Relating dissociation, independence, and matchings

Felix Bock       Johannes Pardey       Lucia D. Penso       Dieter Rautenbach

Institute of Optimization and Operations Research, Ulm University, Ulm, Germany
{felix.bock,johannes.pardey,lucia.penso,dieter.rautenbach}@uni-ulm.de

Abstract

A dissociation set in a graph is a set of vertices inducing a subgraph of maximum degree at most 1. Computing the dissociation number \( \text{diss}(G) \) of a given graph \( G \), defined as the order of a maximum dissociation set in \( G \), is algorithmically hard even when \( G \) is restricted to be bipartite. Recently, Hosseinian and Butenko proposed a simple \( \frac{4}{3} \)-approximation algorithm for the dissociation number problem in bipartite graphs. Their result relies on the inequality \( \text{diss}(G) \leq \frac{4}{3} \alpha(G - M) \) implicit in their work, where \( G \) is a bipartite graph, \( M \) is a maximum matching in \( G \), and \( \alpha(G - M) \) denotes the independence number of \( G - M \). We show that the pairs \((G, M)\) for which this inequality holds with equality can be recognized efficiently, and that a maximum dissociation set can be determined for them efficiently. The dissociation number of a graph \( G \) satisfies \( \max\{\alpha(G), 2\nu_s(G)\} \leq \text{diss}(G) \leq \alpha(G) + \nu_s(G) \leq 2\alpha(G) \), where \( \nu_s(G) \) denotes the induced matching number of \( G \). We show that deciding whether \( \text{diss}(G) \) equals any of the four terms lower and upper bounding \( \text{diss}(G) \) is NP-hard.

Keywords: Dissociation set; independent set; matching; induced matching
1 Introduction

We consider finite, simple, and undirected graphs, and use standard terminology. A set \( I \) of vertices of a graph \( G \) is a dissociation set in \( G \) if the subgraph \( G[I] \) of \( G \) induced by \( I \) has maximum degree at most 1, and the dissociation number \( \text{diss}(G) \) of \( G \) is the order of a maximum dissociation set in \( G \). Dissociation sets and the dissociation number were introduced as a special vertex-deletion problem by Yannakakis [11] who showed that the dissociation number problem, that is, the problem of deciding whether \( \text{diss}(G) \geq k \) for a given pair \((G, k)\), where \( G \) is a graph and \( k \) is a positive integer, is \( \text{NP} \)-complete even restricted to instances where \( G \) is a bipartite graph. This initial hardness result was strengthened in different ways [2, 8]; in particular, the problem remains \( \text{NP} \)-complete for bipartite graphs of maximum degree at most 3.

Recently, Hosseinian and Butenko [5] proposed a simple \( 4/3 \)-approximation algorithm for the maximum dissociation set problem restricted to bipartite graphs. Their result can be derived from the following two simple inequalities: Let \( G \) be a graph and let \( M \) be a maximum matching in \( G \). Since every independent set in \( G - M \) is a dissociation set in \( G \), we have

\[
\text{diss}(G) \geq \alpha(G - M), \tag{1}
\]

where \( \alpha(H) \) denotes the independence number of a graph \( H \), which is the order of a maximum independent set in \( H \). Now, if \( G \) is bipartite, then one can show

\[
\text{diss}(G) \leq \frac{4}{3} \alpha(G - M). \tag{2}
\]

Since a maximum matching \( M \) in a given bipartite graph \( G \) as well as a maximum independent set \( I \) in the bipartite graph \( G - M \) can be determined efficiently, the combination of (1) and (2) implies that returning \( I \) yields a \( 4/3 \)-approximation for the maximum dissociation set problem in \( G \). We will give the simple proof of (2) that is implicit in [5] below. As our first contribution we show that the extremal graphs for (2) have a very restricted structure, which yields the following.

**Theorem 1.** For a given pair \((G, M)\), where \( G \) is a bipartite graph and \( M \) is a maximum matching in \( G \), one can decide in polynomial time whether (2) is satisfied with equality. Furthermore, in case of equality, one can determine in polynomial time a maximum dissociation set in \( G \).

Next to (1) and (2), there are the following relations between the dissociation number \( \text{diss}(G) \), the independence number \( \alpha(G) \), and the induced matching number \( \nu_s(G) \) of a graph \( G \):

\[
\text{diss}(G) \leq 2\alpha(G), \tag{3}
\]
\[
\text{diss}(G) \geq 2\nu_s(G), \tag{4}
\]
\[
\text{diss}(G) \geq \alpha(G), \quad \text{and} \tag{5}
\]
While these inequalities are all straightforward, the extremal graphs are not easy to describe, and we show the following.

**Theorem 2.** For each of the inequalities (3), (4), (5), and (6), it is NP-hard to decide whether a given graph satisfies it with equality.

In view of the special role of bipartite graphs in this context, it makes sense to consider the bipartite extremal graphs for (3) to (6). It is easy to see that a bipartite graph $G$ satisfies $\text{diss}(G) = 2\alpha(G)$ if and only if $G$ is 1-regular. For a bipartite graph $G$, the equality $\text{diss}(G) = \alpha(G)$ holds if and only if $G$ has no induced matching $M$ intersecting every maximum matching in $G$. Inspecting the proofs in [12] reveals that Zenklusen et al. showed that it is NP-complete to decide, for a given pair $(G, k)$, where $G$ is a bipartite graph and $k$ is a positive integer, whether there is an induced matching $M$ in $G$ of size $|M|$ at most $k$ intersecting every maximum matching in $G$. Unfortunately, the size bound is crucial for their reduction. The complexity of the induced matching number is closely tied to the complexity of the dissociation number [8]. In particular, it is hard for bipartite graphs, and the complexity of recognizing the bipartite extremal graphs for (4) and (6) is open. The close relation between dissociation sets, independent sets, and (induced) matchings also reflects in the obvious relation

$$\text{diss}(G) = \max \{ \alpha(G - M) : M \text{ is an induced matching in } G \}.$$  

Before we proceed to the proofs of our results, we briefly mention related research. In fact, bounds on the dissociation number [3, 4], fast exact algorithms [6, 10], randomized approximation algorithms [6], and fixed parameter tractability [9] were studied. As observed in several references, dissociations sets are the dual of so-called 3-path (vertex) covers, cf. also [1].

## 2 Proofs

The following two subsections contain the proofs of Theorem 1 and Theorem 2.

### 2.1 Structure and recognition of the extremal graphs for (2)

We first give a proof of (2), which is implicit in [5]. After that we consider the extremal graphs. Throughout this subsection, let $G$ be a bipartite graph of order $n$ with partite sets $A$ and $B$, and let $M$ be a maximum matching in $G$. Let $I$ be a maximum dissociation set in $G$. Let $E$ be the induced matching spanned by $I$, that is, $E = E(G[I])$. Note that $|E| \leq \frac{|I|}{2}$. Let $A = A_1 \cup A_2 \cup A_3 \cup A_4$ and $B = B_1 \cup B_2 \cup B_3 \cup B_4$ be partitions of $A$ and $B$ such that

- $E \cap M$ is a perfect matching between $A_1$ and $B_1$,
- $E \setminus M$ is a perfect matching between $A_2$ and $B_2$, 

• $A_3 \cup B_3$ is the set of isolated vertices in $G[I]$,
• $A_4 = A \setminus I$, and $B_4 = B \setminus I$.

Let $\ell = |A_2|$. Since $I \setminus A_2$ is an independent set in $G - M$, we have

$$\alpha(G - M) \geq |I \setminus A_2| = |I| - |A_2| = \text{diss}(G) - \ell.$$  

(7)

Gallai’s theorem implies that $\alpha(G)$ and the vertex cover number $\tau(G)$ of $G$ add up to the order $n$ of $G$, and König’s theorem implies that $\tau(G)$ equals the matching number $|M|$ of $G$, which together implies that $\alpha(G) = n - |M|$. Since $M$ contains $|E \cap M| = |E| - |I| \leq \frac{|I|}{2} - \ell$ edges spanned by $I$, and at most $n - |I|$ further edges, one incident with each vertex of $A_4 \cup B_4 = V(G) \setminus I$, we have

$$|M| \leq |E \cap M| + (n - |I|) \leq \left(\frac{|I|}{2} - \ell\right) + (n - |I|) = n - \ell - \frac{\text{diss}(G)}{2},$$

(8)

and, hence,

$$\alpha(G - M) \geq \alpha(G) = n - |M| \geq n - \left(n - \ell - \frac{\text{diss}(G)}{2}\right) = \ell + \frac{\text{diss}(G)}{2}.$$  

(9)

Adding (7) and (9) yields (2).

We now consider the extremal graphs for (2), which leads to a proof of Theorem 1. Therefore, we suppose that (2) is satisfied with equality. This implies that equality holds throughout the inequality chains (7), (8), and (9). Equality throughout (7) and (9) implies $\ell = \frac{\text{diss}(G)}{4}$ and $\alpha(G - M) = \alpha(G)$. Equality in (8) implies $|E| = |E \cap M| + |E \setminus M| = |E \cap M| + \ell = \frac{|I|}{2}$, which implies that $G[I]$ is 1-regular, or, equivalently, $A_3 = B_3 = \emptyset$. In view of the above value of $\ell$, exactly half the edges of $G[I]$ belong to $M$, or, equivalently, $\ell = |A_1| = |A_2| = |B_1| = |B_2|$. Furthermore, equality in (8) also implies that the matching $M$ contains exactly $n - |I| = |A_4| + |B_4|$ further edges, one incident with each vertex of $A_4 \cup B_4$; these edges match all of $A_4$ into $B_2$ as well as all of $B_4$ into $A_2$, in particular, $\ell \geq |A_4|, |B_4|$. The matching $M$ leaves exactly $(|A_2| - |B_2|) + (|B_2| - |A_2|) = 2\ell - (|A_4| + |B_4|)$ vertices unmatched that all lie in $A_2 \cup B_2$. See Figure 1 for an illustration.

Now, let $M'$ be any maximum matching in $G - M$. Since $\alpha(G - M) = \alpha(G)$, the results of Gallai and König imply that $G - M$ and $G$ have the same matching number, which implies $|M| = |M'|$. Since in $G - M$ the vertices in $A_1$ have all their neighbors in $B_4$, and the vertices in $B_1$ have all their neighbors in $A_4$, the matching $M'$ leaves at least $(|A_1| - |B_4|) + (|B_1| - |A_4|) = 2\ell - (|A_4| + |B_4|)$ vertices in $A_1 \cup B_1$ unmatched. Since $|M| = |M'|$, this actually implies that $M'$ consists of $|B_4|$ edges matching all of $B_4$ into $A_1$, $|A_4|$ edges matching all of $A_4$ into $B_1$, and the $\ell$ edges from $E \setminus M$ that form a perfect matching between $A_2$ and $B_2$. Let $H$ be the graph with vertex set $V(G)$ and edge set $M \cup M'$. The components of $H$ are $M$-$M'$-alternating paths and cycles that traverse the sets $A_i$ and $B_i$ respecting the cyclic order illustrated in Figure 2.
Figure 1: The continuous lines illustrate the edges in $M$ while the dashed lines illustrate those in $E \setminus M$. All remaining edges of $G$ are not illustrated and intersect $A_4 \cup B_4$.

Figure 2: The cyclic order respected by the components of $H$, that is, a cycle in $H$ traverses the sets in the cyclic order $A_1, B_1, A_4, B_2, A_2, B_4, A_1, B_1, A_4, B_2, A_2, B_4, \ldots$.

For components $C$ of $H$ that are cycles, this implies that the length of $C$ is a multiple of 6. Furthermore, there are exactly $2\ell - (|A_4| + |B_4|)$ components of $H$ that are paths; they all have length 4 modulo 6, and, starting with an edge from $M$, they go from $A_1 \cup B_1$ to $A_2 \cup B_2$. For a component $P$ of $H$ that is a path, these structural properties allow to decide the location of the individual vertices. If, for example, the path $P : u_1 u_2 \ldots u_{12} \ldots$ starts in a vertex $u_1$ from $A$ not covered by $M'$, then $u_1, u_7, \ldots \in A_1$, $u_2, u_8, \ldots \in B_1$, $u_3, u_9, \ldots \in A_4$, $u_4, u_{10}, \ldots \in B_2$, $u_5, u_{11}, \ldots \in A_2$, and $u_6, u_{12}, \ldots \in B_4$.

We now formulate a satisfiable 2-SAT formula $f$ such that a satisfying truth assignment allows to derive a (suitable) location of all vertices of $G$ on cycle components of $H$. Therefore, let $C_1, \ldots, C_k$ be the components of $H$ that are cycles. For $i$ in $[k]$, let $C_i : a_i^1 b_i^1 a_i^2 b_i^2 \ldots b_i^n a_i^1$, where $a_i^1 \in A$ and $a_i^1 b_i^1 \in M$. Note that exactly one of the three vertices $a_i^1$, $a_i^2$, and $a_i^3$ belongs to $A_4$, which also determines the location of every other vertex on $C_i$. For every $i$ in $[k]$, we introduce three Boolean variables $x_i^1$, $x_i^2$, and $x_i^3$, and add to $f$ the three clauses $\overline{x_i^1} \lor \overline{x_i^2} \lor \overline{x_i^3}$; and $\overline{x_i^1} \lor \overline{x_i^3}$; where $x_i^j$ being true corresponds to $a_i^j \in A_4$. Now, we consider an edge $ab$ of $G$ that does not belong to $M \cup M'$ such that at least one endpoint of $ab$ lies on a cycle component of $H$. The structural properties imply that $a \in A_4$ or $b \in B_4$ (or both). If $a$ lies on a cycle component $C$ of $H$, $b$ lies on a path component of $H$, and $b \notin B_4$, then we add to $f$ the clause $x_i^j$, where $i$ and $j$ are such that $C$ is $C_i$, and the distance of $a_i^j$ and $a$ on $C$ is 0 modulo 6. Note
that, if \( b \in B_4 \), then no clause is added to \( f \). Similarly, if \( a \) lies on a path component of \( H \), \( b \) lies on a cycle component of \( H \), and \( a \not\in A_4 \), then we proceed analogously by adding to \( f \) the clause \( x_i^1 \) corresponding to the condition \( x_i^1 \in A_4 \) that is equivalent to the condition \( b \in B_4 \). Finally, if \( a \) and \( b \) both lie on cycle components of \( H \), say \( a = a_i^j \) and \( b = b_i^j \), then we add to \( f \) the clause \( x_i^j \lor x_i^{j''} \), where \( j'' \) is the uniquely determined index such that \( a_i^{j''} \) is the unique vertex in \( \{a_i^1, a_i^2, a_i^3\} \) whose distance to \( b_i^j \in B_4 \) within \( C_i \) is equivalent to 3 modulo 6. This completes the construction of \( f \). Clearly, setting \( x_i^j \) to true for every \( i \) and \( j \) with \( a_i^j \in A_4 \) and false otherwise, yields a satisfying truth assignment for \( f \), that is, if (2) is satisfied with equality, then \( f \) is satisfiable.

Given a bipartite graph \( G \) and a maximum matching \( M \) in \( G \), one can in polynomial time

- determine a bipartition \( A \) and \( B \) of \( G \),
- determine a maximum matching \( M' \) of \( G - M \),
- check the necessary condition \( |M| = |M'| \),
- construct the auxiliary graph \( H \),
- check the necessary condition that all cycle components of \( H \) have length 0 modulo 6, and all path components of \( H \) have length 4 modulo 6,
- suitably assign the vertices in path components of \( H \) to the sets \( A_i \) and \( B_i \) with \( i \in \{1, 2, 4\} \) as described above,
- check the necessary condition that all edges of \( G \) that do not belong to \( M \cup M' \) and connect vertices in path components of \( H \) intersect \( A_4 \cup B_4 \), and
- set up the 2-SAT formula \( f \) as described above, check its satisfiability, and, in case of satisfiability, determine a satisfying truth assignment.

Note that the construction of \( f \) requires only available knowledge. As we have seen above, if (2) is satisfied with equality, then the necessary conditions mentioned above hold, and \( f \) is satisfiable. Conversely, if the necessary conditions mentioned above hold, and \( f \) is satisfiable, then a satisfying truth assignment allows to suitably assign the vertices in cycle components of \( H \) to the sets \( A_i \) and \( B_i \) with \( i \in \{1, 2, 4\} \) in such a way that all edges of \( G \) that do not belong to \( M \cup M' \) intersect \( A_4 \cup B_4 \), \( A_1 \cup A_2 \cup B_1 \) is a maximum independent set in \( G - M \), and \( A_1 \cup A_2 \cup B_1 \cup B_2 \) is a dissociation set in \( G \), that is, (2) is satisfied with equality. Note that \( H \) may contain cycle components \( C_i \) for which all three variables \( x_i^1 \), \( x_i^2 \), and \( x_i^3 \) are false, even if \( G \) contains edges connecting \( C_i \) to other components of \( H \). In such a case, any of the three vertices \( a_i^1 \), \( a_i^2 \), and \( a_i^3 \) may be located within \( A_4 \), which yields three different valid possibilities. Note furthermore that, in case of equality in (2), the set \( A_1 \cup A_2 \cup B_1 \cup B_2 \), which is efficiently constructible as explained above, is a maximum dissociation set in \( G \).

This complete the proof of Theorem 1.
2.2 Hardness of deciding equality in (3) to (6)

In this subsection, we show Theorem 2.

For the hardness of deciding equality in (3), we suitably adapt Karp’s proof \cite{7} of the NP-completeness of the CLIQUE problem, reducing 3-SAT to the respective problems. Therefore, let \( f \) be an instance of 3-SAT consisting of the clauses \( C_1, \ldots, C_m \) over the Boolean variables \( x_1, \ldots, x_n \).

For the hardness of deciding equality in (3) or (4), we describe the efficient construction of a graph \( G \) such that

\[
\text{\( f \) is satisfiable} \iff \text{diss}(G) = 2\alpha(G) \iff \text{diss}(G) = 2\nu_s(G). \tag{10}
\]

For every clause \( C_i = x \lor y \lor z \) in \( f \), where \( x, y, \) and \( z \) are the three literals in \( C_i \), we introduce the four vertices \( x^i, y^i, z^i, \) and \( c^i \) in \( G \) that induce a clique \( G_i \), where \( x^i, y^i, \) and \( z^i \) are associated with the three literals \( x, y, \) and \( z \) in \( C_i \). Note that \( G \) has order \( 4m \). For every two vertices \( u \) and \( v \) belonging to different cliques \( G_i \) such that the literal associated with \( u \) is the negation of the literal associated with \( v \), we add to \( G \) the edge \( uv \). This completes the construction of \( G \); see Figure 3 for an illustration.

![Figure 3: The graph G for the formula f = C_1 \land C_2 \land C_3 with C_1 = x_1 \lor x_2 \lor x_3, C_2 = \overline{x}_1 \lor x_4 \lor \overline{x}_2, and C_3 = x_1 \lor x_3 \lor x_4.](image)

Clearly, the set \( I = \{c^1, \ldots, c^m\} \) is a maximum independent set of \( G \), in particular, we have \( \alpha(G) = m \). The structure of \( G \) easily implies that \( G \) has a maximum induced matching \( M \) that only contains edges from \( G_1 \cup \ldots \cup G_m \), in fact, any edge in \( M \) between a vertex \( x \) in \( G_i \) and some \( G_j \) for \( i \neq j \) can be replaced by the edge \( xc^i \). Similarly, the graph \( G \) has a maximum dissociation set \( D \) such that all edges induced by \( D \) belong to \( G_1 \cup \ldots \cup G_m \). These observations easily imply that \( G \) satisfies (6) with equality, that is, we have diss\((G) = \alpha(G) + \nu_s(G)\).

As observed by Karp, the formula \( f \) is satisfiable if and only if \( G - I \) has an independent set \( I' \) of order \( m \). If \( f \) is satisfiable, and \( I' \) is as above, then \( I \cup I' \) is a maximum dissociation set in \( G \), and the edges spanned by \( I \cup I' \) form a maximum induced matching in \( G \), that is, we have

\[
\text{diss}(G) = 2\nu_s(G) = 2m = 2\alpha(G).
\]

Conversely, if diss\((G) = 2\alpha(G)\), then \( G \) has a maximum dissociation set \( D \) containing \( I \), and \( D \setminus I \) is an independent set in \( G - I \) of order \( m \), that is, it follows that \( f \) is satisfiable. Similarly, if diss\((G) = 2\nu_s(G)\), then (6) implies \( \nu_s(G) = m \), and
$G$ has a maximum induced matching $M$ covering $I$, and the vertices covered by $M$ not in $I$ form an independent set in $G - I$ of order $m$, that is, again it follows that $f$ is satisfiable. This completes the proof of (10), which shows the NP-hardness of deciding equality in (3) or (4).

For the hardness of deciding equality in (6), we describe an efficient reduction from the NP-complete INDEPENDENT SET problem. Therefore, let $(G, k)$ be an instance of this problem, that is, the problem of deciding whether $\alpha(G) \geq k$. Possibly by adding isolated vertices to $G$ and increasing $k$ for each added vertex by one, we may assume that $2(k - 1) > n \geq 2$, where $n$ is the order of $G$. We describe the efficient construction of a graph $H$ such that $\alpha(G) \geq k$ if and only if $(i, 1)$. As noted above, this is equivalent to the satisfiability of $f$, which shows the NP-hardness of deciding equality in (3).

For the hardness of deciding equality in (5), we describe an efficient construction of a graph $H$ such that $f$ is satisfiable if and only if $\text{diss}(H) = \alpha(H)$. For every clause $C_i = x \lor y \lor z$ in $f$, where $x$, $y$, and $z$ are the three literals in $C_i$, we introduce the six vertices $x^{(i,1)}$, $y^{(i,1)}$, $z^{(i,1)}$, $x^{(i,2)}$, $y^{(i,2)}$, and $z^{(i,2)}$ in $H$ that induce a subgraph $H_i$ that is a clique minus the three edges $x^{(i,1)}x^{(i,2)}$, $y^{(i,1)}y^{(i,2)}$, and $z^{(i,1)}z^{(i,2)}$. Similarly as above, the vertices $x^{(i,1)}$, $y^{(i,1)}$, and $z^{(i,1)}$ in $H_i$ are associated with the three literals $x$, $y$, and $z$ in $C_i$. Note that $H$ has order $6m$. For every two vertices $u$ and $v$ belonging to different subgraphs $H_i$ that are associated with literals such that the literal associated with $u$ is the negation of the literal associated with $v$, we add to $H$ the edge $uv$. This completes the construction of $H$; see Figure 4 for an illustration.

![Figure 4: The graph $H$ for the formula $f = C_1 \land C_2 \land C_3$ with $C_1 = x_1 \lor x_2 \lor x_3$, $C_2 = \bar{x}_1 \lor x_4 \lor \bar{x}_2$, and $C_3 = x_1 \lor x_3 \lor x_4$.](image)

Since every dissociation set in $H$ intersects each $H_i$ in at most two vertices, and selecting two vertices with exponent $(i, 2)$ in $H_i$ for each $i$ in $[m]$ yields a dissociation set in $H$, we have $\text{diss}(H) = 2m$. By the structure of $H$, we have $\text{diss}(H) = \alpha(H)$ if and only if $H$ has an independent set that contains, for every $i$ in $[m]$, exactly one of the vertices with exponent $(i, 1)$. As noted above, this is equivalent to the satisfiability of $f$, which shows the NP-hardness of deciding equality in (3).
as all possible edges between the original vertices of $G$ and the vertices of the copy of $(k - 1)K_2$.

If $V$ denotes the vertex set of $G$, then the vertex set of $H$ is $V \cup V' \cup W$, where $V' = \{u' : u \in V\}$, $W$ is the set of the $2(k - 1)$ vertices of the copy of $(k - 1)K_2$, there are all possible edges between $W$ and $V$, and no edges between $V'$ and $W$. The order of $H$ is $2n + 2(k - 1)$. It is easy to see that $\alpha(H) = n + k - 1$, in fact, the set $V'$ together with one vertex on each of the $k - 1$ edges within $W$ yields a maximum independent set in $H$.

Our next goal is to show $\text{diss}(H) = n + 2(k - 1)$. Since $V' \cup W$ is a dissociation set in $H$, we have $\text{diss}(H) \geq n + 2(k - 1)$. Now, let $D$ be a maximum dissociation set in $H$. If $D$ intersects both $V$ and $W$, then $D$ contains exactly one vertex from $V$, one vertex from $W$, and all but one vertices from $V'$, that is, $|D| \geq 1 + n - 1 = n + 1 < n + 2(k - 1)$, which is a contradiction. If $D$ does not intersect $W$, then $|D| \geq |V \cup V'| = 2n < n + 2(k - 1)$, which is a contradiction. Thus, the set $D$ does not intersect $V$, which, by the choice of $D$, implies $D = V' \cup W$, and, hence, we obtain $\text{diss}(H) = |D| = |V' \cup W| = n + 2(k - 1)$ as desired.

In view of the $k - 1$ independent edges in $W$, we have $\nu_s(H) \geq k - 1$. Since $\text{diss}(H) = \alpha(H) + k - 1$, in order to complete the proof, it suffices to show that $\alpha(G) \geq k$ if and only if $\nu_s(H) \geq k$. If $I$ is an independent set in $G$ of order at least $k$, then $\{uu' : u \in I\}$ is an induced matching in $H$, hence $\alpha(G) \geq k$ implies $\nu_s(H) \geq k$. Now, suppose that $\nu_s(H) \geq k$, and let $M$ be a maximum induced matching in $H$ containing as few edges with both endpoints in $V$ as possible. If $M$ contains an edge with both endpoints in $W$, then all edges in $M$ have both endpoints in $W$, which implies the contradiction $|M| \leq k - 1$. If $M$ contains an edge between $W$ and $V$, then we obtain the contradiction $|M| = 1$. Hence, no edge in $M$ covers any vertex of $W$. If $uv \in M$ for $u, v \in V$, then $M \setminus \{uv\} \cup \{uv'\}$ is a maximum induced matching in $H$ containing fewer edges with both endpoints in $V$ than $M$. Hence, the choice of $M$ implies that the set of $|M| \geq k$ vertices from $V$ covered by an edge from $M$ is an independent set in $G$, that is, $\alpha(G) \geq k$.

This completes the proof of Theorem 2.

3 Conclusion

Our initial motivation to consider the extremal graphs for (2) was the attempt to improve the approximation algorithm of Hosseinian and Butenko [5]. This remains to be done. As explained in the introduction, the complexity of recognizing the bipartite extremal graphs for (4), (5), and (6) remains open. We believe that all three problems are hard. The estimates (11), (2), and (3) imply $\alpha(G - M) \leq \text{diss}(G) \leq 2\alpha(G - M)$ for a given graph $G$ and a given maximum matching $M$ in $G$. Theorem 2 easily implies that the extremal graphs for these two inequalities are also hard to recognize. In fact, if $G$ is a graph with $\alpha(G) \geq 3$, the graph $H$ arises from the disjoint union of $G$ and $K_{n(G)}$ by adding all possible edges between $V(G)$ and $V(K_{n(G)})$, and $M$ is a perfect matching of $H$ using only edges between $V(G)$ and $V(K_{n(G)})$,}
then $\alpha(H) = \alpha(H - M) = \alpha(G)$ and $\text{diss}(H) = \text{diss}(H - M) = \text{diss}(G)$. This implies that, for every $c \in \{1, 2\}$, the graph $G$ satisfies $\text{diss}(G) = c \cdot \alpha(G)$ if and only if the graph $H$ satisfies $\text{diss}(H) = c \cdot \alpha(H - M)$.

References

[1] S. Bessy, P. Ochem, and D. Rautenbach, On the König-Egerváry theorem for $k$-paths, Journal of Graph Theory 91 (2019) 73-87.

[2] R. Boliac, K. Cameron, and V.V. Lozin, On computing the dissociation number and the induced matching number of bipartite graphs, Ars Combinatoria 72 (2004) 241-253.

[3] B. Brešar, F. Kardoš, J. Katrenič, and G. Semanišin, Minimum $k$-path vertex cover, Discrete Applied Mathematics 159 (2011) 1189-1195.

[4] B. Brešar, M. Jakovac, J. Katrenič, G. Semanišin, and A. Taranenko, On the vertex $k$-path cover, Discrete Applied Mathematics 161 (2013) 1943-1949.

[5] S. Hosseinian and S. Butenko, An improved approximation for maximum $k$-dependent set on bipartite graphs, Discrete Applied Mathematics 307 (2022) 95-101.

[6] F. Kardoš, J. Katrenič, and I. Schiermeyer, On computing the minimum 3-path vertex cover and dissociation number of graphs, Theoretical Computer Science 412 (2011) 7009-7017.

[7] R.M. Karp, Reducibility among combinatorial problems, Complexity of computer computations, Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972, 85-103.

[8] Y. Orlovich, A. Dolgui, G. Finke, V. Gordon, and F. Werner, The complexity of dissociation set problems in graphs, Discrete Applied Mathematics 159 (2011) 1352-1366.

[9] D. Tsur, Parameterized algorithm for 3-path vertex cover, Theoretical Computer Science 783 (2019) 1-8.

[10] M. Xiao and S. Kou, Exact algorithms for the maximum dissociation set and minimum 3-path vertex cover problems, Theoretical Computer Science 657 (2017) 86-97.

[11] M. Yannakakis, Node-deletion problems on bipartite graphs, SIAM Journal on Computing 10 (1981) 310-327.

[12] R. Zenklusen, B. Ries, C. Picouleau, D. de Werra, M.-C. Costa, and C. Bentz, Blockers and transversals, Discrete Mathematics 309 (2009) 4306-4314.