Linear time algorithm for computing a small biclique in graphs without long induced paths

Aistis Atminas[1], Vadim V. Lozin[2], and Igor Razgon[3]

[1] DIMAP and Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK. A.Atminas@warwick.ac.uk
[2] DIMAP and Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK. V.Lozin@warwick.ac.uk
[3] Department of Computer Science, University of Leicester, University Road, Leicester, LE1 7RH, UK. ir45@mcs.le.ac.uk

Abstract. The biclique problem asks, given a graph G and a parameter k, whether G has a complete bipartite subgraph of k vertices in each part (a biclique of order k). Fixed-parameter tractability of this problem is a longstanding open question in parameterized complexity that received a lot of attention from the community. In this paper we consider a restricted version of this problem by introducing an additional parameter s and assuming that G does not have induced (i.e. chordless) paths of length s. We prove that under this parameterization the problem becomes fixed-parameter linear. The main tool in our proof is a Ramsey-type theorem stating that a graph with a long (not necessarily induced) path contains either a long *induced* path or a large biclique.

1 Introduction

Overview of our results. Let us call a complete bipartite graph $H = (A, B, E)$ with $|A| = |B| = k$ a *biclique* of order k. Given a graph G and parameter k, the Biclique problem asks if G has a biclique of order k as a subgraph. Fixed-parameter tractability of this problem is a longstanding open question that received significant attention from the parameterized complexity community and is believed to be W[1]-hard (see the abstract of [6]).

In this paper we consider a restricted version of this problem by introducing an additional parameter s and assuming G to be P_s-free, i.e. without induced paths of length s. We show that under this additional parameterization the biclique problem becomes fixed-parameter linear. Let us remark that the parameterization by s alone is not enough for efficient computing of a largest biclique (Proposition 1). Indeed, the construction used by Johnson [16] to establish the NP-hardness of the Biclique problem in fact reduces the Clique problem to an instance of the Biclique problem on P_8-free graphs, that is the Biclique problem is NP-hard on P_s-free graphs for $s \geq 8$. In this sense, the use of s as an additional parameter is meaningful.

The key ingredient in our solution is a combinatorial statement (Theorem 1) claiming the existence of a number $A(k, s)$ such that every P_s-free graph with a path of length at least $A(k, s)$ has a biclique of order k. This result belongs to a large body of 'Ramsey-type' theorems showing that if the given graph is 'large' in a certain sense, then it contains a large subgraph (either induced or not) that belongs to one of the specified families. In our case, the largeness condition is 'long path' and the families are bicliques and induced (i.e. chordless) paths. The proof of this result requires a number of intermediate stages which are proven by a non-trivial use of classical Ramsey's theorem. In particular, we use a 'non-binary' form of Ramsey's theorem with hyperedges of size 3, although we apply it to simple graphs only.

Related work. The Biclique problem appears under the name of 'Balanced Complete Bipartite Subgraph' as problem [GT24] in the famous book of Garey and Johnson, an NP-hardness proof has been further provided by Johnson in [16]. An application of the problem to the VLSI design is described in detail in [1]. The problem has been considered in the context of approximation [12] and exact exponential time algorithms [3]. Polynomial time algorithms for a number of restricted classes of the Biclique problem have been proposed in [1]. To the best of our knowledge, the question regarding fixed-parameter tractability of Biclique was first asked in [9]. The question has been restated as an open problem in a number of subsequent publications, see e.g. [6], where the complexity of a number of parameterized problems is characterised as Biclique-hard. The *induced* Biclique problem is known to be W[1]-hard [7, 13].

Graphs without long induced paths, i.e. P_s-free graphs for a constant s, have been extensively studied in the literature (see e.g. [2, 11, 22]). For small values of s, the structure of P_s-free graphs is simple. For instance, P_3-free graphs are precisely the graphs every connected component of which is a clique. P_4-free graphs also enjoy many nice properties. In particular, the clique-width of P_4-free graphs is bounded by a constant and hence many algorithmic problems that are generally NP-hard admit polynomial time solutions when restricted to P_4-free graphs.

In the class of P_s-free graphs with $s \geq 5$, the situation changes drastically and the computational complexity changes from polynomial-time solvability to NP-hardness for many important algorithmic graph problems. For instance, VERTEX COLORING [18] and MINIMUM DOMINATING SET [17] are NP-hard for P_5-free graphs, and VERTEX 4-COLOURABILITY is NP-hard for P_8-free graphs [5]. For many other problems, the complexity status on graphs without long induced paths is unknown. For instance, the complexity status is unknown for MAXIMUM INDEPENDENT SET in P_s-free graphs with $s \geq 5$ and for VERTEX 3-COLOURABILITY in P_s-free graphs with $s \geq 7$ (for some partial results related to these problems we refer the reader to [15, 20, 21, 24, 25]).

Structure of the paper. Section 2 presents the algorithm for computing a biclique, Section 3 proves the main combinatorial result, Section 4 discusses directions of further research. All graphs in this paper are undirected, without loops and multiple edges.

2 Computing a small biclique in a graph without long induced paths

A biclique of order k is a bipartite graph (A, B, E) with $|A| = |B| = k$ and every $u \in A$ being adjacent to every $v \in B$. A notorious problem in Parameterized Complexity asks, given a parameter k, if the given graph has a biclique of order k. The fixed-parameter tractability of this problem is wide open despite efforts of many researchers. In this paper we consider the following restricted version of this problem (the abbreviation NLIP in the name of this problem stands for 'No Long Induced Paths').

NLIP-BICLIQUE
Input: A graph G
Parameters: k, s
Assumption: G is P_s-free
Output: A biclique of G of size at least k or 'NO' if there is no such biclique.

The following proposition, essentially proven in [16] shows that the choice of parameters is meaningful in the sense that s alone is not enough to compute a maximum biclique efficiently.

Proposition 1. *Computing maximum biclique in a P_s-free graph is NP-hard for $s \geq 8$*

Proof. This is implicitly proven in [16] because an instance of the Clique problem is reduced to an instance of the Biclique problem on a P_8-free graph. Indeed, by construction, given a graph G, a bipartite graph H is constructed in which the first part A corresponds to the edges of G and the second part B corresponds to a superset of its vertices and each vertex of A is adjacent to all vertices of B but those corresponding to the endpoints of the respective edge. It is not hard to see that H is P_8-free. Indeed, let P be a path of length 8. It has one terminal vertex u in A, and 4 vertices $v_1, \dots, v_4$ of B included in it. Three vertices out of $v_1, \dots, v_4$ are non-adjacent to u in P but only two of them may be the endpoints of the respective edge. It follows that u is necessarily adjacent in H to the remaining one, thus producing a chord in P. □

In this paper we prove that the NLIP-BICLIQUE problem is FPT. The central statement towards establishing this is the following.

Theorem 1. *For any natural numbers s and k there is a natural number $P(s,k)$ such that any graph with a path of length $P(s,k)$ has either an induced path of length s or a biclique of size k.*

We prove Theorem 1 in Section 3. Now we use this theorem to establish a corollary that the same long induced path/large biclique statement follows from a large treewidth as well.

Corollary 1. *For any natural numbers s and k there is a natural number $T(s,k)$ such that any graph of treewidth at least $T(s,k)$ either has an induced path of length s or a biclique of order k.*

Proof. It is well known (see Theorem 9 of [14]) that for each natural r there is $Y(r)$ such that if the treewidth of the given graph is at least $Y(r)$, the graph has a path of size at least r. Take $T(s,k) = Y(P(s,k))$ and apply Theorem 1. □

Theorem 2. *For fixed parameters s and k, the* NLIP-BICLIQUE *problem can be solved in a linear time.*

Proof. Let G be the input graph with n vertices. Using the linear time algorithm of Bodlaender [4], test the existence of a path of length $P(s,k)$ and find it, in case it exists.

Assume that such a path P has been found. In this case, the subgraph of G induced by the vertices of P has a biclique of size k as follows from Theorem 1. Since the size of this subgraph depends only on the parameters, the way this biclique is computed does not affect the desired runtime so, we can use the brute force.

If G does not have a path of length $P(s,q)$ then according to Corollary 1, the treewidth is at most $T(s,q)$, therefore, the biclique problem can be solved by standard techniques for graphs of bounded treewidth, say Courcelle's theorem [8]. □

3 Proof of Theorem 1

In order to prove Theorem 1, we modify it in the following way.

Theorem 3. *For every t, q, and s, there is a number $z = Z(s,t,q)$ such that every graph with a path of length at least z contains either K_t or $K_{q,q}$ or P_s as an induced subgraph.*

It is not hard to show that Theorem 1 and Theorem 3 are equivalent. Indeed, assume that Theorem 3 holds. Set $P(s,q) = Z(s,2q,q)$. It follows from Theorem 3 that a P_s-free graph with $P(s,q)$ vertices will have either a clique of size $2q$ or an induced biclique of order q. Clearly, in both cases the graph has a biclique of order q. Conversely, assume that Theorem 1 holds. Then we can just set $Z(s,t,q) = P(s, R(2,2,\max(t,q)))$, where R is the Ramsey number defined below. Thus, the equivalence between Theorem 1 and Theorem 3 has been established.

The proof of Theorem 3 consists of four stages outlined in the following four sections. On the first stage we define a class of graphs called *connecting structures*. We essentially prove that Theorem 1 holds for connecting structures, that is a sufficiently large connecting structure has either a large induced path or a large biclique. On the second stage we consider a class of graphs having a large *grid structure without an independent transversal* and we show that this is a sufficient condition for having a large biclique. On the third stage we define a class of graphs called a *bouquet* and we prove that a sufficiently large (P_s, K_t)-free graph necessarily has a large bouquet. On the final stage, we get the things together. We assume that our graph is (P_s, K_t)-free, using the third stage this immediately leads us to the conclusion that a large bouquet exists. We then show that, appropriately contracting vertices of this bouquet, we can get a large connecting structure as a subgraph. On the resulting connecting structure we consider the possibilities of a long induced path and a large biclique. Based on the assumption that the original graph is P_s-free, in both cases we infer the existence of a large grid structure without independent transversal, which in turn implies the existence of a large biclique.

Before we start the proof itself, we introduce the main tool we use in the proof, namely the fundamental result known as Ramsey's theorem, and provide a few its corollaries.

Theorem 4. *For any k, r and m, there is a number $R = R(k, r, m)$ such that in every coloring of k-subsets of an R-set with r colors there is a monochromatic m-set, i.e. a set of m elements all of whose k-subsets have the same color.*

For $k = 1$, this theorem is known as the Pigeonhole Principle. For $k = r = 2$, the number $R(2, 2, m)$ is frequently referred to as the (symmetric) Ramsey number, i.e. the minimum number such that every graph with at least $R(2, 2, m)$ vertices has either a clique of size m or an independent set of size m. In case of connected graphs, the Ramsey number admits the following generalization (see e.g. Proposition 9.4.1 in [10]).

Lemma 1. *For any t, q and s, there is a number $\ell(t, q, s)$ such that every connected graph with at least $\ell(t, q, s)$ vertices contains either K_t or $K_{1,q}$ or P_s as an induced subgraph.*

The Ramsey number also has a bipartite analog, which can be easily derived with the Pigeonhole Principle and which states that for any q, there is a number $BR(q)$ such that every bipartite graph $G = (V_1, V_2, E)$ with $|V_1| \geq BR(q)$ and $|V_2| \geq BR(q)$ has either a biclique $K_{q,q}$ or its bipartite complement. With a simple induction, this statement can be extended to multipartite graphs as follows.

Lemma 2. *For any k and q, there is a number $MR(k, q)$ such that in every k-partite graph $G = (V_1, V_2, \ldots, V_k, E)$ with $|V_i| \geq MR(k, q)$ $(i = 1, \ldots, k)$ there is a collection of subsets $U_i \subseteq V_i$ of size $|U_i| = q$ $(i = 1, \ldots, k)$ such that every pair of subsets induces either a biclique $K_{q,q}$ or its bipartite complement.*

3.1 Connecting structures

Definition 1. *A bipartite graph $G = (A, B, E)$ is called a* connecting structure *w.r.t. A if there is an injective function f from the set $\{\{u, v\} | u, v \in A\}$ of all the unordered pairs of A to B such that $f(\{u, v\})$ is adjacent to both u and v.*

Put it differently, we call G a connecting structure w.r.t. A if for each pair $\{u, v\} \subseteq A$ we can find a vertex of B adjacent to both u and v so that different pairs are associated with different vertices.

We refer to $|A|$ as the order of the connecting structure G.

Lemma 3. *For every natural numbers s and q, there is a number $L(s, q)$ such that every connecting structure of order at least $L(s, q)$ contains either a biclique of order q or an induced path of size s.*

Proof. Let $M := \max(\lceil s/2 \rceil + 1, 2q)$ and $L(s, q) := R(3, 3, M)$ (R is the Ramsey number). Consider a connecting structure $G = (A, B, E)$ w.r.t. $A = \{a_1, a_2, \ldots, a_l\}$ where $l := L(s, q)$. We color each triple a_i, a_j, a_k $(i < j < k)$ in one of the three colors (breaking any ties between colors 1 and 2 arbitrarily):

- color 1 if a_i is adjacent to $f(\{a_j, a_k\})$,
- color 2 if a_k is adjacent to $f(\{a_i, a_j\})$,
- color 3 if neither a_i is adjacent to $f(\{a_j, a_k\})$ nor a_k is adjacent to $f(\{a_i, a_j\})$.

Then A has a subset $A' = \{a_{i_1}, \ldots, a_{i_M}\}$ of $M \geq 2q$ vertices all of whose triples have the same color. Assume that this color is 1. Then every vertex of $A_1 = \{a_{i_1}, \ldots, a_{i_q}\}$ is adjacent to every vertex of $B_1 = \{f(\{u, v\}) | u, v \in \{a_{i_q}, \ldots, a_{i_{2q}}\}\}$. The adjacency of $u \in A_1$ to $w \in B_1$ follows either from the condition of color 1 or, in case $u = a_{i_q}$ and $w = f(\{a_{i_q}, v\})$ for $v \in \{a_{i_q}, \ldots, a_{i_{2q}}\}\}$, from definition of a connecting structure. Furthermore, observe that $|A_1| = q$ and $|B_1| = \binom{q+1}{2} \geq q$ for all $q \geq 1$. It follows that the subgraph of G induced by A_1 and B_1 contains a biclique of order q.

Assume that all tripes in A' are of color 2. In this case we set $A_2 = \{a_{i_{q+1}}, \ldots, a_{i_{2q}}\}$ and $B_2 = \{f(\{u, v\}) | u, v \in \{a_{i_1}, \ldots, a_{i_{q+1}}\}\}$ and then apply regarding A_2 and B_2 the same reasoning as in the previous paragraph regarding A_1 and B_1.

Assume now that the color of all tripes in A' is 3. Consider the path

$$a_{i_1}, f(\{a_{i_1}, a_{i_2}\}), a_{i_2}, \ldots, a_{i_{M-1}}, f(\{a_{i_{M-1}}, a_{i_M}\}), a_{i_M}$$

By definition of M, the length of this path is at least s. Furthermore, observe that this path is induced. Indeed, the only possible chord is between some a_{i_x} and $f(\{a_{i_y}, a_{i_{y+1}}\})$ such that $x \neq y$ and $x \neq y + 1$. Then either $x < y$ or $x > y + 1$. In both cases such a chord is impossible according to the condition of color 3. □

3.2 Grid structures with large bicliques

In a graph, a (k, t) *grid structure* is a family of $k \times t$ vertex sets $V_{i,j}$, $i = 1, \ldots, k$, $j = 1, \ldots, t$. We call $V_{i,j}$ the set in the i-th row and j-th column. A *transversal* in a grid structure is a collection of sets containing exactly one set from each row. A transversal is *independent* if no two vertices in different sets are adjacent.

Lemma 4. *For each $k \geq 2$, s and q, there is a number $C(k, s, q)$ such that any graph, having a $(k, C(k, s, q))$ grid structure with sets of size at most s and with no independent transversal, has a biclique $K_{q,q}$.*

Proof. For $k = 2$, the statement follows with a double application of the Pigeonhole Principle. In particular, we define $r := R(1, s^q, q)$, $C(2, s, q) := R(1, s^r, q)$ and consider an arbitrary collection A of r sets from the first row. Each set in the second row has a neighbor in each set of the first row, since no transversal is independent. Therefore, the family of the sets in the second row can be colored with at most s^r colors so that all sets of the same color have a common neighbor in each of the r chosen sets of collection A. By the choice of $C(2, s, q)$, one of the color classes contains a collection B of at least q sets. For each set in A, we choose a vertex which is a common neighbor for all sets in B and denote the set of r chosen vertices by U. The vertices of U can be colored with at most s^q colors so that all vertices of the same color have a common neighbor in each of the q sets of collection B. By the choice of r, U contains a color class U_1 of least q vertices. For each set in B, we choose a vertex which is a common neighbor for all vertices of U_1 and denote the set of q chosen vertices by U_2. Then U_1 and U_2 from a biclique $K_{q,q}$.

For $k > 2$, we define $C(k, s, q) := MR(k, C(2, s, q))$ (see Lemma 2 for the definition of the number MR). Since the grid structure has no independent transversal, by

Lemma 2 it must contain two rows with $C(2, s, q)$ sets in each so that the two collections of sets form a $(2, C(2, s, q))$ grid structure with no independent transversal. By the first part of the lemma, this structure contains a biclique $K_{q,q}$. □

3.3 Flowers and bouquets

A *flower centered at* $\{a, b\}$, also called an ab-flower, consists of two distinct vertices a, b and a number of pairwise vertex disjoint induced paths connecting them. Every path in a flower will be called a *petal*. In other words, a petal is an induced path, not including a and b, such that a is adjacent to one terminal vertex of this path and b is adjacent to the other one (of course these terminal vertices may coincide in case of path of length 1). A flower with p petals will be called a p-flower. A *bouquet centered at a set of vertices* B consists of ab-flowers centered at each pair $a, b \in B$ such that no two flowers share a non-central vertex. A bouquet of p-flowers centered at a set of q vertices will be called a (p, q)-bouquet.

In this section we show that every (P_s, K_t)-free graph with a sufficiently large path contains a big bouquet with many petals in each flower. As a step toward this goal, we introduce an auxiliary structure called a *multipattern*.

A *pattern* $Z = (a, P)$ in a graph G is an induced subgraph of G consisting of a (not necessarily chordless) path P and a vertex a outside P such that a is adjacent to the first and the last vertex of P and possibly to some other vertices of P, and the subpath of P between any two consecutive neighbors of a on P is induced (i.e. chordless). If a has at least m neighbors on P, we say that the pattern $Z = (a, P)$ is m-strong.

A *multipattern* of size r in G is a sequence $(Z_1 = (a_1, P_1), \ldots, Z_r = (a_r, P_r))$ of r patterns such that for each $i > 1$, Z_i is a pattern in the subgraph of G induced by the vertices of P_{i-1}. A multipattern is m-strong if each of its patterns is m-strong.

Lemma 5. *For any natural numbers s, t, m, r, there is a number $MP(s, t, m, r)$ such that any (P_s, K_t)-free graph G with a path of length at least $MP(s, t, m, r)$ has an m-strong multipattern $(Z_1, \ldots, Z_r)$.*

Proof. We prove this lemma by induction on r. For $r = 1$, we let $MP(s, t, m, 1)$ be equal $\ell(t, 2m, s)$ (see Lemma 1 for the definition of ℓ) and consider a (P_s, K_t)-free graph G with a path P of length $MP(s, t, m, r)$. Then, by Lemma 1, the subgraph of G induced by the vertices of P must have an induced star $K_{1,2m}$. We denote by a the center of the star. At least m neighbors of a must be located either to the left or to the right of a in the order induced by P. These neighbors together with shortest (i.e. chordless) paths connecting every two consecutive neighbors and together with vertex a create an m-strong pattern in G, which proves the lemma for $r = 1$.

For $r > 1$, we inductively define $MP(s, t, m, r) := MP(s, t, M, r - 1)$, where $M = \max\{m, \ell(t, 2m, s)\}$. Then an (P_s, K_t)-free graph G with a path P of length at least $MP(s, t, m, r)$ has an M-strong multipattern $(Z_1, \ldots, Z_{r-1})$ of size $r - 1$. Since $M \geq \ell(t, 2m, s)$, the path in the pattern Z_{r-1} is of length at least $\ell(t, 2m, s)$. Therefore, as in the basis case $r = 1$, it contains an m-strong pattern Z_r. This completes the proof of the lemma. □

Lemma 6. *For any natural numbers $s, t, p, b \geq 2$, there is a number $B(s, t, p, b)$ such that any (P_s, K_t)-free graph G with a path of length at least $B(s, t, p, b)$ has a (p, b)-bouquet centered at an independent set.*

Proof. Let $B(s, t, p, b) := MP(s, t, s^2pb^2, R(2, 2, \max(t, b)))$ (R is the Ramsey number and MP is defined in Lemma 5). Then, by Lemma 5, any (P_s, K_t)-free graph G with a path of length at least $B(s, t, p, b)$ contains an s^2pb^2-strong multipattern $\mathcal{Z}$ of size $R(2, 2, \max(t, b))$. Since G is K_t-free, $\mathcal{Z}$ contains a sub-multipattern $(Z_1 = (a_1, P_1), \ldots, Z_b = (a_b, P_b))$ such that $\{a_1, \ldots, a_b\}$ is an independent set.

The neighbors of a_1 partition P_1 into vertex disjoint chordless paths each of which has at most $s - 1$ vertices (since G is P_s-free). Let us call these paths *intervals*. Vertex a_2 has neighbors in at least p of these intervals (in fact since each pattern is s^2pb^2 strong and in each interval a_2 can have at most s neighbors, the number of such neighboring intervals is at least spb^2), and each of them can be used to form a petal in the flower centered at $\{a_1, a_2\}$. This proves the lemma for $b = 2$.

For $b > 2$, assume by induction that $(Z_2 = (a_2, P_2), \ldots, Z_b = (a_b, P_b))$ contains a $(p, b-1)$-bouquet centered at vertices $\{a_2, \ldots, a_b\}$. Vertices $\{a_2, \ldots, a_b\}$ are adjacent to at most $sp(b-1)(b-2)$ vertices of the bouquet, these adjacent vertices intersect (use) at most $sp(b-1)(b-2) \leq sp(b-1)^2$ intervals of P_1. Since each interval consists of at most s vertices, vertex a_2 can have at most $s^2p(b-1)^2$ neighbors in these intervals, and since the total number of neighbors of a_2 on P_1 is at least s^2pb^2, it also has neighbors in at least p of the unused intervals (at least $s^2p(b^2 - (b-1)^2) > sp$ of unused neighbors with at most s neighbors per interval), and each of them can be used to form a petal in the flower centered at $\{a_1, a_2\}$. For $2 < i \leq b$, we assume by induction that flowers centered at $\{a_1, a_2\}, \ldots, \{a_1, a_{i-1}\}$ have been added to the bouquet. Collectively, all flowers in the bouquet use at most $sp(b-1)^2 + p(i-1) \leq sp(b-1)^2 + p(b-1) \leq spb(b-1)$ intervals of P_1. Since each interval consists of at most s vertices, vertex a_i can have at most $s^2pb(b-1)$ neighbors in these intervals, and since the total number of neighbors of a_i on P_1 is s^2pb^2, it also has neighbors in at least p of the unused intervals, and each of them can be used to form a petal in the flower centered at $\{a_1, a_i\}$. □

3.4 Proof of Theorem 3

We define $c := C(\lceil s/2 \rceil, s, q)$, $a := C(2, sc, q)$, $b := L(s, a)$, and $z := B(s, t, c, b)$ (for the definitions of numbers C, L and B see Lemmas 4, 3 and 6, respectively). Let G be a graph with a path of length z. If G contains a clique K_t or an induced path P_s, then we are done. So assume G is (K_t, P_s)-free.

By Lemma 6, G contains a (c, b)-bouquet centered at an independent set B of size b. Contract the non-central vertices of each uv-flower ($u, v \in B$) to a single vertex, called the *uv-connecting* vertex. Let X be the set of all connecting vertices. Consider a bipartite graph S with the bipartition B, X where $u \in B$ and $w \in X$ are adjacent if and only if in the graph G vertex u has a neighbour in the set of vertices contracted to w. Clearly S is a connecting structure with $f(\{u, v\})$ being the uv-connecting vertex. By Lemma 3, S has either an induced path of length s or a biclique of order a.

Assume first that S contains an induced path P of length s. This path contains at most $\lceil s/2 \rceil$ vertices of X and each of them represents a set of c petals of size at most

s each. Consider an arbitrary transversal containing one petal from each set. If this transversal is independent (i.e. no two vertices in different petals are adjacent), then by replacing the vertices of X in P by the respective petals of the transversal, we obtain an induced path of length at least s in the original graph G, which is impossible. Therefore, each transversal has at least one edge and hence by Lemma 4 G has a biclique of size q.

Suppose now that the connecting structure S has a biclique of order a. Each connecting vertex of this biclique represents a set of c petals of size at most s each. Therefore, this biclique represents a $(2, a)$ grid structure of G with sets of size at most sc and with no independent transversal. Therefore, by Lemma 4, G has a biclique of size q.

4 Directions of future research

In this paper, we have shown that computing a biclique of order k in a P_s-free graph is fixed-parameter tractable when parameterized by k and s. The main ingredient of the proposed method is Theorem 1 that establishes connection between the Biclique problem and a W[1]-hard problem Induced Path. This might give a hope of a possibility of establishing W[1]-hardness of the Biclique problem by a reduction from the Induced path problem. However, it is not clear how such a reduction would work in the presence of a *large* biclique.

Intuitively, Biclique problem is 'similar' to the Clique problem (but much more resistant to the attempts of proving W[1]-hardness). Does this similarity preserve in the case of P_s-free graphs? In particular, what is the complexity of k-Clique problem in P_s-free graphs where k and s as parameters? In fact it would be a strong result even if this problem shown FPT on k with the power of polynomial depending on s: it will show, for instance, that the Clique problem is FPT for P_5-free graphs, where it is known to be NP-hard by a reduction from the Independent set problem on graphs with large girth [23]. Dániel Marx suggested that an interesting intermediate problem between Biclique and Clique is the Tripartite Clique, i.e. finding out if the given P_s-free graph has a complete 3-partite subgraph with k vertices in each part. Although these problems are closely related to the result proposed in this paper, it is not clear how Theorem 1 can help in their resolution: P_s-free graphs of large treewidth cannot be guaranteed to have a large clique, not even a large tripartite clique, because of the failure of such potential theorems on bipartite graphs. Therefore, if any of these problems is FPT, new methods would be required to establish this.

Finally, it is interesting to see how Theorem 1 can be modified and extended. In particular, assume that the given graph *does not* have a large biclique. In this case if the largeness condition is a *large average degree* then the consequences are very strong: as shown in [19], the considered graph will have an induced subdivision of *any* graph. In this paper, we prove that if the largeness condition is a *long path*, then the absence of a large biclique implies 'just' a long induced path. Can we claim more than that? What if the 'long path' condition is replaced by a stronger 'large treewidth' assumption?

Acknowledgments

Research of the first two authors was supported by the Centre for Discrete Mathematics and Its Applications (DIMAP) at the University of Warwick. The second author also acknowledges a support from EPSRC, grant EP/I01795X/1. Part of this research was done when the third author visited the Institute of Informatics of the Humboldt University of Berlin (hosted by Martin Grohe and Dániel Marx) and the Department of Informatics of the University of Bergen (hosted by Fedor Fomin). Both of the visits took place in August 2011. We would like to thank Dániel Marx for interesting conversations regarding possible further research related to the proposed result. We would also like to thank anonymous reviewers for noticing a number of minor mistakes in the preliminary version of this paper.

References

1. Claudio Arbib and Raffaele Mosca. Polynomial algorithms for special cases of the balanced complete bipartite subgraph problem. *J. Combin. Math. Combin. Comput.*, 39:3–22, 1999.
2. Gabór Bascó and Zsolt Tuza. A characterization of graphs without long induced paths. *J. of Graph Theory*, 14:455–464, 1990.
3. Daniel Binkele-Raible, Henning Fernau, Serge Gaspers, and Mathieu Liedloff. Exact exponential-time algorithms for finding bicliques. *Inf. Process. Lett.*, 111(2):64–67, 2010.
4. Hans L. Bodlaender. On linear time minor tests with depth-first search. *J. Algorithms*, 14(1):1–23, 1993.
5. Hajo Broersma, Petr A. Golovach, Daniël Paulusma, and Jian Song. Updating the complexity status of coloring graphs without a fixed induced linear forest. *Theor. Comput. Sci.*, 414(1):9–19, 2012.
6. Andrei A. Bulatov and Dániel Marx. Constraint satisfaction parameterized by solution size. In *ICALP (1)*, pages 424–436, 2011.
7. Yijia Chen, Marc Thurley, and Mark Weyer. Understanding the complexity of induced subgraph isomorphisms. In *ICALP (1)*, pages 587–596, 2008.
8. Bruno Courcelle, Johann A. Makowsky, and Udi Rotics. Linear time solvable optimization problems on graphs of bounded clique width. In *WG*, pages 1–16, 1998.
9. Erik Demaine, Gregory Z. Gutin, Daniel Marx, and Ulrike Stege. 07281 open problems – structure theory and FPT algorithmcs for graphs, digraphs and hypergraphs. In Erik Demaine, Gregory Z. Gutin, Daniel Marx, and Ulrike Stege, editors, *Structure Theory and FPT Algorithmics for Graphs, Digraphs and Hypergraphs*, number 07281 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2007. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.
10. Reinhard Diestel. *Graph Theory*. Springer-Verlag, third edition, 2005.
11. Jinquan Dong. Some results on graphs without long induced paths. *J. of Graph Theory*, 22:23–28, 1996.
12. Uriel Feige and Shimon Kogan. Hardness of approximation of the balanced complete bipartite subgraph problem. Technical Report MCS04-04, Weizmann Institute of Science, 2004.
13. Michael Fellows, Serge Gaspers, and Frances Rosamond. Multivariate complexity theory. In Edward K. Blum and Alfred V. Aho, editors, *Computer Science: The Hardware, Software and Heart of It*, pages 269–293. Springer, 2011.
14. Michael R. Fellows and Michael A. Langston. On search, decision and the efficiency of polynomial-time algorithms (extended abstract). In *STOC*, pages 501–512, 1989.

15. Petr A. Golovach, Daniël Paulusma, and Jian Song. Coloring graphs without short cycles and long induced paths. In *FCT*, pages 193–204, 2011.
16. David S. Johnson. The NP-completeness column: An ongoing guide. *J. Algorithms*, 8(3):438–448, 1987.
17. Dmitry Korobitsyn. On the complexity of determining the domination number in monogenic classes of graphs. *Diskretnaya Matematika (in Russian)*, 2(3):90–96, 1990.
18. Daniel Král, Jan Kratochvíl, Zsolt Tuza, and Gerhard J. Woeginger. Complexity of coloring graphs without forbidden induced subgraphs. In *WG*, pages 254–262, 2001.
19. Daniela Kühn and Deryk Osthus. Induced subdivisions in $K_{s,s}$-free graphs of large average degree. *Combinatorica*, 24(2):287–304, 2004.
20. Van Bang Le, Bert Randerath, and Ingo Schiermeyer. On the complexity of 4-coloring graphs without long induced paths. *Theor. Comput. Sci.*, 389(1-2):330–335, 2007.
21. Vadim V. Lozin and Raffaele Mosca. Maximum independent sets in subclasses of P_5-free graphs. *Inf. Process. Lett.*, 109(6):319–324, 2009.
22. Vadim V. Lozin and Dieter Rautenbach. Some results on graphs without long induced paths. *Inf. Process. Lett.*, 88(4):167–171, 2003.
23. Owen J. Murphy. Computing independent sets in graphs with large girth. *Discrete Applied Mathematics*, 35(2):167–170, 1992.
24. Bert Randerath and Ingo Schiermeyer. 3-colorability in P for P_6-free graphs. *Discrete Applied Mathematics*, 136(2-3):299–313, 2004.
25. Gerhard J. Woeginger and Jiri Sgall. The complexity of coloring graphs without long induced paths. *Acta Cybernetica*, 15(1):107–117, 2001.