{m} \left(\frac{p}{1-p}\right)^m \leqslant \exp[-\varepsilon \log_b n(1 + o(1)) + (r + 2\mu)\ln x_1] = o(1) \tag{6}$$

as needed.

Now let us switch to the second summand in (5). Let $k \geqslant x_1 + 1$. Then

$$\binom{n}{k} \sum_{m<\mu+\frac{r-1}{2}(k-x_0)} \binom{\binom{k}{2}}{m} p^m (1-p)^{\binom{k}{2}-m}$$

$$\leqslant \left(\frac{ne}{k}\right)^k \left[\mu + \frac{r-1}{2}(k-x_0)\right] \left(\frac{\binom{k}{2}e}{\mu + \left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor}\right)^{\mu+\left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor} \left(\frac{p}{1-p}\right)^{\mu+\left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor} (1-p)^{\binom{k}{2}}$$

$$= \left[\left(\frac{ne}{x_1}\right)^{x_1} (1-p)^{\binom{x_1}{2}}\right] (ne)^{k-x_1} \frac{x_1^{x_1}}{k^k} \left[\mu + \frac{r-1}{2}(k-x_0)\right] \times$$

$$\times \left(\frac{\binom{k}{2}e}{\mu + \left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor}\right)^{\mu+\left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor} \left(\frac{p}{1-p}\right)^{\mu+\left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor} (1-p)^{\binom{k}{2}-\binom{x_1}{2}}.$$

Note that the factor $\left(\frac{ne}{x_1}\right)^{x_1}(1-p)^{\binom{x_1}{2}}$ does not exceed 1, because this is an upper bound for $\mathsf{E}\xi_0(x_1)$ (the expected number of independent sets of size x_1), and x_1 is chosen exactly in a way such that this bound approaches 0 (see [3, Remark 7.3]). Then the last expression can be estimated from above for large enough n as

$$\frac{(ne)^{k-x_1} x_1^{x_1}}{k^k} \left[\mu + \frac{r-1}{2}(k-x_0)\right] \left(\frac{\binom{k}{2}e}{\mu + \left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor}\right)^{\mu+\left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor} \times$$

$$\times \left(\frac{p}{1-p}\right)^{\mu+\left\lfloor \frac{r-1}{2}(k-x_0)\right\rfloor} (1-p)^{\binom{k}{2}-\binom{x_1}{2}} =: f(k). \quad (7)$$

Then

$$\ln f(k) \sim (k-x_1)\ln n - (k-x_1)\ln k + \left[\mu + \frac{r-1}{2}(k-x_0)\right] \ln \binom{k}{2} -$$

$$- \left[\mu + \frac{r-1}{2}(k-x_0)\right] \ln\left(\mu + \frac{r-1}{2}(k-x_0)\right) - \left(\binom{k}{2} - \binom{x_1}{2}\right) \ln \frac{1}{1-p}$$

$$\leqslant (k-x_1)\ln n + \left[\mu + \frac{r-1}{2}(k-x_0)\right] \ln \binom{k}{2} - \left(\binom{k}{2} - \binom{x_1}{2}\right) \ln \frac{1}{1-p}$$

$$\sim (k-x_1)\ln n + (r-1)(k-x_1)\ln k - \frac{(k-x_1)(k+x_1)}{2} \ln \frac{1}{1-p}.$$

Denote $c := k - x_1$, then

$$\ln f(k) \leqslant c\left(\ln n + (r-1)\ln k - \frac{2x_1 + c}{2} \ln \frac{1}{1-p}\right)(1+o(1)).$$

Differentiating the function inside the brackets with respect to k, we get $\left(\frac{r-1}{k} - \frac{1}{2}\ln\frac{1}{1-p}\right)$. Since $k \to \infty$, then $\frac{r-1}{k} - \frac{1}{2}\ln\frac{1}{1-p} < 0$ for sufficiently large n, and hence the function itself is decreasing when $k \geqslant x_1$. And since it decreases, then its maximum value is reached at the smallest possible $k = x_1$. Note that $\ln x_1 = O(\ln\ln n)$, and it does not affect asymptotics. Summing up, we get that

$$\ln f(k) \leqslant (k-x_1)\left(\ln n - x_1 \ln \frac{1}{1-p}\right)(1+o(1)) \sim -\ln n(k-x_1)(1+o(1)).$$

Therefore, the second summand in (5) is bounded from above as follows:

$$\sum_{k\geqslant x_1+1} \binom{n}{k} \sum_{m<\mu+\frac{r-1}{2}(k-x_0)} \binom{\binom{k}{2}}{m} p^m (1-p)^{\binom{k}{2}-m} \leqslant \sum_{c\geqslant 1} e^{-c\ln n(1+o(1))} \to 0. \quad (8)$$

Due to (6) and (8), $\mathbb{E}X \to 0$ as $n \to \infty$ as needed.

4. Upper bound

Assume first that $\mu \leqslant r'$. Then whp there exists a set with $\mu+1$ edges and x_0 vertices (see the discussion after the statement of Theorem 2), and we let V_1 to be such a set. If $\mu = r'+1$, then V_1 is an independent set on x_0-1 vertices (it exists whp due to (1)).

Let us show that whp there exists a subgraph in G such that

1. V_1 induces the same set of edges in this subgraph as in G,
2. each vertex from $V_2 := [n] \setminus V_1$ has degree exactly $r-1$, and
3. there is at most 1 edge between V_1 and V_2.

Let us show that whp there exists a subgraph in G such that each vertex from $V_2 := [n] \setminus V_1$ has degree exactly $r-1$, and there is at most 1 edge between V_1 and V_2. Clearly, such a subgraph is saturated in G.

According to Lemma 1, whp $G|_{V_2}$ contains the $(2r-2)$-th power of a Hamilton cycle. Let us preserve in this cycle a disjoint union of cliques K_r of size r and a clique K^* with at least r and at most $2r-1$ vertices such that this union covers all vertices of V_2.

It remains to show that we may turn K^* into an $(r-1)$-regular graph by deleting some edges of K^* and drawing at most 1 edge from K^* to V_1.

Let us first remove some edges from K^* so that only a simple cycle containing all its vertices remains.

If $r-1 = 2s$ is even, then join each vertex in this cycle with s nearest neighbors. We get $(r-1)$-regular graph. If $r-1 = 2s+1$ is odd and $|V(K^*)|$ is even then join each vertex in the cycle with s nearest neighbors and also with the opposite vertex. And finally, if $r-1 = 2s+1$ is odd and $|V(K^*)|$ is odd as well, then we arbitrarily choose a single vertex v_1 in the cycle. Note that V_1 is either an independent set with maximum possible size or a set with $\mu+1$ edges with maximum possible size, and therefore v_1 has a neighbor in V_1. Draw a single edge of G from v_1 to V_1. Inside the cycle, we join each vertex with its s nearest neighbors. Then, eventually, consider the cyclic order on $V(K^*) \setminus \{v_1\}$ which is exactly the order induced by the initial cycle from which we exclude the vertex v_1. Draw from every vertex in $V(K^*) \setminus \{v_1\}$ the edge to the opposite vertex in this order.

The desired saturated subgraph of G is constructed, it has exactly $\left\lceil \frac{(r-1)(n-x_0)}{2} \right\rceil + \mu + 1$ edges if $\mu \leqslant r'$ and $\left\lceil \frac{(r-1)(n-x_0)}{2} \right\rceil + \mu$ edges if $\mu = r'+1$ due to the choice of r' implying the upper bound in Theorem 1.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Appendix A. Proof of Lemma 1

Let Q be the property that a graph contains the ℓ-th power of a Hamilton cycle. According to Theorem 3, the threshold probability of this property is $\hat{p} = n^{-\frac{1}{\ell}}$. Let $p_0 = n^{-\frac{1}{\ell}} \ln n$. Then $\mathsf{P}(G(n, p_0) \notin Q) \leqslant \frac{1}{2}$ for large enough n.

Let x be the maximum integer such that $1-p \leqslant (1-p_0)^x$. Let us take logarithm of the both sides of this inequality and find the asymptotic behavior of x:

$$(x + O(1)) \ln(1-p_0) = \ln(1-p). \tag{9}$$

Since $p_0 \to 0$ for $n \to \infty$, then $\ln(1-p_0) \sim -p_0$ and (9) can be written as

$$x \sim \frac{1}{p_0} \ln\left(\frac{1}{1-p}\right) = \frac{n^{\frac{1}{\ell}}}{\ln n} \ln\left(\frac{1}{1-p}\right).$$

Consider a union of x independent copies $G_1, \ldots, G_x$ of $G(n, p_0)$. Clearly, there exists a coupling such that this union G is a subgraph of $G(n, p)$. Therefore, if $G(n, p)$ does not contain the ℓ-th power of a Hamilton cycle, then G does not contain it as well, and the same applies to each of $G_1, \ldots, G_x$. Then

$$\mathsf{P}(G(n, p) \notin Q) \leqslant (\mathsf{P}(G(n, p_0) \notin Q))^x \leqslant \left(\frac{1}{2}\right)^{n^{1/\ell}(\ln b + o(1))/\ln n}.$$

By the union bound, the probability that there exists a set W of size at most $2\log_b n$ such that $G(n, p)|_{[n]\setminus W}$ does not contain the ℓ-th power of a Hamilton cycle is at most

$$2\log_b n\ n^{2\log_b n}\,\mathsf{P}\,(G(n(1-o(1)),p)\notin Q) \leqslant 2\log_b n\ n^{2\log_b n}\,e^{-\ln 2\ln bn^{1/\ell}(\ln n)^{-1}(1+o(1))} \to 0,\quad n\to\infty$$

as needed.

References

[1] P. Erdős, A. Hajnal, J. Moon, A problem in graph theory, Am. Math. Mon. 71 (1964) 1107–1110.
[2] G.R. Grimmett, C.J.H. McDiarmid, On colouring random graphs, Math. Proc. Camb. Philos. Soc. 77 (1975) 313–324.
[3] S. Janson, T. Łuczak, A. Ruciński, Random Graphs, Wiley, New York, 2000.
[4] J. Kahn, B. Narayanan, J. Park, The threshold for the square of a Hamilton cycle, Proc. Am. Math. Soc. 149 (2021) 3201–3208.
[5] D. Kamaldinov, A. Skorkin, M. Zhukovskii, Maximum sparse induced subgraphs of the binomial random graph with given number of edges, Discrete Math. 344 (2021) 112205.
[6] L. Kászonyi, Zs. Tuza, Saturated graphs with minimal number of edges, J. Graph Theory 10 (1986) 203–210.
[7] D. Korándi, B. Sudakov, Saturation in random graphs, Random Struct. Algorithms 51 (1) (2016) 169–181.
[8] D. Kühn, D. Osthus, On Pósas conjecture for random graphs, SIAM J. Discrete Math. 26 (2012) 1440–1457.
[9] D.W. Matula, On the complete subgraphs of a random graph, in: Proceedings of the 2nd Chapel Hill Conference on Combinatorial Mathematics and Its Applications, Chapel Hill, N.C., 1970, pp. 356–369.
[10] D.W. Matula, The employee party problem, Not. Am. Math. Soc. 19 (2) (1972) A-382.
[11] D.W. Matula, The largest clique size in a random graph, Tech. Rep. 1987, Department of Computer Science, Southern Methodist University, Dallas, Texas, 1976.
[12] A. Mohammadian, B. Tayfeh–Rezaie, Star saturation number of random graphs, Discrete Math. 341 (4) (2018) 1166–1170.
[13] M. Zito, Small maximal matchings in random graphs, Theor. Comput. Sci. 297 (2003) 487–507.
[14] A.A. Zykov, On some properties of linear complexes, Mat. Sb. 24 (66) (1949) 163–188 (in Russian).