

Disjoint Covering of Bipartite Graphs with s -Clubs

Angelo Monti¹  Blerina Sinaimeri² 

¹Department of Computer Science, Sapienza University, Rome, Italy

²AI, Data and Decision Sciences Department, Luiss University of Rome, Italy

Submitted: Sept. 2025 Accepted: Sept. 2026 Published: October 2026

Article type: Regular paper Communicated by: P. Kindermann

Abstract. For a positive integer s , an s -club in a graph G is a set of vertices inducing a subgraph with diameter at most s . As generalizations of cliques, s -clubs offer a flexible model for real-world networks. This paper addresses the problems of partitioning and disjoint covering of vertices with s -clubs on bipartite graphs. First we consider the (k, s) -PC problem where ask whether the vertices of G can be partitioned into at most k disjoint s -clubs. We prove that for any fixed $k \geq 2$ and for any fixed odd $s \geq 3$ or even $s \geq 8$, the (k, s) -PC problem is NP-complete even for bipartite graphs. Note that our NP-completeness result is stronger than the one in Abbas and Stewart (1999), as we assume that both s and k are constants and not part of the input.

Additionally, we study the Maximum Disjoint (t, s) -Club Covering problem $((t, s)$ -MAX-DCC), which aims to find a collection of vertex-disjoint (t, s) -clubs (i.e. s -clubs with at least t vertices) that covers the maximum number of vertices in G . We prove that it is NP-hard to achieve an approximation factor of $\frac{95}{94}$ for $(t, 3)$ -MAX-DCC for any fixed $t \geq 8$ and for $(t, 2)$ -MAX-DCC for any fixed $t \geq 5$ even for bipartite graphs. Previously, results were known only for $(3, 2)$ -MAX-DCC. Finally, we provide a polynomial-time algorithm for $(2, 2)$ -MAX-DCC resolving an open problem from Dondi et al. (2019).

1 Introduction

For a positive integer s , an s -club in a graph G is a set of vertices that induces a subgraph of G of diameter at most s . Clubs are generalizations of cliques (1-clubs are exactly cliques) and offer a wider and more practical way to model real-world interactions [17, 18, 16, 15, 21, 22]. Partitioning a graph into cliques is important for clustering and community detection. Consequently, there has been research into partitioning graphs into s -clubs as a way to extend these methods to more flexible and realistic groupings. In this paper we focus exclusively on bipartite graphs. Bipartite graphs are of particular interest due to their wide range of applications in various fields such as

Part of this work has appeared in [19]

E-mail addresses: monti@di.uniroma1.it (Angelo Monti) bsinaimeri@sluiss.it (Blerina Sinaimeri)



This work is licensed under the terms of the [CC-BY](https://creativecommons.org/licenses/by/4.0/) license.

scheduling, matching problems, recommendation systems and network flow optimization [2]. They also provide a particularly interesting setting for s -club partitioning and covering. Indeed, since bipartite graphs contain no triangles, clique-based notions of cohesion are limited in this setting, whereas s -clubs can still capture tightly connected groups through short alternating paths. This is especially relevant in bipartite networks, where preserving the underlying two-mode structure is often important for community detection and network analysis [3, 5]. Thus, s -clubs provide a natural framework for modeling cohesive substructures in bipartite graphs.

We examine two closely related problems involving the partitioning and covering of a graph's vertices using s -clubs. The first problem is the Minimum Partition s -Club problem, where the objective is to partition the vertices of a graph G into the minimum number of disjoint s -clubs. This problem is NP-hard for $s \geq 2$ [10], even when restricted to bipartite or chordal graphs [1, 6]. Clearly from these results we have that the decision version of this problem, where for a fixed s , given a graph G and an integer k , we decide whether it is possible to partition the vertices of G into at most k disjoint s -clubs, is NP-complete. However, this does not address the complexity of the decision problem when k is also fixed (that is both s and k are not part of the input). We thus, consider the problem below.

k -Partition s -Club problem $((k, s)$ -PC)

Instance: A graph $G = (V, E)$.

Question: Is there a partition of V into at most k vertex disjoint s -clubs?

Previous studies have explored the complexity of this problem for some fixed values of s and k . For $k = 1$, the problem is equivalent to determining the diameter of a graph and thus is trivially solvable in polynomial time. For $s = 1$ the problem is equivalent to determine whether there exists a partition of the vertices into k cliques, that is equivalent to determine whether there exists a k -coloring of the complement graph. This problem is NP-complete for any $k \geq 3$ (see [14]) and polynomial for $k = 2$. More recently, Bazgan et al. [4] showed that when we restrict to split graphs the problem $(2, 2)$ -PC is NP-complete. When we restrict to bipartite graphs, $s = 1$ corresponds to the problem of determining whether there exists a perfect matching of size k in a graph and is thus polynomial. Again for bipartite graphs and $s = 2$ and $k = 2$ the problem is polynomial¹[13]. This is a distinction with the split graphs where already $(2, 2)$ -PC problem is NP-complete. In this paper we show that for bipartite graphs and any fixed odd $s \geq 3$ or even $s \geq 8$ and for any fixed $k \geq 2$ the (k, s) -PC problem is NP-complete. Note that our NP-completeness result is stronger than the one in [1] as we assume that both s and k are constants and not part of the input.

In some real-world applications, it may not be feasible to partition all vertices into s -clubs. To address such cases, a variant of this problem, known as the Maximum Disjoint (t, s) -Club Covering problem $((t, s)$ -MAX-DCC), was introduced by Dondi et al. [12]. This problem seeks to find, given a graph G , a collection of disjoint (t, s) -clubs that covers the maximum number of vertices in G . A (t, s) -club is an s -club with at least t vertices. The concept of (t, s) -clubs extends that of s -clubs by adding a minimum size constraint, and is motivated by applications where identifying large, well-connected subgraphs is important (see, e.g., [16, 12]). Hence, the second problem we consider is the following:

¹Notice that if we do not require the s -clubs to be disjoint, then the case $s = 2$, $k = 2$ is shown to be NP-complete for general graphs [11].

Maximum Disjoint (t, s) -Club Covering problem $((t, s)$ -MAX-DCC)

Instance: A graph $G = (V, E)$.

Required: A collection of vertex disjoint (t, s) -clubs that covers the maximum number of vertices in V .

To the best of our knowledge the only cases considered in literature are the cases $s = 2$ and $s = 3$ [12]. In particular, for $s = 2$ in [12] it is proved that $(3, 2)$ -MAX-DCC is APX-hard and the case $(2, 2)$ -MAX-DCC is left open. Here we show that $(t, 2)$ -MAX-DCC is APX-hard for any fixed $t \geq 5$ even for bipartite graphs and $(2, 2)$ -MAX-DCC can be solved in polynomial time for general graphs. For $s = 3$ in [12] it is shown that $(2, 3)$ -MAX-DCC is polynomial and the case $(3, 3)$ -MAX-DCC is left as an open problem. We show that $(t, 3)$ -MAX-DCC is APX-hard for any fixed $t \geq 8$, even for bipartite graphs.

The paper is organized as follows: In Section 2 we introduce the definitions we need for the paper. In Section 3 we show that for any fixed $k \geq 2$ and for any fixed $s \geq 3$ with $s \neq 4, s \neq 6$ the (k, s) -PC problem is NP-complete. In Section 4 we prove that it is NP-hard to achieve an approximation factor of $\frac{95}{94}$ for $(t, 3)$ -MAX-DCC for any fixed $t \geq 8$ and for $(t, 2)$ -MAX-DCC for any fixed $t \geq 5$ even for bipartite graphs. On the positive side we provide a polynomial-time algorithm for $(2, 2)$ -MAX-DCC. Finally, in Section 5 we conclude with some open problems.

2 Preliminaries

All the graphs we consider here are undirected and simple (with no loops or multiple edges). For a graph $G = (V, E)$ and a subset $V' \subseteq V$ we denote by $G[V']$ the subgraph induced by the vertices in V' . For any two vertices $u, v \in V$ we denote by $u \text{ --- } v$ a path connecting u and v in G . To clarify notation and avoid confusion, we will slightly abuse notation by writing (u, v) for the edge between u and v , instead of $\{u, v\}$. A subset V' of vertices is called an s -club if the diameter of $G[V']$ is at most s . In other words, every pair of vertices in the s -club can be connected by a path of length at most s within the subgraph. An (t, s) -club is an s -club of at least t vertices.

The *closed neighborhood* of a vertex v in a graph $G = (V, E)$, denoted by $N[v]$, is the set consisting of the vertex v itself and all vertices adjacent to v . Formally, $N[v] = \{v\} \cup \{u \in V \mid (u, v) \in E\}$.

A graph $G = (V, E)$ is *bipartite* if its vertices can be partitioned into two independent sets. We denote it as $G = (V_1, V_2, E)$, where V_1 and V_2 are the independent sets. A tree T_G is said to be a *spanning tree* of a connected graph G if T_G is a subgraph of G (not necessarily induced) and T_G contains all vertices of G . A *rooted tree* is a tree with a special vertex labelled as the *root* of the tree. In a tree, a vertex v is said to be at *level* l if v is at a distance l from the root. The *height* of a tree is the maximum level which occurs in the tree. The *parent* of a vertex v is the vertex connected to v on the path to the root.

For any integer k we denote by $[k]$ the set $\{1, 2, \dots, k\}$. We denote by $2^{[k]}$ the family of all possible subsets of $[k]$.

3 NP-completeness of (k, s) -PC problem

It is claimed in [13] that, for every fixed $k \geq 3$, the corresponding bipartite biclique vertex-partition problem is NP-complete, via a reduction from the k -List Coloring problem. Since, on

bipartite graphs, induced 2-clubs correspond to complete bipartite induced subgraphs, this claim would imply the NP-completeness of $(k, 2)$ -PC. However, the correctness proof of the reduction contains a gap. In the converse direction of Lemma 4, the authors consider a biclique vertex-cover $\mathcal{S} = G_1, \dots, G_k$ of the constructed graph H . They then state that the special edges $u_i v_i$, for $i \in [k]$, belong to distinct bicliques, and conclude that, without loss of generality, one may assume $u_i v_i \in E(G_i)$ for every $i \in [k]$. However, this step is not justified: a biclique vertex-cover is required to cover vertices, not edges. The proof implicitly moves from the fact that the two vertices u_i and v_i are covered to the conclusion that the edge $u_i v_i$ is covered by one biclique. This implication is not valid for a vertex-cover by bicliques. Indeed, in a biclique vertex-cover, the vertices u_i and v_i must each belong to some biclique, but they are not forced to belong to the same biclique. Consequently, the edge $u_i v_i$ need not be contained in any biclique of the cover. This is crucial for the reduction, because the converse argument uses the pair u_i, v_i as an anchor for color i . If a biclique G_i contains both u_i and v_i , then any original vertex placed in G_i can be associated with color i , using the adjacencies to u_i and v_i to infer that this color belongs to its list. If, instead, u_i and v_i are covered by different bicliques, this anchoring is lost. Consequently, the sets obtained by deleting u_i and v_i from the bicliques do not necessarily define the color classes of a valid k -list coloring. Hence the implication from the existence of a biclique vertex-cover of H to the existence of a corresponding k -list coloring of the original graph is not established by the proof as written.

In this section we prove the NP-completeness of (k, s) -PC for every $k \geq 2$ and odd $s \geq 3$ and even $s \geq 8$. We start by proving that $(2, 3)$ -PC is NP-complete even for bipartite graphs. We will reduce from the Monotone 3-SAT problem [14]. The reduction is inspired by the proof in [4] where it is shown that the $(2, 2)$ -PC problem is NP-complete for split graphs.

Monotone 3-SAT

Instance: A set X of variables, a collection C of clauses over X , which contains either only negated variables or only positive variables and such that for each clause $c \in C$, $|c| = 3$.

Question: Is there a satisfying truth assignment for C ?

Lemma 1. *The $(2, 3)$ -PC problem is NP-complete even when restricted to bipartite graphs.*

Proof: Membership in NP is immediate: given a partition (X, Y) of $V(G)$, one can check in polynomial time whether both $G[X]$ and $G[Y]$ have diameter at most 3.

For NP-hardness, we reduce from MONOTONE 3-SAT. Let ϕ be an instance of MONOTONE 3-SAT and let

$$X(\phi) = \{x_1, \dots, x_n\} \quad \text{and} \quad C(\phi) = \{c_1, \dots, c_m\}$$

be the sets of variables and clauses, respectively. We write

$$C^+ := \{c_i \in C(\phi) \mid c_i \text{ is positive}\}, \quad C^- := \{c_i \in C(\phi) \mid c_i \text{ is negative}\}.$$

Notice that we may assume that $|C^+| \geq 2$ and $|C^-| \geq 2$. Indeed, if all clauses have the same sign, then ϕ is trivially satisfiable: if all clauses are positive we set every variable to TRUE, and if all clauses are negative we set every variable to FALSE. Otherwise, if ϕ contains exactly one positive clause (respectively exactly one negative clause), we simply duplicate that clause. This does not affect satisfiability, and after this modification we have $|C^+| \geq 2$ and $|C^-| \geq 2$. From ϕ we construct a graph $G = G(\phi)$ as follows.

Vertices.

- For each variable x_i , create a vertex v_i . Let

$$V_{\text{var}} := \{v_1, \dots, v_n\}.$$

- For each positive clause $c_i \in C^+$, create two vertices a_i^+ and b_i^+ ; for each negative clause $c_i \in C^-$, create two vertices a_i^- and b_i^- . Let

$$Q_i^+ := \{a_i^+, b_i^+\}, \quad Q_i^- := \{a_i^-, b_i^-\},$$

$$V_{\text{cl}}^+ := \bigcup_{c_i \in C^+} Q_i^+, \quad V_{\text{cl}}^- := \bigcup_{c_i \in C^-} Q_i^-, \quad V_{\text{cl}} := V_{\text{cl}}^+ \cup V_{\text{cl}}^-.$$

- For every pair of distinct positive clauses $c_i, c_j \in C^+$ with $i < j$, create four vertices

$$P_{i,j}^+ := \{p_{a_i a_j}^+, p_{a_i b_j}^+, p_{b_i a_j}^+, p_{b_i b_j}^+\}.$$

Similarly, for every pair of distinct negative clauses $c_i, c_j \in C^-$ with $i < j$, create four vertices

$$P_{i,j}^- := \{p_{a_i a_j}^-, p_{a_i b_j}^-, p_{b_i a_j}^-, p_{b_i b_j}^-\}.$$

Let

$$V_{\text{pair}}^+ := \bigcup_{\substack{c_i, c_j \in C^+ \\ i < j}} P_{i,j}^+, \quad V_{\text{pair}}^- := \bigcup_{\substack{c_i, c_j \in C^- \\ i < j}} P_{i,j}^-, \quad V_{\text{pair}} := V_{\text{pair}}^+ \cup V_{\text{pair}}^-.$$

- Add four special vertices u^+, u^-, v^+, v^- , and let

$$V_{\text{sp}} := \{u^+, u^-, v^+, v^-\}.$$

Finally, set

$$V(G) = V_{\text{var}} \cup V_{\text{cl}} \cup V_{\text{pair}} \cup V_{\text{sp}}.$$

Edges.

- Suppose a clause c_i contains the variables x_j, x_ℓ, x_r . If $c_i \in C^+$, then connect each of a_i^+, b_i^+ to each of v_j, v_ℓ, v_r . If $c_i \in C^-$, then connect each of a_i^-, b_i^- to each of v_j, v_ℓ, v_r . Let E_{cv} denote the set of all these clause-variable edges.
- For every pair $c_i, c_j \in C^+$ with $i < j$, connect each pair-vertex to the two clause vertices indicated by its subscript:

$$\begin{array}{cccc} p_{a_i a_j}^+ a_i^+, & p_{a_i a_j}^+ a_j^+, & p_{a_i b_j}^+ a_i^+, & p_{a_i b_j}^+ b_j^+, \\ p_{b_i a_j}^+ b_i^+, & p_{b_i a_j}^+ a_j^+, & p_{b_i b_j}^+ b_i^+, & p_{b_i b_j}^+ b_j^+. \end{array}$$

Define the edges for $P_{i,j}^-$ analogously. Let E_{pc} be the union of all these pair-clause edges.

- Connect v^+ to every vertex of V_{pair}^+ , and connect v^- to every vertex of V_{pair}^- .
- Connect u^+ to every vertex of $V_{\text{pair}}^+ \cup V_{\text{var}}$, and connect u^- to every vertex of $V_{\text{pair}}^- \cup V_{\text{var}}$.

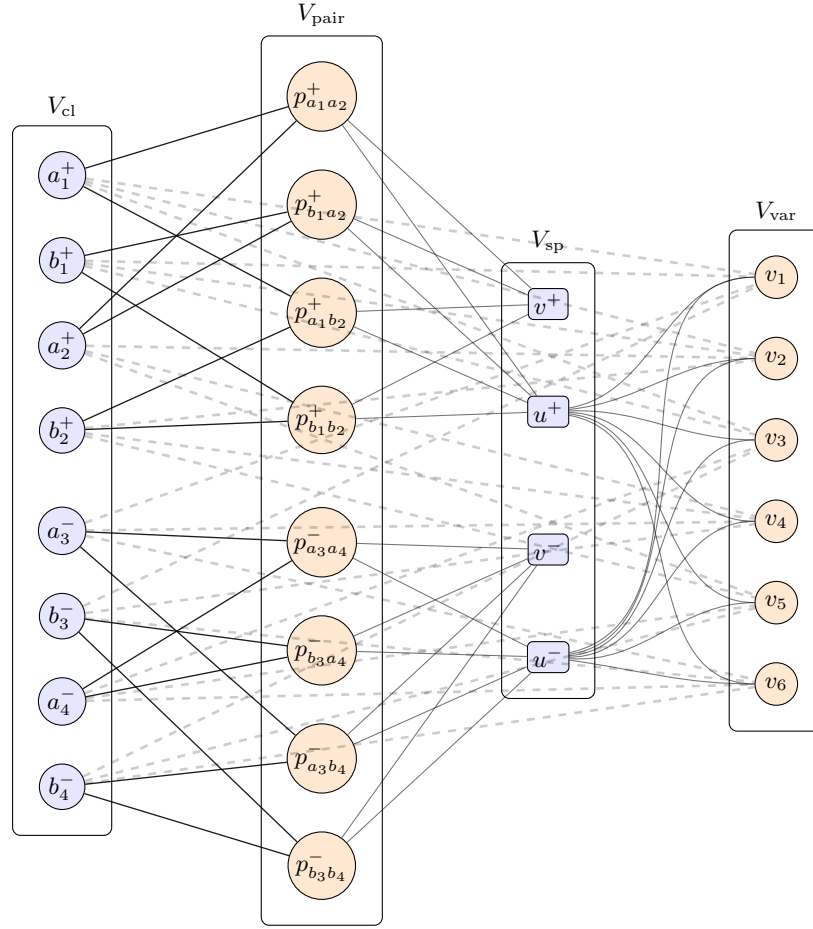


Figure 1: An example of the graph $G(\phi)$ for $\phi = (x_1 \vee x_2 \vee x_3) \wedge (x_2 \vee x_4 \vee x_5) \wedge (\overline{x_1} \vee \overline{x_4} \vee \overline{x_6}) \wedge (\overline{x_3} \vee \overline{x_5} \vee \overline{x_6})$. For readability, the clause-variable edges are drawn as dashed lines. The two colors of the vertices indicate the bipartition of $G(\phi)$.

Let $E(G)$ be the union of all edges defined above.

Figure 1 illustrates the construction for a small instance.

By construction, G is bipartite with bipartition

$$L := V_{\text{var}} \cup V_{\text{pair}} \quad \text{and} \quad R := V_{\text{cl}} \cup V_{\text{sp}}.$$

We now prove that ϕ is satisfiable if and only if $V(G)$ can be partitioned into two sets whose induced subgraphs both have diameter at most 3.

($\Rightarrow$) Assume that ϕ is satisfiable, and let τ be a satisfying truth assignment.

We partition the variable vertices according to τ :

$$V_{\text{var}}^+ := \{v_i \mid \tau(x_i) = \text{TRUE}\}, \quad V_{\text{var}}^- := \{v_i \mid \tau(x_i) = \text{FALSE}\}.$$

We then define

$$\begin{aligned} X &:= V_{\text{var}}^+ \cup V_{\text{cl}}^+ \cup V_{\text{pair}}^+ \cup \{u^+, v^+\}, \\ Y &:= V_{\text{var}}^- \cup V_{\text{cl}}^- \cup V_{\text{pair}}^- \cup \{u^-, v^-\}. \end{aligned}$$

Clearly, (X, Y) is a partition of $V(G)$.

We show that $G[X]$ has diameter at most 3; the argument for $G[Y]$ is symmetric.

Observe first that in $G[X]$:

- u^+ is adjacent to every vertex of $V_{\text{var}}^+ \cup V_{\text{pair}}^+$;
- v^+ is adjacent to every vertex of V_{pair}^+ .

Hence:

- any two vertices of V_{var}^+ are at distance at most 2 via u^+ ;
- any vertex of V_{var}^+ and any vertex of V_{pair}^+ are at distance at most 2 via u^+ ;
- any two vertices of V_{pair}^+ are at distance at most 2 via either u^+ or v^+ ;
- u^+ and v^+ are at distance 2 through any vertex of V_{pair}^+ .

It remains to consider the vertices in V_{cl}^+ . Let c_i be a positive clause and let $z \in \{a_i^+, b_i^+\}$ be one of its clause vertices. Since the assignment τ satisfies ϕ , the clause c_i contains some variable x_j with $\tau(x_j) = \text{TRUE}$. Hence $v_j \in V_{\text{var}}^+$ and, by construction of G , the vertex z is adjacent to v_j . Therefore

$$u^+ - v_j - z$$

is a path of length 2 in $G[X]$. Thus every vertex of V_{cl}^+ is at distance at most 2 from u^+ .

Moreover, recall that we assumed $|C^+| \geq 2$. Hence, besides c_i , there exists another positive clause. By construction, this implies that there is a vertex $p \in V_{\text{pair}}^+$ adjacent to z . Since v^+ is adjacent to every vertex of V_{pair}^+ , we obtain a path

$$z - p - v^+$$

of length 2 in $G[X]$. Therefore every vertex of V_{cl}^+ is also at distance at most 2 from v^+ . It follows that every vertex of V_{cl}^+ is at distance at most 3 from every vertex of $V_{\text{var}}^+ \cup V_{\text{pair}}^+ \cup \{u^+, v^+\}$.

It remains to consider distances between two vertices of V_{cl}^+ . Let $z, z' \in V_{\text{cl}}^+$. If $z \in Q_i^+$ and $z' \in Q_j^+$ with $i \neq j$, then by construction there is a pair vertex in $P_{i,j}^+$ adjacent to both z and z' . Hence

$$\text{dist}_{G[X]}(z, z') \leq 2.$$

If instead $z, z' \in Q_i^+$, then z and z' are the two clause vertices of the same positive clause c_i . Since τ satisfies c_i , one of the variables of c_i is set to TRUE; let x_j be such a variable. Then $v_j \in V_{\text{var}}^+$, and both z and z' are adjacent to v_j . Therefore

$$z - v_j - z'$$

is a path of length 2 in $G[X]$, and thus

$$\text{dist}_{G[X]}(z, z') \leq 2.$$

Consequently, every two vertices of $G[X]$ are at distance at most 3, and hence

$$\text{diam}(G[X]) \leq 3.$$

By symmetry, $\text{diam}(G[Y]) \leq 3$.

($\Leftarrow$) Suppose that $V(G)$ can be partitioned into two sets (X, Y) such that both $G[X]$ and $G[Y]$ have diameter at most 3. Without loss of generality, assume that $v^+ \in X$.

Claim 1. v^- belongs to Y .

Proof: The neighbors of v^+ are exactly the vertices of V_{pair}^+ , and the neighbors of v^- are exactly the vertices of V_{pair}^- . Since there are no edges between V_{pair}^+ and V_{pair}^- , every path from v^+ to v^- must pass through at least another vertex. Hence every such path has length strictly greater than 3. Therefore $\text{dist}_G(v^+, v^-) \geq 4$, so v^+ and v^- cannot lie in the same part of a 3-club partition. Since $v^+ \in X$, we obtain $v^- \in Y$. $\square$

Claim 2. All vertices of V_{cl}^+ belong to X .

Proof: Suppose that some vertex $z \in V_{\text{cl}}^+$ belongs to Y . Since $v^- \in Y$ and $G[Y]$ has diameter at most 3, we must have $\text{dist}_{G[Y]}(z, v^-) \leq 3$. But this is impossible: from a vertex of V_{cl}^+ one can only move to a variable vertex or to a vertex of V_{pair}^+ , while the only neighbors of v^- are the vertices of V_{pair}^- and moreover there exists no edge connecting a vertex in $V_{\text{pair}}^+ \cup V_{\text{cl}}^+$ and one in V_{pair}^- . Hence every path from z to v^- has length strictly greater than 3. This contradiction shows that $z \notin Y$, and therefore $z \in X$. $\square$

Claim 3. Let $c_i \in C^+$ be a positive clause, and let x, y, z be the variable vertices corresponding to the three variables of c_i . Then at least one of x, y, z belongs to X .

Proof: Let a_i^+, b_i^+ be the two clause vertices representing c_i . By construction the only common neighbors of a_i^+ and b_i^+ are precisely the three variable vertices x, y, z . If all three belonged to Y , then none of the common neighbors of a_i^+ and b_i^+ would belong to X . Therefore, in the induced subgraph $G[X]$, there would be no path of length 2 between a_i^+ and b_i^+ . Moreover, by the previous claim, both a_i^+ and b_i^+ belong to X . Then, since G is bipartite and a_i^+ and b_i^+ lie in the same side of the bipartition, every path between them has even length. In particular, they cannot be connected by a path of length 3. Hence, in $G[X]$, the vertices a_i^+ and b_i^+ could not be at distance at most 3, contradicting the fact that $G[X]$ has diameter at most 3. Thus at least one of x, y, z lies in X . $\square$

Now define the truth assignment τ by setting

$$\tau(x_i) = \begin{cases} \text{TRUE} & \text{if } v_i \in X, \\ \text{FALSE} & \text{if } v_i \in Y. \end{cases}$$

By Claim 3, every positive clause contains a variable whose vertex lies in X , and hence a variable set to TRUE by τ . By the symmetric argument, every negative clause contains a variable whose vertex lies in Y , and hence a variable set to FALSE by τ . Therefore τ satisfies ϕ . $\square$

We next extend this hardness result in two steps. First, we show how to pass from partitioning in 3-clubs to any fixed diameter $s \geq 3$ and $s \neq 4$. Then we show how to lift the number of partitions k from 2 to any fixed $k \geq 2$.

Lemma 2. *For every odd $s \geq 3$ and every even $s \geq 8$, $(2, s)$ -PC is NP-complete on bipartite graphs.*

Proof: Membership in NP is immediate. We know by Lemma 1 that $(2, 3)$ -PC is NP-complete on bipartite graphs so we will reduce from that. Let G be an instance of $(2, 3)$ -PC on bipartite graphs and fix an integer $s \geq 3$, $s \neq 4, s \neq 6$. We construct in polynomial time a bipartite graph G' such that G is a yes-instance of $(2, 3)$ -PC if and only if G' is a yes-instance of $(2, s)$ -PC. We distinguish two cases.

Case 1: s is odd. Then $s = 2q + 3$ for some integer $q \geq 0$. Construct G' from G by attaching to every vertex $v \in V(G)$ a pendant path of length q and we denote by $R(v)$ the nodes in this path. Since adding pendant paths preserves bipartiteness, G' is bipartite.

For every vertex $u \in V(G')$, define $\rho(u) \in V(G)$ as the unique vertex $v \in V(G)$ such that either $u = v$ or u lies on the path $R(v)$ attached to v .

Assume first that G is a yes-instance of $(2, 3)$ -PC, and let (X, Y) be a partition of $V(G)$ into two 3-clubs. Set

$$X' := X \cup \bigcup_{v \in X} R(v), \quad Y' := Y \cup \bigcup_{v \in Y} R(v).$$

Clearly (X', Y') is a partition of $V(G')$.

We claim that $G'[X']$ is an s -club. Let $u, v \in X'$. In $G'[X']$ we can go from u to $\rho(u)$ along the attached path, then from $\rho(u)$ to $\rho(v)$ inside $G[X]$, and finally from $\rho(v)$ to v along the attached path. Hence

$$\text{dist}_{G'[X']}(u, v) = \text{dist}_{G'[X']}(u, \rho(u)) + \text{dist}_{G[X]}(\rho(u), \rho(v)) + \text{dist}_{G'[X']}(\rho(v), v) \leq 2q + 3 = s.$$

Therefore $G'[X']$ has diameter at most s , so it is an s -club. By the same argument, $G'[Y']$ is also an s -club. Hence G' is a yes-instance of $(2, s)$ -PC.

Conversely, assume that G' is a yes-instance of $(2, s)$ -PC, and let (X', Y') be a partition of $V(G')$ into two s -clubs.

For every $v \in V(G)$, let v^* denote the leaf of the path attached to v , that is, the vertex of $R(v)$ farthest from v . Define

$$X := \{v \in V(G) : v^* \in X'\}, \quad Y := \{v \in V(G) : v^* \in Y'\}.$$

Since each attached leaf belongs to exactly one between X', Y' , then (X, Y) is a partition of $V(G)$.

We show that $G[X]$ is a 3-club. Let $u, v \in X$, $u \neq v$. Since by definition $u^*, v^* \in X'$ and $G'[X']$ is connected, every $u^* \text{ --- } v^*$ path in $G'[X']$ must pass through u and v (the attached paths are pendant). Thus

$$\text{dist}_{G'[X']}(u^*, v^*) = \text{dist}_{G[X]}(u, v) + 2q.$$

Because $G'[X']$ is an s -club and $s = 2q + 3$, we obtain

$$\text{dist}_{G[X]}(u, v) = \text{dist}_{G'[X']}(u^*, v^*) - 2q \leq s - 2q = 3.$$

Hence $G[X]$ is a 3-club. Similarly, $G[Y]$ is a 3-club. Therefore G is a yes-instance of $(2, 3)$ -PC.

Case 2: s is even. Since $s \geq 8$, we have that $s = 2q + 6$ for some integer $q \geq 1$. Construct G' by first subdividing every edge (u, v) of G once, adding a node $x_{u,v}$, we denote by S the set of all such vertices and then, similarly as in the previous case, attaching to every vertex $v \in V(G)$

a proper path of length q . We denote by $R(v)$ the nodes in this path. Since, both operations preserve bipartiteness G' is bipartite.

Assume first that G is a yes-instance of $(2, 3)$ -PC, and let (X, Y) be a partition of $V(G)$ into two 3-clubs. Define

$$S_Y := \{x_{uv} \in S : u, v \in Y\},$$

Then set

$$X' := X \cup \bigcup_{v \in X} R(v) \cup (S \setminus S_Y), \quad Y' := Y \cup \bigcup_{v \in Y} R(v) \cup S_Y.$$

Clearly (X', Y') is a partition of $V(G')$. For every vertex $u \in X'$, let $\rho(u)$ be a vertex of X at minimum distance from u in G' . If there is more than one such vertex, choose $\rho(u)$ arbitrarily among them. Notice that

$$\text{dist}_{G'[X']}(u, \rho(u)) = \begin{cases} 0, & \text{if } u \in X, \\ 1, & \text{if } u \in (S \setminus S_Y), \\ \text{dist}_{G'[X]}(u, v) \leq q, & \text{if } u \in R(v) \text{ for some } v \in X. \end{cases}$$

Hence, $\text{dist}_{G'[X']}(u, \rho(u)) \leq q$ as $q \geq 1$. Moreover, for any two vertices $u, v \in X'$, we have that $\text{dist}_{G'[X']}(u, v) = 2 \text{dist}_{G[X]}(\rho(u), \rho(v)) \leq 6$ as by construction both $\rho(u), \rho(v) \in X$ and $G[X]$ is a 3-club.

Let $u, v \in X'$. We have

$$\text{dist}_{G'[X']}(u, v) \leq \text{dist}_{G'[X']}(u, \rho(u)) + \text{dist}_{G'[X']}(u, \rho(v)) + \text{dist}_{G'[X']}(u, v) \leq 2q + 6 = s.$$

Therefore $G'[X']$ has diameter at most s , so it is an s -club. By the same argument, $G'[Y']$ is also an s -club. Hence G' is a yes-instance of $(2, s)$ -PC.

Conversely, assume that G' is a yes-instance of $(2, s)$ -PC, and let (X', Y') be a partition of $V(G')$ into two s -clubs.

For every $v \in V(G)$, let v^* denote the leaf of the path attached to v , that is, the vertex of $R(v)$ farthest from v . Define

$$X := \{v \in V(G) : v^* \in X'\}, \quad Y := \{v \in V(G) : v^* \in Y'\}.$$

Since each attached leaf belongs to exactly one between X', Y' , then (X, Y) is a partition of $V(G)$.

We show that $G[X]$ is a 3-club. Let $u, v \in X$, $u \neq v$. Since by definition $u^*, v^* \in X'$ and $G'[X']$ is connected, every $u^* \text{---} v^*$ path in $G'[X']$ must pass through u and v (the attached paths are pendant). Thus

$$\text{dist}_{G'[X']}(u^*, v^*) = \text{dist}_{G'[X']}(u, v) + 2q = 2 \text{dist}_{G[X]}(u, v) + 2q$$

where the last equality follows by the construction of G' from G . Because $G'[X']$ is an s -club and $s = 2q + 6$, we obtain $\text{dist}_{G[X]}(u, v) \leq 3$. Hence $G[X]$ is a 3-club. Similarly, $G[Y]$ is a 3-club. Therefore G is a yes-instance of $(2, 3)$ -PC. $\square$

Theorem 1. *For every fixed $k \geq 2$ and every fixed odd $s \geq 3$ or even $s \geq 8$, the (k, s) -PC problem is NP-complete.*

Proof: Membership in NP is immediate. We know by Lemma 2 that $(2, s)$ -PC is NP-complete on bipartite graphs so we will reduce from that. If disconnected input graphs are allowed, then the result follows immediately by adding $k - 2$ isolated vertices, each of which must form a singleton part in any feasible partition. If one wants to keep the graph connected, we use a slightly different gadget.

Let $G = (V, E)$ be a bipartite graph. Construct a graph $G' = (V', E')$ as follows:

- Choose an arbitrary vertex $v \in V$.
- Attach to v a path P on $(s + 1)(k - 2)$ vertices, denoted by

$$P = w_1, w_2, \dots, w_{(s+1)(k-2)},$$

by adding the edge $(w_{(s+1)(k-2)}, v)$.

The resulting graph G' remains bipartite, and the construction is polynomial.

($\Rightarrow$) Suppose G admits a $(2, s)$ -partition X, Y such that $G[X], G[Y]$ are s -clubs. In G' we keep X and Y for the vertices of $V(G)$ and assign consecutive segments of length s (i.e. of $s + 1$ vertices) of the path P to $k - 2$ partitions. Each segment has diameter s so all induced subgraphs have diameter at most s . Thus G' admits a (k, s) -partition.

($\Leftarrow$) Conversely, suppose G' admits a (k, s) -partition $\mathcal{P} = \{V_1, \dots, V_k\}$. Consider the path P and its vertices w_i with position $(s + 1)i + 1$, with $0 \leq i \leq k - 3$. Any two of these vertices are at a distance greater than s and thus these $k - 2$ vertices must belong to different s -clubs. Moreover, no vertex of G can be in the same partition with any of these vertices. Indeed, the smallest distance from a vertex of G to these w_i is between v and $w_{(s+1)(k-3)+1}$ which is $s + 1$. Thus $k - 2$ sets of the partition are used for P , leaving 2 sets for the vertices of G . Restricting the partition $\mathcal{P}$ to V gives a $(2, s)$ -partition of G .

Hence G is a yes-instance of $(2, s)$ -PC if and only if G' is a yes-instance of (k, s) -PC. Therefore (k, s) -PC is NP-complete on bipartite graphs $\square$

4 Hardness of (t, s) -MAX-DCC problem

4.1 The $(t, 3)$ -MAX-DCC problem

In [12] it was proven that $(2, 3)$ -MAX-DCC can be solved in polynomial time and the complexity of $(3, 3)$ -MAX-DCC was posed as an open problem. In this section we prove that $(t, 3)$ -MAX-DCC is APX-hard for any fixed $t \geq 8$, even in bipartite graphs. We will use the following problem, which was proven to be NP-hard to approximate in [7].

Maximum 2 Bounded 3-Dimensional Matching problem (Max-2B3DM)

Instance: A set $M \subseteq X \times Y \times Z$ of ordered triples where X, Y and Z are disjoint sets and the number of occurrences in M of an element in X, Y or Z is bounded by constant 2.

Required: The largest matching $M' \subseteq M$, that is, the largest subset such that no two elements of M agree in any coordinate.

To prove that $(t, 3)$ -MAX-DCC for $t \geq 8$ is NP-hard, and to derive its inapproximability bound, we give an L -reduction from Max-2B3DM.

To make the paper self-contained, we recall the definition of an L -reduction and the approximation-preserving property that we will use. For more details on L -reductions, see [20].

Given an instance I of an optimization problem A , let $\text{opt}(I)$ denote the value of an optimal solution for I .

Definition 1. Let A and B be two optimization problems with objective functions c_1 and c_2 , respectively. We say that there is an L -reduction from A to B if there exist constants $\alpha, \beta > 0$ such that:

1. for each instance I_1 of A , one can compute in polynomial time an instance I_2 of B ;
2. if I_1 is an instance of A and I_2 is the corresponding instance of B , then

$$\text{opt}(I_2) \leq \alpha \text{opt}(I_1);$$

3. given any feasible solution s_2 for I_2 , one can compute in polynomial time a feasible solution s_1 for I_1 such that

$$|\text{opt}(I_1) - c_1(s_1)| \leq \beta |\text{opt}(I_2) - c_2(s_2)|.$$

For maximization problems, we say that an algorithm is an r -approximation, with $r \geq 1$, if it always returns a feasible solution of value at least $\text{opt}(I)/r$. We will use the following theorem, which is equivalent to Theorem 16.5 in [20].

Theorem 2. Suppose there is an L -reduction with parameters α and β from a maximization problem A to a maximization problem B . If there is an r -approximation algorithm for B , then there is a $\frac{1}{1-\alpha\beta(1-\frac{1}{r})}$ -approximation algorithm for A .

We begin by describing the reduction. Let $M = \{C_1, C_2, \dots, C_m\}$ be an instance of the Max-2B3DM problem, where each C_i is an ordered triple in $X \times Y \times Z$. Fix a constant $t \geq 8$. We construct a bipartite graph $G_{M,t,3} = (V_1, V_2, E)$, as follows:

$$\begin{aligned} V_1 &= X \cup Y \cup Z \cup \{a_i \mid i \in [m]\} \\ V_2 &= \{c_i \mid i \in [m]\} \cup \{h_{i,j} \mid i \in [m], j \in [t-5]\} \\ E &= \{(c_i, a_i) \mid i \in [m]\} \cup \{(a_i, h_{i,j}) \mid i \in [m], j \in [t-5]\} \bigcup_{i \in [m]} E_{C_i} \end{aligned}$$

where $E_{C_i} = \{(c_i, x), (c_i, y), (c_i, z)\}$ for each triple $C_i = (x, y, z)$ in M . As an example see Figure 2.

Claim 4. The graph $G_{M,t,3}$ can be constructed in polynomial time, is bipartite and has maximum degree $t-4$.

Proof: It is easy to see that $G_{M,t,3}$ can be constructed in polynomial time. Then by construction V_1 and V_2 form a bipartition as there are no edges within the sets V_1 and V_2 . Furthermore, for every $i \in [m]$, the vertex c_i has degree 4, a_i has degree $t-4$, and for every $j \in [t-5]$, the vertex $h_{i,j}$ has degree 1. Finally, every vertex $u \in X \cup Y \cup Z$ has degree 2, by the definition of the Max-2B3DM problem. $\square$

We need the following lemma.

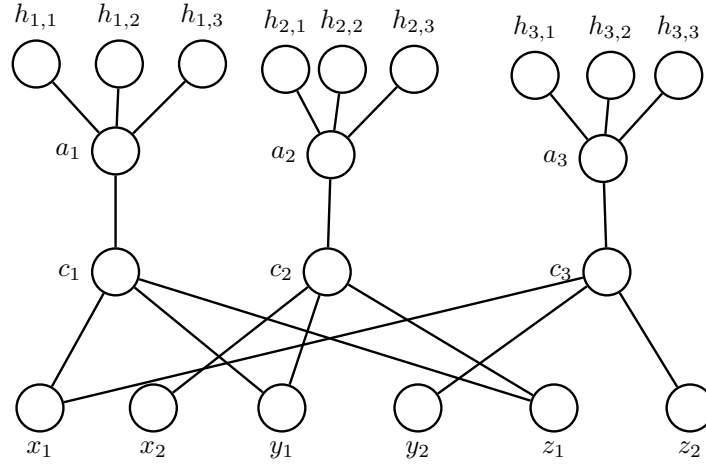


Figure 2: The graph $G_{M,8,3}$ obtained when the instance of Max-2B3DM problem is the set $M = \{(x_1, y_1, z_1), (x_2, y_1, z_1), (x_1, y_2, z_2)\}$.

Lemma 3. *In the graph $G_{M,t,3}$, all 3-clubs are of size at most t , with $t \geq 8$, and the only 3-clubs of size exactly t are of the form $N[c_i] \cup \{h_{i,j} \mid j \in [t-5]\}$ for $1 \leq i \leq m$.*

Proof: To prove this lemma, we will show that any 3-club in $G_{M,t,3}$ either contains exactly t vertices or has at most 7 vertices (with $7 < t$). Clearly the subgraphs of $G_{M,t,3}$ induced by $N[c_i] \cup \{h_{i,j} \mid j \in [t-5]\}$ for $1 \leq i \leq m$ are 3-clubs of size t . To simplify the analysis, we classify the vertices in $G_{M,t,3}$ into four distinct sets: the vertices in $W = X \cup Y \cup Z$, the vertices in $C = \{c_1, \dots, c_m\}$, the vertices in $A = \{a_1, \dots, a_m\}$, and the vertices in $H = \{h_{i,j} \mid i \in [m], j \in [t-5]\}$.

Let S be a 3-club in $G_{M,t,3}$. We will now show that if S contains at least two vertices from C then $|S| \leq 7 < t$. We make the following considerations regarding the composition of the vertices in S .

- *the number of vertices in $S \cap H$ is zero.* Note that such 3-club cannot include any vertices from H , as a vertex in H would be at a distance of at least 4 from one of the two vertices in C .
- *the number of vertices in $S \cap C$ is at most 3.* Suppose on the contrary there are 4 vertices from C , say c_1, c_2, c_3 and c_4 . For S to be a 3-club, there must be a vertex from $S \cap W$ adjacent to each pair of the c_i ($i \in [4]$) vertices. Let u be the vertex $S \cap W$ adjacent to c_1 and c_2 , and v be the vertex from $S \cap W$ adjacent c_3 and c_4 . The vertices u and v are at distance at least 4 in $G_{M,t,3}[S]$ because $N[u] \cap N[v] = \emptyset$ as all vertices from W have degree 2. This contradicts the assumption that S is a 3-club.
- *The number of vertices in $S \cap W$ is at most 4.* Suppose on the contrary that $|S \cap W| \geq 5$, and let $U \subseteq S \cap W$ be a set of five such vertices. At least one vertex $u \in U$ has degree 1 in $G_{M,t,3}[S]$. Indeed, otherwise the vertices of U would contribute at least 10 edges to vertices in $S \cap C$, while, as shown above, $|S \cap C| \leq 3$ and each vertex of C is adjacent to at most three vertices of W .

Let c be the unique neighbor of u in $S \cap C$. Since $G_{M,t,3}$ is bipartite and u and all vertices of W lie on the same side of the bipartition (since $W \subset V_1$), any path from u to another

vertex of W has even length. Therefore, for a vertex of $S \cap W$ to be within distance at most 3 from u , it must in fact be at distance at most 2 from u , and hence it must be adjacent to c . But by construction c is adjacent to at most three vertices of W in total. Thus at most three vertices of $S \cap W$ can be within distance at most 3 from u , contradicting $|S \cap W| \geq 5$. Hence $|S \cap W| \leq 4$.

Given the above, if $S \cap A = \emptyset$ then trivially the 3-club cannot have more than 7 vertices. Now assume that $S \cap A \neq \emptyset$. Notice that it must hold $|S \cap A| = 1$, since any two vertices in $S \cap A$ would be at distance at least 4 from each other. Let $a_i \in S \cap A$, then the vertices in W that a_i can reach with a path of length at most three are the ones adjacent to c_i . Thus $|S \cap W| \leq 3$ and the total number of vertices in the 3-club remains bounded by $1 + 3 + 3 = 7$. $\square$

Theorem 3. *Let t be a constant with $t \geq 8$. It is NP-hard to approximate the solution of $(t, 3)$ -MAX-DCC within a factor of $\frac{95}{94}$, even for bipartite graphs with a constant maximum degree of $t - 4$.*

Proof: Let $G_{M,t,3}$ be the graph obtained from an instance M of Max-2B3DM, for a given fixed t , with $t \geq 8$. By Claim 4 we have that $G_{M,t,3}$ can be constructed in polynomial time, has maximum degree of $t - 4$ and is bipartite. We prove now that from a matching of $k \geq 1$ triples in M , we can always obtain, in polynomial time, a disjoint cover with $(t, 3)$ -clubs that covers $t \cdot k$ vertices in $G_{M,t,3}$. Moreover, for every disjoint cover of $G_{M,t,3}$ with $(t, 3)$ -clubs that covers k vertices in $G_{M,t,3}$ it is possible to obtain in polynomial time a matching of $\frac{k}{t}$ triples in M (note that by Lemma 3 we have that k is a multiple of t).

Let $S = \{C_1, \dots, C_k\}$ be a subset of k triples from M that correspond to a matching. We define for each $i \in [k]$ the following set in $G_{M,t,3}$.

$$S'_i = \{x, y, z, c_i, a_i, h_{i,1} \dots h_{i,t-5}\}$$

Notice that S'_i is obviously a 3-club of size t and, since the triples in S form a matching, it follows that $S'_i \cap S'_j = \emptyset$ for any $i \neq j$. Thus, $S'_1, S'_2, \dots, S'_k$ form a disjoint cover with $(t, 3)$ -clubs that covers $t \cdot k$ vertices of $G_{M,t,3}$.

Let now $S_1, S_2, \dots, S_r$ be a disjoint cover with $(t, 3)$ -clubs that covers k vertices in $G_{M,t,3}$. By Lemma 3, we know that this cover is the union of $\frac{k}{t}$ disjoint 3-clubs, each containing t vertices of the form $N[c_i] \cup \{h_{i,j} \mid j \in [t-5]\}$ for some i , $1 \leq i \leq |M|$ where c_i corresponds to a triple (x, y, z) in M . To obtain the $\frac{k}{t}$ triples in M corresponding to a matching, it is enough to take the triples corresponding to the $\frac{k}{t}$ disjoint 3-clubs from S .

We therefore have an L -reduction with $\alpha = t$ and $\beta = \frac{1}{t}$ from Max-2B3DM to $(t, 3)$ -MAX-DCC. In [8], it was proven that for Max-2B3DM, achieving an approximation factor better than $\frac{95}{94}$ is NP-hard. Therefore we deduce the same result for $(t, 3)$ -MAX-DCC (see [20] for more information about L -reductions and in particular Theorem 2). $\square$

4.2 The $(t, 2)$ -MAX-DCC problem

In [12] it was proven that $MAX-DCC(3, 2)$ is APX-hard and the complexity of $MAX-DCC(2, 2)$ was posed as an open problem. In this section we prove that $MAX-DCC(t, 2)$ is APX-hard for any fixed $t \geq 5$, even in bipartite graphs and that $MAX-DCC(2, 2)$ can be solved in polynomial time.

In order to prove that $DCC(t, 2)$ problem with $t \geq 5$ is APX-hard, we present an L -reduction from Max-2B3DM. We begin by describing the reduction. Let $M = \{C_1, C_2 \dots C_m\}$ be an instance of Max-2B3DM where each C_i is an ordered triple in $X \times Y \times Z$.

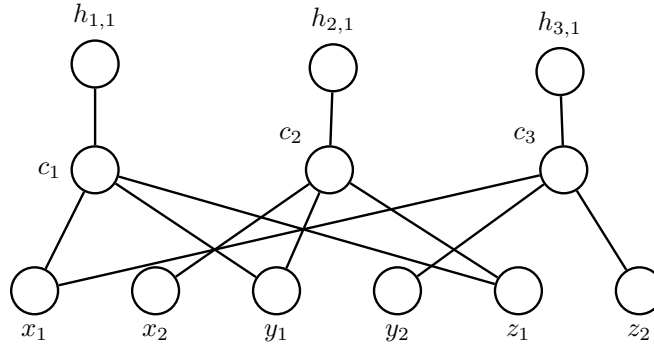


Figure 3: The graph $G_{M,5,2}$ obtained when the instance of Max-2B3DM problem is the set $M = \{(x_1, y_1, z_1), (x_2, y_1, z_1), (x_1, y_2, z_2)\}$.

Fix a constant t with $t \geq 5$. We construct a bipartite graph $G_{M,t,2} = (V_1, V_2, E)$, as follows:

$$\begin{aligned} V_1 &= X \cup Y \cup Z \cup \{h_{i,j} \mid 1 \leq i \leq |M|, j \in [t-4]\} \\ V_2 &= \{c_i \mid 1 \leq i \leq |M|\} \\ E &= \{(c_i, h_{i,j}) \mid i \in [m], j \in [t-4]\} \cup \bigcup_{i \in [m]} E_{C_i} \end{aligned}$$

where $E_{C_i} = \{(c_i, x), (c_i, y), (c_i, z)\}$ for each triple $C_i = (x, y, z)$ in M .

As an example see Figure 3.

Claim 5. *The graph $G_{M,t,2}$ can be constructed in polynomial time, is bipartite and has maximum degree $t-1$.*

Proof: It is easy to see that $G_{M,t,2}$ can be constructed in polynomial time. Then by construction V_1 and V_2 form a bipartition as there are no edges within the sets V_1 and V_2 . Furthermore, for every $i \in [m]$, the vertex c_i has degree $t-1$ and for every $j \in [t-4]$, the vertex $h_{i,j}$ has degree 1. Finally, every vertex $u \in X \cup Y \cup Z$ has degree 2, by the definition of the Max-2B3DM problem. $\square$

Lemma 4. *In the graph $G_{M,t,2}$, all the 2-clubs are of size at most t , with $t \geq 5$ and the only 2-clubs of size exactly t are of the form $N[c_i]$ for $1 \leq i \leq |M|$.*

Proof: To prove the claim, we will show that any 2-club in $G_{M,t,2}$ either contains exactly t vertices or has at most 4 vertices (with $4 < t$). Clearly the subgraphs of $G_{M,t,2}$ induced by $N[c_i]$ for $1 \leq i \leq m$ are 2-clubs of size t .

To simplify the analysis, we classify the vertices in $G_{M,t,2}$ into three distinct sets: the vertices in $W = X \cup Y \cup Z$, the vertices in $C = \{c_1 \dots c_m\}$ and the vertices in $H = \{h_{i,j} \mid i \leq [m], j \in [t-4]\}$. Let S be a 2-club in $G_{M,t,2}$. We will now show that if S contains at least two vertices from C then $|S| \leq 4 < t$.

Note that the number of vertices in $S \cap H$ is zero as a vertex in H of $G_{M,t,2}[S]$ would be at distance of at least 3 from one of the two vertices in C . Furthermore, in $G_{M,t,2}[S]$ each vertex in $S \cap C$ must be adjacent to each vertex in $S \cap W$. Consequently, S forms a complete bipartite graph having

$S \cap C$ and $S \cap W$ as independent sets. The degree of the vertices in $S \cap W$ is at most two thus S can have at most 4 vertices (two in $S \cap C$ and two in $S \cap W$). $\square$

Theorem 4. *Let t be a constant with, $t \geq 5$. It is NP-hard to approximate the solution of $(t, 2)$ -MAX-DCC to within $\frac{95}{94}$ even for bipartite graphs of degree at most $t - 1$.*

Proof: Let $G_{M,t,2}$ be the graph obtained from an instance M of Max-2B3DM, for a given fixed t , with $t \geq 5$. By Claim 5 we have that $G_{M,t,2}$ can be constructed in polynomial time, has a maximum degree of $t - 1$ and is bipartite.

We prove now that from a matching of $k \geq 1$ triples in M , we can always obtain in polynomial time, a disjoint cover with $(t, 2)$ -clubs that cover $t \cdot k$ vertices in $G_{M,t,2}$. Moreover, for every disjoint cover of $G_{M,t,2}$ it is possible to obtain in polynomial time a matching of $\frac{k}{t}$ triples in M (note that by Lemma 4 we have that k is a multiple of t).

Let $S = \{C_1, \dots, C_k\}$ be a subset of k triples from M that correspond to a matching. We define for each $i \in [k]$ the following set in $G_{M,t,2}$

$$S'_i = \{x, y, z, c_i, h_{i,1} \dots h_{i,t-4}\}$$

Note that S'_i is obviously a 2-club of size t and, since the triples in S form a matching, it follows that $S'_i \cap S'_j = \emptyset$ for any $i \neq j$. Thus $S'_1, S'_2, \dots, S'_k$ form a disjoint cover with $(t, 2)$ -clubs that cover $t \cdot k$ vertices of $G_{M,t,2}$.

Let now $S_1, S_2, \dots, S_r$ be a disjoint cover with $(t, 2)$ -clubs that cover k vertices in $G_{M,t,2}$. By Lemma 4, we know that this cover is the union of $\frac{k}{t}$ disjoint 2-clubs, each containing t vertices of the form $N[c_i]$, for some i , $1 \leq i \leq |M|$, where c_i corresponds to a triple (x, y, z) in M . To obtain the $\frac{k}{t}$ triples in M corresponding to a matching, is enough to take the triples corresponding to the $\frac{k}{t}$. To obtain the $\frac{k}{t}$ triples it is enough to take the triples corresponding to the $\frac{k}{t}$ disjoint 2-clubs from S .

We therefore have an L -reduction with $\alpha = t$ and $\beta = \frac{1}{t}$ from Max-2B3DM to $(t, 2)$ -MAX-DCC. In [8], it was proven that for Max-2B3DM, it is NP-hard to achieve an approximation factor of $\frac{95}{94}$. Thus, from Theorem 2, we deduce the same result for $(t, 2)$ -MAX-DCC. $\square$

In [12], it is shown that for any $s \geq 3$ $(2, s)$ -MAX-DCC is solvable in linear time. The authors leave the problem of determining the complexity of $(2, s)$ -MAX-DCC as an open problem. We prove the following ideas similar to the one in [12].

Theorem 5. *$(2, 2)$ -MAX-DCC can be solved in linear time.*

Proof: Let G be the input graph for $(2, 2)$ -MAX-DCC. We will prove the claim by describing a linear algorithm that produces a cover of disjoint $(2, 2)$ -clubs that covers the maximum number of vertices of G . Consider the graph G' , obtained by removing the isolated vertices from G . Note that isolated vertices cannot be covered by a $(2, 2)$ -club, so they can be disregarded. We can assume that G' is connected as otherwise, we apply the following procedure to each connected component.

From G' , construct a rooted spanning tree $T_{G'}$. This can be done in linear time see [9]. Notice that $T_{G'}$ has a height of at least one as G' has no isolated vertices. Thus, let x be a vertex of $T_{G'}$ at maximum level and let y be the parent of x . Consider the subtree T_y rooted at y . Clearly, T_y contains at least 2 vertices. Moreover, as x has maximum level then T_y is of height 1 and all vertices in this subtree are at a distance at most 2 from each other. Hence, the set S_1 of vertices in T_y form a $(2, 2)$ -club. We add S_1 to the solution and remove T_y from $T_{G'}$. The remaining tree

is still connected, so we can iterate the process until either we reach an empty tree or we reach a tree consisting of a single vertex. In the latter case, let s be this vertex and let S_i be the $(2, 2)$ -club added to the solution by the last iteration. Notice that s is at distance 1 from the root of this subtree and at distance 2 from any vertex of S_i . Therefore, we can add s to S_i and still have a $(2, 2)$ -club. In the end the solution obtained is a cover of *all* the vertices of G' with disjoint $(2, 2)$ -clubs and thus is a solution for $(2, s)$ -MAX-DCC on G . $\square$

5 Open problems

Several problems remain open in the context of bipartite graphs. Specifically, the cases of $(t, 4)$ -PC and $(t, 6)$ -PC for any $t \geq 2$ remain unsolved.

Regarding the second problem, for $s = 3$, the complexity of the $(t, 3)$ -MAX-DCC problem is still open for $3 \leq t \leq 7$. Similarly, for $s = 2$, the case of $(4, 2)$ -MAX-DCC remains unsolved.

Acknowledgements

The authors thank both the reviewers for carefully reading of the paper and in particular Reviewer 2 for identifying a flaw in the previous proof of the complexity of (t, s) -PC, which gave us the opportunity to reconsider the argument and strengthen the result.

This research was supported in part by MUR PRIN Project EXPAND, grant number 2022TS4Y3N.

References

- [1] N. Abbas and L. Stewart. Clustering bipartite and chordal graphs: Complexity, sequential and parallel algorithms. *Discrete Applied Mathematics*, 91(1–3):1–23, Jan. 1999. doi:10.1016/S0166-218X(98)00094-8.
- [2] A. S. Asratian, T. M. J. Denley, and R. Häggkvist. *Bipartite Graphs and their Applications*. Cambridge University Press, Cambridge, UK, July 1998. URL: <http://dx.doi.org/10.1017/CB09780511984068>, doi:10.1017/cbo9780511984068.
- [3] M. J. Barber. Modularity and community detection in bipartite networks. *Phys. Rev. E*, 76:066102, Dec 2007. URL: <https://link.aps.org/doi/10.1103/PhysRevE.76.066102>, doi:10.1103/PhysRevE.76.066102.
- [4] C. Bazgan, P. Heggernes, A. Nichterlein, and T. Pontoizeau. On the hardness of problems around s -clubs on split graphs. *Discrete Applied Mathematics*, 364:247–254, 2025. URL: <https://www.sciencedirect.com/science/article/pii/S0166218X25000290>, doi:10.1016/j.dam.2025.01.023.
- [5] G. Calderer and M. L. Kuijjer. Community detection in large-scale bipartite biological networks. *Frontiers in Genetics*, 12, Apr. 2021. URL: <http://dx.doi.org/10.3389/fgene.2021.649440>, doi:10.3389/fgene.2021.649440.
- [6] J.-M. Chang, J.-S. Yang, and S.-L. Peng. On the complexity of graph clustering with bounded diameter. In *2014 International Computer Science and Engineering Conference (ICSEC)*, volume 10, page 18–22, Khon Kaen, Thailand, July 2014. IEEE. doi:10.1109/icsec.2014.6978122.

- [7] M. Chlebík and J. Chlebíková. Approximation hardness for small occurrence instances of np-hard problems. In *Italian Conference on Algorithms and Complexity*, pages 152–164. Springer, 2003.
- [8] M. Chlebík and J. Chlebíková. Complexity of approximating bounded variants of optimization problems. *Theoretical Computer Science*, 354(3):320–338, 2006. Foundations of Computation Theory (FCT 2003). doi:10.1016/j.tcs.2005.11.029.
- [9] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. The MIT Press, Cambridge, MA, 2nd edition, 2001.
- [10] J. S. Deogun, D. Kratsch, and G. Steiner. An approximation algorithm for clustering graphs with dominating diametral path. *Information Processing Letters*, 61(3):121–127, Feb. 1997. doi:10.1016/S0020-0190(97)81663-8.
- [11] R. Dondi and M. Lafond. On the tractability of covering a graph with 2-clubs. *Algorithmica*, 85(4):992–1028, 2023. doi:10.1007/S00453-022-01062-3.
- [12] R. Dondi, G. Mauri, and I. Zoppis. On the tractability of finding disjoint clubs in a network. *Theoretical Computer Science*, 777:243–251, July 2019. doi:10.1016/j.tcs.2019.03.045.
- [13] H. Fleischner, E. Mujuni, D. Paulusma, and S. Szeider. Covering graphs with few complete bipartite subgraphs. *Theoretical Computer Science*, 410(21–23):2045–2053, May 2009. doi:10.1016/j.tcs.2008.12.059.
- [14] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, San Francisco, CA, 1979.
- [15] C. Komusiewicz. Multivariate algorithmics for finding cohesive subnetworks. *Algorithms*, 9(1):21, Mar. 2016. URL: <http://dx.doi.org/10.3390/a9010021>, doi:10.3390/a9010021.
- [16] S. Laan, M. Marx, and R. J. Mokken. Close communities in social networks: boroughs and 2-clubs. *Social Network Analysis and Mining*, 6(1), Apr. 2016. doi:10.1007/s13278-016-0326-0.
- [17] R. J. Mokken. Cliques, clubs and clans. *Quality and Quantity*, 13(2):161–173, Apr. 1979. URL: <http://dx.doi.org/10.1007/BF00139635>, doi:10.1007/bf00139635.
- [18] R. J. Mokken, E. M. Heemskerk, and S. Laan. Close communication and 2-clubs in corporate networks: Europe 2010. *Social Network Analysis and Mining*, 6(1), June 2016. URL: <http://dx.doi.org/10.1007/s13278-016-0345-x>, doi:10.1007/s13278-016-0345-x.
- [19] A. Monti and B. Sinaireri. Disjoint covering of bipartite graphs with s-clubs. In R. Kráľovič and V. Kůrková, editors, *SOFSEM 2025: Theory and Practice of Computer Science*, pages 198–210, Cham, 2025. Springer Nature Switzerland. doi:10.1007/978-3-031-82697-9_15.
- [20] D. P. Williamson and D. B. Shmoys. *The design of approximation algorithms*. Cambridge university press, Cambridge, UK, 2011.
- [21] I. Zoppis, R. Dondi, E. Santoro, G. Castelnovo, F. Sicurello, and G. Mauri. Optimizing social interaction. In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies*, page 651–657, Funchal, Madeira, Portugal, 2018. SCITEPRESS - Science and Technology Publications. doi:10.5220/0006730606510657.

- [22] P. Zou, H. Li, W. Wang, C. Xin, and B. Zhu. Finding disjoint dense clubs in a social network. *Theoretical Computer Science*, 734:15–23, 2018. Selected papers from the The Tenth International Frontiers of Algorithmics Workshop (FAW 2016). URL: <https://www.sciencedirect.com/science/article/pii/S0304397517307466>, doi:10.1016/j.tcs.2017.10.018.