On factors of independent transversals in k-partite graphs

Raphael Yuster*
Department of Mathematics
University of Haifa
Haifa 3498838, Israel
raphael.yuster@gmail.com

Submitted: Jun 26, 2021; Accepted: Oct 26, 2021; Published: Nov 19, 2021

Abstract

A $[k, n, 1]$-graph is a k-partite graph with parts of order n such that the bipartite graph induced by any pair of parts is a matching. An independent transversal in such a graph is an independent set that intersects each part in a single vertex. A factor of independent transversals is a set of n pairwise-disjoint independent transversals. Let $f(k)$ be the smallest integer n_0 such that every $[k, n, 1]$-graph has a factor of independent transversals assuming $n \geqslant n_0$. Several known conjectures imply that for $k \geqslant 2$, $f(k) = k$ if k is even and $f(k) = k + 1$ if k is odd. While a simple greedy algorithm based on iterating Hall's Theorem shows that $f(k) \leqslant 2k - 2$, no better bound is known and in fact, there are instances showing that the bound $2k - 2$ is tight for the greedy algorithm. Here we significantly improve upon the greedy algorithm bound and prove that $f(k) \leqslant 1.78k$ for all k sufficiently large, answering a question of MacKeigan.

Mathematics Subject Classifications: 05C35, 05C69

1 Introduction

Given a k-partite graph, a *transversal* is a set of vertices containing a single vertex from each part. An *independent transversal* is an independent set which is also a transversal. A *factor of transversals* is a set of pairwise-disjoint transversals covering all vertices. The problem of finding sufficient conditions for the existence of independent transversals and factors of independent transversals in k-partite graphs was studied by several researchers [1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 13, 14, 15, 16, 17] not least because it is strongly related to

*Supported by Israel Science Foundation (grant No. 1082/16).

the Hajnal-Szemerédi Theorem and to concepts such as the strong chromatic number and list coloring.

In this paper we consider sufficient conditions for a factor of independent transversals in very sparse k-partite graphs. To this end, define a $[k, n, 1]$-*graph* to be a k-partite graph with parts of order n such that the bipartite graph induced by any pair of parts is a matching (more generally, in a $[k, n, \Delta]$-graph the bipartite graph induced by any pair of parts has maximum degree at most Δ). Erdős, Gyárfás and Łuczak [5] first considered $[k, n, 1]$-graphs where the matching between each pair is a single edge, and asked for the smallest $n = n(k)$ such that there is always an independent transversal. They proved that $n \geqslant \sqrt{k}(1 - o(1))$, proved an upper bound which is larger by a constant factor, and conjectured that the lower bound is asymptotically tight. The author [16] improved the upper bound of [5] and generalized the problem to arbitrary $[k, n, 1]$-graphs. Very recently, the conjecture of [5] was solved independently by Glock and Sudkaov [7] and by Kang and Kelly [9].

As for a sufficient condition guaranteeing a factor of independent transversals in $[k, n, 1]$-graphs, the requirement of n is obviously more demanding. Formalizing it, let $f(k)$ be the smallest integer n_0 such that every $[k, n, 1]$-graph has a factor of independent transversals assuming $n \geqslant n_0$. To see that $f(k) \geqslant k$, let the parts of a $[k, n, 1]$-graph be denoted by $V_1, \ldots, V_k$ with $V_i = \{v_{i,1}, \ldots, v_{i,n}\}$. Let G be the $[k, k-1, 1]$-graph consisting of the edges $v_{i,1}v_{j,1}$ for all $1 \leqslant i < j \leqslant k$. Then every independent transversal has a single vertex of the form $v_{\star,1}$, so G does not have a factor of independent transversals, proving that $f(k) \geqslant k$. When $k \geqslant 3$ is odd, the following construction of Catlin [4] shows, in fact, that $f(k) \geqslant k + 1$. Let G be the $[k, k, 1]$-graph consisting of the edges $v_{i,t}v_{j,t}$ for all $1 \leqslant i < j \leqslant k$ and for all $1 \leqslant t \leqslant k-2$ and of the edges $v_{i,k-1}v_{j,k}$ and $v_{i,k}v_{j,k-1}$ for all $1 \leqslant i < j \leqslant k$. Then each independent transversal contains exactly one vertex of the form $v_{\star,t}$ for all $1 \leqslant t \leqslant k-2$ and exactly two vertices of the form $v_{\star,k-1}$ or exactly two vertices of the form $v_{\star,k}$. But since k is odd, this means that we can have at most $k-1$ disjoint independent transversals, showing that $f(k) \geqslant k+1$. As we shall see below, the following conjecture, which is implied by special cases of several known conjectures, asserts that the above constructions are the worst cases.

Conjecture 1. Let $k \geqslant 2$. Then $f(k) = k$ if k is even and $f(k) = k + 1$ if k is odd.

It is trivial that $f(2) = 2$ and (see below) easy to show that $f(3) = 4$. A computer-assisted proof (see Section 4) shows that $f(4) = 4$ but in general, Conjecture 1 is wide-open. To motivate this conjecture, let us consider the more general case of $[k, n, \Delta]$-graphs. Recall that the Hajnal-Szemerédi Theorem states that if a graph with nk vertices has maximum degree less than n, then it has n pairwise-disjoint independent sets of size k each. Fischer [6] considered the analogue of the Hajnal-Szemerédi Theorem in $[k, n, \Delta]$-graphs and raised the following intriguing conjecture: Every $[k, n, \Delta]$-graph has a factor of independent transversals as long as $\Delta \leqslant n/k$. Notice that if $k = n$ then Fischer's Conjecture becomes equivalent to the conjecture $f(k) = k$. Fischer's conjecture is false when k is odd due to Catlin's construction but, as we shall see evident, it stands a wide chance of being true in the sense that Catlin's construction (and its generalization

for larger Δ, see [11]) is the only counter-example. In fact, this version (call it the "modified Fischer's Conjecture") is explicitly conjectured by Kühn and Osthus in [11]. In particular, the modified Fischer's Conjecture implies Conjecture 1. The modified Fischer's Conjecture has been solved for $k = 3$ and large n by Magyar and Martin [13], for $k = 4$ and large n by Martin and Szemerédi [14] and, finally, for every fixed k and sufficiently large $n = n(k)$ by Keevash and Mycroft [10]. However, as all of these proofs require n to be very large compared to k, they do not imply Conjecture 1 nor an upper bound close to the conjecture. Lastly, we mention that the author in [16] considered the refinement of $f(k)$ to $[k, n, 1]$-graphs in which every matching consists of at most s independent edges (so the case $s = n$ is the general case of $[k, n, 1]$-graphs). Denoting the refined parameter by $f(k, s)$ we observe by the construction above that already $f(k, 1) \geqslant k$ and it is proved in [16] that $f(k, 1) = f(k, 2) = k$. Furthermore, it is conjectured there that $f(k, s) = k$ for all $1 \leqslant s \leqslant k$, which again states that $f(k) = k$, ignoring Catlin's obstacle (this was also observed in [12]), and so if the modified Fischer Conjecture holds (namely: Catlin's construction is the only obstacle), so does the modified conjecture on $f(k, s) = k$ unless $k \geqslant 3$ is odd and $s = k$ in which case $f(k, s) = k + 1$.

There is, however, a simple upper bound of $f(k) \leqslant 2k - 2$ observed by MacKeigan [12] which follows by greedily applying Hall's algorithm. Assume that G is a $[k, n, 1]$-graph with $n \geqslant 2k - 2$. Suppose we have already constructed a factor F of independent transversals on the induced subgraph G' of G consisting of all but the last set V_k. Then G' is a $[k-1, n, 1]$-graph. Now consider the bipartite graph B with one part being F and the other part being V_k, and there is an edge between $I \in F$ and $v \in V_k$ if v is not a neighbor in G of any vertex of I. Then B has n vertices in each part and minimum degree at least $n-(k-1) \geqslant n/2$, so B has a perfect matching by Hall's Theorem, implying that F can be extended to a factor of independent transversals of G. Notice that this already shows that $f(3) \leqslant 4$ so together with Catlin's example, $f(3) = 4$. Unfortunately, there are examples where this greedy algorithm of repeatedly applying Hall's theorem fails to produce a factor of independent transversals if $n = 2k-3$. Indeed we can construct examples of F as above such that it has a subset of $k - 1$ elements $F^* = \{I_1, \ldots, I_{k-1}\}$, and such that each of the vertices in $V^* = \{v_{k,1}, \ldots, v_{k,k-1}\} \subset V_k$ has exactly one neighbor in G belonging to I_j for each $1 \leqslant j \leqslant k - 1$ (for example, any Latin square of order $k - 1$ corresponds to such a construction). Then in the bipartite graph B, the neighborhood of V^* is only of size $n - (k - 1) = k - 2$, violating Hall's condition, so B has no perfect matching. A problem therefore raised by MacKeigan asks whether it is possible to improve upon the greedy algorithm. Our main result shows that indeed this is the case.

Theorem 2. *$f(k) \leqslant 1.78k$ for all k sufficiently large.*

Our main idea in the proof of Theorem 2 is to apply semi-randomness to the perfect matchings at each stage of the greedy algorithm (hence it is no longer greedy) so as to guarantee that as we come close to the end of the algorithm (close to Stage k) we can still guarantee Hall's condition with high probability.

The rest of this paper continues as follows. In Section 2 we state and prove some preliminary lemmas that are useful in the proof of the main result. Theorem 2 is then

proved in Section 3. In Section 4 we prove that $f(4) = 4$ thereby verifying Conjecture 1 in the first non-elementary case. Throughout the paper, we assume, whenever necessary, that k and n are sufficiently large.

2 Preliminary setup

Let us first observe that if we can show that every $[k, n, 1]$-graph has a factor of independent transversals, then $f(k) \leqslant n$. Indeed, if G is a $[k, n', 1]$-graph with $n' > n$ we can just take an induced subgraph consisting of n vertices of each part observing that this subgraph is an $[n, k, 1]$-graph. Being such, it has a (factor of) independent transversal(s), so we can just apply this repeatedly to find $n' - n$ pairwise disjoint independent transversals in G. Removing them from G we remain with an induced subgraph G' of G which is an $[n, k, 1]$-graph and find a factor of independent transversals in G', hence, altogether, a factor of independent transversals in G.

In our proof of Theorem 2 it will be more convenient to fix the number of vertices in each part and to parameterize on the number of parts. Let, therefore $g(n)$ denote the largest integer k such that every $[k, n, 1]$-graph has a factor of independent transversals. We therefore have:

Corollary 3. *If $g(n) \geqslant 0.562n$ for all n sufficiently large, then $f(k) \leqslant 1.78k$ for all k sufficiently large, so Theorem 2 holds.*

As we prove in Lemma 11 that $g(n) \geqslant 0.562n$ for all n sufficiently large, we have by the last corollary that Theorem 2 holds.

For a graph G and a set of its vertices X, let $N_G(X)$ denote the set of vertices of G that have a neighbor in X. The following is an immediate corollary of Hall's Theorem.

Lemma 4. *Let B be a bipartite graph with vertex parts X, Y of order m each. Assume that the degree of every vertex is at least $m - t$ and, furthermore, every $W \subseteq Y$ with $m - t \leqslant |W| \leqslant t$ has $|N_B(W)| \geqslant |W|$. Then B has a perfect matching.* □

Lemma 5. *Let $\delta > 0$. Then there exists a constant $C = C(\delta) > 0$ such that the following holds. Let B be a bipartite graph with vertex parts $X = \{x_1, \ldots, x_n\}, Y = \{y_1, \ldots, y_n\}$ and suppose the degree of each vertex of X is precisely $n - t$. Furthermore, suppose that $X^* \subseteq X$ and $Y^* \subseteq Y$ are given. Let π be a random permutation of the vertices of Y and let $M = \{(x_i, y_{\pi(i)})\}_{i=1}^n$. Then with probability at least $1 - C/n$ the following hold:*

(i) The number of pairs in M that are edges of B is at least $n - t - \delta n$ and at most $n - t + \delta n$.

(ii) There is a subset $M^* \subseteq M$ with $|M^*| \geqslant |X^*|(|Y^*| - t)/n - \delta n$ such that every pair $(x_i, y_{\pi(i)}) \in M^*$ is an edge of B and furthermore, $x_i \in X^*$ and $y_{\pi(i)} \in Y^*$.

Proof. We start with (ii). Suppose without loss of generality, that the vertices of X^* are $\{x_1, \ldots, x_q\}$ where $q = |X^*|$. Let d_i be the number of neighbors of x_i in Y^* and

observe that $d_i \geqslant (n-t) - (n - |Y^*|) = |Y^*| - t$. Let R_i be the indicator random variable for the event $A_i =$ "$y_{\pi(i)} \in Y^*$ and $(x_i, y_{\pi(i)})$ is an edge of B". Let $R = \sum_{i=1}^{q} R_i$ count the number of pairs that satisfy the condition of being in M^*. Now, $Pr[A_i] = d_i/n \geqslant (|Y^*| - t)/n$. It follows that $\mathbb{E}[R] \geqslant q(|Y^*| - t)/n$. It remains to upper bound the probability that R deviates from its expectation by more than δn. Now, if $q \leqslant \delta n$ then (ii) is trivial since $R \leqslant q$. Otherwise, R is the sum of $\Theta(n)$ indicator random variables. We claim that $Var[R] = O(n)$. Indeed, $Var[R_i] \leqslant 1$ as it is an indicator variable. We estimate $Cov(R_i, R_j) = \Pr[A_i \cap A_j] - d_i d_j/n^2$. Now, given that A_i occurred, there are two possibilities. If $y_{\pi(i)}$ is a neighbor of x_j, then A_j occurred with probability $(d_j - 1)/(n-1)$. If $y_{\pi(i)}$ is not a neighbor of x_j, then A_j occurred with probability $d_j/(n-1)$. In any case,

$$\frac{d_i(d_j - 1)}{n(n-1)} \leqslant \Pr[A_i \cap A_j] \leqslant \frac{d_i d_j}{n(n-1)}$$

So $Cov(R_i, R_j) \leqslant \frac{2}{n-1}$ implying that $Var[R] = O(n)$. It now follows from Chebyshev's inequality that the probability that R deviates from its expected value by more than δn is $O(1/n)$.

For (i), apply the proof above for the special case that $X^* = X$ and $Y^* = Y$ (so now $q = n$ and $|Y^*| = n$). Then the corresponding indicator random variable R_i has $\Pr[R_i] = (n-t)/n$ since the degree of x_i is precisely $n - t$. Hence, $\mathbb{E}[R] = n - t$ and, as we have shown, the probability that R deviates from its expectation by more than δn is $O(1/n)$. □

We need the next definition that formalizes the following procedure. Suppose we are given a bipartite graph B with n vertices in each part and a (not necessarily perfect) matching M of B. If we take an induced subgraph B^* of B consisting of all the vertices not in M and some pairs of matched vertices of M, we can consider a maximum matching in B^*. If this maximum matching of B^* is a perfect matching of B^*, then (obviously) we obtain a perfect matching of B as well.

Definition 6 (**(B, M, s)**-reshuffling; leftover graph)**.** Let $n \geqslant s \geqslant 0$. Let B be a bipartite graph with n vertices in each part, and suppose that M is a matching of B consisting of $m \geqslant s$ edges. Consider an induced bipartite subgraph B^* of B obtained by taking all $2n - 2m$ vertices not in M and taking $2s$ vertices of M obtained by randomly selecting s edges of M and taking their endpoints (observe that B^* has $n - m + s$ vertices in each part). We call B^* the *leftover graph*. Now take a maximum matching M^* of B^*. Then M^* together with the $m - s$ edges of M not belonging to B^* is a matching of B of size $|M^*| + m - s$. In particular, if M^* is a perfect matching of B^* then we obtain a perfect matching of B as well. We call this procedure a (B, M, s)*-reshuffling*. We say that the reshuffling is *successful* if M^* is a perfect matching of B^*.

We emphasize that when constructing B^* in Definition 6, it will be very important to select the s edges of M at random, as stated, and not just arbitrarily.

In our proof we will need to ascertain that the constant we choose satisfies the following constraint.

Lemma 7. *Let $0 < c \leqslant 1$ satisfy $2c^2 \ln(\frac{1+c}{c}) \geqslant 1$ (note: any $0.778 \leqslant c \leqslant 1$ satisfies this). Then the following holds for all $0 \leqslant \mu \leqslant \frac{1}{1+c}$.*

$$\int_0^{\mu} \left[1 - (1 - c\mu - x)\left(1 - \frac{cx}{1-x}\right)\right] dx \leqslant c\mu \ .$$

Furthermore, for every $\epsilon > 0$ there exists $\gamma > 0$ such that if $0 \leqslant \mu \leqslant \frac{1-\epsilon}{1+c}$, the definite integral is at most $(c - \gamma)\mu$.

Proof. The definite integral equals $\mu^2(c^2 + \frac{3c}{2} + \frac{1}{2}) + \mu c^2 \ln(1-\mu)$. Hence, dividing by μ, it suffices to prove that the function

$$f(\mu) = \mu(c^2 + \frac{3c}{2} + \frac{1}{2}) + c^2 \ln(1-\mu) - c$$

is non-positive in the range $0 \leqslant \mu \leqslant \frac{1}{1+c}$. The derivative of f is

$$f'(\mu) = c^2 + \frac{3c}{2} + \frac{1}{2} - \frac{c^2}{1-\mu} \ .$$

So, f is strictly monotone increasing if $\mu < 1 - c^2/(c^2 + 3c/2 + 1/2)$ and indeed this is easily verified to hold for all $0 \leqslant \mu \leqslant \frac{1}{1+c}$. Since $f(0) = -c$, to prove the non-negativity of f in the specified range it suffices to prove that $f(\frac{1}{1+c}) \leqslant 0$. Now, rearranging the terms it is easily verified that

$$f\left(\frac{1}{1+c}\right) \leqslant 0 \quad \Leftrightarrow \quad 2c^2 \ln\left(\frac{1+c}{c}\right) \geqslant 1 \ .$$

For the second part of the lemma, define $\gamma = h(\epsilon) = -f\left(\frac{1-\epsilon}{1+c}\right)$ and observe that $h(0) \geqslant 0$ and that h is a continuous strictly monotone increasing function, since f is. □

3 Proof of the main result

In this section we assume that G is a $[k, n, 1]$-graph with vertex parts $V_1, \ldots, V_k$ of order n each. We let G_t denote the subgraph of G induced by $\cup_{i=1}^{t} V_i$, so G_t is a $[t, n, 1]$-graph.

Definition 8 (***t*-partial factor of independent transversals**)**.** *A t-partial factor of transversals of G* is a factor of transversals of G_t. If each element in the factor is an independent set, it is called a *t-partial factor of independent transversals of G*.

In our proof we construct a sequence $F_1, \ldots, F_k$ such that F_t is a t-partial factor of (hopefully independent) transversals of G. Let $F_t = \{I(t,1), \ldots, I(t,n)\}$ where $I(t,j)$ is a (not necessarily independent) transversal of G_t. Our sequence will have the property that for every $1 \leqslant t \leqslant k-1$, F_{t+1} *extends* F_t, meaning that $I(t+1,j) \setminus V_{t+1} = I(t,j)$. To facilitate the construction of F_{t+1} as an extension of F_t we need the following definition.

Definition 9 (The auxiliary bipartite graph $\boldsymbol{B_t}$)**.** Suppose that F_t is t-partial factor of transversals. The *auxiliary bipartite graph* B_t is defined by one part being F_t (so the vertices of this part are $I(t,j)$ for $1 \leqslant j \leqslant n$) and the other part being V_{t+1}. There is an edge of B_t connecting $I(t,j) \in F_t$ and $v \in V_{t+1}$ if all t vertices of $I(t,j)$ are not neighbors of v in G.

Observe that a perfect matching in B_t corresponds to F_{t+1} by defining $I(t+1,j)$ to be the union of $I(t,j)$ and its match. Notice that indeed F_{t+1} extends F_t and notice that if F_t is a t-partial factor of *independent* transversals, then F_{t+1} is also a $(t+1)$-partial factor of independent transversals.

Lemma 10. *The minimum degree of B_t is at least $n-t$.*

Proof. Observe first that $I(t,i) \in F_t$ has t vertices, and each such vertex has at most one neighbor (in G) belonging to V_{t+1}. Hence, the number of non-neighbors of $I(t,i)$ in B_t is at most t, so the degree of $I(t,i)$ in B_t is at least $n-t$. Observe next that each $v \in V_{t+1}$ has at most one neighbor (in G) in V_j for $j = 1,\ldots,t$. So, for each $1 \leqslant j \leqslant t$, there is at most one element of F_t whose unique vertex belonging to V_j is a neighbor (in G) of v. So, each $1 \leqslant j \leqslant t$ contributes at most one non-neighbor of v in B_t. Hence, the degree of v in B_t is at least $n-t$. □

The following is our main lemma.

Lemma 11. *For all sufficiently large n it holds that $g(n) \geqslant 0.562n$.*

Proof. Let $0.778 \leqslant c \leqslant 1$ and recall that the conditions of Lemma 7 hold. (Note: our proof works for every c in this range, but we will eventually optimize by using $c = 0.778$.) Suppose $\epsilon > 0$ is a given small constant. Let $\gamma = \gamma(\epsilon)$ be the constant from Lemma 7 and choose constants $0 < \delta \ll \eta \ll \gamma$ (by $x \ll y$ we mean that x is a small function of y, small enough to satisfy the claimed inequalities that will follow). Let $C = C(\delta)$ be the constant from Lemma 5. Throughout the proof we assume that n is sufficiently large as a function of ϵ. Let $k = \lfloor n(1-\epsilon)/(c+1) \rfloor$ and let G be a $[k,n,1]$-graph with vertex parts $V_1,\ldots,V_k$ of order n each. We will prove that G has a factor of independent transversals.

We sequentially construct partial factors of transversals $F_1,\ldots,F_k$ starting with $F_1 = \{I(1,1),\ldots,I(1,n)\}$ where $I(1,j)$ is just a single vertex of V_1. Trivially, F_1 is a 1-partial factor of independent transversals of G. We next define stages $t = 1,\ldots,k-1$ where at Stage t we construct (partly using a probabilistic argument) F_{t+1} as an extension of F_t. Note that at this point we only assume that F_t is a t-partial factor of transversals. However, we will prove that with positive probability, all elements of F_t, for all $t = 1,\ldots,k$ are in fact independent sets. Hence we will obtain with positive probability that the final F_k is a factor of independent transversals of G, as required. We now describe Stage t in detail, assuming all previous stages have been completed (for completeness, Stage 0 constructs the trivial F_1 as above).

Consider B_t in which one part is the already constructed $F_t = \{I(t,1),\ldots,I(t,n)\}$ and the other part is $V_{t+1} = \{y_1,\ldots,y_n\}$. Recall by Lemma 10 that each vertex of B_t has degree at least $n-t$. So, remove edges from B_t until each vertex of the part

F_t of B_t has degree exactly $n-t$ and denote the resulting spanning subgraph by B'_t (observe that in B'_t some vertices in the part V_{t+1} may now have degree smaller than $n-t$). Randomly select a permutation of V_{t+1}, denoting it by π_t. Now consider the set of all pairs $M(\pi_t) = \{(I(t,i), y_{\pi_t(i)})\}_{i=1}^{n}$. Unfortunately, not each pair $(I(t,i), y_{\pi_t(i)})$ is necessarily an edge of B_t and thus, moreover, not necessarily an edge of B'_t (recall, to be an edge of B_t we should have that all t vertices of $I(t,i)$ are not neighbors in G of $y_{\pi_t(i)}$). Let $M'(\pi_t)$ denote the set of elements of $M(\pi_t)$ that are edges of B'_t and set $m_t = |M'(\pi_t)|$. We say that $M(\pi_t)$ is *good* if $|m_t - (n-t)| \leqslant \delta n$.

Claim 12. *$M(\pi_t)$ is good with probability at least $1 - C/n$.*

Proof. We apply Lemma 5 with $B = B'_t$, $X = F_t$, $Y = V_{t+1}$ and $\pi = \pi_t$. According to case (i) of that lemma, with probability at least $1 - C/n$, $M(\pi_t)$ contains at least $(n-t) - \delta n$ edges of B'_t and at most $(n-t) + \delta n$ edges of B'_t. □

It is important to note that the stated probability $1 - C/n$ of $M(\pi_t)$ being good does not depend on the goodness of $M(\pi_\ell)$ for any $\ell < t$, as the proof of Lemma 5 does not assume anything except that the degree of each vertex of B'_t is $n-t$ and that the permutation π_t is chosen at random.

Now, if $M(\pi_t)$ is good (which happens with very high probability by Claim 12) we proceed by performing a $(B_t, M'(\pi_t), s_t)$-reshuffling where $s_t = \lfloor ct + \eta n \rfloor$ (recall Definition 6). To see that the parameters in Definition 6 fit, we must show that $s_t \leqslant m_t = |M'(\pi_t)|$. Indeed this holds since $s_t \leqslant ct + \eta n \leqslant n - t - \delta n \leqslant m_t$ where we have used that $t(c+1) \leqslant k(c+1) \leqslant n(1-\epsilon) \leqslant n(1-\delta-\eta)$. Now, if this $(B_t, M'(\pi_t), s_t)$-reshuffling is successful (recall again Definition 6), then B_t has a perfect matching, and we define F_{t+1} using this perfect matching as shown in the paragraph after Definition 9. Recall that F_{t+1} extends F_t and that if F_t is a t-partial factor of *independent* transversals, then F_{t+1} is also a $(t+1)$-partial factor of independent transversals. We can now define the notion of a successful stage.

Definition 13 (Successful stage)**.** We say that Stage t is *successful* if $M(\pi_t)$ is good and furthermore, the corresponding $(B_t, M'(\pi_t), s_t)$-reshuffling is successful.

As an immediate corollary we obtain:

Corollary 14. *If all stages are successful then F_k is a factor of independent transversals of G.* □

Now, Corollary 23 below asserts that indeed, the probability that all stages are successful is positive, so given Corollary 23, we have that F_k is a factor of independent transversals of G. Now, if we choose $c = 0.778$ then $k \geqslant 0.5624n(1-\epsilon)$, so Lemma 11 holds. □

It remains to prove Corollary 23 below, namely that with positive probability, all stages are successful. To help with the analysis, it would be convenient (i.e. more uniform) to also say how we construct F_{t+1} as an extension of F_t in case that Stage t is not successful, either because $M(\pi_t)$ is not good or because the $(B_t, M'(\pi_t), s_t)$-reshuffling

failed. While this may seem artificial (since if some stage failed, why proceed to the next stage?) the analysis becomes more uniform as it will help us lower-bound success probability of stages regardless of successes or failures of previous stages. So, if Stage t is unsuccessful, we simply use $M(\pi_t)$ (which might certainly not be a perfect matching of B_t) to define F_{t+1}. In other words, we define $I(t+1,i) = I(t,i) \cup \{y_{\pi_t(i)}\}$ for $i = 1, \ldots, n$ and $F_{t+1} = \{I(t+1,i)\}_{i=1}^{n}$. Observe that even if F_t is a factor of independent transversals, F_{t+1} might not be, since if $(I(t+1,i), y_{\pi_t(i)})$ is not an edge of B_t, then $I(t+1,i)$ is not an independent set of G_{t+1}. Nevertheless, F_{t+1} is still a factor of transversals that extends F_t. We have now defined how to perform all stages $1 \leqslant t \leqslant k-1$ (some of which may be successful, while others might not be).

Corollary 15. *With probability at least $\frac{1}{2}e^{-C}$ we have that $M(\pi_t)$ is good for all stages $1 \leqslant t \leqslant k-1$.*

Proof. By Claim 12, for each t it holds that $M(\pi_t)$ is good with probability $1 - C/n$ and recall that this probability bound holds for stage t regardless of the goodness of stages other than t. Hence, the probability that $M(\pi_t)$ is good for all $1 \leqslant t \leqslant k-1$ is at least $(1 - C/n)^k \geqslant (1 - C/n)^n \geqslant \frac{1}{2}e^{-C}$. □

Given that $M(\pi_t)$ is good, we would like to ascertain some property whose existence guarantees that the corresponding $(B_t, M'(\pi_t), s_t)$-reshuffling is successful. One way to do that is to prove that the leftover graph ${B_t}^*$ of the $(B_t, M'(\pi_t), s_t)$-reshuffling satisfies the conditions of Lemma 4 (so B_t^* has a perfect matching implying that the $(B_t, M'(\pi_t), s_t)$-reshuffling is successful, hence the entire Stage t is successful). We start by showing that if t is very small (namely, in the first few stages) then, given that $M(\pi_t)$ is good, surely (i.e. with probability 1) the corresponding $(B_t, M'(\pi_t), s_t)$-reshuffling is successful.

Lemma 16. *For $1 \leqslant t \leqslant \eta n/3$, if $M(\pi_t)$ is good, then the corresponding $(B_t, M'(\pi_t), s_t)$-reshuffling is successful, namely Stage t is successful.*

Proof. We need to prove that the leftover graph B_t^* has a perfect matching. Recall (Definition 6) that B_t^* has $n - m_t + s_t$ vertices in each part. On the other hand, B_t^* is an induced subgraph of B_t and by Lemma 10 the latter has minimum degree at least $n - t$. Hence the minimum degree of B_t^* is at least $n - m_t + s_t - t$. Since by Hall's Theorem, every bipartite graph with x vertices in each side and minimum degree at least $x/2$ has a perfect matching, it suffices to prove that $n - m_t + s_t - t \geqslant (n - m_t + s_t)/2$. Equivalently, it suffices to prove that $t \leqslant (n - m_t + s_t)/2$. Indeed, this holds since

$$\frac{n - m_t + s_t}{2} \geqslant \frac{s_t}{2} = \frac{\lfloor ct + \eta n \rfloor}{2} \geqslant \frac{\eta n}{3} \geqslant t\,. \qquad \square$$

In contrast with Lemma 16, proving that (with high probability) the reshuffling corresponding to larger t is successful is more involved. We start with the following lemma that gives a sufficient condition for a reshuffling to be successful when $t \geqslant \eta n/3$.

Lemma 17. *Let $t \geqslant \eta n/3$ and suppose that $M(\pi_t)$ is good. If each $W \subset V_{t+1}$ with $|W| = \lfloor ct \rfloor$ has $|N_{B_t}(W)| \geqslant n - \lfloor \eta n/7 \rfloor$ then the $(B_t, M'(\pi_t), s_t)$-reshuffling is successful.*

Proof. As in the proof of Lemma 16 we observe that B_t^* has $n - m_t + s_t$ vertices in each part and its minimum degree is at least $n - m_t + s_t - t$. Let $W \subseteq V_{t+1}$ with $t \geqslant |W| \geqslant n - m_t + s_t - t$. If we can show that $|N_{B_t^*}(W)| \geqslant |W|$ then by Lemma 4, B_t^* has a perfect matching. It thus suffices to prove the stronger statement, that each $W \subseteq V_{t+1}$ with $|W| = n - m_t + s_t - t$ has $|N_{B_t^*}(W)| \geqslant t$. To prove the latter, it suffices to prove the even stronger statement that $|N_{B_t}(W)| \geqslant t + m_t - s_t$. Now, as $M(\pi_t)$ is good we have that $m_t \leqslant n - t + \delta n$ and recall that $s_t = \lfloor ct + \eta n \rfloor \geqslant ct$. Thus, we have:

$$\begin{aligned} t + m_t - s_t &\leqslant t + (n - t) + \delta n - ct \\ &= n + \delta n - ct \leqslant n + \delta n - c\eta n/3 \\ &\leqslant n + \delta n - \eta n/6 \\ &\leqslant n - \lfloor \eta n/7 \rfloor \end{aligned}$$

where we have used that $c \geqslant \frac{1}{2}$ and $\delta \ll \eta$. We also have that

$$\begin{aligned} n - m_t + s_t - t &\geqslant n - (n - t + \delta n) + s_t - t \\ &= s_t - \delta n \\ &= \lfloor ct + \eta n \rfloor - \delta n \\ &\geqslant \lfloor ct \rfloor \, . \end{aligned}$$

Hence, if $W \subset V_{t+1}$ with $|W| = \lfloor ct \rfloor$ has $|N_{B_t}(W)| \geqslant n - \lfloor \eta n/7 \rfloor$ then B_t^* has a perfect matching, implying that the $(B_t, M'(\pi_t), s_t)$-reshuffling is successful. □

Proving that the conditions of Lemma 17 hold with high probability, namely that with high probability each $W \subset V_{t+1}$ with $|W| = \lfloor ct \rfloor$ has $|N_{B_t}(W)| \geqslant n - \lfloor \eta n/7 \rfloor$, is rather technical and requires a few additional definitions, notations, and lemmas.

For $Q \subseteq [n]$ and for a t-partial factor of transversals $F_t = \{I(t,i)\}_{i=1}^n$, let $F_t(Q) = \{I(t,i) \mid i \in Q\}$. For $W \subseteq V_{t+1}$ and $1 \leqslant \ell \leqslant t$, let $Y_\ell(W) = V_\ell \setminus N_G(W)$ (i.e. the non-neighbors of W in G that belong to V_ℓ). Also define $Y_{t+1}(W) = V_{t+1} \setminus W$. Observe that $|Y_\ell(W)| \geqslant n - |W|$ holds for all $1 \leqslant \ell \leqslant t+1$ since each vertex of W has at most one neighbor in G that belongs to V_ℓ.

We need the notions of $F_t(Q)$ and $Y_\ell(W)$ from the previous paragraph in the following definition, which is an important ingredient in the remainder of the proof.

Definition 18 (**(Q, ℓ, W, r)**-intersection)**.** Let $Q \subseteq [n]$, let $W \subseteq V_{t+1}$ and let $1 \leqslant \ell \leqslant t$. We say that a (Q, ℓ, W, r)-*intersection* exists if
(i) $M(\pi_\ell)$ is good (so we proceed to doing a $(B_\ell, M'(\pi_\ell), s_\ell)$-reshuffling),
(ii) there are at least r edges of $M'(\pi_\ell)$ where in each of these r edges, one endpoint is from $F_\ell(Q)$ and the other endpoint is from $Y_{\ell+1}(W)$,
(iii) the $2r$ endpoints of these r edges from (ii) are *not* selected to the leftover graph B_ℓ^*.

Until the end of this section, let

$$q = \lceil \eta n/7 \rceil \qquad r_{\ell,t} = \frac{q(n - ct - \ell)}{n} \left(1 - \frac{c\ell}{n - \ell}\right) - 5\eta q \, .$$

Lemma 19. *Let $1 \leqslant t \leqslant k-1$, let $T = \{1, \dots, t\}$, let $W \subseteq V_{t+1}$ with $|W| = \lfloor ct \rfloor$, and let $Q \subset [n]$ with $|Q| = q$. Then with probability at least $1 - 5^{-n}$ the following holds. There exists $T^* \subseteq T$ with $|T^*| \geqslant t - \delta n$ such that for all $\ell \in T^*$, a $(Q, \ell, W, r_{\ell,t})$-intersection exists.*

Proof. Suppose that t, W, Q are given as in the statement of the lemma. Fixing some $\ell \in T$ we estimate the probability that a $(Q, \ell, W, r_{\ell,t})$-intersection exists. We recap what we are doing in Stage ℓ. At the beginning of that stage, we have F_ℓ, the already constructed ℓ-partial factor of transversals (note: we do not assume anything about success or failure of prior stages) as one part of B_ℓ and we have $V_{\ell+1}$ as the other part. We remove some edges of B_ℓ to obtain B'_ℓ. We then take a random permutation π_ℓ of $V_{\ell+1}$ yielding $M(\pi_\ell)$ and its subset $M'(\pi_\ell)$ which is a matching of B'_ℓ. Observe that $F_\ell(Q)$ is a subset of F_ℓ of order q, and that $Y_{\ell+1}(W)$ is a subset of $V_{\ell+1}$ of order at least $n - |W|$.

Now, apply Lemma 5 with $B = B'_t$, $X = F_\ell$, $Y = V_{\ell+1}$, $X^* = F_\ell(Q)$, $Y^* = Y_{\ell+1}(W)$, $\pi = \pi_\ell$, $M = M(\pi_\ell)$. We obtain from that lemma that with probability $1 - C/n$, both items (i) and (ii) of Lemma 5 hold. Recall that item (i) of Lemma 5 says that $M(\pi_\ell)$ is good, since $|m_\ell - (n - \ell)| \leqslant \delta n$ where $m_\ell = |M'(\pi_\ell)|$, so already item (i) of Definition 18 holds. Item (ii) of Lemma 5 says that there is a subset $M^* \subset M(\pi_\ell)$ of size

$$|M^*| \geqslant \frac{|F_\ell(Q)|(|Y_{\ell+1}(W)| - \ell)}{n} - \delta n \geqslant \frac{q(n - |W| - \ell)}{n} - \delta n \geqslant \frac{q(n - ct - \ell)}{n} - \delta n \quad (1)$$

such that every pair $(x_i, y_{\pi(i)}) \in M^*$ is and edge of B'_t (hence an edge of $M'(\pi_\ell)$) and furthermore, $x_i \in F_\ell(Q)$ and $y_{\pi_\ell(i)} \in Y_{\ell+1}(W)$.

Notice that if $r_{\ell,t} \geqslant |M^*|$ then Item (ii) of Definition 18 holds, but this is not enough for us, as we also want Item (iii) to hold. In the reshuffling, we create B^*_ℓ by randomly selecting s_ℓ out of the m_ℓ edges of $M'(\pi_\ell)$ and add the $2s_\ell$ endpoints of the selected edges to the vertices not matched by $M'(\pi_\ell)$. So, the probability that a single edge (and its two endpoints) from M^* is selected to B^*_ℓ is precisely s_ℓ/m_ℓ. Hence, if $Z \subseteq M^*$ denotes the set of edges of M^* that are *not* selected to the leftover graph B^*_ℓ then $\mathbb{E}[|Z|] = |M^*|(1 - \frac{s_\ell}{m_\ell})$. Observe that $|Z|$ has hypergeometric distribution (as its distribution is identical to having $|M^*|$ red balls, $m_t - |M^*|$ blue balls, and we select without replacement $m_\ell - s_\ell$ balls and ask for the number of red balls that are selected). As trivially, $|Z| \leqslant n$, we have by Chebyshev's inequality, that the probability that $|Z|$ deviates from its expected value by more than δn is $O(1/n)$ (since the variance in our hypergeometric distribution is $O(n)$). So overall, we obtain that with probability $1 - O(1/n)$,

$$|Z| \geqslant |M^*| \left(1 - \frac{s_\ell}{m_\ell}\right) - \delta n \, .$$

To prove that with probability $1 - O(1/n)$, a $(Q, \ell, W, r_{\ell,t})$-intersection exists we just need to make sure that $r_{\ell,t}$ is not larger than the right-hand side of the last inequality. Indeed, using (1), using Lemma 20 below which upper bounds s_ℓ/m_ℓ and using $\delta \ll \eta$ we obtain

$$|M^*| \left(1 - \frac{s_\ell}{m_\ell}\right) - \delta n \geqslant \left(\frac{q(n - ct - \ell)}{n} - \delta n\right) \left(1 - \frac{s_\ell}{m_\ell}\right) - \delta n$$

$$
\begin{aligned}
&\geqslant \frac{q(n-ct-\ell)}{n}\left(1-\frac{s_\ell}{m_\ell}\right)-2\delta n\\
&\geqslant \frac{q(n-ct-\ell)}{n}\left(1-\frac{c\ell}{n-\ell}-4\eta\right)-2\delta n\\
&\geqslant \frac{q(n-ct-\ell)}{n}\left(1-\frac{c\ell}{n-\ell}\right)-4\eta q-2\delta n\\
&\geqslant \frac{q(n-ct-\ell)}{n}\left(1-\frac{c\ell}{n-\ell}\right)-5\eta q\\
&= r_{\ell,t}\ .
\end{aligned}
$$

We have proved that with probability $1-O(1/n)$, a $(Q,\ell,W,r_{\ell,t})$-intersection exists but we still need to prove that the claimed T^* exists with high probability. First observe that if $t\leqslant \delta n$ then T^* exists vacuously so assume that $t\geqslant \delta n$. For T^* not to exist, there should be a set of $\lceil \delta n\rceil$ indices ℓ with $1\leqslant \ell\leqslant t$, such that a $(Q,\ell,W,r_{\ell,t})$-intersection does not exist. Fix a set $L\subseteq T$ of $\lceil\delta n\rceil$ indices. We have proved that the probability that for $\ell\in L$, a $(Q,\ell,W,r_{\ell,t})$-intersection does not exist is $O(1/n)$ regardless of the outcome of any other $\ell'\in L$. Hence, the probability that all $\ell\in L$ are such that the corresponding $(Q,\ell,W,r_{\ell,t})$-intersection does not exist is $O(1/n)^{\delta n}$. As there are less than 2^n possible sets L to consider, we have by the union bound that the probability that T^* does not exist is at most

$$2^nO(1/n)^{\delta n}\ll 5^{-n}\ .$$ □

We next prove the upper bound for $\frac{s_\ell}{m_\ell}$ that we have used in Lemma 19.

Lemma 20. *For all $1\leqslant\ell\leqslant k-1$, if $M(\pi_\ell)$ is good then*

$$\frac{s_\ell}{m_\ell}\leqslant\frac{c\ell}{n-\ell}+4\eta\ .$$

Proof. Recall that $s_\ell=\lfloor c\ell+\eta n\rfloor$ and that $M(\pi_\ell)$ being good implies that $m_\ell\geqslant n-\ell-\delta n$. Also notice that $\ell<k<n/(1+c)\leqslant 2n/3$ and that $\delta\ll\eta$. Hence,

$$
\begin{aligned}
\frac{s_\ell}{m_\ell}&\leqslant\frac{c\ell+\eta n}{n-\ell-\delta n}\\
&=\frac{c\ell-\delta n}{n-\ell-\delta n}+\frac{(\delta+\eta)n}{n-\ell-\delta n}\\
&\leqslant\frac{c\ell}{n-\ell}+\frac{(\delta+\eta)n}{n-\ell-\delta n}\\
&\leqslant\frac{c\ell}{n-\ell}+\frac{(\delta+\eta)n}{n/3-\delta n}\\
&\leqslant\frac{c\ell}{n-\ell}+4\eta\ .
\end{aligned}
$$ □

Lemma 22 below proves that if all stages satisfy a certain condition, then all stages are successful. The next lemma describes this condition and proves that it holds with positive probability.

Lemma 21. *Let $\mathcal{A}$ denote the intersection of the two following events:*

(i) $M(\pi_t)$ is good for all stages $1 \leqslant t \leqslant k-1$.

(ii) For all $1 \leqslant t \leqslant k-1$, for all $W \subseteq V_{t+1}$ with $|W| = \lfloor ct \rfloor$, and for all $Q \subset [n]$ with $|Q| = q$ there exists $T^* \subseteq \{1, \dots, t\}$ with $|T^*| \geqslant t - \delta n$ such that for all $\ell \in T^*$, a $(Q, \ell, W, r_{\ell,t})$-intersection exists.

Then $\mathcal{A}$ holds with positive probability.

Proof. By Corollary 15, (i) holds with probability at least $\frac{1}{2}e^{-C}$. As for (ii), observe that the number of choices for Q is less than 2^n and that for a given t, the number of possible choices for W is less than 2^n. Hence, the number of possible triples (t, Q, W) is less than $n4^n$. For each such triple, the probability that there exists T^* as specified is at least $1-5^{-n}$ by Lemma 19. So, the probability that (ii) is satisfied is at least $1 - n4^n5^{-n} > 1 - 1/n$. Hence the event $\mathcal{A}$ holds with probability at least $\frac{1}{2}e^{-C} - 1/n > 0$. □

Lemma 22. *If $\mathcal{A}$ holds then all stages are successful.*

Proof. Assume by contradiction that $\mathcal{A}$ holds, yet there is an unsuccessful stage t. First observe that we must have $t \geqslant \eta n/3$, since by Lemma 16, for smaller t, Stage t is successful as $M(\pi_t)$ is good (since we assume that $\mathcal{A}$ holds). Now, the only reason that Stage t is unsuccessful is because the $(B_t, M'(\pi_t), s_t)$-reshuffling is not successful. Hence, by Lemma 17, there exists $W \subset V_{t+1}$ with $|W| = \lfloor ct \rfloor$ such that $|N_{B_t}(W)| < n - \lfloor \eta n/7 \rfloor$. This means that there are at least $q = \lfloor \eta n/7 \rfloor$ elements of F_t that are non-neighbors in B_t of all elements of W. Let, therefore, $Q \subset [n]$ with $|Q| = q$ be such that $F_t(Q) = \{I(t,i) \,|\, i \in Q\} \subset F_t$ and for each pair $(I(t,i), y)$ where $I(t,i) \in F_t(Q)$ and $y \in W$, $I(t,i)$ and y are non-adjacent in B_t. Being non-adjacent in B_t means that for each $y \in W$ and for each $I(t,i) \in F_t(Q)$, there is at least one vertex $x \in I(t,i)$ such that xy is an edge of G. Let $P = \cup_{i \in Q} I(t,i)$ and let $P^* \subseteq P$ be such that $x \in P^*$ if and only if there exists $y \in W$ such that xy is an edge of G. So we have that $|P| = qt$ and

$$|P^*| \geqslant q|W| = q\lfloor ct \rfloor \geqslant q(ct-1) \tag{2}$$

since every $y \in W$ has at least q neighbors in G that belong to P^*, and these neighbors are unique for y since G is an $[n, k, 1]$-graph (every vertex in P^* has exactly one neighbor in W). We will now count the number of elements of P^* in another way, and obtain that it is less than $q(ct-1)$, hence reaching a contradiction with (2).

For every $1 \leqslant \ell \leqslant t$, let $P_\ell = P \cap V_\ell$ and $P^*_\ell = P^* \cap V_\ell$. Clearly $|P_\ell| = q$ so $|P^*_\ell| \leqslant q$, but we can obtain a better upper bound for $|P^*_\ell|$ for most values of ℓ. In order to show this, recall that as we assume that $\mathcal{A}$ holds, there exist $T^* \subseteq \{1, \dots, t\}$ with $|T^*| \geqslant t - \delta n$ such that for all $\ell \in T^*$, a $(Q, \ell, W, r_{\ell,t})$-intersection exists.

Consider first $\ell \in T \setminus T^*$. For such ℓ we will use the trivial bound $|P^*_{\ell+1}| \leqslant q$ but since there are at most δn such ℓ, we have that

$$\sum_{\ell \in T \setminus T^*} |P^*_{\ell+1}| \leqslant \delta n q \,.$$

Consider next $\ell \in T^*$ in which case a $(Q, \ell, W, r_{\ell,t})$-intersection exists. First observe that by the definition of $P_{\ell+1}$, when we perform Stage ℓ extending F_ℓ to $F_{\ell+1}$, the elements of $F_\ell(Q)$ are matched with the vertices of $P_{\ell+1}$ to obtain $F_{\ell+1}(Q)$. Let us recall what it means for our Q, W that a $(Q, \ell, W, r_{\ell,t})$-intersection exists. By Definition 18 applied to Stage ℓ, Q and W, we see that Case (i) of that definition holds since $\mathcal{A}$ holds. Now, Cases (ii) and (iii) of Definition 18 say that when we extend $F_\ell(Q)$ to $F_{\ell+1}(Q)$, at least $r_{\ell,t}$ elements of $F_\ell(Q)$ are matched to vertices of $Y_{\ell+1}(W)$ and notice that $Y_{\ell+1}(W) \cap P^*_{\ell+1} = \emptyset$ since the elements in $Y_{\ell+1}(W)$ are non-neighbors in G of the vertices of W while the elements of $P^*_{\ell+1}$ are neighbors in G of the vertices of W. Hence we have

$$|P_{\ell+1} \setminus P^*_{\ell+1}| \geqslant r_{\ell,t}$$

or equivalently,

$$|P^*_{\ell+1}| \leqslant q - r_{\ell,t} \leqslant q - \frac{q(n-ct-\ell)}{n}\left(1 - \frac{c\ell}{n-\ell}\right) + 5\eta q \ .$$

Altogether, we have

$$|P^*| = \sum_{\ell=1}^{t} |P^*_\ell| \leqslant |P_1| + \delta n q + \sum_{\ell=1}^{t-1}\left(q - \frac{q(n-ct-\ell)}{n}\left(1 - \frac{c\ell}{n-\ell}\right) + 5\eta q\right) \ .$$

Observe that in the last inequality we are summing over all $1 \leqslant \ell \leqslant t-1$ although it would have been enough to sum over the elements of T^*. To obtain the desired contradiction to (2) it therefore suffices to prove that

$$q + \delta n q + \sum_{\ell=1}^{t-1}\left(q - \frac{q(n-ct-\ell)}{n}\left(1 - \frac{c\ell}{n-\ell}\right) + 5\eta q\right) < q(ct-1) \ .$$

Dividing both sides by q, our remaining task is to prove

$$\delta n + \sum_{\ell=1}^{t-1}\left(1 - \frac{(n-ct-\ell)}{n}\left(1 - \frac{c\ell}{n-\ell}\right) + 5\eta\right) < ct - 2 \ .$$

Rearranging the terms, it suffices to show that

$$\sum_{\ell=1}^{t-1}\left(1 - \frac{(n-ct-\ell)}{n}\left(1 - \frac{c\ell}{n-\ell}\right)\right) < ct - 2 - \delta n - 5\eta t \ .$$

Recall that $t \geqslant \eta n/3$ and that $\delta \ll \eta$ so $ct - 2 - \delta n - 5\eta t > t(c - 6\eta)$. Now, define $\mu = t/n$ and define $x = \ell/n$. So it suffices to prove that

$$\frac{1}{n}\left[\sum_{x \in \{\frac{\ell}{n} \,|\, 1 \leqslant \ell \leqslant t-1\}}\left(1 - (1 - c\mu - x)\left(1 - \frac{cx}{1-x}\right)\right)\right] \leqslant \mu(c - 6\eta).$$

As the (real) function $f(x) = 1-(1-c\mu-x)\left(1-\frac{cx}{1-x}\right)$ is monotone increasing in $[0,\mu]$ and $f(x) \in [0,1]$ in that range, the sum in square brackets is at most n times the corresponding integral of $f(x)$ in the range $[0,\mu]$, hence it suffices to prove that

$$\int_0^{\mu} \left[1-(1-c\mu-x)\left(1-\frac{cx}{1-x}\right)\right] dx \leqslant \mu(c-6\eta).$$

But now $\mu = t/n \leqslant k/n \leqslant \frac{1-\epsilon}{1+c}$ and $\eta \leqslant \gamma/6$ so the last inequality indeed holds by Lemma 7. □

From the last two lemmas we immediately obtain:

Corollary 23. *With positive probability, all stages are successful.* □

4 $f(4) = 4$

Here we provide a computer-assisted proof that $f(4) = 4$. Recall from the introduction that $f(4) \geqslant 4$ and recall from the first paragraph in Section 2 that to prove $f(4) \leqslant 4$ it suffices to prove that every $[4,4,1]$-graph has a factor of independent transversals.

Consider the set of $[4,4,1]$-graphs with vertex parts $V_i = \{x_{i,1}, x_{i,2}, x_{i,3}, x_{i,4}\}$ for $1 \leqslant i \leqslant 4$. As adding edges only makes the problem more difficult, we may assume that each of the 6 pairs (V_i, V_j) with $1 \leqslant i < j \leqslant 4$ induces a perfect matching $M_{i,j}$. By relabeling vertices, we may assume that the perfect matching (V_1, V_j) for $j = 2,3,4$ is

$$M_{1,j} = \{(x_{1,1}, x_{j,1}), (x_{1,2}, x_{j,2}), (x_{1,3}, x_{j,3}), (x_{1,4}, x_{j,4})\}$$

as the 12 edges of these three matching are a spanning forest. Thus, the set of graphs that we must construct and check correspond to constructing the remaining three perfect matchings $M_{2,3}$, $M_{2,4}$, $M_{3,4}$. There are $4! = 24$ choices for each of them, each corresponding to a permutation of one of the vertex classes involved in the matching. Overall, there are $24^3 = 13824$ graphs constructed this way (some may be isomorphic).

Checking each constructed graph for a factor of independent transversals proceeds as follows. Take three permutations π_2, π_3, π_4 of V_2, V_3, V_4 respectively, and check if the following factor $\{\{x_{1,j}, x_{2,\pi_2(j)}, x_{3,\pi_3(j)}, x_{4,\pi_4(j)}\} \mid j = 1,2,3,4\}$ is a factor of independent transversals. Note that there are only $(4!)^3 = 13824$ choices for π_2, π_3, π_4, so there are only 13824 checks to perform for each of the 13824 constructed graphs. It turns out that each of these graphs has a factor of independent transversals. A simple program implementing this search and verifying its conclusion appears in Appendix A.

References

[1] Ron Aharoni, Eli Berger, and Ran Ziv. Independent systems of representatives in weighted graphs. *Combinatorica*, 27(3):253–267, 2007.

[2] Noga Alon. The strong chromatic number of a graph. *Random Structures & Algorithms*, 3(1):1–7, 1992.

[3] Béla Bollobás, Paul Erdős, and Endre Szemerédi. On complete subgraphs of r-chromatic graphs. *Discrete Mathematics*, 13(2):97–107, 1975.

[4] Paul Catlin. On the Hajnal-Szemerédi Theorem on disjoint cliques. *Utilitas Math*, 17:163–177, 1980.

[5] Paul Erdős, András Gyárfás, and Tomasz Łuczak. Independent transversals in sparse partite hypergraphs. *Combinatorics, Probability & Computing*, 3(3):293–296, 1994.

[6] Eldar Fischer. Variants of the Hajnal-Szemerédi Theorem. *Journal of Graph Theory*, 31(4):275–282, 1999.

[7] Stefan Glock and Benny Sudakov. An average degree condition for independent transversals. arXiv:2003.01683, 2020.

[8] Penny E. Haxell. On the strong chromatic number. *Combinatorics, Probability & Computing*, 13(6):857, 2004.

[9] Ross J. Kang and Tom Kelly. Colourings, transversals and local sparsity. *Random Structures & Algorithms*, to appear, https://doi.org/10.1002/rsa.21051.

[10] Peter Keevash and Richard Mycroft. A multipartite Hajnal-Szemerédi Theorem. *Journal of Combinatorial Theory, Series B*, 114:187–236, 2015.

[11] Daniela Kühn and Deryk Osthus. Embedding large subgraphs into dense graphs. In *Surveys in Combinatorics*, pages 137–167. Cambridge University Press, 2009.

[12] Kyle MacKeigan. Independent coverings and orthogonal colourings. *Discrete Mathematics*, 344(8):112431, 2021.

[13] Csaba Magyar and Ryan R. Martin. Tripartite version of the Corrádi-Hajnal theorem. *Discrete mathematics*, 254(1-3):289–308, 2002.

[14] Ryan Martin and Endre Szemerédi. Quadripartite version of the Hajnal-Szemerédi Theorem. *Discrete mathematics*, 308(19):4337–4360, 2008.

[15] Tibor Szabó and Gábor Tardos. Extremal problems for transversals in graphs with bounded degree. *Combinatorica*, 26(3):333–351, 2006.

[16] Raphael Yuster. Independent transversals and independent coverings in sparse partite graphs. *Combinatorics, Probability & Computing*, 6(1):115–125, 1997.

[17] Raphael Yuster. Independent transversals in r-partite graphs. *Discrete Mathematics*, 176(1-3):255–261, 1997.

A Source code determining $f(4)$

```
const int N = 4; const int F = 24;
int permutations[F][N] = {
  {0,1,2,3},{1,0,2,3},{2,0,1,3},{0,2,1,3},{1,2,0,3},{2,1,0,3},
  {2,1,3,0},{1,2,3,0},{3,2,1,0},{2,3,1,0},{1,3,2,0},{3,1,2,0},
  {3,0,2,1},{0,3,2,1},{2,3,0,1},{3,2,0,1},{0,2,3,1},{2,0,3,1},
  {1,0,3,2},{0,1,3,2},{3,1,0,2},{1,3,0,2},{0,3,1,2},{3,0,1,2} };

int g[N][N][N][N];

bool check()
{
  int p[N]; // p[i] determines the order of vertices of set i
  p[0] = 0;

  for (p[1] = 0; p[1] < F; p[1]++)
    for (p[2] = 0; p[2] < F; p[2]++)
      for (p[3] = 0; p[3] < F; p[3]++)
      {
        bool current = true;
        for (int i = 0; i < N; i++)
          for (int j = 0; j < N; j++)
            for (int k = j + 1; k < N; k++)
              if (g[j][permutations[p[j]][i]][k][permutations[p[k]][i]]==1)
                current = false;
        if (current)
          return true;
      }
  return false;
}

int main()
{
  for (int i = 0; i < N; i++)
    for (int j = 0; j < N; j++)
      for (int k = 0; k < N; k++)
        for (int l = 0; l < N; l++)
          g[i][j][k][l] = 0;
  for (int i = 0; i < N; i++)
    for (int k = 1; k < N; k++)
      g[0][i][k][i] = 1;
  int p[7];
  for (p[1]=0; p[1] < F; p[1]++) //p1 determines matching between sets 1,2
    for (p[2] = 0; p[2] < F; p[2]++) //p2 determines matching between 1,3
      for (p[3] = 0; p[3] < F; p[3]++) //p3 determines matching between 2,3
      {
        /* reset the graph */
        for (int i = 1; i < N; i++)
          for (int j = 0; j < N; j++)
            for (int k = 1; k < N; k++)
```

```
                for (int l = 0; l < N; l++)
                  g[i][j][k][l] = 0;
          g[1][0][2][permutations[p[1]][0]] = 1;
          g[1][1][2][permutations[p[1]][1]] = 1;
          g[1][2][2][permutations[p[1]][2]] = 1;
          g[1][3][2][permutations[p[1]][3]] = 1;
          g[1][0][3][permutations[p[2]][0]] = 1;
          g[1][1][3][permutations[p[2]][1]] = 1;
          g[1][2][3][permutations[p[2]][2]] = 1;
          g[1][3][3][permutations[p[2]][3]] = 1;
          g[2][0][3][permutations[p[3]][0]] = 1;
          g[2][1][3][permutations[p[3]][1]] = 1;
          g[2][2][3][permutations[p[3]][2]] = 1;
          g[2][3][3][permutations[p[3]][3]] = 1;
          if (!check())
            printf("failed\n,");
        }
    printf("Ended\n");
    return 0;
}
```