Rainbow matchings and cycle-free partial transversals of Latin squares

András Gyárfás\textsuperscript{a}, Gábor N. Sárközy\textsuperscript{a,b,*}

\textsuperscript{a} Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, P.O. Box 127, Budapest, H-1364, Hungary
\textsuperscript{b} Computer Science Department, Worcester Polytechnic Institute, Worcester, MA, 01609, USA

\textbf{Abstract}

In this paper we show that properly edge-colored graphs $G$ with $|V(G)| \geq 4\delta(G) - 3$ have rainbow matchings of size $\delta(G)$; this gives the best known bound for a recent question of Wang. We also show that properly edge-colored graphs $G$ with $|V(G)| \geq 2\delta(G)$ have rainbow matchings of size at least $\delta(G) - 2\delta(G)^2/3$. This result extends (with a weaker error term) the well-known result that a factorization of the complete bipartite graph $K_{n,n}$ has a rainbow matching of size $n - o(n)$, or equivalently that every Latin square of order $n$ has a partial transversal of size $n - o(n)$ (an asymptotic version of the Ryser–Brualdi conjecture). In this direction we also show that every Latin square of order $n$ has a cycle-free partial transversal of size $n - o(n)$.

\section{Introduction—Rainbow matchings in proper colorings}

In this paper we consider properly edge-colored graphs, i.e. two edges with the same color cannot share an endpoint, so each color class is a matching. A matching is rainbow if its edges have different colors. The minimum degree of a graph is denoted by $\delta(G)$. Recently, Wang [7] proposed to find the largest guaranteed size of a rainbow matching in terms of the minimum degree in a properly edge-colored graph. More precisely, [7] raised the following problem.

\textbf{Problem 1.} Does there exist a function $f$ such that $|V(G)| \geq f(\delta(G))$ implies that a properly edge-colored graph $G$ contains a rainbow matching of size $\delta(G)$?

Positive answers to Problem 1 were given in [8,2]. In [2] it was proved that $f(k) = \frac{98k^2}{27}$ is sufficient. In this paper we give a better bound, namely $4\delta(G) - 3$.

\textbf{Theorem 1.} Every properly edge-colored graph $G$ with at least $4\delta(G) - 3$ vertices contains a rainbow matching of size $\delta(G)$.

We note that after submission we learned that Lo and Tan [4] independently proved Theorem 1.

Wang notices that the "best" function in his problem must be greater than $2\delta(G)$ because certain Latin squares have no transversals. For $\delta(G) \in \{2, 3\}$ Theorem 1 is best possible, as shown by a properly 2-edge-colored $C_4$ and by two vertex disjoint copies of a 1-factorization of $K_4$. Our next result shows that if $|V(G)| \geq 2\delta(G)$, then we can find a rainbow matching almost as large as the desired $\delta(G)$.
Theorem 2. Every properly edge-colored graph $G$ with at least $2\delta(G)$ vertices contains a rainbow matching of size at least $\delta(G) - 2(\delta(G))^{2/3}$.

Theorem 2 relates to partial transversals of Latin squares. A Latin square of order $n$ is an $n \times n$ array $[a_{ij}]$ in which each symbol occurs exactly once in each row and exactly once in each column. A partial transversal of a Latin square is a set of distinct symbols, with at most one in each row or column. Latin squares can be also viewed as 1-factorizations of the complete bipartite graph $K_{n,n}$, by mapping rows and columns to vertex classes $R, C$ of $K_{n,n}$ and considering the symbol $[a_{ij}]$ to be the color of the edge $ij$, for $i \in R$ and $j \in C$. The color classes then form a 1-factorization of $K_{n,n}$, and partial transversals become rainbow matchings. A well-known conjecture of Ryser [5] states that for odd $n$ every 1-factorization of $K_{n,n}$ has a rainbow matching of size $n$. The companion conjecture, attributed to Brualdi, is that for every $n$, every 1-factorization of $K_{n,n}$ has a rainbow matching of size at least $n - 1$. These conjectures are known to be true in an asymptotic sense, i.e., every 1-factorization of $K_{n,n}$ has a rainbow matching containing $n - o(n)$ symbols. For the $o(n)$ term, Woodbright [9] and independently Brouwer et al. [1] proved $\sqrt{n}$. Shor [6] improved this to $5.518(\log n)^2$, but his proof had an error, which was corrected in [3]. Theorem 2 extends these results in two senses. It allows proper colorings (instead of factorizations) of arbitrary graphs (instead of complete bipartite graphs). The price we pay is that our error term is weaker than the logarithmic one of Hatami and Shor [3].

We also prove that Latin squares have a large partial transversal without short cycles. A cycle of length $l$ in a Latin square $L$ is a set $(i_1, j_1), (i_2, j_2), \ldots, (i_l, j_l)$ of positions such that $j_1 = i_2, j_2 = i_3, \ldots, j_l = i_1$ and no row index or column index is repeated. For example, a cycle of length 1 is a diagonal position in $L$, a cycle of length 2 is a pair of positions symmetric to the main diagonal, etc. In the complete bipartite graph formulation, considering C as a permutation $\pi$ of $K$, a cycle of length $l$ is an $l$-cycle of $\pi$. There is another reformulation of Latin squares, explaining further the notion of cycles; in fact, Theorem 3 will be proved in that form. Associate the symbol $a_{ij} \in L$ as a color to the edge $ij$ of the complete directed graph $K_n$, where we have a loop $(i, i)$ at each vertex $i$ and joining any two distinct vertices $i$ and $j$ there are two oriented edges $(i, j)$ and $(j, i)$. Then we have an $n$-coloring on the edges of $\overrightarrow{K_n}$, where each color class is a 1-regular subgraph and a partial transversal becomes a rainbow subgraph with maximum indegree and outdegree 1. In this representation an $l$-cycle indeed becomes a directed cycle of length $l$.

Theorem 3. For positive integers $n$ and $k \geq 2$, every Latin square of order $n$ has a partial transversal with at least $n - 6n^{k+1}$ elements containing no cycle of length at most $k$.

Using the reformulation above, Theorem 3 has the following form.

Theorem 4. For positive integers $n$ and $k \geq 2$, consider an edge coloring of $\overrightarrow{K_n}$ with $n$ colors where each color class is a 1-regular digraph. Then there is a rainbow subgraph with maximum indegree and outdegree 1 that has at least $n - 6n^{k+1}$ edges and has no directed cycle of length at most $k$.

Applying Theorem 3 with

$$k = \left\lfloor \frac{\log n}{3 \log \log n} \right\rfloor,$$

there is a partial transversal with at least $n - 6n^{k+1}$ elements that does not contain a cycle of length $l$ for $l \leq k$. From each cycle (of length at least $k+1$) we remove an arbitrary element of the transversal. In the resulting partial transversal we have at least

$$n - \frac{n}{k+1} - 6n^{k+1} \geq n - \frac{3n \log \log n}{\log n} - 6n^{k+1} \geq \left(1 - \frac{4 \log \log n}{\log n}\right)n$$

elements and we get the following.

Corollary 1. Every Latin square of order $n$ has a partial transversal $T$ of order $\left(1 - \frac{4 \log \log n}{\log n}\right)n$ such that $T$ has no cycles at all.

Notice that the error term in the corollary is much worse than in Theorem 2. It is possible that the corollary holds in the following strong form (in the spirit of the Ryser–Brualdi conjecture).

Conjecture 1. Any Latin square of order $n$ has a cycle-free partial transversal of order $n - 2$.

Conjecture 1 would be best for $n = 4$, as shown by the Latin square $L$ with rows 1234, 2143, 3412, 4321. (Any use of symbol 1 would form a cycle of length 1, and only two of (2, 3, 4) can be selected to avoid a 3-cycle.)

2. Proofs

2.1. Proof of Theorem 1

Consider a properly edge-colored graph $G$ with $|V(G)| \geq 4\delta(G) - 3$. Let $c(e)$ denote the color of edge $e$ and let $\delta = \delta(G)$. We start from a “good” configuration $H = M_1 \cup M_2 \cup M_3 \cup \bigcup_{i=1}^{6} F_i$ defined as follows.
• For some integer \( k \geq 0 \), \( M_1 = \{ e_i : i = 1, 2, \ldots, k \} \) and \( M_2 = \{ f_j : i = 1, 2, \ldots, k \} \) form two vertex disjoint rainbow matchings in \( G \), \( c(e_i) = c(f_j) \).

• \( M_3 = \{ g_k : i = k + 1, \ldots, \delta - 1 \} \) is a rainbow matching, vertex disjoint from \( M_1 \cup M_2 \) and \( c(g_i) \neq c(e_j) \) for \( 1 \leq j \leq k \) and \( k + 1 \leq i \leq \delta - 1 \). Thus \( M_1 \cup M_3 \) (likewise \( M_2 \cup M_3 \)) is a rainbow matching of size \( \delta - 1 \).

• \( F_1 = \{ h_t : i = k + 1, \ldots, t_1 \} \) is a matching, vertex disjoint from \( M_1 \cup M_2 \), and \( h_t \cap M_3 = \{ v_i \} \in g_i \). Moreover, \( c(h_{k+1}) \notin c(e) : e \in M_1 \cup M_3 \) and for \( t_1 \geq i > k + 1 \),

\[
c(h_i) \in \bigcup_{k+1 \leq t \leq i} c(g_t).
\]

We call \( F_1 \) a chain. Note that \( F_1 \) is not necessarily rainbow, for example, \( c(h_i) = c(g_{k+1}) \) for \( k + 1 < i \leq t_1 \) satisfies the definition.

• We allow several further disjoint chains \( F_2, \ldots, F_t \), where for \( s \geq j \geq 2 \), \( F_j = \{ h_t : i = t_j - 1, \ldots, t_j \} \) is a matching, vertex disjoint from \( M_1 \cup M_2 \), and \( h_t \cap M_3 = \{ v_i \} \in g_i \). Moreover, as before, \( c(h_{t_j+1}) \notin \{ c(e) : e \in M_1 \cup M_3 \} \) and for \( t_j \geq i > t_{j-1} + 1 \),

\[
c(h_i) \in \bigcup_{t_{j-1} + 1 \leq i \leq t_j} c(g_t).
\]

One can easily see that a good configuration exists. Indeed, by induction there is a rainbow matching \( M \) with \( \delta - 1 \) colors. Let \( v \) be a vertex not in \( V(M) \), and select an edge \( uv \) of \( G \) such that \( c(uv) \notin \{ c(e) : e \in M \} \). If \( w \) is not in \( V(M) \), then \( uv \) extends \( M \) to a rainbow matching of size \( \delta \) and the proof is finished. Otherwise with \( k = 0, t_1 = 1, M_1 = M_2 = \emptyset, M_3 = M \) and \( F_1 = \{ uv \} \), we have a good configuration.

Select a good configuration \( \mathcal{H} \) with the largest possible \( k \). Then select maximal chains \( F_1, F_2, \ldots, F_t \) to cover the maximum number of vertices of \( M_3 \) by \( \cup_{i=1}^t F_i \). If \( k = \delta - 1 \), i.e., \( M_3 = F_1 = \emptyset \), then select any vertex \( v \) not in \( V(H) \) and an edge \( vw \) such that \( c(uv) \notin \{ c(e) : e \in M_1 \} \). Since every color appears in both \( M_1 \) and \( M_2 \), we find a rainbow matching of size \( \delta \). Thus we may assume \( k < \delta - 1 \). Recall that \( v_i = g_i \cap h_i \) for \( i \in \{ k + 1, \ldots, t_1 \} \). Consider a vertex \( v \notin V(H) \) and an edge \( e = vw \) such that \( c(uv) \notin \{ c(f) : f \in M_1 \cup \{ v_i \} \} \) and \( w \neq v_i \) for \( i \in \{ k + 1, \ldots, t_1 \} \). There is such an \( e \) since we have precisely \( \delta - 1 \) restrictions on the choice of \( w \) and since \( \delta(v) \geq \delta \).

**Case 1.** \( w \in (M_1 \cup V(M_2)) \). Note that \( c(uv) \notin \{ c(f) : f \in M_1 \} \) by construction. If \( f = c(uv) \notin \{ c(f) : f \in M_1 \cup M_2 \} \), then we obtain a rainbow matching of size \( \delta \) by adding \( uv \) to \( M_2 \cup M_3 \) if \( w \notin V(M_1) \) (similarly if \( w \in V(M_2) \)). Otherwise, the choice of \( w \) yields \( j = c(uv) = c(g_i) \) for some \( i \) with \( t_i + 1 > i > k \). We can now define a rainbow matching of size \( \delta \) as follows: for \( 1 \leq i \leq k \) take either \( e_i \) or \( f_i \) so that the selected edges are not incident to \( v \) or \( w \). This gives a matching with colors 1, \ldots, \( k \) and color \( j \). Remove (the \( j \)-colored) \( g_i \) from \( M_3 \) and add \( h_i \) (from the chain \( F_i \) covering \( v_i \)). By definition of the chain, the color \( c(h_i) \) equals \( c(g_i) \) for some \( i \) with \( t_i + 1 \leq i \leq t_1 \). Remove \( g_i \) and add \( h_i \) from \( F_i \) continue the procedure. Eventually we add \( h_{t_i+1} \), and the resulting matching is a rainbow matching of size \( \delta \).

**Case 2.** \( w \in \bigcup_{i=1}^t V(f_i) \). For \( c(uv) \notin \{ c(f) : f \in M_1 \cup M_2 \} \) (i.e., \( c(uv) \notin \{ c(f) : f \in M_3 \} \)), this contradicts the choice of \( k \), since \( uv \) can be added to the matching \( M_1 \cup M_2 \cup M_3 \) to get a new repeated color. For \( c(uv) \notin \{ c(f) : f \in M_1 \cup M_2 \} \) we can add \( uv \) to \( M_2 \cup M_3 \) to get a matching of size \( \delta \).

**Case 3.** \( w \in \bigcup_{i=k+1}^t V(g_i) \), say \( g_i = wv_i \) (by the choice of \( c(uv) \), \( w \neq v_i \)). Since \( v_i \) is in some chain \( F_i \), we can add the edge \( uv \), delete \( g_i \), and add \( h_i \) to \( F_i \) repeatedly until we end up adding an edge of the chain \( F_i \) that has color not in the color set \( \{ c(f) : f \in M_1 \cup M_2 \} \). Thus we get either a new good configuration with \( \delta - 1 \) colors in which the color \( c(uv) \) is repeated or a matching with at least \( \delta - 1 \) colors. The latter case finishes the proof, and the former contradicts the choice of \( k \).

**Case 4.** \( w \in \bigcup_{i=t_1+1}^t V(g_i) \). This contradicts the maximality of the chain cover. Indeed, if \( c(uv) \notin \{ c(f) : f \in M_1 \cup M_3 \} \) then we can start a new chain, and if \( c(uv) \in \{ c(f) : f \in M_3 \} \) then we can continue an existing chain.

Since the good configurations involved have at most \( 4(\delta - 1) \) vertices and one further vertex \( w \) is required to get the rainbow matching of size \( \delta(G) \) from it, \( |V(G)| \geq 4\delta(G) - 3 \) is indeed a sufficient condition.

\[ \Box \]

2.2. **Proof of Theorem 2**

Let \( M_1 = \{ e_1, \ldots, e_k \} \) be a largest rainbow matching in a properly edge-colored graph \( G \). Assume to the contrary that \( k < \delta(G) - 2(\delta(G))^3/2 \). Set \( \delta = \delta(G), R = V(G) \setminus V(M_1) \) and let \( C_1 \) be the set of “unused” colors, i.e., colors not used on \( M_1 \). Since \( |V(G)| \geq 2\delta(G) \) we have

\[
|R| > 4\delta^{2/3}.
\]  \hspace{1cm} (1)

Select an arbitrary \( v \in R \). Since \( \deg(v) \geq \delta \) and \( M_1 \) is maximum, at least \( \delta - k > 2\delta^{2/3} \) edges in colors \( C_1 \) go from \( v \) to \( M_1 \):

\[
\deg_{C_1}(v, V(M_1)) > 2\delta^{2/3}.
\]  \hspace{1cm} (2)

Indeed, otherwise we could increase the size of our matching \( M_1 \). This implies in particular that \( \delta - 2\delta^{2/3} > \delta^{2/3} \) (indeed, since \( |M_1| \) has to be at least \( \delta^{2/3} \), so \( |M_1| = k \geq \delta^{2/3} \) but on the other hand we have \( k < \delta - 2\delta^{2/3} \)), i.e.

\[
\delta^{1/3} > 3.
\]  \hspace{1cm} (3)
Furthermore, this also implies that for the number of edges in $C_1$ between $M_1$ and $R$ we have the following lower bound

$$|E_{C_1}(R, V(M_1))| > 2\delta^{2/3}|R|.$$ (4)

In order to define the sets $M_2$ and $C_2$ in the next iteration we do the following. We partition the edges $e_i$ in $M_1$ into two classes. We put $e_i = x_iy_i$ into $M'_i$ if and only if

$$\deg_{C_1}(x_i, R) + \deg_{C_1}(y_i, R) \geq 4\delta^{1/3},$$ (5)

which is greater than 12 by using (3).

We define $M_2 = M_1 \setminus M'_i$ and $C_2 = C_1 \cup \{c(e_i): e_i \in M'_i\}$, where again $c(e_i)$ denotes the color of edge $e_i$. We have the following two crucial claims.

**Claim 1.** $|M'_1| \geq \frac{\delta^{2/3}}{2}$, i.e. $|M_2| \leq |M_1| - \frac{\delta^{2/3}}{2}$.

Otherwise, using (1) we get

$$|E_{C_1}(R, V(M_1))| \leq |M'_1|(2|R|) + |M_2|(4\delta^{1/3}) < \delta^{2/3}|R| + 4\delta^{4/3} < 2\delta^{2/3}|R|,$$

which contradicts (4).

**Claim 2.** For every vertex $v \in R$, we have

$$\deg_{C_2}(v, V(M_2)) > 2\delta^{2/3}.$$

For the proof of this claim observe first that if $e_i = x_iy_i \in M'_i$, then all $C_1$-edges to $R$ incident to this edge must be incident to the same endpoint (say $x_i$) since otherwise we could increase $M_1$ (using (5)). Denote by $X_i$ the set of these $x_i$ endpoints in $M'_i$ and by $Y_i$ the set of other endpoints. Thus there is no $C_1$-edge between $Y_i$ and $R$, and for every $x_i \in X_i$ there is at least $4\delta^{1/3}$-edges from $x_i$ to $R$.

Consider an arbitrary $v \in R$ and an edge $vw$ with $c(vw) \in C_2$. First note that $w \not\in R$. Otherwise, if $c(vw) \in C_1$, then we can clearly increase $M_1$; if $c(vw) = c(e_i)$ for some $e_i \in M'_i$, then we can increase $M_1$ by exchanging $e_i$ with $vw$ and adding a $C_1$-edge from $x_i$ to a free neighbor in $R$ (using (5) again). Here and henceforth, a free neighbor is a neighbor that is not covered by the current matching, so in particular here it is a neighbor in $R$ outside $\{v, w\}$.

Thus $w \in V(M_1)$. Next we show that $w \not\in Y_i$. Assume otherwise that $w = y_j$ for some $y_j \in Y_i$. If $c(vw) \in C_1$, then again we can increase $M_1$ by exchanging $e_i$ with $vw$ and adding another $C_1$-edge from $x_i$ to a free neighbor in $R$ such that this edge has a different color from $c(vw)$ (using (5)). If $c(vw) = c(e_i)$ for some $e_i \in M'_i$, then we could increase $M_1$ by deleting $e_i$ and $v_j$, adding $vw$ and adding one $C_1$-edge from $x_i$, one $C_1$-edge from $x_j$ to free neighbors in $R$ such that the two edges have different colors.

Thus if $w \in V(M'_i)$, then $w \in X_i$ and this implies Claim 2, since by using (2) we get

$$\deg_{C_2}(v, V(M_2)) \geq \deg_{C_1}(v, V(M_1)) + |M'_1| - |M'_i| > 2\delta^{2/3}.$$  

Indeed, we get $\deg_{C_1}(v, V(M_1)) + |M'_1|$ from the $C_1$-edges and the edges using colors from $M'_1$, and there is at most one $C_2$-edge from $v$ to each edge in $M'_i$.

Suppose now that $M_j$ and $C_j$ are already defined for $j \geq 2$ such that the two claims are true for $j$, i.e.

$$|M_j| \leq |M_{j-1}| - \frac{\delta^{2/3}}{2},$$ (6)

and

$$\deg_{C_j}(v, V(M_j)) > 2\delta^{2/3}.$$ (7)

In order to define $M_{j+1}$ and $C_{j+1}$ we put the edges $e_i = x_iy_i \in M_j$ into $M'_j$ if and only if

$$\deg_{C_j}(x_i, R) + \deg_{C_j}(y_i, R) \geq 4\delta^{1/3}.$$

We define $M_{j+1} = M_j \setminus M'_j$ and $C_{j+1} = C_j \cup \{c(e_i): e_i \in M'_j\}$.

We have to show that the two claims remain true for $j + 1$. The proof of Claim 1 for $j + 1$ is identical (replacing indices 1, 2 by $j, j + 1$). The proof of Claim 2 for $j + 1$ is also similar, but we will have longer exchange sequences. First we show again that if $e_i = x_iy_i \in M'_i$, then all $C_j$-edges to $R$ must be incident to the same endpoint (say $x_i$). Assume otherwise that we have two $C_j$-edges of different colors $x_iy_i$ and $y_1v_2$, where $v_1, v_2 \in R$.

We “trace back” both edges to a $C_1$-edge. If $c(x_iy_i) \in C_1$ (and similarly for $y_1v_2$), then we are done. Otherwise, by definition, there exists $j_1 < j$ such that there exists an edge $x_{i_1}y_{i_1} \in M'_{j_1}$ with $c(x_{i_1}y_{i_1}) = c(x_iy_i)$. We find a $C_{j_1}$-edge $x_{i_1}v_1$, such that $v_1$ is a free neighbor of $x_{i_1}$ in $R$. If $c(x_{i_1}v_1) \in C_1$, then we are done. Otherwise, we trace this edge back further until we find an edge $x_{i_2}y_{i_2} \in M'_{j_2}$, for which there is a free $C_{j_2}$-neighbor $v_{i_2}$ of $x_{i_2}$ in $R$. We proceed similarly for $y_1v_2$, but we always select $C_{j_2}$-edges in unused colors to free vertices in $R$. At this point we can define a larger rainbow matching $M^*$ from $M$ by deleting
the edges \(x_i y_i, x_i y_t\) for \(t \in \{1, \ldots, s\}\) and adding the edges \(x_i v_t, x_i v_t\) for \(t \in \{1, \ldots, s\}\) and similarly for \(y_i v_2\). Note that the above procedure succeeds if the number of available neighbors in \(R\) is at least \(2(j + 1)\). Since the number of available neighbors is at least \(4\delta^{1/3}\), the above works for \(j + 1\) as long as \(j + 1 \leq 2\delta^{1/3}\). Let us denote again the set of these \(x_i\) endpoints in \(M_j^t\) by \(X_j\) and the set of other endpoints by \(Y_j\). Thus there is no \(C_j\)-edge between \(Y_j\) and \(R\), and for every \(x_i \in X_j\) there are at least \(4\delta^{1/3} C_j\)-edges from \(x_j\) to \(R\).

Consider again an arbitrary \(v \in R\) and an edge \(vw\) with \(c(vw) \in C_{j+1}\). First we show again that \(w \not\in R\). As above we trace \(vw\) back to a \(C_j\)-edge. If \(c(vw) \in C_1\), then we are done. Otherwise, as above for \(t \in \{1, \ldots, s\}\) we find edges \(x_i y_t, x_i v_t\) and \(x_t y_t, v_t v_t\) in \(R\) such that

\[
c(vw) = c(x_i y_i) \in C_{j+1}, c(x_t y_t) = c(x_i y_{t+1}) \in C_j \quad \text{for} \ t \in \{1, \ldots, s - 1\}
\]

and

\[
c(x_i v_t) \in C_1.
\]

Again we can define a larger rainbow matching \(M^*\) from \(M\) by deleting the edges \(x_i y_t\) for \(t \in \{1, \ldots, s\}\) and adding the edges \(vw\) and \(x_t v_t\) for \(t \in \{1, \ldots, s\}\). Again this works for \(j + 1\) when \(j + 1 \leq 2\delta^{1/3}\).

Thus \(w \in V(M_j^t)\). Finally we show again that \(w \not\in Y_j\). Assume otherwise that \(w = y_t\) for some \(y_t \in Y_j\). As above we can trace back this \(vw\) edge to a \(C_1\)-edge and thus we could increase our matching. Thus if \(w \in V(M_j^t)\), then \(w \in X_j\), and this implies 

Claim 2 for \(j + 1\) (assuming \(j + 1 \leq 2\delta^{1/3}\)), since by using (7) we get

\[
deg_{C_j}(v, V(M_j^t)) \geq \deg_{C_j}(v, V(M_j^t)) + |M_j^t| - |M_j^t| > 2\delta^{2/3}.
\]

Indeed, we get \(\deg_{C_j}(v, V(M_j^t)) + |M_j^t|\) from the \(C_j\)-edges and the edges using colors from \(M_j^t\), and there is at most one \(C_{j+1}\)-edge from \(v\) to each edge in \(M_j^t\).

However, applying Claims 1 and 2 with \(l = \lceil 2\delta^{1/3}\rceil\) we get

\[
\deg_{C_j}(v, V(M_j^t)) > 2\delta^{2/3},
\]

while

\[
|M_j^t| \leq |M_j^t| - (l - 1) \frac{\delta^{2/3}}{2} < \delta - 2\delta^{2/3} - (2\delta^{1/3} - 2) \frac{\delta^{2/3}}{2} = -\delta^{2/3} < 0,
\]

a contradiction. \(\square\)

2.3. Proof of Theorem 4

Consider a coloring of \(\overrightarrow{K_n}\) where each color class is a 1-regular directed graph, we may refer to it as a 1-factorization. Subgraphs of 1-regular digraphs will be called linear digraphs. We start from a rainbow linear digraph \(G_1\) on \(n\) vertices with the most edges that does not contain a directed cycle with length at most \(k\). Let \(t\) be the number of edges in \(G_1\).

We will show that

\[
t \geq n - 6n^{k-1}.
\]

Thus \(G_1\) is a collection of directed cycles with length greater than \(k\), directed paths, and isolated vertices. We consider isolated vertices as paths of length 0. If the directed edge \(uv\) is a part of a cycle or a path in \(G_1\), then we call \(v\) the forward neighbor of \(u\). Thus every vertex other than that of an endpoint of a path has a forward neighbor.

We will define two nested sequences \(A_1 \subseteq A_2 \subseteq \cdots\) and \(B_1 \subseteq B_2 \subseteq \cdots\) of sets. Define \(A_1\) as the set of beginning vertices of the paths and \(B_1\) as the set of end vertices of the paths. We clearly have

\[
|A_1| = |B_1| = n - t.
\]

Consider the edges with the head in \(A_1\) and having one of the \(n - t\) colors not used in \(G_1\), which we call new colors. Denote the set of these edges by \(E_1\). We will designate some edges in \(E_1\) as forbidden edges. For a beginning vertex \(u \in A_1\) of a path \(P\) in \(G_1\) the edge \(vu\) is forbidden if \(v\) is a vertex on the path \(P\) at a distance \(l\) from \(u\), where \(2 \leq l \leq k - 1\). Indeed, these edges may potentially create short rainbow cycles, which are not allowed. Thus altogether we have at most \((k - 2)(n - t)\) forbidden edges in \(E_1\). This implies that there is a new color (denoted by \(c_1\)) that contains at most \(k - 2\) forbidden edges. Let \(f_1 = k - 2 f_2, f_3, \ldots\) will be defined later). Consider those edges in \(E_1\) that have color \(c_1\), and remove the at most \(f_1\) forbidden edges. Denote the resulting edge set by \(E_1^{c_1}\). We have the following claim.

Claim 3. If \(vu \in E_1^{c_1}\) with \(u \in A_1\), then \(v \not\in B_1\).

Indeed, otherwise we would get a rainbow linear subgraph with \(t + 1\) edges that does not contain \(C_t\) with \(l \leq k\) (since the forbidden edges were removed), contradicting that \(t\) was maximum.

Now we are ready to define \(A_2\) and \(B_2\). Since we have a 1-factorization, every vertex is the ending point of an edge colored with \(c_1\), and thus

\[
|E_1^{c_1}| \geq n - t - f_1.
\]
Consider the set \( \{ v : \ vu \in E_i^t, \ u \in A_i \} \). By Claim 3 this is disjoint from \( B_1 \). Define
\[
B_2 = B_1 \cup \{ v : \ vu \in E_i^t, \ u \in A_i \}.
\]
By (8), we have \( |B_2 \setminus B_1| \geq n - t - f_1 \). The set \( A_2 \setminus A_1 \) is the forward neighbors of the vertices in \( B_2 \setminus B_1 \). These forward neighbors all exist since the vertices in \( B_2 \setminus B_1 \) cannot be ends of the paths (those vertices are in \( B_1 \)). Furthermore, clearly this set is disjoint from \( A_1 \), since we are moving away from the beginning vertices of the paths. Finally, define \( G_2 \) as the union of \( G_1 \) and the edge set \( E_i^{t-1} \). If \( G_2 \) contains a rainbow cycle \( C \) of length at most \( k \), then \( C \) contains exactly one edge colored with \( c_1 \) (since \( G_1 \) has no rainbow cycle with length at most \( k \)), but then this edge is forbidden and was removed, a contradiction.

At this point we have the following four properties for \( i = 2 \) (with \( f_1 = k - 2 \)):

1. \( A_{i-1} \subseteq A_1 \text{ and } B_{i-1} \subseteq B_1 \).
2. \( n - t \geq |A_i \setminus A_{i-1}| = |B_i \setminus B_{i-1}| \geq n - t - f_{i-1} \).
3. \( G_i \) does not contain a rainbow cycle with length at most \( k \).
4. For every \( u \in A_i \) there is a linear rainbow subdigraph of \( G_i \) with \( t \) edges such that the set of endpoints of the paths is exactly \( B_1 \) and one of the paths begins at \( u \).

We continue in this fashion, maintaining these properties with a suitable \( f_{i-1} \). Assume that \( A_1, \ldots , A_{i-1} \text{ and } B_1, \ldots , B_{i-1} \) are already defined for some \( i \geq 3 \). Consider the edges with the ending point in \( A_{i-1} \) in one of the \( n - t - (i - 2) \) new colors (not used in \( G_{i-1} \)), denote the set of these edges by \( E_i \). Again we will identify some edges in \( E_i \) as forbidden edges. For a vertex \( u \in A_i \) the edge \( uv \in E_i \) is forbidden if there is a rainbow path of length at most \( k - 1 \) from \( u \) to \( v \) in \( G_{i-1} \), where the last edge is from \( G_1 \). Indeed, these edges may potentially create short rainbow cycles, which are not allowed.

For \( u \in A_{i-1} \), a crude upper bound on the number of these rainbow paths of length at most \( k - 1 \) (and thus for the number of \( uv \in E_i \) forbidden edges) is \((k - 2)t^{k-2} \). Indeed, for each of the edges before the last one we have at most \( i \) possibilities (one from \( G_1 \) and one for each of the \( i - 2 \) added colors), and for the last edge we have only one possibility, since it must be in \( G_1 \).

Thus altogether we have at most \( ki^{k-1}(n - t) \) forbidden edges in \( E_{i-1} \). This implies that there is a new color (denoted by \( c_{i-1} \)) that contains at most
\[
f_{i-1} = \frac{ki^{k-1}(n - t)}{n - t - (i - 2)}
\]
forbidden edges. Consider those edges in \( E_{i-1} \) that have color \( c_{i-1} \) and remove these forbidden edges. Denote the resulting edge set by \( E_i^{t-1} \). We have the following claim.

**Claim 4.** If \( vu \in E_i^{t-1} \) with \( u \in A_{i-1} \), then \( v \not\in B_1 \).

Otherwise from Property 4 the edge \( vu \) would join two paths or create a cycle, yielding a rainbow linear subgraph with \( t + 1 \) edges, contradicting that \( t \) was maximum. We also would have no rainbow \( l \)-cycle with \( l \leq k \), since otherwise this \( l \)-cycle must contain exactly one edge colored with \( c_{i-1} \), namely \( uv \) (since \( G_{i-1} \) does not contain a rainbow \( l \)-cycle with \( l \leq k \)), but then this edge is forbidden and was removed, a contradiction.

Now we are ready to define \( A_i \) and \( B_i \). Since we have a 1-factorization, every vertex is the ending point of an edge colored with \( c_{i-1} \). Define
\[
B_i \setminus B_{i-1} = \{ v : \ vu \in E_i^{t-1}, \ u \in A_i, \ v \not\in B_{i-1} \}.
\]
By Claim 4 we have Property 2 for \( |B_i \setminus B_{i-1}| \), since
\[
|B_i \setminus B_{i-1}| \geq |A_{i-1}| - f_{i-1} - |B_{i-1} \setminus B_i| = |A_{i-1}| - f_{i-1} - |A_{i-1} \setminus A_i| = |A_i| - f_{i-1} = n - t - f_{i-1}.
\]
The set \( A_i \setminus A_{i-1} \) is the forward neighbors of the vertices in \( B_i \setminus B_{i-1} \). These forward neighbors all exist since the vertices in \( B_i \setminus B_{i-1} \) cannot be ends of the paths (those vertices are in \( B_1 \)). Furthermore, this set is indeed disjoint from \( A_{i-1} \) since in \( G_1 \) the in-degree of every vertex is at most 1.

Finally, define \( G_i \) as the union of \( G_{i-1} \) and those edges in \( E_i^{t-1} \) that start in vertices in \( B_i \setminus B_{i-1} \). Notice that Property 4 is maintained from the definition of the vertices of \( A_i \setminus A_{i-1}, B_i \setminus B_{i-1} \).

Property 3 above also holds for \( G_i \). Otherwise, there is a rainbow \( l \)-cycle in \( G_i \) with \( l \leq k \). Then this \( l \)-cycle contains exactly one edge colored with \( c_{i-1} \) (since \( G_{i-1} \) does not contain a rainbow \( l \)-cycle with \( l \leq k \)), but then this edge is forbidden and was removed, a contradiction. Note again that from the construction the last edge is from \( G_1 \) on the rainbow path of length at most \( k - 1 \) connecting the two endpoints of this edge.

Next we claim that
\[
|A_i \setminus A_{i-1}| = |B_i \setminus B_{i-1}| \geq \frac{n - t}{i} \quad \text{for } i \leq \left( \frac{n - t}{4k} \right)^{\frac{1}{k-1}}.
\]
In fact, for such \( i \) Property 2 yields
\[
|A_i \setminus A_{i-1}| = |B_i \setminus B_{i-1}| \geq n - t - \frac{ki^{k-1}(n - t)}{n - t - (i - 2)} = n - t - 2ki^{k-1} \geq \frac{n - t}{2}
\]
for some \( i \).
Thus we must have
\[
\frac{n - t}{2} \left( \frac{n - t}{4k} \right)^{\frac{1}{k-1}} \leq n,
\]
and therefore, using \( k \geq 2 \),
\[
n - t \leq 2(4k)^{\frac{1}{k-1}} n^{\frac{k-1}{k}} \leq 6n^{\frac{k-1}{k}}.
\]
From this we get that
\[
t \geq n - 6n^{\frac{k-1}{k}},
\]
as desired. \( \square \)

Acknowledgment

We appreciate Doug West’s remarks, which improved the presentation.

References