4-9-2015

Elegantly colored paths and cycles in edge colored random graphs

Lisa Espig
Carnegie Mellon University

Alan Frieze
Carnegie Mellon University, af1p@andrew.cmu.edu

Michael Krivelevich
Tel Aviv University

Elegantly colored paths and cycles in edge colored random graphs

Lisa Espig * Alan Frieze † Michael Krivelevich‡

April 9, 2015

Abstract

We first consider the following problem. We are given a fixed perfect matching M of $[n]$ and we add random edges one at a time until there is a Hamilton cycle containing M. We show that w.h.p. the hitting time for this event is the same as that for the first time there are no isolated vertices in the graph induced by the random edges. We then use this result for the following problem. We generate random edges and randomly color them black or white. A path/cycle is said to *zebraic* if the colors alternate along the path. We show that w.h.p. the hitting time for a zebraic Hamilton cycle coincides with every vertex meeting at least one edge of each color. We then consider some related problems and extend to multiple colors.

1 Introduction

This paper studies the existence of nicely structured objects in (randomly) colored random graphs. Our basic interest will be in what we call *zebraic* paths and cycles. We assume that the edges of a graph G have been colored black or white. A path or cycle will be called *zebraic* if the edges alternate in color along the path. We view this as a variation on the usual theme of *rainbow* paths and cycles that have been well-studied. Rainbow Hamilton cycles in edge colored complete graphs were first studied in Erdős, Nešetřil and Rödl [7]. Colorings were constrained by the number of times, k, that an individual color could be used. Such a coloring is called k-bounded. They showed that allowing k to be any constant, there was always a rainbow Hamilton cycle. Hahn and Thomassen [15] were the next people to consider this problem and they showed that k could grow as fast as $n^{1/3}$ and conjectured that the growth rate of k could in fact be linear. In unpublished work Rödl and Winkler [18] in 1984 improved this to $n^{1/2}$. Frieze and Reed [14] improved this to $k = \Omega(n/\log n)$ and finally Albert, Frieze and Reed [2] improved the bound on k to $\Omega(n)$. In another line of research, Cooper and Frieze [5] discussed the existence of rainbow Hamilton cycles in the random graph $G^{(q)}{}_{n,p}$ where each edge is independently and randomly given one of q colors. They showed that if $p \geq \frac{21\log n}{n}$ and $q \geq 21n$ then with high probability (w.h.p.) i.e. probability $1-o(1)$,

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, Research supported in part by NSF grant DMS-6721878, e-mail lespig@andrew.cmu.edu

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, Research supported in part by NSF grant DMS-6721878, e-mail alan@random.math.cmu.edu

‡School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Research supported in part by a USA-Israel BSF grant and by a grant from the Israel Science Foundation. e-mail krivelev@post.tau.ac.il

there is a rainbow colored Hamilton cycle. Frieze and Loh [12] improved this to $p \geq \frac{(1+o(1))\log}{n}$ and $q \geq n + o(n)$. Bal and Frieze [3] considered the case $q = n$ and showed that $p \geq \frac{K \log n}{n}$ suffices for large enough k. Ferber, Kronenberg, Mousset and Shikhelman [9] proved that if $p \gg \frac{\log n}{n}$ then w.h.p. $G_{n,p}$ contains cnp edge disjoint rainbow Hamilton cycles, for some constant $c > 0$.

In this paper we study the existence of other colorings of paths and cycles. Our first result does not at first sight fit into this framework. Let n be even and let M_0 be an arbitrary perfect matching of the complete graph K_n. Now consider the random graph process $G_m = ([n], E_m)$ where $E_m = \{e_1, e_2, \ldots, e_m\}$ is obtained from E_{m-1} by adding a random edge $e_m \notin E_{m-1}$, for $m = 0, 1, \ldots, N = \binom{n}{2}$.

Let

$$\tau_1 = \min \{m : \delta(G_m) \geq 1\}$$

where δ denotes minimum degree. Then let

$$\tau_H = \min \{m : G_m \text{ contains a Hamilton cycle } H \supseteq M_0\}.$$

Theorem 1 $\tau_1 = \tau_H$ *w.h.p.*[1]

In actual fact there are two slightly different versions. One where we insist that $M_0 \cap E_m = \emptyset$ and one where E_m is chosen completely independently of M_0. The theorem holds in both cases.

We note that Robinson and Wormald [17] considered a similar problem with respect to random regular graphs. They showed that one can choose $o(n^{1/2})$ edges at random, orient them and then w.h.p. there will be a Hamilton cycle containing these edges and following the orientations.

Theorem 1 has an easy corollary that fits our initial description. Let $G_m^{(r)}$ be an r-colored version of the graph process. This means that $G_m^{(r)}$ is obtained from $G_{m-1}^{(r)}$ by adding a random edge and then giving it a random color from $[r]$. Let $E_{m,i}$ denote the edges of color i for $i = 1, 2, \ldots, r$. When $r = 2$ denote the colors by *black* and *white* and let $E_{m,b} = E_{m,1}, E_{m,w} = E_{m,2}$. Then let $G_m^{(b)}$ be the subgraph of $G_m^{(2)}$ induced by the black edges and let $G_m^{(w)}$ induced by the white edges. Let

$$\tau_{1,1} = \min \left\{m : \delta(G_m^{(b)}), \delta(G_m^{(w)}) \geq 1\right\}$$

and let

$$\tau_{ZH} = \min \{m : G_m \text{ contains a zebraic Hamilton cycle.}\}.$$

Corollary 1 $\tau_{1,1} = \tau_{ZH}$ *w.h.p.*

Our next result is a zebraic analogue of *rainbow connection*. For a connected graph G, its rainbow connection $rc(G)$, is the minimum number r of colors needed for the following to hold: The edges of G can be r-colored so that every pair of vertices is connected by a rainbow path, i.e. a path in which no color is repeated. Recently, there has been interest in estimating this parameter for various classes of graph, including random graphs. By analogy, we say that a two-coloring of a connected graph provides a *zebraic connection* if there is a zebraic path joining every pair of vertices.

Theorem 2 *At time τ_1, a random black-white coloring of G_{τ_1} provides a zebraic connection, w.h.p.*

[1] A sequence of events $\mathcal{E}_n$ is said to occur *with high probability* (w.h.p.) if $\lim_{n\to\infty} \mathbf{Pr}(\mathcal{E}_n) = 1$

We consider now how we can extend our results to more than two colors. Suppose we have r colors $[r]$ and that $r \mid n$. We would like to consider the existence of Hamilton cycles where the ith edge has color $(i \mod r) + 1$. Call such a cycle *r-zebraic*. Our result for this case is not as tight as for the case of two colors. We are not able to prove a hitting time version. We will instead satisfy ourselves with a result for $G_{n,p}^{(r)}$. Let

$$p_r = \frac{r}{\alpha_r} \frac{\log n}{n}$$

where

$$\alpha_r = \left\lceil \frac{r}{2} \right\rceil.$$

Here and in the rest of the paper all logarithms will have base e unless explicitly stated otherwise.

Theorem 3 *Let $\varepsilon > 0$ be an arbitrary positive constant.*

$$\lim_{n\to\infty} \mathbf{Pr}(G_{n,p}^{(r)} \text{ contains an } r\text{-zebraic Hamilton cycle}) = \begin{cases} 0 & p \leq (1-\varepsilon)p_r \\ 1 & p \geq (1+\varepsilon)p_r \end{cases}.$$

The proofs of Theorems 1-3 will be given in Sections 4-6.

2 Notation

For a graph $G = (V, E)$ and $S, T \subseteq V$ we let $e_G(S)$ denote the number of edges contained in S, $e_G(S, T)$ denote the number of edges with one end in S and the other in T and $N_G(S)$ denote the set of neighbors of S that are not in S.

We will use certain values throughout our proofs. We list most of them here for easy reference. The reader is encouraged to skip reading this section and to just refer back as necessary.

$$t_0 = \frac{n}{2}(\log n - 2\log\log n) \text{ and } t_1 = \frac{n}{2}(\log n + 2\log\log n)$$

$$t_2 = \frac{t_0}{10} \text{ and } t_3 = \frac{t_0}{5} \text{ and } t_4 = \frac{9t_0}{10}.$$

$$\tau_i = t_i - t_{i-1} \text{ for } i = 3, 4.$$

$$p_i = \frac{t_i}{\binom{n}{2}},\ i = 0, 1, 2.$$

$$n_0 = \frac{n}{\log^2 n} \text{ and } n_0' = \frac{n_0}{20\log n} \text{ and } n_1 = \frac{n}{10\log n}.$$

$$n_b = \frac{n\log\log\log n}{\log\log n} \text{ and } n_c = \frac{200n}{\log n}.$$

$$L_0 = \frac{\log n}{100} \text{ and } L_1 = \frac{\log n}{\log\log n}.$$

$$\ell_0 = \frac{\log n}{200} \text{ and } \ell_1 = \frac{2\log n}{3\log\log n} \text{ and } \nu_L = \ell_0^{\ell_1} = n^{2/3+o(1)}.$$

We will also define some graphs and edge sets. For a graph $G = (V = V(G), E = E(G))$ and $v \in V$ we let $d_G(v)$ denote the degree of v in G. Also, for $S \subseteq V$ we let $N(S) = \{w \notin S : \exists v \in S, \{v, w\} \in E\}$.

The following graphs and sets of vertices are used.

$$\begin{aligned}
&\Psi_0 = G_{t-1}.\\
&V_0 = \{v \in [n] : d_{\Psi_0}(v) \leq L_0\}.\\
&\Psi_1 = \Psi_0 \cup \{e \in E_{t_1} \setminus E_{t_2} : e \cap V_0 \neq \emptyset\}.\\
&V_\lambda = \{v \in [n] :\ v \text{ is large}\}.\\
&V_\sigma = [n] \setminus V_\lambda.\\
&E_B = \{e \in E_{t_4} \setminus E_{t_3} : e \cap V_0 = \emptyset\}.\\
&V_\tau = \left\{v \in [n] \setminus V_0 : \deg_{E_B}(v) \leq L_0\right\}.
\end{aligned}$$

The definition of "large" depends on which theorem we are proving.

3 Probabilistic Inequalities

We will need standard estimates on the tails of various random variables.

Chernoff Bounds: Let $B(n,p)$ denote the binomial random variable where n is the number of trials and p is the probability of success.

$$\mathbf{Pr}(|B(n,p) - np| \geq \varepsilon np) \leq 2e^{-\varepsilon^2 np/3} \qquad \text{for } 0 \leq \varepsilon \leq 1. \tag{1}$$

$$\mathbf{Pr}(B(n,p) \geq anp) \leq \left(\frac{e}{a}\right)^{anp} \qquad \text{for } a > 0. \tag{2}$$

For proofs, see the appendix of Alon and Spencer [1].

McDiarmid's Inequality: Let $Z = Z(Y_1, Y_2, \ldots, Y_n)$ be a random variable where $Y_1, Y_2, \ldots, Y_n$ are independent for $i = 1, 2, \ldots, n$. Suppose that

$$|Z(Y_1, \ldots, Y_{i-1}, Y_i, Y_{i+1}, \ldots, Y_n) - Z(Y_1, \ldots, Y_{i-1}, \widehat{Y}_i, Y_{i+1}, \ldots, Y_n)| \leq c_i$$

for all $Y_1, Y_2, \ldots, Y_n, \widehat{Y}_i$ and $1 \leq i \leq n$. Then

$$\mathbf{Pr}(|Z - \mathbf{E}(Z)| \geq t) \leq \exp\left\{-\frac{t^2}{c_1^2 + c_2^2 + \cdots + c_n^2}\right\}. \tag{3}$$

For a proof see for example [1], [4], [10], or [16].

4 Proof of Theorem 1

4.1 Outline of proof

It is well known (see for example [4], [10], [16]) that w.h.p. we have $t_0 \leq \tau_1 \leq t_1$.

Our strategy for proving this is broadly in line with the 3-phase algorithm described in [6].

(a) We will take the first t_2 edges plus all the edges incident to vertices that have a low degree in G_{t_2}. We argue that w.h.p. this contains a perfect matching M_1 that is disjoint from M_0. The union of M_0, M_1 will then have $O(\log n)$ components w.h.p.

(b) $M_0 \cup M_1$ induces a 2-factor made up of alternating cycles. We then use about t_3 edges to make the minimum cycle size $\Omega(n/\log n)$.

(c) We then create a Hamilton cycle containing M_0, where we use the final $\approx t_2$ edges to close cycles in an second moment calculation.

We are working in a different model to that in [6] and there are many more conditioning probelms to be overcome.

4.2 Phase 1: Building M_1

We begin with $\Psi_0 = G_{t_2}$. Then let V_0 denote the set of vertices that have degree at most L_0 in Ψ_0. Now create $\Psi_1 = ([n], E_1)$ by adding those edges in $E_{t_1} \setminus E_{t_2}$ that are incident with V_0. We argue that w.h.p. Ψ_1 is a random graph with minimum degree one in which almost all vertices have degree $\Omega(\log n)$. Furthermore, we will show that w.h.p. Ψ_1 is an expander and then it will not be difficult to show that it contains the required perfect matching M_1.

Let a vertex be *large* if its degree in G_{t_1} is at least L_0 and *small* otherwise. Let V_λ denote the set of large vertices and let V_σ denote the set of small vertices.

The calculations for the next lemma will simplify if we observe the following: Suppose that $m = Np$. It is known that for any monotone increasing property of graphs

$$\mathbf{Pr}(G_m \in \mathcal{P}) \leq 3\,\mathbf{Pr}(G_{n,p} \in \mathcal{P}). \tag{4}$$

In general we have for not necessarily monotone properties:

$$\mathbf{Pr}(G_m \in \mathcal{P}) \leq 3m^{1/2}\,\mathbf{Pr}(G_{n,p} \in \mathcal{P}). \tag{5}$$

For proofs of (4), (5) see Bollobás [4] or Frieze and Karoński [10] or Janson, Łuczak and Ruciński [16].

The properties in the next lemma will be used to show that w.h.p. Ψ_1 is an expander. For technical reasons, we require the failure probabilities to be $O(n^{-0.51})$. Precisely, this is still $o(1)$ even after inflating by $n^{1/2+o(1)}$ and this will mean that the lower bound proved in (27) is large enough so that for any relevant event A we can use a crude estimate $\mathbf{Pr}(A \mid B) \leq \mathbf{Pr}(A)/\mathbf{Pr}(B)$ to handle conditioning on the event B describd in (27).

Lemma 2 *The following holds with probability* $1 - O(n^{-0.51})$*:*

(a) $|V_0| \leq n^{11/12}$.

(b) *If* $x, y \in V_\sigma$ *then the distance between them in* G_{t_1} *is at least 10.*

(c) *If* $S \subseteq [n]$ *and* $|S| \leq n_0$ *then* $e_{G_{t_1}}(S) \leq 10|S|$.

(d) *If* $S \subseteq [n]$ *and* $|S| = s \in [n_0', n_1]$ *then* $|N_{\Psi_0}(S)| \geq s\log n/25$.

(e) *No cycle of length 4 in* G_{t_1} *contains a small vertex.*

(f) *The maximum degree in* G_{t_1} *is less than* $10\log n$.

Proof (a) Suppose that the sequence $x_1, x_2, \ldots, x_{2t_2}$ is chosen randomly from $[n]^{2t_2}$ and we let Γ_{t_2} denote the multigraph with edge-set (x_{2i-1}, x_{2i}), $i = 1, 2, \ldots, t_2$. *After* we remove repeated edges and loops we can couple what remains with a subgraph H of G_{t_2}. Let Z_1 denote the number of loops and let Z_2 denote the number of repeated edges in Γ_{t_2}. Let V_0' denote the set of vertices of degree at most L_0 in Γ_{t_2}. Then $|V_0| \leq Z_1 + 2Z_2 + |V_0'|$. This is because if $v \in V_0 \setminus V_0'$ then it must lie in a loop or a multiple edge.

Now Z_1 is distributed as $\mathrm{Bin}(t_2, 1/n)$ and then the Chernoff bound (2) implies that

$$\mathbf{Pr}(Z_1 \geq \log^2 n) \leq e^{-\log^2 n}. \tag{6}$$

We are doing more than usual here, because we need probability $o(n^{-0.51})$, rather than just probabilty $o(1)$.

Now Z_2 is dominated by $\mathrm{Bin}(t_2, t_2/N)$ and then the Chernoff bound (2) implies that

$$\mathbf{Pr}(Z_2 \geq \log^3 n) \leq e^{-\log^3 n}. \tag{7}$$

Now,

$$\begin{aligned}
\mathbf{Pr}\left(v \in V_0'\right) &\leq \sum_{k=0}^{L_0} \binom{2t_2}{k} n^{-k} \left(1 - \frac{1}{n}\right)^{2t_2-k} \\
&\leq \; 2\binom{2t_2}{L_0} n^{-L_0} e^{-(2t_2-L_0)/n} \\
&\leq \; 2\left(\frac{2et_2}{nL_0}\right)^{L_0} n^{-1/10+o(1)} \\
&\leq n^{-1/11}.
\end{aligned}$$

It follows, that $\mathbf{E}(|V_0'|) \leq n^{10/11}$. We now use inequality (3) to finish the proof. Indeed, changing one of the x_i's can change $|V_0'|$ by at most one. Hence, for any $u > 0$,

$$\mathbf{Pr}(|V_0'| \geq \mathbf{E}(|V_0'|) + u) \leq \exp\left\{-\frac{2u^2}{t_2}\right\}.$$

Putting $u = n^{4/7}$ into the above and using (6), (7) finishes the proof of (a).

(b) We do not have room to apply (5) here. We need the inequality

$$\frac{\binom{N-a}{t-b}}{\binom{N}{t}} \leq \left(\frac{t}{N}\right)^b \left(\frac{N-t}{N-b}\right)^{a-b} \tag{8}$$

for $b \leq a \leq t \leq N$. Verification of (8) is straightforward and can be found for example in Chapter 21.1 of [10]. We will now and again use the notation $A \leq_b B$ in place of $A = O(B)$ when it suits our aesthetic taste.

$$\begin{aligned}
\mathbf{Pr}(\exists\ x,y) &\leq \sum_{k=2}^{11}\binom{n}{k}k!\sum_{\ell_1,\ell_2=0}^{L}\binom{n-k}{\ell_1}\binom{n-k}{\ell_2}\frac{\binom{N-(2n-k+1)}{t_1-k+1-\ell_1-\ell_2}}{\binom{N}{t_1}}\\
&\leq_b \sum_{k=2}^{11} n^k \sum_{\ell_1,\ell_2=0}^{L}\left(\frac{ne}{\ell_1}\right)^{\ell_1}\left(\frac{ne}{\ell_2}\right)^{\ell_2}\left(\frac{t_1}{N}\right)^{\ell_1+\ell_2+k-1}\left(\frac{N-t_1}{N-(\ell_1+\ell_2+k-1)}\right)^{2n-(\ell_1+\ell_2)}\\
&\leq_b n\sum_{k=2}^{11}\log^{k-1} n \sum_{\ell_1,\ell_2=0}^{L}\left(\frac{3\log n}{\ell_1}\right)^{\ell_1}\left(\frac{3\log n}{\ell_2}\right)^{\ell_2} n^{-2+o(1)}\\
&= o(n^{-0.51}).
\end{aligned}$$

(c) We can use (4) here. If $s = |S|$, then in G_{n,p_1} where $p_1 = t_1/N$,

$$\mathbf{Pr}(e_{G_{t_1}}(S) > 10|S|) \leq \binom{\binom{s}{2}}{10s}p_1^{10s} \leq \left(\frac{s^2 e}{10s}\cdot\frac{\log n + 2\log\log n}{n-1}\right)^{10s} \leq \left(\frac{s\log n}{n}\right)^{10s}.$$

So,

$$\mathbf{Pr}(\exists\ S) \leq \sum_{s=10}^{n_0}\binom{n}{s}\left(\frac{s\log n}{n}\right)^{10s} \leq \sum_{s=10}^{n_0}\left(\frac{ne}{s}\right)^{s}\left(\frac{s\log n}{n}\right)^{10s} = \sum_{s=10}^{n_0}\left(\left(\frac{s}{n}\right)^{9}\log^{10} n\right)^{s} = o(n^{-0.51}).$$

(d) We can use (4) here with $p_2 = t_2/N$. For $v \in V$, $\mathbf{Pr}(v \in N(S)) = 1-(1-p_2)^s \geq \frac{sp_2}{2}$ for $s \leq n_1$. So $|N(S)|$ stochastically dominates $\mathrm{Bin}(n-s, \frac{sp_2}{2})$. Now $(n-s)\frac{sp_2}{2} \sim \frac{s\log n}{20}$ and so using the Chernoff bound (1) with $\varepsilon \sim 1/5$,

$$\mathbf{Pr}(|N_{\Psi_0}(S)| < s\log n/25) \leq e^{-s\log n/1001}.$$

So,

$$\mathbf{Pr}(\exists\ S) \leq \sum_{s=n_0'}^{n_1}\binom{n}{s}e^{-s\log n/1001} \leq \sum_{s=n_0}^{n_1}\left(\frac{ne}{s}\cdot n^{-1/1001}\right)^{s} = o(n^{-0.51}).$$

(e) The expected number of such cycles is bounded by

$$\begin{aligned}
\binom{n}{4}\frac{3!}{2}\sum_{k=0}^{L_0}4\binom{n-4}{k}\frac{\binom{N-n-3}{t_1-4-k}}{\binom{N}{t_1}} &\leq n^4\sum_{k=0}^{L_0}\left(\frac{ne}{k}\right)^{k}\left(\frac{t_1}{N}\right)^{k+4}\left(\frac{N-t_1}{N-k-4}\right)^{n+k-1}\\
&\leq_b \log^4 n\sum_{k=0}^{L_0}\left(\frac{e^{1+o(1)}\log n}{k}\right)^{k} n^{-1+o(1)}\\
&= o(n^{-0.51}).
\end{aligned}$$

(f) We apply (4) and find that the probability of having a vertex of degree exceeding $10\log n$ is at most

$$3n\binom{n-1}{10\log n}\left(\frac{\log n + \log\log n}{n-1}\right)^{10\log n} \leq 3n\left(\frac{e^{1+o(1)}}{10}\right)^{10\log n} = o(n^{-0.51}).$$

$\square$

We will sometimes use (f) without comment in what follows.

Lemma 2 implies the following:

Lemma 3 *With probability* $1 - o(n^{-0.51})$,

$$S \subseteq [n] \text{ and } |S| \leq n/2000 \text{ implies } |N_{\Psi_1}(S)| \geq |S| \text{ in } \Psi_1, \tag{9}$$

Proof Assume that the conditions described in Lemma 2 hold. Let $N(S) = N_{\Psi_1}(S)$. We first argue that if $S \subseteq V_\lambda$ and $|S| \leq n/2000$ then

$$|N(S)| \geq 4|S|. \tag{10}$$

From the lemma, we only have to concern ourselves with $|S| \leq n_0'$ or $|S| \in [n_1, n/2000]$.

If $|S| \leq n_0'$ and $T = N(S)$ then in Ψ_1 we have, using Lemma 2(f),

$$e(S \cup T) \geq \frac{|S| \log n}{200} \text{ and } |S \cup T| \leq |S| \, (1 + 10 \log n) \leq n_0. \tag{11}$$

It is important to note that to obtain (11) we use the fact that vertices in $V_0 \setminus V_\sigma$ are given all their edges in Ψ_1.

Equation (11) and Lemma 2(c) imply that $\frac{|S| \log n}{200} \leq 10|S \cup T|$ and so (10) holds with room to spare.

If $|S| \in [n_1, n/2000]$ then we choose $S' \subseteq S$ where $|S'| = n_1$ and use

$$|N(S)| \geq |N(S')| - |S| \geq \frac{\log n}{25} \cdot \frac{200|S|}{\log n} - |S|.$$

This yields (10), again with room to spare.

Now let $S_0 = S \cap V_\sigma$ and $S_1 = S \setminus S_0$. Then we have

$$|N(S)| \geq |N(S_0)| + |N(S_1)| - |N(S_0) \cap S_1| - |N(S_1) \cap S_0| - |N(S_0) \cap N(S_1)|. \tag{12}$$

But $|N(S_0)| \geq |S_0|$. This follows from (i) Ψ_1 has no isolated vertices, and (ii) Lemma 2(b) means that S_0 is an independent set and no two vertices in S_0 have a common neighbor. Equation (10) implies that $|N(S_1)| \geq 4|S_1|$. We next observe that trivially, $|N(S_0) \cap S_1| \leq |S_1|$. Then we have $|N(S_1) \cap S_0| \leq |S_1|$, for otherwise some vertex in S_1 has two neighbors in S_0, contradicting Lemma 2(b.) Finally, we also have $|N(S_0) \cap N(S_1)| \leq |S_1|$. If for a vertex in S_1 there are two distinct paths of length two to S_0 then we violate one of the conditions Lemma 2(b) or (e).

So, from (12) we have

$$|N(S)| \geq |S_0| + 4|S_1| - |S_1| - |S_1| - |S_1| = |S|.$$

□

Next let $G = (V, E)$ be a graph with an even number of vertices that does not contain a perfect matching. Let v be a vertex not covered by some maximum matching and let

$$A_G(v) = \{w : \exists \text{ a maximum matching of } G \text{ that does not cover both } v \text{ and } w.\}$$

Lemma 4 *If $A = A_G(v)$ for some v, G, then $|N_G(A)| < |A|$.*

Proof Let $v \in V$, let M be a maximum matching that isolates v, and let $S_0 \neq \emptyset$ be the set of vertices, other than v, that are isolated by M. Let $S_1 \supseteq S_0$ be the set of vertices reachable from S_0 by a non-empty even length alternating path with respect to M. Let $x \in N_G(S_1)$ and let $y \in S_1$ be a neighbor of x. Then x is covered by M, as otherwise we can get a larger matching by using an alternating path from v to y, and then the edge $\{y, x\}$.

Let y_1 satisfy $(x, y_1) \in M$. We show that $y_1 \in S_1$ and this implies that $|N_G(S_1)| \leq |S_1|$ as M defines a mapping $x \to y_1$ of $N_G(S_1)$ into S_1. Let P be an even length alternating path from v terminating at y. If P contains (x, y_1) we can truncate it to terminate with (x, y_1), otherwise we can extend it using edges $\{y, x\}$ and (x, y_1).

Finally, observe that $A(v) = S_1 \cup \{v\}$. $\square$

Now consider the edge set

$$E_A = E_{t_3} \setminus E(\Psi_1) = \{f_1, f_2, \ldots, f_\rho\},$$

where with probability $1 - o(n^{-0.51})$ we have

$$\tau_3 \geq \rho \geq \tau_3 - 10n^{11/12} \log n \sim \frac{n \log n}{20}.$$

Lemma 5 *Given ρ, E_A is a random ρ-subset of $\binom{W}{2}$, where $W = [n] \setminus V_0$.*

Proof This follows from the fact that if we remove any f_i and replace it with any other edge from $\binom{[n]\setminus V_0}{2}$ then V_0 is unaffected. $\square$

Now consider the sequence of graphs $H_0 = \Psi_1, H_1, \ldots, H_\rho$ where H_i is obtained from H_{i-1} by adding the edge f_i. We claim that if μ_i denotes the size of a largest matching in H_i, then

$$\mathbf{Pr}(\mu_i \geq \mu_{i-1} + 1 \mid \mu_{i-1} < n/2,\ f_1, \ldots, f_{i-1},\ (\Psi_1 \text{ satisfies } (9))) \geq 10^{-7}. \tag{13}$$

To see this, let M_{i-1} be a matching of size μ_{i-1} in H_{i-1} and suppose that v is a vertex not covered by M_{i-1}. It follows from (9) and Lemma 4 that if $A_{H_{i-1}}(v) = \{g_1, g_2, \ldots g_r\}$ then $r \geq n/2000$. Now consider the pairs $\{g_j, x\}$, $j = 1, \ldots, r$, $x \in A_{H_{i-1}}(g_i)$. There are at least $\binom{n/2000}{2}$ such pairs and if f_i lies in this collection, then $\mu_i = \mu_{i-1} + 1$. Equation (13) follows from this and the fact that E_A is a random set. In fact, given the condition in Lemma 2(a) and a maximum degree of at most $10 \log n$ in G_{t_1}, the probability in question is at least

$$\frac{\binom{n/2000}{2} - 10n^{11/12} \log n - \rho}{\binom{n}{2}} > 10^{-7}.$$

It follows from (13) that

$$\mathbf{Pr}(H_\mu \text{ has no perfect matching}) \leq o(n^{-0.51}) + \mathbf{Pr}\left(\text{Bin}\left(\rho, 10^{-7}\right) \leq n/2\right) = o(n^{-0.51}). \tag{14}$$

So with probability $1 - o(n^{-0.51})$, $\Psi_2 = H_\rho$ has a perfect matching M_1.

Remark 6 M_1 is uniformly random, independently of M_0, and so the inclusion-exclusion formula gives

$$\mathbf{Pr}(M_0 \cap M_1 = \emptyset) = \sum_{i=0}^{n/2} (-1)^i \binom{n/2}{i} \frac{(n-2i)!}{(n/2-i)!2^{n/2-i}} \frac{2^{n/2}(n/2)!}{n!}. \tag{15}$$

Here we use the fact that there are $(2m)!/(m!2^m)$ perfect matchings in K_{2m}.

Now if u_i denotes the summand in (15) then we have $u_0 = 1$ and

$$\frac{|u_{i+1}|}{|u_i|} = \frac{1}{2(i+1)}\left(1 + \frac{1}{n-2i}\right).$$

So if $m = \Theta(\log n)$ say, then by the Bon-Ferroni inequalities,

$$\begin{aligned}\mathbf{Pr}(M_0 \cap M_1 = \emptyset) &\geq \sum_{i=0}^{2m+1} u_i \\ &= \sum_{i=0}^{2m+1} (-1)^i \frac{1}{2^i i!} + O\left(\frac{\log n}{n}\right) \\ &= e^{-1/2} + o(1).\end{aligned}$$

It follows that M_1 exists with probability $1 - o(n^{-0.51})$, even if we insist that it be disjoint from M_0. Indeed, conditioning on $M_0 \cap M_1 = \emptyset$ can only increase the probability of some "unlikely" event by a factor of at most $e^{1/2} + o(1)$.

We will need the following properties of the 2-factor $\Pi_0 = M_0 \cup M_1$.

Lemma 7 *The following hold with probability* $1 - o(n^{-0.51})$*:*

(a) *$M_0 \cup M_1$ has at most $10 \log_2 n$ components.*

(b) *There are at most n_b vertices on components of size at most n_c.*

Proof

(a) Following the argument in [13] we note tht if C is the cycle of $M_0 \cup M_1$ that contains vertex 1 then

$$\mathbf{Pr}(|C| = 2k) < \prod_{i=1}^{k-1}\left(\frac{n-2i}{n-2i+1}\right)\frac{1}{n-2k+1} < \frac{1}{n-2k+1}. \tag{16}$$

Indeed, consider M_0-edge $\{1 = i_1, i_2\} \in C$ containing vertex 1. Let $\{i_2, i_3\} \in C$ be the M_1-edge containing i_2. Then $\mathbf{Pr}(i_3 \neq 1) = \frac{n-2}{n-1}$. Assume $i_3 \neq 1$ and let $\{i_3, i_4 \neq 1\} \in C$ be the M_0 edge containing i_3. Let $\{i_4, i_5\} \in C$ be the M_1-edge containing i_4. Then $\mathbf{Pr}(i_5 \neq 1) = \frac{n-4}{n-3}$ and so on.

Having chosen C, the remaining cycles come from the union of two (random) matchings on the complete graph $K_{n-|C|}$. It follows from this, by summing (16) over $k \leq n/4$ that $\mathbf{Pr}(|C| < n/2) \leq 1/2$. Hence,

$$\mathbf{Pr}(\neg(a)) \leq \mathbf{Pr}(Bin(10\log_2 n, 1/2) \leq \log_2 n) = 2^{-10\log_2 n}\sum_{i=0}^{\log_2 n}\binom{10\log_2 n}{i} \leq 2^{-5\log_2 n} = o(n^{-0.51}).$$

(b) It follows from (16) that

$$\mathbf{Pr}(|C| \le n_c) \le \frac{201}{\log n}.$$

If we generate cycle sizes as in (a) then up until there are fewer than $n_b/2$ vertices left, $\log \nu \sim \log n$ where ν is the number of vertices that need to be partitioned into cycles. It follows that the probability we generate more than $k = \frac{\log\log\log n \times \log n}{1000 \log\log n}$ cycles of size at most n_c up to this time is bounded by

$$o(n^{-0.51}) + \mathbf{Pr}\left(Bin\left(10\log_2 n, \frac{201}{\log n}\right) \ge k\right) \le o(n^{-0.51}) + \left(\frac{3000e}{k}\right)^k = o(n^{-0.51}).$$

Thus with probability $1 - o(n^{-0.51})$, we have at most

$$\frac{n_b}{2} + kn_c \le n_b$$

vertices on cycles of length at most n_b. □

4.3 Phase 2: Increasing cycle size

In this section, we will use the edges in

$$E_B = \{e \in E_{t_4} \setminus E_{t_3} : e \cap V_0 = \emptyset\}$$

to create a 2-factor that contains M_0 and in which each cycle has size at least n_c. Note that

$$E_B \cap E(\Psi_1) = \emptyset.$$

We eliminate the small cycles (of length less than n_c) one by one (more or less). Let C be a small cycle. We remove an edge $\{u_0, v_0\} \notin M_0$ of C. We then try to join u_0, v_0 by a sufficiently long M_0 alternating path P that begins and ends with edges not in M_0. This is done in such a way that the resulting 2-factor contains M_0 but has at least one less small cycle. The search for P is done in a breadth first manner from both ends, creating $n^{2/3+o(1)}$ paths that begin at v_0 and another $n^{2/3+o(1)}$ paths that end at u_0. We then argue that with sufficient probability, we can find a pair of paths that can be joined by an edge from E_B to create the required alternating path.

We proceed to a detailed description. Let

$$V_\tau = \left\{v \in [n] \setminus V_0 : \deg_{E_B}(v) \le L_0\right\},$$

where for a set of edges X and a vertex x, $\deg_X(x)$ is the number of edges in X that are incident with x.

Lemma 8 *The following hold with probability* $1 - o(n^{-0.51})$*:*

(a) $|V_\tau| \le n^{2/5}$.

(b) *No vertex has 10 or more G_{t_1} neighbors in V_τ.*

(c) *If C is a cycle with $|C| \le n_c$ then $|C \cap V_\tau| \le |C|/200$ in G_{t_1}.*

Proof

(a) We follow a similar argument to that in Lemma 2(a). We condition on $|V_0| \le n^{11/12}$ and maximum degree $10 \log n$ in G_{t_0} and generate a random sequence from $[n - n^{11/12}]^{7t_0/10 - 10n^{11/12}\log n}$. The argument is now almost identical to that in Lemma 2(a).

(b) This time we can condition on $\nu = n - |V_0|$ and $\mu = |\{e \in E_{t_4} \setminus E_{t_3} : e \cap V_0 \neq \emptyset\}| \le n^{11/12} \times 10 \log n$. We write

$$\mathbf{Pr}(v \text{ violates (b)}) \le \sum_{S \in \binom{[n-1]}{10}} \mathbf{Pr}(\mathcal{A}(v,S))\, \mathbf{Pr}(\mathcal{B}(v,S) \mid \mathcal{A}(v,S))$$

where

$$\begin{aligned}
\mathcal{A}(v,S) &= \{N(v) \supseteq S, \text{ in } G_{t_1}\}, \\
\mathcal{B}(v,S) &= \{w \text{ has at most } L_0\ E_B\text{-neighbors in } [n] \setminus (S \cup \{v\}), \forall w \in S\}.
\end{aligned}$$

Applying (4) we see that $\mathbf{Pr}(\mathcal{A}(v,S)) \le 3\binom{n}{10} p_1^{10}$ and then using (4) with

$$p = \frac{t_4 - t_3 - \mu}{\binom{\nu}{2}} \sim \frac{7 \log n}{10n} \tag{17}$$

we see that

$$\mathbf{Pr}(\mathcal{B}(v,S) \mid \mathcal{A}(v,S)) \le 3 \left(\sum_{k=0}^{L_0} \binom{\nu - 11}{k} p^k (1-p)^{\nu - 11} \right)^{10}$$

and so

$$\begin{aligned}
\mathbf{Pr}(v \text{ violates (b)}) &\le_b \binom{n}{10} p_1^{10} \left(\sum_{k=0}^{L_0} \binom{\nu - 11}{k} p^k (1-p)^{\nu - 11} \right)^{10} \\
&\le (e^{o(1)} \log n \cdot n^{1/20 - 7/10 + o(1)})^{10} \\
&= o(n^{-6}).
\end{aligned}$$

Now use the Markov inequality.

(c) Let Z denote the number of cycles violating the requied property. Using (4) and ν as in (b) and p as in (17), we have

$$\begin{aligned}
\mathbf{E}(Z) &\le_b \sum_{k=3}^{n_c} \binom{n}{k} k! p_1^k \binom{k}{\lceil \frac{k}{200} \rceil} \left(\sum_{\ell=0}^{L_0} \binom{\nu - k}{\ell} p^\ell (1-p)^{\nu - \ell} \right)^{\lceil k/200 \rceil} \\
&\le \sum_{k=3}^{n_c} (2n)^k \left(\frac{\log n + \log\log n}{n-1} \right)^k n^{-3\lceil k/200 \rceil / 5} \\
&= o(n^{-0.51}).
\end{aligned}$$

$\square$

Now consider the distribution of the edges in E_B.

Lemma 9 *Let $V_1 = [n] \setminus V_0$ and $A \subseteq \binom{V_1}{2}$ with $|A| = a = O(\log n)$. Let X be a subset of E_B that is disjoint from A. Suppose that $|X| = O(n^{11/12} \log n)$. Then*

$$\begin{aligned}
&\mathbf{Pr}(E_B \supseteq A \mid |E_B| = \mu = \alpha n \log n, |V_1| = \nu \geq n - n^{11/12}, E_B \supseteq X) \\
&= \frac{\binom{\binom{\nu}{2}-a-|X|}{\mu-a-|X|}}{\binom{\binom{\nu}{2}-|X|}{\mu-|X|}} && (18) \\
&= (1 + o(n^{-1/13})) \left(\frac{2\alpha \log n}{n}\right)^a. && (19)
\end{aligned}$$

Proof Equation (18) follows from Lemma 5. For equation (19), we write

$$\frac{\binom{\binom{\nu}{2}-a-|X|}{\mu-a-|X|}}{\binom{\binom{\nu}{2}-|X|}{\mu-|X|}} = \left(\frac{\mu - |X|}{\binom{\nu}{2} - |X|}\right)^a \left(1 + O\left(\frac{a}{\mu - |X|}\right)\right)^a =$$

$$\left(\frac{\mu}{\binom{\nu}{2}}\right)^a \left(1 + O\left(\frac{a}{\mu - |X|}\right) + O\left(\frac{|X|}{\mu}\right)\right)^a.$$

the fact that in general, if $s^2 = o(N)$ then

$$\frac{\binom{N-s}{M-s}}{\binom{N}{M}} = \left(\frac{M}{N}\right)^s \left(1 + O\left(\frac{s^2}{N}\right)\right).$$

□

By construction, we can apply this lemma to the graph induced by E_B with

$$\alpha = \frac{7 + o(1)}{20}.$$

Let a cycle C of Π_0 be *small* if its length $|C| < n_c$ and *large* otherwise. Define a near 2-factor to be a graph that is obtained from a 2-factor by removing one edge. A near 2-factor Γ consists of a path $P(\Gamma)$ and a collection of vertex disjoint cycles. A 2-factor or a near 2-factor are *proper* if they contain M_0. We abbreviate proper near 2-factor to PN2F.

We will describe a process for eliminating small cycles from Ψ_0. In this process we create intermediate proper 2-factors. Let Γ_0 be a 2-factor and suppose that it contains a small cycle C. To begin the elimination of C we choose an arbitrary edge $\{u_0, v_0\}$ in $C \setminus M_0$, where $u_0, v_0 \notin V_\tau$. This is always possible, see Lemma 8(c). We delete it, obtaining a PN2F Γ_1. Here, $P(\Gamma_1) \in \mathcal{P}(v_0, u_0)$, the set of M_0-alternating paths in G from v_0 to u_0. Here an M_0-alternating path must begin and end with an edge of M_0. The initial goal will be to create a large set of PN2Fs such that each Γ in this set has path $P(\Gamma)$ of length at least n_c and the small cycles of Γ are a strict subset of the small cycles of Γ_0. Then we will show that with probability $1 - o(n^{-0.51})$, the endpoints of one of the paths in some such Γ can be joined by an edge to create a proper 2-factor with at least one fewer small cycle than Π.

This process can be divided into two stages. In a generic step of Stage I, we take a PN2F Γ as above with $P(\Gamma) \in \mathcal{P}(u_0, v)$ and construct a new PN2F with the same starting point u_0 for its path. We do this by considering edges from E_B incident to v. Suppose $\{v, w\} \in E_B$ and that the non-M_0 edge in Γ containing vertex w is $\{w, x\}$. Then $\Gamma' = \Gamma \cup \{v, w\} \setminus \{w, x\}$ is a PN2F with $P(\Gamma') \in \mathcal{P}(u_0, x)$. We say that $\{v, w\}$ is *acceptable* if $x, w \notin W$ (W defined immediately below) and $P(\Gamma')$ has length at least n_c and any new cycle created (in Γ' but not Γ) has at least n_c edges.

There is an unlikely technicality to be faced. If Γ has no non-M_0 edge (x, w), then $w = u_0$ and this is accepted if $P(\Gamma')$ has at least n_c edges. This would prematurely end an iteration. The probability that we close a cycle at such a step is $O(1/n)$ and so we can safely ignore this possibility.

In addition we define a set W of *used* vertices, where

$$W = V_\sigma \cup V_\tau \text{ at the beginning of Phase 2}$$

and whenever we look at edges $\{v, w\}, \{w, x\}$ (that is, consider using that edge to create a new Γ'), we add v, w, x to W. Additionally, we maintain $|W| = O(n^{11/12})$, or fail if we cannot.

We will build a tree T of PN2Fs, breadth-first, where each non-leaf vertex Γ yields PN2F children Γ' as above. We stop building T when we have $\nu_L = n^{2/3+o(1)}$ leaves. This will end Stage 1 for the current cycle C being removed.

We'll restrict the set of PN2F's which could be children of Γ in T_0 as follows: We restrict our attention to $w \notin W$ with $\{v, w\} \in E_B$ and $\{v, w\}$ acceptable as defined above. Also, we only construct children from the first $\ell_0 = L_0/2$ acceptable $\{v, w\}$'s at a vertex v. Furthermore we only build the tree down to $\ell_1 = \frac{2 \log n}{3 \log \log n}$ levels. We denote the nodes in the ith level of the tree by S_i. Thus $S_0 = \{\Gamma_1\}$ and S_{i+1} consists of the PN2F's that are obtained from S_i using acceptable edges. In this way we define a tree of PN2F's with root Γ_1 that has branching factor at most ℓ_0. Thus,

$$|S_{\ell_1}| \leq \nu_L = \ell_0^{\ell_1}. \tag{20}$$

On the other hand, if we let $\mathcal{E}_0$ denote the intersection of the high probability events of Lemmas 2, 7 and 8, then:

Lemma 10 *Conditional on the event $\mathcal{E}_0$,*

$$|S_{\ell_1}| = \nu_L$$

with probability $1 - o(n^{-3})$.

Proof If $P(\Gamma)$ has endpoints u_0, v and $e = \{v, w\} \in E_B$ and e is unacceptable then i) w lies on $P(\Gamma)$ and is too close to an endpoint or (ii) $x \in W$ or $w \in W$ or (iii) w lies on a small cycle. Ab initio, there are at least L_0 choices for w and we must bound the number of unacceptable choices.

The probability that at least $L_0/10$ vertices are unacceptable due to (iii) is by Lemmas 7 and 9 at most

$$(1 + o(1)) \binom{n_b}{L_0/10} \left(\frac{7 \log n}{(10 + o(1))n} \right)^{L_0/10} \leq \left(\frac{10 e n_b \log n}{L_0 n} \right)^{L_0/10}$$
$$\leq \left(\frac{1000 e \log \log \log n}{\log \log n} \right)^{L_0/10} = O(n^{-K}) \tag{21}$$

for any constant $K > 0$.

A similar argument deals with conditions (i) and (ii).

Thus, with (conditional) probability $1 - o(n^{-4})$, there are at least

$$\left(\frac{\log n}{100} - \frac{3\log n}{2000}\right)|S_t| \geq \frac{\log n}{200}|S_t|$$

acceptable edges, for all t. So, with (conditional) probability $1 - o(n^{-3})$ we have

$$|S_{\ell_1}| = \nu_L \tag{22}$$

as desired. $\square$

Having built T, if we have not already made a cycle, we have a tree of PN2Fs and the last level, ℓ_0 has leaves $\Gamma_i,\ i = 1, ..., \nu_L$, each with a path $P(\Gamma_i)$ of length at least n_c. Now, perform a second stage which will be like executing ν_L-many *Stage 1*'s *in parallel* by constructing trees $T_i,\ i = 1, ..., \nu_L$, where the root of T_i is Γ_i. Suppose for each i, $P(\Gamma_i) \in \mathcal{P}(u_0, v_i)$; we fix the vertex v_i and build paths by first looking at neighbors of u_0, for all i (so in tree T_i, very Γ will have path $P(\Gamma) \in \mathcal{P}(u, v_i)$ for some u).

Construct these ν_L trees in the Stage 2 by only enforcing the conditions that $w \notin W$. This change will allow the PN2Fs to have small paths and cycles. We will not impose a bound on the branching factor either. As a result of this and the fact that each tree T_i begins by considering edges from E_B adjacent to u_0, the sets of endpoints of paths (that are not the $v_i s$) of PN2Fs at the same level are the same in each of the trees $T_i, i = 1, 2, \ldots, \nu_L$. That is, if Γ'_i is a node at level ℓ of tree T_i and Γ'_j is a node at level ℓ of tree T_j, $P(\Gamma'_i) \in \mathcal{P}(w, v_i)$ and $P(\Gamma'_j) \in \mathcal{P}(w, v_j)$ for some $w \in V_0$. This can be proved by induction, see [5]. Indeed, let $L_{i,\ell}$ denote the set of end vertices, other than v_i, of the paths associated with the nodes at depth ℓ of the tree T_i, $i = 1, 2 \ldots, \nu_\ell, \ell = 0, 1, \ldots, \ell_1$. Thus $L_{i,0} = \{u_0\}$ for all i. We can see inductively that $L_{i,\ell} = L_{j,\ell}$ for all i, j, ℓ. In fact if $v \in L_{i,\ell} = L_{j,\ell}$ then $\{v, w\} \in E_B$ is acceptable for some i means that $w \notin W$ (at the start of the construction of level $\ell + 1$ and hence if $\{w, x\}$ is the non-M_0 edge for this i then $x \notin W$ and it is the non-M_0 edge for all j. In which case $\{v, w\}$ is acceptable for all i and we have $L_{i,\ell+1} = L_{1,\ell+1}$.

The trees $T_i, i = 1, ..., \nu_L$, will be succesfully constructed with probability $1 - o(1/n^3)$ and with a similar probability the number of nodes in each tree is at most $(10 \log n)^{\ell_1} = n^{2/3+o(1)}$. Here we use the fact that the maximum degree in $G_{t_1} \leq 10 \log n$ with this probability. However, some of the trees may not follow all of the conditions listed initially, and so we will "prune" the trees by disallowing any node Γ that was constructed in violation of any of those conditions. Call tree T_i GOOD if it still has at least L_0 leaves remaining after pruning and BAD otherwise. Notice that

$$\mathbf{Pr}(\exists\ i : T_i \text{ is BAD}) = o\left(\frac{\nu_L}{n^3}\right) = o(n^{-2}).$$

Finally, consider the probability that there is no E_B edge from any of the $n^{2/3+o(1)}$ endpoints found in Stage 1 to any of the $n^{2/3-o(1)}$ endpoints found in Stage 2. At this point we will have only exposed the edges of Π_0 incident with these endpoints. So if for some $k \leq \nu_L$ we examine the (at least) $\log n/200$ edges incident to $v_1, v_2, \ldots, v_k$ but not W then the probability we fail to close a cycle and produce a proper 2-factor is at most

$$\left(1 - \frac{1}{n^{1/3+o(1)}}\right)^{k \log n/200}.$$

Thus taking $k = n^{1/3+o(1)}$ suffices to make the failure probability $o(n^{-2})$. Also, this final part of the construction only contributes $n^{1/3+o(1)}$ to W.

Therefore, the probability that we fail to eliminate a particular small cycle C is $o(n^{-2})$ and then given $\mathcal{E}_0$, the probability that Phase 2 fails is $o(\log n/n^2) = o(1)$.

We should check now that w.h.p. $|W| = O(n^{11/12})$ throughout Phase 2. It starts out with at most $n^{11/12} + n^{2/5}$ vertices (see Lemmas 2(a) and 8(a)) and we add $O(n^{2/3+o(1)} \times \log n)$ vertices altogether in this phase.

Lemma 11 *The probability that Phase 2 fails to produce a proper 2-factor with minimum cycle length at least n_c is $o(n^{-0.51})$.*

□

4.4 Phase 3: Creating a Hamilton cycle

By the end of Phase 2, we will with probability $1 - o(n^{-0.51})$ have found a proper 2-factor with all cycles of length at least n_c. Call this subgraph Π^*.

In this section, we will use the edges in

$$E_C = \{e \in E_{t_0} \setminus (E_{t_4} \cup E(\Psi_1)) : e \cap V_0 = \emptyset\}$$

to turn Π^* into a Hamilton cycle that contains M_0, w.h.p. It is basically a second moment calculation with a twist to keep the variance under control. We note that Lemma 9 continues to hold if we replace E_B by E_C.

Arbitrarily assign an orientation to each cycle. Let $C_1, ..., C_k$ be the cycles of Π^* (note that if $k = 1$ we are done) and let $c_i = |C_i \setminus W|/2$. Then $c_i \geq \frac{n_c}{2} - O(n^{11/12}) \geq \frac{99n}{\log n}$ for all i. Let $a = \frac{n}{\log n}$ and $m_i = 2\lfloor \frac{c_i}{a} \rfloor + 1$ for all i and $m = \sum_{i=1}^k m_i$. From each C_i, we will consider choosing m_i vertices $v \in C_i \setminus W$ that are heads of non-M_0 arcs after the arbitrary ordering of all cycles, deleting these m arcs and replacing them with m others to create a proper Hamilton cycle.

Given such a deletion of edges, re-label the broken arcs as $(v_i, u_i), i \in [m]$ as follows: in cycle C_i identify the lowest numbered vertex $x_i \in [n]$ which loses a cycle edge directed out of it. Put $v_1 = x_1$ and then go round C_1 defining $v_2, v_3, \ldots v_{m_1}$ in order. Then let $v_{m_1+1} = x_2$ and so on. We thus have m path sections $P_j \in \mathcal{P}(u_{\phi(j)}, v_j)$ in Π^* for some permutation ϕ.

It is our intention to rejoin these path sections of Π^* to make a Hamilton cycle using E_C, if we can. Suppose we can. This defines a permutation ρ on $[m]$ where $\rho(i) = j$ if P_i is joined to P_j by $(v_i, u_{\phi(j)})$, where $\rho \in H_m$, the set of cyclic permutations on $[m]$. We will use the second moment method to show that a suitable ρ exists w.h.p. A technical problem forces a restriction on our choices for ρ. This will produce a variance reduction in a second moment calculation.

Given ρ define $\lambda = \phi\rho$. In our analysis we will restrict our attention to $\rho \in R_\phi = \{\rho \in H_m : \phi\rho \in H_m\}$. If $\rho \in R_\phi$ then we have not only constructed a Hamilton cycle in $\Pi^* \cup E_C$, but also in the *auxillary digraph* Λ, whose edges are $(i, \lambda(i))$.

The following lemma is from [6]. The content is in the lower bound. It shows that there are still many choices for ρ and it is needed to show that the expected number of possible re-arrangements of path sections, grows with n.

Lemma 12 $(m-2)! \leq |R_\phi| \leq (m-1)!$

Let H be the graph induced by the union of Π^* and E_C.

Lemma 13 *H contains a Hamilton cycle w.h.p.*

Proof Let X be the number of Hamilton cycles in G that can be obtained by removing the edges described above and rearranging the path segments generated by ϕ according to those in $\rho \in R_\phi$ and connecting the path segments using edges in H.

We will use the inequality $\mathbf{Pr}(X > 0) \geq \frac{\mathbb{E}(X)^2}{\mathbb{E}(X^2)}$ to show that such a Hamilton cycle exists with the required probability.

The definition of m_i gives us $\frac{2n}{a} - k \leq m \leq \frac{2n}{a} + k$ and so $1.99 \log n \leq m \leq 2.01 \log n$. Additionally we will use $k \leq \frac{n}{n_c} = \frac{\log n}{200}$, $m_i \geq 199$ and $\frac{c_i}{m_i} \geq \frac{a}{2.01}$ for all i.

From Lemmas 9 and 12, with $\alpha = 1/(10 + o(1))$

$$\mathbb{E}(X) \geq (1 - o(1)) \left(\frac{2\alpha \log n}{n}\right)^m (m-2)! \prod_{i=1}^{k} \binom{c_i}{m_i} \tag{23}$$

$$\geq \frac{1 - o(1)}{m^{3/2}} \left(\frac{2m\alpha \log n}{en}\right)^m \prod_{i=1}^{k} \left(\left(\frac{c_i e^{1-1/10m_i}}{m_i^{1+(1/2m_i)}}\right)^{m_i} \left(\frac{1 - 2m_i^2/c_i}{\sqrt{2\pi}}\right)\right) \tag{24}$$

where to go from (23) to (24) we have used the approximation $(m-2)! \geq m^{-3/2}(m/e)^m$ and

$$\binom{c_i}{m_i} \geq \frac{c_i^{m_i}(1 - 2m_i^2/c_i)}{m_i!} \text{ and } m_i! \leq \sqrt{2\pi m_i} \left(\frac{m_i}{e}\right)^{m_i} e^{1/10m_i}.$$

Explanation of (23)**:** We choose the arcs to delete in $\prod_{i=1}^{k} \binom{c_i}{m_i}$ ways and put them together as explained prior to Lemma 12 in at least $(m-2)!$ ways. The probability that the required edges exist in E_C is $(1 + o(1)) \left(\frac{2\alpha \log n}{n}\right)^m$, from Lemma 9.

Continuing, we have

$$\begin{aligned}
\mathbb{E}(X) &\geq \frac{(1-o(1))(2\pi)^{-m/398}e^{-k/10}}{m^{3/2}}\left(\frac{2m\alpha\log n}{en}\right)^m\prod_{i=1}^k\left(\frac{c_ie}{(1.02)m_i}\right)^{m_i}\\
&\geq \frac{(1-o(1))(2\pi)^{-m/398}}{n^{1/2000}m^{3/2}}\left(\frac{2m\alpha\log n}{en}\right)^m\left(\frac{ea}{2.01\times 1.02}\right)^m\\
&\geq \frac{(1-o(1))(2\pi)^{-m/398}}{n^{1/2000}m^{3/2}}\left(\frac{\log n}{6}\right)^m\\
&\to\infty. \qquad (25)
\end{aligned}$$

Let M, M' be two sets of selected edges which have been deleted in Π^* and whose path sections have been re-arranged into Hamilton cycles according to ρ, ρ' respectively. Let N, N' be the corresponding sets of edges which have been added to make the Hamilton cycles. Let Ω denote the set of choices for M (and M'.)

Let $s = |M\cap M'|$ and $t = |N\cap N'|$. Now $t\leq s$ since if $(v,u)\in N\cap N'$ then there must be a unique $(\tilde{v},u)\in M\cap M'$ which is the unique Π^*-edge into u. It is shown in [6] that $t=s$ implies $t=s=m$ and $(M,\rho)=(M',\rho')$. (This removes a large term from the second moment calculation). Indeed, Suppose then that $t=s$ and $(v_i,u_i)\in M\cap M'$. Now the edge $(v_i,u_{\lambda(i)})\in N$ and since $t=s$ this edge must also be in N'. But this implies that $(v_{\lambda(i)},u_{\lambda(i)})\in M'$ and hence in $M\cap M'$. Repeating the argument we see that $(v_{\lambda^k(i)},u_{\lambda^k(i)})\in M\cap M'$ for all $k\geq 0$. But λ is cyclic and so our claim follows.

If $\langle s,t\rangle$ denotes the case where $s=|M\cap M'|$ and $t=|N\cap N'|$, then

$$\begin{aligned}
\mathbb{E}(X^2) &\leq \mathbb{E}(X) + (1+o(1)\sum_{M\in\Omega}\left(\frac{2\alpha\log n}{n}\right)^m\sum_{\substack{M'\in\Omega\\ N'\cap N=\emptyset}}\left(\frac{2\alpha\log n}{n}\right)^m\\
&\quad + (1+o(1))\sum_{M\in\Omega}\left(\frac{2\alpha\log n}{n}\right)^m\sum_{s=2}^m\sum_{t=1}^{s-1}\sum_{\substack{M'\in\Omega\\ \langle s,t\rangle}}\left(\frac{2\alpha\log n}{n}\right)^{m-t}\\
&= \mathbb{E}(X)+E_1+E_2 \text{ say.}
\end{aligned}$$

Note that $E_1\leq(1+o(1))\mathbb{E}(X)^2$.

Now, with σ_i denoting the number of common $M\cap M'$ edges selected from C_i,

$$E_2\leq E(X)^2\sum_{s=2}^m\sum_{t=1}^{s-1}\binom{s}{t}\left[\sum_{\sigma_1+\ldots+\sigma_k=s}\prod_{i=1}^k\frac{\binom{m_i}{\sigma_i}\binom{c_i-m_i}{m_i-\sigma_i}}{\binom{c_i}{m_i}}\right]\frac{(m-t-1)!}{(m-2)!}\left(\frac{n}{2\alpha\log n}\right)^t.$$

Some explanation: There are $\binom{s}{t}$ choices for $N\cap N'$, given s and t. Given σ_i there are $\binom{m_i}{\sigma_i}$ ways to choose $M\cap M'$ and $\binom{c_i-m_i}{m_i-\sigma_i}$ ways to choose the rest of $M'\cap C_i$. After deleting M' and adding $N\cap N'$ there are at most $(m-t-1)!$ ways of putting the segments together to make a Hamilton cycle.

We see that

$$\frac{\binom{c_i-m_i}{m_i-\sigma_i}}{\binom{c_i}{m_i}}\leq\frac{\binom{c_i}{m_i-\sigma_i}}{\binom{c_i}{m_i}}=\frac{m_i(m_i-1)\cdots(m_i-\sigma_i+1)}{(c_i-m_i+1)\cdots(c_i-m_i+\sigma_i)}\leq(1+o(1))\left(\frac{2.01}{a}\right)^{\sigma_i}\exp\left\{-\frac{\sigma_i(\sigma_i-1)}{2m_i}\right\}.$$

Also,

$$\sum_{i=1}^{k} \frac{\sigma_i^2}{2m_i} \geq \frac{s^2}{2m} \text{ for } \sigma_1 + ... + \sigma_k = s$$

and

$$\sum_{i=1}^{k} \frac{\sigma_i}{2m_i} \leq \frac{k}{2} \text{ and } \sum_{\sigma_1+...+\sigma_k=s} \prod_{i=1}^{k} \binom{m_i}{\sigma_i} = \binom{m}{s}.$$

Using these approximations, we have

$$\sum_{\sigma_1+...+\sigma_k=s} \prod_{i=1}^{k} \frac{\binom{m_i}{\sigma_i}\binom{c_i-m_i}{m_i-\sigma_i}}{\binom{c_i}{m_i}} \leq (1+o(1))e^{k/2} \exp\left\{-\frac{s^2}{2m}\right\} \left(\frac{2.01}{a}\right)^s \binom{m}{s}.$$

So we can write

$$\frac{E_2}{\mathbb{E}(X)^2} \leq (1+o(1))e^{k/2} \sum_{s=2}^{m} \sum_{t=1}^{s-1} \binom{s}{t} \exp\left\{-\frac{s^2}{2m}\right\} \left(\frac{2.01}{a}\right)^s \binom{m}{s} \frac{(m-t-1)!}{(m-2)!} \left(\frac{n}{2\alpha \log n}\right)^t.$$

We approximate

$$\binom{m}{s} \frac{(m-t-1)!}{(m-2)!} \leq C_1 \frac{m^s}{s!} \left(\frac{m-t-1}{e}\right)^{m-t-1} \left(\frac{e}{m-2}\right)^{m-2} \leq C_2 \frac{m^s}{s!} \frac{e^t}{m^{t-1}},$$

for some constants $C_1, C_2 > 0$.

Substituting this in, we obtain,

$$\begin{aligned}
\frac{E_2}{\mathbb{E}(X)^2} &\leq_b n^{1/400} m \sum_{s=2}^{m} \left(\frac{2.01}{a}\right)^s \frac{m^s}{s!} \exp\left\{-\frac{s^2}{2m}\right\} \sum_{t=1}^{s-1} \binom{s}{t} \left(\frac{en}{2\alpha m \log n}\right)^t \\
&\leq (1+o(1) \left(\frac{m^2}{5en^{.99}}\right) \sum_{s=2}^{m} \left(\frac{(2.01)en \exp\{-s/2m\}}{2\alpha a \log n}\right)^s \frac{1}{s!} \\
&\leq n^{-9/10}.
\end{aligned}$$

To see this, notice that

$$\sum_{t=1}^{s-1} \binom{s}{t} \left(\frac{en}{2\alpha m \log n}\right)^t \leq m \left(\frac{en}{2\alpha m \log n}\right)^{s-1}$$

and

$$\sum_{s=2}^{m} \left(\frac{(2.01)en \exp\{-s/2m\}}{2\alpha a \log n}\right)^s \frac{1}{s!} \leq \sum_{s=2}^{m} \frac{30^s}{s!} \leq e^{30}.$$

Combining things, we get

$$\begin{aligned}
&\mathbb{E}(X^2) \leq \mathbb{E}(X) + \mathbb{E}(X)^2(1+o(1)) + \mathbb{E}(X)^2 n^{-.9} \quad \text{so} \\
&\frac{(\mathbb{E}X)^2}{\mathbb{E}(X^2)} \geq \frac{1}{\frac{1}{\mathbb{E}X} + 1 + o(1) + n^{-.9}} \\
&\qquad\quad \longrightarrow 1
\end{aligned}$$

as $n \to \infty$, as desired. $\square$

4.5 Proof of Corollary 1

We begin the proof by replacing the sequence $E_0, E_1, \ldots, E_m, \ldots$ by $E_0', E_1', \ldots, E_m', \ldots,$ where the edges of $E_m' = \{e_1', e_2', \ldots, e_m'\}$ are randomly chosen *with replacement.* This means that e_m is allowed to be a member of E_{m-1}'. We let G_m' be the graph $([n], E_m')$.

If an edge appears a second time, it will be randomly re-colored. We let R denote the set of edges that get repeated. Note that if $\tau_{1,1} = \mu$ and $e_\mu = (v,w) \in R$ then v or w is isolated in $G_m'^{(b)}$ or $G_m'^{(w)}$.

$$\mathbf{Pr}(e_{\tau_{1,1}} \in R) \leq 4\,\mathbf{Pr}(\exists e = (v,w) \in R : v \text{ has black degree } 1) = o(1). \tag{26}$$

Explanation: The factor 4 comes from v or w has black or white degree one. Next suppose first that $e_\mu = (v,w)$ and that v has black degree zero in $G_{\mu-1}$ and w also has zero black degree in $G_{\mu-1}$. An argument similar to that for Lemma 2(b) shows that w.h.p. there is no white edge joining v and w and so $e_\mu \notin R$.

Now suppose that $e_\mu = (v,w)$ and that v has black degree zero in $G_{\mu-1}$ and w has positive black degree in $G_{\mu-1}$. An argument similar to that given for Lemma 2(f) shows that w.h.p. the maximum white degree in G_m' is $O(\log n)$. There are $n-1$ choices for w, of which $O(\log n)$ put e_μ into R. So e_μ has an $O(\log n/n)$ chance of being in R.

At time $m = \tau_{1,1}$ the graphs $G_m^{(b)'}, G_m^{(w)'}$ will w.h.p. contain perfect matchings, see [8]. That paper does not allow repeated edges, but removing them enables one to use the result claimed. We choose random perfect matchings M_0, M_1 from $G_{\tau_{1,1}}^{(b)'}$, $G_{\tau_{1,1}}^{(w)'}$.

We couple the sequence $G_1, G_2, \ldots,$ with the sequence $G_1', G_2', , \ldots,$ by ignoring repeated edges in the latter. Thus $G_1', G_2', \ldots, G_m'$ is coupled with a sequence $G_1, G_2, \ldots, G_{m'}$ where $m' \leq m$. It follows from (26) that w.h.p. the coupled processes stop with the same edge. Furthermore, they stop with two independent matchings M_0, M_1. We can then begin analysing Phase 2 and Phase 3 within this context.

We will prove that

$$\mathbf{Pr}(M_1 \cap R = \emptyset) \geq n^{-1/2-o(1)}. \tag{27}$$

Corollary 1 follows from this. Indeed, it follows from (27) and the fact that Phases 1 and 2 succeed with probability $1 - O(n^{-0.51})$ that they succeed w.h.p. conditional on $M_1 \cap R = \emptyset$.

Phase 3 succeeds w.h.p. even if we avoid using edges in R. We have already carried out calculations with an arbitrary set of $O(n^{11/12})$ edges that must be avoided. The size of R is dominated by a binomial $Bin(O(n \log n), O(n^{-1} \log n)$ and so $|R| = O(\log^2 n)$ w.h.p. So avoidng R does not change any calculation in any significant way. In other words, we can w.h.p. find a zebraic Hamilton cycle in G_m'.

Finally note that the Hamilton cycle we obtain is zebraic.

Proof of (27)**:** R is a random set and it is independent of M_1. Let t_B be the number of black edges then

$$\mathbf{Pr}(M_1 \cap R = \emptyset \mid t_B) \geq \left(1 - \frac{n/2}{N}\right)^{t_B - n/2} \geq \exp\left\{-t_B\left(\frac{1}{n} + \frac{1}{n^2}\right)\right\}.$$

To remove the conditioning, we take expectations and then by convexity

$$\mathbf{E}\left(\exp\left\{-t_B\left(\frac{1}{n}+\frac{1}{n^2}\right)\right\}\right) \geq \left(\exp\left\{-\mathbf{E}(t_B)\left(\frac{1}{n}+\frac{1}{n^2}\right)\right\}\right) \geq n^{-1/2-o(1)}$$

since $\mathbf{E}(t_B) \sim \frac{1}{2}n\log n$. This proves (27).

5 Proof of Theorem 2

For a vertex $v \in [n]$ we let its *black* degree $d_b(v)$ be the number of black edges incident with v in G_{t_0}. We define its *white* degree $d_w(v)$ analogously. Let a vertex be *large* if $d_b(v), d_w(v) \geq L_0$ and *small* otherwise.

We first show how to construct zebraic paths between a pair x, y of large vertices. We can in fact construct paths, even if we decide on the color of the edges incident with x and y. We do breadth first searches from each vertex, alternately using black and white edges, constructing search trees T_x, T_y. We build trees with $n^{2/3+o(1)}$ leaves and then argue that we coonect the leaves with a correctly colored edge. We then find paths between small vertices and other vertices by piggybacking on the large to large paths.

We will need the following structural properties:

Lemma 14 *The following hold w.h.p.:*

(a) *No set S of at most 10 vertices that is connected in G_{t_1} contains three small vertices.*

(b) *Let a be a positive integer, independent of n. No set of vertices S, with $|S| = s \leq aL_1$, contains more than $s + a$ edges in G_{t_1}.*

(c) *There are at most $n^{2/3}$ small vertices.*

(d) *There are at most $\log^3 n$ isolated vertices in G_{t_0}.*

Proof (a) We say that a vertex is a *low color vertex* if it is incident in G_{t_1} to at most $L_\varepsilon = (1+\varepsilon)L_0$ edges of one of the colors, where ε is some sufficiently small positive constant. Furthermore, it follows from (4) that

$$\begin{aligned}
&\mathbf{Pr}(\exists \text{ a connected } S \text{ in } G_{n,t_1} \text{ with three low color vertices})\\
&\leq \sum_{k=3}^{10}\binom{n}{k}k^{k-2}\frac{\binom{N-k+1}{t_1-k+1}}{\binom{N}{t_1}}\binom{k}{3}\mathbf{Pr}(\text{vertices 1,2,3 are low color}) && (28)\\
&\leq_b \sum_{k=3}^{10}\binom{n}{k}k^{k-2}\frac{\binom{N-k+1}{t_1-k+1}}{\binom{N}{t_1}}\binom{k}{3}\left(2\sum_{\ell=0}^{L_\varepsilon}\binom{n-k}{\ell}\left(\frac{p_1}{2}\right)^\ell\left(1-\frac{p_1}{2}\right)^{n-k-\ell}\right)^3 && (29)\\
&\leq_b \sum_{k=3}^{10} n^k\left(\frac{t_1}{N}\right)^{k-1}(n^{-.45})^3\\
&\leq_b \sum_{k=3}^{10} n^k\left(\frac{\log n}{n}\right)^{k-1}(n^{-.45})^3\\
&= o(1).
\end{aligned}$$

Explanation of (28),(29)**:** Having chosen our tree, $\frac{\binom{N-k+1}{t_1-k+1}}{\binom{N}{t_1}}$ is the probability that this tree exists in G_{t_1}. Condition on this and choose three vertices. The final $(\cdots)^3$ in (29) bounds the probability of the event that 1,2,3 are low color vertices in G_{n,p_1}. This event is monotone decreasing, given the conditioning, and so we can use (4) to replace G_{n,t_1} by G_{n,p_1} here.

Now a simple first moment calculation shows that w.h.p. each vertex in $[n]$ is incident with $o(\log n)$ edges of $E_{t_1} \setminus E_{t_0}$. Hence, for (a) to fail, there would have to be a relevant set S with three vertices, each incident in G_{t_1} with at most $(1+o(1))L_0$ edges of one of the colors, contradicting the above.

(b) We will prove something slightly stronger. Suppose that $p = \frac{K \log n}{n}$ where $K > 0$ is arbitrary. We will show this result for $G_{n,p}$. The result for this lemma follows from $K = 1 + o(1)$ and (4). We get

$$\begin{aligned}
\mathbf{Pr}(\exists\ S) &\leq_b \sum_{s \geq 4}^{aL_1} \binom{n}{s} \binom{\binom{s}{2}}{s+a+1} p^{s+a+1} \\
&\leq_b \sum_{s \geq 4}^{aL_1} \left(\frac{ne}{s} \cdot \frac{sep}{2} \right)^s (sep)^{a+1} \\
&\leq_b (Ke^2 \log n)^{aL_1} \left(\frac{\log^2 n}{n} \right)^{a+1} \\
&\leq n^{o(1)} \left(\frac{\log^{3+L_1} n}{n} \right)^a \frac{\log^2 n}{n} \\
&= o(1).
\end{aligned}$$

(c) Using (4) we see that if Z denotes the number of small vertices then

$$\mathbf{E}(Z) \leq_b n \sum_{k=0}^{L_0} (p_0/2)^k (1-p_0/2)^{n-1-k} \leq n^{1/2+o(1)}.$$

We now use the Markov inequality.

(d) Using (4) we see that the expected number of isolated vertices in G_{t_0} is $O(\log^2 n)$. We now use the Markov inequality. $\square$

Now fix a pair of large vertices $x < y$. We will define sets $S_i^{(b)}(z), S_i^{(w)}(z), i = 0, 1, \ldots, \ell_1,\ z = x, y$. Assume w.l.o.g. that ℓ_1 is even. We let $S_0^{(b)}(x) = S_0^{(w)}(x) = \{x\}$ and then $S_1^{(b)}(x)$ (resp. $S_1^{(w)}(x)$) is the set consisting of the first ℓ_0 black (resp. white) neighbors of x. We will use the notation $S_{\leq i}^{(c)}(x) = \bigcup_{j=1}^{i} S_j^{(c)}(x)$ for $c = b, w$. We now iteratively define for $i = 0, 1, \ldots, (\ell_1 - 2)/2$.

$$\hat{S}_{2i+1}^{(b)}(x) = \left\{ v \notin S_{\leq 2i}^{(b)}(x) : v \neq y \text{ is joined by a black } G_{t_0}\text{-edge to a vertex in } S_{2i}^{(b)}(x) \right\}.$$
$$S_{2i+1}^{(b)}(x) = \text{the first } \ell_0 \text{ members of } \hat{S}_{2i+1}^{(b)}(x).$$
$$\hat{S}_{2i+2}^{(b)}(x) = \left\{ v \notin S_{\leq 2i+1}^{(b)} : v \neq y \text{ is joined by a white } G_{t_0}\text{-edge to a vertex in } S_{2i+1}^{(b)}(x) \right\}.$$
$$S_{2i+2}^{(b)}(x) = \text{the first } \ell_0 \text{ members of } \hat{S}_{2i+2}^{(b)}(x) :$$

We then define, for $i = 0, 1, \ldots, (\ell_1 - 2)/2$.

$\hat{S}^{(w)}_{2i+1}(x) = \left\{ v \notin (S^{(b)}_{\leq \ell_1}(x) \cup S^{(w)}_{\leq 2i}(x)) : v \neq y \text{ is joined by a white } G_{t_0}\text{-edge to a vertex in } S^{(w)}_{2i}(x) \right\}$
$S^{(w)}_{2i+1}(x) = $ the first ℓ_0 members of $\hat{S}^{(w)}_{2i+1}(x)$.
$\hat{S}^{(w)}_{2i+2}(x) = \left\{ v \notin (S^{(b)}_{\leq \ell_1}(x) \cup S^{(w)}_{\leq 2i+1}(x)) : v \neq y \text{ s joined by a black } G_{t_0}\text{-edge to a vertex in } S^{(w)}_{2i+1}(x) \right\}$
$S^{(w)}_{2i+2}(x) = $ the first ℓ_0 members of $\hat{S}^{(w)}_{2i+2}(x)$:

Lemma 15 *If $1 \leq i \leq \ell_1$, then in G_{t_0}, for $c = b, w$,*

$$\mathbf{Pr}(|\hat{S}^{(c)}_{i+1}(x)| \leq \ell_0 |S^{(c)}_i(x)| \mid |S^{(c)}_j(x)| = \ell_0^j,\ 0 \leq j \leq i) = O(n^{-anyconstant}).$$

Proof This follows easily from (5) and the Chernoff bounds. Each random variable $\hat{S}^{(c)}(x)$ is binomially distributed with parameters $n - o(n)$ and $1 - (1 - p_0/2)^{\ell_0^i}$. The mean is therefore asymptotically $\frac{1}{2}\ell_0^i \log n = \Omega(\log^2 n)$ and we are asking for the probability that it is much less than half its mean. $\square$

It follows from this lemma, that w.h.p., we may define $S^{(b)}_0(x), S^{(b)}_1(x), \ldots, S^{(b)}_{\ell_1}(x)$ where $|S^{(b)}_i(x)| = \ell_0^i$ such that for each j and $z \in S^{(b)}_j(x)$ there is a zebraic path from x to z that starts with a black edge. For $S^{(w)}_{\ell_1}(x)$ we can say the same except that the zebraic path begins with a white edge.

Having defined the $S^{(c)}_i(x)$ etc., we define sets $S^{(c)}_i(y), i = 1, 2 \ldots, \ell_1$, $c = b, w$. We let $S^{(b)}_0(y) = S^{(w)}_0(y) = \{y\}$ and then $S^{(b)}_1(y)$ (resp. $S^{(w)}_1(y)$) is the set consisting of the first ℓ_0 black (resp. white) neighbors of y that are not in $S^{(b)}_{\leq \ell_1}(x) \cup S^{(w)}_{\leq \ell_1}(x)$. We note that for $c = b, w$ we have $|S^{(c)}_1(y)| \geq L_0 - 18 > \ell_0$. This follows from Lemma 14(b). Suppose that y has ten neighbors T in $S^{(w)}_{\leq \ell_1}(x)$. Let S be the set of vertices in the paths from T to x in $S^{(w)}_{\leq \ell_1}(x)$. If $|S| = s$ then $S \cup \{y\}$ contains at least $s + 9$ edges. This is because every neighbour after the first adds an additional k vertices and $k + 1$ edges to the subgraph of G_{t_0} spanned by $S \cup \{y\}$, for some $k \leq \ell_1$. Now $s + 1 \leq 10\ell_1 + 1 \leq 7L_1$ and the $s + 9$ edges contradict the condition in the lemma, with $a = 7$.

We make a slight change in the definitions of the $\hat{S}^{(c)}_i(y)$ in that we keep these sets disjoint from the $S^{(c')}_i(x)$. Thus we take for example

$$\hat{S}^{(w)}_{2i+1}(y) = \\ \left\{ v \notin (S^{(w)}_{\leq 2i}(y) \cup S^{(b)}_{\leq \ell_1}(x) \cup S^{(w)}_{\leq \ell_1}(x)) : v \text{ is joined by a black } G_{t_0}\text{-edge to a vertex in } S^{(w)}_{2i}(y) \right\}.$$

Then we note that excluding $o(n)$ extra vertices has little effect on the proof of Lemma 15 which remains true with x replaced by y. We can then define the $S^{(c)}_i(y)$ by taking the first ℓ_0 vertices.

Suppose now that we condition on the sets $S^{(c)}_i(x), S^{(c)}_i(y)$ for $c = b, w$ and $i = 0, 1, \ldots, \ell_1$. The edges between the sets with $c = b$ and $i = \ell_1$ and those with $c = w$ and $i = \ell_1$ are unconditioned. Let

$$\Lambda = \ell_0^{2\ell_1} = n^{4/3 - o(1)}.$$

Then, for example, using (4),

$$\mathbf{Pr}(\nexists \text{ a black } G_{t_0} \text{ edge joining } S_{\ell_1}^{(b)}(x), S_{\ell_1}^{(b)}(y)) \leq 3\left(1-\frac{\log n}{(2+o(1))n}\right)^{\Lambda} = O(n^{-anyconstant}). \quad (30)$$

Thus w.h.p. there is a zebraic path with both terminal edges black between every pair of large vertices. A similar argument using $S_{\ell_1}^{(w)}(x), S_{\ell_1}^{(w)}(y)$ shows that w.h.p. there is a zebraic path with both terminal edges white between every pair of large vertices.

If we want a zebraic path with a black edge incident with x and a white edge incident with y then we argue that there is a white G_{t_0} edge between $S_{\ell_1}^{(b)}(x)$ and $S_{\ell_1-1}^{(w)}(y)$.

We now consider the small vertices. Let V_σ be the set of small vertices that have a large neighbor in G_{τ_1}. The above analysis shows that there is a zebraic path between $v \in V_\sigma$ and $w \in V_\sigma \cup V_\lambda$, where V_λ is the set of large vertices. Indeed if v is joined by a black edge to a vertex $w \in V_\lambda$ then we can continue with a zebraic path that begins with a white edge and we can reach any large vertex and choose the color of the terminating edge to be either black or white. This is useful when we need to continue to another vertex in V_σ.

We now have to deal with small vertices that have no large neighbors at time τ_1. It follows from Lemma 14(a) that such vertices have degree one or two in G_{τ_1} and that every vertex at distance two from such a vertex is large.

Lemma 16 *All vertices of degree at most two in G_{t_0} are w.h.p. at distance greater than 10 in G_{t_1},*

Proof Simpler than Lemma 2(b). We use (5) and then

$$\mathbf{Pr}(\exists \text{ such a pair of vertices}) \leq_b t_1^{1/2} \sum_{k=0}^{9} n^k p_1^{k-1} \left((1-p_0)^{n-k-1} + (n-k)p_0(1-p_0)^{n-k-2}\right)^2 = o(1).$$

$\square$

Let Z_i be the number of vertices of degree $0 \leq i \leq 2$ in G_{t_0} that are adjacent in G_{τ_1} to vertices that are themselves only incident to edges of one color.

First consider the case $i = 1, 2$. Here we let Z_i' be the number of vertices of degree i in G_{t_0} that are adjacent in G_{t_0} to vertices that are themselves only incident to edges of one color. Note that

$Z_i \le Z_i'$. Then we have, with the aid of (8),

$$\mathbf{E}(Z_1') \le n\binom{n-1}{1}\frac{\binom{N-n+1}{t_0-1}}{\binom{N}{t_0}}\sum_{k=0}^{n-2}\binom{n-2}{k}\frac{\binom{N-2n+3}{t_0-1-k}}{\binom{N-n+1}{t_0-1}}2^{-(k-1)}. \tag{31}$$

$$\begin{aligned}
&\le_b n^2\frac{t_0}{N}\left(\frac{N-t_0}{N-1}\right)^{n-2}\sum_{k=0}^{n-2}\binom{n-2}{k}2^{-k}\left(\frac{t_0-1}{N-n+1}\right)^k\left(\frac{N-n-t_0+2}{N-n-k+1}\right)^{n-2-k}\\
&\le_b n\log n\exp\left\{-\frac{(n-2)(t_0-1)}{N-1}\right\}\sum_{k=0}^{n-2}\binom{n-2}{k}\left(\frac{t_0-1}{2(N-n+1)}\right)^k\left(\frac{N-n-t_0+2}{N-n-k+1}\right)^{n-2-k}\\
&\le n\log n\exp\left\{-\frac{(n-2)(t_0-1)}{N-1}\right\}\sum_{k=0}^{n-2}\binom{n-2}{k}\left(\frac{t_0}{2(N-n)}\right)^k\left(\frac{N-n-2t_0/3}{N-n}\right)^{n-2-k}\\
&\le_b \log^3 n\left(\frac{t_0}{2(N-n)}+\frac{N-n-2t_0/3}{N-n}\right)^{n-2}\\
&\le \log^3 n\left(\frac{N-t_0/6}{N-n}\right)^{n-2}\\
&= o(1).
\end{aligned}\tag{32}$$

Explanation for (31)**:** We choose a vertex v of degree one and its neighbor w in $n\binom{n-1}{1}$ ways. The probability that v has degree one is $\frac{\binom{N-n+1}{t_0-1}}{\binom{N}{t_0}}$. We fix the degree of w to be $k+1$. This now has probability $\frac{\binom{N-2n+3}{t_0-k-1}}{\binom{N-n+1}{t_0-1}}$. The final factor $2^{-(k-1)}$ is the probability that w only sees edges of one color.

In order to deal with Z_2', we next eliminate the possibility of a vertex of degree two in G_{t_0} being in a triangle of G_{t_1}. First, using (4), the expected number of vertices of degree two in G_{t_0} is at most

$$3n\binom{n-1}{2}p_0^2\,(1-p_0)^{n-3} = O(\log^4 n).$$

So, w.h.p. there are fewer than $\log^5 n$.

Using (5), we see that the expected number of triangles of G_{t_0} containing a vertex of degree two is at most

$$O(t_0^{1/2})\times O(\log^4 n)\times n^3p_0^3(1-p_0)^{n-3} = o(1).$$

So, w.h.p. there are no such triangles.

Then the probability that there is an edge of $G_{t_1}-G_{t_0}$ that joins the two neighbors of a vertex of degree two in G_{t_0} is at most

$$o(1)+\log^5 n\times\frac{t_1-t_0}{N} = o(1).$$

Now we can proceed to estimate $\mathbf{E}(Z_2')$, ignoring the possibility of such a triangle. In which case,

$$\begin{aligned}
&\mathbf{E}(Z_2') \\
&\leq_b n\binom{n-1}{2}\frac{\binom{N-n+1}{t_0-2}}{\binom{N}{t_0}}\sum_{k,l=0}^{n-3}\binom{n-3}{k}\binom{n-3}{l}\frac{\binom{N-3n+6}{t_0-2-k-l}}{\binom{N-n+1}{t_0-2}}2^{-k-l} \qquad (33)\\
&\leq n^3\left(\frac{t_0}{N}\right)^2\left(\frac{N-t_0}{N-2}\right)^{n-3}\times \\
&\qquad\sum_{k,l=0}^{n-3}\binom{n-3}{k}\binom{n-3}{l}\left(\frac{t_0-2}{N-n+1}\right)^{k+l}\left(\frac{N-n-t_0+3}{N-n-k-l+1}\right)^{2n-5-k-l}2^{-k-l}\\
&\leq_b \log^4 n\left(\sum_{k=0}^{n-3}\binom{n-3}{k}\left(\frac{t_0}{2(N-n)}\right)^k\left(\frac{N-t_0}{N-2n}\right)^{n-3-k}\right)^2\\
&\leq \log^4 n\left(1-\frac{t_0-2}{2(N-2n)}\right)^{2(n-3)}\\
&= o(1). \qquad (34)
\end{aligned}$$

Finally, consider Z_0. Condition on G_{t_0} and assume that Properties (c),(d) of Lemma 14 hold. The first edge incident with an isolated vertex of G_{t_0} will have a random endpoint. It follows immediately that

$$\mathbf{E}(Z_0) \leq o(1) + \log^3 n \times n^{-1/3} = o(1). \qquad (35)$$

Here the $o(1)$ accounts for Properties (c),(d) of Lemma 14 and $\log^3 n \times n^{-1/3}$ bounds the expected number of "first edges" that choose small endpoints.

Equations (31), (33) and (35) show that $Z_0 + Z_1 + Z_2 = 0$ w.h.p. In which case it will be possible to find zebraic paths starting from small vertices. Indeed, we now know that w.h.p. any small vertex v will be adjacent to a vertex w that is incident with edges of both colors and that any other neighbor of w is large.

6 Proof of Theorem 3

The case $r = 2$ is implied by Corollary 1 and so we can assume that $r \geq 3$.

6.1 $p \leq (1-\varepsilon)p_r$

For a vertex v, let

$$C_v = \{i : v \text{ is incident with an edge of color } i\}.$$
$$I_v = \{i : \{i, i+1\} \subseteq C_v\}.$$

Let v be *bad* if $I_v = \emptyset$. The existence of a bad vertex means that there are no r-zebraic Hamilton cycles. Let Z_B denote the number of bad vertices. Now if r is odd and $C_v \subseteq \{1, 3, \ldots, 2\lfloor r/2\rfloor - 1\}$ or r is even and $C_v \subseteq \{1, 3, \ldots, r-1\}$ then $I_v = \emptyset$. Hence,

$$\mathbf{E}(Z_B) \geq n\left(1 - \frac{\alpha_r p}{r}\right)^{n-1} = n^{\varepsilon - o(1)} \to \infty.$$

A straightforward second moment calculation shows that $Z_B \neq 0$ w.h.p. and this proves the first part of the theorem.

6.2 $p \geq (1+3\varepsilon)p_r$

Note the replacement of ε by 3ε here, for convenience. Note also that ε is assumed to be sufficiently small for some inequalities below to hold.

Write $1-p = (1-p_1)(1-p_2)^2$ where $p_1 = (1+\varepsilon)p_r$ and $p_2 \sim \varepsilon p_r$. Thus $G_{n,p}$ is the union of G_{n,p_1} and two independent copies of G_{n,p_2}. If an edge appears more than once in $G_{n,p}$, then it retains the color of its first occurence.

Now for a vertex v let $d_i(v)$ denote the number of edges of color i incident with v in G_{n,p_1}. Let

$$J_v = \{i : d_i(v) \geq \eta_0 \log n\}$$

where $\eta_0 = \varepsilon^2/r$.

Let v be *poor* if $|J_v| < \beta_r$ where $\beta_r = \lfloor r/2 \rfloor + 1$. Observe that $\alpha_r + \beta_r = r+1$. Then let Z_P denote the number of poor vertices in G_{n,p_1}. A simple calcluation shows that w.h.p. the minimum degree in G_{n,p_1} is at least L_0 and that the maximum degree is at most $6 \log n$. Then

$$\begin{aligned}
\mathbf{E}(Z_P) &\leq o(1) + n \sum_{k=L_0}^{n-1} \binom{n-1}{k} p_1^k (1-p_1)^{n-1-k} \sum_{l=r-\beta_r+1}^{r} \binom{r}{l} \binom{k}{l\eta_0 \log n} \left(1 - \frac{l}{r}\right)^{k - r\eta_0 \log n} \\
&\leq o(1) + n \sum_{k=0}^{n-1} \binom{n-1}{k} p_1^k (1-p_1)^{n-1-k} 2^r \binom{6 \log n}{r\eta_0 \log n} \left(\frac{\beta_r - 1}{r}\right)^k \left(\frac{r}{\beta_r - 1}\right)^{r\eta_0 \log n} \\
&\leq o(1) + 2^r n^{1 + r\eta_0 \log(6e/\eta_0)} (1-p_1)^{n-1} \left(1 + \frac{(\beta_r - 1)p_1}{r(1-p_1)}\right)^{n-1} \\
&\leq o(1) + 2^r n^{1 + r\eta_0 \log(6e/\eta_0)} \left(1 - \frac{\alpha_r p_1}{r}\right)^{n-1} \\
&= o(1).
\end{aligned}$$

We can therefore assert that w.h.p. there are no poor vertices. This means that

$$K_v = \{i : d_i(v), d_{i+1}(v) \geq \eta_0 \log n\} \neq \emptyset \text{ for all } v \in [n].$$

The proof now follows our general 3-phase procedure of (i) finding an r-zebraic 2-factor, (ii) removing small cycles so that we have a 2-factor in which every cycle has length $\Omega(n/\log n)$ and then (iii) using a second moment calculation to show that this 2-factor can be re-arranged into an r-zebraic Hamilton cycle.

6.2.1 Finding an r-zebraic 2-factor

We partition $[n]$ into r sets $V_i = [(i-1)n/r + 1, in]$ of size n/r. Now for each i and each vertex v let

$$\begin{aligned}
d_i^+(v) &= |\{w \in V_{i+1} : (v,w) \text{ is an edge of } G_{n,p_1} \text{ of color } i+1\}|. \\
d_i^-(v) &= |\{w \in V_{i-1} : (v,w) \text{ is an edge of } G_{n,p_1} \text{ of color } i-1\}|.
\end{aligned}$$

(Here 1-1 is interpreted as r and $r+1$ is interpreted as 1).

We now let a vertex $v \in V_i$ be *i-large* if $d_i^+(v), d_i^-(v) \geq \eta \log n$ where $\eta = \min\{\eta_0, \eta_1, \eta_2\}$ and η_1 is the solution to

$$\eta_1 \log\left(\frac{e(1+\varepsilon)}{r\eta_1\alpha_r}\right) = \frac{1}{r\alpha_r}$$

and η_2 is the solution to

$$\eta_2 \log\left(\frac{3er(1+\varepsilon)}{\eta_2\alpha_r}\right) = \frac{1}{3\alpha_r}.$$

Let v be *large* if it is i-large for all i. Let v be *small* otherwise. (Note that $d_i^+(v), d_i^-(v)$ are defined for all v, not just for $v \in V_i$, $i \in [r]$).

Let V_λ, V_σ denote the sets of large and small vertices respectively.

Lemma 17 *W.h.p., in G_{n,p_1},*

(a) $|V_\sigma| \leq n^{1-\theta}$ *where* $\theta = \frac{\varepsilon}{2r\alpha_r}$.

(b) *No connected subset of size at most* $\log\log n$ *contains more than* $\mu_0 = r\alpha_r$ *members of* V_σ.

(c) *If* $S \subseteq [n]$ *and* $|S| \leq n_0 = n/\log^2 n$ *then* $e(S) \leq 100|S|$.

Proof
(a) If $v \in V_\sigma$ then there exists i such that $d_i^+(v) \leq \eta \log n$ or $d_i^-(v) \leq \eta \log n$. So we have

$$\mathbf{E}(|V_\sigma|) \leq 2rn \sum_{k=0}^{\eta \log n} \binom{n/r}{k} \left(\frac{p_1}{r}\right)^k \left(1 - \frac{p_1}{r}\right)^{n/r-k} \tag{36}$$

$$\leq 3\left(\frac{(1+\varepsilon)e}{r\eta\alpha_r}\right)^{\eta \log n} n^{1-(1+\varepsilon+o(1))/r\alpha_r} \tag{37}$$

$$\leq n^{1-2\theta+o(1)}. \tag{38}$$

Part (a) follows from the Markov inequality. Note that we can lose the factor 2 in (36) since $d_i^+(v) = d_{i+2}^-(v)$.

(b) The expected number of connected sets S of size $2\log\log n$ containing μ_0 members of V_σ can be bounded by

$$\sum_{s=\mu_0}^{2\log\log n} \binom{n}{s} s^{s-2} p_1^{s-1} \binom{s}{\mu_0} \left(r \sum_{k=0}^{\eta \log n} \binom{n/r - s}{k} \left(\frac{p_1}{r}\right)^k \left(1 - \frac{p_1}{r}\right)^{n/r-s-k}\right)^{\mu_0}. \tag{39}$$

Explanation: We choose s vertices for S and a tree to connect up the vertices of S. We then choose μ_0 members $A \subseteq S$ to be in V_σ. We multiply by the probability that for each vertex in A, there is at least one j such that v has few neighbors in $V_j \setminus S$ connected to v by edges of color j.

The sum in (39) can be bounded by

$$n \sum_{s=\mu_0}^{2\log\log n} (4e\log n)^s n^{-\mu_0(1+\varepsilon+o(1))/r\alpha_r} = o(1).$$

(c) This is proved in the same manner as Lemma 2(c). $\square$

For $v \in V_\sigma$ we let $\phi(v) = \min\{i : v \text{ is } i\text{-large}\}$ and then let $X_i = \{v \in V_\sigma : \phi(v) = i\}$ for $i \in [r]$. Now let

$$W_i = (V_i \setminus V_\sigma) \cup \{v \in V_\sigma : \phi(v) = i\}, \quad i = 1, 2, \ldots, r.$$

Suppose that $w_i = |W_i| - n/r$ for $i \in [r]$ and let $w_i^+ = \max\{0, w_i\}$ for $i \in [r]$. We now remove w_i^+ randomly chosen large vertices from each W_i and then randomly assign $w_i^- = -\min\{0, w_i\}$ of them to each $W_i, i \in [r]$. Thus we obtain a partition of $[n]$ into r sets $Z_i, i = 1, 2, \ldots, r$, of size n/r for $i \in [r]$.

Let H_i be the bipartite graph induced by W_i, W_{i+1} and the edges of color i in G_{n,p_1}. We now argue that

Lemma 18 *H_i has minimum degree at least $\frac{1}{2}\eta \log n$ w.h.p.*

Proof It follows from Lemma 17(b) that no vertex in $W_i \cap V_i$ loses more than μ_0 neighbors from the deletion of V_σ. Also, we move $v \in V_\sigma$ to a W_i where it has large degree in V_{i-1} and V_{i+1}. Its neighborhood may have been affected by the deletion of V_σ, but only by at most μ_0. Thus for every i and $v \in X_i$, v has at least $\eta \log n - \mu_0$ neighbors in W_{i-1} connected to v by an edge of color $i-1$. Similarly w.r.t. $i+1$.

Now consider the random re-shuffling to get sets of size n/r. Fix a $v \in V_i$. Suppose that it has $d = \Theta(\log n)$ neighbors in W_{i+1} connected by an edge of color $i+1$. Now randomly choose $w_{i+1} \leq |V_\sigma|$ to delete from W_{i+1}. The number ν_v of neighbors of v chosen is dominated by $\text{Bin}\left(w_{i+1}, \frac{d}{n/r - w_{i+1}}\right)$. This follows from the fact that if we choose these w_{i+1} vertices one by one, then at each step, the chance that the chosen vertex is a neighbor of v is bounded from above by $\frac{d}{n/r - w_{i+1}}$. So, given the condition in Lemma 17(a) we have

$$\mathbf{Pr}(\nu_v \geq 2/\theta) \leq \binom{n^{1-\theta}}{2/\theta} \left(\frac{dr}{n - o(n)}\right)^{2/\theta} \leq \left(\frac{n^{1-\theta} edr\theta}{n}\right)^{2/\theta} = o(n^{-1}).$$

$\square$

We can now verify the existence of perfect matchings w.h.p.

Lemma 19 *W.h.p., each H_i contains a perfect matching $M_i, i = 1, 2, \ldots, r$.*

Proof Fix i. We use Hall's theorem and consider the existence of a set $S \subseteq W_i$ that has fewer than $|S|$ H_i-neighbors in W_{i+1}. Let $s = |S|$ and let $T = N_{H_i}(S)$ and $t = |T| < s$. We can rule out $s \leq n_0 = n/\log^2 n$ through Lemma 17(c). This is because we have $e(S \cup T)/|S \cup T| \geq \frac{1}{4}\eta \log n$ in this case. Let $n_\sigma = |V_\sigma|$ and now consider $n/\log^2 n \leq s \leq n/2r$. Given such a pair S, T we deduce that there exist $S_1 \subseteq S \subseteq V_i, |S_1| \geq s - n_\sigma$ and $T_1 \subseteq T \subseteq V_{i+1}$ and $U_1 \subseteq V_{i+1}, |U_1| \leq n_\sigma$ such that there are at least $m_s = (s\eta/2 - 6n_\sigma)\log n$ edges between S_1 and T_1 and no edges between S_1 and

$V_{i+1} \setminus (T_1 \cup U_1)$. There is no loss of generality in increasing the size of T to s. We can then write

$$\begin{aligned}\mathbf{Pr}(\exists\, S,T) &\leq \sum_{s=n_0}^{n/2r} \binom{n/r-n_\sigma}{s}^2 \binom{s^2}{m_s} p_1^{m_s}(1-p_1)^{(s-n_\sigma)(n/r-s-n_\sigma)} \\ &\leq \sum_{s=n_0}^{n/2r} \left(\frac{ne}{rs}\right)^{2s} \left(\frac{s^2p_1e}{m_s}\right)^{m_s} e^{-(s-n_\sigma)(n/r-s-n_\sigma)p_1} \\ &\leq \left(\left(\frac{s}{n}\right)^{\eta\log n/3} \left(\frac{3er(1+\varepsilon)}{\alpha_r\eta}\right)^{\eta\log n/2} n^{-(1-o(1))/2\alpha_r}\right)^s \\ &= o(1).\end{aligned}$$

For the case $s \geq n/2r$ we look for subsets of V_{i+1} with too few neighbors. □

It follows from symmetry considerations that the M_i are independent of each other. Analogously to Lemma 7, we have

Lemma 20 *The following hold w.h.p.:*

(a) $\bigcup_{i=1}^r M_i$ *has at most* $10\log n$ *components. (Components are* r*-zebraic cycles of length divisible by* r*.)*

(b) *There are at most* n_b *vertices on components of size at most* n_c*.*

Proof The matchings induce a permutation π on W_1. Suppose that $x \in W_1$. We follow a path via a matching edge to W_2 and then by a matching edge to W_3 and so on until we return to a vertex $\pi(x) \in W_1$. π can be taken to be a random permutation and then the lemma follows from Lemma 7. □

The remaining part of the proof is similar to that described in Sections 4.3, 4.4. We use the edges of the first copy G_{n,p_2} of color 1 to make all cycles have length $\Omega(n/\log n)$ and then we use the edges of the second copy of G_{n,p_2} of color 1 to create an r-zebraic Hamilton cycle. The details are left to the reader.

References

[1] N. Alon and J. Spencer, The Probabilistic Method, Third Edition, John Wiley and Sons, 2008.

[2] M. Albert, A.M. Frieze and B. Reed, Multicoloured Hamilton cycles, *Electronic Journal of Combinatorics* **2** (1995), R10.

[3] D. Bal and A.M. Frieze, Rainbow Matchings and Hamilton Cycles in Random Graphs, see arxiv.org. *Random Structures and Algorithms* 22 (2001) 33–42.

[4] B. Bollobás, Random Graphs, Second Edition, Cambridge University Press, 2001.

[5] C.D. Cooper and A. M. Frieze, Multi-coloured Hamilton cycles in randomly coloured random graphs, *Combinatorics, Probability and Computing* 11 (2002) 129–134.

[6] C.D. Cooper and A. M. Frieze, Hamilton cycles in a class of random directed graphs, *Journal of Combinatorial Theory B* 62 (1994) 151–163.

[7] P. Erdős, J. Nešetřil, and V. Rödl, Some problems related to partitions of edges of a graph, in: *Graphs and Other Combinatorial Topics*, Teubner, Leipzig (1983), 54–63.

[8] P. Erdős and A. Rényi, On the existence of a factor of degree one of a connected random graph, *Acta. Math. Acad. Sci. Hungar.* 17 (1966), 359–368.

[9] A. Ferber, G. Kronenberg, F. Mousset and C. Shikhelman, Packing a randomly edge-colored random graph with rainbow k-outs, see arxiv.org.

[10] A.M. Frieze and M. Karoński, Introduction to Random Graphs, Cambridge University Press, 2015.

[11] A.M. Frieze and M. Krivelevich, On packing Hamilton Cycles in ϵ-regular Graphs, *Journal of Combinatorial Theory B* 94 (2005) 159–172. *Random Structures and Algorithms* 41 (2012), 1–22.

[12] A.M. Frieze and P. Loh, Rainbow Hamilton cycles in random graphs, *Random Structures and Algorithms* 44 (2014) 328-354.

[13] A. M. Frieze and T. Łuczak, Hamiltonian cycles in a class of random graphs: one step further, *in Proceedings of Random Graphs '87, Edited by M.Karonski, J.Jaworski and A.Rucinski* (1990), 53–59.

[14] A.M.Frieze and B.A.Reed, *Polychromatic Hamilton cycles*, Discrete Mathematics 118, (1993) 69-74.

[15] G. Hahn and C. Thomassen, Path and cycle sub-Ramsey numbers, and an edge colouring conjecture, *Discrete Mathematics.* **62** (1986), 29–33.

[16] S. Janson, T. Łuczak and A. Ruciński, Random Graphs. John Wiley and Sons, 2000.

[17] R. Robinson and N. Wormald, Hamilton cycles containing randomly selected edges in random cubic graphs, *Random Structures and Algorithms* (2001), 128–147.

[18] P.Winkler, Private Communication.