# On fault-tolerant low-diameter clusters in graphs

Yajun Lu
Department of Management & Marketing, Jacksonville State University, ylu@jsu.edu

Hosseinali Salemi
Department of Industrial & Manufacturing Systems Engineering, Iowa State University, hsalemi@iastate.edu

Balabhaskar Balasundaram, Austin Buchanan
School of Industrial Engineering & Management, Oklahoma State University, baski@okstate.edu, buchanan@okstate.edu

August 18, 2021

Cliques and their generalizations are frequently used to model "tightly knit" clusters in graphs, and identifying such clusters is a popular technique used in graph-based data mining. One such model is the $s$-club, which is a vertex subset that induces a subgraph of diameter at most $s$. This model has found use in a variety of fields because low-diameter clusters have practical significance in many applications. As this property is not hereditary on vertex-induced subgraphs, the diameter of a subgraph could increase upon the removal of some vertices and the subgraph could even become disconnected. For example, star graphs have diameter two but can be disconnected by removing the central vertex. The pursuit of a fault-tolerant extension of the $s$-club model has spawned two variants that we study in this article: hereditary $s$-clubs and robust $s$-clubs. We analyze the complexity of the verification and optimization problems associated with these variants. Then, we propose cut-like integer programming formulations for both variants whenever possible and investigate the separation complexity of the cut-like constraints. We demonstrate through our extensive computational experiments that the algorithmic ideas we introduce enable us to solve the problems to optimality on benchmark instances with several thousand vertices. This work lays the foundations for effective mathematical programming approaches for finding fault-tolerant $s$-clubs in large-scale networks.

*Key words*: integer programming; hereditary $s$-clubs; robust $s$-clubs; branch-and-cut; social network analysis

---

## 1. Introduction

Modeling data entities and their pairwise relationships as a graph is a popular approach to visualizing and mining information from datasets in a variety of fields (Cook and Holder 2006). An established technique in this setting involves the detection of clusters—either by finding those of the largest cardinality or weight, finding those that cover or partition the graph, or enumerating all inclusionwise maximal clusters.

Clique, a subset of pairwise adjacent vertices, is often viewed as an idealized representation of a cluster. However, in the presence of errors in the data upon which the graph is based, the clique requirement may be too restrictive, resulting in small clusters that miss key members. Graph-theoretic clique generalizations based on the principle of relaxing elementary structural properties of a clique have been proposed in diverse fields to describe clusters of interest (Pattillo et al. 2013). Such clique relaxations are less sensitive to edges missed due to erroneous or incomplete data underlying the graph representation. Next we introduce the notations followed by the clique relations of interest.

We consider simple, unweighted graphs in this article. We denote by $G = (V, E)$ an $n$-vertex graph with vertex set $V = \{1, 2, \ldots, n\}$ and edge set $E \subseteq \{e \subseteq V \mid |e| = 2\}$ containing $m$ edges. We denote by $G[S]$ the subgraph induced by a subset of vertices $S$, and by $G - S$ we denote the graph obtained by deleting the vertices in $S$ and incident edges from $G$. The set of neighbors of a vertex $u$ in graph $G$ is denoted by $N_G(u) \coloneqq \{v \in V \mid \{u, v\} \in E\}$. The *closed* neighborhood of vertex $u$ includes itself, and is denoted by $N_G[u] = N_G(u) \cup \{u\}$. The *distance* between a pair of vertices $u$ and $v$ in $G$, denoted by $\text{dist}_G(u, v)$, is the minimum number of edges on a path from $u$ to $v$ in $G$. The *diameter* of $G$ is the maximum distance between any pair of vertices in $G$ and is denoted by $\text{diam}(G)$. Given a positive integer $s$, the set of distance-$s$ neighbors of $u$ is denoted by $N_G^s(u)$ and is defined as $N_G^s(u) \coloneqq \{v \in V \mid 1 \leq \text{dist}_G(u, v) \leq s\}$. The *closed* distance-$s$ neighborhood of $u$ is denoted by $N_G^s[u] = N_G^s(u) \cup \{u\}$. We use the short form $uv$ for an edge $\{u, v\}$ and drop the subscript $G$ when the graph under consideration is understood without any ambiguity. We recall two distance-based clique relaxations from the literature: $s$-clique and $s$-club.

Definition 1 (Luce (1950)). Given a positive integer $s$, a subset of vertices $S$ is called an *$s$-clique* if $\text{dist}_G(u, v) \leq s$ for every pair of vertices $u, v \in S$.

Definition 2 (Mokken (1979)). Given a positive integer $s$, a subset of vertices $S$ is called an *$s$-club* if $\text{diam}(G[S]) \leq s$.

Clearly, the special case $s = 1$ in both definitions corresponds to the clique. The fundamental difference between an $s$-clique and an $s$-club is that the distance bound is applicable to the original graph in the former, and to the induced subgraph in the latter. Hence, every $s$-club is an $s$-clique, but not vice versa. Figure 1 illustrates this difference (Alba 1973).

Arguably, the $s$-club model is more cohesive because it guarantees that the length-bounded paths between vertices are completely contained within the induced subgraph. Originally introduced to model cohesive subgroups in social networks (Mokken 1979), $s$-clubs can be used to model low-diameter clusters for small values of $s$. In particular, the

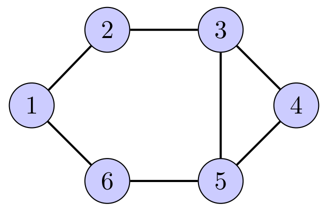

**Figure 1** **The set $S = \{2,3,4,5,6\}$ is a 2-clique that is not a 2-club as the distance between vertices 2 and 6 in the graph induced by $S$ is more than 2. Note that $\text{dist}_G(2,6) = 2$, which uses vertex 1 that is not in $S$.**

2-club represents clusters in which every pair of its vertices are either adjacent, or have a common neighbor inside the cluster. Hence, 2-clubs formalize the notion of a friend-of-a-friend social group in which members may be directly acquainted or related through a mutual acquaintance in the group.

### 1.1. Fault-Tolerant Clubs

Although $s$-clubs can ensure low pairwise distances inside the cluster, they may not be *fault-tolerant* in the sense that deleting a single vertex could increase the distances or even disconnect the graph. For example, in the graph in Figure 1 the set $S_1 = \{2,3,4,5\}$ is a 2-club, but $S_1 \setminus \{3\}$ induces a disconnected subgraph. Yezerska et al. (2017) refer to this as a "fragile" 2-club, and focus instead on finding 2-clubs that induce biconnected subgraphs, e.g., the 2-club $S_2 = \{1,2,3,5,6\}$ in Figure 1. Nonetheless, $S_2$ is not fault-tolerant as the diameter of $G[S_2]$ increases if any single vertex in $S_2$ is deleted.

It is important to note here that the $s$-club property for $s \geq 2$ is not hereditary in the sense of Lewis and Yannakakis (1980). One cannot guarantee that the diameter bound will be preserved under vertex deletion in general, even if the induced subgraph remains connected. This has led researchers to devise notions of "strong attack tolerance", wherein the graph property in question (e.g., diameter) persists under a small number of vertex and/or edge failures. These observations motivated Veremyev and Boginski (2012b) to introduce the following fault-tolerant variant of $s$-clubs.

Definition 3 (Veremyev and Boginski (2012a)). Given a graph $G$ and positive integers $r$ and $s$, a subset of vertices $S$ is called an *$r$-robust $s$-club* if between every pair of distinct vertices in $S$ there are at least $r$ internally vertex-disjoint paths of length at most $s$ in $G[S]$.

By Definition 3, every $r$-robust $s$-club $S$ must contain at least $r+1$ vertices except when the definition is trivially satisfied, i.e., $|S| \leq 1$. Furthermore, the only $r$-robust $s$-clubs that contain exactly $r+1$ vertices are cliques of that size. As long as an $r$-robust $s$-club contains two vertices that are not adjacent, it must contain at least $r+2$ vertices. While this definition bestows an $s$-club with fault tolerance by ensuring redundant short paths, it is not the only approach to achieve that effect. Consider the following variant introduced by Pattillo et al. (2013).

Definition 4 (Pattillo et al. (2013)). Given a graph $G$ and positive integers $t$ and $s$, a subset of vertices $S$ is called a *$t$-hereditary $s$-club* if $S \setminus T$ is an $s$-club for every deletion set $T \subseteq S$ containing fewer than $t$ vertices.

If we have $t = r = 1$, Definitions 3 and 4 coincide with Definition 2 for every positive integer $s$, i.e., every $s$-club is both 1-hereditary and 1-robust. We note here that Definition 4 deviates slightly from the original definition of Pattillo et al. (2013), which allowed deletion sets up to (and equal to) size $t$. Our redefinition is more convenient when working with both models simultaneously. Lemma 1 that follows states (without proof) a general relationship between $r$-robust $s$-clubs and $t$-hereditary $s$-clubs that can be easily verified.

Lemma 1. *Every $r$-robust $s$-club is also an $r$-hereditary $s$-club.*

The converse of Lemma 1 is not true. For example, a 4-cycle is a 2-hereditary 2-club that is not a 2-robust 2-club. The distance between adjacent vertices in a $t$-hereditary $s$-club remains one after any other vertex is deleted and hence, adjacent pairs of vertices are not subjected to any additional requirements. By contrast, an adjacent pair of vertices in an $r$-robust $s$-club still need to be connected by at least $r-1$ additional vertex-disjoint paths of length $s$ or less. This is one key difference that can impact the type of fault-tolerant cluster detected in practice. For example, the largest 3-robust 3-club found in the `dolphins` graph from the DIMACS Clustering Challenge benchmarks (Bader et al. 2013) contains 14 vertices, while the largest 3-hereditary 3-club found contains 17 vertices.

### 1.2. Prior Work and Our Contributions

The focus of this article is on combinatorial optimization problems seeking a maximum cardinality $s$-club that also satisfies an additional property of robustness or heredity, following Definitions 3 and 4. We refer to these as the maximum $r$-robust $s$-club problem (MRCP) and the maximum $t$-hereditary $s$-club problem (MHCP). We briefly review the limited literature currently available related to these problems before outlining our contributions.

Komusiewicz et al. (2019) showed that the decision version of the maximum $t$-hereditary and $r$-robust 2-club problems are NP-complete for every pair of fixed integers $t, r \geq 2$. The hardness of the problem for *arbitrary* $s$ follows immediately from their results because any algorithm for either problem where $s$ is arbitrary (meaning $s$ is specified in the input) must also solve the problems for $s = 2$. By contrast, the complexity of these problems where $s$ is a *fixed* constant in the problem definition, e.g., $t$-hereditary 3-club, is not explicitly addressed by this result. We answer these open complexity questions in this article.

Veremyev and Boginski (2012a) proposed a compact integer programming (IP) formulation for a relaxation of the MRCP in which the $r$ paths of length at most $s$ are only required to be distinct, and not necessarily vertex-disjoint. However, when $s = 2$, the $r$ distinct paths must also be vertex-disjoint, and therefore their formulation correctly models this special case. Almeida and Carvalho (2014) provided a compact IP formulation for $r$-robust 3-clubs. No general formulations are currently available for the MRCP when $s \geq 4$.

Salemi and Buchanan (2020) introduced a cut-like formulation for the maximum $s$-club problem and explained how it can be modified to correctly formulate the MHCP. This formulation also generalizes the IP formulation of the maximum $t$-hereditary 2-club problem proposed by Komusiewicz et al. (2019). We prove the correctness of this formulation along similar lines as suggested by Salemi and Buchanan (2020) in this article.

Our contributions are summarized as follows. In Section 2, we establish the NP-completeness of the decision counterparts of the MRCP and the MHCP for all *fixed* integers $s, r, t \geq 2$. We also establish the conditions on the parameters under which the problem of verifying whether a subset of vertices is an $r$-robust $s$-club is NP-complete and verifying if it is a $t$-hereditary $s$-club is coNP-complete. In Section 3, we present cut-like formulations based on length-bounded vertex separators for the MRCP and MHCP. Our cut-like formulations are compared with existing formulations in the literature wherever possible. In light of the worst-case exponential size of the cut-like formulations, we establish whether or not $t$-hereditary $s$-clubs and $r$-robust $s$-clubs admit polynomial-size formulations depending on the values of $t$ (or $r$) and $s$ in Section 4. In order to speed up solving the MRCP and the MHCP using our cut-like formulations in a branch-and-cut algorithm, we introduce several preprocessing and graph decomposition techniques in Section 5. We report our computational experience solving the MRCP and the MHCP for $s \in \{2, 3, 4\}$ in Section 6. The results demonstrate the effectiveness of our proposed approaches. In addition, these are the first reported numerical results for the MRCP and MHCP when $s \in \{3, 4\}$. Our codes are shared publicly on GitHub at `https://github.com/yajun668/FaultTolerantClubs`.

We conclude the paper in Section 7 with a summary and remarks for future research on these and related problems.

## 2. Problem Complexity

In this section, we establish the intractability of the decision and verification versions of the MRCP and MHCP, formally stated next.

**Problem**: $s$-Club/$t$-Hereditary $s$-Club/$r$-Robust $s$-Club (positive integers $s,t,r$)
**Question**: Given a graph $G$ and positive integer $c$, does $G$ contain an $s$-club/$t$-hereditary $s$-club/$r$-robust $s$-club of size at least $c$?

Bourjolly et al. (2002) established that the $s$-Club problem is NP-complete for every fixed integer $s \geq 2$, and it remains NP-complete even when restricted to graphs of diameter $s+1$ (Balasundaram et al. 2005). Testing inclusionwise maximality of $s$-clubs is also coNP-complete (Pajouh and Balasundaram 2012). The $s$-Club problem remains NP-hard on 4-chordal graphs for every positive integer $s$ (Golovach et al. 2014). Komusiewicz et al. (2019) showed that the $t$-Hereditary 2-Club and $r$-Robust 2-Club problems are NP-complete for every fixed integer $t \geq 2$ and $r \geq 2$, respectively. Table 1 summarizes the known complexity results related to the $s$-Club problem and its fault-tolerant extensions.

We prove the NP-completeness of the decision counterparts of MHCP and MRCP for every fixed integer $s,r,t \geq 2$ on general graphs, and obtain complexity results on some special graph classes as corollaries. We then show that even the verification problems (checking whether a given subset of vertices is an $r$-robust or $t$-hereditary $s$-club) are intractable in certain circumstances where the parameters involved are treated as part of the input. The complexity results pertaining to the verification problems are especially important due to their implications for algorithm development and for the existence of compact IP formulations.

### 2.1. NP-Hardness of Optimization

Theorems 1 and 2 that follow establish that $t$-Hereditary $s$-Club and $r$-Robust $s$-Club are NP-complete using reductions from $s$-Club. Note that the problems are trivially NP-hard when parameters $t,r,s$ are not fixed in the problem definition, as they all include Clique as a special case where $s=1$. The proofs of the results presented in this section are included in Appendix A.

**Table 1 Main complexity results related to $s$-clubs and other variants**

| Problem | Key results |
|---|---|
| $s$-CLUB | NP-complete for every positive integer $s$ (Bourjolly et al. 2002), even when restricted to graphs of diameter $s+1$ (Balasundaram et al. 2005). |
| | NP-complete on 4-chordal graphs for every positive integer $s$ (Golovach et al. 2014), on bipartite graphs for every fixed $s \geq 3$, and on chordal graphs for every *even* fixed integer $s \geq 2$ (Asahiro et al. 2010). |
| | Testing maximality by inclusion is coNP-complete for every fixed integer $s \geq 2$ (Pajouh and Balasundaram 2012). |
| | NP-hard to approximate within a factor of $n^{1/2-\epsilon}$ in general graphs for any $\epsilon > 0$ and a fixed $s \geq 2$ (Asahiro et al. 2018). |
| | Polynomial-time solvable on the following graph classes: trees, interval graphs, and graphs with bounded treewidth or cliquewidth for every fixed $s \geq 1$ (Schäfer 2009); chordal bipartite, strongly chordal and distance hereditary graphs for every fixed $s \geq 1$; weakly chordal graphs for every fixed *odd* $s$ (Golovach et al. 2014). |
| | $O(n^{1/2})$-approximable for fixed $s \geq 2$ (Asahiro et al. 2018). |
| | Fixed-parameter tractable when parameterized by solution size (Schäfer et al. 2012). |
| 2-CLUB | NP-hard on the following graph classes: split graphs (Asahiro et al. 2010); graphs with clique cover number three and diameter three; graphs with domination number two and diameter three (Hartung et al. 2015). |
| | Polynomial-time solvable on bipartite graphs in $O(n^5)$ (Schäfer 2009). |
| | Approximable by a factor of $n^{1/3}$ on split graphs (Asahiro et al. 2010). |
| $r$-ROBUST and $t$-HEREDITARY 2-CLUB | NP-complete for every pair of fixed positive integers $r, t \geq 1$ (Komusiewicz et al. 2019). Fixed-parameter tractable when parameterized by $\ell = \lvert V \rvert - k$ where $k$ is solution size; does not admit a $(2-\epsilon)^{\ell} n^{O(1)}$-time algorithm for any $\epsilon > 0$ if the Strong Exponential Time Hypothesis is true (Komusiewicz et al. 2019). |

THEOREM 1. $t$-HEREDITARY $s$-CLUB *is NP-complete for every pair of fixed integers* $s \geq 2$ *and* $t \geq 2$*, even on graphs with domination number one.*

THEOREM 2. $r$-ROBUST $s$-CLUB *is NP-complete for every pair of fixed integers* $s \geq 2$ *and* $r \geq 2$*, even on graphs with domination number one.*

Chordal graphs, which contain no chordless cycles of length four or more, are a subclass of perfect graphs with interesting and desirable properties for clique detection (Rose et al. 1976). For every nonnegative integer $k$, a $k$-chordal graph contains no chordless cycles of length greater than $k$. So, 3-chordal graphs are precisely the classical chordal graphs. Golovach et al. (2014) proved that $s$-CLUB is NP-complete on 4-chordal graphs for every fixed integer $s \geq 1$, and it remains NP-complete on the subclass of chordal graphs for every fixed *even* integer $s \geq 2$ (Asahiro et al. 2010). Golovach et al. (2014) also proved that 2-CLUB is NP-hard on graphs with clique cover number three, i.e., on graphs whose vertex sets can be covered using three cliques. These results allow us to show the NP-hardness of MRCP and MHCP even on these restricted graph classes.

COROLLARY 1. *For every pair of fixed integers $r, t \geq 2$, $t$-*HEREDITARY *$s$-*CLUB *and $r$-*ROBUST *$s$-*CLUB *remain NP-complete,*

1. *on 4-chordal graphs for every fixed integer $s \geq 1$, and*
2. *on chordal graphs for every fixed even integer $s \geq 2$.*

COROLLARY 2. *For every pair of fixed integers $r, t \geq 2$, $t$-*HEREDITARY 2-CLUB *and $r$-*ROBUST 2-CLUB *remain NP-complete on graphs with clique cover number three.*

### 2.2. Hardness of Verification

We begin this section by recalling the definition of a length-bounded vertex separator (Lovász et al. 1978, Salemi and Buchanan 2020).

DEFINITION 5. Given a pair of nonadjacent vertices $u$ and $v$ in graph $G = (V, E)$, a subset of vertices $C \subseteq V \backslash \{u, v\}$ is called a length-$s$ $u,v$-separator if $\text{dist}_{G-C}(u, v) > s$.

Proposition 1 below, established by Lovász et al. (1978), relates the size of length-bounded vertex separators to the number of length-bounded vertex disjoint paths. More importantly, this proposition offers a length-bounded counterpart of Menger's Theorem for lengths in the set $\{2, 3, 4\}$ (Lawler 1976, Menger 1927). For a pair of vertices $u, v$ in $G$, let $\rho_s(G; u, v)$ denote the maximum number of internally vertex-disjoint $u,v$-paths of length at most $s$ in $G$, and for nonadjacent vertices $u$ and $v$, let $\kappa_s(G; u, v)$ denote the minimum cardinality of a length-$s$ $u,v$-separator.

PROPOSITION 1 **(Lovász et al. (1978))**. *Consider a graph $G$ with $n$ vertices containing nonadjacent vertices $u, v$, and a positive integer $s$. Then,*

$$\rho_s(G; u, v) \leq \kappa_s(G; u, v) \leq \left\lfloor \frac{n}{2} \right\rfloor \rho_s(G; u, v).$$

*Furthermore, for $s \in \{2, 3, 4\}$,*

$$\rho_s(G; u, v) = \kappa_s(G; u, v).$$

Figure 2 provides an example illustrating that $\rho_s(G; u, v)$ could be strictly smaller than $\kappa_s(G; u, v)$ when $s \geq 5$. Note that $\text{dist}_G(10, 11) = 4$. Although several paths of length five and one of length four exist between vertices 10 and 11, no more than one can be included in a vertex disjoint collection of paths of length at most 5. After deleting any single vertex from the set {2,5,8} we can still find a length-5 path in this graph between vertices 10 and 11. Hence, $\rho_5(G; 10, 11) = 1$, but $\kappa_5(G; 10, 11) = 2$.

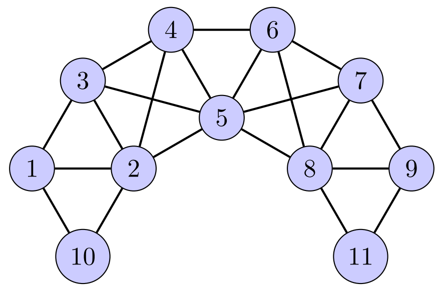


**Figure 2** **In this graph** $G$**,** $\rho_5(G;10,11)=1$**, but** $\kappa_5(G;10,11)=2$**, while** $\rho_4(G;10,11)=1=\kappa_4(G;10,11)$ **as** $\{2\}$ **is a length-4 separator for vertices 10 and 11.**

When $s=2$, verifying if a vertex subset $S$ is an $r$-robust 2-club amounts to checking if every adjacent pair of vertices in $S$ have at least $r-1$ common neighbors in $G[S]$ and every nonadjacent pair have at least $r$ common neighbors in $G[S]$. For distinct vertices $u$ and $v$ that are adjacent in $G[S]$, clearly $\rho_2(G[S];u,v)=|N_G(u)\cap N_G(v)\cap S|+1$. If they are not adjacent, then $\rho_2(G[S];u,v)=|N_G(u)\cap N_G(v)\cap S|$. Moreover, for nonadjacent vertices $u$ and $v$, the set of common neighbors $N_G(u)\cap N_G(v)\cap S$ is the unique minimum cardinality length-2 $u,v$-separator in $G[S]$. Hence, we can verify if $S$ is a $t$-hereditary 2-club by checking if every nonadjacent pair of vertices have at least $t$ common neighbors in $G[S]$.

For $s\in\{3,4\}$, we can compute $\rho_s(G;u,v)$ in $O(|E|\sqrt{|V|})$ time (Lovász et al. 1978, Itai et al. 1982), which along with Proposition 1 can be used to verify if $S$ is an $r$-robust $s$-club or a $t$-hereditary $s$-club for arbitrary $r$ and $t$ in polynomial time. That leaves open the tractability of the verification problem for $s\geq 5$, in which case the length-bounded counterpart of Menger's Theorem does not hold. The discussion that follows addresses these questions depending on whether the parameters $s$, $r$, and $t$ are fixed or arbitrary. The new complexity results on optimization and verification are summarized in Table 2.

**Table 2** **Summary of complexity results established in Section 2**

| **Problem** | **Parameter(s) fixed in the problem** | **Parameter specified in the input** | **Complexity** |
|---|---|---|---|
| $r$-Robust $s$-Club | $s\geq 2$ and $r\geq 2$ | | NP-complete |
| $t$-Hereditary $s$-Club | $s\geq 2$ and $t\geq 2$ | | NP-complete |
| | $s\geq 5$ | $r$ | NP-complete |
| Is $r$-Robust $s$-Club | $s\leq 4$ | $r$ | polynomial-time |
| | $r\geq 2$ | $s$ | NP-complete |
| | $s\geq 5$ | $t$ | coNP-complete |
| Is $t$-Hereditary $s$-Club | $s\leq 4$ | $t$ | polynomial-time |
| | $t\geq 2$ | $s$ | polynomial-time |

**Problem**: Is $r$-Robust $s$-Club (positive integers $s, r$)
**Question**: Given a graph $G = (V, E)$ and a subset $S \subseteq V$, is $S$ an $r$-robust $s$-club in $G$?

Theorem 3. Is $r$-Robust $s$-Club *is NP-complete for every fixed integer* $s \geq 5$ *and arbitrary positive integer* $r$.

Theorem 4. Is $r$-Robust $s$-Club *is NP-complete for every fixed integer* $r \geq 2$ *and arbitrary positive integer* $s$.

Theorems 3 and 4 establish that the verification problem is NP-complete if one of the two parameters $s \geq 5$ and $r \geq 2$ is arbitrary, while the other is fixed. Verification of $t$-hereditary $s$-clubs is also difficult as stated in Theorem 5 that follows, when $t$ is arbitrary; but it is easy when $t$ is a fixed constant. The proofs of these results are included in Appendix B.

**Problem**: Is $t$-Hereditary $s$-Club (positive integers $s, t$)
**Question**: Given a graph $G = (V, E)$ and a subset $S \subseteq V$, is $S$ a $t$-hereditary $s$-club in $G$?

Theorem 5. Is $t$-Hereditary $s$-Club *is coNP-complete for every fixed integer* $s \geq 5$ *and arbitrary positive integer* $t$.

Remark 1. If instead, parameter $t$ is fixed in the problem and $s$ is specified in the input, verifying whether $S$ is a $t$-hereditary $s$-club can be completed in polynomial time by enumerating every pair of vertices $u, v \in S$ and every deletion set $T \subseteq S \setminus \{u, v\}$ of size less than $t$; and then verifying if $\operatorname{diam}(G[S \setminus T]) \leq s$. If the diameter bound is not satisfied for some $T$, then $S$ is not a $t$-hereditary $s$-club, and otherwise, it is.

## 3. Integer Programming Formulations

We introduce cut-like formulations based on length-bounded vertex separators for the MRCP and MHCP and compare their strength with existing formulations in the literature whenever alternate formulations are available. We also consider the complexity of the associated separation problems due to the worst-case exponential size of our formulations.

### 3.1. Cut-Like Formulation for the MHCP

The cut-like formulations we introduce are based on vertex separators that disconnect all length-bounded paths between a specified pair of vertices. Similar ideas have been used to impose connectivity constraints in other settings; see for instance (Salemi and Buchanan 2020, Wang et al. 2017, Carvajal et al. 2013).

In the IP formulation of the MHCP proposed next, we use the following additional notations. Given a graph $G = (V, E)$, denote its complement by $\overline{G} = (V, \overline{E})$, where the edge set is defined as $\overline{E} \coloneqq \{e \subseteq V \mid |e| = 2, e \notin E\}$. For each edge $uv \in \overline{E}$, let $\mathcal{C}_{uv}(G)$ denote the collection of all length-$s$ $u,v$-separators in $G$. In formulation (1), binary variable $x_i$ equals one if and only if vertex $i \in V$ is included in the $t$-hereditary $s$-club.

$$\max \sum_{i \in V} x_i \tag{1a}$$

$$\text{s.t.} \quad t(x_u + x_v - 1) \leq \sum_{i \in C} x_i \qquad \forall C \in \mathcal{C}_{uv}(G), \forall uv \in \overline{E} \tag{1b}$$

$$x_i \in \{0, 1\} \qquad \forall i \in V. \tag{1c}$$

PROPOSITION 2. *Given a graph $G = (V, E)$, a subset of vertices $S$ is a $t$-hereditary $s$-club if and only if its characteristic vector $x^S$ satisfies the constraints of formulation* (1).

*Proof.* ($\Longrightarrow$) Let $S \subseteq V$ and suppose that $t(x_u^S + x_v^S) - \sum_{i \in C} x_i^S > t$ for some $C \in \mathcal{C}_{uv}(G)$. This implies that $u, v \in S$ and $|C \cap S| < t$. Let $D = C \cap S$. Then $S$ violates the definition of a $t$-hereditary $s$-club as $|D| < t$, and $S \setminus D$ is not an $s$-club as there is no $u,v$-path of length at most $s$ in $G[S \setminus C]$.

($\Longleftarrow$) Suppose that $S \subseteq V$ is not a $t$-hereditary $s$-club in $G$. Then, it contains a deletion set $D \subset S$ with $|D| < t$ such that $S \setminus D$ is not an $s$-club. Hence, there exist vertices $u, v \in S \setminus D$ such that the distance between them in $G[S \setminus D]$ is greater than $s$. So, $D \cup (V \setminus S)$ is a length-$s$ $u,v$-separator in $G$, and it can be verified that $x^S$ violates the corresponding constraint (1b). $\square$

Clearly, it is sufficient to only consider length-$s$ $u,v$-separators in (1b) that are minimal by inclusion, and $\mathcal{C}_{uv}(G)$ can be safely redefined to only contain minimal members. In particular, when $s = 2$, $\mathcal{C}_{uv}(G) = \{N(u) \cap N(v)\}$ as the common neighbors form the unique minimal length-2 $u,v$-separator. As a result, formulation (1) generalizes the formulation of $t$-hereditary 2-clubs presented by Komusiewicz et al. (2019).

Although we are only required to consider minimal length-bounded separators in a complete and correct formulation, there can still be prohibitively many such sets to enumerate. In a practical implementation of this cut-like formulation, we would typically work with a relaxation that uses only a subset of these constraints and employ a delayed constraint generation scheme where a violated cut-like constraint is detected on-the-fly and added to strengthen the relaxation during the progress of an algorithm. In this regard, it helps to know the complexity of the associated separation problem, i.e., identifying a constraint (1b) violated by a given solution $x^* \in [0,1]^n$ to the relaxation or conclude that all such constraints are satisfied (Grötschel et al. 1993).

**Problem**: Separation of length-$s$ $u,v$-separator inequalities (1b).
**Input**: A graph $G=(V,E)$, $x^* \in [0,1]^n$, and positive integers $t$ and $s$.
**Output**: If any exist, nonadjacent vertices $u,v \in V$ and a length-$s$ $u,v$-separator $C \subseteq V \setminus \{u,v\}$ such that $t(x_u^* + x_v^* - 1) > \sum_{i \in C} x_i^*$.

To solve this separation problem, we can treat $x_i^*$ for $i \in V$ as vertex weights, and for each $uv \in \overline{E}$ find a length-$s$ $u,v$-separator $C$ of minimum weight $\sum_{i \in C} x_i^*$ in $G$. If we find a pair $uv \in \overline{E}$ and a minimum weight separator $C$ such that $t(x_u^* + x_v^* - 1) > \sum_{i \in C} x_i^*$, we have identified a violated constraint; or, after checking all non-adjacent vertex pairs, we may conclude that no violated constraint exists.

A minimum-weight length-$s$ $u,v$-separator can be found in polynomial time when $s \in \{2,3,4\}$, but the problem is NP-hard when $s \geq 5$ (Lovász et al. 1978, Itai et al. 1982). Consequently, Salemi and Buchanan (2020) show that when $t=1$, for each $s \geq 5$, determining whether a given point $x^*$ satisfies all length-$s$ $u,v$-separator inequalities (1b) is coNP-complete, and their result applies for every $t \geq 2$ after a slight modification.

PROPOSITION 3. *For every pair of fixed integers $s \geq 5$ and $t \geq 1$, it is coNP-complete to determine whether a given $x^* \in \mathbb{R}^n$ satisfies all length-$s$ $u,v$-separator inequalities* (1b).

### 3.2. Cut-Like Formulation for the MRCP for $s \in \{2,3,4\}$

Unlike a $t$-hereditary $s$-club, every pair of *adjacent* vertices in an $r$-robust $s$-club is still required to satisfy additional requirements, i.e., they need to be connected by at least $r-1$ other vertex-disjoint paths of length $s$ or less. We use the notation $\binom{V}{2}$ to denote all 2-element subsets of $V$ and the short-form $G-uv$ to denote the graph obtained by deleting the edge $uv$, *if it exists.* For every pair of vertices $uv \in \binom{V}{2}$, we use $\mathbb{1}_E(u,v)$ as the edge indicator function, i.e., $\mathbb{1}_E(u,v)=1$ if $uv \in E$ and 0 otherwise. Binary variable $x_i$ equals one if and only if vertex $i \in V$ is included in the $r$-robust $s$-club. An analogous cut-like formulation for MRCP is proposed next, which we show is correct when $s \in \{2,3,4\}$ in

Theorem 6. As alluded to in Section 2.2, this is a consequence of Proposition 1 offering a length-bounded Mengerian theorem only when $s \in \{2,3,4\}$, but not when $s \geq 5$.

$$\max \sum_{i \in V} x_i \tag{2a}$$

$$\text{s.t.} \quad (r - \mathbb{1}_E(u,v))(x_u + x_v - 1) \leq \sum_{i \in C} x_i \qquad \forall C \in \mathcal{C}_{uv}(G - uv), \forall uv \in \binom{V}{2} \tag{2b}$$

$$x_i \in \{0,1\} \qquad \forall i \in V. \tag{2c}$$

THEOREM 6. *Given a graph $G = (V, E)$ and parameter $s \in \{2,3,4\}$, a subset of vertices $S$ is an $r$-robust $s$-club if and only if its characteristic vector $x^S$ satisfies the constraints of formulation* (2).

*Proof.* ($\Longrightarrow$) Let $S \subseteq V$ and suppose $(r - \mathbb{1}_E(u,v))(x_u^S + x_v^S - 1) > \sum_{i \in C} x_i^S$ for some $C \in \mathcal{C}_{uv}(G - uv)$. It implies that $u, v \in S$ and thus $\sum_{i \in C} x_i^S < r - \mathbb{1}_E(u,v)$. Let $C' = S \cap C$, then $|C'| \leq r - 1 - \mathbb{1}_E(u,v)$. Because $C \in \mathcal{C}_{uv}(G - uv)$, $C'$ is a length-$s$ $u,v$-separator in $G[S] - uv$. Hence, $S$ is not an $r$-robust $s$-club based on the following chain of inequalities:

$$\begin{aligned}\rho_s(G[S]; u, v) = \rho_s(G[S] - uv; u, v) + \mathbb{1}_E(u,v) &\leq \kappa_s(G[S] - uv; u, v) + \mathbb{1}_E(u,v) \\ &\leq |C'| + \mathbb{1}_E(u,v) \leq r - 1.\end{aligned}$$

(Note that the foregoing inequality does not make use of the length-bounded Menger's theorem for $s \in \{2,3,4\}$, only the dual relationship between $\rho_s(\cdot)$ and $\kappa_s(\cdot)$. Therefore, the length-$s$ $u,v$-separator inequality (2b) will be satisfied by the characteristic vector of every $r$-robust $s$-club even when $s \geq 5$.)

($\Longleftarrow$) Suppose $S$ is not an $r$-robust $s$-club. It follows that there exist two vertices $u, v \in S$ such that $\rho_s(G[S]; u, v) \leq r - 1$. Then, it follows from Proposition 1 that for $s \in \{2,3,4\}$,

$$\kappa_s(G[S] - uv; u, v) + \mathbb{1}_E(u,v) = \rho_s(G[S] - uv; u, v) + \mathbb{1}_E(u,v) = \rho_s(G[S]; u, v) \leq r - 1.$$

Now consider a minimum size length-$s$ $u,v$-separator $C'$ in $G[S] - uv$. Then, $|C'| \leq r - 1 - \mathbb{1}_E(u,v)$. As before, $C' \cup (V \setminus S)$ belongs to $\mathcal{C}_{uv}(G - uv)$, and the corresponding constraint (2b) is violated by the characteristic vector of $S$. $\square$

As noted in the proof of Theorem 6, formulation (2) is a relaxation of the feasible region of the MRCP when $s \geq 5$. In this case, formulation (2) may be satisfied by binary vectors that do not correspond to $r$-robust $s$-clubs. For instance, the vertex set of the graph in Figure 2 is not a 2-robust 5-club as $\rho_5(G;10,11)=1$. However, we can satisfy all constraints (2b) by setting $x_i=1$ for every vertex $i$ in the graph in Figure 2. In particular, as $\kappa_5(G;10,11)=2$, every $C\in\mathcal{C}_{10,11}(G-\{10,11\})$ contains at least two vertices and the left-hand side of constraints (2b) is at most two.

In any practical implementation of the cut-like formulation (2) we would need to use delayed constraint generation as exponentially many minimal length-bounded separators can exist in the worst case. To that end, given $x^*\in[0,1]^n$ we would solve the separation problem by finding a minimum weight length-$s$ $u,v$-separator $C\subseteq V\setminus\{u,v\}$ in $G-uv$ iterating over $\{u,v\}\in\binom{V}{2}$ and checking whether $(r-\mathbb{1}_E(u,v))(x_u^*+x_v^*-1)>\sum_{i\in C}x_i^*$. As discussed in Section 3.1 this problem can be solved in polynomial time for $s\in\{2,3,4\}$.

We can further strengthen formulation (2) using the "conflict" inequalities given below:

$$x_u+x_v\leq 1,\quad \forall uv\in\binom{V}{2} \text{ such that } \rho_s(G;u,v)\leq r-1. \tag{3}$$

The validity of these inequalities follows from the observation that a pair of distinct vertices $u$ and $v$ that do not have an adequate number of vertex disjoint paths of length at most $s$ cannot be simultaneously included in an $r$-robust $s$-club. In addition to strengthening the linear programming (LP) relaxation of formulation (2), these inequalities are candidates for initializing a master relaxation that can be used in the aforementioned delayed constraint generation framework. We describe this approach in greater detail in Section 5.4.

The MRCP has been formulated for $s=2$ and $s=3$ in the literature. Proposition 4 that follows, establishes that the cut-like formulation (2) has a tighter LP relaxation than the formulation of the maximum $r$-robust 2-club problem presented by Veremyev and Boginski (2012a). In Proposition 5, we establish that the LP relaxations of the maximum $r$-robust 3-club problem formulation proposed by Almeida and Carvalho (2014) and that of the cut-like formulation (2) strengthened by inequalities (3) are incomparable. The proofs of these results are included in Appendix C.

PROPOSITION 4. *The cut-like formulation* (2) *has a tighter LP relaxation than that of the Veremyev and Boginski (2012a) formulation of the maximum $r$-robust 2-club problem when $r\geq 2$.*

PROPOSITION 5. *The LP relaxations of the formulation of the maximum $r$-robust 3-club problem proposed by Almeida and Carvalho (2014) and the cut-like formulation* (2) *strengthened by inequalities* (3) *are incomparable.*

Despite Proposition 5, the computational results in Section 6 show that a decomposition branch-and-cut algorithm employing formulation (2) with the master problem initialized by inequalities (3) is faster than solving the formulation of Almeida and Carvalho (2014).

## 4. The Existence of Compact Formulations

In light of the exponential size of the cut-like formulations in the worst case, it is reasonable to ask whether the problems admit compact (i.e., polynomial-size) formulations in general. In this section, we study whether or not $t$-hereditary $s$-clubs and $r$-robust $s$-clubs admit compact formulations depending on the values of $t$ (or $r$) and $s$. The results established in this section are summarized in Table 3. This may explain why previous works (Veremyev and Boginski 2012a, Almeida and Carvalho 2014) and ours have failed to create compact formulations for general values of parameter $s$.

**Table 3** **Existence of compact formulations**

| Model | Constant fixed in the problem | Parameter specified in the input | Existence |
|---|---|---|---|
| $t$-hereditary $s$-club | $s \le 4$ | $t$ | Exist |
| | $s \ge 5$ | $t$ | Unlikely |
| | $t = O(1)$ | $s$ | Exist |
| $r$-robust $s$-club | $s \le 4$ | $r$ | Exist |
| | $s \ge 5$ | $r$ | Unlikely |
| | $r \ge 2$ | $s$ | Unlikely |

### 4.1. Existence of Compact Formulations for $t$-Hereditary $s$-Clubs

We begin with an impossibility result for each $s \ge 5$. It is based on the hardness of Is $t$-Hereditary $s$-Club established in Theorem 5. Following this result we discuss the existence of compact formulations for $s \le 4$.

Proposition 6. *If P $\neq$ NP, then for every fixed integer $s \ge 5$ and arbitrary positive integer $t$ there is no polynomial-time procedure for constructing IP formulations for $t$-hereditary $s$-clubs.*

*Proof.* Suppose there were a polynomial-time procedure for constructing an IP formulation for $t$-hereditary $s$-clubs (for some fixed integer $s \ge 5$ and when $t$ is part of the input). Then, one could verify whether a given subset of vertices is a $t$-hereditary $s$-club in polynomial time (by plugging its characteristic vector into the IP formulation). By Theorem 5, this would imply that P = coNP, which is equivalent to P = NP. □

Despite the foregoing negative result, we point out that there exist compact formulations for $t$-hereditary $s$-clubs when $t=O(1)$, although the proposed approach is not expected to be of practical interest in solving large-scale instances of the problem.

PROPOSITION 7. *There exist size $O(tsn^{t+1})$ IP formulations for $t$-hereditary $s$-clubs.*

*Proof.* Recall that Veremyev and Boginski (2012a) introduced IP formulations for (1-hereditary) $s$-clubs that use $b$ binary variables and $c$ constraints (henceforth "VB formulation"), where $b=O(sn^2)$ and $c=O(sn^2)$. In other words, for any graph $G$, there exists a set $Q_s(G)\subseteq\mathbb{R}^{n+p}$ defined by $c$ linear inequalities such that $x$ is the characteristic vector of an $s$-club of $G$ if and only if there exists $y\in\{0,1\}^p$ such that $(x,y)\in Q_s(G)$.

We can assume $t\geq 2$ in the formulation we propose next. For every $T\subseteq V$ with $0\leq |T|<t$, construct the VB formulation $Q_s(G-T)$ for graph $G-T$ with respect to variables $(x|_{\overline{T}},y^T)$. Here, $x|_{\overline{T}}$ denotes the subvector of $x$ restricted to those variables indexed by $\overline{T}:=V\setminus T$, and each $y^T$ is a new set of $b_T=O(sn^2)$ variables. Thus, the formulation can be described as:

$$(x|_{\overline{T}},y^T)\in Q_s(G-T) \qquad \forall T\subseteq V \text{ such that } 0\leq|T|<t \tag{4a}$$

$$y^T\in\{0,1\}^{b_T} \qquad \forall T\subseteq V \text{ such that } 0\leq|T|<t \tag{4b}$$

$$x\in\{0,1\}^n. \tag{4c}$$

So, the number of variables and constraints in formulation (4) is $O\left(sn^2\sum_{i=0}^{t-1}\binom{n}{i}\right)=O(tsn^{t+1})$. Now we claim that the formulation is correct, i.e., $\bar{x}$ is the characteristic vector of a $t$-hereditary $s$-club in $G$ if and only if there exists a collection of 0-1 vectors $\{\bar{y}^T\}_T$ such that $(\bar{x},\{\bar{y}^T\}_T)$ satisfies formulation (4).

($\Longrightarrow$) Suppose $\bar{x}$ is the characteristic vector of a $t$-hereditary $s$-club $D\subseteq V$ in $G$. Let $T\subseteq V$ be an arbitrary vertex subset with $0<|T|<t$. Let $T':=T\cap D$, thus $|T'|\leq|T|<t$. By the definition of a $t$-hereditary $s$-club, $D\setminus T'$ is an $s$-club of $G$, and also of $G-T$. Thus, by the correctness of the VB formulation, there exists a binary vector $\bar{y}^T$ such that $(\bar{x}|_{\overline{T}},\bar{y}^T)\in Q_s(G-T)$. Thus, there is an appropriate collection $\{\bar{y}^T\}_T$ of 0-1 vectors such that $(\bar{x},\{\bar{y}^T\}_T)$ satisfies formulation (4).

($\Longleftarrow$) Suppose that $(\bar{x},\{\bar{y}^T\}_T)$ satisfies formulation (4). By constraint (4c), $\bar{x}$ is the characteristic vector of some set $D\subseteq V$. To show that $D$ is a $t$-hereditary $s$-club in $G$, consider some $F\subseteq D$ with $|F|<t$. We are to show that $D\setminus F$ is an $s$-club in $G$. By

constraints (4a) with $T := F$ (and by correctness of the VB formulation), $D \setminus F$ is an $s$-club in $G - T$, and is thus a $t$-hereditary $s$-club in $G$, as $F$ is arbitrary. □

As noted in Section 3.1, when $s = 2$, formulation (1) is compact because the common neighbors of nonadjacent vertices $u$ and $v$ form a unique minimal length-$s$ $u,v$-separator. We argue that compact extended formulations also exist for $t$-hereditary $s$-clubs when $s \in \{3,4\}$. From the compact formulation of Almeida and Carvalho (2014) for $r$-robust 3-clubs (see formulation (10) in Appendix C), we can drop the constraints on adjacent pairs of vertices to obtain a compact formulation for $t$-hereditary 3-clubs. Recall that the separation problem of the MHCP for $s \in \{3,4\}$ can be reduced to a min-cut problem (Lovász et al. 1978). Based on the technique introduced by Martin (1991), we can create a compact extended formulation for the the MHCP when $s = 4$. The construction of compact formulations in this manner is not very practical for our problems, so we do not discuss it here any further. Interested readers can refer to Carr and Lancia (2002) on building compact formulations provided that the separation problems admit compact LP formulations.

### 4.2. Existence of Compact Formulations for $r$-Robust $s$-Clubs

Based on Theorems 3 and 4 we can immediately deduce the unlikelihood of compact formulations for $s \geq 5$ and $r \geq 2$ as the following two propositions state (without proof). However, contrast the unlikelihood result of Proposition 9 with the affirmative result of Proposition 7; this is of course a consequence of Theorem 4 establishing the hardness of verifying $r$-robust $s$-clubs even for fixed constant $r$.

PROPOSITION 8. *If P ≠ NP, then for every fixed integer $s \geq 5$ and arbitrary positive integer $r$ there is no polynomial-time procedure for constructing an IP formulation for the MRCP.*

PROPOSITION 9. *If P ≠ NP, then for every fixed integer $r \geq 2$ and arbitrary positive integer $s$ there is no polynomial-time procedure for constructing an IP formulation for the MRCP.*

Compact formulations for $r$-robust 2-clubs and $r$-robust 3-clubs already exist in the literature (Veremyev and Boginski 2012a, Almeida and Carvalho 2014). As the separation problem of MRCP for $s = 4$ can be reduced to a min-cut problem, analogous to our discussion with regards to $t$-hereditary 4-clubs (Lovász et al. 1978), the techniques presented by Martin (1991) can be used to construct a compact formulation for $r$-robust 4-clubs.

## 5. Recursive Block Decomposition Algorithm

In this section, we turn our focus to computational techniques that are effective in solving the problems using the cut-like formulations introduced in Section 3. Given a graph $G = (V, E)$, a *block* is a maximal biconnected[1] subgraph of $G$, and the *block decomposition* of $G$ is the collection of all the blocks of $G$ (see Figure 3). Note that every vertex of $G$ belongs to some block in the decomposition, and two distinct blocks of $G$ can share at most one vertex. Every $r$-robust ($t$-hereditary) $s$-club, whenever $r, t \geq 2$, must be contained within a single block of $G$. Based on this observation, we present a block decomposition approach to solve the MRCP and MHCP.

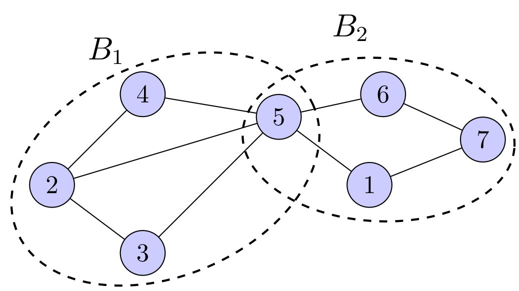


**Figure 3** **A graph $G$ that decomposes into two blocks $B_1$ and $B_2$**

The main idea is to decompose the original graph $G$ into many smaller blocks so we can restrict our attention to one block at a time. Furthermore, with the help of a feasible solution obtained using a heuristic we can apply preprocessing techniques on each block. As a result, the preprocessed "blocks" may no longer be biconnected, entailing further decomposition into even smaller blocks. This motivates our *recursive block decomposition* approach to solve the MRCP and MHCP, described in Algorithm 1. We note here that maximal $r$-connected subgraphs of $G$ could be used in this approach instead of blocks, whenever $r \geq 3$. However, our preliminary experiments indicated that repeatedly finding them was time consuming and the trade-off was not favorable in terms of finding potentially smaller subgraphs on which to solve the problems. Pertinently, the computational complexity of finding all blocks in a graph $G$ is $O(m+n)$ (Hopcroft and Tarjan 1973) while, to our best knowledge, finding all maximal $r$-connected subgraphs of $G$ can only be performed in $O(mn^2 \min\{r, \sqrt{n}\})$ (Carmesin et al. 2014, Matula 1978).

[1] a cut-vertex, as well as a bridge and its end-vertices are considered biconnected subgraphs

**Algorithm 1:** Recursive Block Decomposition for the MRCP and MHCP

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

**Output:** A maximum cardinality $r$-robust ($t$-hereditary) $s$-club $K$.

1 find the block decomposition $\mathcal{B}$ of $G$
2 $K \leftarrow$ a heuristic solution of MRCP (MHCP) on the largest block in $\mathcal{B}$
3 **while** $\mathcal{B} \neq \emptyset$ **do**
4     pick $D \in \mathcal{B}$ with the most vertices
5     **if** $|D| \leq |K|$ **then**
6         **return** $K$
7     $\mathcal{B} \leftarrow \mathcal{B} \setminus \{D\}$
8     preprocess block $D$ by vertex peeling using solution $K$
9     find the block decomposition $\mathcal{F}$ of $D$
10     **if** $|\mathcal{F}|=1$ **then**
11         $K' \leftarrow$ a maximum $r$-robust ($t$-hereditary) $s$-club in $D$
12         **if** $|K'|>|K|$ **then**
13             $K \leftarrow K'$
14     **else**
15         $\mathcal{B} \leftarrow \mathcal{B} \cup \mathcal{F}$
16 **return** $K$

We choose a "greedy" strategy for solving the MRCP (MHCP) on a block with most vertices first, and then update the current best solution as needed after each block is considered. If the block with the most vertices has fewer vertices than the current largest solution, the algorithm is terminated, and the current best solution is indeed optimal. In line 1 of Algorithm 1, we find the block decomposition of $G$ in $O(m+n)$ time (Hopcroft and Tarjan 1973). On a block with the most vertices we find a heuristic solution (line 2) that serves as a lower bound. In the while-loop, we preprocess the current block (line 8), decompose it into smaller blocks if possible (line 9), and add them to collection $\mathcal{B}$ for future consideration (line 15). If it cannot be further decomposed after preprocessing, we solve the MRCP (MHCP) on this block (line 11). The algorithm terminates when the largest unexplored block contains fewer vertices than the current best objective value.

Next we discuss ideas for reducing the computational burden involved in computing $\rho_s$, which is a frequent task in the heuristic used in line 2, preprocessing used in line 8, and implementing the initial relaxation used in the decomposition branch-and-cut algorithm in line 11 of Algorithm 1. We discuss these steps in greater detail in Sections 5.2, 5.3, and 5.4.

### 5.1. Computing Bounds on $\rho_s$

Recall from Section 2.2 that $\rho_s(G;u,v)$ denotes the maximum number of internally vertex-disjoint $u,v$-paths of length at most $s$ in $G$. When $s=2$, we know that $\rho_s(G;u,v) = |N(u)\cap N(v)| + \mathbb{1}_E(u,v)$, which can be computed in $O(n)$ time.

When $s \in \{3,4\}$, we can use the approach introduced by Itai et al. (1982) that applies max flow–min cut theorem to an auxiliary flow network to compute $\rho_s(G-uv;u,v)$ and find a minimum cardinality length-$s$ $u,v$-separator. Hence, for a given pair of vertices $u,v\in V$ we can check if $\rho_s(G;u,v)\geq r$ using the Ford Jr. and Fulkerson (1956) algorithm in $O(rm)$ time. As we only consider small values of $r$ in our experiments, the approach essentially takes $O(m)$ time, if we treat $r$ as a constant. When we compute $\rho_s(G;u,v)$ using a flow augmenting algorithm, we can terminate early once we confirm that it is at least $r$ (before computing its actual value). Limiting the number of $u,v$-pairs for which we need to compute $\rho_s$ can be even more significant from a computational perspective. For example, if $|N(u)\cap N(v)|\geq r$, then $\rho_s(G;u,v)\geq r$. Likewise, if $u$ and $v$ are in different components, or even if they are in the same component but in different *blocks*, we know that $\rho_s(G;u,v)\leq 1$. These observations motivate us to explore techniques to quickly or incrementally compute upper and lower bounds of $\rho_s$ that can be used to reduce the overall computational overhead. Given the value of $\rho_2$, we can calculate upper bounds of $\rho_s$ for any $s\geq 3$ based on a recursive relationship between $\rho_{s-1}$ and $\rho_s$ using Lemma 2.

Lemma 2. *Given a graph $G=(V,E)$, a pair of vertices $uv\in\binom{V}{2}$, and a positive integer $s\geq 3$, we have*

$$\rho_s(G;u,v)\leq \mathbb{1}_E(u,v)+\sum_{w\in N(v)\setminus\{u\}}\min\{1,\rho_{s-1}(G;u,w)\}.$$

*Proof.* Let $P_1,P_2,\ldots,P_{\rho_s}$ be a set of vertex-disjoint paths of length at most $s$ between vertices $u$ and $v$ (see Figure 4). Each of these paths must use a distinct vertex in $N(v)$ to reach $v$, and for each such vertex $w\in N(v)\setminus\{u\}$ we have $\rho_{s-1}(G;u,w)\geq 1$. Thus, any vertex $w\in N(v)\setminus\{u\}$ such that $\rho_{s-1}(G;u,w)=0$ cannot be on any such path. The claim follows by observing that the right-hand side of the inequality is the same as deducting the number of such $w \in N(v) \setminus \{u\}$ from $\deg(v)$. □

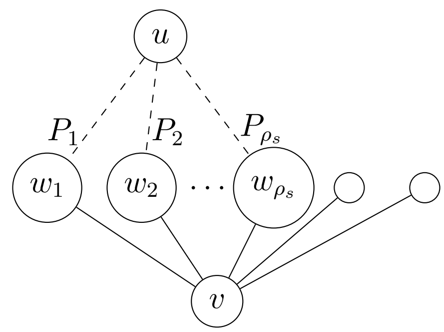


**Figure 4** **A set of maximum number of length-bounded vertex-disjoint $u$-$v$ paths**

Consider any valid upper bound $\bar{\rho}_s(G;u,v) \geq \rho_s(G;u,v)$. We will refer to $\bar{\rho}_s(G;u,\cdot)$ as the *single-source upper bounds.* Given $\bar{\rho}_{s-1}(G;u,\cdot)$, we can compute upper bounds $\bar{\rho}_s(G;u,v)$ in $O(\deg(v))$ time and single-source upper bounds $\bar{\rho}_s(G;u,\cdot)$ in $O(m)$ time using the recursion:

$$\bar{\rho}_s(G;u,v) = \mathbb{1}_E(u,v) + \sum_{w \in N(v)\setminus\{u\}} \min\{1, \bar{\rho}_{s-1}(G;u,w)\}\,. \tag{5}$$

For each pair of vertices $uv \in \binom{V}{2}$, Algorithm 4 in Appendix D describes a simple heuristic to obtain a lower bound of $\rho_s(G;u,v)$ for any $s \geq 3$, which we denote by $\hat{\rho}_3(G;u,v)$. Essentially, Algorithm 4 constructs a matching in the bipartite graph $G_{uv} = (V_{uv}, E_{uv})$ where $E_{uv} \coloneqq \{\{p,q\} \in E \mid p \in N(u) \setminus N[v], q \in N(v) \setminus N[u]\}$ and $V_{uv}$ is the union of endpoints of edges in $E_{uv}$. The size of the matching then gives us a lower bound on the number of disjoint length-3 paths between vertices $u$ and $v$, because each edge in the matching is the inner edge of such a path. Algorithm 4 can be implemented to run in $O(m)$ and is usually fast in practice. Previous empirical studies also report that the simple greedy matching heuristic used in Algorithm 4 can usually produce a solution at least 90% the size of an optimum, even though it is only known to guarantee a 2-approximation (Langguth et al. 2010, Möhring and Müller-Hannemann 1995, Magun 1998). Furthermore, although $\rho_3$ can be computed in $O(rm)$, it requires additional effort to construct the auxiliary flow network on which we need to solve the maximum flow problem. These factors make Algorithm 4 helpful in practice for reducing running times when solving the MRCP and MHCP.

If the lower bound $\hat{\rho}_s(G;u,v)$ is at least $r$ or the upper bound $\bar{\rho}_s(G;u,v)$ is at most $r-1$, there is no need to run the Ford–Fulkerson algorithm. Using this observation significantly decreases the number of pairs of vertices that require the application of the Ford–Fulkerson algorithm and the overall running time taken to check if $\rho_s(G;u,v) \geq r$, as reported in Table 5 of Appendix E. Taking the instance PGP as an example, it suffices to verify $\rho_4 \geq 2$ for 727,213 out of 57,025,860 pairs of vertices using the Ford–Fulkerson algorithm, reducing the running time required from 155.08 seconds to 5.12 seconds (which includes the time to compute the lower and upper bounds).

### 5.2. Heuristics

In this section, we discuss heuristics for finding a feasible solution that we subsequently use for preprocessing in Section 5.3. The first heuristic described in Algorithm 2 for finding an $r$-robust $s$-club, generalizes the greedy vertex elimination heuristic proposed by Bourjolly et al. (2000) for finding an $s$-club.

A pair of vertices $i,j \in V$ cannot be included in the same $r$-robust $s$-club if $\rho_s(G;i,j) \leq r-1$. We call a pair of vertices $i$ and $j$ *compatible* if they satisfy $\rho_s(G;i,j) \geq r$. Our heuristic first builds a maximal subset $S \subseteq V$ that is pairwise compatible. Essentially, we seek a maximal clique $S$ in the compatibility graph $G^c = (V, E^c)$ where $E^c := \left\{ ij \in \binom{V}{2} \mid \rho_s(G;i,j) \geq r \right\}$. This is similar to a technique used for $s$-clubs by Salemi and Buchanan (2020) where they begin by first constructing its (weakly) hereditary counterpart (cf. Pattillo et al. (2013)) i.e., an $s$-clique on an analogous compatibility graph. Note that although $\rho_s(G;i,j) \geq r$, it is possible that $\rho_s(G[S];i,j) \leq r-1$ in the vertex subset $S$ used in Algorithm 2. Hence, we need to check if $S$ is an $r$-robust $s$-club in $G$; if not, we select a vertex $v \in S$ that has the most vertices $w$ such that $\rho_s(G[S];v,w) \leq r-1$ and remove it from $S$. We repeat this step until $S$ is an $r$-robust $s$-club.

When $s \in \{3,4\}$, line 1 in Algorithm 2 can be completed in $O(n^2rm)$, and line 2 can be implemented to run in $O(|V| + |E^c|)$ time (Walteros and Buchanan 2020); we use the implementation provided by Salemi and Buchanan (2020) for this step. The while-loop in line 3 may require at most $\omega(G^c)$ (the clique number) iterations to complete, and the for-loop (line 5) in each iteration requires at most $O(rn^2m)$ time to complete.

The computational effort needed by this heuristic is dominated by the computation of pairwise $\rho_s$ values. So when constructing the compatibility graph in line 1, we take as much advantage of the bounds discussed in Section 5.1 as possible. We first check if the lower bound $\hat{\rho}_s(G;i,j)$ is at least $r$, in which case we create edge $ij$. Next we check if the upper bound $\bar{\rho}_s(G;i,j)$ is at most $r-1$, in which case we can conclude that the pair

**Algorithm 2:** A heuristic for finding an $r$-robust $s$-club

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

**Output:** An $r$-robust $s$-club $S$.

1 create compatibility graph $G^c \leftarrow (V, E^c)$, where $E^c := \left\{ij \in \binom{V}{2} \,\middle|\, \rho_s(G; i, j) \geq r\right\}$
2 $S \leftarrow$ a maximal clique in $G^c$
3 **while** $S \neq \emptyset$ **do**
4     $\tau_i \leftarrow 0, \forall i \in S$
5     **for** $ij \in \binom{S}{2}$ **do**
6         **if** $\rho_s(G[S]; i, j) \leq r - 1$ **then**
7             $\tau_i \leftarrow \tau_i + 1$
8             $\tau_j \leftarrow \tau_j + 1$
9     $v \leftarrow \arg\max_{i \in S} \tau_i$
10     **if** $\tau_v \geq 1$ **then**
11         $S \leftarrow S \setminus \{v\}$
12     **else**
13         **return** $S$

is incompatible. The lower and upper bounds are similarly used in line 6 to check the condition of the if-statement faster. The Ford–Fulkerson algorithm is used to exactly verify the conditions only in cases where verification using the bounds has been inconclusive.

As a result, the running times for our heuristic are reasonable in practice for a one-time application. As reported in Table 7 in Appendix E, the longest time taken by the heuristic was 9.58 seconds for the instance `email` with $s = 4$ and $r = 3$ and the average time taken across all instances is 0.43 seconds. Furthermore, as reported in Table 9 in Appendix E, $r$-robust $s$-clubs found by this heuristic were subsequently proved to be optimal in 50 out of 126 instances. We can extend this heuristic to find $t$-hereditary $s$-clubs for $s \in \{2, 3, 4\}$ by slightly modifying lines 1, 5, and 6 (see Algorithm 5 in Appendix D). The performance of this heuristic is reported in Tables 8 and 10 in Appendix E.

### 5.3. Preprocessing

Vertex peeling is a generic term applied to techniques in which we delete vertices from the graph based on a heuristic solution, without affecting the optimality guarantee and correctness of a subsequent exact algorithm. We discuss a vertex peeling technique applicable for the MRCP and then extend this idea to the MHCP for $s \in \{2,3,4\}$.

Given a solution of size $\ell$ for the MRCP, we can delete a vertex that has fewer than $\ell$ distance-$s$ neighbors, as it cannot be part of a solution whose size is greater than $\ell$. This technique is often used when solving the maximum $s$-club problem (Veremyev and Boginski 2012b, Moradi and Balasundaram 2018, Lu et al. 2018, Salemi and Buchanan 2020). For the MRCP, we can strengthen this idea by deleting a vertex $v \in V$ if it has fewer than $\ell$ *compatible* distance-$s$ neighbors, i.e., $|T_v| < \ell$ where $T_v := \{u \in N_G^s(v) \mid \rho_s(G;v,u) \geq r\}$.

Consider the graph on the left in Figure 5 that contains the 2-robust 2-club $\{1,2,3\}$, i.e., $\ell = 3$. For each vertex $v \in \{1,2,3,4,5\}$, we can see that $|T_v| \geq 3$ and therefore they are not removed by vertex peeling. However, for each $v \in \{6,7,8\}$, we have $|T_v| < 3$. In particular, although the distance-2 neighborhood of vertex 7 contains three vertices, the set $T_7 = \{5\}$ contains only one vertex. Therefore, vertices $\{6,7,8\}$ and their incident edges are removed by vertex peeling shown on the right in Figure 5.


**Figure 5 An illustration of vertex peeling for the MRCP.**

In addition, vertices of degree less than $r$ can also be removed from $G$ as the degree of every vertex in an $r$-robust $s$-club must be at least $r$. In other words, every $r$-robust $s$-club in $G$ is contained within its $r$-core (the maximal induced subgraph of $G$ with minimum degree at least $r$). We can recursively implement these ideas as each vertex $v$ that is removed may affect the size of the distance-$s$ neighborhood or the degree of another vertex. The pseudocode for vertex peeling that we propose is described in Algorithm 3.

The $r$-core of $G$ in line 2 of Algorithm 3 can be be found using an $O(m+n)$ algorithm (Matula and Beck 1983, Batagelj and Zaversnik 2003). The repeat-until loop may execute at most $n$ times and each iteration can be completed in $O(n^2rm)$ time if we

**Algorithm 3:** Vertex peeling based on an $r$-robust $s$-club of size $\ell$

```
Input: A graph G = (V, E) and a lower bound ℓ.
Output: Preprocessed graph G.
repeat
    G ← the r-core of G
    S ← ∅
    for v ∈ V(G) do
        if |N_G^s(v)| < ℓ or |T_v| < ℓ then
            S ← S ∪ {v}
    if S ≠ ∅ then
        G ← G − S
until S = ∅
return G
```

exhaustively verify $\rho_s(\cdot) \geq r$ for every vertex pair. But depite what the worst-case complexity suggests, our vertex peeling implementation is reasonably quick on our test bed of instances. As mentioned before it is not always necessary to compute $\rho_s(G;u,v)$ if we can determine that $\hat{\rho}_s(G;u,v) \geq r$ or $\bar{\rho}_s(G;u,v) \leq r-1$. The longest time taken by vertex peeling is 10.73 seconds for the instance `PGP` with $s=4$ and $r=2$, and the procedure took 0.44 seconds on average across our test bed. Approximately 90% of the instances in our test bed were preprocessed in less than one second (see Table 11 of Appendix E for details). Furthermore, we first decompose the graph into blocks and then apply vertex peeling on these blocks. This is also very effective in reducing the number of vertex pairs we need to consider.

Given a $t$-hereditary $s$-club of size $\ell$, we can extend the vertex peeling ideas discussed above to the MHCP as well. Recall from Proposition 1 that verifying if $S$ is a $t$-hereditary $s$-club, for $s \in \{2,3,4\}$, is equivalent to checking if $\rho_s(G[S];u,v) \geq t$ for every pair of nonadjacent vertices $u$ and $v$ in $S$. Hence, a vertex $v$ may be deleted if $\rho_s(G;u,v) \leq t-1$ for a sufficient number of its nonadjacent distance-$s$ neighbors $u$. In other words, vertex $v$ can be deleted if $|W_v| + |N_G(v)| < \ell$, where $W_v := \{u \in N_G^s(v) \setminus N_G(v) \mid \rho_s(G;u,v) \geq t\}$. The pseudocode of vertex peeling for the MHCP when $s \in \{2,3,4\}$ is presented in Algorithm 6 in Appendix D. Its performance is also comparable to the MRCP counterpart (see Table 12).

### 5.4. Delayed Constraint Generation

In this section, we describe decomposition approaches for solving the MHCP and the MRCP when $s \in \{2,3,4\}$. We employ the following relaxation based on conflict inequalities (3) at the root node of the branch-and-cut (BC) tree for the MRCP when $s \in \{2,3,4\}$.

$$\max \sum_{i \in V} x_i \tag{6a}$$

$$\text{s.t.} \quad x_u + x_v \le 1 \qquad \forall uv \in \binom{V}{2} \text{ such that } \rho_s(G; u, v) \le r-1 \tag{6b}$$

$$x_i \in \{0,1\} \qquad \forall i \in V. \tag{6c}$$

The BC algorithm starts by solving the initial relaxation (6) at the root node and branches when the LP relaxation optimum is fractional. It also prunes the search tree as usual when the node LP relaxation is infeasible and when the incumbent solution has a better objective value than the node LP bound. If we obtain an integral optimum $x^* \in \{0,1\}^n$ at some node of the BC tree, we check if the selected vertices $S := \{i \in V \mid x_i^* = 1\}$ form an $r$-robust $s$-club. Specifically, for each pair of vertices $u, v \in S$, we have to check if $\rho_s(G[S]; u, v) \ge r$. If $S$ is an $r$-robust $s$-club, then that node can be pruned and the incumbent is updated if necessary.

If we detect a pair of vertices $u, v \in S$ such that $\rho_s(G[S]; u, v) \le r-1$, then $S$ is not an $r$-robust $s$-club, and we construct a length-$s$ $u,v$-separator that corresponds to a constraint (2b) violated by $x^*$. If $s = 2$, then $N(u) \cap N(v)$ is the unique minimal length-2 $u,v$-separator in $G - uv$. If $s \in \{3,4\}$, we first identify a minimum cardinality length-$s$ $u,v$-separator $C$ in $G[S] - uv$ using the max flow–min cut theorem on the auxiliary flow network construction described by Itai et al. (1982). The set $C \cup (V \setminus S)$ is then a length-$s$ $u,v$-separator in $G - uv$, which is made minimal using the Minimalize algorithm described by Salemi and Buchanan (2020).

Using Proposition 1 when $s \in \{2,3,4\}$, we can tackle the MHCP using a decomposition BC algorithm that starts by solving the relaxation presented below.

$$\max \sum_{i \in V} x_i \tag{7a}$$

$$\text{s.t.} \quad x_u + x_v \leq 1 \qquad \forall uv \in \overline{E} \text{ such that } \rho_s(G;u,v) \leq t-1 \tag{7b}$$

$$x_i \in \{0,1\} \qquad \forall i \in V. \tag{7c}$$

As before, if we encounter an integral solution $x^* \in \{0,1\}^n$ at some node of the BC tree, we need to check if the selected vertices $S := \{i \in V \mid x_i^* = 1\}$ form a $t$-hereditary $s$-club. Specifically, we check if $\rho_s(G[S];u,v) \geq t$ for every pair of nonadjacent vertices $u$ and $v$ in the induced subgraph $G[S]$. If $\rho_s(G[S];u,v) \leq t-1$ for some pair $u$ and $v$, we add a length-$s$ $u,v$-separator inequality $t(x_u + x_v - 1) \leq \sum_{i \in C} x_i$ violated by $x^*$, in a manner analogous to the foregoing discussion for the MRCP.

During implementation, we take advantage of the lower and upper bounds introduced in Section 5.1 to quickly build the conflict constraints in the initial relaxations (6) and (7), by limiting the number of times we exactly verify if $\rho_s(G;u,v)$ is small enough. Furthermore, we solve the MRCP and the MHCP using the recursive block decomposition algorithm. The BC algorithms described above are only applied on blocks that are irreducible by vertex peeling, which helps reduce the scale of the instance solved by the BC algorithms.

## 6. Computational Study

The goal of the computational experiments is to assess the effectiveness of the cut-like IP formulations, preprocessing techniques, and the recursive block decomposition algorithm for solving the MHCP and the MRCP. We selected large-scale real-life graphs from the Tenth DIMACS Implementation Challenge on Clustering (Bader et al. 2013) that are frequently used as benchmarks for the maximum $s$-club problem. Numerical results are reported and discussed for the MRCP and the MHCP for the parameters $r \in \{2,3,4\}$ and $t \in \{2,3,4\}$, respectively. An instance in our DIMACS-10 test bed is thus defined by a graph from the collection and a value for the parameter $r$ or $t$. Note that our approaches are applicable for any positive integer-valued parameter $t$ or $r$, although we require $s \in \{2,3,4\}$ in our experiments. Formulation (2) for the MRCP is valid only for these values of $s$. Although formulation (1) of the MHCP is valid for every positive integer $s$, our separation procedure that is based on computing $\rho_s(\cdot)$ is only valid for $s \in \{2,3,4\}$.

All algorithms evaluated in this computational study are implemented in C++, and Gurobi™ Optimizer 9.0.1 (Gurobi Optimization 2020) is employed to solve the IP formulations in its default settings, other than the use of "lazy cuts" feature to implement the decomposition BC algorithm described in Section 5.4. We impose a one hour wall-clock time limit per instance for all solvers. If an instance was not solved to optimality within the time limit, we report the relative optimality gap calculated as (best bound − best objective)/best objective × 100%. In addition, we set the Gurobi cut-off parameter to the size of the largest $r$-robust ($t$-hereditary) $s$-club known at the time of calling the Gurobi optimization solver, informing the solver that we are only interested in solutions with better objective values. We conduct all numerical experiments on a 64-bit Linux® compute node running a dual Intel® Skylake 6130 processor with 32 cores and 96 GB RAM. We use parallel programming with the OpenMP library (Dagum and Enon 1998) when we implement the computation of length-bounded vertex-disjoint paths. Specifically, the tasks of finding the lower bound $\hat{\rho}_s(\cdot)$, the upper bound $\bar{\rho}_s(\cdot)$, and the exact value $\rho_s(\cdot)$ are parallelized. We observe 8-fold speed-up with OpenMP using all 32 cores over a single-threaded implementation; Table 6 in Appendix E contains more details.

### 6.1. Assessing the Cut-Like Formulations, Recursive Block Decomposition, and Preprocessing When $s = 2$

In this section, we focus on the case $s = 2$ as it admits comparison between multiple competing mathematical programming approaches. Specifically, we assess the performance of the recursive block decomposition Algorithm 1 to solve the MRCP and the MHCP using the delayed constraint generation scheme from Section 5.4 (labeled as "BCUT" in the tables), by comparing it against an implementation of the delayed constraint generation scheme without block decomposition, preprocessing, or speed-ups achieved using $\rho_s$-bounds (labeled as "CUT" in the tables). This comparison serves to highlight the impact of the graph decomposition and IP model decomposition techniques that we introduce, along with preprocessing and other ideas for achieving better performance. A direct "monolithic" implementation of the common neighbor formulation also serves as a baseline solver in this study (labeled as "CN" in the tables). The comparison between CUT and CN serves to demonstrate the benefits of using delayed constraint generation for the MHCP and the MRCP even when the formulation is compact.

Figure 6 shows performance profiles (Dolan and Moré 2002, Gould and Scott 2016) based on the wall-clock running times of solvers CUT and BCUT for the maximum $r$-robust 2-club and $t$-hereditary 2-club problems across all instances in our test bed. For each solver $i$ we plot $f_i(\tau)$—the fraction of the test instances for which the running time required by solver $i$ is at most a factor $\tau$ of the running time of the fastest solver for that instance. Following convention, we take the solution time to be equal to the time limit for instances that terminated by reaching the time limit (Dolan and Moré 2002). The performance profiles reflect the dominant performance of BCUT, which solved all the instances of the MRCP and the MHCP in this test bed to optimality when $s = 2$. From the details reported in Tables 13 and 14 in Appendix E for the MRCP and the MHCP respectively, we can see that BCUT outperforms CN and CUT on all instances with more than 150 vertices, demonstrating the effectiveness of the recursive block decomposition algorithm and preprocessing. For example, for the maximum $r$-robust 2-club problem, BCUT solves instance PGP with $r = 4$ to optimality in 0.1 seconds, while CUT takes 3,074.4 seconds, and CN fails to solve this instance under the time limit. Similarly, for the maximum $t$-hereditary 2-club problem, BCUT solves instance PGP with $t = 4$ to optimality in 0.1 seconds, while both CUT and CN fail to solve this instance under the time limit.

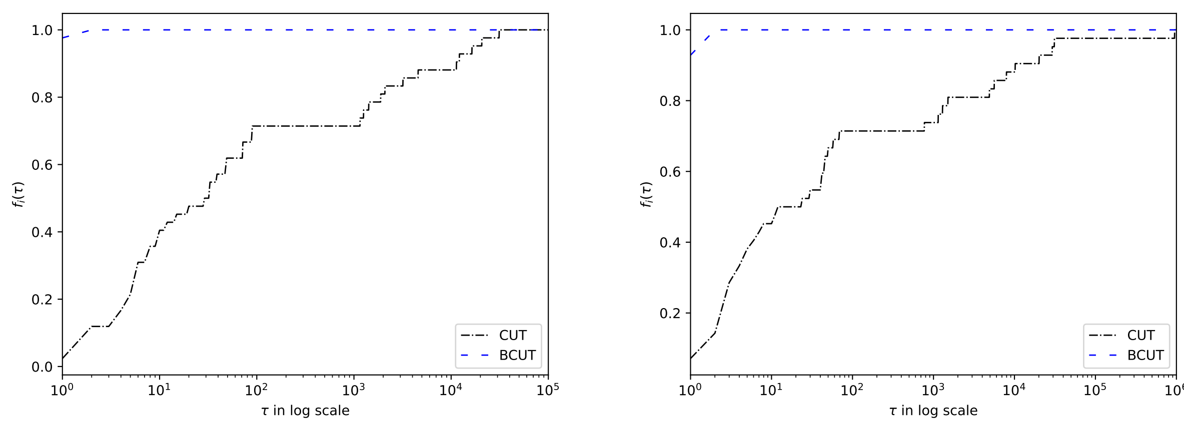


**Figure 6** **Performance profiles of solvers CUT and BCUT for the maximum $r$-robust 2-club (left) and $t$-hereditary 2-club (right) problems.**

### 6.2. Assessing the Cut-Like Formulation For the MRCP When $s = 3$

For the case $s = 3$, although no competing formulation is available for the MHCP, the formulation introduced by Almeida and Carvalho (2014) for the MRCP can be compared against the cut-like formulation (2). To isolate the effect of the formulation used in the recursive block decomposition algorithm, we change the exact approach used in line 11 of Algorithm 1, and compare running time performance between the AC formulation (10) in Appendix C and the delayed constraint generation approach from Section 5.4. These results are reported in Table 16 of Appendix E.

The performance profiles in Figure 7 based on the wall-clock running times of these solvers show that the recursive block decomposition algorithm (with the same heuristic and preprocessing for both solvers) performs better when the cut-like formulation is used, compared to AC formulation (10). For example, BCUT solves the instance `hep-th` for $r = 2$ to optimality in 16.8 seconds, while AC fails to solve this instance under the time limit. For some challenging instances, the optimality gap using BCUT is smaller when both solvers fail to solve to optimality.

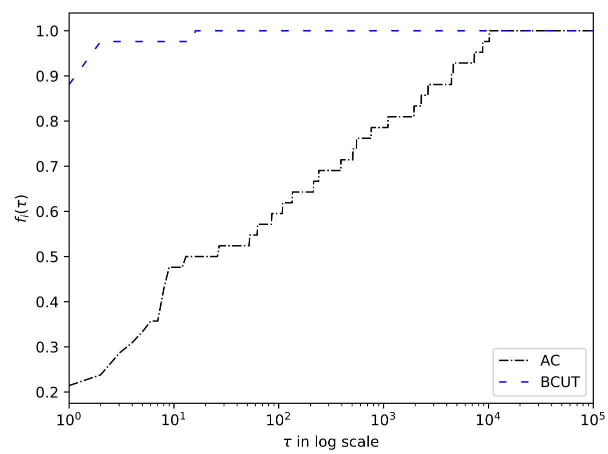


**Figure 7 Performance profiles of solvers AC and BCUT for the maximum $r$-robust 3-club problem.**

### 6.3. Assessing the Impact of Fault-Tolerance on Solution Size

At this time, no numerical results for the MRCP and the MHCP are available in the literature for $s \geq 3$ with the exception of the AC formulation for the MRCP when $s = 3$, which we compared against the cut-like formulation in the foregoing section. In Table 17 of Appendix E, we report results obtained using only the BCUT solver for the MRCP and the MHCP for $r, t \in \{2, 3, 4\}$ and $s \in \{3, 4\}$. All instances in the test bed except the graph `email` for $s = r = t = 3$ were solved to optimality.

Moreover, to understand the impact of the fault-tolerance requirements of the MRCP and the MHCP on the solution, we compare the size of the largest $t$-hereditary $s$-club, $r$-robust $s$-club, and the "relaxed" $r$-robust $s$-club—requiring $r$ *distinct* paths of length at most $s$ for all vertex pairs—considered by Veremyev and Boginski (2012a) as found on

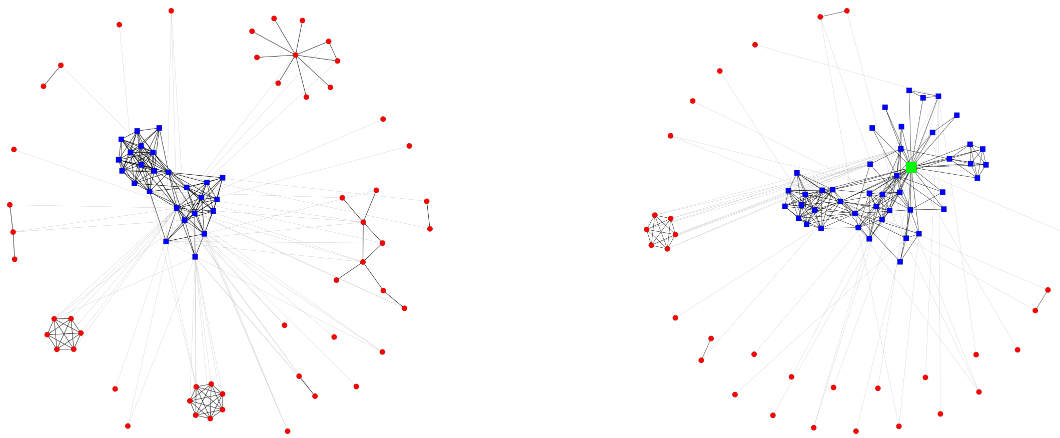

**Figure 8** **In the graph `lesmis` pictured on the left, square dots (blue vertices) identify the maximum 3-robust (also 3-hereditary) 3-club containing 25 vertices that we found. In the same graph visualized on the right, square dots identify the maximum 3-distinct-path 3-club containing 45 vertices that we found. The larger square dot (green vertex) is a cut-vertex in the subgraph induced by the solution. Images were generated using the igraph package (Csardi and Nepusz 2006).**

our test bed. The distinct-path relaxation of the MRCP is found by directly solving the formulation of Veremyev and Boginski (2012a).

The best objective values of the MRCP and the MHCP are the same for most instances, and when they are different, the $t$-hereditary $s$-club is larger than the $r$-robust $s$-club (for $r = t$), consistent with Lemma 1. However, it can also be seen from Table 17 that the $r$-distinct-path $s$-clubs found are on average 60.64% larger than the $r$-robust $s$-clubs found, and on average 19.71% smaller than the maximum $s$-club size. These observations indicate that requiring $r$ distinct length-$s$ paths between vertices in $s$-clubs may not be sufficient if we seek fault-tolerant $s$-clubs. As an example, Figure 8 illustrates the maximum 3-robust (and 3-hereditary) 3-club found by our solver (the same 25 vertices are optimal to the MRCP and the MHCP) and the optimal solution for the 3-distinct-path relaxation (which contains 45 vertices) in the instance `lesmis`. It is evident from the figure (and verified computationally) that the 3-distinct-path 3-club that was found can be disconnected by targeted vertex deletion, specifically the central vertex highlighted in the figure.

## 7. Conclusion

The $r$-robust $s$-club and $t$-hereditary $s$-club models formalize the notion of fault-tolerance in $s$-clubs, a desirable property when seeking reliable low-diameter clusters. In this article we establish the NP-hardness of the associated optimization problems on arbitrary and restricted graph classes for integer constants $t, r, s \geq 2$. Furthermore, we show that it is

NP-complete to verify if a vertex subset is an $r$-robust $s$-club when $r \geq 2$ is fixed and $s$ is a part of the input, and so is its counterpart for fixed $s \geq 5$ and $r$ part of the input. We also show that it is coNP-complete to verify if a vertex subset is a $t$-hereditary $s$-club when $s \geq 5$ is fixed and $t$ is a part of the input.

We propose cut-like formulations for the MRCP for $s \in \{2, 3, 4\}$ and the MHCP for every integer $s \geq 2$ based on length-bounded vertex separators. This is the first IP formulation of the maximum $r$-robust 4-club problem to appear in the literature. For $s \in \{2, 3, 4\}$, we establish the polynomial-time solvability of the associated separation problem. For each $s \geq 5$, we show that it is coNP-complete to determine whether a given solution satisfies all length-bounded vertex separator inequalities used to formulate the MHCP.

We introduce a graph decomposition approach based on finding maximal biconnected components (blocks) that enables us to solve the IP on several smaller subgraphs of the input graph. This block decomposition algorithm is recursive and incorporates preprocessing techniques based on a heuristic solution that aims to further reduce the size of the subgraph on which the IP is solved. We also propose lower and upper bounds on the number of length-bounded vertex-disjoint paths between any given a pair of vertices, which enables us to avoid the exact computation of this quantity used very frequently in several steps of the overall algorithm. We devise a decomposition BC algorithm to solve the cut-like IP formulations of the MRCP and the MHCP when $s \in \{2, 3, 4\}$. The computational gains were empirically evaluated on a test bed of real-life instances from the Tenth DIMACS Implementation Challenge. Our computational studies include the first reported numerical results for the MRCP and the MHCP when $s \in \{3, 4\}$.

This line of research could be continued by targeting the cases involving $s \geq 5$, where the length-bounded counterpart of Menger's Theorem no longer applies. Our results concerning the unlikelihood of compact IP formulations of the MRCP and the MHCP for $s \geq 5$ can be informative in that regard. The foundations laid in this article, such as the recursive block decomposition algorithm, can be helpful to any exact approach to these problems.

## Acknowledgments

The computing for this project was performed at the High Performance Computing Center at Oklahoma State University supported in part through the National Science Foundation grant OAC-1531128. This material is based upon work supported by the National Science Foundation under Grant Nos. 1662757 and 1942065. We would like to thank Alexander Veremyev for sharing his code for graph visualization.

## References

Alba RD (1973) A graph-theoretic definition of a sociometric clique. *Journal of Mathematical Sociology* 3(1):113–126.

Almeida MT, Carvalho FD (2014) An analytical comparison of the LP relaxations of integer models for the $k$-club problem. *European Journal of Operational Research* 232(3):489–498.

Asahiro Y, Doi Y, Miyano E, Samizo K, Shimizu H (2018) Optimal approximation algorithms for maximum distance-bounded subgraph problems. *Algorithmica* 80(6):1834–1856.

Asahiro Y, Miyano E, Samizo K (2010) Approximating maximum diameter-bounded subgraphs. López-Ortiz A, ed., *LATIN 2010: Theoretical Informatics*, 615–626 (Berlin, Heidelberg: Springer Berlin Heidelberg).

Bader DA, Meyerhenke H, Sanders P, Wagner D, eds. (2013) *Graph Partitioning and Graph Clustering: Tenth* DIMACS *Implementation Challenge Workshop*, volume 588 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* (Providence, RI: American Mathematical Society), URL `https://www.cc.gatech.edu/dimacs10/index.shtml`.

Balasundaram B, Butenko S, Trukhanov S (2005) Novel approaches for analyzing biological networks. *Journal of Combinatorial Optimization* 10(1):23–39.

Batagelj V, Zaversnik M (2003) An $O(m)$ algorithm for cores decomposition of networks.

Bourjolly JM, Laporte G, Pesant G (2000) Heuristics for finding $k$-clubs in an undirected graph. *Computers & Operations Research* 27:559–569.

Bourjolly JM, Laporte G, Pesant G (2002) An exact algorithm for the maximum $k$-club problem in an undirected graph. *European Journal of Operational Research* 138:21–28.

Carmesin J, Diestel R, Hamann M, Hundertmark F (2014) $k$-blocks: a connectivity invariant for graphs. *SIAM Journal on Discrete Mathematics* 28(4):1876–1891.

Carr RD, Lancia G (2002) Compact vs. exponential-size LP relaxations. *Operations Research Letters* 30(1):57–65.

Carvajal R, Constantino M, Goycoolea M, Vielma JP, Weintraub A (2013) Imposing connectivity constraints in forest planning models. *Operations Research* 61(4):824–836.

Cook DJ, Holder LB (2006) *Mining Graph Data* (John Wiley & Sons).

Csardi G, Nepusz T (2006) The igraph software package for complex network research. *InterJournal* Complex Systems:1695, URL `http://igraph.org`.

Dagum L, Enon R (1998) OpenMP: An industry standard API for shared-memory programming. *IEEE Computational Science & Engineering* 5(1):46–55.

Dolan ED, Moré JJ (2002) Benchmarking optimization software with performance profiles. *Mathematical Programming* 91(2):201–213.

Ford Jr L, Fulkerson DR (1956) Maximal flow through a network. *Canadian Journal of Mathematics* 8:399–404.

Golovach PA, Heggernes P, Kratsch D, Rafiey A (2014) Finding clubs in graph classes. *Discrete Applied Mathematics* 174:57–65.

Golovach PA, Thilikos DM (2011) Paths of bounded length and their cuts: Parameterized complexity and algorithms. *Discrete Optimization* 8(1):72–86.

Gould N, Scott J (2016) A note on performance profiles for benchmarking software. *ACM Transactions on Mathematical Software (TOMS)* 43(2):1–5.

Grötschel M, Lovász L, Schrijver A (1993) *Geometric Algorithms and Combinatorial Optimization* (Berlin: Springer-Verlag), 2nd edition.

Gurobi Optimization L (2020) Gurobi optimizer reference manual. URL `http://www.gurobi.com`.

Hartung S, Komusiewicz C, Nichterlein A, Suchý O (2015) On structural parameterizations for the 2-club problem. *Discrete Applied Mathematics* 185:79–92.

Hopcroft J, Tarjan R (1973) Algorithm 447: Efficient algorithms for graph manipulation. *Communications of the ACM* 16(6):372–378.

Itai A, Perl Y, Shiloach Y (1982) The complexity of finding maximum disjoint paths with length constraints. *Networks* 12(3):277–286.

Komusiewicz C, Nichterlein A, Niedermeier R, Picker M (2019) Exact algorithms for finding well-connected 2-clubs in sparse real-world graphs: Theory and experiments. *European Journal of Operational Research* 275(3):846–864.

Langguth J, Manne F, Sanders P (2010) Heuristic initialization for bipartite matching problems. *Journal of Experimental Algorithmics (JEA)* 15:1–1.

Lawler E (1976) *Combinatorial Optimization: Networks and Matroids* (New York: Holt, Rinehart, and Winston).

Lewis JM, Yannakakis M (1980) The node-deletion problem for hereditary properties is NP-complete. *Journal of Computer and System Sciences* 20(2):219 – 230.

Li CL, McCormick ST, Simchi-Levi D (1990) The complexity of finding two disjoint paths with min-max objective function. *Discrete Applied Mathematics* 26(1):105–115.

Lovász L, Neumann-Lara V, Plummer M (1978) Mengerian theorems for paths of bounded length. *Periodica Mathematica Hungarica* 9(4):269–276.

Lu Y, Moradi E, Balasundaram B (2018) Correction to: Finding a maximum $k$-club using the $k$-clique formulation and canonical hypercube cuts. *Optimization Letters* 12(8):1959–1969.

Luce RD (1950) Connectivity and generalized cliques in sociometric group structure. *Psychometrika* 15(2):169–190.

Magun J (1998) Greedy matching algorithms, an experimental study. *ACM Journal of Experimental Algorithmics* 3:6–es.

Martin RK (1991) Using separation algorithms to generate mixed integer model reformulations. *Operations Research Letters* 10(3):119–128.

Matula DW (1978) $k$-blocks and ultrablocks in graphs. *Journal of Combinatorial Theory, Series B* 24(1):1–13.

Matula DW, Beck LL (1983) Smallest-last ordering and clustering and graph coloring algorithms. *Journal of the ACM (JACM)* 30(3):417–427.

Menger K (1927) Zur allgemeinen kurventheorie. *Fundamenta Mathematicae* 10:95–115.

Möhring R, Müller-Hannemann M (1995) Cardinality matching: Heuristic search for augmenting paths. Technical Report 439, Technische Universität Berlin.

Mokken RJ (1979) Cliques, clubs and clans. *Quality and Quantity* 13(2):161–173.

Moradi E, Balasundaram B (2018) Finding a maximum $k$-club using the $k$-clique formulation and canonical hypercube cuts. *Optimization Letters* 12(8):1947–1957.

Pajouh FM, Balasundaram B (2012) On inclusionwise maximal and maximum cardinality $k$-clubs in graphs. *Discrete Optimization* 9(2):84–97.

Pattillo J, Youssef N, Butenko S (2013) On clique relaxation models in network analysis. *European Journal of Operational Research* 226(1):9–18.

Rose D, Tarjan R, Lueker G (1976) Algorithmic aspects of vertex elimination on graphs. *SIAM Journal on Computing* 5(2):266–283.

Salemi H, Buchanan A (2020) Parsimonious formulations for low-diameter clusters. *Mathematical Programming Computation* 12(3):493–528.

Schäfer A (2009) *Exact algorithms for s-club finding and related problems*. Master's thesis, Diplomarbeit, Institut für Informatik, Friedrich-Schiller-Universität Jena.

Schäfer A, Komusiewicz C, Moser H, Niedermeier R (2012) Parameterized computational complexity of finding small-diameter subgraphs. *Optimization Letters* 6(5):883–891.

Validi H, Buchanan A (2020) The optimal design of low-latency virtual backbones. *INFORMS Journal on Computing* 32(4):952–967.

Veremyev A, Boginski V (2012a) Identifying large robust network clusters via new compact formulations of maximum $k$-club problems. *European Journal of Operational Research* 218(2):316–326.

Veremyev A, Boginski V (2012b) Robustness and strong attack tolerance of low-diameter networks. Sorokin A, Murphey R, Thai MT, Pardalos PM, eds., *Dynamics of Information Systems: Mathematical Foundations*, volume 20 of *Springer Proceedings in Mathematics & Statistics*, 137–156 (Springer New York).

Walteros JL, Buchanan A (2020) Why is maximum clique often easy in practice? *Operations Research* 68(6):1866–1895.

Wang Y, Buchanan A, Butenko S (2017) On imposing connectivity constraints in integer programs. *Mathematical Programming* 166(1):241–271.

Yezerska O, Pajouh FM, Butenko S (2017) On biconnected and fragile subgraphs of low diameter. *European Journal of Operational Research* 263(2):390–400.

## Appendix A: Proofs of Complexity Results: Optimization

The additional notations we use in this section are as follows. The complete graph on $n$ vertices is denoted by $K_n$. Given two disjoint graphs $H$ and $G$, the graph join operation, denoted as $H * G$, produces the graph that "combines" $G$ and $H$ by joining every vertex from $G$ and every vertex from $H$ using a new edge. Let us use $V(G)$ and $E(G)$ to denote the vertex and edge sets of $G$, respectively. Formally, $V(H*G) = V(H) \cup V(G)$ and $E(H*G) = E(H) \cup E(G) \cup \{uv \mid u \in V(H), v \in V(G)\}$. If two graphs $G$ and $H$ are isomorphic, we denote that by $G \simeq H$.

*Proof of Theorem 1.* We show a polynomial-time reduction from $s$-Club. Given an instance $\langle G, c\rangle$ of $s$-Club, construct the instance $\langle G', c'\rangle$ of $t$-Hereditary $s$-Club as follows: let $G' := G * K_{t-1}$ and $c' := c + t - 1$. Suppose $S \subseteq V(G)$ is an $s$-club in $G$ of size at least $c$. We claim that $S \cup V(K_{t-1})$ is a $t$-Hereditary $s$-Club in $G'$ (of size at least $c'$). Consider a deletion set $T \subseteq S \cup V(K_{t-1})$ such that $|T| \le t-1$ and a pair of vertices $u, v \in S \cup V(K_{t-1}) \setminus T$. If $T = V(K_{t-1})$, then $G'[S \cup V(K_{t-1}) \setminus T] \simeq G[S]$, is an $s$-club. Otherwise, $S \cup V(K_{t-1}) \setminus T$, which contains some vertex from $K_{t-1}$ that dominates $G'[S \cup V(K_{t-1}) \setminus T]$, is a 2-club. Conversely, suppose $S' \subseteq V(G')$ is a $t$-Hereditary $s$-club in $G'$ of size at least $c'$. Then, $S' \setminus V(K_{t-1})$ of size at least $c$ is an $s$-club in $G'$ by Definition 4. Since $G'[S' \setminus V(K_{t-1})] \simeq G[S' \setminus V(K_{t-1})]$, it is an $s$-club in $G$ as well. Finally, $t$-Hereditary $s$-Club belongs to class NP as $t$ is a fixed constant and verification can be completed in polynomial time by enumerating all possible deletion sets of size at most $t-1$. For the final part, observe that every vertex in $V(K_{t-1}) \neq \emptyset$ dominates $G'$. □

*Proof of Theorem 2.* We show a polynomial-time reduction from $s$-Club. Given an instance $\langle G, c\rangle$ of $s$-Club, construct the instance $\langle G', c'\rangle$ of $r$-Robust $s$-Club where $G' := G * K_{r-1}$ and $c' := c + r - 1$. Suppose $S \subseteq V(G)$ is an $s$-club in $G$ of size at least $c$. We claim that $S \cup V(K_{r-1})$ is an $r$-robust $s$-club of size at least $c + r - 1$ in $G'$ for the nontrivial case when $c \ge 2$.

For any two vertices $u, v \in S$, there exists a $u,v$-path of length at most $s$ between them in $G[S]$, and there exist $r-1$ $u,v$-paths of length at most two through their common neighbors in $V(K_{r-1})$; these constitute $r$ internally vertex-disjoint paths of length at most $s$ between $u$ and $v$ in $G'[S \cup V(K_{r-1})]$.

For any two vertices $u, v \in V(K_{r-1})$ (if $r \ge 3$), given that $c \ge 2$ there are at least two paths of length two via their common neighbors in $S$, and $r-3$ $u,v$-paths of length two via their common neighbors in $V(K_{r-1})$. These paths, along with the edge $\{u,v\} \in E(G')$ constitute at least $r$ internally vertex-disjoint paths of length at most $s$ between $u$ and $v$ in $G'[S \cup V(K_{r-1})]$.

Finally, consider $u \in S$ and $v \in V(K_{r-1})$. There are $r-2$ $u,v$-paths of length two in $G'$ via vertices in $V(K_{r-1}) \setminus \{v\}$. Because $S$ is an $s$-club in $G$ containing at least two vertices, there exists a vertex $w \in S \setminus \{u\}$ that is a common neighbor of $u$ and $v$ in $G'$. Since $\{u,v\} \in E(G')$, there are $r$ vertex disjoint paths between every pair of vertices in $G'[S \cup V(K_{r-1})]$.

Conversely, suppose $S' \subseteq V(G')$ is an $r$-robust $s$-club of size at least $c + r - 1$ in $G'$. Since $|S' \cap V(K_{r-1})| \le r-1$, after deleting all vertices in $S' \cap V(K_{r-1})$, a path of length at most $s$ still exists between any two vertices in subgraph $G'[S' \setminus V(K_{r-1})]$. Hence, $S' \setminus V(K_{r-1})$ is an $s$-club of size at least $c$ in $G$, since $G'[S' \setminus V(K_{r-1})] \simeq G[S' \setminus V(K_{r-1})]$. Hence, $r$-Robust $s$-Club is NP-hard.

Golovach and Thilikos (2011) showed that verifying whether or not a graph with $n$ vertices and $m$ edges contains $r$ vertex-disjoint $(u,v)$-paths of length at most $s$ between distinct vertices $u$ and $v$ can be answered in $O(2^{O(rs)} m \log n)$ time. Using their algorithm we can verify if $S' \subseteq V'$ is an $r$-robust $s$-club in $G'$ in polynomial time for constant $r$ and $s$. Hence, $r$-Robust $s$-Club belongs to class NP. For the final part, observe that every vertex in $V(K_{r-1}) \neq \emptyset$ dominates $G'$. □

*Proof of Corollary 1.* It suffices to show that $H := G * K_\ell$ is $k$-chordal if $G$ is $k$-chordal for $\ell \geq 1$. Suppose that $G$ is $k$-chordal and $C \subseteq V(H)$ is a cycle of length strictly greater than $k \geq 3$. If $C$ contains no vertices from $K_\ell$, then $C$ contains a chord between two vertices in $V(G)$ since $G$ is $k$-chordal by assumption. If it contains at least one vertex from $K_\ell$, then that vertex is adjacent to every other vertex in the $C$, and at least one such adjacent vertex in $C$ creates a chord in $C$. □

*Proof of Corollary 2.* This follows from the result of Golovach et al. (2014) who proved that 2-Club is NP-hard on graphs with clique cover number three, and our constructions. □

## Appendix B: Proofs of Complexity Results: Verification

First we state a complexity result of Validi and Buchanan (2020). The construction technique used in their polynomial-time reduction can also be used in establishing several of the complexity results in this section.

**Problem:** Diameter-$s$ Interdiction by Node Deletion (positive integer $s$).

**Question:** Given a graph $G = (V, E)$ and constant $q$, is there a subset $C \subseteq V$ of $q$ vertices such that $\text{diam}(G - C) \geq s + 1$?

Note that $G - C := G[V \setminus C]$, i.e., $G - C$ is the graph obtained by deleting all vertices in $C$ and incident edges.

Proposition 10 **(Validi and Buchanan (2020))**. Diameter-$s$ Interdiction by Node Deletion *is NP-complete for every fixed integer $s \geq 5$.*

Itai et al. (1982) established the following hardness result that serves as the source problem to show that the verification problem for $r$-robust $s$-clubs is NP-hard for arbitrary $r$ and fixed integer $s \geq 5$.

**Problem:** $r$-Vertex-Disjoint $s$-Paths (positive integers $s, r$).

**Question:** Given a graph $G = (V, E)$ and a pair of vertices $a, b \in V$, does $G$ contain at least $r$ vertex disjoint $a,b$-paths of length at most $s$?

Proposition 11 **(Itai et al. (1982))**. $r$-Vertex-Disjoint $s$-Paths *is NP-complete for every fixed integer $s \geq 5$ and arbitrary positive integer $r$.*

*Proof sketch for Theorem 3.* Given an instance $\langle G = (V, E), a, b, r\rangle$ of $r$-Vertex-Disjoint $s$-Paths, we can assume without loss of generality that $\text{dist}_G(a, v) \leq s - 1$ and $\text{dist}_G(b, v) \leq s - 1$ for every vertex $v \in V \setminus \{a, b\}$, because only such vertices can be internal to any $a$-$b$ path of length at most $s$. We can further assume that $a$ and $b$ are not adjacent as we can decrease the value of $r$ by one and delete the edge $ab$ from $G$ if that is not the case. For the polynomial time reduction from such an instance $\langle G = (V, E), a, b, r\rangle$ of $r$-Vertex-Disjoint $s$-Paths to an instance $\langle G' = (V', E'), S', r\rangle$ of Is $r$-Robust $s$-Club, we can follow the same approach used by Validi and Buchanan (2020) to prove Proposition 10.

The proof of Validi and Buchanan (2020) uses a different construction depending on the parity of $s$ to obtain an instance of Diameter-$s$ Interdiction by Node Deletion, namely a graph $G' = (V', E')$ and an integer $q$, from a given graph $G = (V, E)$ containing distinct nonadjacent vertices $a$ and $b$. We can employ the same technique to construct an instance $\langle G' = (V', E'), S', r\rangle$ of Is $r$-Robust $s$-Club by using $r$-vertex complete graphs in place of complete graphs with $q+1$ vertices in their construction and setting $S' = V'$. Under this construction $\langle G = (V, E), a, b, r\rangle$ is a "yes" instance for $r$-Vertex-Disjoint $s$-Paths if and only if $\langle G', S', r\rangle$ is a "yes" instance for Is $r$-Robust $s$-Club. We refer the reader to Validi and Buchanan (2020) for more details regarding this construction.

This problem belongs to NP since a polynomially verifiable witness is a collection of $r$ (internally) vertex-disjoint paths of length at most $s$ between every pair of vertices. □

**Problem**: $r$-Robust Vertex-Disjoint $s$-Paths (positive integers $r, s$)

**Question:** Given a graph $G = (V, E)$ and a pair of vertices $a, b \in V$, does $G$ contain at least $r$ vertex-disjoint $a, b$-paths of length at most $s$?

Li et al. (1990) established the NP-completeness of $r$-Robust Vertex-Disjoint $s$-Paths for fixed $r = 2$ and arbitrary $s$. This result can be used to show that the problem is NP-complete for every fixed integer $r \geq 3$ using a simple reduction connecting vertices $a$ and $b$ using $r-2$ additional vertex disjoint paths of length 2. We can then show that Is $r$-Robust $s$-Club is NP-hard when $r \geq 2$ is a fixed positive integer and $s$ is arbitrary, using a polynomial-time reduction from $r$-Robust Vertex-Disjoint $s$-Paths.

*Proof sketch for Theorem 4.* The same arguments as Theorem 3, using $r$-Robust Vertex-Disjoint $s$-Paths as the source problem. □

*Proof of Theorem 5.* Membership in coNP is clear, because a deletion set of size less than $t$ that increases the diameter higher than $s$ is a certificate for "no" instances of Is $t$-Hereditary $s$-Club. Hardness follows from the observation that every "no" instance of Is $t$-Hereditary $s$-Club is a "yes" instance of Diameter-$s$ Interdiction by Node Deletion with input $G[S]$ and constant $c := t-1$; and vice versa. □

## Appendix C: Proofs of MRCP Formulation Strength

When $s = 2$, the set of common neighbors $N(u) \cap N(v)$ form the unique minimal length-2 $u, v$-separator in $G - uv$. Therefore, constraints (2b) reduce to the following:

$$(r - \mathbb{1}_E(u, v))(x_u + x_v - 1) \leq \sum_{i \in N(u) \cap N(v)} x_i \qquad \forall uv \in \binom{V}{2}. \tag{8}$$

The only existing formulation of the maximum $r$-robust 2-club problem is the one below that was proposed by Veremyev and Boginski (2012a).

$$\textbf{(VB)} \quad \max \sum_{i \in V} x_i \tag{9a}$$

$$\text{s.t. } r(x_u + x_v - 1) \leq \mathbb{1}_E(u, v) + \sum_{i \in N(u) \cap N(v)} x_i \qquad \forall uv \in \binom{V}{2} \tag{9b}$$

$$x_i \in \{0, 1\} \qquad \forall i \in V. \tag{9c}$$

*Proof of Proposition 4.* The claim that the linear programming (LP) relaxation of our cut-like formulation is contained in the LP relaxation of formulation VB, follows immediately from the fact that $\mathbb{1}_E(u, v)(x_u + x_v - 1) \leq \mathbb{1}_E(u, v)$ $\forall x \in [0, 1]^{|V|}$. Figure 9 shows that the inclusion is strict when $r \geq 2$. □

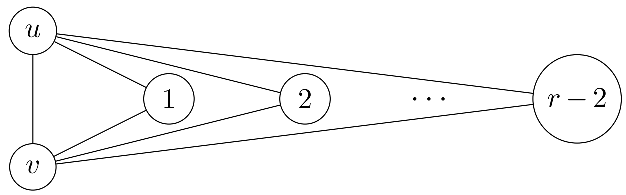


**Figure 9** **When $r \geq 2$, the point $\bar{x}$ constructed as described next is feasible to the LP relaxation of the VB formulation, but not to the cut-like formulation; $\bar{x}$ is obtained by setting $\bar{x}_u = \frac{5}{8}, \bar{x}_v = \frac{5}{8}$, and $\bar{x}_i = \frac{1}{4}, \forall i \in [r-2]$. Note that when $r = 2$, we take $N(u) \cap N(v)$ to be empty.**

Next, we consider the formulation of the maximum $r$-robust 3-club problem proposed by Almeida and Carvalho (2014).

$$\textbf{(AC)} \quad \max \sum_{i \in V} x_i \tag{10a}$$

$$\text{s.t.} \quad x_a + x_b \leq 1 \qquad \forall ab \in \binom{V}{2} : \text{dist}_G(a,b) \geq 4 \tag{10b}$$

$$(r - \mathbb{1}_E(a,b))(x_a + x_b - 1) \leq \sum_{\ell \in N(a) \cap N(b)} x_\ell + \sum_{pq \in E^{ab}} y_{pq}^{ab} \qquad \forall ab \in \binom{V}{2} : \text{dist}_G(a,b) \leq 3 \tag{10c}$$

$$y_{pq}^{ab} \leq \min\{x_a, x_b\} \qquad \forall ab \in \binom{V}{2} : \text{dist}_G(a,b) \leq 3, \forall pq \in E^{ab} \tag{10d}$$

$$\sum_{q: pq \in E^{ab}} y_{pq}^{ab} \leq x_p \qquad \forall ab \in \binom{V}{2} : \text{dist}_G(a,b) \leq 3, \forall p \in V^{ab} \setminus N[b] \tag{10e}$$

$$\sum_{p: pq \in E^{ab}} y_{pq}^{ab} \leq x_q \qquad \forall ab \in \binom{V}{2} : \text{dist}_G(a,b) \leq 3, \forall q \in V^{ab} \setminus N[a] \tag{10f}$$

$$x_i \in \{0,1\} \qquad \forall i \in V \tag{10g}$$

$$y_{pq}^{ab} \in \{0,1\} \qquad \forall ab \in \binom{V}{2} : \text{dist}_G(a,b) \leq 3, \forall pq \in E^{ab}. \tag{10h}$$

In formulation (10), for each pair of vertices $\{a,b\} \in \binom{V}{2}$, we let

$$E^{ab} := \left\{ pq \in E \;\middle|\; p \in N(a) \setminus N[b], q \in N(b) \setminus N[a] \right\},$$

represent the set of inner edges of length-3 paths that connect vertices $a$ and $b$. Vertex set $V^{ab}$ denotes the set of endpoints of edges in $E^{ab}$. Each variable $y_{pq}^{ab}$ is associated with an edge in $E^{ab}$ for every distinct pair of vertices $a$ and $b$.

*Proof of Proposition 5.* The point $(\hat{x}, \hat{y})$ given below is feasible to the LP relaxation of the AC formulation (10) for the graph $G$ in Figure 10, but $\hat{x}$ is not feasible to the LP relaxation of the cut-like formulation (2) strengthened by inequalities (3) when $s = 3$ and $r = 2$.

Let $\hat{x}_6 = 0.5, \hat{x}_7 = 0.75$, and $\hat{x}_i = 1, \forall i \in \{1,2,3,4,5\}$. Note that we only need to consider $\hat{y}_{pq}^{ab}$ for a pair of vertices $a$ and $b$ for which $E^{ab}$ is not empty. We set $\hat{y}_{1,2}^{5,6} = \hat{y}_{2,3}^{5,6} = \hat{y}_{3,4}^{5,6} = \hat{y}_{3,4}^{6,7} = 0.5$, $\hat{y}_{2,5}^{1,7} = \hat{y}_{3,4}^{1,7} = \hat{y}_{3,4}^{2,7} = \hat{y}_{2,5}^{3,7} = 0.75$, and $\hat{y}_{2,5}^{1,4} = \hat{y}_{3,4}^{1,5} = \hat{y}_{4,5}^{2,3} = \hat{y}_{3,4}^{2,5} = \hat{y}_{2,5}^{3,4} = \hat{y}_{2,3}^{4,5} = 1$. We can now verify by direct substitution that the point $(\hat{x}, \hat{y})$ is feasible to the LP relaxation of the AC formulation (10). However, because $\rho_3(G; 6, 7) = 1$, the conflict inequality (3) for the pair $\{6,7\}$ is violated by $\hat{x}$ as $\hat{x}_6 + \hat{x}_7 = 1.25 \nleq 1$. Therefore, $\hat{x}$ is not feasible to the LP relaxation of the cut-like formulation (2) strengthened by conflict inequalities (3).

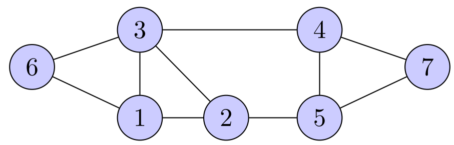

**Figure 10** **When $s=3$ and $r=2$, there exists a $\hat{y}$ such that the point $\hat{x}$ obtained by setting $\hat{x}_6=0.5, \hat{x}_7=0.75$ and $\hat{x}_i=1, \forall i \in \{1,2,3,4,5\}$ is feasible to the AC formulation but $\hat{x}$ is not feasible to the LP relaxation of the cut-like formulation (2) strengthened by conflict inequalities (3).**

Next we show that the point $\bar{x}$ given below is feasible to the LP relaxation of cut-like formulation (2) strengthened by inequalities (3) for the graph in Figure 11, but no $\bar{y}$ exists such that $(\bar{x}, \bar{y})$ is feasible to the LP relaxation of the AC formulation (10) when $s=3$ and $r=2$. The point $\bar{x}$ is obtained by setting $\bar{x}_1=0.2, \bar{x}_i=0.8, \forall i \in \{2,3,4,8\}$, and $\bar{x}_i=1, \forall i \in \{5,6,7\}$ (see Figure 11).

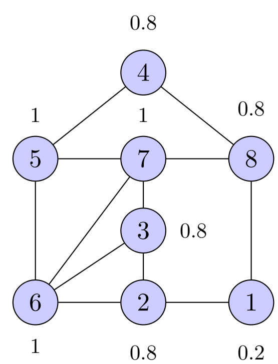

**Figure 11** **When $s=3$ and $r=2$, the point $\bar{x}$ obtained by setting $\bar{x}_1=0.2, \bar{x}_i=0.8, \forall i \in \{2,3,4,8\}$, and $\bar{x}_i=1, \forall i \in \{5,6,7\}$ is feasible to the LP relaxation of the cut-like formulation strengthened by conflict inequalities (3), but no $\bar{y}$ exists such that $(\bar{x}, \bar{y})$ is feasible to the LP relaxation of the AC formulation.**

We can verify that $\bar{x}$ satisfies constraints (2b) using direct substitution by enumerating every pair of distinct vertices $a$ and $b$. This step is made easier by using the following facts. We only need to consider constraints with a positive left-hand side. Every length-3 $a,b$-separator must include all the common neighbors of $a$ and $b$. For the small number of cases where the total variable value of the common neighbors does not exceed the left-hand side, that is

$$(r - \mathbb{1}_E(a,b))(\bar{x}_a + \bar{x}_b - 1) > \sum_{i \in N(a) \cap N(b)} \bar{x}_i,$$

we can add the minimum vertex value on each path from the two vertex disjoint paths of length three between $a$ and $b$ listed in Table 4, in order to calculate a larger lower bound on the right-hand side of constraint (2b) that exceeds the left-hand side.

For the $a$-$b$-pairs $\{1,2\}, \{1,3\}, \{1,4\}$, and $\{1,8\}$, conflict constraints (3) are trivially satisfied because $\bar{x}_a + \bar{x}_b = 1$. For the remaining pairs of distinct vertices $a$ and $b$, at least two vertex-disjoint paths of length at most 3 exist (shown in Table 4) and hence, there are no constraints of type (3) associated with them.

To see that no $\bar{y}$ exists such that $(\bar{x}, \bar{y})$ is feasible to the LP relaxation of the AC formulation (10), we consider the pair of vertices $\{2,4\}$. The edge set $E^{2,4} = \{\{1,8\},\{5,6\}\}$. From constraint (10c) we have

$\bar{y}^{2,4}_{1,8} + \bar{y}^{2,4}_{5,6} \geq 1.2$ and from constraints (10e) and (10f) we have $\bar{y}^{2,4}_{1,8} \leq 0.2$ and $\bar{y}^{2,4}_{5,6} \leq 1$. Taken together, $\bar{y}^{2,4}_{1,8} = 0.2$ and $\bar{y}^{2,4}_{5,6} = 1$. Then, constraint (10d) for this pair is violated as $\bar{y}^{2,4}_{5,6} = 1 \not\leq \min\{\bar{x}_2, \bar{x}_4\} = 0.8$. □

**Table 4** **Two vertex-disjoint paths of length at most 3 between selected vertex-pairs in the graph in Figure 11.**

| Pairs | Paths | Pairs | Paths | Pairs | Paths | Pairs | Paths | Pairs | Paths | Pairs | Paths |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1, 5 | 1–2–6–5 | 2, 4 | 2–6–5–4 | 2, 8 | 2–1–8 | 3, 7 | 3–7 | 4, 7 | 4–5–7 | 5, 8 | 5–7–8 |
| | 1–8–7–5 | | 2–1–8–4 | | 2–6–7–8 | | 3–6–7 | | 4–8–7 | | 5–4–8 |
| 1, 6 | 1–2–6 | 2, 5 | 2–6–5 | 3, 4 | 3–6–5–4 | 3, 8 | 3–7–8 | 4, 8 | 4–8 | 6, 7 | 6–7 |
| | 1–8–7–6 | | 2–3–7–5 | | 3–7–8–4 | | 3–2–1–8 | | 4–5–7–8 | | 6–5–7 |
| 1, 7 | 1–2–3–7 | 2, 6 | 2–6 | 3, 5 | 3–6–5 | 4, 5 | 4–5 | 5, 6 | 5–6 | 6, 8 | 6–7–8 |
| | 1–8–7 | | 2–3–6 | | 3–7–5 | | 4–8–7–5 | | 5–7–6 | | 6–5–4–8 |
| 2, 3 | 2–3 | 2, 7 | 2–3–7 | 3, 6 | 3–6 | 4, 6 | 4–5–6 | 5, 7 | 5–7 | 7, 8 | 7–8 |
| | 2–6–3 | | 2–6–7 | | 3–2–6 | | 4–8–7–6 | | 5–6–7 | | 7–5–4–8 |

Remark 2. Another interesting observation that we made whilst creating the counterexamples used in the arguments above is that the LP relaxation of AC formulation (10) without constraints (10d) has the same projection to the $x$ variable space as the LP relaxation of the cut-like formulation (2) (without the conflict constraints (3)). This claim can be proved using a flow-model constructed using the ideas in (Lovász et al. 1978, Almeida and Carvalho 2014).

## Appendix D: Pseudocodes

**Algorithm 4:** Find $\hat{\rho}_3(G;u,v)$, a valid lower bound on $\rho_s(G;u,v)$ for all $s \geq 3$

**Input:** Graph $G=(V,E)$, pair of distinct vertices $u$ and $v$, and $r$.

**Output:** $\hat{\rho}_3(G;u,v)$.

$C \leftarrow N(u) \cap N(v)$<br>
$\hat{\rho}_3(G;u,v) \leftarrow \mathbb{1}_E(u,v) + |C|$<br>
**if** $\hat{\rho}_3(G;u,v) \geq r$ **then**<br>
 **return** $\hat{\rho}_3(G;u,v)$<br>
visited$[i] \leftarrow$ `false` $\forall i \in V$<br>
is-v-nbr$[i] \leftarrow$ `false` $\forall i \in V$<br>
visited$[i] \leftarrow$ `true` $\forall i \in C \cup \{u,v\}$<br>
is-v-nbr$[i] \leftarrow$ `true` $\forall i \in N(v)$<br>
**for** $p \in N(u)$ **do**<br>
 **if** visited$[p] =$ `false` **then**<br>
  **for** $q \in N(p)$ **do**<br>
   **if** visited$[q] =$ `false` *and* is-v-nbr$[q] =$ `true` **then**<br>
    $\hat{\rho}_3(G;u,v) \leftarrow \hat{\rho}_3(G;u,v) + 1$<br>
    visited$[p] \leftarrow$ `true`<br>
    visited$[q] \leftarrow$ `true`<br>
    **if** $\hat{\rho}_3(G;u,v) \geq r$ **then**<br>
     **return** $\hat{\rho}_3(G;u,v)$<br>
    **break**<br>
**return** $\hat{\rho}_3(G;u,v)$

**Algorithm 5:** A heuristic for finding a $t$-hereditary $s$-club for $s \in \{2,3,4\}$

**Input:** A graph $G=(V,E)$, $t$, and $s$.

**Output:** A $t$-hereditary $s$-club.

```
E^c ← { ij ∈ (V choose 2) | ij ∈ E or ρ_s(G; i, j) ≥ t }
create compatibility graph G^c ← (V, E^c)
S ← a maximal clique in G^c
while S ≠ ∅ do
    τ_i ← 0 ∀i ∈ S
    for {i, j} ∈ (S choose 2) \ E do
        if ρ_s(G[S]; i, j) ≤ t − 1 then
            τ_i ← τ_i + 1
            τ_j ← τ_j + 1
    v ← arg max_{i∈S} τ_i
    if τ_v ≥ 1 then
        S ← S \ {v}
    else
        return S
```

**Algorithm 6:** Vertex peeling based on a $t$-hereditary $s$-club of size $\ell$

**Input:** A graph $G = (V, E)$, $t, s$, and a lower bound $\ell$.

**Output:** Preprocessed graph $G$.

```
repeat
    G ← the t-core of G
    D ← ∅
    for each v ∈ V(G) do
        if |N_G^s(v)| < ℓ or |W_v| + |N(v)| < ℓ then
            D ← D ∪ {v}
    if D ≠ ∅ then
        G ← G − D
until D = ∅
return G
```

## Appendix E: Detailed Numerical Results

**Table 5 The impact of using lower and upper bounds when computing $\rho_s$ on DIMACS-10 instances: the number of vertex-pairs to be considered before ($p$) and after ($p'$) filtering, and the running time needed in seconds before ($t$) and after ($t'$) filtering are reported.**

| | | | | $s=3$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $r=2$ | | | $r=3$ | | | $r=4$ | | |
| Graph | $n$ | $m$ | $p$ | $t$ | $p'$ | $t'$ | $t$ | $p'$ | $t'$ | $t$ | $p'$ | $t'$ |
| karate | 34 | 78 | 561 | 0.00 | 46 | 0.00 | 0.00 | 32 | 0.00 | 0.00 | 15 | 0.00 |
| dolphins | 62 | 159 | 1,891 | 0.00 | 24 | 0.00 | 0.00 | 70 | 0.00 | 0.00 | 78 | 0.00 |
| lesmis | 77 | 254 | 2,926 | 0.00 | 228 | 0.00 | 0.00 | 303 | 0.00 | 0.00 | 231 | 0.00 |
| polbooks | 105 | 441 | 5,460 | 0.01 | 93 | 0.00 | 0.01 | 325 | 0.00 | 0.01 | 396 | 0.00 |
| adjnoun | 112 | 425 | 6,216 | 0.01 | 68 | 0.00 | 0.01 | 236 | 0.00 | 0.01 | 367 | 0.00 |
| football | 115 | 613 | 6,555 | 0.01 | 39 | 0.00 | 0.01 | 178 | 0.00 | 0.01 | 365 | 0.00 |
| jazz | 198 | 2742 | 19,503 | 0.03 | 63 | 0.00 | 0.04 | 324 | 0.00 | 0.05 | 710 | 0.00 |
| celegans | 453 | 2025 | 102,378 | 0.08 | 10,683 | 0.01 | 0.09 | 20,974 | 0.02 | 0.10 | 16,469 | 0.02 |
| email | 1133 | 5451 | 641,278 | 0.34 | 4,252 | 0.05 | 0.41 | 8,404 | 0.06 | 0.40 | 10,124 | 0.06 |
| polblogs | 1490 | 16715 | 1,109,305 | 0.95 | 4,129 | 0.06 | 1.29 | 7,590 | 0.08 | 1.48 | 10,115 | 0.08 |
| netscience | 1589 | 2742 | 1,261,666 | 0.25 | 2,212 | 0.04 | 0.27 | 1,753 | 0.04 | 0.28 | 1,030 | 0.04 |
| power | 4941 | 6594 | 12,204,270 | 2.82 | 4,165 | 0.30 | 3.08 | 2,501 | 0.29 | 3.15 | 1,061 | 0.27 |
| hep-th | 8361 | 15751 | 34,948,980 | 11.15 | 30,822 | 0.71 | 11.05 | 23,429 | 0.70 | 11.37 | 13,471 | 0.74 |
| PGP | 10680 | 24316 | 57,025,860 | 20.43 | 33,796 | 1.19 | 19.43 | 40,177 | 1.09 | 19.46 | 39,155 | 1.51 |

| | | | | $s=4$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $r=2$ | | | $r=3$ | | | $r=4$ | | |
| Graph | $n$ | $m$ | $p$ | $t$ | $p'$ | $t'$ | $t$ | $p'$ | $t'$ | $t$ | $p'$ | $t'$ |
| karate | 34 | 78 | 561 | 0.00 | 161 | 0.00 | 0.00 | 102 | 0.00 | 0.00 | 50 | 0.00 |
| dolphins | 62 | 159 | 1,891 | 0.01 | 333 | 0.00 | 0.01 | 316 | 0.00 | 0.01 | 305 | 0.00 |
| lesmis | 77 | 254 | 2,926 | 0.01 | 721 | 0.00 | 0.01 | 700 | 0.00 | 0.01 | 548 | 0.00 |
| polbooks | 105 | 441 | 5,460 | 0.02 | 1,374 | 0.01 | 0.02 | 1,729 | 0.01 | 0.02 | 1,708 | 0.01 |
| adjnoun | 112 | 425 | 6,216 | 0.03 | 896 | 0.01 | 0.03 | 1,042 | 0.01 | 0.04 | 1,143 | 0.01 |
| football | 115 | 613 | 6,555 | 0.03 | 1,302 | 0.01 | 0.04 | 2,625 | 0.02 | 0.04 | 3,868 | 0.02 |
| jazz | 198 | 2742 | 19,503 | 0.17 | 1,210 | 0.02 | 0.22 | 1,743 | 0.03 | 0.24 | 2,554 | 0.05 |
| celegans | 453 | 2025 | 102,378 | 0.56 | 30,802 | 0.19 | 0.58 | 50,332 | 0.30 | 0.62 | 35,724 | 0.26 |
| email | 1133 | 5451 | 641,278 | 2.21 | 238,000 | 0.84 | 2.23 | 199,893 | 0.94 | 2.45 | 165,313 | 1.01 |
| polblogs | 1490 | 16715 | 1,109,305 | 20.74 | 94,085 | 1.53 | 23.90 | 81,429 | 1.82 | 27.80 | 72,585 | 2.61 |
| netscience | 1589 | 2742 | 1,261,666 | 1.02 | 7,964 | 0.07 | 1.02 | 4,988 | 0.07 | 1.01 | 2,768 | 0.07 |
| power | 4941 | 6594 | 12,204,270 | 16.41 | 19,550 | 0.37 | 16.47 | 8,596 | 0.29 | 16.56 | 3,270 | 0.32 |
| hep-th | 8361 | 15751 | 34,948,980 | 74.50 | 372,611 | 2.07 | 74.42 | 211,676 | 1.54 | 74.69 | 120,801 | 1.30 |
| PGP | 10680 | 24316 | 57,025,860 | 155.08 | 727,213 | 5.12 | 155.79 | 429,349 | 4.16 | 155.67 | 315,819 | 3.72 |

**Table 6** Running time (in seconds) for computing $\rho_s$, single threaded (ST) vs multi-threaded (MT) with 32 threads using OpenMP, after filtering based on lower and upper bounds of $\rho_s$ on DIMACS-10 instances.

| | | | $s=3$ | | | | | | $s=4$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $r=2$ | | $r=3$ | | $r=4$ | | $r=2$ | | $r=3$ | | $r=4$ | |
| Graph | $n$ | $m$ | ST | MT | ST | MT | ST | MT | ST | MT | ST | MT | ST | MT |
| karate | 34 | 78 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| dolphins | 62 | 159 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.02 | 0.00 | 0.02 | 0.00 |
| lesmis | 77 | 254 | 0.00 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 | 0.02 | 0.00 | 0.03 | 0.00 | 0.02 | 0.00 |
| polbooks | 105 | 441 | 0.01 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 | 0.05 | 0.01 | 0.08 | 0.01 | 0.09 | 0.01 |
| adjnoun | 112 | 425 | 0.01 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 | 0.04 | 0.01 | 0.07 | 0.01 | 0.10 | 0.01 |
| football | 115 | 613 | 0.01 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 | 0.09 | 0.01 | 0.18 | 0.02 | 0.29 | 0.02 |
| jazz | 198 | 2742 | 0.03 | 0.00 | 0.03 | 0.00 | 0.04 | 0.00 | 0.20 | 0.02 | 0.34 | 0.03 | 0.44 | 0.05 |
| celegans | 453 | 2025 | 0.10 | 0.01 | 0.22 | 0.02 | 0.25 | 0.02 | 1.73 | 0.19 | 3.97 | 0.30 | 3.16 | 0.26 |
| email | 1133 | 5451 | 0.21 | 0.05 | 0.25 | 0.06 | 0.26 | 0.06 | 10.18 | 0.84 | 11.96 | 0.94 | 13.85 | 1.01 |
| polblogs | 1490 | 16715 | 0.61 | 0.06 | 0.66 | 0.08 | 0.74 | 0.08 | 16.77 | 1.53 | 23.16 | 1.82 | 31.84 | 2.61 |
| netscience | 1589 | 2742 | 0.13 | 0.04 | 0.12 | 0.04 | 0.12 | 0.04 | 0.37 | 0.07 | 0.28 | 0.07 | 0.25 | 0.07 |
| power | 4941 | 6594 | 1.21 | 0.30 | 1.25 | 0.29 | 1.16 | 0.27 | 2.46 | 0.37 | 2.15 | 0.29 | 1.70 | 0.32 |
| hep-th | 8361 | 15751 | 4.07 | 0.71 | 3.95 | 0.70 | 3.75 | 0.74 | 29.36 | 2.07 | 21.30 | 1.54 | 14.65 | 1.30 |
| PGP | 10680 | 24316 | 6.93 | 1.19 | 6.96 | 1.09 | 7.46 | 1.51 | 80.33 | 5.12 | 62.43 | 4.16 | 52.39 | 3.72 |

**Table 7** Running time (in seconds) of the heuristic for $r$-robust $s$-clubs on DIMACS-10 instances.

| | | | $s=2$ | | | $s=3$ | | | $s=4$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Graph | $n$ | $m$ | $r=2$ | $r=3$ | $r=4$ | $r=2$ | $r=3$ | $r=4$ | $r=2$ | $r=3$ | $r=4$ |
| karate | 34 | 78 | 0.03 | 0.00 | 0.02 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 |
| dolphins | 62 | 159 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 |
| lesmis | 77 | 254 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 |
| polbooks | 105 | 441 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.02 |
| adjnoun | 112 | 425 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.07 |
| football | 115 | 613 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.04 | 0.18 | 0.49 |
| jazz | 198 | 2742 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.03 | 0.05 | 0.10 |
| celegans | 453 | 2025 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.02 | 0.23 | 0.77 | 0.90 |
| email | 1133 | 5451 | 0.01 | 0.01 | 0.01 | 0.08 | 0.08 | 0.06 | 6.15 | 9.58 | 8.93 |
| polblogs | 1490 | 16715 | 0.03 | 0.02 | 0.03 | 0.13 | 0.11 | 0.13 | 2.26 | 2.87 | 3.81 |
| netscience | 1589 | 2742 | 0.00 | 0.00 | 0.00 | 0.08 | 0.00 | 0.00 | 0.09 | 0.11 | 0.04 |
| power | 4941 | 6594 | 0.06 | 0.00 | 0.00 | 0.07 | 0.00 | 0.00 | 0.16 | 0.00 | 0.00 |
| hep-th | 8361 | 15751 | 0.06 | 0.05 | 0.01 | 0.09 | 0.08 | 0.03 | 1.70 | 1.46 | 1.02 |
| PGP | 10680 | 24316 | 0.05 | 0.02 | 0.01 | 0.13 | 0.08 | 0.07 | 4.19 | 3.19 | 3.33 |

**Table 8** Running time (in seconds) of the heuristic for $t$-hereditary $s$-clubs on DIMACS-10 instances.

| | | | $s=2$ | | | $s=3$ | | | $s=4$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Graph | $n$ | $m$ | $t=2$ | $t=3$ | $t=4$ | $t=2$ | $t=3$ | $t=4$ | $t=2$ | $t=3$ | $t=4$ |
| karate | 34 | 78 | 0.05 | 0.00 | 0.04 | 0.07 | 0.01 | 0.07 | 0.07 | 0.37 | 0.01 |
| dolphins | 62 | 159 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.51 | 0.01 |
| lesmis | 77 | 254 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.11 | 0.00 |
| polbooks | 105 | 441 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.35 | 0.02 |
| adjnoun | 112 | 425 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.17 | 0.06 |
| football | 115 | 613 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.86 | 0.47 |
| jazz | 198 | 2742 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.04 | 0.14 | 0.12 |
| celegans | 453 | 2025 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.02 | 0.25 | 1.12 | 1.01 |
| email | 1133 | 5451 | 0.02 | 0.01 | 0.01 | 0.11 | 0.07 | 0.06 | 5.54 | 11.05 | 8.89 |
| polblogs | 1490 | 16715 | 0.03 | 0.02 | 0.04 | 0.09 | 0.12 | 0.13 | 2.12 | 2.72 | 3.62 |
| netscience | 1589 | 2742 | 0.00 | 0.00 | 0.00 | 0.08 | 0.00 | 0.07 | 0.01 | 0.10 | 0.00 |
| power | 4941 | 6594 | 0.08 | 0.00 | 0.00 | 0.10 | 0.00 | 0.00 | 0.20 | 0.17 | 0.00 |
| hep-th | 8361 | 15751 | 0.10 | 0.07 | 0.02 | 0.15 | 0.10 | 0.04 | 1.78 | 4.10 | 1.00 |
| PGP | 10680 | 24316 | 0.08 | 0.05 | 0.02 | 0.15 | 0.09 | 0.09 | 4.03 | 3.80 | 3.50 |

**Table 9** The size of the $r$-robust $s$-clubs found by the heuristic on DIMACS-10 instances. Entries highlighted in bold were subsequently proved to be optimal.

| | | | $s=2$ | | | $s=3$ | | | $s=4$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Graph | $n$ | $m$ | $r=2$ | $r=3$ | $r=4$ | $r=2$ | $r=3$ | $r=4$ | $r=2$ | $r=3$ | $r=4$ |
| karate | 34 | 78 | **12** | **6** | **6** | **21** | **11** | **9** | **26** | 12 | **10** |
| dolphins | 62 | 159 | 7 | **7** | 1 | 20 | 9 | 5 | **32** | 23 | 13 |
| lesmis | 77 | 254 | **18** | **14** | **13** | **35** | **25** | **21** | **51** | **34** | **25** |
| polbooks | 105 | 441 | 19 | 13 | 10 | **39** | 28 | 22 | 57 | 43 | 33 |
| adjnoun | 112 | 425 | 21 | 10 | 5 | 62 | **47** | 26 | **94** | **81** | 64 |
| football | 115 | 613 | **14** | **13** | **12** | 15 | 12 | 7 | 111 | 97 | 19 |
| jazz | 198 | 2742 | 76 | **73** | **65** | **158** | **145** | 132 | **186** | **181** | 173 |
| celegans | 453 | 2025 | **104** | **54** | 25 | **234** | 140 | 92 | **378** | **291** | 204 |
| email | 1133 | 5451 | 25 | 18 | 13 | 118 | 55 | 30 | 497 | 387 | 326 |
| polblogs | 1490 | 16715 | 179 | 181 | 153 | 669 | 604 | 553 | **1000** | **913** | 851 |
| netscience | 1589 | 2742 | **22** | 15 | 9 | **24** | 20 | 9 | **29** | **21** | 9 |
| power | 4941 | 6594 | 6 | 4 | **6** | 16 | 4 | **12** | 28 | 4 | **13** |
| hep-th | 8361 | 15751 | **33** | 19 | 14 | 50 | 36 | **32** | 165 | 93 | 56 |
| PGP | 10680 | 24316 | 94 | 70 | **64** | **239** | 168 | 121 | 444 | 305 | 215 |

**Table 10** **The size of the $t$-hereditary $s$-clubs found by the heuristic on DIMACS-10 instances. Entries highlighted in bold were subsequently proved to be optimal.**

| | | | $s=2$ | | | $s=3$ | | | $s=4$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Graph | $n$ | $m$ | $t=2$ | $t=3$ | $t=4$ | $t=2$ | $t=3$ | $t=4$ | $t=2$ | $t=3$ | $t=4$ |
| karate | 34 | 78 | **12** | 4 | **6** | **21** | **11** | **9** | **26** | 12 | **10** |
| dolphins | 62 | 159 | 5 | 6 | 5 | 20 | **17** | 4 | **32** | 23 | 13 |
| lesmis | 77 | 254 | **18** | **14** | **13** | **35** | **25** | **21** | **51** | **34** | **25** |
| polbooks | 105 | 441 | 19 | 13 | 10 | **39** | 28 | 22 | 57 | 43 | 33 |
| adjnoun | 112 | 425 | 21 | 10 | 5 | 62 | **47** | 26 | **94** | **81** | 64 |
| football | 115 | 613 | 12 | **13** | 5 | 10 | 13 | 6 | **115** | 97 | 18 |
| jazz | 198 | 2742 | 76 | **73** | **65** | **158** | **145** | 132 | **186** | **181** | 173 |
| celegans | 453 | 2025 | **104** | **54** | 25 | **234** | 140 | 92 | **378** | **291** | 204 |
| email | 1133 | 5451 | 26 | 18 | 15 | 120 | 57 | 30 | 497 | 387 | 326 |
| polblogs | 1490 | 16715 | 178 | 180 | 152 | 669 | 604 | 554 | **1000** | **913** | 851 |
| netscience | 1589 | 2742 | **22** | 15 | 9 | **24** | 20 | 9 | **29** | **21** | 9 |
| power | 4941 | 6594 | 7 | 4 | **6** | 16 | 3 | **12** | 28 | 3 | **13** |
| hep-th | 8361 | 15751 | **33** | 19 | 14 | 50 | 36 | **32** | 165 | 93 | 63 |
| PGP | 10680 | 24316 | 94 | 70 | **64** | **239** | 168 | 121 | 444 | 305 | 215 |

**Table 11** **Vertex peeling time (in seconds) for $r$-robust $s$-clubs on DIMACS-10 instances.**

| | | | $s=2$ | | | $s=3$ | | | $s=4$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Graph | $n$ | $m$ | $r=2$ | $r=3$ | $r=4$ | $r=2$ | $r=3$ | $r=4$ | $r=2$ | $r=3$ | $r=4$ |
| karate | 34 | 78 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| dolphins | 62 | 159 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 |
| lesmis | 77 | 254 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| polbooks | 105 | 441 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.03 | 0.02 | 0.02 |
| adjnoun | 112 | 425 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.01 | 0.01 | 0.01 | 0.00 |
| football | 115 | 613 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| jazz | 198 | 2742 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.01 | 0.07 | 0.03 | 0.04 |
| celegans | 453 | 2025 | 0.01 | 0.02 | 0.01 | 0.04 | 0.04 | 0.03 | 0.46 | 0.40 | 0.33 |
| email | 1133 | 5451 | 0.12 | 0.13 | 0.08 | 0.13 | 0.12 | 0.22 | 2.29 | 3.53 | 3.06 |
| polblogs | 1490 | 16715 | 0.12 | 0.16 | 0.33 | 0.43 | 0.39 | 0.46 | 5.15 | 3.14 | 4.23 |
| netscience | 1589 | 2742 | 0.00 | 0.00 | 0.00 | 0.14 | 0.00 | 0.00 | 0.00 | 0.05 | 0.00 |
| power | 4941 | 6594 | 0.39 | 0.00 | 0.00 | 0.13 | 0.00 | 0.00 | 0.53 | 0.00 | 0.00 |
| hep-th | 8361 | 15751 | 0.59 | 0.20 | 0.04 | 0.54 | 0.47 | 0.12 | 3.26 | 1.36 | 1.06 |
| PGP | 10680 | 24316 | 0.59 | 0.17 | 0.06 | 0.85 | 0.22 | 0.17 | 10.73 | 4.50 | 3.26 |

**Table 12 Vertex peeling time (in seconds) for $t$-hereditary $s$-clubs on DIMACS-10 instances.**

| | | | $s=2$ | | | $s=3$ | | | $s=4$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Graph | $n$ | $m$ | $t=2$ | $t=3$ | $t=4$ | $t=2$ | $t=3$ | $t=4$ | $t=2$ | $t=3$ | $t=4$ |
| karate | 34 | 78 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| dolphins | 62 | 159 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.06 | 0.00 |
| lesmis | 77 | 254 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 | 0.00 |
| polbooks | 105 | 441 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.00 | 0.03 | 0.12 | 0.02 |
| adjnoun | 112 | 425 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.06 | 0.00 |
| football | 115 | 613 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| jazz | 198 | 2742 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 | 0.07 | 0.06 | 0.04 |
| celegans | 453 | 2025 | 0.01 | 0.01 | 0.01 | 0.04 | 0.04 | 0.02 | 0.52 | 0.55 | 0.37 |
| email | 1133 | 5451 | 0.12 | 0.16 | 0.12 | 0.21 | 0.15 | 0.24 | 2.26 | 3.19 | 3.09 |
| polblogs | 1490 | 16715 | 0.12 | 0.16 | 0.20 | 0.48 | 0.47 | 0.47 | 4.93 | 3.20 | 4.06 |
| netscience | 1589 | 2742 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.10 | 0.00 | 0.12 | 0.00 |
| power | 4941 | 6594 | 0.46 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.53 | 0.00 | 0.00 |
| hep-th | 8361 | 15751 | 0.66 | 0.27 | 0.04 | 0.55 | 0.55 | 0.14 | 3.88 | 1.51 | 0.90 |
| PGP | 10680 | 24316 | 0.48 | 0.26 | 0.07 | 0.85 | 0.23 | 0.21 | 10.85 | 4.90 | 3.40 |

**Table 13 The wall-clock running time (in seconds) for solving the maximum $r$-robust 2-club problem. If an instance was not solved to optimality under the time limit, the optimality gap is reported (highlighted in bold). The entry "LPNS" means that the root LP relaxation was not solved to optimality under the 1-hour time limit.**

| | | | Wall-clock running time | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $r=2$ | | | $r=3$ | | | $r=4$ | | |
| Graph | $n$ | $m$ | CN | CUT | BCUT | CN | CUT | BCUT | CN | CUT | BCUT |
| karate | 34 | 78 | 0.05 | 0.03 | 0.03 | 0.04 | 0.03 | 0.00 | 0.03 | 0.02 | 0.02 |
| dolphins | 62 | 159 | 0.15 | 0.06 | 0.03 | 0.08 | 0.06 | 0.00 | 0.06 | 0.05 | 0.04 |
| lesmis | 77 | 254 | 0.15 | 0.08 | 0.00 | 0.83 | 0.08 | 0.00 | 0.18 | 0.09 | 0.00 |
| polbooks | 105 | 441 | 0.36 | 0.14 | 0.03 | 3.25 | 0.13 | 0.03 | 0.33 | 0.20 | 0.03 |
| adjnoun | 112 | 425 | 0.29 | 0.15 | 0.03 | 0.16 | 0.18 | 0.02 | 0.09 | 0.39 | 0.03 |
| football | 115 | 613 | 0.31 | 0.10 | 0.11 | 0.31 | 0.07 | 0.00 | 0.26 | 0.07 | 0.00 |
| jazz | 198 | 2742 | 0.87 | 0.19 | 0.06 | 1.24 | 0.20 | 0.05 | 1.00 | 0.20 | 0.04 |
| celegans | 453 | 2025 | 3.30 | 1.41 | 0.02 | 71.68 | 1.38 | 0.02 | 30.02 | 1.11 | 0.03 |
| email | 1133 | 5451 | 318.11 | 109.48 | 7.38 | 121.89 | 38.12 | 0.53 | 37.63 | 13.40 | 0.28 |
| polblogs | 1490 | 16715 | 1382.21 | 22.39 | 5.25 | **4.52%** | 56.15 | 7.69 | **3.21%** | 61.49 | 6.61 |
| netscience | 1589 | 2742 | 34.78 | 22.64 | 0.00 | 37.21 | 19.97 | 0.01 | 39.24 | 15.23 | 0.01 |
| power | 4941 | 6594 | **240.00%** | 625.26 | 0.50 | 1363.12 | 53.24 | 0.02 | 790.46 | 41.31 | 0.00 |
| hep-th | 8361 | 15751 | **136.84%** | 1299.56 | 0.69 | **144.44%** | 1284.72 | 0.28 | **70.83%** | 897.76 | 0.07 |
| PGP | 10680 | 24316 | LPNS | 1479.27 | 0.71 | LPNS | LPNS | 0.22 | LPNS | 3074.40 | 0.10 |

**Table 14** The wall-clock running times (in seconds) for solving the maximum $t$-hereditary 2-club problem. If an instance was not solved to optimality, the optimality gap is reported (highlighted in bold). The entry "LPNS" means that the root LP relaxation was not solved to optimality under the 1-hour time limit.

| | | | Wall-clock running time | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $t=2$ | | | $t=3$ | | | $t=4$ | | |
| Graph | $n$ | $m$ | CN | CUT | BCUT | CN | CUT | BCUT | CN | CUT | BCUT |
| karate | 34 | 78 | 0.04 | 0.05 | 0.05 | 0.03 | 0.06 | 0.04 | 0.03 | 0.04 | 0.04 |
| dolphins | 62 | 159 | 0.12 | 0.05 | 0.05 | 0.21 | 0.05 | 0.02 | 0.12 | 0.06 | 0.03 |
| lesmis | 77 | 254 | 0.15 | 0.07 | 0.00 | 0.12 | 0.06 | 0.00 | 0.10 | 0.06 | 0.00 |
| polbooks | 105 | 441 | 0.27 | 0.09 | 0.02 | 1.15 | 0.10 | 0.03 | 1.16 | 0.09 | 0.04 |
| adjnoun | 112 | 425 | 0.29 | 0.09 | 0.03 | 1.25 | 0.38 | 0.05 | 1.08 | 0.31 | 0.15 |
| football | 115 | 613 | 0.21 | 0.09 | 0.10 | 0.17 | 0.09 | 0.00 | 0.25 | 0.09 | 0.09 |
| jazz | 198 | 2742 | 0.58 | 0.18 | 0.06 | 0.98 | 0.21 | 0.05 | 0.94 | 0.24 | 0.04 |
| celegans | 453 | 2025 | 3.30 | 1.26 | 0.02 | 65.61 | 1.08 | 0.02 | 41.52 | 1.10 | 0.05 |
| email | 1133 | 5451 | 460.84 | 109.72 | 9.79 | 260.83 | 45.89 | 1.03 | 67.91 | 14.50 | 0.36 |
| polblogs | 1490 | 16715 | 1484.41 | 18.38 | 5.08 | **55.71%** | 53.69 | 8.57 | **54.13%** | 76.73 | 7.18 |
| netscience | 1589 | 2742 | 38.74 | 18.53 | 0.00 | 34.82 | 18.43 | 0.02 | 40.08 | 19.72 | 0.01 |
| power | 4941 | 6594 | 1762.25 | 444.17 | 0.58 | 2411.53 | 656.91 | 0.02 | 2351.34 | 1919.99 | 0.00 |
| hep-th | 8361 | 15751 | **130.00%** | 1034.27 | 0.80 | **136.84%** | 1832.96 | 0.37 | **131.58%** | 1679.40 | 0.08 |
| PGP | 10680 | 24316 | LPNS | LPNS | 0.64 | LPNS | LPNS | 0.35 | LPNS | LPNS | 0.12 |

**Table 15** Optimal objective values for the maximum $r$-robust 2-club and $t$-hereditary 2-club problems found by BCUT.

| | | | MRCP, $s=2$ | | | MHCP, $s=2$ | | |
|---|---|---|---|---|---|---|---|---|
| Graph | $n$ | $m$ | $r=2$ | $r=3$ | $r=4$ | $t=2$ | $t=3$ | $t=4$ |
| karate | 34 | 78 | 12 | 6 | 6 | 12 | 6 | 6 |
| dolphins | 62 | 159 | 9 | 7 | 6 | 9 | 7 | 6 |
| lesmis | 77 | 254 | 18 | 14 | 13 | 18 | 14 | 13 |
| polbooks | 105 | 441 | 20 | 15 | 12 | 20 | 15 | 13 |
| adjnoun | 112 | 425 | 23 | 12 | 6 | 23 | 12 | 9 |
| football | 115 | 613 | 14 | 13 | 12 | 14 | 13 | 13 |
| jazz | 198 | 2742 | 79 | 73 | 65 | 79 | 73 | 65 |
| celegans | 453 | 2025 | 104 | 54 | 30 | 104 | 54 | 30 |
| email | 1133 | 5451 | 27 | 23 | 19 | 27 | 23 | 20 |
| polblogs | 1490 | 16715 | 232 | 182 | 158 | 232 | 182 | 159 |
| netscience | 1589 | 2742 | 22 | 21 | 20 | 22 | 21 | 20 |
| power | 4941 | 6594 | 9 | 7 | 6 | 9 | 7 | 6 |
| hep-th | 8361 | 15751 | 33 | 24 | 24 | 33 | 24 | 24 |
| PGP | 10680 | 24316 | 96 | 71 | 64 | 96 | 71 | 64 |

**Table 16 The wall-clock running times (in seconds) of solvers AC and BCUT using Algorithm 1 with the same heuristic and preprocessing steps for solving the maximum $r$-robust 3-club problem. If an instance was not solved to optimality, the optimality gap is reported. The entry "MEM" means that the solver ran out of memory.**

| | | | Objective | | | Wall-clock running time | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | $r=2$ | | $r=3$ | | $r=4$ | |
| Graph | $n$ | $m$ | $r=2$ | $r=3$ | $r=4$ | AC | BCUT | AC | BCUT | AC | BCUT |
| karate | 34 | 78 | 21 | 11 | 9 | 0.03 | 0.01 | 0.01 | 0.01 | 0.07 | 0.01 |
| dolphins | 62 | 159 | 22 | 14 | 7 | 0.34 | 0.04 | 0.35 | 0.06 | 0.17 | 0.13 |
| lesmis | 77 | 254 | 35 | 25 | 21 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| polbooks | 105 | 441 | 39 | 31 | 24 | 0.80 | 0.03 | 4.29 | 0.05 | 1.56 | 0.03 |
| adjnoun | 112 | 425 | 63 | 47 | 31 | 9.63 | 0.04 | 15.26 | 0.03 | 54.63 | 0.14 |
| football | 115 | 613 | 40 | 27 | 17 | 94.02 | 0.70 | 55.85 | 0.89 | 9.94 | 1.34 |
| jazz | 198 | 2742 | 158 | 145 | 136 | 65.95 | 0.06 | 438.29 | 0.06 | 276.30 | 0.06 |
| celegans | 453 | 2025 | 234 | 141 | 99 | 121.50 | 0.16 | 1627.88 | 0.16 | 1142.98 | 0.13 |
| email | 1133 | 5451 | 138 | 88 | 66 | **45.76%** | 403.02 | **125.42%** | **9.09%** | **243.33%** | 1130.71 |
| polblogs | 1490 | 16715 | 672 | 605 | 557 | MEM | 1.36 | MEM | 1.58 | MEM | 1.85 |
| netscience | 1589 | 2742 | 24 | 21 | 20 | 0.02 | 0.32 | 0.10 | 0.02 | 0.08 | 0.01 |
| power | 4941 | 6594 | 17 | 12 | 12 | 0.82 | 0.32 | 0.26 | 0.02 | 0.00 | 0.00 |
| hep-th | 8361 | 15751 | 52 | 38 | 32 | **8%** | 16.80 | 76.80 | 0.71 | 0.18 | 0.18 |
| PGP | 10680 | 24316 | 239 | 170 | 124 | 1.08 | 1.12 | 231.37 | 0.42 | 1815.75 | 0.41 |

**Table 17 Comparison of the best objective values found using the BCUT solver for the MRCP, the MHCP, and by directly solving the formulation for the maximum $r$-distinct-path $s$-club problem proposed by Veremyev and Boginski (2012a). Wall-clock running times (in seconds) are also reported for the BCUT solver. If an instance was not solved to optimality under the time limit, the optimality gap is reported (highlighted in bold).**

| | | | $r=t=2, s=3$ | | | | | $r=t=2, s=4$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Best objective | | | Wall-clock time | | Best objective | | | Wall-clock time | |
| Graph | $n$ | $m$ | Robust | Hereditary | Distinct-paths | Robust | Hereditary | Robust | Hereditary | Distinct-paths | Robust | Hereditary |
| karate | 34 | 78 | 21 | 21 | 22 | 0.01 | 0.07 | 26 | 26 | 33 | 0.01 | 0.07 |
| dolphins | 62 | 159 | 22 | 22 | 24 | 0.04 | 0.03 | 32 | 32 | 36 | 0.01 | 0.01 |
| lesmis | 77 | 254 | 35 | 35 | 49 | 0.00 | 0.00 | 51 | 51 | 65 | 0.01 | 0.01 |
| polbooks | 105 | 441 | 39 | 39 | 46 | 0.03 | 0.03 | 58 | 58 | 64 | 0.11 | 0.08 |
| adjnoun | 112 | 425 | 63 | 63 | 73 | 0.04 | 0.04 | 94 | 94 | 104 | 0.08 | 0.06 |
| football | 115 | 613 | 40 | 40 | 43 | 0.70 | 0.60 | 113 | 115 | 115 | 0.32 | 0.02 |
| jazz | 198 | 2742 | 158 | 158 | 165 | 0.06 | 0.06 | 186 | 186 | 192 | 0.10 | 0.11 |
| celegans | 453 | 2025 | 234 | 234 | 353 | 0.16 | 0.18 | 378 | 378 | 429 | 0.69 | 0.77 |
| email | 1133 | 5451 | 138 | 138 | 168 | 403.02 | 296.03 | 505 | 505 | $\geq$582 | 42.86 | 37.72 |
| polblogs | 1490 | 16715 | 672 | 672 | 715 | 1.36 | 1.30 | 1000 | 1000 | $\geq$1000 | 18.59 | 18.74 |
| netscience | 1589 | 2742 | 24 | 24 | 36 | 0.32 | 0.10 | 29 | 29 | 68 | 0.10 | 0.01 |
| power | 4941 | 6594 | 17 | 17 | 22 | 0.32 | 0.33 | 29 | 29 | 41 | 0.81 | 0.84 |
| hep-th | 8361 | 15751 | 52 | 52 | 76 | 16.80 | 9.99 | 177 | 177 | $\geq$177 | 66.77 | 80.39 |
| PGP | 10680 | 24316 | 239 | 239 | $\geq$239 | 1.12 | 1.14 | 446 | 446 | $\geq$446 | 34.73 | 30.67 |

| | | | $r=t=3, s=3$ | | | | | $r=t=3, s=4$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Best objective | | | Wall-clock time | | Best objective | | | Wall-clock time | |
| Graph | $n$ | $m$ | Robust | Hereditary | Distinct-paths | Robust | Hereditary | Robust | Hereditary | Distinct-paths | Robust | Hereditary |
| karate | 34 | 78 | 11 | 11 | 19 | 0.01 | 0.01 | 13 | 13 | 32 | 0.05 | 0.39 |
| dolphins | 62 | 159 | 14 | 17 | 23 | 0.06 | 0.00 | 24 | 24 | 34 | 0.11 | 0.68 |
| lesmis | 77 | 254 | 25 | 25 | 45 | 0.00 | 0.00 | 34 | 34 | 64 | 0.01 | 0.17 |
| polbooks | 105 | 441 | 31 | 31 | 41 | 0.05 | 0.06 | 44 | 44 | 60 | 0.21 | 0.69 |
| adjnoun | 112 | 425 | 47 | 47 | 67 | 0.03 | 0.04 | 81 | 81 | 101 | 0.14 | 0.36 |
| football | 115 | 613 | 27 | 27 | 36 | 0.89 | 1.24 | 99 | 103 | 115 | 0.96 | 1.42 |
| jazz | 198 | 2742 | 145 | 145 | 162 | 0.06 | 0.06 | 181 | 181 | 191 | 0.32 | 0.46 |
| celegans | 453 | 2025 | 141 | 141 | 321 | 0.16 | 0.18 | 291 | 291 | 426 | 3.23 | 3.69 |
| email | 1133 | 5451 | 88 | 88 | 130 | **9.09%** | **11.49%** | 407 | 407 | 558 | 83.68 | 87.10 |
| polblogs | 1490 | 16715 | 605 | 605 | 677 | 1.58 | 1.72 | 913 | 913 | $\geq$913 | 22.50 | 22.27 |
| netscience | 1589 | 2742 | 21 | 21 | 35 | 0.02 | 0.02 | 21 | 21 | 63 | 0.16 | 0.34 |
| power | 4941 | 6594 | 12 | 12 | 17 | 0.02 | 0.02 | 17 | 17 | 37 | 0.07 | 0.52 |
| hep-th | 8361 | 15751 | 38 | 38 | 65 | 0.71 | 0.87 | 109 | 109 | $\geq$109 | 174.38 | 262.87 |
| PGP | 10680 | 24316 | 170 | 170 | 251 | 0.42 | 0.44 | 308 | 308 | $\geq$308 | 14.29 | 19.30 |

| | | | $r=t=4, s=3$ | | | | | $r=t=4, s=4$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Best objective | | | Wall-clock time | | Best objective | | | Wall-clock time | |
| Graph | $n$ | $m$ | Robust | Hereditary | Distinct-paths | Robust | Hereditary | Robust | Hereditary | Distinct-paths | Robust | Hereditary |
| karate | 34 | 78 | 9 | 9 | 16 | 0.01 | 0.07 | 10 | 10 | 31 | 0.00 | 0.01 |
| dolphins | 62 | 159 | 7 | 7 | 20 | 0.13 | 0.17 | 17 | 17 | 33 | 0.06 | 0.06 |
| lesmis | 77 | 254 | 21 | 21 | 40 | 0.00 | 0.00 | 25 | 25 | 63 | 0.00 | 0.01 |
| polbooks | 105 | 441 | 24 | 24 | 40 | 0.03 | 0.04 | 35 | 35 | 59 | 0.06 | 0.06 |
| adjnoun | 112 | 425 | 31 | 32 | 64 | 0.14 | 0.11 | 67 | 67 | 100 | 0.47 | 0.49 |
| football | 115 | 613 | 17 | 17 | 30 | 1.34 | 1.41 | 65 | 65 | 115 | 7.23 | 13.07 |
| jazz | 198 | 2742 | 136 | 136 | 158 | 0.06 | 0.08 | 174 | 174 | 191 | 0.51 | 0.55 |
| celegans | 453 | 2025 | 99 | 99 | 295 | 0.13 | 0.13 | 207 | 207 | 424 | 5.38 | 3.08 |
| email | 1133 | 5451 | 66 | 66 | 116 | 1130.71 | 2563.28 | 340 | 340 | $\geq$522 | 92.21 | 83.22 |
| polblogs | 1490 | 16715 | 557 | 558 | 657 | 1.85 | 1.79 | 852 | 852 | $\geq$852 | 50.94 | 47.81 |
| netscience | 1589 | 2742 | 20 | 20 | 33 | 0.01 | 0.21 | 20 | 20 | 56 | 0.06 | 0.02 |
| power | 4941 | 6594 | 12 | 12 | 16 | 0.00 | 0.00 | 13 | 13 | 35 | 0.00 | 0.00 |
| hep-th | 8361 | 15751 | 32 | 32 | 51 | 0.18 | 0.20 | 70 | 70 | $\geq$70 | 118.55 | 98.71 |
| PGP | 10680 | 24316 | 124 | 124 | 229 | 0.41 | 0.45 | 227 | 227 | $\geq$227 | 12.72 | 13.36 |