Hardness of Relating Dissociation, Independence, and Matchings

Lucia Draque Penso — Ulm University
Felix Bock — Ulm University
Mitre Dourado — Federal University of Rio de Janeiro
Johannes Pardey — Ulm University
Dieter Rautenbach — Ulmm University

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. 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 always NP-hard for general graphs, and in some cases also for bipartite graphs, but not always.

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 [5] 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 NP-complete even restricted to instances where G is a bipartite graph. This initial hardness result was strengthened in different ways [2, 4]; in particular, the problem remains NP-complete for bipartite graphs of maximum degree at most 3.

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:

$$\begin{aligned}
\text{diss}(G) &\leq 2\alpha(G), & (1)\\
\text{diss}(G) &\geq 2\nu_s(G), & (2)\\
\text{diss}(G) &\geq \alpha(G), \text{ and} & (3)\\
\text{diss}(G) &\leq \alpha(G) + \nu_s(G). & (4)
\end{aligned}$$

While these inequalities are all straightforward, the extremal graphs are not easy to describe, and we show the following.

Theorem 1. *For each of the inequalities (1), (2), (3), and (4), 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 (1) to (4). 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, and in fact it is NP-hard to decide, from recent work together with Mitre Dourado in Rio de Janeiro, to be shown in the talk but not in the extended abstract due to lack of space. The complexity of the induced matching number is closely tied to the complexity of the dissociation number [4]. The close relation between dissociation sets, independent sets, and (induced) matchings also reflects in

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

Before we proceed to the proofs of our results, we briefly mention that dissociations sets are the dual of so-called 3-*path (vertex) covers*, cf. also [1].

2 Hardness of deciding equality in (1) to (4)

In this section, we show Theorem 1. For the hardness of deciding equality in (1), (2), or (3), we suitably adapt Karp's proof [3] 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 $C_1, \dots, C_m$ over the Boolean variables $x_1, \dots, x_n$.

For the hardness of deciding equality in (1) or (2), we describe the efficient construction of a graph G such that

$$f \text{ is satisfiable} \quad \Leftrightarrow \quad \mathrm{diss}(G) = 2\alpha(G) \quad \Leftrightarrow \quad \mathrm{diss}(G) = 2\nu_s(G). \tag{5}$$

For every clause $C_i = x \vee y \vee z$ 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 1.

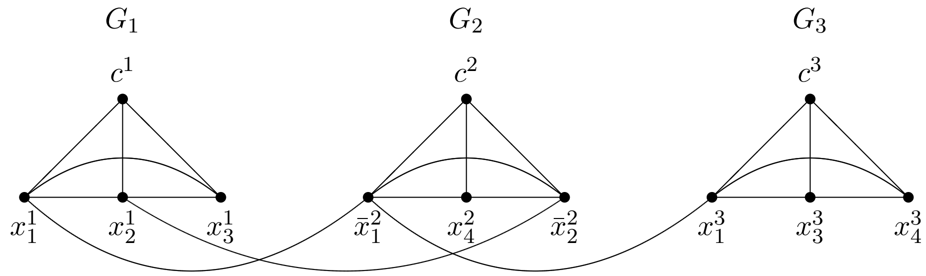

Figure 1: The graph G for the formula $f = C_1 \wedge C_2 \wedge C_3$ with $C_1 = x_1 \vee x_2 \vee x_3$, $C_2 = \bar{x}_1 \vee x_4 \vee \bar{x}_2$, and $C_3 = x_1 \vee x_3 \vee x_4$.

Clearly, the set $I = \{c^1, \dots, c^m\}$ 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 (4) with equality, that is, we have $\mathrm{diss}(G) = \alpha(G) + \nu_s(G)$.

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 $\mathrm{diss}(G) = 2\nu_s(G) = 2m = 2\alpha(G)$. Conversely, if $\mathrm{diss}(G) = 2\alpha(G)$, 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 $\mathrm{diss}(G) = 2\nu_s(G)$, then (4) 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 (5), which shows the NP-hardness of deciding equality in (1) or (2).

For the hardness of deciding equality in (3), we describe the efficient construction of a graph H such that f is satisfiable if and only if $\mathrm{diss}(H) = \alpha(H)$. For every clause $C_i = x \vee y \vee z$ 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 2 for an illustration.

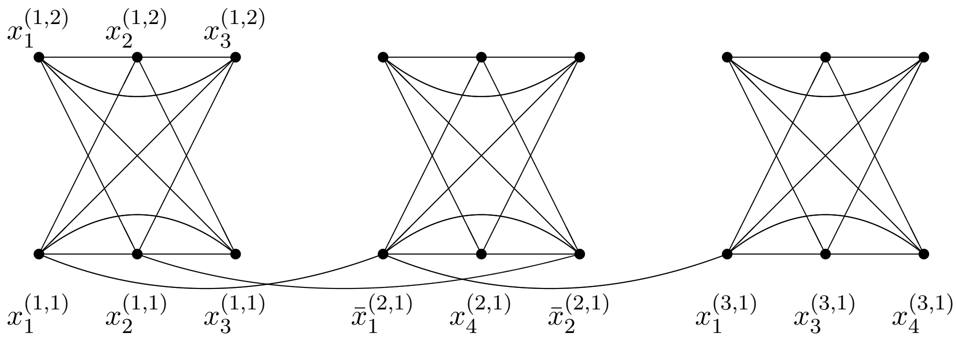

Figure 2: The graph H for the formula $f = C_1 \wedge C_2 \wedge C_3$ with $C_1 = x_1 \vee x_2 \vee x_3$, $C_2 = \bar{x}_1 \vee x_4 \vee \bar{x}_2$, and $C_3 = x_1 \vee x_3 \vee x_4$.

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 $\mathrm{diss}(H) = 2m$. By the structure of H, we have $\mathrm{diss}(H) = \alpha(H)$ 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).

For the hardness of deciding equality in (4), 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 (4) does *not* hold with equality. The graph H arises from G

- by adding, for every vertex u of G, a new vertex u' as well as the edge uu', and
- by adding a disjoint copy of $(k-1)K_2$, that is, $k-1$ further independent edges, as well 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, $|D| \leq 1+1+n-1 = n+1 < n+2(k-1)$, which is a contradiction. If D does not intersect W, then $|D| \leq |V \cup V'| = 2n < n+2(k-1)$, 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 $M \setminus \{uv\} \cup \{uu'\}$ 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 1.

Interestingly, the speaker was born in the year ICGT was created. Therefore, quite a good year!

References

[1] S. Bessy, P. Ochem, and D. Rautenbach, On the Kőnig-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] 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.

[4] 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.

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