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 $2\sigma$ coincides with the hitting time for $\sigma$ 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 $\sigma$ colors so that at time $\tau_{2\sigma}$, each color class is Hamiltonian.

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 $\omega$ 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 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 \( H_\sigma \) of having \( \sigma \) 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 \( H_\sigma \) when the minimum degree becomes \( 2\sigma \). It took quite a while, but this result was extended to the more difficult case of growing \( \sigma \) 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 \( \sigma \) edge disjoint Hamilton cycles in an online fashion by the time the minimum degree becomes \( 2\sigma \)? We answer this question affirmatively in the case \( \sigma = O(1) \).

**Theorem 1.1.** For a fixed integer \( \sigma \geq 2 \), let \( r_{2\sigma} \) denote the hitting time for the random graph process \( G_i, i = 1, 2, \ldots \) to have minimum degree \( 2\sigma \). Then w.h.p. we can color the edges of \( G_i, i = 1, 2, \ldots \) online with \( \sigma \) colors so that \( G_{r_{2\sigma}} \) contains \( \sigma \) 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 \( \Gamma^*_c \) be the subgraph of \( G_{r_{2\sigma}} \) 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 \( \Gamma^*_c = \Gamma^*_c \cup \Gamma^*_c+c+\sigma \) 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 $\Gamma_c$.

**Definition 2.1.** Let $\text{Full}$ 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 $\epsilon$ is some sufficiently small constant depending only on the constant $q$. The actual value of $\epsilon$ needed will depend on certain estimates below being valid, in particular equation (14). A vertex is $\text{Full}$ if is lies in $\text{Full}$. Similarly, let $\text{Full}' \subseteq \text{Full}$ denote the set of vertices with degree at least $\frac{\epsilon \log n}{1000q}$ in every color at time $t^* := \frac{1}{2} \epsilon n \log n$.

This definition only makes sense if $t_\epsilon$ 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 $C_v(t) := \{ c \in [q] : d_{c,v}(t) = 0 \}$ 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**

1. **Step 1** Let $e_t = uv$.

2. **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 $E^*_c$.

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

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

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

### 3
3 Structural properties

Let
\[ p = \log n + (q - 1) \log \log n - \omega \]
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:

\[ \mathbb{P}(G_{n,m} \in \mathcal{P}) \leq 10m^{1/2} \mathbb{P}(G_{n,p} \in \mathcal{P}). \quad (1) \]
\[ \mathbb{P}(G_{n,m} \in \mathcal{P}) \leq 3 \mathbb{P}(G_{n,p} \in \mathcal{P}), \quad \text{if \( \mathcal{P} \) is monotone.} \quad (2) \]

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

**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 \( \phi : H \to G_{n,m} \) such that \( \phi(S(H)) \subseteq \text{SMALL} \). 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:

\[
\mathbb{P}(G_{n,p} \text{ contains } H) \leq \binom{n}{h} h! p^f \left( \sum_{i=0}^{\log n} \binom{n-h}{i} p^i (1-p)^{n-h-i} \right)^s \\
\leq n^h \left( \frac{\log n}{n} \right)^f \left( \sum_{i=0}^{\log n} \left( \frac{(e + o(1)) \log n}{i} \right)^i e^{-\log n - (q-1) \log \log n + o(1)} \right)^s \\
\leq n^h \left( \frac{\log n}{n} \right)^f \left( \frac{(300q \log n)}{n \log n^{q+1-o(1)}} \right)^s 
\]
\[ = o(n^{h-f-s+1/4}) = o(n^{-3/4}). \]

(We used the notation \( A \lesssim B \) in place of \( A \leq (1 + o(1))B. \) 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 \( \log n^{100q} \).

Lemma 3.3. W.h.p., for every \( k \in [q - 1, \log n^{100q}] \), 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. \( \nu_k \) is increasing in \( k \) for this range, and for the largest \( k = \log n^{100q} \) we have \( \nu_k \lesssim n^{\log(100eq^{100q})} \).

Proof. Fix \( k \) and then we have

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

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,

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

Applying (2) we see that

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

which is stronger than required.

Lemma 3.5. With probability \( 1 - o(n^{-10}) \), \( G_{n,m} \) has no vertices of degree \( \geq 20 \log n \).
Proof. We will prove that w.h.p. \( G_{n,p} \) has the stated property. We can then obtain the lemma by applying (2).

\[
\mathbb{P}(\exists v : d(v) \geq 20 \log n) \leq n \left( \frac{n-1}{20 \log n} \right)^{20 \log n} \\
\leq n \left( \frac{en}{20 \log n} \right)^{20 \log n} \\
\leq n \left( \frac{e}{10} \right)^{20 \log n} \\
= o(n^{-10}).
\]

\[ \square \]

4 Analysis of COL

Let \( \Gamma = G_m \) and let \( d(v) \) denote the degree of \( v \in [n] \) in \( \Gamma \). 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. \( t(e_i) = i \). Then 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 at time \( t_1 \). 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 \( S \) of \( k - q + 2 + \theta_v \) edges incident with \( v \) which fail to contribute to \( C_v^{(m)} \). Let \( I_S = [2,k] \setminus S \) and \( T_S \) denote \( \{ t_i : i \in I_S \} \), and 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). Note that \( \mathbb{P}(\mathcal{L}) = 1 - o(n^{-10}) \).

Taking a union bound over all \( S \),

\[
\mathbb{P}(|C_v^{(m)}| > \theta_v | \mathcal{L}) \leq \sum_{S \subset [2,k]} \sum_{|S|=q-2-\theta_v} \frac{1}{m} \mathbb{P}\left( \bigwedge_{i \in [2,k]\setminus S} A_i \bigg| T_S, \mathcal{L} \right) \\
= \sum_{S \subset [2,k]} \sum_{|S|=q-2-\theta_v} \frac{t_{i_2} - 1}{m} \mathbb{P}\left( \bigwedge_{i \in [2,k]\setminus S} A_i \bigg| T_S, \mathcal{L} \right),
\]

since there are \( t_{i_2} - 1 \) choices for \( i_1 \). Here

\[
A_i := \left\{ C_{u_i}^{(t_i)} \neq \emptyset \right\},
\]
and \( \{ 1 = i_1 < i_2 < \cdots < i_s \} \) are the elements of \( [k] \setminus S \) in increasing order for a fixed \( S \), where 
\[ s := k - |S| = k - q + 2 + \theta_v. \]
Next let
\[
Y_t = \{ \text{edges of } u_i \text{ that appeared before } t_i \text{ excluding edges contained in } N^{(m)}(v) \},
\]
\[
d_r = d(u_r) \text{ and } Z_r := |Y_r| \text{ for } r = 2, 3, \ldots, s,
\]
\[
D_S = \{ d_i : i \in I_S \}.
\]

Now fix \( S \) and \( T_S \) and \( D_S \).

**Remark 4.2.** Going back to Algorithm COL, we observe that Step 5 implies that if \( C^{(t)}_v \neq \emptyset \) then \( uv \) is colored with a color in \( C^{(t)}_v \) with probability at least \( \frac{1}{q} \). This holds regardless of the previous history of the algorithm and also holds conditional on \( T_S, D_S \). 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 \( T_S, D_S \).

Then,
\[
\mathbb{P}(A_{i_2} \land \cdots \land A_{i_s} \mid T_S, D_S, \mathcal{L}) = \sum_{z_s} \mathbb{P}(A_{i_{s-1}} \mid A_{i_2}, \ldots, A_{i_{s-1}}, Z_{s-1} = z_{s-1}, Z_s = z_s, T_S, D_S, \mathcal{L}) \mathbb{P}(A_{i_2}, \ldots, A_{i_s} = z_s \mid T_S, D_S, \mathcal{L})
\]
\[
\leq \sum_{z_s} g(z_s) \sum_{z_{s-1}} \mathbb{P}(A_{i_{s-1}} \mid A_{i_2}, \ldots, A_{i_{s-1}}, Z_{s-1} = z_{s-1}, Z_s = z_s, T_S, D_S, \mathcal{L})
\]
\[
\times \mathbb{P}(A_{i_2}, \ldots, A_{i_{s-1}}, Z_{s-1} = z_{s-1}, Z_s = z_s \mid T_S, D_S, \mathcal{L})
\]
\[
\leq \sum_{z_s, z_{s-1}} g(z_s) g(z_{s-1}) \mathbb{P}(Z_r = z_r, r = 2, \ldots, s \mid T_S, D_S, \mathcal{L}) \quad \text{(by induction)}
\]
\[
(4)
\]
Here \( g(z) := \mathbb{P}((\text{Bin}(z, q^{-1}) \leq q - 1) \text{ for any } z \geq 0.

**Claim 4.3.**
\[
\mathbb{P}(Z_r = z_r, r = 2, 3, \ldots, s \mid T_S, D_S, \mathcal{L}) \leq \left( 1 + \tilde{O}(n^{-1}) \right) \prod_{r=2}^{s} \frac{t_r}{(m-t_r)} \frac{(m-t_r)}{(m-d_r)},
\]
where \( \tilde{O} \) hides polylog factors.

**Proof** Fix \( \frac{\log n}{100q} \leq d_1, d_2, \ldots, d_s = O(\log n) \) and \( t_2, t_3, \ldots, t_s \). Then, for every \( 1 < r \leq s \),
\[
\mathbb{P}(Z_r = z_r \mid Z_{r-1} = z_{r-1}, \ldots, Z_2 = z_2, T_S, D_S, \mathcal{L}) \leq \left( 1 + o(n^{-10}) \right) \frac{(t_r)}{(m-t_r)} \frac{(m-t_r)}{(m-d_r - d_{r-1} - s)}
\]
\[
\leq \left( 1 + \tilde{O}(n^{-1}) \right) \frac{(t_r)}{(m-t_r)} \frac{(m-t_r)}{(m-d_r)}. \quad \text{(5)}
\]

**Explanation for (5):** 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 \( z_2, z_3, \ldots, z_{r-1} \). 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 $d_2, d_3, \ldots, d_{r-1}$. We subtract the extra $s$ to (over)count for edges from $v$ to $u_{r+1}, \ldots, u_s$. The factor $(1 + o(n^{-10}))$ accounts for the conditioning on $L$.

Expanding $P(Z_r = z_r, r = 2, \ldots, s \mid T_S, D_S, L)$ 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 $d_2, d_3, \ldots, d_s$,

$$P(A_{d_2} \land \cdots \land A_{d_s} \mid T_S, D_S, L)$$

$$\leq s \prod_{r=2}^{d_r} \sum_{z_r=0}^{d_r} \left( P(Bin(z_r, q^{-1}) \leq q - 1) \times \binom{t_r}{z_r} \frac{d_r - z_r}{m} \right)$$

$$\leq s \prod_{r=2}^{d_r} \sum_{z_r=0}^{d_r} \left( C_2 \left( \min\{z_r, q - 1\} \right) \left( 1 - \frac{1}{q} \right)^{z_r} \times \binom{t_r}{z_r} \frac{d_r - z_r}{m} \right)$$

$$\leq s \prod_{r=2}^{d_r} \sum_{z_r=0}^{d_r} \left( C_2 z_r^{q-1} e^{-z_r/q} \times \binom{t_r}{z_r} \frac{d_r - z_r}{m} \right).$$

Here, $C_2 = C_2(q)$ depends only on $q$. We will use constants $C_3, C_4, \ldots$ in a similar fashion without further comment.

**Justification for (6):** If $z_r \leq q - 1$ then $P(Bin(z_r, q^{-1}) \leq q - 1) = 1$ and $C_2 = eq^q$ will suffice. If $q \leq z_r \leq 10q$ we use

$$P(Bin(z_r, q^{-1}) \leq q - 1) \leq 1 \text{ and } \left( \frac{z_r}{q - 1} \right) \frac{1}{q - 1} \geq \left( \frac{z}{q - 1} \right)^{10q}$$

and $C_2 = e^{20q^q}$ will suffice in this case.

If $z_r > 10q$ then putting $a_i := P(Bin(z_r, q^{-1}) = i) = \binom{z}{i} \frac{1}{q^i} \left( 1 - \frac{1}{q} \right)^{z_i - i}$ for $i \leq q - 1$ we see that

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

So here

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

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

and thus $C_2 = (5q)^q$ suffices.

Now, writing $(t)_j$ for the falling factorial $t! / (t - z)! = t(t - 1)(t - 2) \cdots (t - z + 1)$,

$$\binom{t_r}{z_r} \frac{d_r - z_r}{m} = \binom{d_r}{z_r} \frac{(t_r)_{z_r} (m - t_r)_{d_r - z_r}}{(m)_{d_r}}$$

$$= \binom{d_r}{z_r} \prod_{i=0}^{r-1} \frac{t_r - i}{m - (d_r - z_r) - i} \cdot \prod_{i=0}^{r-1} \frac{t_r - i}{m - i}.$$
It follows from (3) and (10) that,

\[
\leq \left(1 + O\left(\frac{d_r^2}{m}\right)\right) \left(\frac{d_r}{z_r}\right) \left(\frac{t_r}{m}\right) z_r \left(1 - \frac{t_r}{m}\right)^{d_r - z_r}.
\]  

(8)

Observe next that if \( z_r \geq q^2 \) then

\[
(z_r)^{q-1} = \left(z_r^q\right)^{1 - \frac{i}{z_r}} \geq z_r^q \left(1 - \frac{q^2}{2z_r}\right) \geq \frac{z_r^{q-1}}{2}.
\]  

(9)

It follows from (8) and (9) that

\[
\sum_{d_r} C_2 z_r^{q-1} e^{-z_r/q} \times \frac{(t_r)(m-t_r)}{(d_r-z_r)} \leq 2C_2 \sum_{z_r=q^2}^{d_r} (z_r)^{q-1} \left(\frac{d_r}{z_r}\right) \left(\frac{t_r}{m}\right) z_r \left(1 - \frac{t_r}{m}\right)^{d_r - z_r}
\]

\[
\leq 2C_2 (d_r/t_r) z_r^{q-1} \left(\frac{t_r}{m}\right) \sum_{z_r=q^2}^{d_r} (d_r - q + 1) \left(\frac{t_r}{m}\right) z_r^{q-1} \left(1 - \frac{t_r}{m}\right)^{d_r - z_r}
\]

\[
\leq 2C_2 \left(\frac{d_r t_r}{m}\right)^{q-1} \left(1 - \frac{t_r}{m}\right) \left(1 - e^{-1/q}\right)^{d_r - q+1}
\]

\[
\leq 2C_2 \left(\frac{d_r t_r}{m}\right)^{q-1} \exp \left\{ -\left(\frac{d_r - q + 1}{m}\right) t_r \left(1 - e^{-1/q}\right) \right\}.
\]

Furthermore, not forgetting

\[
\sum_{z_r=0}^{q^2-1} \left(1 + \frac{C_2}{z_r + 1}\right) z_r^{q-1} e^{-z_r/q} \times \frac{(t_r)(m-t_r)}{(d_r-z_r)} \leq C_3 \sum_{z_r=1}^{q^2-1} \left(\frac{t_r}{z_r}\right) \left(\frac{m-t_r}{d_r}\right) \leq C_4 \sum_{z_r=1}^{q^2-1} \frac{t_r^{z_r}}{(m-t_r)^{d_r-z_r}} \frac{d_r!}{m^{d_r}}.
\]

\[
\leq C_4 \sum_{z_r=1}^{q^2-1} \frac{t_r^{z_r}}{(m-t_r)^{d_r-z_r}} \frac{d_r!}{m^{d_r}}.
\]

\[
\leq C_4 \sum_{z_r=1}^{q^2-1} \left(\frac{d_r t_r}{m}\right)^{z_r} e^{-\left(\frac{d_r - z_r}{d_r}\right) t_r/m}
\]

\[
\leq C_5 \psi \left(\frac{d_r t_r}{m}\right).
\]

where \( \psi(x) = e^{-x} \sum_{z_r=1}^{q^2-1} x^z \). (Now \( z_r \leq q^2 \) and so the factor \( e^{-z_r t_r/m} \leq e^{q^2} \) can be absorbed into \( C_5 \).) Going back to (7) we have

\[
\mathbb{P}(A_{i_2} \wedge \cdots \wedge A_{i_s} \mid T_S, D_S, \mathcal{L}) \leq C_6 \prod_{r=2}^{s} \left(\frac{d_r t_r}{m}\right)^{q-1} \exp \left\{ -\frac{d_r t_r}{m} \left(1 - e^{-1/q}\right) \right\} + \psi \left(\frac{d_r t_r}{m}\right).
\]  

(10)

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

\[
p_u := \mathbb{P}(|C_v^{(m)}| > \theta_u \mid T_S, D_S, \mathcal{L})
\]
Replacing a sum of products by a product of sums and dividing by \((s-1)!\) to account for repetitions, we get

\[
p_v \leq \sum_{S \subseteq [2,k]} \sum_{|S| = q-2-\theta_v} \frac{C_6^S}{(m^S_s)} (s-1)! \prod_{r=3}^{s} \left( \int_{t=0}^{\infty} \left( \frac{d_r t_r}{m} \right)^{q-1} \exp \left\{ - \frac{d_r t_r}{m} (1 - e^{-1/q}) \right\} + \psi \left( \frac{d_r t_r}{m} \right) \right) dt \times \left( \int_{t=0}^{\infty} \left( \frac{d_2 t_2}{m} \right)^{q-1} \exp \left\{ - \frac{d_2 t_2}{m} (1 - e^{-1/q}) \right\} + t \psi \left( \frac{d_2 t_2}{m} \right) \right) dt \]

We now replace the sums by integrals. This is valid seeing as the summands have a bounded number of extrema, and we replace \(C_6\) by \(C_7\) to absorb any small error factors.

\[
p_v \leq \sum_{S \subseteq [2,k]} \sum_{|S| = q-2-\theta_v} \frac{C_7^S}{(m^S_s)} (s-1)! \prod_{r=3}^{s} \left( \int_{t=0}^{\infty} \left( \frac{d_r t_r}{m} \right)^{q-1} \exp \left\{ - \frac{d_r t_r}{m} (1 - e^{-1/q}) \right\} + \psi \left( \frac{d_r t_r}{m} \right) \right) dt \times \left( \int_{t=0}^{\infty} \left( \frac{d_2 t_2}{m} \right)^{q-1} \exp \left\{ - \frac{d_2 t_2}{m} (1 - e^{-1/q}) \right\} + \psi \left( \frac{d_2 t_2}{m} \right) \right) dt \]

\[
= \sum_{S \subseteq [2,k]} \sum_{|S| = q-2-\theta_v} \frac{C_7^S}{(m^S_s)} (s-1)! \prod_{r=3}^{s} \left( \int_{t=0}^{\infty} \left( \frac{d_r t_r}{m} \right)^{q-1} \exp \left\{ - \frac{d_r t_r}{m} (1 - e^{-1/q}) \right\} + \psi \left( \frac{d_r t_r}{m} \right) \right) dt \times \left( \int_{t=0}^{\infty} \left( \frac{d_2 t_2}{m} \right)^{q-1} \exp \left\{ - \frac{d_2 t_2}{m} (1 - e^{-1/q}) \right\} + \psi \left( \frac{d_2 t_2}{m} \right) \right) dt \]

\[
\leq \sum_{S \subseteq [2,k]} \sum_{|S| = q-2-\theta_v} \frac{C_7^S}{(m^S_s)} (s-1)! \left( \frac{C_8 m}{\min_r \{d_r^s \}} \right)^s \leq \frac{C_9^S k^{q-2-\theta_v}}{(\log n)^{k-\theta_v}}.
\]

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

\[
\mathbb{P}(\exists v : |C_v^{(m)}| > \theta_v) \leq \mathbb{P}(\mathcal{L}) + \sum_{k=q-1}^{k_0} \frac{e^{2\omega} (\log n)^{k-\theta_v}}{(k-1)!} \times \frac{C_9^k k^{q-2+\theta_v}}{(\log n)^{k-\theta_v}} + \sum_{k=k_0}^{20\log n} \frac{C_9^k k^{q-2}}{(\log n)^{k-\theta_v}} \leq o(1) + \frac{e^{2\omega}}{\log n} \sum_{k=q-1}^{k_0} \frac{C_9^k k^{q-2+\theta_v}}{(k-1)!} + \sum_{k=k_0}^{20\log n} \frac{(20C_9)^k}{(\log n)^{k/2}} \leq o(1) + \frac{e^{2\omega}}{\log n} \sum_{k=q-1}^{k_0} \frac{(C_9 e)^k}{k^{k-\theta_v}}.
\]
Theorem 4.4. There exists $\epsilon = \epsilon(q) > 0$ such that w.h.p. on completion of COL every $v \in \text{LARGE}$ has $d^*_c(v) \geq \frac{\epsilon \log n}{100q}$ for all $c \in [q]$.

Suppose we define a vertex to be $small_c$ if it has $d_c(v) \leq \frac{\epsilon \log n}{100q}$. Theorem 4.4 says w.h.p. the set of $small_c$ vertices $\text{SMALL}_c \subset \text{SMALL}$ 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

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

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

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

The first inequality follows from the fact that the ratio of successive 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 $d(t_\epsilon/2)(v) \geq \frac{\epsilon \log n}{100}$. 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 $\geq \frac{\epsilon \log n}{100}$ 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 $d(t_\epsilon/2)(v) \geq \frac{\epsilon \log n}{100}$ and $\frac{\epsilon \log n}{200}$ edges of fully random color. For such a $v$, and any color $c$,

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

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

$$|Full'| \geq n - \frac{201qn}{\epsilon \log n} - n^{1-\epsilon/2000q} \geq n - \frac{202qn}{\epsilon \log n}. \tag{14}$$

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.

$$= o(1). \tag{15}$$
Now, for \( v \notin \text{Full} \), let \( S(v) := \text{Full} \setminus N^{(t_e/2)}(v) \). Since \( d^{(t_e/2)}(v) \leq d(v) \leq 20 \log n \), we have \(|S(v)| \geq n - \frac{203m}{\epsilon \log n} \). Furthermore, every \( w \in S(v) \) is no longer needy, and so among the next \( t_e/2 \) 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 \( |e(v, S(v))| \geq \frac{\epsilon \log n}{100} \), we have \( \mathbb{P}(v \notin \text{Full}) \leq qn^{-\epsilon/1000q} \) by a similar calculation to (13) and so \(|\text{Full}| \geq n - n^{1-\epsilon/2000q} \) with probability \( \geq 1 - O(n^{-\epsilon/2000q}) \) by the Markov inequality.

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

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

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

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

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

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

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

\[
\mathbb{P} \left( d(v) \geq \frac{\log n}{200q} \right) \leq \left( \frac{n}{\log n/200q} \right)^{\log n/200q} \leq (400q\epsilon)^{\log n/200q} = o(n^{-2}),
\]

for \( \epsilon \) sufficiently small.

The result follows by taking a union bound over choices of \( v \) and using (2) (again noting \( \binom{n}{2} p_2 = t_e \to \infty \)).

**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 \( \text{Full} \) that occur after time \( t_e \). These edges will provide all needed edges of all colors.

It is known that w.h.p. \( m \leq \tau_q \leq m' = m + 2\omega n \), 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 \( \tau_q \) 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 \( \text{LARGE} \). 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 \( \text{LARGE} \).
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

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

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

$$2\omega n \times e^{2\omega} \times n^{1/3} = o(1).$$

We remind the reader that $q = 2\sigma$ where we only use $\sigma$ 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 $d^*_c(v) \geq 2$ for all $v \in [n]$.

**Proof.** We can see from the above that w.h.p. at time $\tau_{2\sigma}$ we have that $d_c(v) \geq 2$ 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.)

From now on we think in terms of $\sigma$ colors.

### 4.1 Expansion

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

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

Let

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

**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 $G_{n,p}$:

$$P \left( \exists |R| \leq \frac{n}{(\log n)^3} : |E(G[R])| > 2|R| \right)$$

$$\leq \sum_{r=4}^{n/(\log n)^3} \binom{n}{r} \binom{r}{2r} p^{2r}$$

13
\[ \leq \frac{n/(\log n)^3}{2} \left( \frac{ne}{r} \left( \frac{re^{1+o(1)} \log n}{4n} \right) \right)^r \]
\[ \leq \frac{n/(\log n)^3}{2} \left( \frac{r}{n} \cdot \frac{e^{3+o(1)}(\log n)^2}{16} \right)^r = o(n^{-3}). \]

**Proof of Lemma 4.10:**

**Case 1:** \(|S| \leq \frac{n}{(\log n)^3} \).

We may assume that \( S \cup N^*_c(S) \) is small enough for Claim 4.11 to apply (otherwise \(|N^*_c(S)| \geq \frac{n}{(\log n)^3} - \frac{n}{(\log n)^3} \delta \) 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,

\[ \frac{\epsilon \log n}{1000q} \leq \sum_{v \in S} d^*_c(v) = 2e_c(S) + e_c(S, N^*_c(S)) \leq 4|S| + 2|N^*_c(S) \cup S|, \]

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:** \( \frac{n}{(\log n)^3} \leq |S| \leq \frac{n}{50(\log n)}. \)

Let

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

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

**Claim 4.12.** \(|E^+_c|, |E^*_c| \geq m_+ \text{ w.h.p.}\)

**Proof.** Once \( \text{Full} \) has been formed, it follows from Lemma 4.5, that at most \((n^{1-\delta(n - n^{1-\delta})) + \binom{n^{1-\delta}}{2} < 2n^{-\delta}(\frac{n}{2})^2} \) spaces remain in \( E(\text{Full}, V \setminus \text{Full}) \) or \( E(V \setminus \text{Full}) \). For each of the \( m - t_+ \sim (\frac{1}{2} - \epsilon)n \log n \) edges appearing thereafter, since \( \leq n \log n < n^{-\delta}(\frac{n}{2}) \) edges have been placed already, each has a probability \( \geq 1 - 4n^{-\delta} \) of having both ends in \( \text{Full} \), 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 \( \text{Full} \) is at most \( e^{-\Omega(n \log n)} \). We remind the reader that every edge with both endpoints in \( \text{Full} \) is randomly colored and placed in \( E^+ \) or \( E^*_c \) in Step 3 of \( \text{COL} \).

So, we may assume there are at least \( \frac{1}{3q}n \log n \) of these edges in \( E^+ \cup E^*_c \) of color \( c \) in expectation and then the Chernoff bounds imply that there are at least \( \frac{1}{8q}n \log n = m_+ \text{ w.h.p.} \) in both \( E^+ \) and \( E^*_c \).

Suppose there exists \( S \) as above with \(|N^*_c(S)| < \frac{\log n}{1000q}|S| \). For \( F := S \cap \text{Full} \), note that \(|F| \geq |S| - n^{1-\delta} = |S|(1 - o(1))\). Therefore \(|N^*_c(F) \cap \text{Full}| < \frac{\log n}{100q}|S| \leq \frac{\log n}{999q}|F| \). We will show that w.h.p. there are no such \( F \subseteq \text{Full} \).

We consider the graphs \( H_1 = G_{|\text{Full}|, m_+} \setminus E_t \) and the corresponding independent model \( H_2 = G_{|\text{Full}|, p_+} \setminus E_t \) where \( p_+ \sim \frac{\log n}{4qm} \). We will show that w.h.p. \( H_2 \) contains no set \( F \) of the postulated size and small neighborhood. Together with (2) and \((|\text{Full}|)p_+ \to \infty\) 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
$E_t$ are incident with $F$. This calculation is relevant because $E^*$'s only dependence on $E_t$ is that it is disjoint from it.

Hence, in $H_2$,

$$
\mathbb{P}(\exists F) \leq \sum_{f=(n-o(n))/\log n}^{n/50 \log n} \sum_{k=0}^{\log n \times f 999} (|Full| f)^k \left(1 - p_+\right)^{(f-20f \log n)} \leq \sum_{f=(n-o(n))/\log n}^{n/50 \log n} \sum_{k=0}^{\log n \times f 999} \left(\frac{ne}{f} \cdot (999)^{\log n \times f 999} n^{-1/5q}\right)^f \leq 2n \left(3(\log n)^4 n^{-1/10q}\right)^{(1-o(1)) n \log n} = o(1).
$$

Here, the ratio

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

Therefore,

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

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

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

The following corollary applies to the subgraph of $G_{2^\omega}$ 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 $d_v^*(v) \geq 2$. Let $S_2 = S \cap \text{LARGE}$, $S_1 = S \setminus S_2$. Then

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

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 $\text{SMALL}_c \subset \text{SMALL}$ 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 $|N_c^*(S)| \geq 2|S_1| + 19|S_2| - 3|S_2| \geq 2|S|$. \(\square\)

Recall $\Gamma_c^*$ is the subgraph induced by edges of color $c$ that are not in $E^+$.\(\square\)
Corollary 4.14. W.h.p. \( \Gamma_c^* \) is connected for every \( c \in [q] \).

Proof. If \([S, V \setminus S]\) is a cut in \( \Gamma_c^* \) then Corollary 4.13 implies that \(|S|, |V \setminus S| \geq an\). Let \( F = S \cap \text{Full} \). Since \(|V \setminus \text{Full}| \leq n^{1-\delta} < \frac{an}{2}\) we see that \(|F|, |\text{Full} \setminus F| \geq \frac{an}{2}\). 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 \( p_+ \sim \frac{\log n}{4qn} \).

\[
\mathbb{P}(\exists F) \leq 2 \sum_{f=\frac{an}{2}}^{\frac{|\text{Full}|-an}{2}} \binom{|\text{Full}|}{f} (1-p_+)^f (|\text{Full}|-f-t),
\]

\[
\leq n^{2n(1-p_+)} \frac{an}{2}(n-n^{1-\delta}-\frac{2n}{4}-o(n^2))
\]

\[
\leq n^{2n}e^{-an\log n/10q} = o(1).
\]

We subtract \( t \) from \( f(|\text{Full}|-f) \) because we do not include the first \( t \) edges in this calculation. This is because \( \text{Full} \) depends on them. \( \square \)

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 \( x_ix_k, 2 \leq i \leq k - 2 \) 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 \( \text{END}(a) \), 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 \(|N_c(\text{END}(a))| < 2|\text{END}(a)|\). 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. \(|\text{END}(a)| \geq an\). For each \( b \in \text{END}(a) \) there will be a path \( P_b \) of the same length as \(|P| \) with endpoints \( a, b \). We let \( \text{END}(b) \) 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. \(|\text{END}(b)| \geq an\) for all \( b \in \text{END}(a) \). Let \( \text{END}(P) = \{a\} \cup \text{END}(a) \).

An edge \( u = \{x, y\} \) of color \( c \) with \( y \in \text{END}(x) \) is called a booster. Let \( P_{x,y} \) be the path of length \(|P| \) from \( x \) to \( y \) implied by \( y \in \text{END}(x) \). 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 \( \Gamma_c^* \) 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 \( \Gamma_c^* \) and let \( E_c^+ = \{f_1, f_2, \ldots, f_t\} \) where w.h.p. \( t \geq m_+ = \frac{n\log n}{8q} \), 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 \( \text{END}(P) \) and look for a booster in \( E_c^+ \). In the search for boosters we start from \( f_i \) assuming that we have already examined \( f_1, f_2, \ldots, f_{i-1} \) in previous rounds. Now \( f_i \) is chosen uniformly from \( (1-o(1))(\frac{n}{2}) \) 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 \( \mathbb{P}(\text{Bin}(m_+, \beta) \leq n) = o(1) \). We can inflate this \( o(1) \) by \( \sigma \) 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 $\sigma$ 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 $\sigma$ edge disjoint ones.

The paper [4] shows that at the hitting time $\tau_{2\sigma+1}$ there will w.h.p. be $\sigma$ 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. $E^c, E^+, c = 1, 2, \sigma+1$ such that $E^c \cup E^+$ induce Hamiltonian graphs for $1 \leq c \leq \sigma$ and $d^c_{\sigma+1}(v) \geq 1$ for $v \in [n]$. For color $\sigma + 1$, we replace the statement of Corollary 4.13 by

\[ \text{w.h.p. } |N^*_{\sigma+1}(S)| \geq |S| \text{ for all } S \subset V(G) \text{ with } |S| \leq \alpha n. \]  

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 $\Gamma^*_\sigma$. 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 $uw \in E^+_{\sigma+1}$ 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 $\alpha n$ 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 $k = k_1 + \ldots + k_s$ 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 $\tau_k$ 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.

**References**


