A Hall-Type Condition for Path Covers in Bipartite Graphs

Mikhail Lavrov Jennifer Vandenbussche

Submitted: Oct 10, 2023; Accepted: May 28, 2024; Published: Jul 12, 2024 © The authors. Released under the CC BY-ND license (International 4.0).

Abstract

Let G be a bipartite graph with bipartition (X, Y). Inspired by a hypergraph problem posed by Kostochka et al. (2021), we seek an upper bound on the number of disjoint paths needed to cover all the vertices of X. We conjecture that a Halltype sufficient condition holds based on the maximum value of $|S| - |\Lambda(S)|$, where $S \subseteq X$ and $\Lambda(S)$ is the set of all vertices in Y with at least two neighbors in S. This condition is also a necessary one for a hereditary version of the problem, where we delete vertices from X and try to cover the remaining vertices by disjoint paths. The conjecture holds when G is a forest, has maximum degree 3, or is regular with high girth, and we prove those results in this paper.

Mathematics Subject Classifications: 05C38, 05C65

1 Introduction

1.1 Path covers of bipartite graphs

Problems regarding path covers of graphs are ubiquitous in graph theory. A path cover of G is a collection of vertex-disjoint paths in G where the union of the vertices of the paths is V(G). Certainly the most well-studied example looks for a single path covering all vertices of G, i.e. a Hamiltonian path. Graphs with such a path are also called *traceable*. See [8] for a survey of results in this area. Determining whether a graph has a Hamiltonian path is NP-complete even for very restrictive classes of graphs; for example, Akiyama et al. [1] prove that it is NP-complete for 3-regular bipartite graphs.

In graphs that are not traceable, we may seek a path cover with as few paths as possible. For example, Magnant and Martin [13] conjecture that a *d*-regular graph *G* can be covered with at most |V(G)|/(d+1) paths, and prove this when $d \leq 5$. Feige and Fuchs [7] extend the result to d = 6. In [14], Magnant et al. conjecture that a graph with maximum degree Δ and minimum degree δ needs at most max $\left\{\frac{1}{\delta+1}, \frac{\Delta-\delta}{\Delta+\delta}\right\} \cdot |V(G)|$ paths

Department of Mathematics, Kennesaw State University, Kennesaw, Georgia, USA (mlavrov@kennesaw.edu, jvandenb@kennesaw.edu)

to cover its vertices, which they verify for $\delta \in \{1, 2\}$ and which Kouider and Zamime [11] prove for $\Delta \ge 2\delta$. For dense *d*-regular bipartite graphs, Han [9] proves that a collection of |V(G)|/(2d) vertex-disjoint paths covers all but o(|V(G)|) vertices.

In this paper, we focus on a variant of the path cover problem for bipartite graphs: collections of vertex-disjoint paths that cover one partite set of the bipartite graph. Let an (X, Y)-bigraph be a bipartite graph with a specified ordered bipartition (X, Y). If G is an (X, Y)-bigraph, a path X-cover of G is a set of pairwise vertex-disjoint paths in G that cover all of X.

We seek a Hall-type condition for the existence of a path X-cover of G with at most k paths. For any $S \subseteq X$, let $\Lambda_G(S)$ be the set of all vertices in Y that have at least two neighbors in S; in cases where there is only one graph G under consideration, we will write $\Lambda_G(S)$ simply as $\Lambda(S)$. We define the Λ -deficiency of S to be Λ -def $(G,S) := |S| - |\Lambda(S)|$, and the Λ -deficiency of G to be

 $\Lambda\text{-def}(G) := \max\{\Lambda\text{-def}(G, S) : S \subseteq X\}.$

Note that Λ -def $(G) \ge 1$ for all graphs G, since Λ -def $(G, \{x\}) = 1$ for any vertex $x \in X$. We conjecture the following:

Conjecture 1. Every (X, Y)-bigraph G has a path X-cover by at most Λ -def(G) paths.

If this conjecture holds, then for every $S \subseteq X$, there is a set of at most Λ -def(G) vertex-disjoint paths whose intersection with X is precisely S. To see this, just delete all the vertices in X - S from G, which can only decrease the Λ -deficiency.

Conversely, suppose it is true that for every $S \subseteq X$, there is a set of at most k vertex-disjoint paths whose intersection with X is precisely S. Then for every S, these paths have at least |S| - k internal vertices in Y that are all elements of $\Lambda(S)$; therefore $|\Lambda(S)| \ge |S| - k$ for all S, which implies that Λ -def $(G) \ge k$. It follows that the condition in our conjecture is a *necessary* one if we would like to draw the stronger conclusion in the preceding paragraph.

Our conjecture is a slightly weakened form of a conjecture on cycle covers proposed in [10]:

Conjecture 2 (Conjecture 5 in [10]). Let G be an (X, Y)-bigraph with the property that for all $S \subseteq X$ with $|S| \ge 3$, Λ -def $(G, S) \le 0$ and the induced subgraph between S and $\Lambda(S)$ is 2-connected. Then G contains a cycle that covers all of X.

The statement of Conjecture 2 considers only sets S with $|S| \ge 3$ because these are more natural in the hypergraph setting described in Subsection 1.3. In the case of graphs, it is more natural to start at $|S| \ge 2$, and then the 2-connectivity assumption is unnecessary. This modification to the conjecture was proposed by Salia in [16]:

Conjecture 3 (Conjecture 1.8.23 in [16]). Let G be an (X, Y)-bigraph with the property that for all $S \subseteq X$ with $|S| \ge 2$, Λ -def $(G, S) \le 0$. Then G contains a cycle that covers all of X.

We claim that Conjecture 2 implies Conjecture 3, which in turn implies Conjecture 1. To see the first implication, let G satisfy the hypotheses of Conjecture 3. Then for all $S \subseteq X$ with $|S| \ge 3$, the induced subgraph of G between S and $\Lambda(S)$ is 2-connected: if any vertex is deleted, then any two remaining vertices $x, x' \in S$ still have at least one common neighbor (since $|\Lambda(\{x, x'\}| \ge 2)$ and any vertex in $\Lambda(S)$ still has at least one neighbor in S. Therefore G satisfies the hypotheses of Conjecture 2.

To see the second implication, let G satisfy the hypothesis of Conjecture 1, and let H be the graph obtained from G by adding Λ -def(G) more vertices to Y, each of which is adjacent to every vertex in X. Then for all $S \subseteq X$ with $|S| \ge 3$ (and even with |S| = 2), we have Λ -def $(H, S) \le 0$, since all the new vertices of H are in $\Lambda_H(S)$. Therefore H satisfies the hypotheses of Conjecture 3. Now a cycle in H covering all of X yields a path X-cover of G by at most Λ -def(G) paths by deleting all the new vertices.

1.2 Previous results for these conjectures

In [10], Kostochka et al. prove that Conjecture 2 holds if $|X| \leq 6$, which implies that the other two conjectures hold if $|X| \leq 6$ as well. They also prove that Conjecture 2 holds if all vertices in X have degree at least $\max\{|X|, \frac{|Y|+10}{4}\}$.

Conjecture 3 is considered by Barát et al. in [3], where it is proved in a number of special cases, with the degree of vertices in Y restricted: either if all vertices in Y have degree at most $\frac{1}{4}\sqrt{|X|}$, for |X| sufficiently large, or if all vertices in Y have degree either 2 or |X|. The first of these results does not imply any corollary for Conjecture 1, since the reduction in the previous subsection adds many vertices of degree |X| to Y. However, the second result implies that Conjecture 1 holds if all vertices in Y have degree 2. Barát et al. also prove that if the condition of Conjecture 3 holds, then G has a collection of pairwise vertex-disjoint cycles covering X (but with no restriction on their number).

1.3 Hypergraphs and the Gallai–Milgram theorem

The setting of Conjecture 1 can be translated into the language of hypergraphs and Berge paths in hypergraphs, and here we see the motivation for focusing on a path cover of X.

Following the terminology of Berge [4], a hypergraph H consists of a set of vertices V(H) and a set of edges E(H) where each edge $e \in E(H)$ is a subset of V(H). (We allow edges of any size.) The subhypergraph of H generated by a set $S \subseteq V(H)$ is the hypergraph with V(H) = S and

$$E(H) = \{ e \cap S : e \in E(H), e \cap S \neq \emptyset \}.$$

There are several notions of paths in hypergraphs that generalize paths in graphs. One such notion is that of a *Berge path*: a sequence

$$(v_0, e_1, v_1, e_2, v_2, \dots, e_\ell, v_\ell)$$

where v_0, v_1, \ldots, v_ℓ are distinct vertices in $V(H), e_1, e_2, \ldots, e_\ell$ are distinct edges in E(H), and $\{v_{i-1}, v_i\} \subseteq e_i$ for all $i = 1, \ldots, \ell$. Given a hypergraph H, we can define its *incidence graph* to be the (X, Y)-bigraph G with X = V(H) and Y = E(H) such that $xy \in E(G)$ if and only if $x \in X, y \in Y$, and $x \in y$. Berge paths in H correspond to paths in G that begin and end in X; these are vertex-disjoint in G if and only if they are both vertex-disjoint and edge-disjoint in H.

If we define a *Berge path cover* of the hypergraph H to be a set of pairwise vertexand edge-disjoint paths that cover all of V(H), then Conjecture 1 proposes a sufficient condition for H to have a Berge path cover of size at most k. Moreover, the proposed sufficient condition is a necessary condition for every subhypergraph of H to have a Berge path cover of size at most k.

This statement is reminiscent of the Gallai–Milgram theorem ([6], p. 298 in [4]), which states that the vertices of any directed graph D can be covered by at most $\alpha(D)$ disjoint paths, where $\alpha(D)$ is the independence number of D. (The weaker statement for undirected graphs clearly follows.) For a hypergraph H, let a set $I \subseteq V(H)$ be strongly independent (following the terminology of Berge) if $|e \cap I| \leq 1$ for all $e \in E(H)$; let $\alpha(H)$, the strong independence number of H, be the size of a largest strongly independent set in H. It would be natural to hope that H has a path cover by at most $\alpha(H)$ pairwise-disjoint paths. In [15], Müller proves such a generalization of the Gallai–Milgram theorem (and, in fact, a generalization of it to directed hypergraphs), but in a slightly different setting: Müller does not require the edges of a path to be distinct, and does not require the paths in the cover to be edge-disjoint, merely vertex-disjoint.

In our setting, the corresponding generalization is false. Translating from hypergraphs back into the language of graphs: a set $I \subseteq V(H)$ is strongly independent if and only if, in the incidence graph of H, $\Lambda(I) = \emptyset$. Generalizing to an arbitrary (X, Y)-bigraph G, let $S \subseteq X$ be Λ -independent if $\Lambda(S) = \emptyset$, and let the Λ -independence number $\alpha_{\Lambda}(G)$ be the size of a largest Λ -independent set. Note that if S is Λ -independent, then Λ -def(G, S) =|S|, so $\alpha_{\Lambda}(G)$ is always at most Λ -def(G).

To see that an (X, Y)-bigraph G may not have a path X-cover with at most $\alpha_{\Lambda}(G)$ paths, even if G is balanced and has a high connectivity, consider the following family of examples. Fix an integer k between 1 and n, and let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$ with $x_i y_j \in E(G)$ when $i \leq k$ or $j \leq k$. Then $\alpha_{\Lambda}(G) = 1$, since any two vertices share the neighbor y_1 , but Λ -def(G) = n - 2k + 1 (choose $S = \{x_k, x_{k+1}, \ldots, x_n\}$), and in fact it can be checked that a minimum path X-cover contains n - 2k + 1 paths.

However, in all the cases of Conjecture 1 we consider where G is a regular graph, $\alpha_{\Lambda}(G)$ paths suffice for a path X-cover of G. Whether this holds for all regular bigraphs Gis an open question that would have far-reaching consequences. For example, a result of Singer [17] states that the incidence graph of any classical projective plane is Hamiltonian. The proof relies on algebra over finite fields, but the claim above would give a purely graphtheoretic reason that these incidence graphs are always traceable, since the incidence graph G of any projective plane must have $\alpha_{\Lambda}(G) = 1$.

More generally, a hypergraph H is *covering* if every pair of vertices of H lie on a common edge: in other words, $\alpha(H) = 1$. Lu and Wang [12] prove that every $\{1, 2, 3\}$ -uniform covering hypergraph has a Hamiltonian Berge cycle. This implies Conjecture 1 for (X, Y)-bigraphs G with maximum degree 3 in Y and $\alpha_{\Lambda}(G) = 1$.

1.4 Our results

Our first result states that Conjecture 1 holds for forests:

Proposition 4. If G is an (X, Y)-bigraph with no cycles, then G has a path X-cover of size at most Λ -def(G).

To strengthen Proposition 4, we go in two directions: we consider graphs with low maximum degree and graphs with high girth. In the first case, we begin by proving:

Theorem 5. If G is a 3-regular (X, Y)-bigraph, then G has a path X-cover of size at most $\alpha_{\Lambda}(G)$.

The proof of Theorem 5 begins by taking a 2-factor of G, covering the graph (and, in particular, X) with pairwise vertex-disjoint cycles. If we generalize to graphs with maximum degree 3, we are unable to do this, but if we cover as much of G with cycles as possible, we are left with a forest. Once we deal with the interaction between the forest and the cycles, we can combine the arguments of Theorem 5 with Proposition 4 to prove a result for all graphs with maximum degree 3:

Theorem 6. If G is an (X, Y)-bigraph with maximum degree at most 3, then G has a path X-cover of size at most Λ -def(G).

It is particularly interesting to strengthen Theorem 5 to Theorem 6 because if G has maximum degree at most 3, then so does every subgraph of G. As a result, we obtain a necessary and sufficient condition for an (X, Y)-bigraph G of maximum degree 3 to have the property that for all $S \subseteq X$, there is a set of at most Λ -def(G) pairwise vertex-disjoint paths whose intersection with X is precisely S.

Conjecture 1 holds for regular bigraphs of any degree if we add a condition on the girth of G, that is, the length of the shortest cycle in G.

Theorem 7. Let G be an (X, Y)-bigraph with maximum degree at most d and girth at least $4ed^2 + 1$, and assume that there exists a collection of pairwise vertex-disjoint cycles in G that cover all of X. (In particular, such a collection is guaranteed to exist if G is d-regular.) Then G has a path X-cover of size at most $\alpha_{\Lambda}(G)$.

2 Forests

Proof of Proposition 4. We may assume that G has no leaves in Y, since a vertex in Y of degree 1 does not contribute to Λ -def(G, S) for any S, and it does not help cover more of X by paths. We may also assume that G is a tree; if G has multiple components, we can solve the problem on each component separately.

We induct on |X|. When |X| = 1, we have Λ -def $(G) = \Lambda$ -def(G, X) = 1, and we can cover X by a single path of length 0.

When |X| > 1, consider G as a rooted tree with an arbitrary root in X. Let $x \in X$ be a leaf of G at the furthest distance possible from the root, and let $y \in Y$ be the parent vertex of x.

Case 1: *y* has other children.

Let x_1, \ldots, x_k be all the children of y (including x); by the case, $k \ge 2$. Since x was chosen to be as far from the root as possible, each x_i must be a leaf. Delete x_1, x_2, \ldots, x_k, y from G to get G'.

Let $S \subseteq X - \{x_1, \ldots, x_k\}$ be the set such that Λ -def $(G') = \Lambda$ -def(G', S). We claim that

$$\Lambda\text{-def}(G, S \cup \{x_1, \dots, x_k\}) = \Lambda\text{-def}(G', S) + k - 1.$$

On one hand, $|S \cup \{x_1, \ldots, x_k\}| = |S| + k$. On the other hand, $\Lambda_G(S \cup \{x_1, \ldots, x_k\}) = \Lambda_{G'}(S) \cup \{y\}$, so $|\Lambda_G(S \cup \{x_1, \ldots, x_k\})| = |\Lambda_{G'}(S)| + 1$. In particular, Λ -def $(G) \ge \Lambda$ -def(G', S) + k - 1.

By the inductive hypothesis, G' has a path X-cover by at most Λ -def(G', S) paths. Add k-1 more paths to that set: the path (x_1, y, x_2) and the length-0 paths $(x_3), \ldots, (x_k)$. This is a path X-cover of G by at most Λ -def $(G', S) + k - 1 \leq \Lambda$ -def(G) paths, completing the case.

Case 2: *y* has no other children.

Let $x^* \in X$ be the parent vertex of y.

Case 2a: $\deg(x^*) \leq 2$ (this includes the case where x^* is the root and $\deg(x^*) = 1$). Delete x and y from G to get G'. For every $S \subseteq X - \{x\}$, we have Λ -def $(G, S) = \Lambda$ -def(G', S), since y cannot be in $\Lambda_G(S)$ and all other vertices of Y are still in G'. Therefore Λ -def $(G) \geq \Lambda$ -def(G').

By the inductive hypothesis, G' has a path X-cover by at most Λ -def(G') paths. By the case, x^* is a leaf of G' (or an isolated vertex), so the path that covers x^* must begin or end at x^* . Extend that path to go through y and x, and we get a path X-cover of Gby Λ -def $(G') \leq \Lambda$ -def(G) paths, completing the case.

Case 2b: $\deg(x^*) \ge 3$.

Let y_1, \ldots, y_k be all of the children of x^* (including y); by the case, $k \ge 2$. No vertices of Y are leaves, so each has a child. By our choice of x, those children are all as far from the root as possible, so they must all be leaves. If any of y_1, \ldots, y_k have multiple children, then we can proceed as in **Case 1**, so assume each y_i has a single child x_i . Delete x_k and y_k from G to get G'.

Let $S \subseteq X - \{x_k\}$ be the set such that Λ -def $(G') = \Lambda$ -def(G', S). We may assume that $x^* \notin S$ by one of the following modifications:

- If $x^* \in S$ and $x_1 \notin S$, replace S by $S' = S \cup \{x_1\} \{x^*\}$. Then |S'| = |S| and $|\Lambda_{G'}(S')| \leq |\Lambda_{G'}(S)|$: y_1 is in neither $\Lambda_{G'}(S)$ nor $\Lambda_{G'}(S')$, and no other vertices in Y have any neighbors in S' that they did not have in S. So Λ -def $(G', S') \geq \Lambda$ -def(G', S).
- If $x^* \in S$ and $x_1 \in S$, replace S by $S' := S \{x^*\}$. Then |S'| = |S| 1, but $|\Lambda_{G'}(S')| \leq |\Lambda_{G'}(S)| - 1$ as well, since $y_1 \in \Lambda_{G'}(S)$ but $y_1 \notin \Lambda_{G'}(S')$. (No other vertices in Y have any neighbors in S' that they did not have in S.) So Λ -def $(G', S') \geq \Lambda$ -def(G', S).

When $x^* \notin S$, we have Λ -def $(G, S \cup \{x_k\}) = \Lambda$ -def(G', S) + 1, because $|S \cup \{x_k\}| = |S| + 1$, while $\Lambda_G(S \cup \{x_k\}) = \Lambda_{G'}(S)$. Therefore Λ -def $(G) \ge \Lambda$ -def(G') + 1.

By the inductive hypothesis, G' has a path X-cover by at most Λ -def(G') paths. Add the path (x_k) to get a path X-cover of G by Λ -def $(G') + 1 \leq \Lambda$ -def(G) paths, completing the case and the proof.

3 3-regular graphs

It is a standard result (Corollary 3.1.13 in [18]) that every regular bipartite graph has a perfect matching. Removing a perfect matching from a d-regular bipartite graph leaves a (d-1)-regular bipartite graph, which also has a perfect matching. The union of the two matchings provides a cover of G by vertex-disjoint cycles, giving the following lemma (which is also well-known):

Lemma 8. If G is a regular bipartite graph, then G has a cycle cover.

The existence of this lemma is the primary reason that this proof is simpler than the proof of Theorem 6 in the next section. That proof begins with the same ideas, but must deal with vertices of X that are not part of the initially chosen collection of cycles.

Proof of Theorem 5. By Lemma 8, we can take a cycle cover \mathcal{C} of G.

Of all Λ -independent subsets of X containing at most one vertex of each cycle in C, let S be one with the maximum possible size. Since $|S| \leq \alpha_{\Lambda}(G)$, it suffices to construct a path cover of G with exactly |S| paths. We give an algorithm for this below.

Let H be the subgraph of G consisting of the |S| cycles in \mathcal{C} which contain a vertex of S; together, these cycles include all vertices of S. Over the course of the algorithm, we will modify H, ultimately extending it to a spanning subgraph of G, while maintaining two properties: (1) H has |S| components, and (2) each component of H is traceable.

In the first step of the algorithm, choose a cycle $C \in \mathcal{C}$ that is not yet contained in H, and $x(C) \in V(C) \cap X$. By the maximality of S, we have $\Lambda(S \cup \{x(C)\}) \neq \emptyset$, so there is some vertex $s \in S$ such that x(C) and s have a common neighbor y. Let C(s) be the cycle in \mathcal{C} containing s. The vertex y must lie on either C or C(s), since otherwise y would have four neighbors: x(C), s, and its two neighbors on the cycle in \mathcal{C} containing y. We extend H by adding cycle C to H, and either the edge x(C)y (if y lies on C(s)) or sy (if y lies on C). This ends one step of the algorithm.

This step maintains the property that H has |S| components, since cycle C has been joined to an existing component of H. Furthermore, after this first step, each component of H remains traceable.

Subsequent steps of the algorithm iterate this step, although we must choose C carefully to maintain the property that each component of H is traceable. To do this, we define the idea of a *dangerous* vertex. Consider an arbitrary $s \in S$; let C(s) be the cycle of C containing s, and let y_1, y_2 be the two neighbors of s along C(s). Initially, the component of H containing s is just C(s). There are three ways that C(s) can potentially be extended in H, namely via an edge from any of s, y_1 , or y_2 going to another cycle in C.

The component remains traceable if any one of these edges is used to extend it: in that case, we can extend that edge to a Hamiltonian path by going the long way around both cycles. The component also remains traceable if it is extended both using an edge from s and using an edge from y_1 . In that case, delete edge sy_1 , obtaining a long path containing s and y_1 joining two cycles; extend that path by going the long way around both of those cycles. The same is true if y_1 is replaced by y_2 .

However, we must ensure that the component of H containing s is never extended by using edges from both y_1 and y_2 . Suppose that a step of the algorithm extended the component of H containing s via an external edge to y_1 , and y_2 has a neighbor w in some $C \in \mathcal{C}$ not yet contained in H. In this situation, say w is *dangerous*, and we select C in the subsequent step of the algorithm, and choose $x(C) \neq w$. This ensures the component of H containing s cannot be extended using edges from both y_1 and y_2 , because one of those edges goes to C, and C will become part of H in the next step of the algorithm and hence will not be considered at later stages of the algorithm. As a result, no component of H is ever prevented from being traceable.

If there is no dangerous cycle at a particular step of the algorithm, then we may select C arbitrarily.

At the end of the algorithm, we have a spanning subgraph H with |S| traceable components. By taking a Hamiltonian path in each component, we obtain a path cover of G with |S| paths, completing the proof.

4 Graphs with maximum degree 3

Proof of Theorem 6. We will prove the theorem by describing an algorithm that constructs a path X-cover \mathcal{P} and a set $S \subseteq X$ with $|\mathcal{P}| = \Lambda$ -def(G, S).

To begin the algorithm, let C be a collection of vertex-disjoint cycles in G satisfying the following conditions:

- 1. The union of the cycles contains as many vertices of G as possible.
- 2. Subject to condition 1, there are as few cycles as possible.

As a consequence of condition 1, deleting the vertices in C from G leaves a forest, which we call F.

In the next phase of the algorithm, we process the cycles in \mathcal{C} , one at a time. This phase has two goals. First, for each $C \in \mathcal{C}$, we will choose a designated vertex $x(C) \in V(C) \cap X$. Intuitively, x(C) will be the only vertex of C which may become part of the high- Λ deficiency set S we construct. We define $y^+(C)$ and $y^-(C)$ to be the two neighbors of x(C) along C. Second, we will split \mathcal{C} into three sets: \mathcal{C}_{good} , \mathcal{C}_{bad} , and \mathcal{C}_{ugly} . Intuitively, if $C \in \mathcal{C}_{good}$, then x(C) is far from any problems; if $C \in \mathcal{C}_{bad}$, then x(C) is too close to the forest F; finally, if $C \in \mathcal{C}_{ugly}$, then x(C) is too close to x(D) for some $D \in \mathcal{C}_{good} \cup \mathcal{C}_{bad}$.

As in the proof of Theorem 5, processing a cycle in C will sometimes yield a *dangerous* vertex w on an unprocessed cycle C which must never become x(C). In this case, we select C to be processed next, and we choose an arbitrary vertex in $V(C) \cap X$ other than

w to be x(C); after this is done, w is no longer dangerous. If there is no dangerous vertex (such as, for example, at the start of this phase), then we choose an unprocessed C and a vertex $x(C) \in V(C) \cap X$ arbitrarily.

To decide what to do with a cycle C as we process it, we consider the following cases, *in order*, choosing the first that applies:

Case 1: x(C) has a common neighbor with x(D) for some $D \in \mathcal{C}_{good} \cup \mathcal{C}_{bad}$, and that common neighbor lies on either C or D. In other words, at least one edge

$$e(C) \in \{x(C)y^+(D), x(C)y^-(D), y^+(C)x(D), y^-(C)x(D)\}$$

must exist in G. (If multiple choices of D or of e(C) are possible, then fix one of them.) In this case, we place C in C_{ugly} ; we say that C attaches to D at u, where u is the endpoint of e(C) in D. We save the edge e(C) for reference; later, we will use it to extend a path covering D to also cover C.

Additionally, if $e(C) = x(C)y^{\pm}(D)$ and the vertex $y^{\mp}(D)$ (that is, whichever of $y^{+}(D), y^{-}(D)$ is not an endpoint of e(C)) is adjacent to a vertex w on an unprocessed cycle, we mark w as dangerous, and choose the next cycle to be processed accordingly (that is, process w's cycle C' next, and choose $x(C') \neq w$).

Case 2: At least one of x(C), $y^+(C)$, or $y^-(C)$ has a neighbor in F. In this case, we place C in $\mathcal{C}_{\mathsf{bad}}$. Additionally, if $y^{\pm}(C)$ has a neighbor in F and $y^{\mp}(C)$ has a neighbor w on an unprocessed cycle, we mark w as dangerous, and choose the next cycle to be processed accordingly.

Case 3: Neither case 1 nor case 2 occurs. In this case, we simply place C in \mathcal{C}_{good} .

This concludes the second phase (or the *processing phase*) of the algorithm.

In the third phase of the algorithm, we create an auxiliary graph F^* (which is not precisely a subgraph of G) containing F and some extra vertices representing the elements of $\mathcal{C}_{\mathsf{bad}}$. For each $C \in \mathcal{C}_{\mathsf{bad}}$:

- We add x(C) to F^* , together with the edge to its neighbor in F, if there is one.
- We add an artificial vertex $y^*(C)$ to F^* that is adjacent to x(C) and to the neighbors of both $y^+(C)$ and $y^-(C)$ in F, if these exist.

Before we continue, we must show that F^* is a forest. Suppose for the sake of contradiction that F^* contains a cycle. Since F is acyclic, this cycle must contain either x(C) or $y^*(C)$ for at least one $C \in \mathcal{C}_{bad}$.

First, consider the case that the cycle only includes the vertex $y^*(C)$ for a single $C \in \mathcal{C}_{\mathsf{bad}}$. This means that there is a path P from $y^+(C)$ to $y^-(C)$ of length at least 3, whose internal vertices are in F. Now we can modify C, replacing x(C) and the edges $y^+(C)x(C), x(C)y^-(C)$ by P. The resulting cycle contains more vertices that C, violating condition 1 in the definition of \mathcal{C} .

Similarly, if the cycle in F^* includes only the vertices x(C) and $y^*(C)$ for a single $C \in \mathcal{C}_{\mathsf{bad}}$, we can expand C to include some vertices in F. This also violates condition 1 in the definition of \mathcal{C} .

Finally, consider the case that the cycle in F^* includes vertices x(C) and/or $y^*(C)$ for multiple $C \in \mathcal{C}_{bad}$. In this case, we can extend it to a cycle in G: every time the cycle in F^* visits $y^*(C)$, we can replace that visit by a path that enters C via $y^{\pm}(C)$, goes around C, and leaves via either $y^{\mp}(C)$ or x(C). This cycle in G contains at least as many vertices as the cycles from \mathcal{C} it uses: it misses at most the vertex x(C) from each of them, but includes a vertex in F between any two of the cycles in \mathcal{C} . Therefore, we can replace multiple cycles in \mathcal{C} by a single cycle through at least as many vertices, violating condition 2 in the definition of \mathcal{C} .

In all cases, we arrive at a contradiction, so we can conclude that F^* is a forest. We give it the structure of an (X', Y')-bigraph by defining:

$$X' = (X \cap V(F)) \cup \{x(C) : C \in \mathcal{C}_{\mathsf{bad}}\},\$$

$$Y' = (Y \cap V(F)) \cup \{y^*(C) : C \in \mathcal{C}_{\mathsf{bad}}\}.$$

By Proposition 4, we can find a path X'-cover \mathcal{P}' of F^* and a set $S' \subseteq X'$ such that Λ -def $(F^*, S') \ge |\mathcal{P}'|$. We may assume that both endpoints of every path in \mathcal{P}' are in X', not Y'.

Let S'' be obtained from S' by deleting each vertex of the form x(C) for which $y^*(C) \in \Lambda_{F^*}(S')$. Because $X' \subseteq X$, we have $S'' \subseteq S' \subseteq X$ as well.

Claim 9. $\Lambda_G(S'') \subseteq V(F)$.

Proof. Suppose for the sake of contradiction that $y \in \Lambda_G(S'')$ but $y \notin V(F)$. Then y lies on some cycle $C \in \mathcal{C}$, and therefore has at most one neighbor not on C; in order to have $y \in \Lambda_G(S'')$, we must have $y \in \{y^+(C), y^-(C)\}$, with $x(C) \in S''$. Additionally, y must have a second neighbor $x' \in S''$ outside C. To have $x' \in S''$, either x' = x(D) for some other cycle $D \in \mathcal{C}$, or $x' \in V(F)$.

If x' = x(D) for some other cycle $D \in C$, then since $x(C), x(D) \in S'' \subseteq X'$, we must have $C, D \in \mathcal{C}_{\mathsf{bad}}$. But then, whichever of C and D is considered second, it would have fallen under Case 1 when processed due to the common neighbor y, and then it would be placed in $\mathcal{C}_{\mathsf{ugly}}$ instead. Therefore this cannot occur. If $x' \in V(F)$, then x(C) and x'have a common neighbor in F^* as well: vertex $y^*(C)$. But then, by the definition of S'', $x(C) \notin S''$, so this also cannot occur.

Claim 9 implies that $\Lambda_G(S'') \subseteq \Lambda_{F^*}(S')$: if $y \in V(F)$, then the neighbors of y are the same in G and F^* , so the vertices of S'' adjacent to y in G are vertices of S' adjacent to y in F^* . Moreover, there are at least |S'| - |S''| vertices in $\Lambda_{F^*}(S')$ that are not in $\Lambda_G(S'')$: for each vertex x(C) in S' but not in $S'', y^*(C) \in \Lambda_{F^*}(S')$, but $y^*(C) \notin V(G)$, so $y^*(C) \notin \Lambda_G(S'')$.

It follows that $|\Lambda_G(S'')| \leq |\Lambda_{F^*}(S')| - |S'| + |S''|$, and therefore Λ -def $(G, S'') \geq |\mathcal{P}'|$.

We are now ready to construct the path X-cover \mathcal{P} in G and a set $S \subseteq X$ with $|\mathcal{P}| \leq \Lambda$ -def(G, S). Let $S = S'' \cup \{x(C) : C \in \mathcal{C}_{good}\}$. The vertices $\{x(C) : C \in \mathcal{C}_{good}\}$ have no common neighbors with each other or with any vertex in S''. This is ensured by Case 1 and Case 2 of the processing phase, where any cycle C for which x(C) did have

THE ELECTRONIC JOURNAL OF COMBINATORICS 31(3) (2024), #P3.1

such a common neighbor would be placed in \mathcal{C}_{bad} or \mathcal{C}_{ugly} instead. Therefore Λ -def $(G, S) = \Lambda$ -def $(G, S'') + |\mathcal{C}_{good}|$. We transform \mathcal{P}' into \mathcal{P} with $|\mathcal{C}_{good}|$ additional paths.

For each $C \in \mathcal{C}_{good}$, we add a single path to \mathcal{P} that covers the vertices of X on both C and any cycles in \mathcal{C}_{ugly} that attach to C. There are three possibilities:

- If there are no cycles in \mathcal{C}_{ugly} that attach to C, we take a path that goes around C.
- If there is one cycle $C_1 \in C_{ugly}$ that attaches to C, we obtain a path that covers both C and C_1 by taking both cycles, adding edge $e(C_1)$, and deleting an edge incident on $e(C_1)$ from both cycles.
- If there are two cycles C₁, C₂ ∈ C_{ugly} that attach to C, then e(C₁) and e(C₂) cannot both have endpoints in {y⁺(C), y⁻(C)}, because once one such cycle is added to C_{ugly} in Case 1, we mark the appropriate vertex as dangerous to prevent a second cycle of this type from appearing. (This is also why three cycles in C_{ugly} cannot attach to C.) Without loss of generality, C₁ and C₂ attach to C at x(C) and y⁺(C). We obtain a path that covers C, C₁, and C₂ by taking all three cycles, adding edges e(C₁) and e(C₂), deleting edge x(C)y⁺(C) from C, and deleting an edge from each of C₁ and C₂ incident on e(C₁) and e(C₂), respectively.

Finally, we must transform the paths in \mathcal{P}' that use the vertices x(C) or $y^*(C)$ for some $C \in \mathcal{C}_{\mathsf{bad}}$ into paths in G that cover the cycles in $\mathcal{C}_{\mathsf{bad}}$, as well as any cycles in $\mathcal{C}_{\mathsf{ugly}}$ that attach to them.

We do this by considering the cycles in $\mathcal{C}_{\mathsf{bad}}$ one at a time. For each $C \in \mathcal{C}_{\mathsf{bad}}$, there are either one or two paths in \mathcal{P}' that need to be transformed: the path $P_x \in \mathcal{P}'$, and possibly a path $P_y \in \mathcal{P}'$ that passes through $y^*(C)$, if there is such a path and it is distinct from P_x .

To transform P_x and P_y , we replace x(C) in P_x and $y^*(C)$ in P_y by segments in G that cover C and any cycles in C_{ugly} that attach to C. There are multiple possibilities for how the transformation is done, depending on how x(C) and $y^*(C)$ appear on paths in \mathcal{P}' :

• P_x contains x(C) but not $y^*(C)$, and no path in \mathcal{P}' uses the vertex $y^*(C)$.

In this case, x(C) must be an endpoint of P_x , since x(C) has only one neighbor in F^* other than $y^*(C)$. We extend P_x to go around the cycle C, ending at either $y^+(C)$ or $y^-(C)$.

If there is a cycle $C_1 \in \mathcal{C}_{ugly}$ that attaches to C at $y^{\pm}(C)$, we choose P_x to go around C so that it ends at the endpoint of $e(C_1)$; then, extend P_x to use $e(C_1)$ and go around C_1 . As before, there cannot be two cycles $C_1, C_2 \in \mathcal{C}_{ugly}$ that attach to C at $y^+(C)$ and $y^-(C)$, since we mark the appropriate vertex as dangerous to prevent this.

• P_x contains both x(C) and $y^*(C)$.

In this case, if we orient P_x to visit x(C) before $y^*(C)$, then it must use edge $x(C)y^*(C)$ to arrive at $y^*(C)$: there is no other $x(C) - y^*(C)$ path, since F^* is a forest.

The edge P_x uses to leave $y^*(C)$ comes from an edge in G from $y^{\pm}(C)$ to F; we modify P_x to go around the cycle C from x(C) to $y^{\pm}(C)$. If there is a cycle $C_1 \in \mathcal{C}_{ugly}$ that attaches to C at x(C), then x(C) cannot have a neighbor in F, so it is an endpoint of P_x . We extend x(C) in the other direction, prepending a path that goes around C_1 and takes edge $e(C_1)$ to x(C). There cannot be a cycle $C_2 \in \mathcal{C}_{ugly}$ that attaches to C at $y^{\mp}(C)$, again because we mark a vertex as dangerous to prevent it.

• P_x contains x(C), and a different path P_y contains $y^*(C)$.

In this case, x(C) must be an endpoint of P_x . We leave P_x unchanged, unless there is a cycle $C_1 \in C_{ugly}$ that attaches to C at x(C). In this case, x(C) has no neighbors in F, so P_x must be a path of length 0. We replace P_x by a path that covers C_1 and ends with $e(C_1)$, also covering x(C).

Meanwhile, P_y must enter and leave $y^*(C)$ by edges other than $x(C)y^*(C)$; in G, these correspond to two edges between $\{y^+(C), y^-(C)\}$ and F. We modify P_y , replacing $y^*(C)$ by the $y^+(C), y^-(C)$ -path that goes around C, covering all its vertices except the previously covered x(C).

Once a path $P \in \mathcal{P}'$ has been transformed once for every cycle $C \in \mathcal{C}_{bad}$ such that P contains x(C) or $y^*(C)$, it is a path in G (since it no longer contains any vertex $y^*(C)$ for $C \in \mathcal{C}_{bad}$) and we add it to \mathcal{P} .

After all transformations of \mathcal{P}' are complete, the collection \mathcal{P} is a path X-cover. The $|\mathcal{C}_{good}|$ paths first added to \mathcal{P} ensure that \mathcal{P} covers all vertices of X on \mathcal{C}_{good} . The paths transformed from paths in \mathcal{P}' ensure that \mathcal{P} covers all vertices of X on \mathcal{C}_{bad} , as well as all vertices of X'. Finally, for each $C \in \mathcal{C}_{good} \cup \mathcal{C}_{bad}$, the path in \mathcal{P} that covers C also covers each cycle in \mathcal{C}_{ugly} that attaches to C. This completes the proof, since $|\mathcal{P}| = |\mathcal{P}'| + |\mathcal{C}_{good}| \leq \Lambda\text{-def}(G, S') + |\mathcal{C}_{good}| = \Lambda\text{-def}(G, S).$

5 Graphs with high girth

We will use the Lovász Local Lemma [5] to prove Theorem 7, in the form stated below.

Lemma 10 (The Local Lemma; Lemma 5.1.1 in [2]). Let A_1, A_2, \ldots, A_N be events in an arbitrary probability space. A directed graph D = (V, E) on the set of vertices V = $\{1, 2, \ldots, N\}$ is called a dependency digraph for the events A_1, \ldots, A_N if for each i, $1 \leq i \leq N$, the event A_i is mutually independent of all the events $\{A_j : (i, j) \notin E\}$. Suppose that D = (V, E) is a dependency digraph for the above events and suppose there are real numbers x_1, \ldots, x_N such that $0 \leq x_i < 1$ and $\Pr[A_i] \leq x_i \prod_{(i,j) \in E} (1 - x_j)$ for all $1 \leq i \leq N$. Then

$$\Pr\left[\bigwedge_{i=1}^{N} \overline{A_i}\right] \geqslant \prod_{i=1}^{N} (1-x_i).$$

In particular, with positive probability, no event A_i holds.

A symmetric version of Lemma 10 is often used, where $\Pr[A_i] = p$ for all i; by setting $x_i = e \cdot \Pr[A_i]$ for all i, and using the inequality $(1 - \frac{1}{x+1})^x \ge \frac{1}{e}$, valid for all $x \ge 0$, the hypotheses of the lemma are satisfied. In our case, the probabilities of our events A_i will vary, but we will pursue mostly the same strategy. We will still set $x_i = e \cdot \Pr[A_i]$ for all i; because event A_i will depend on two types of other events, we will use the inequality $(1 - \frac{1}{2x+1})^x \ge e^{-1/2}$ (also valid for all $x \ge 0$) on two parts of the product, instead.

Proof of Theorem 7. Let G be an (X, Y)-bigraph with maximum degree at most d and girth $g \ge 4ed^2 + 1$. Additionally, let \mathcal{C} be a collection of pairwise vertex-disjoint cycles in G that cover X. This collection exists either by assumption or (if G is taken to be d-regular) by Lemma 8. The girth condition on G guarantees that the cycles in \mathcal{C} are relatively long, so there cannot be too many of them; in fact, we will show that there cannot be more than $\alpha_{\Lambda}(G)$ of them.

If $v \in X$, we will write C(v) for the cycle in C containing v. Furthermore, if $v_1, v_2 \in X$, we will write $v_1 \sim v_2$ to mean that $C(v_1) \neq C(v_2)$ and v_1 and v_2 have a common neighbor in Y.

We choose a set S randomly, by selecting one vertex of X uniformly at random from each cycle in \mathcal{C} . Let $\{u_1, v_1\}, \ldots, \{u_N, v_N\}$ be an enumeration of the (unordered) pairs of vertices in X such that $u_i \sim v_i$. For each $i, 1 \leq i \leq N$, we let A_i be the event that $u_i \in S$ and $v_i \in S$. As a result, the conjunction $\overline{A_1} \wedge \cdots \wedge \overline{A_n}$ is exactly the claim that S is Λ -independent. If we can satisfy the hypotheses of Lemma 10 for the events A_1, \ldots, A_N , then this conjunction occurs with positive probability, and therefore $|\mathcal{C}| = |S| \leq \alpha_{\Lambda}(G)$, proving the theorem.

We define the dependency digraph D to include an edge (i, j) whenever the four vertices u_i, v_i, u_j , and v_j do not lie on four distinct cycles. If $C(u_i)$ has length $2\ell_1$ and $C(v_i)$ has length $2\ell_2$, we define $x_i = \frac{e}{\ell_1 \ell_2}$; for reference, $\Pr[A_i] = \frac{1}{\ell_1 \ell_2}$.

We will show that the hypotheses of Lemma 10 hold with the choices made above. To do this, we must put a lower bound on

$$x_i \prod_{(i,j)\in E(D)} (1-x_j)$$

for an arbitrary event A_i .

This product consists of two types of events A_j . The first type consists of those A_j for which either u_j or v_j lies on $C(u_i)$ (including $u_j = u_i$ or $v_j = u_i$). Since $C(x_i)$ contains ℓ_1 vertices of X, there are at most $\ell_1 d^2$ events A_j of this type. For each of them, one of $C(u_j), C(v_j)$ is the same as $C(u_i)$ and has length $2\ell_1$, and the other has length $2\ell' \ge g$. Therefore $1 - x_j = 1 - \frac{e}{\ell_1 \ell'} \ge 1 - \frac{2e}{\ell_1 g}$, for an overall product of at least $(1 - \frac{2e}{\ell_1 g})^{\ell_1 d^2}$.

The second type of events A_j such that $(i, j) \in E(D)$ consists of those A_j for which either u_j or v_j lies on $C(v_i)$. By a similar argument, there are at most $\ell_2 d^2$ events A_j of this type, and for each of them, $1 - x_j \ge 1 - \frac{2e}{\ell_2 g}$, so

$$x_i \prod_{(i,j)\in E(D)} (1-x_j) \ge \frac{e}{\ell_1 \ell_2} \left(1 - \frac{2e}{\ell_1 g} \right)^{\ell_1 d^2} \left(1 - \frac{2e}{\ell_2 g} \right)^{\ell_2 d^2}$$

THE ELECTRONIC JOURNAL OF COMBINATORICS 31(3) (2024), #P3.1

Recall that $g \ge 4ed^2 + 1$; a lower bound on $\frac{\ell_1 g}{2e} \ge 2\ell_1 d^2 + \frac{\ell_1}{2e}$ is $2\ell_1 d^2 + 1$. Applying $(1 - \frac{1}{2x+1})^x \ge e^{-1/2}$, we get

$$\left(1 - \frac{2e}{\ell_1 g}\right)^{\ell_1 d^2} \ge \left(1 - \frac{1}{2\ell_1 d^2 + 1}\right)^{\ell_1 d^2} \ge e^{-1/2}$$

and similarly $\left(1 - \frac{2e}{\ell_2 g}\right)^{\ell_2 d^2} \ge e^{-1/2}$. Therefore

$$\frac{e}{\ell_1 \ell_2} \left(1 - \frac{2e}{\ell_1 g} \right)^{\ell_1 d^2} \left(1 - \frac{2e}{\ell_2 g} \right)^{\ell_2 d^2} \geqslant \frac{e}{\ell_1 \ell_2} \cdot e^{-1/2} \cdot e^{-1/2} = \frac{1}{\ell_1 \ell_2} = \Pr[A_i]$$

and the conditions of Lemma 10 are satisfied.

We conclude that with positive probability, S is Λ -independent, and therefore $|\mathcal{C}| = |\mathcal{S}| \leq \alpha_{\Lambda}(G)$. The theorem follows, since we can find a path X-cover of G of size $|\mathcal{C}|$ by removing a vertex of Y from each cycle in \mathcal{C} .

Acknowledgements

We thank both referees for their helpful comments.

References

- T. Akiyama, T. Nishizeki, and N. Saito. NP-completeness of the Hamiltonian cycle problem for bipartite graphs. J. Inform. Process., 3(2):73-76, 1980.
- [2] N. Alon and J. Spencer. The Probabilistic Method. John Wiley & Sons, Hoboken, NJ, 2008.
- [3] J. Barát, A. Grzesik, A. Jung, Z. Lóránt Nagy, and D. Pálvölgyi. The double Hall property and cycle covers in bipartite graphs. *Discrete Math.*, 347(9):114079, 2024. doi:10.1016/j.disc.2024.114079.
- [4] C. Berge. *Graphs and hypergraphs*. North-Holland Publishing Company, Amsterdam-London, 1973.
- [5] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. *Infinite and finite sets*, 10(2):609–627, 1975.
- [6] T. Gallai and A. Norton Milgram. Verallgemeinerung eines graphentheoretischen Satzes von Rédei. Acta Sc. Math, 21:181–186, 1960.
- [7] U. Feige and E. Fuchs. On the path partition number of 6-regular graphs. J. Graph Theory, 101(3):345-378, 2022. doi:10.1002/jgt.22830.
- [8] R.J. Gould. Advances on the Hamiltonian problem—a survey. Graphs Combin., 19(1):7–52, 2003. doi:10.1007/s00373-002-0492-x.
- [9] J. Han. On vertex-disjoint paths in regular graphs. *Electron. J. Combin.*, 25(2):#P1.12, 2018. doi:10.37236/7109.

- [10] A. Kostochka, M. Lavrov, R. Luo, and D. Zirlin. Conditions for a bigraph to be super-cyclic. *Electron. J. Combin.*, 28(1):P1.2, 2021. doi:10.37236/9683.
- [11] M. Kouider and M. Zamime. On the path partition of graphs. arXiv:2212.12793, 2022.
- [12] L. Lu and Z. Wang. On Hamiltonian Berge cycles in [3]-uniform hypergraphs. Discrete Math., 344(8):112462, 2021. doi:10.1016/j.disc.2021.112462.
- [13] C. Magnant and D. M. Martin. A note on the path cover number of regular graphs. Australas. J. Combin., 43:211-217, 2009.
- [14] C. Magnant, H. Wang and S. Yuan. Path partitions of almost regular graphs. Australas. J. Combin., 64:334–340, 2016.
- [15] H. Müller. Oriented hypergraphs, stability numbers and chromatic numbers. Discrete Math., 34(3):319–320, 1981. doi:10.1016/0012-365X(81)90011-X.
- [16] N. Salia. Extremal problems for paths and cycles. Doctoral dissertation, Central European University.
- [17] J. Singer. A theorem in finite projective geometry and some applications to number theory. Trans. Amer. Math. Soc., 43(3):377–385, 1938. doi:10.2307/1990067.
- [18] D.B. West. Introduction to Graph Theory. Prentice Hall, Englewood Cliffs, NJ, 1996.