Packing Hamilton Cycles Online

Joseph Briggs[*], Alan Frieze[†], Michael Krivelevich[‡],
Po-Shen Loh[§], Benny Sudakov[¶]

Abstract

It is known that w.h.p. the hitting time $\tau_{2\sigma}$ for the random graph process to have minimum degree 2σ coincides with the hitting time for σ edge disjoint Hamilton cycles, [4], [13], [9]. In this paper we prove an online version of this property. We show that, for a fixed integer $\sigma \geq 2$, if random edges of K_n are presented one by one then w.h.p. it is possible to color the edges online with σ colors so that at time $\tau_{2\sigma}$, each color class is Hamiltonian.

AMS Classification Codes: 05C80, 05C15.

1 Introduction

The celebrated *random graph process*, introduced by Erdős and Rényi [5] in the 1960's, begins with an empty graph on n vertices, and at every step $t = 1, \ldots, \binom{n}{2}$ adds to the current graph a single new edge chosen uniformly at random out of all missing edges. Taking a snapshot of the random graph process after m steps produces the distribution $G_{n,m}$. An equivalent "static" way of defining $G_{n,m}$ would be: choose m edges uniformly at random out of all $\binom{n}{2}$ possible ones. One advantage in studying the random graph process, rather than the static model, is that it allows for a higher resolution analysis of the appearance of monotone graph properties (a graph property is monotone if it is closed under edge addition).

A *Hamilton cycle* of a graph is a simple cycle that passes through every vertex of the graph, and a graph containing a Hamilton cycle is called *Hamiltonian*. Hamiltonicity is one of the most fundamental notions in graph theory, and has been intensively studied in various contexts, including random graphs. The earlier results on Hamiltonicity of random graphs were obtained by Pósa [15], and Korshunov [10]. Improving on these results, Bollobás [3], and Komlós and Szemerédi [11] proved that if $m' = \frac{1}{2}n \log n + \frac{1}{2}n \log\log n + \omega n$, then $G_{n,m'}$ is Hamiltonian w.h.p. Here ω is any function of n tending to infinity together with n. One obvious necessary condition for the graph to be Hamiltonian is for the minimum degree to be at least 2, and the above result indicates that the events of being Hamiltonian and of having all degrees at least two are indeed bundled together

[*]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA. `jbriggs@andrew.cmu.edu`.

[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA. `alan@random.math.cmu.edu`. Research supported in part by NSF Grants DMS1362785, CCF1522984 and a grant(333329) from the Simons Foundation.

[‡]School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 6997801, Israel. `krivelev@post.tau.ac.il`. Research supported in part by a USA-Israel BSF Grant and by a grant from Israel Science Foundation.

[§]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213. `ploh@cmu.edu`. Research supported in part by NSF grant DMS-1455125.

[¶]Department of Mathematics, ETH, 8092 Zurich, Switzerland. `benjamin.sudakov@math.ethz.ch`. Research supported in part by SNSF grant 200021-149111.

closely. Bollobás [3], and independently, Ajtai, Komlós, and Szemerédi [1], further strengthened this by proving that w.h.p. the random graph process becomes Hamiltonian when the last vertex of degree one disappears. A more general property $\mathcal{H}_\sigma$ of having σ edge disjoint Hamilton cycles was studied by Bollobás and Frieze [4]. They showed that if $\sigma = O(1)$ then w.h.p. the random graph process satisfies $\mathcal{H}_\sigma$ when the minimum degree becomes 2σ. It took quite a while, but this result was extended to the more difficult case of growing σ in the $G_{n,m}$ context by Knox, Kühn and Osthus [9] and Krivelevich and Samotij [13].

Recently, quite a lot of attention and research effort has been devoted to controlled random graph processes. In processes of this type, an input graph or a graph process is usually generated fully randomly, but then an algorithm has access to this random input and can manipulate it in some well defined way (say, by dropping some of the input edges, or by coloring them), aiming to achieve some preset goal. There is usually the so-called *online* version where the algorithm must decide on its course of action based only on the history of the process so far and without assuming any familiarity with future random edges. For example, in the so-called *Achlioptas process* the random edges arrive in batches of size k. An online algorithm chooses one of them and puts it into the graph. By doing this one can attempt to accelerate or to delay the appearance of some property. Hamiltonicity in Achlioptas processes was studied in [12]. Another online result on Hamiltonicity was proved in [14]. There, it was shown that one can orient the edges of the random graph process so that w.h.p. the resulting graph has a directed Hamilton cycle exactly at the time when the underlying graph has minimum degree two.

Here we consider a Ramsey-type version of controlled random processes. In this version, the incoming random edge, when it is exposed, is irrevocably colored by an algorithm in one of r colors, for a fixed $r \geq 2$. The goal of the algorithm is to achieve or to maintain a certain monotone graph property in all of the colors. For example, in [2] the authors considered the problem of creating a linear size (so-called *giant*) component in every color.

The above mentioned result of Bollobás and Frieze [4] gives rise to the following natural question. Can one typically construct σ edge disjoint Hamilton cycles in an online fashion by the time the minimum degree becomes 2σ? We answer this question affirmatively in the case $\sigma = O(1)$.

Theorem 1.1. For a fixed integer $\sigma \geq 2$, let $\tau_{2\sigma}$ denote the hitting time for the random graph process $G_i, i = 1, 2, \ldots$ to have minimum degree 2σ. Then w.h.p. we can color the edges of $G_i, i = 1, 2, ..$ online with σ colors so that $G_{\tau_{2\sigma}}$ contains σ Hamilton cycles $C_1, C_2, \ldots, C_\sigma$, where the edges of cycle C_j all have color j.

2 Description of the coloring procedure

We describe our coloring procedure in terms of $q = 2\sigma$ colors we aim to color the edges so that each vertex has degree at least one in each color. Think of colors 1 and $1 + \sigma$ being light red and dark red, say, and then that each vertex is incident with at least two red edges. This may appear cumbersome, but it does make some of the description of the analysis a little easier.

In the broadest terms, we construct two sets of edges E^+ and E^*. Let Γ_c^* be the subgraph of $G_{\tau_{2\sigma}}$ induced by the edges of color c in E^*. We ensure that w.h.p. this has minimum degree at least one for all c. We then show that w.h.p. after merging colors c and $c + \sigma$ for $c \in [\sigma]$ the subgraph $\Gamma_c^{**} = \Gamma_c^* \cup \Gamma_{c+\sigma}^*$ has sufficient expansion properties so that standard arguments using Pósa rotations can be applied. For every color c, the edges of E_c^* are used to help create a good expander, and produce a backbone for rotations. And the edges in E_c^+ are used to close cycles in this argument.

Notation. "At time t" is taken to mean "when t edges have been revealed".

Notation. Let $N^{(t)}(v)$ denote the set of neighbors of v in G_t and let $d_v^{(t)} = |N^{(t)}(v)|$.

For color $c \in [q]$, write $d_c = d_{c,t}, N_c = N_{c,t}$ for the degrees and neighborhoods of vertices and sets in Γ_c.

Definition 2.1. Let *Full* denote the set of vertices with degree at least $\frac{\epsilon \log n}{1000q}$ in every color at time

$$t_\epsilon := \epsilon n \log n\,,$$

where ϵ is some sufficiently small constant depending only on the constant q. The actual value of ϵ needed will depend on certain estimates below being valid, in particular equation (14). A vertex is *Full* if is lies in *Full*. Similarly, let $Full' \subseteq Full$ denote the set of vertices with degree at least $\frac{\epsilon \log n}{1000q}$ in every color at time $\frac{1}{2}\epsilon n \log n$.

This definition only makes sense if t_ϵ is an integer. Here and below we use the following convention. If we give an expression for an integer quantity that is not clearly an integer, then rounding the expression up or down will give a value that can be used to satisfy all requirements.

2.1 Coloring Algorithm COL

We now describe our algorithm for coloring edges as we see them. At any time t, vertex v has a list $C_v^{(t)} := \{c \in [q] : d_c^{(t)}(v) = 0\}$ of colors currently not present among edges incident to v; "the colors that v needs". A vertex is *needy* at time t if $C_v^{(t)} \neq \emptyset$. If the next edge to color contains a needy vertex then we try to reduce the need of this vertex. Otherwise, we make choices to guarantee expansion in E^*, needed to generate many endpoints in the rotation phase, and to provide edges for E^+, which are used to close cycles, if needed.

FOR $t = 1, 2, \ldots, \tau_q$ **DO**
BEGIN

Step 1 Let $e_t = uv$.

Step 2 If $C_v^{(t)} \cup C_u^{(t)} = \emptyset$, $t > t_\epsilon$, and precisely one of $\{u, v\}$ (WLOG u) is *Full*, then give uv the color c that minimises $d_c(v)$ (breaking ties arbitrarily). Add uv to E_c^*.

Step 3 If $C_v^{(t)} \cup C_u^{(t)} = \emptyset$, $t > t_\epsilon$ and *both* $u, v \in Full$, give uv a color c uniformly at random from $[q]$. Then add this edge to E_c^+ or E_c^*, each with probability 1/2.

Step 4 If $C_v^{(t)} \cup C_u^{(t)} = \emptyset$ but $t \leq t_\epsilon$ or *both* $u, v \notin Full$, then color uv with color c chosen uniformly at random from $[q]$. Add uv to E_c^*.

Step 5 Otherwise, color uv with color c chosen uniformly at random from $C_u^{(t)} \cup C_v^{(t)}$. Add uv to E_c^*.

END
Let

$$E^* = \bigcup_{c \in [q]} E_c^* \text{ and } E^+ = \bigcup_{c \in [q]} E_c^+.$$

3 Structural properties

Let
$$p = \frac{\log n + (q-1)\log\log n - \omega}{n} \text{ and } m = \binom{n}{2}p$$
where
$$\omega = \omega(n) \to \infty, \omega = o(\log\log n).$$
We will use the following well-known properties relating $G_{n,p}$ and $G_{n,m}$, see for example [7], Chapter 1. Let $\mathcal{P}$ be a graph property. It is monotone increasing if adding an edge preserves it, and is monotone decreasing if deleting an edge preserves it. We have:
$$\mathbb{P}(G_{n,m} \in \mathcal{P}) \leq 10m^{1/2}\mathbb{P}(G_{n,p} \in \mathcal{P}). \tag{1}$$
$$\mathbb{P}(G_{n,m} \in \mathcal{P}) \leq 3\mathbb{P}(G_{n,p} \in \mathcal{P}), \text{ if } \mathcal{P} \text{ is monotone.} \tag{2}$$

A vertex $v \in [n]$ is *small* if its degree $d(v)$ in $G_{n,m}$ satisfies $d(v) < \frac{\log n}{100q}$. It is *large* otherwise. The set of small vertices is denoted by $SMALL$ and the set of large vertices is denoted by $LARGE$.

Definition 3.1. A subgraph H of $G_{n,m}$ with a subset $S(H) \subset V(H)$ is called a *small structure* if
$$|E(H)| + |S(H)| - |V(H)| \geq 1.$$
We say that $G_{n,m}$ *contains* H if there is an injective homomorphism $\phi : H \hookrightarrow G_{n,m}$ such that $\phi(S(H)) \subseteq SMALL$. The important examples of H include:

- A single edge between 2 *small* vertices.
- A path of length at most five between two *small* vertices.
- A copy of C_3 or C_4 with at least one *small* vertex.
- Two distinct triangles sharing at least one vertex.

Lemma 3.2. For any fixed small structure H of constant size,
$$\mathbb{P}(G_{n,m} \text{ contains } H) = o(n^{-1/5}).$$

Proof. We will prove that
$$\mathbb{P}(G_{n,p} \text{ contains } H) = o(n^{-3/4}).$$

This along with (1) implies the lemma.
Let $h = |V(H)|, f = |E(H)|, s = |S(H)|$ so that $f + s \geq h + 1$. Then:
$$\mathbb{P}\left(G_{n,p} \text{ contains } H\right) \leq \binom{n}{h}h!p^f\left(\sum_{i=0}^{\frac{\log n}{100q}}\binom{n-h}{i}p^i(1-p)^{n-h-i}\right)^s$$
$$\lesssim n^h\left(\frac{\log n}{n}\right)^f\left(\sum_{i=0}^{\frac{\log n}{100q}}\left(\frac{(e+o(1))\log n}{i}\right)^i e^{-\log n-(q-1)\log\log n+\omega+o(1)}\right)^s$$
$$\leq n^h\left(\frac{\log n}{n}\right)^f\left(\frac{(300q)^{\frac{\log n}{100q}}}{n(\log n)^{q-1-o(1)}}\right)^s$$

$$= o(n^{h-f-s+1/4}) = o(n^{-3/4}).$$

(We used the notation $A \lesssim B$ in place of $A \le (1+o(1))B$.) In the calculation above, in the first line we placed the vertices of H and decided about the identity of s vertices falling into $SMALL$, then required that all f edges of H are present in $G_{n,p}$, and finally required that for each of the s vertices in $SMALL$, their degree outside the copy of H is at most $\frac{\log n}{100q}$.

□

Lemma 3.3. W.h.p., for every $k \in \left[q-1, \frac{\log n}{100q}\right]$, there are less than $\nu_k = \frac{e^{2\omega}(\log n)^{k-q+1}}{(k-1)!}$ vertices of degree k in $G_{n,m}$.

Remark 3.4. ν_k is increasing in k for this range, and for the largest $k = \frac{\log n}{100q}$ we have $\nu_k \lesssim n^{\frac{\log(100eq)}{100q}}$.

Proof. Fix k and then we have

$$\begin{aligned}
&\mathbb{P}(G_{n,p} \text{ has at least } \nu_k \text{ vertices of degree at most } k)\\
&\le \binom{n}{\nu_k}\left(\sum_{\ell=0}^{k}\binom{n-\nu_k}{\ell}p^\ell(1-p)^{n-\nu_k-\ell}\right)^{\nu_k}\\
&= \binom{n}{\nu_k}\left((1+o(1))\binom{n-\nu_k}{k}p^k(1-p)^{n-\nu_k-k}\right)^{\nu_k}\\
&\le \left(\frac{ne}{\nu_k}\times\frac{n^k}{k!}\left(\frac{\log n+(q-1)\log\log n-\omega}{n}\right)^k e^{-\log n-(q-1)\log\log n+\omega+o(1)}\right)^{\nu_k}\\
&\le \left(\frac{e^{\omega+O(1)}}{(\log n)^{q-1}}\frac{(\log n+q\log\log n)^k}{k!\nu_k}\right)^{\nu_k}\\
&= \left(\frac{e^{-\omega+O(1)}}{k}\left(1+\frac{q\log\log n}{\log n}\right)^k\right)^{\nu_k}\\
&\le \left(e^{-\omega+O(1)}\frac{(\log n)^{kq/\log n}}{k}\right)^{\nu_k}.
\end{aligned}$$

The function $f(k) = \frac{(\log n)^{kq/\log n}}{k}$ is log-convex, and so f is maximised at the extreme values of k (specifically $f(q-1) = e^{O(1)} > f\left(\frac{\log n}{100q}\right) = o(1)$). Hence,

$$\mathbb{P}(\exists k : G_{n,p} \text{ has at least } \nu_k \text{ vertices of degree } k) \le \sum_{k=q-1}^{\frac{\log n}{100q}} e^{-\omega\nu_k/2} = o(1).$$

Applying (2) we see that

$$\mathbb{P}(\exists k : G_{n,m} \text{ has at least } \nu_k \text{ vertices of degree } k) = o(1),$$

which is stronger than required. □

Lemma 3.5. With probability $1-o(n^{-10})$, $G_{n,m}$ has no vertices of degree $\ge 20\log n$.

Proof. We will prove that w.h.p. $G_{n,p}$ has the stated property. We can then obtain the lemma by applying (2).

$$\begin{aligned}\mathbb{P}(\exists v : d(v) \geq 20\log n) &\leq n\binom{n-1}{20\log n}p^{20\log n}\\ &\leq n\left(\frac{en}{20\log n}\frac{2\log n}{n}\right)^{20\log n}\\ &\leq n\left(\frac{e}{10}\right)^{20\log n}\\ &= o(n^{-10}).\end{aligned}$$

□

4 Analysis of COL

Let $\Gamma = G_m$ and let $d(v)$ denote the degree of $v \in [n]$ in Γ. Let

$$\theta_v = \begin{cases} 0 & d(v) \geq q. \\ 1 & d(v) = q-1. \end{cases}$$

Lemma 4.1. Suppose we run COL as described above. Then w.h.p. $|C_v^{(m)}| = \theta_v$ for all $v \in [n]$.

In words, Lemma 4.1 guarantees that the algorithm COL typically performs so that at time m, each vertex of degree at least q has all colors present at its incident edges, while each vertex of degree $q-1$ has exactly one color missing. (It is well known that w.h.p. $\delta(G_m) = q-1$, see for example [7], Section **4.2**.)

Proof. Fix v and suppose v has k neighbours in LARGE, via edges $\{f_i = vu_i\}_{i=1}^k$. Then in general $d(v)-1 \leq k \leq d(v)$ as small vertices do not share a path of length two. Also, when v is small, $k = d(v)$. Write $t(e)$ for the time $t \in [1, m]$ at which an edge e appears in the random graph process, i.e. $t(e_i) = i$. Let $t_i = t(f_i)$ and assume that $t_i < t_{i+1}$ for $i > 0$. We omit $i = 1$ in the next consideration since v will always get a color it needs by time t_1. (It may get a color before t_1 through an edge vw where w is not in LARGE.) Every time an $f_i, i \geq 2$, appears while u_i needs no additional colors, v gets a color it needs. So for v to have $|C_v^{(m)}| > \theta_v$ at the end of the process, this must happen at most $q - 2 - \theta_v$ times, so there is certainly some set

$$S = \{i_1 < i_2 < \cdots < i_s\} \subseteq [2, k] \text{ of } s = k - q + 1 + \theta_v \text{ indices,}$$

whose corresponding edges $\{f_i, i \in S\}$ incident with v satisfy $C_{u_i}^{(t_i)} \neq \emptyset$. Let $\mathbf{T_S}$ denote $\{t_i : i \in S\}$ and $\mathbf{U}$ denote the sequence $u_1, u_2, \ldots, u_k$. In the following we will sum over S and condition on the choices for $\mathbf{T_S}$ and then estimate the probability that $C_{u_i}^{(t_i)} \neq \emptyset$ for $i \in S$. For a fixed S there will be at least $\binom{m-k}{|S|+1}$ equally likely choices for the set $\{t_i, i \in \{1\} \cup S\}$. (We do not condition on t_1. The factor $t_{i_1} - 1$ in (3) below will allow for the number of choices for t_1.) Let $\mathcal{L}$ denote the occurrence of the bound of $20\log n$ on the degree of v and its neighbors (see Lemma 3.5), and note that $\mathbb{P}(\mathcal{L}) = 1 - o(n^{-10})$.

Taking a union bound over S, and letting

$$A_i := \left\{C_{u_i}^{(t_i)} \neq \emptyset\right\},$$

we have

$$\begin{aligned}\mathbb{P}(|C_v^{(m)}| > \theta_v \mid \mathcal{L}, \mathbf{U}) &\leq \sum_{\substack{S\subset[2,k]\\|S|=s}} \sum_{t_i : i\in\{1\}\cup S} \frac{1}{\binom{m-k}{k-q+2+\theta_v}} \mathbb{P}\bigg(\bigwedge_{i\in S} A_i \bigg| \mathbf{T_S}, \mathbf{U}, \mathcal{L} \bigg) \\ &\approx \sum_{\substack{S\subset[2,k]\\|S|=s}} \sum_{t_i : i\in S} \frac{t_{i_1}-1}{\binom{m}{k-q+2+\theta_v}} \mathbb{P}\bigg(\bigwedge_{i\in S} A_i \bigg| \mathbf{T_S}, \mathbf{U}, \mathcal{L} \bigg), \qquad (3)\end{aligned}$$

since there are $t_{i_1} - 1$ choices for t_1 and $k^2 = o(m)$, implying $\binom{m-k}{k-q+2+\theta_v} \approx \binom{m}{k-q+2+\theta_v}$, given $\mathcal{L}$. Next let

$$\begin{aligned}&Y_i = \{\text{edges of } u_i \text{ that appeared before } t_i \text{ excluding edges contained in } N^{(m)}(v)\}, \\ &d_r = d(u_r) \text{ and } Z_r := |Y_r| \text{ for } r = 1, 2, \dots, s, \\ &\mathbf{D_S} = \{d_i : i \in S\}\,.\end{aligned}$$

Now fix $\mathbf{U}$ and S and $\mathbf{T_S}$ and $\mathbf{D_S}$.

Remark 4.2. Going back to Algorithm COL, we observe that Step 5 implies that if $C_v^{(t)} \neq \emptyset$ then uv is colored with a color in $C_v^{(t)}$ with probability at least $\frac{1}{q}$. This holds regardless of the previous history of the algorithm and also holds conditional on $\mathbf{T_S}, \mathbf{U}, \mathbf{D_S}$. Indeed, the random bits used in Step 5 are independent of the history and are distinct from those used to generate the random graphs. The latter explains why we can condition on the future by fixing $\mathbf{T_S}, \mathbf{U}, \mathbf{D_S}$. We condition on $\mathcal{L}$ in order to control s as $O(\log n)$.

Then,

$$\begin{aligned}&\mathbb{P}\left(A_{i_1} \wedge \cdots \wedge A_{i_s} \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}\right) \\ &= \sum_{z_s} \underbrace{\mathbb{P}(A_{i_s} \mid A_{i_1}, \dots, A_{i_{s-1}}, Z_s = z_s, \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L})}_{\leq \mathbb{P}(\mathrm{Bin}(z_s, q^{-1}) \leq q-1) \text{ by Remark 4.2}} \mathbb{P}(A_{i_1}, \dots, A_{i_{s-1}}, Z_s = z_s \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}) \\ &\leq \sum_{z_s} g(z_s) \sum_{z_{s-1}} \mathbb{P}(A_{i_{s-1}} \mid A_{i_1}, \dots, A_{i_{s-2}}, Z_{s-1} = z_{s-1}, Z_s = z_s, \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}) \\ &\qquad\qquad \times \mathbb{P}(A_{i_1}, \dots, A_{i_{s-2}}, Z_{s-1} = z_{s-1}, Z_s = z_s \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}) \\ &\leq \sum_{z_s, z_{s-1}} g(z_s) g(z_{s-1}) \mathbb{P}(A_{i_1}, \dots, A_{i_{s-2}}, Z_{s-1} = z_{s-1}, Z_s = z_s \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}) \\ &\leq \sum_{z_s, \dots, z_1} g(z_s) \cdots g(z_2) \mathbb{P}(Z_r = z_r, r = 1, \dots, s \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}) \text{ (by induction)} \qquad (4)\end{aligned}$$

Here $g(z) := \mathbb{P}(\mathrm{Bin}(z, q^{-1}) \leq q - 1)$ for any $z \geq 0$.

Claim 4.3.

$$\mathbb{P}(Z_r = z_r, r = 1, 2, \dots, s \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}) \leq \left(1 + \tilde{O}(n^{-1})\right) \prod_{r=1}^{s} \frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}},$$

where $\tilde{O}$ hides polylog factors.

Proof Fix $\frac{\log n}{100q} \le d_1, d_2, \dots, d_s = O(\log n)$ and $t_1, t_2, \dots, t_s$. Then, for every $1 \le r \le s$,

$$\mathbb{P}(Z_r = z_r \mid Z_{r-1} = z_{r-1}, \dots, Z_1 = z_1, \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}) \le (1 + o(n^{-10})) \frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m-d_2-\cdots-d_{r-1}-s}{d_r}} \tag{5}$$
$$\le \left(1 + \tilde{O}(n^{-1})\right) \frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}}.$$

Explanation for (5)**:** The the first binomial coefficient in the numerator in (5) bounds the number of choices for the z_r positions in the sequence where an edge contributing Y_r occurs. This holds regardless of $z_1, z_2, \dots, z_{r-1}$. The second binomial coefficient bounds the number of choices for the $d_r - z_r$ positions in the sequence where we choose an edge incident with u_r after time t_r. Conversely, the denominator in (5) is a lower bound on the number of choices for the d_r positions where we choose an edge incident with u_r, given $d_1, d_2, \dots, d_{r-1}$. We subtract the extra s to (over)count for edges from v to $u_{r+1}, \dots, u_s$. The factor $(1 + o(n^{-10}))$ accounts for the conditioning on $\mathcal{L}$.
Expanding $\mathbb{P}(Z_r = z_r, r = 1, \dots, s \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L})$ as a product of $s = O(\log n)$ of these terms completes the proof of Claim 4.3. □

Going back to (4) we see that given $d_1, d_2, \dots, d_s$,

$$\mathbb{P}\left(A_{i_1} \wedge \cdots \wedge A_{i_s} \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}\right)$$
$$\lesssim \prod_{r=1}^{s} \sum_{z_r=0}^{d_r} \left(\mathbb{P}(Bin(z_r, q^{-1}) \le q-1) \times \frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}} \right)$$
$$\le \prod_{r=1}^{s} \sum_{z_r=0}^{d_r} \left(C_1 \binom{z_r}{\min\{z_r, q-1\}} \frac{1}{q^{q-1}} \left(1 - \frac{1}{q}\right)^{z_r} \times \frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}} \right) \tag{6}$$
$$\le \prod_{r=1}^{s} \sum_{z_r=0}^{d_r} \left(C_1 \max\left\{1, z_r^{q-1}\right\} e^{-z_r/q} \times \frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}} \right). \tag{7}$$

Here, $C_1 = C_1(q)$ depends only on q. We will use constants $C_2, C_3, \dots$ in a similar fashion without further comment.
Justification for (6)**:** If $z_r \le q-1$ then $\mathbb{P}(Bin(z_r, q^{-1}) \le q-1) = 1$ and $C_1 = eq^q$ will suffice.
If $q \le z_r \le 10q$ we use

$$\mathbb{P}(Bin(z_r, q^{-1}) \le q-1) \le 1 \text{ and } \binom{z_r}{q-1} \frac{1}{q^{q-1}} \left(1 - \frac{1}{q}\right)^{z_r} \ge \frac{1}{q^{q-1}} \left(1 - \frac{1}{q}\right)^{10q}$$

and $C_1 = e^{20} q^q$ will suffice in this case.
If $z_r > 10q$ then putting $a_i := \mathbb{P}(Bin(z_r, q^{-1}) = i) = \binom{z_r}{i} \frac{1}{q^i} \left(1 - \frac{1}{q}\right)^{z_r - i}$ for $i \le q-1$ we see that

$$\frac{a_i}{a_{i-1}} = \frac{z_r - i + 1}{i} \cdot \frac{1}{q-1} \ge \frac{z_r - q}{q^2} > \frac{z_r}{2q^2} \ge \frac{5}{q}.$$

So here

$$\mathbb{P}(\mathrm{Bin}(z_r, q^{-1}) \le q-1) = \sum_{i=0}^{q-1} a_i \le a_{q-1}\left(1 + \frac{2q^2}{z_r} + \cdots + \left(\frac{2q^2}{z_r}\right)^{q-2}\right) \le$$

$$\left(1-\frac{1}{q}\right)^{1-q}\left(\binom{z_r}{q-1}\frac{1}{q^{q-1}}\left(1-\frac{1}{q}\right)^{z_r}\right)\frac{\left(\frac{q}{5}\right)^{q-1}-1}{\frac{q}{5}-1},$$

and thus $C_1 = (5q)^q$ suffices.
This completes the verification of (6).
Now, writing $(t)_z$ for the *falling factorial* $t!/(t-z)! = t(t-1)(t-2)\dots(t-z+1)$,

$$\begin{aligned}
\frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}} &= \binom{d_r}{z_r}\frac{(t_r)_{z_r}(m-t_r)_{d_r-z_r}}{(m)_{d_r}} \\
&= \binom{d_r}{z_r}\prod_{i=0}^{z_r-1}\frac{t_r-i}{m-(d_r-z_r)-i}\cdot\prod_{i=0}^{d_r-z_r-1}\frac{m-t_r-i}{m-i} \\
&\le \left(1+O\left(\frac{d_r^2}{m}\right)\right)\binom{d_r}{z_r}\left(\frac{t_r}{m}\right)^{z_r}\left(1-\frac{t_r}{m}\right)^{d_r-z_r}. \qquad (8)
\end{aligned}$$

Observe next that if $z_r \ge q^2$ then

$$(z_r)_{q-1} = z_r^{q-1}\prod_{i=0}^{q-1}\left(1-\frac{i}{z_r}\right) \ge z_r^{q-1}\left(1-\frac{q^2}{2z_r}\right) \ge \frac{z_r^{q-1}}{2}. \qquad (9)$$

It follows from (8) and (9) that

$$\begin{aligned}
&\sum_{z_r=q^2}^{d_r} C_1 z_r^{q-1}e^{-z_r/q} \times \frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}} \\
&\le 2C_1\sum_{z_r=q-1}^{d_r}(z_r)_{q-1}\binom{d_r}{z_r}\left(\frac{t_re^{-1/q}}{m}\right)^{z_r}\left(1-\frac{t_r}{m}\right)^{d_r-z_r} \\
&\le 2C_1(d_r)_{q-1}\left(\frac{t_r}{m}\right)^{q-1}\sum_{z_r=q-1}^{d_r}\binom{d_r-q+1}{z_r-q+1}\left(\frac{t_re^{-1/q}}{m}\right)^{z_r-q+1}\left(1-\frac{t_r}{m}\right)^{d_r-z_r} \\
&\le 2C_1\left(\frac{d_rt_r}{m}\right)^{q-1}\left(1-\frac{t_r}{m}\left(1-e^{-1/q}\right)\right)^{d_r-q+1} \\
&\le 2C_1\left(\frac{d_rt_r}{m}\right)^{q-1}\exp\left\{-\frac{(d_r-q+1)t_r}{m}\left(1-e^{-1/q}\right)\right\}.
\end{aligned}$$

Furthermore, not forgetting

$$\begin{aligned}
\sum_{z_r=0}^{q^2-1} C_1\max\left\{1, z_r^{q-1}\right\}e^{-z_r/q} \times \frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}} &\le C_2\sum_{z_r=0}^{q^2-1}\frac{\binom{t_r}{z_r}\binom{m-t_r}{d_r-z_r}}{\binom{m}{d_r}} \\
&\le C_3\sum_{z_r=0}^{q^2-1} t_r^{z_r}\cdot\frac{(m-t_r)^{d_r-z_r}}{(d_r-z_r)!}\cdot\frac{d_r!}{m^{d_r}} \\
&\le C_3\sum_{z_r=0}^{q^2-1}\left(\frac{d_rt_r}{m}\right)^{z_r} e^{-(d_r-z_r)t_r/m}
\end{aligned}$$

$$\leq C_4 \psi\left(\frac{d_r t_r}{m}\right),$$

where $\psi(x) = e^{-x}\sum_{z=0}^{q^2-1} x^z$. (Now $z_r \leq q^2$ and so the factor $e^{z_r t_r/m} \leq e^{q^2}$ can be absorbed into C_4.) Going back to (7) we have

$$\mathbb{P}\left(A_{i_1} \wedge \cdots \wedge A_{i_s} \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}\right) \leq \\ C_5^s \prod_{r=1}^{s}\left(\left(\frac{d_r t_r}{m}\right)^{q-1} \exp\left\{-\frac{d_r t_r}{m}\left(1-e^{-1/q}\right)\right\} + \psi\left(\frac{d_r t_r}{m}\right)\right). \quad (10)$$

It follows from (3) and (10) that,

$$p_v := \mathbb{P}(|C_v^{(m)}| > \theta_v \mid \mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}) \\ \leq \sum_{\substack{S\subset[2,k]\\|S|=s}} \sum_{t_i : i\in S} \frac{t_{i_1} C_5^s}{\binom{m}{s+1}} \prod_{r=1}^{s}\left(\left(\frac{d_r t_r}{m}\right)^{q-1} \exp\left\{-\frac{d_r t_r}{m}(1-e^{-1/q})\right\} + \psi\left(\frac{d_r t_r}{m}\right)\right).$$

Replacing a sum of products by a product of sums and dividing by $s!$ to account for repetitions, we get

$$p_v \leq \sum_{\substack{S\subset[2,k]\\|S|=s}} \frac{C_5^s}{\binom{m}{s+1}s!} \prod_{r=2}^{s}\left(\sum_{t=1}^{m}\left(\left(\frac{d_r t}{m}\right)^{q-1} \exp\left\{-\frac{d_r t}{m}(1-e^{-1/q})\right\} + \psi\left(\frac{d_r t}{m}\right)\right)\right) \\ \times\left(\sum_{t=1}^{m}\left(t\left(\frac{d_1 t}{m}\right)^{q-1} \exp\left\{-\frac{d_1 t}{m}(1-e^{-1/q})\right\} + t\psi\left(\frac{d_1 t}{m}\right)\right)\right).$$

We now replace the sums by integrals. This is valid seeing as the summands have a bounded number of extrema, and we replace C_5 by C_6 to absorb any small error factors.

$$p_v \leq \sum_{\substack{S\subset[2,k]\\|S|=s}} \frac{C_6^s}{\binom{m}{s+1}s!} \prod_{r=2}^{s}\left(\int_{t=0}^{\infty}\left[\left(\frac{d_r t}{m}\right)^{q-1} \exp\left\{-\frac{d_r t}{m}(1-e^{-1/q})\right\} + \psi\left(\frac{d_r t}{m}\right)\right]dt\right) \\ \times\left(\int_{t=0}^{\infty}\left(t\left(\frac{d_1 t}{m}\right)^{q-1} \exp\left\{-\frac{d_1 t}{m}(1-e^{-1/q})\right\} + \psi\left(\frac{d_1 t}{m}\right)\right)dt\right)$$

$$= \sum_{\substack{S\subset[2,k]\\|S|=s}} \frac{C_6^s}{\binom{m}{s+1}s!} \prod_{r=2}^{s}\left(\frac{m}{d_r}\int_{x=0}^{\infty}(x^{q-1}e^{-(1-e^{-1/q})x} + \psi(x))dx\right) \\ \times \frac{m^2}{d_1^2}\left(\int_{x=0}^{\infty}\left(x^q \exp\left\{-(1-e^{-1/q})x\right\} + x\psi(x)\right)dx\right)$$

$$\leq \sum_{\substack{S\subset[2,k]\\|S|=s}} \frac{C_6^s}{\binom{m}{s+1}s!} \cdot \left(\frac{C_7 m}{\min_r\{d_r\}}\right)^{s+1}$$

$$\leq \frac{C_8^k}{(\log n)^{k-q+2+\theta_v}}.$$

Applying Lemmas 3.3 and 3.5 and removing the conditioning on $\mathbf{T_S}, \mathbf{U}, \mathbf{D_S}, \mathcal{L}$ we see that with $k_0 = \frac{\log n}{100q}$,

$$\begin{aligned}
&\mathbb{P}(\exists v : |C_v^{(m)}| > \theta_v) \\
&\leq \mathbb{P}(\neg\mathcal{L}) + \sum_{k=q-1}^{k_0} \frac{e^{2\omega}(\log n)^{k-q+1}}{(k-1)!} \times \frac{C_8^k}{(\log n)^{k-q+2+\theta_v}} + n \sum_{k=k_0}^{20\log n} \frac{C_8^k}{(\log n)^{k-q+2}} \\
&\leq o(1) + \frac{e^{2\omega}}{\log n} \sum_{k\geq q-1} \frac{C_8^k}{(k-1)!} + n \sum_{k=k_0}^{20\log n} \frac{C_8^k}{(\log n)^{k/2}} \\
&\leq o(1) + \frac{C_9 e^{2\omega+C_9}}{\log n} \\
&= o(1).
\end{aligned}$$

□

We show next that at time m, w.h.p. sets of size up to $\Omega(n)$ have large neighbourhoods in every color.
We first prove that typically "large-degree vertices have large degree in every color": let $d_c^*(v)$ denote the number of edges incident with v that COL colors c, except for those edges that are colored in Step 3.

Theorem 4.4. *There exists $\epsilon = \epsilon(q) > 0$ such that w.h.p. on completion of COL every $v \in LARGE$ has $d_c^*(v) \geq \frac{\epsilon \log n}{1000q}$ for all $c \in [q]$.*

Suppose we define a vertex to be $small_c$ if it has $d_c(v) \leq \frac{\epsilon \log n}{1000q}$. Theorem 4.4 says w.h.p. the set of $small_c$ vertices $SMALL_c \subset SMALL$ so that by Lemma 3.2, G does not contain any $small_c$ structures of constant size. Here a $small_c$ structure is a small structure made up of $small_c$ vertices. The proof of Theorem 4.4 will follow from Lemmas 4.5, 4.6 and 4.7 below.

Lemma 4.5. There exists $\delta = \delta(q) > 0$ such that the following holds w.h.p.: Let $Full', Full$ be as in Definition 2.1. Then $|Full'| \geq n - \frac{203qn}{\epsilon \log n}$, and $|Full| \geq n - n^{1-\delta}$.

Proof. We first note that for $v \in [n]$, that if $t_\epsilon = \epsilon n \log n$ then

$$\mathbb{P}\left(d^{(t_\epsilon/2)}(v) < \lambda_0 := \frac{\epsilon \log n}{100}\right) \leq 3n^{-\epsilon/2} < n^{-\epsilon/3}. \tag{11}$$

Indeed, with $p_1 = \frac{t_\epsilon/2}{\binom{n}{2}}$ we see that, in the random graph model G_{n,p_1}:

$$\mathbb{P}\left(d(v) < \lambda_0\right) = \sum_{i=0}^{\lambda_0-1} \binom{n}{i} p_1^i (1-p_1)^{n-i} \leq 2\binom{n}{\lambda_0} p_1^{\lambda_0}(1-p_1)^{n-\lambda_0} \leq 2\left(\frac{nep_1}{\lambda_0}\right)^{\lambda_0} n^{-\epsilon+o(1)} \leq n^{-\epsilon/2}. \tag{12}$$

The first inequality follows from the fact that the ratio of succesive summands in the sum is at least $(n-\lambda_0)p_1/\lambda_0 > 50$.
Equation (11) now follows from (2) (with p replaced by p_1) and (12).

Thus the Markov inequality shows that with probability at least $1-n^{-\epsilon/3}$, at least $n-n^{1-\epsilon/6}$ of the vertices v have $d^{(t_\epsilon/2)}(v) \ge \frac{\epsilon \log n}{100}$. Now note that at most qn of the first $t_\epsilon/2$ edges were restricted in color by being incident to at least one needy vertex. This is because each time a needy vertex gets an edge incident to it, the total number of needed colors decreases by at least one. Therefore at most $\frac{200qn}{\epsilon \log n}$ of these vertices v have fewer than $\frac{\epsilon \log n}{200}$ of their $\ge \frac{\epsilon \log n}{100}$ initial edges colored completely at random, as in Step 4 of COL. Hence, there are at least $n - \frac{201qn}{\epsilon \log n}$ vertices v which have $d^{(t_\epsilon/2)}(v) \ge \frac{\epsilon \log n}{100}$ and $\frac{\epsilon \log n}{200}$ edges of fully random color. For such a v, and any color c,

$$\mathbb{P}\left(d_c^{(t_\epsilon/2)}(v) < \frac{\epsilon \log n}{1000q}\right) \le \mathbb{P}\left(Bin\left(\frac{\epsilon \log n}{200}, \frac{1}{q}\right) \le \frac{\epsilon \log n}{1000q}\right) \le \exp\left\{-\frac{1}{2}\cdot\frac{16}{25}\cdot\frac{\epsilon \log n}{200q}\right\} \le n^{-\epsilon/1000q}. \tag{13}$$

So $\mathbb{P}(v \notin Full') \le qn^{-\epsilon/1000q}$, and the Markov inequality shows that w.h.p.

$$|Full'| \ge n - \frac{201qn}{\epsilon \log n} - n^{1-\epsilon/2000q} \ge n - \frac{202qn}{\epsilon \log n}.$$

Now, for $v \notin Full'$, let $S(v) := Full' \setminus N^{(t_\epsilon/2)}(v)$. Since $d^{(t_\epsilon/2)}(v) \le d(v) \le 20 \log n$, we have $|S(v)| \ge n - \frac{203qn}{\epsilon \log n}$. Furthermore, every $w \in S(v)$ is no longer needy, and so among the next $t_\epsilon/2$ edges, at most q of the edges between v and $S(v)$ have their choice of color restricted by v, and the rest are colored randomly as in Step 4 of COL. Now $\mathbb{P}(|e(v,S(v))| < \frac{\epsilon \log n}{100}) = O(n^{-\epsilon/3})$ by a similar calculation to (12). Conditioning on $|e(v,S(v))| \ge \frac{\epsilon \log n}{100}$ we have $\mathbb{P}(v \notin Full) \le qn^{-\epsilon/1000q}$ by a similar calculation to (13) and so $|Full| \ge n - n^{1-\epsilon/2000q}$ with probability $\ge 1 - O(n^{-\epsilon/2000q})$ by the Markov inequality. □

We are working towards showing that vertices with low degree in some color must have have a low overall degree. The point is that all *Full* vertices no longer need additional colors later than $t_\epsilon = \epsilon n \log n$, so any new edge connecting *Full* to $V \setminus Full$ after time t_ϵ has its color determined by the vertex not in *Full*, as in Step 2 of COL. Indeed, suppose a vertex $v \notin Full$ has at least $\frac{\epsilon \log n}{400}$ edges to *Full* after time t_ϵ. Then v gets at least $\frac{\epsilon \log n}{400q} > \frac{\epsilon \log n}{1000q}$ edges of every color incident with it.

Lemma 4.6. W.h.p. there are no vertices $v \notin Full$ with at least $\frac{\epsilon \log n}{200}$ edges after time t_ϵ i.e., $d^{(m)}(v) - d^{(t_\epsilon)}(v) \ge \frac{\epsilon \log n}{200}$ but with at most $\frac{\epsilon \log n}{400}$ of these edges to *Full*.

Proof. Take any vertex $v \notin Full$ and consider the first $\frac{\epsilon \log n}{200}$ edges incident to v after time t_ϵ. We must estimate the probability that at least half of these edges are to vertices not in *Full*. We bound this by

$$\binom{\epsilon \log n/200}{\epsilon \log n/400}\left(\frac{n^{1-\delta}}{n - 20\log n}\right)^{\epsilon \log n/400} = o(1/n).$$

We subtract off a bound of $20 \log n$ on the number of edges from v to *Full* in E_{t_ϵ}. Note that we do not need to multiply by the number of choices for *Full*, as *Full* is defined by the first t_ϵ edges. There at most n choices for v and so the lemma follows. □

Lemma 4.7. There are no *large* vertices v with $d^{(m)}(v) - d^{(t_\epsilon)}(v) < \frac{\epsilon \log n}{200}$.

Proof. Any v satisfying these conditions must have $d^{(t_\epsilon)}(v) \ge \frac{\log n}{200q}$, if $\epsilon \le 1/q$ say. However, with $p_2 = \frac{t_\epsilon}{\binom{n}{2}}$ we have that in the random graph model G_{n,p_2},

$$\mathbb{P}\left(d(v) \ge \frac{\log n}{200q}\right) \le \binom{n}{\log n/200q} p_2^{\log n/200q} \le (400qe\epsilon)^{\log n/200q} = o(n^{-2}), \tag{14}$$

for ϵ sufficiently small.
The result follows by taking a union bound over choices of v and using (2) (again noting $\binom{n}{2}p_2 = t_\epsilon \to \infty$). $\square$

Proof of Theorem 4.4: It follows from Lemmas 4.6, 4.7 that every large vertex has at least $\frac{\epsilon \log n}{400}$ edges to $Full$ that occur after time t_ϵ. These edges will provide all needed edges of all colors. $\square$

It is known that w.h.p. $m \le \tau_q \le m' = m + 2\omega n$, see Erdős and Rényi [6]. We have shown that at time m all vertices, other than vertices of degree $q-1$, are incident with edges of all colors. Furthermore, vertices of degree $q-1$ are only missing one color. As we add the at most $2\omega n$ edges needed to reach τ_q we find (see Claim 4.8 below) that w.h.p. the edges we add incident to a vertex v of degree $q-1$ have their other end in LARGE. As such COL will now give vertex v its missing color.

Claim 4.8. W.h.p. an edge of $E_{m'} \setminus E_m$ that meets a vertex of degree $q-1$ in G_m has its other end in LARGE.

Proof. It follows from Lemma 3.3 that at time m and later there are w.h.p. at most $e^{2\omega}$ vertices of degree $q-1$ and at most

$$M = \sum_{k=q-1}^{\log n/100q} \frac{e^{2\omega}(\log n)^{k-q+1}}{(k-1)!} \le 2e^{2\omega}\left(\frac{e\log n}{\log n/100q}\right)^{-q+1+\log n/100q} \le n^{1/3}$$

vertices in SMALL. (The first inequality follows from the fact that the summands grow by a factor of at least $100q$.)
Thus the probability that there is an edge contradicting the claim is at most

$$2\omega n \times \frac{e^{2\omega} \times n^{1/3}}{\binom{n}{2} - m'} = o(1).$$

$\square$

We remind the reader that $q = 2\sigma$ where we only use σ colors. We apply the above analysis by identifying colors $mod\ \sigma$. We therefore have the following:

Corollary 4.9. W.h.p. the algorithm COL applied to $G_{\tau_{2\sigma}}$ yields a coloring for which $d_c^*(v) \ge 2$ for all $v \in [n]$.

Proof. We can see from the above that w.h.p. at time $\tau_{2\sigma}$ we have that $d_c(v) \ge 2$ for all $c \in [\sigma], v \in [n]$. Furthermore, by construction, for each $c \in [q], v \in [n]$ the first edge incident with v that gets color c will be in E_c^*. (The only time we place an edge in E_c^+ is when it joins two full vertices.) $\square$

From now on we think in terms of σ colors.

4.1 Expansion

For a set $S \subseteq [n]$ we let

$$N_c^*(S) = \{v \notin S : \exists u \in S\ s.t.\ uv \in E_c^*\} \subset N_c(S).$$

Let

$$\alpha = \frac{1}{10^6 q}.$$

Lemma 4.10. Then w.h.p. $|N_c^*(S)| \geq 19|S|$ for all $S \subset LARGE$, $|S| \leq \alpha n$.

Claim 4.11. At time m, for every $R \subset V(G)$ with $|R| \leq \frac{n}{(\log n)^3}$, there are w.h.p. at most $2|R|$ edges within R of color c for every $c \in [q]$.

Proof of Claim: We will show that w.h.p. every such R does not have this many edges irrespective of color. Note that the desired property is monotone decreasing, so it suffices to use (2) and show this occurs w.h.p. in $G_{n,p}$:

$$
\begin{aligned}
&\mathbb{P}\left(\exists |R| \leq \frac{n}{(\log n)^3} : |E(G[R])| > 2|R|\right) \\
&\leq \sum_{r=4}^{n/(\log n)^3} \binom{n}{r}\binom{\binom{r}{2}}{2r} p^{2r} \\
&\leq \sum_{r=4}^{n/(\log n)^3} \left(\frac{ne}{r}\left(\frac{re^{1+o(1)}\log n}{4n}\right)^2\right)^r \\
&\leq \sum_{r=4}^{n/(\log n)^3} \left(\frac{r}{n}\cdot\frac{e^{3+o(1)}(\log n)^2}{16}\right)^r = o(n^{-3}).
\end{aligned}
$$

□

Proof of Lemma 4.10:
Case 1: $|S| \leq \frac{n}{(\log n)^4}$.
We may assume that $S \cup N_c^*(S)$ is small enough for Claim 4.11 to apply (otherwise $|N_c^*(S)| \geq \frac{n}{(\log n)^3} - \frac{n}{(\log n)^4}$ so that S actually has logarithmic expansion in color c). Then, using e_c to denote the number of edges in color c, and using Theorem 4.4,

$$\frac{\epsilon \log n}{1000q}|S| \leq \sum_{v\in S} d_c^*(v) = 2e_c(S) + e_c(S, N_c^*(S)) \leq 4|S| + 2|N_c^*(S) \cup S|.$$

Hence,

$$|N_c^*(S)| \geq \frac{\epsilon \log n}{2001q}|S| \geq 19|S|,$$

which verifies the truth of the lemma for this case.

Case 2: $\frac{n}{(\log n)^4} \leq |S| \leq \frac{n}{50\log n}$.
Let

$$m_+ := \frac{n\log n}{8q}.$$

Let E_c^+, E_c^* denote the edges of E^+, E^* respectively, which are colored c. We begin by proving

Claim 4.12. $|E_c^+|, |E_c^*| \geq m_+$ w.h.p.

Proof. Once $Full$ has been formed, it follows from Lemma 4.5, that at most $(n^{1-\delta}(n - n^{1-\delta})) + \binom{n^{1-\delta}}{2} < 2n^{-\delta}\binom{n}{2}$ spaces remain in $E(Full, V \setminus Full)$ or $E(V \setminus Full)$. For each of the $m - t_\epsilon \sim (\frac{1}{2} - \epsilon)n\log n$ edges appearing thereafter, since $\lesssim n\log n < n^{-\delta}\binom{n}{2}$ edges have been placed already, each has a probability $\geq 1 - 4n^{-\delta}$ of having both ends in $Full$, independently of what has happened previously. Applying the Chernoff bounds (see for example [7], Chapter 21.4) we see that the probability that fewer than $\frac{1}{3}n\log n$ of these $(\frac{1}{2} - \epsilon)n\log n$ edges were between vertices in $Full$ is at

most $e^{-\Omega(n\log n)}$. We remind the reader that every edge with both endpoints in $Full$ is randomly colored and placed in E^+ or E^* in Step 3 of COL.
So, we may assume there are at least $\frac{1}{3q}n\log n$ of these edges in $E^+\cup E^*$ of color c in expectation and then the Chernoff bounds imply that there are at least $\frac{1}{8q}n\log n = m^+$ w.h.p. in both E^+ and E^*. $\square$

Suppose there exists S as above with $|N_c^*(S)| < \frac{\log n}{1000q}|S|$. For $F := S\cap Full$, note that $|F| \geq |S| - n^{1-\delta} = |S|(1-o(1))$. Therefore $|N_c^*(F)\cap Full| < \frac{\log n}{1000q}|S| \leq \frac{\log n}{999q}|F|$. We will show that w.h.p. there are no such $F\subseteq Full$.
We consider the graphs $H_1 = G_{|Full|,m_+}\setminus E_{t_\epsilon}$ and the corresponding independent model $H_2 = G_{|Full|,p_+}\setminus E_{t_\epsilon}$ where $p_+ \sim \frac{\log n}{4qn}$. We will show that w.h.p. H_2 contains no set F of the postulated size and small neighborhood. Together with (2) (and $\binom{|Full|}{2}p_+\to\infty$) this implies that w.h.p. H_1 has no such set either. Note that by Lemma 3.5, we see that w.h.p. at most $20|F|\log n$ edges of E_{t_ϵ} are incident with F. This calculation is relevant because $(E^*\setminus E_{t^*})$'s only dependence on E_{t_ϵ} is that it is disjoint from it.
Hence, in H_2,

$$\mathbb{P}(\exists F) \leq \sum_{f=(n-o(n))/(\log n)^4}^{n/50\log n}\ \sum_{k=0}^{\frac{\log n}{999q}f}\binom{|Full|}{f}\binom{|Full|}{k}f^k p_+^k(1-p_+)^{(|Full|-k)f-20f\log n}$$
$$\leq \sum_{f=(n-o(n))/(\log n)^4}^{n/50\log n}\ \sum_{k=0}^{\frac{\log n}{999q}f}\underbrace{\left(\frac{ne}{f}\right)^f\left(\frac{nf}{k}\cdot\frac{\log n}{qn}\right)^k e^{-nfp_+(1-o(1))}}_{u_{f,k}}.$$

Here, the ratio

$$\frac{u_{f,k+1}}{u_{f,k}} = \frac{f\log n}{q(k+1)}\left(\frac{k}{k+1}\right)^k \geq 999/e.$$

Therefore,

$$\mathbb{P}(\exists F) \leq 2\sum_{f\sim n/(\log n)^4}^{n/50\log n}\left(\frac{ne}{f}\cdot(999)^{\frac{\log n}{999q}}n^{-1/5q}\right)^f \leq 2n\left(3(\log n)^4 n^{-1/10q}\right)^{\frac{(1-o(1))n}{(\log n)^4}} = o(1).$$

Case 3: $\frac{n}{50\log n}\leq |S| \leq \frac{n}{10^6 q}$.
Choose any $S_1\subset S$ of size $\frac{n}{50\log n}$, then

$$|N_c^*(S)| \geq |N_c^*(S_1)| - |S| \geq \frac{\log n}{1000q}\cdot\frac{n}{50\log n} - \frac{n}{10^6 q} = 19\alpha n \geq 19|S|.$$

$\square$

The following corollary applies to the subgraph of $G_{\tau_{2\sigma}}$ induced by E_c^*.

Corollary 4.13. W.h.p. $|N_c^*(S)|\geq 2|S|$ for all $S\subset V(G)$ with $|S|\leq \alpha n$.

Proof. We know from Corollary 4.9 that w.h.p. at time m every vertex v has $d_c^*(v)\geq 2$. Let $S_2 = S\cap LARGE$, $S_1 = S\setminus S_2$. Then

$$|N_c^*(S)| = |N_c^*(S_1)| + |N_c^*(S_2)| - |N_c^*(S_1)\cap S_2| - |N_c^*(S_2)\cap S_1| - |N_c^*(S_1)\cap N_c^*(S_2)|$$

$$\geq |N_c^*(S_1)| + |N_c^*(S_2)| - |S_2| - |N_c^*(S_2) \cap S_1| - |N_c^*(S_1) \cap N_c^*(S_2)|.$$

Clearly, $|S_2| \leq |S| \leq \alpha n$, and so Lemma 4.10 gives $|N_c^*(S_2)| \geq 19|S_2|$, w.h.p. Also, recall from Lemma 3.2 that w.h.p. there are no small structures in G_m and since $SMALL_c \subset SMALL$ w.h.p., this means there aren't any small-c-structures either. In particular,

- No $small_c$ vertices are adjacent and there is no path of length two between $small_c$ vertices which implies that $|N_c^*(S_1)| \geq 2|S_1|$ and $|N_c^*(S_2) \cap S_1| \leq |S_2|$.
- In addition, there is no C_4 containing a $small_c$ vertex, and no path of length 4 between $small_c$ vertices. This means that $|N_c^*(S_1) \cap N_c^*(S_2)| \leq |S_2|$.

We deduce that $|N_c^*(S)| \geq 2|S_1| + 19|S_2| - 3|S_2| \geq 2|S|$. □

Recall Γ_c^* is the subgraph induced by edges of color c that are not in E^+.

Corollary 4.14. W.h.p. Γ_c^* is connected for every $c \in [q]$.

Proof. If $[S, V \setminus S]$ is a cut in Γ_c^* then Corollary 4.13 implies that $|S|, |V \setminus S| \geq \alpha n$. Let $F = S \cap Full$. Since $|V \ \setminus Full| \leq n^{1-\delta} < \frac{\alpha n}{2}$ we see that $|F|, |Full \setminus F| \geq \frac{\alpha n}{2}$. As in **Case 2** of Lemma 4.10 we show that w.h.p. no such F exists by doing the relevant computation in H_2 with $p_+ \sim \frac{\log n}{4qn}$:

$$\begin{aligned}
\mathbb{P}(\exists F) &\leq 2 \sum_{f=\frac{\alpha n}{2}}^{|Full|-\frac{\alpha n}{2}} \binom{|Full|}{f} (1-p_+)^{f(|Full|-f)-t_\epsilon} \\
&\leq n 2^n (1-p_+)^{\frac{\alpha n}{2}(n-n^{1-\delta}-\frac{\alpha}{2}n)-o(n^2)} \\
&\leq n 2^n e^{-\alpha n \log n/10q} = o(1).
\end{aligned}$$

We subtract t_ϵ from $f(|Full| - f)$ because we do not include the first t_ϵ edges in this calculation. This is because $Full$ depends on them. □

5 Rotations

We now use E_c^+ to build the Hamiltonian cycles for every color c using Pósa rotations. We let G_c denote the graph induced by the edges of color c. Given a path $P = (x_1, x_2, \ldots, x_k)$ and an edge $x_i x_k, 2 \leq i \leq k-2$ we say that the path $P' = (x_1, \ldots, x_i, x_k, \ldots, x_{i+1})$ is obtained from P by a rotation with x_1 as the fixed endpoint.

For a path P in G_c with endpoint a denote by $END(a)$, the set of all endpoints of paths obtainable from P by a sequence of Pósa rotations with a as the fixed endpoint. In this context, Pósa [15] shows that $|N_c(END(a))| < 2|END(a)|$. This is assuming that in the course of executing the rotations, no simple extension of our path is found. It follows from Corollary 4.13 that w.h.p. $|END(a)| \geq \alpha n$. For each $b \in END(a)$ there will be a path P_b of the same length as $|P|$ with endpoints a, b. We let $END(b)$ denote the set of all endpoints of paths obtainable from P_b by a sequence of Pósa rotations with b as the fixed endpoint. It also follows from Corollary 4.13 that w.h.p. $|END(b)| \geq \alpha n$ for all $b \in END(a)$. Let $END(P) = \{a\} \cup END(a)$.

An edge $u = \{x, y\}$ of color c with $y \in END(x)$ is called a *booster*. Let $P_{x,y}$ be the path of length $|P|$ from x to y implied by $y \in END(x)$. Adding the edge u to $P_{x,y}$ will either create a Hamilton cycle or imply the existence of a path of length $|P|+1$ in G_c, after using Corollary 4.14. Indeed, if the cycle C created is not a Hamilton cycle, then the connectivity of Γ_c^* implies that there is an

edge $u = xy$ of color c with $x \in V(c)$ and $y \notin V(C)$. Then adding u and removing an edge of C incident to x creates a path of length $|P| + 1$.
We start with a longest path in Γ_c^* and let $E_c^+ = \{f_1, f_2, \ldots, f_\ell\}$ where w.h.p. $\ell \geq m_+ = \frac{n \log n}{8q}$, see Claim 4.12. A round consists of an attempt to find a longer path than the current one or to close a Hamilton path to a cycle. Suppose we start a round with a path P of length k. We use rotations and construct many paths. If one of these paths has an endpoint with a neighbor outside the path then we add this neighbor to the current path and start a new round with a path of length $k + 1$. Here we only use edges not in E_c^+. Failing this we compute $END(P)$ and look for a booster in E_c^+. In the search for boosters we start from f_r assuming that we have already examined $f_1, f_2, \ldots, f_{r-1}$ in previous rounds. Now f_r is chosen uniformly from $(1 - o(1))\binom{n}{2}$ pairs and so the probability it is a booster is at least $\beta = (1 - o(1))\alpha^2$. It is clear that at most n boosters are needed to create a Hamilton cycle. Adding a booster increases the length of the current path by one, or creates a Hamilton cycle. So the probability we fail to find a Hamilton cycle of color c is at most $\mathbb{P}(Bin(m_+, \beta) \leq n) = o(1)$. We can inflate this $o(1)$ by σ to show that w.h.p. we find a Hamilton cycle in each color, completing the proof of Theorem 1.1.

6 Concluding remarks

In this paper we studied a very natural variant of the classical problem of the appearance of σ edge disjoint Hamilton cycles in a random graph process. We showed that one can color the edges of the process online so that every color class has a Hamilton cycle exactly at the moment when the underlying graph has σ edge disjoint ones.
The paper [4] shows that at the hitting time $\tau_{2\sigma+1}$ there will w.h.p. be σ edge disjoint Hamilton cycles plus an edge disjoint matching of size $\lfloor n/2 \rfloor$. It is straightforward to extend this result to the online situation. It should be clear that at time $\tau_{2\sigma+1}$ COL can be used to construct w.h.p. $E_c^*, E_c^+, c = 1, 2, \sigma + 1$ such that $E_c^* \cup E_c^+$ induce Hamiltonian graphs for $1 \leq c \leq \sigma$ and $d_{\sigma+1}^*(v) \geq 1$ for $v \in [n]$. For color $\sigma + 1$, we replace the statement of Corollary 4.13 by

$$\text{W.h.p. } |N_{\sigma+1}^*(S)| \geq |S| \text{ for all } S \subset V(G) \text{ with } |S| \leq \alpha n. \tag{15}$$

We then replace rotations by alternating paths, using $E_{\sigma+1}^+$ as boosters. The details are as described in Chapter 6 of [7]. In outline, let $G = (V, E)$ be a graph without a matching of size $\lfloor |V(G)|/2 \rfloor$. For $v \in V$ such that v is isolated by some maximum matching, let

$$A(v) = \{w \in V : w \neq v \text{ and } \exists \text{ a maximum matching of } G \text{ that isolates } v \text{ and } w\}.$$

We use the following lemma

Lemma 6.1. Let G be a graph without a matching of size $\lfloor |V(G)|/2 \rfloor$. Let M be a maximum matching of G. If $v \in V$ and $A(v) \neq \emptyset$ then $|N_G(A(v))| < |A(v)|$.

We start with a maximum matching M of $\Gamma_{\sigma+1}^*$. Suppose that v is not covered by M. Using (15), we see that w.h.p. $|A(v)| \geq \alpha n$. Further, if $u \in A(v)$ and $uv \in E_{\sigma+1}^+$ then adding this edge gives a larger matching. Also, because u is isolated by a maximum matching, there is a corresponding set A_u of size at least αn such if $w \in A_u$ and $uw \in E_{\sigma+1}^+$ then we can find a larger matching. Therefore we have $\Omega(n^2)$ boosters and the proof is similar to that for Hamilton cycles.
There are several related problems which can likely be treated using our approach. One potential application for our technique is to show that for any fixed positive integer k and any decomposition $k = k_1 + ... + k_s$ into the sum of s positive integers, there is an online algorithm, coloring the edges

of a random graph process in s colors so that exactly at the hitting time τ_k the i-th color forms a k_i-connected spanning graph for $i = 1, \ldots, s$. In general, one can generate many more interesting problems by considering the online Ramsey version of other results in the theory of random graphs.
Acknowledgement: We thank the referee(s) for their comments.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi, First occurrence of Hamilton cycles in random graphs, *in* Cycles in graphs (Burnaby, B.C., 1982), North-Holland Mathematical Studies **115**, North-Holland, Amsterdam (1985), 173–178.

[2] T. Bohman, A. Frieze, M. Krivelevich, P. Loh and B. Sudakov, Ramsey games with giants, *Random Structures and Algorithms* 38 (2011), 68–99.

[3] B. Bollobás, The evolution of sparse graphs, *in* "Graph theory and combinatorics, Proceedings of a Cambridge Combinatorial Conference in Honour of Paul Erdős, 1984" (B. Bollobas, Ed.), 35–57.

[4] B. Bollobás and A.M. Frieze, On matchings and hamiltonian cycles in random graphs, *Annals of Discrete Mathematics* 28 (1985), 23–46.

[5] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* 5A (1960), 17–61.

[6] P. Erdős and A. Rényi, On the strength of connectedness of a random graph, *Acta. Math. Acad. Sci. Hungar.*, 8 (1961), 261–267.

[7] A.M. Frieze and M. Karoński, **Introduction to Random Graphs**, Cambridge University Press, 2016.

[8] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association* 58 (1963), 13–30.

[9] F. Knox, D. Kühn and D. Osthus, Edge-disjoint Hamilton cycles in random graphs, *Random Structures and Algorithms* 46 (2015), 397–445.

[10] A. Korshunov, Solution of a problem of Erdős and Rényi on Hamilton cycles non-oriented graphs, *Soviet Math. Dokl.*, 17 (1976), 760–764.

[11] J. Komlós and E. Szemerédi, Limit distribution for the existence of Hamilton cycles in random graphs, *Discrete Math*, 43 (1983), 55–63.

[12] M. Krivelevich, E. Lubetzky and B. Sudakov, Hamiltonicity thresholds in Achlioptas processes, *Random Structures and Algorithms* 37 (2010), 1–24.

[13] M. Krivelevich and W. Samotij, Optimal packings of Hamilton cycles in sparse random graphs, *SIAM Journal on Discrete Mathematics* 26 (2012), 964–982.

[14] C. Lee, B. Sudakov and D. Vilenchik, Getting directed Hamilton cycle twice faster, *Combinatorics, Probability and Computing* 21 (2012), 773–801.

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