Subquadratic Kernels for Implicit 3-Hitting Set and 3-Set Packing Problems

Tien-Nam Le* Daniel Lokshtanov† Saket Saurabh‡ Stéphan Thomassé§ Meirav Zehavi¶

Abstract

We consider four well-studied NP-complete packing/covering problems on graphs: Feedback Vertex Set in Tournaments (FVST), Cluster Vertex Deletion (CVD), Triangle Packing in Tournaments (TPT) and Induced P_3-Packing. For these four problems kernels with $\mathcal{O}(k^2)$ vertices have been known for a long time. In fact, such kernels can be obtained by interpreting these problems as finding either a packing of k pairwise disjoint sets of size 3 (3-Set Packing) or a hitting set of size at most k for a family of sets of size at most 3 (3-Hitting Set). In this paper, we give the first kernels for FVST, CVD, TPT and Induced P_3-Packing with a subquadratic number of vertices. Specifically, we obtain the following results.

- FVST admits a kernel with $\mathcal{O}(k^{\frac{3}{2}})$ vertices.
- CVD admits a kernel with $\mathcal{O}(k^{\frac{5}{3}})$ vertices.
- TPT admits a kernel with $\mathcal{O}(k^{\frac{3}{2}})$ vertices.
- Induced P_3-Packing admits a kernel with $\mathcal{O}(k^{\frac{5}{3}})$ vertices.

Our results resolve an open problem from WorKer 2010 on the existence of kernels with $\mathcal{O}(k^{2-\epsilon})$ vertices for FVST and CVD. All of our results are based on novel uses of old and new "expansion lemmas", and a weak form of crown decomposition where (i) *almost all* of the head is used by the solution (as opposed to *all*), (ii) *almost none* of the crown is used by the solution (as opposed to *none*), and (iii) if H is removed from G, then there is *almost no* interaction between the head and the rest (as opposed to *no* interaction at all).

*École Normale Supérieure de Lyon, Lyon, France. tien-nam.le@ens-lyon.fr

†University of Bergen, Bergen, Norway. daniello@ii.uib.no

‡University of Bergen, Bergen, Norway, and The Institute of Mathematical Sciences, HBNI, Chennai, India. saket@imsc.res.in

§École Normale Supérieure de Lyon, Lyon, France. stephan.thomasse@ens-lyon.fr

¶University of Bergen, Bergen, Norway. meirav.zehavi@ii.uib.no

1 Introduction

Kernelization, a subfield of Parameterized Complexity, provides a mathematical framework to analyze the performance of polynomial time preprocessing. It makes it possible to quantify the degree to which polynomial time algorithms succeed at reducing input instances of an NP-hard problem. More formally, every instance of a parameterized problem Π is associated with an integer k, which is called the *parameter*, and Π is said to admit a *kernel* if there is a polynomial-time algorithm, called a *kernelization algorithm*, that reduces the input instance of Π down to an equivalent instance of Π whose size is bounded by a function $f(k)$ of k. (Here, two instances are equivalent if both of them are either Yes-instances or No-instances.) Such an algorithm is called an $f(k)$-*kernel* for Π. If $f(k)$ is a polynomial function of k, we say that the kernel is a polynomial kernel. Over the last decade, kernelization has become an active field of study, especially with the development of complexity-theoretic lower bound tools for kernelization. These tools can be used to show that a polynomial kernel [5, 14, 21, 23], or a kernel of a specific size [10, 11, 24] for concrete problems would imply an unlikely complexity-theoretic collapse. We refer to the surveys [19, 22, 27, 29], as well as the books [9, 13, 17, 32], for a detailed treatment of the area of kernelization.

One of the most well known examples of a polynomial kernel is a kernel with $\mathcal{O}(k^d)$ sets and elements for d-Hitting Set using the Erdös-Rado Sunflower lemma.[1] In this problem, the input consists of a universe U, a family $\mathcal{F}$ containing sets of size at most d over U, and in integer k. The objective is to determine whether there exists a set $S \subseteq U$ of size at most k that intersects every set in $\mathcal{F}$. Abu-Khzam [2] gave an improved kernel for d-Hitting Set, still with $\mathcal{O}(k^d)$ sets, but with $\mathcal{O}(k^{d-1})$ elements.

The importance of the d-Hitting Set problem stems from the number of other problems that can be re-cast in terms of it. For example, in the Feedback Vertex Set in Tournaments (FVST) problem, the input is a tournament T together with an integer k. The task is to determine whether there exists a subset S of vertices of size at most k such that the sub-tournament $T - S$ obtained from T by removing S is acyclic. It turns out that FVST is a 3-Hitting Set problem, where the vertices of T are the universe, and the family $\mathcal{F}$ is the family containing the vertex set of every directed cycle on three vertices (triangle) of T. Indeed, it can easily be shown that for every vertex set S, $T-S$ is acyclic if and only if S is a hitting set for $\mathcal{F}$. Another example is the Cluster Vertex Deletion (CVD) problem. Here, the input is a graph G and an integer k, and the task is to determine whether there exists a subset S of at most k vertices such that every connected component of $G - S$ is a clique (such graphs are called *cluster* graphs). Also this problem can be formulated as a 3-Hitting Set problem where the family $\mathcal{F}$ contains the vertex sets of all *induced* P_3*'s* of G. An induced P_3 is a path on three vertices where the first and last vertex are non-adjacent in G. The kernel with $\mathcal{O}(k^2)$ elements for d-Hitting Set [2] can be adapted to obtain kernels with $\mathcal{O}(k^2)$ vertices for Feedback Vertex Set in Tournaments [12] and for Cluster Vertex Deletion [25].

The formulation of problems in terms of 3-Hitting Set is useful not only in the context of kernelization, but within several paradigms for dealing with NP-hardness. Indeed, the $2.076^k n^{\mathcal{O}(1)}$ time parameterized algorithm of Wahlström [34], the $\mathcal{O}(1.519^{n+o(n)})$ time exact exponential time algorithm of Fomin et al. [18], and the folklore factor 3-approximation algorithm for 3-Hitting Set, all *immediately* translate to algorithms with the same performance for FVST and CVD.

Still, as one translates graph problems into 3-Hitting Set, some structure is lost. This structure can often be exploited to obtain algorithms with better performance than the corre-

[1] The origins of this result are unclear. The first kernel with $\mathcal{O}(k^d)$ sets appeared in the work by Fellows et al. [15], but they do not make use of the Sunflower Lemma. To the best of our knowledge, the first exposition of the kernel based on the Sunflower Lemma appears in the book of Flum and Grohe [17].

sponding 3-Hitting Set algorithm. In particular, for FVST, Cai et al. [8] gave a factor 2.5 approximation algorithm. This has recently been improved to 7/3 by Mnich et al. [30]. For CVD, You et al. [35] gave a factor 2.5 approximation algorithm, which later was improved to 7/3 by Fiorini et al. [16]. In the realm of parameterized algorithms, the graph problems also seem more tractable than the general 3-Hitting Set. For FVST, Dom et al. [12] designed a $2^k n^{\mathcal{O}(1)}$ time algorithm, which was recently improved by Kumar and Lokshtanov [28] to a $1.619^k n^{\mathcal{O}(1)}$ time algorithm. For CVD, Hüffner et al. [25] gave a $2^k n^{\mathcal{O}(1)}$ time algorithm, which in turn was improved by Boral et al. [7] to a $1.911^k n^{\mathcal{O}(1)}$ time algorithm. Finally, by the result of Fomin et al. [18] that translates parameterized algorithms for subset problems into exact exponential time algorithms in a black box fashion, the improvements in parameterized algorithms percolate to the realm of exact exponential time algorithms. In particular, FVST and CVD have algorithms with running times $\mathcal{O}(1.382^n)$ and $\mathcal{O}(1.476^n)$, respectively, outperforming the $\mathcal{O}(1.519^n)$ time algorithm [18] for 3-Hitting Set.

Remarkably, from the perspective of kernelization, FVST and CVD have so far seemed to be as difficult as 3-Hitting Set in the sense that no kernel with $\mathcal{O}(k^{2-\epsilon})$ vertices, for some fixed $\epsilon > 0$, has been found for either of these two problems. Whether FVST and CVD admit such kernels was first posed as an open problem in WorKer 2010 [6, page 4], and variants of this question have been re-stated several times after that [12, 35, 4].

In this paper, we give the first kernels for FVST and CVD with a subquadratic number of vertices. Specifically, we obtain the following results.

- FVST admits a kernel with $\mathcal{O}(k^{\frac{3}{2}})$ vertices.
- CVD admits a kernel with $\mathcal{O}(k^{\frac{5}{3}})$ vertices.

The Sunflower Lemma based kernel for d-Hitting Set and the improvement of Abu-Khzam [2] can also be applied to the d-Set Packing problem [1]. Here, the input consists of a universe U and a family $\mathcal{F}$ of sets of size d over U, together with an integer k. The task is to determine whether there exists a subfamily $\mathcal{F}'$ of k pairwise disjoint sets. The d-Set Packing problem is *dual* to d-Hitting Set in several ways, among others in the sense that the dual of the linear programming relaxation of the d-Hitting Set problem is exactly the linear programming relaxation of d-Set Packing, and vice versa.

In the same way that d-Hitting Set is an archetypal "covering" problem that generalizes many such problems, d-Set Packing generalizes many "packing" problems. For example, it generalizes the Triangle Packing in Tournaments (TPT) and Induced P_3-Packing problems. In Triangle Packing in Tournaments, the input is a tournament T and an integer k, and the task is to determine whether T contains k pairwise vertex-disjoint triangles. In Induced P_3-Packing, the input is a graph G and an integer k, and the task is to determine whether G contains k pairwise vertex-disjoint induced P_3's. These problems are the duals of FVST and CVD, respectively.

Just like the insights that led to a kernel for d-Hitting Set also led to a kernel for d-Set Packing, our insights from the improved kernelization algorithms for FVST and CVD yield improved kernelization algorithms for Triangle Packing in Tournaments (TPT) and Induced P_3-Packing. Specifically, we obtain the following results.

- TPT admits a kernel with $\mathcal{O}(k^{\frac{3}{2}})$ vertices.
- Induced P_3-Packing admits a kernel with $\mathcal{O}(k^{\frac{5}{3}})$ vertices.

We remark that, while the underlying philosophy of the kernels for Triangle Packing in Tournaments (TPT) and Induced P_3-Packing is borrowed from the kernels for FVST and CVD, obtaining the kernels for TPT and Induced P_3-Packing requires significant additional insights. However, for the sake of exposition, we next only focus (in the introduction) on our methods in the context of FVST and CVD.

Overview and Our Methods. Our kernelization algorithms for both FVST and CVD begin by employing trivial factor 3 polynomial time approximation algorithms.[2] We use these algorithms to obtain approximate solutions of size at most $3k$, or conclude that no solution of size at most k exists. So, let us now assume that we have solutions S of size at most $3k$. In what follows, for both FVST and CVD, we aim to understand which "subpart" of the problem is similar to the VERTEX COVER problem.

Let us first focus on our approach to specifically solve FVST. To this end, let (T, k) be an instance of FVST. Given the approximate solution S, our analysis starts by introducing the notion of a *strong arc*. Formally, an arc $xy \in E(T)$ is *strong* if **(i)** at least one vertex among x and y belongs to S, and **(ii)** there are at least $k+2$ vertices $z \in V(T)$ such that xyz is a triangle. Let F be the set of all the strong arcs of T. Observe that any solution of size at most $k+1$ must be a vertex cover of F. Before we analyze F, we need to examine S as described below.

Now, we try to "fit" every vertex $s \in S$ into the unique topological ordering, $\prec$, of $X = T - S$. Towards this, for $s \in S$ and $x \in V(X)$, define $f_s^-(x) = \big|\{y \in V(X) : y \preceq x,\ sy \in E(T)\}\big|$, and $f_s^+(x) = \big|\{y \in V(X) : y \succ x,\ ys \in E(T)\}\big|$. Intuitively, the functions $f_s^-(x)$ and $f_s^+(x)$ measure how many arcs would have been in the "wrong direction" (with respect to the ordering $\prec$) if we inserted s into the position immediately after x in X. Using a simple "sliding argument", we show that for each $s \in S$, there exists $x_s \in V(X)$ such that $0 \leq f_s^-(x_s) - f_s^+(x_s) \leq 1$. Then, for each $s \in S$, the smallest vertex (with respect to $\prec$) satisfying the property that $0 \leq f_s^-(x_s) - f_s^+(x_s) \leq 1$ is denoted by $\varphi(s)$. Observe that if for some $s \in S$ and $x \in X$ we have that $f_s^-(x), f_s^+(x) \geq k+2$, then s participates in $k+1$ triangles whose pairwise intersection is exactly s. This implies that s must be part of every solution of size at most k. Thus, $f_s^-(\varphi(s)), f_s^+(\varphi(s)) \leq k+1$.

Next, we separately investigate the structure of triangles that contain a *strong arc*, and triangles that do not contain any strong arc. Formally, we call a triangle *local* if it does not contain any strong arc. In particular, we show that the vertices of any local triangle cannot lie "too far apart" in the ordering $\prec$ (of course, for a vertex $s \in S$, we use $\varphi(s)$ to measure the distance with respect to $\prec$). Having this claim at hand, FVST can be thought of as the problem of simultaneously hitting local triangles and strong arcs.

To take care of the two sets of objects to be hit simultaneously, we define a variant of Expansion Lemma [33, 9], which we call Double Expansion Lemma. To (roughly) describe it here, let $\ell > 0$ and G be a bipartite graph with vertex sets A, S, and $\widehat{S} \subseteq S$ and $\widehat{A} \subseteq A$. We say that $\widehat{S}$ has an *ℓ-expansion* into $\widehat{A}$ in G if $|N_G(Y) \cap \widehat{A}| \geq \ell|Y|$ for every $Y \subseteq \widehat{S}$. In addition, we would like to ensure that $N_G(\widehat{A}) \subseteq \widehat{S}$. In Double Expansion Lemma, we consider a scheme where we have one "global" bipartite graph, as well as d vertex-disjoint "local" bipartite graphs, and we would like to find a vertex set that exhibits the expansion and neighborhood containment properties in all of the graphs simultaneously (see Section 3 for details).

To design the subquadratic kernel for FVST, we apply Double Expansion Lemma where one "part" is S, and the other "part" is derived by first defining a set of "carefully selected subintervals" of X, say $Y_1, \ldots, Y_p$, trimming their ends to obtain yet another set of subintervals, $Y_1', \ldots, Y_p'$, and then further partitioning each trimmed subinterval Y_i' into a more refined set of subintervals, say $Y_{i,1}, \ldots, Y_{i,q}$. To be somewhat more precise, let us note that we have a global graph, G, with vertex bipartition $(\{Y_{i,j} : i \in \{1, \ldots, p\}, j \in \{1, \ldots, q\}\}, S)$,[3] as well as local bipartite graphs, H_i, with vertex bipartition $(\{Y_{i,j} : j \in \{1, \ldots, q\}\}, S_i)$, where S_i are those vertices in S that were determined to "fit" Y_i. The graphs H_i take care of local triangles, and the global graph G takes care of vertex cover constraints (that is, the edges in F). We apply the Double Expansion Lemma appropriately, and show that if $|V(X)| \geq \zeta k^{3/2}$, for some constant

[2]We could have also used approximation algorithms with better approximation ratios, but this modification would not result in better kernels.

[3]More precisely, here we mean that each subinterval $Y_{i,j}$ is represented by a unique single vertex in G.

ζ, then we can find an irrelevant vertex in X (that is, a vertex whose removal preserves the answer). This together with the fact that $|S| \leq 3k$ implies that we have a kernel of size $\mathcal{O}(k^{3/2})$.

Now, let us describe our approach to solve CVD. To this end, recall that we have an approximate solution S of size at most $3k$. Our kernelization algorithm begins with a simple application of the classical Expansion Lemma to bound the number of cliques in $G \setminus S$. Having bounded the number of cliques, we repeatedly apply a marking procedure called Mark, whose sequential set of applications is of the flavor of an Expansion Lemma, and can be thought of as a weak form of a crown decomposition, as we explain after its description. Roughly speaking, one run of Mark is executed as follows. Initially, all the vertices in S are "alive". For $k+1$ stages, Mark examines every vertex $s \in S$ that is still alive, and attempts to associate an edge of a clique of $G \setminus S$ to it. Here, the association can be done only if s is adjacent to exactly one vertex of the edge, and no vertex of that edge belongs to an already associated edge. If the attempt is successful, the vertex remains alive also for the next stage. If there exists a vertex that is alive after stage $k+1$, then this vertex is part of $k+1$ induced P_3's that intersect only at it, and hence we can apply a reduction rule. Supposing that this "lucky" situation does not occur, we say that the procedure was successful if roughly $k^{2/3}$ vertices were still alive at stage (roughly) $k^{2/3}$. If the run was indeed successful in this sense, we mark all of the vertices alive at stage $k^{2/3}$, and rerun the procedure on the graph G from which all marked vertices, which belong to S, are removed (only for the sake of applying Mark again).

Let $\widehat{U}$ denote the set of all the vertices in S that were marked across all successful runs. Furthermore, denote $L = S \setminus \widehat{U}$. Now, let us explain how the sets S, $V(G) \setminus (S \cup L)$ and L can be thought of as a weak form of a crown decomposition.[4] Here, the Head is $\widehat{U}$, and we indeed prove that *almost all* of the vertices of $\widehat{U}$ should be part of any solution (as opposed to *all* vertices as in a standard crown decomposition). Second, the Crown is $V(G) \setminus (S \cup L)$, and as a consequence of the fact that most of $\widehat{U}$ is present in every solution and as $V(G) \setminus (S \cup L)$ is significantly larger than k (else we already have a kernel), we can (roughly) say that *most* of the vertices in $V(G) \setminus (S \cup L)$ are not present in any solution (as opposed to *none*). Third, the Rest (or Royal Body) is L, and we prove (in the sense explained below) that the "interaction" between the Head and the Rest is structured (as opposed to non-existent as in a standard crown decomposition). Let us now elaborate on the meaning of our last claim. Here, we compute a "small" subset $M \subseteq V(G) \setminus (S \cup L)$ (specifically, this is the set of vertices associated to the vertices of L in the last unsuccessful run of Mark) such that every clique in $G \setminus S$ becomes a module with respect to L once we remove the vertices in M from it.

Having the decomposition described above, the situation is more complicated that it usually is when we have a standard crown decomposition. To analyze this situation, we first classify the cliques in $G \setminus S$ using three definitions. First, we classify these cliques as small, large or huge, and "throw away" the small cliques. Next, we also classify these cliques as either heavy or light, which corresponds to whether the fraction of vertices of the cliques that belong to M is large or small, respectively; in this step, we also throw away the heavy cliques, which can be done safely as M is shown to be small. Then, we also classify the cliques as either visible or hidden, corresponding to whether many or few vertices from L are adjacent to many vertices in these cliques, respectively. We show that not too many cliques can be visible, else a reduction rule can be applied, which allows us to throw away also large (but not huge) visible cliques. Next, we focus on good cliques, which are either large or huge, light, and either hidden or huge.

Our analysis proceeds by defining, for every vertex $s \in S$, a small and a large side with respect to every clique. Roughly speaking, a side is the set of either all neighbors or all non-neighbors of s in that clique. Then, in the context of these sides, we prove (using an exchange argument) that good cliques exhibit a vertex cover-like behavior. That is, for any vertex $s \in S$

[4] A crown decomposition is among the most classical and well-known tools in parameterized complexity. Readers unfamiliar with this notion (which we use only in the introduction) are referred to books such as [9].

and good clique, every solution either picks s or the entire small side of that clique with respect to s. This claim gives rise to the definition of a bipartite graph where one side is S and the other side is the set of vertices of the good cliques. Here, there is an edge between $s \in S$ and a vertex v in a good clique C if v belongs to the small side of C with respect to s. Using the Expansion Lemma, if we find a large enough expansion in this graph, we prove that it is safe to select the vertices in S corresponding to that expansion. Let us remark that this proof is non-trivial as the edges of the bipartite graph are *not* necessarily edges in the input graph G. Finally, if no large expansion can be found, it means that the bipartite graph contains many isolated vertices, which belong to the good cliques. However, because these vertices are isolated, we can observe that they form sets that are modules with respect to the entire graph G (rather than only with respect to L), which allows us to employ a reduction rule that decreases their number.

Finally, we say a few words about our kernels for packing problems, that is, for TPT and INDUCED P_3-PACKING. In both of these kernels, we start by finding a greedy packing, $\mathcal{S}$ of either triangles or induced paths on 3 vertices, depending on the problem we are dealing with. If the greedy collection is large, then we already have the answer. Else, the vertices present in any set in $\mathcal{S}$, say S, form a hitting set. That is, $G - S$ is a cluster graph and $T - S$ is a transitive tournament. We exploit this structure in a manner similar to the way we exploited it to design subquadratic kernels for the hitting problems. Specifically, we make reduction rules that are, in some sense, "dual" to those given for FVST and CVD, and use the appropriate variants of Expansion Lemma to find an irrelevant vertex to delete. However, as we currently deal with packing problems, there are also major deviations required to design the new kernels. For example, for INDUCED P_3-PACKING, the last stage of the kernelization algorithm, which lies at the heart of its correctness, is completely different from the last stage of the kernelization algorithm for CVD. Here, the difference stems from the following crucial observation: in INDUCED P_3-PACKING, we need to present structural claims that hold for at least one solution, rather than for all solutions as in CVD, but these structural claims have to be stronger than the ones presented for CVD as the solution itself has a more complicated structure (being a set of paths rather than a set of vertices). This crucial observation also holds for TPT, posing difficulties of the same nature.

Additional Related Works. It is known that unless $\mathsf{NP} \subseteq \frac{\mathsf{co\text{-}NP}}{\mathsf{poly}}$, for any $d \geq 2$ and for any $\epsilon > 0$, d-HITTING SET and d-SET PACKING do not admit a kernel with $\mathcal{O}(k^{d-\epsilon})$ sets [10, 11]. In [10], Dell and Marx studied several matching and packing problems, and provided non-trivial lower bounds as well as non-trivial upper bounds for packing some specific graphs such as matchings, P_4's (here, the packing need not be induced) and $K_{1,d}$'s (stars with d leaves). Moser et al. [31] studied the problem of packing a fixed connected graph H on ℓ vertices in an input graph G (that is, determining whether there exist k vertex disjoint copies of H in G) and designed a kernel with $\mathcal{O}(k^{\ell-1})$ vertices. In this context, it is also worth to point out the dichotomy result of Jansen and Marx [26] regarding packing a fixed graph H. Finally, very recently Bessy et al. [3] studied FVST where the input tournament is restricted to be a *sparse tournament*, that is, a tournament where the feedback arc set is a matching. For this special case, they presented a linear-vertex kernel, and remarked that their methods do not extend to handle general tournaments.

Reading Guide. On the one hand, our kernels for FVST and CVD are independent of each other. On the other hand, the kernels for TPT and INDUCED P_3-PACKING borrow some of their ideas from the corresponding hitting set kernels, and therefore we recommend to read them after reading our kernels for FVST and CVD. Section 3 gives the old and new Expansion Lemmas used in this paper. In Section 4, we give our $\mathcal{O}(k^{5/3})$-vertex kernel for CVD, followed by a kernel of $\mathcal{O}(k^{3/2})$ vertices for FVST in Section 5. In Sections 6 and 7, we give our kernels for INDUCED P_3-PACKING and TPT with $\mathcal{O}(k^{5/3})$ and $\mathcal{O}(k^{3/2})$ vertices, respectively. We conclude the paper with some remarks and open problems in Section 8. To get a detailed idea of our techniques, a reader can first read the statements of Expansion Lemmas in Section 3, and then proceed to read our kernels for CVD and FVST. The proofs of the new Expansion Lemmas and the kernels for the packing problems can be read afterwards.

2 Preliminaries

Graph Theory. Given a graph G (or digraph D), we let $V(G)$ $(V(D))$ and $E(G)$ $(E(D))$ denote its vertex-set and edge-set (arc-set), respectively. We use $\{u, v\}$ to denote an edge in an undirected graph and uv to denote an arc in a digraph. The *open neighborhood*, or simply the *neighborhood*, of a vertex $v \in V(G)$ is defined as $N_G(v) = \{w \mid \{v, w\} \in E(G)\}$. The *closed neighborhood* of v is defined as $N_G[v] = N_G(v) \cup \{v\}$. The *degree* of v is defined as $d_G(v) = |N_G(v)|$. We can extend the definition of neighborhood of a vertex to a set of vertices as follows. Given a subset $U \subseteq V(G)$, $N_G(U) = \bigcup_{u\in U} N_G(u)$ and $N_G[U] = \bigcup_{u\in U} N_G[u]$. The *induced subgraph* $G[U]$ is the graph with vertex-set U and edge-set $\{\{u, u'\} \mid u, u' \in U$, and $\{u, u'\} \in E(G)\}$. Moreover, we define $G \setminus U$ as the induced subgraph $G[V(G) \setminus U]$. We omit subscripts when the graph G is clear from context. We use P_ℓ to denote a path in a graph on ℓ vertices. A path $P = uvw$ in a graph G is called *an induced* path if there is no edge between u and v in $E(G)$. An induced P_3-packing is a set of vertex disjoint induced P_3's. A subset X of $V(G)$ is called a *module* if every vertex in X has same set of neighbors in $V(G) \setminus X$. For a collection of graph $\mathcal{H}$, by $V(\mathcal{H})$ we denote $\bigcup_{H\in\mathcal{H}} V(H)$.

A tournament is a directed graph T such that for every pair of vertices $u, v \in V(T)$, exactly one of uv or vu is a directed arc of T. For any three vertices $x, y, z \in V(T)$, we say that xyz is a *triangle* if arcs xy, yz and zx form a directed cycle. A tournament in which there is no directed cycle is called a *transitive tournament*.

Reduction Rules. Kernelization algorithms often rely on the design of *reduction rules*. The rules are numbered, and each rule consists of a condition and an action. We always apply the first rule whose condition is true. Given a problem instance (I, k), the rule computes (in polynomial time) an instance (I', k') of the same problem where $k' \leq k$. Typically, $|I'| < |I|$, where if this is not the case, it should be argued why the rule can be applied only polynomially many times. We say that the rule *safe* if the instances (I, k) and (I', k') are equivalent.

3 Tool: Expansion Lemmas

In this section we give the classical Expansion Lemma as well as some two new Expansion Lemmas that we make use of in our kernels. We start with some preliminaries. Let ℓ be a positive integer. An *ℓ-star* is a graph on $\ell + 1$ vertices where one vertex, called the *center*, has degree ℓ, and all other vertices are adjacent to the center and have degree one. A *bipartite graph* is a graph whose vertex-set can be partitioned into two independent sets. Such a partition of the vertex-set is called a *bipartition* of the graph. Let G be a bipartite graph with bipartition (A, S). A subset of edges $M \subseteq E(G)$ is called *ℓ-expansion of S into A* if

(i) every vertex of S is incident to exactly ℓ edges of M,

(ii) and M saturates exactly $\ell|S|$ vertices in A.

Note that an ℓ-expansion saturates all vertices of S, and for each $u \in S$ the set of edges in M incident on u form an ℓ-star. The following lemma allows us to compute an ℓ-expansion in a bipartite graph. It captures a certain property of neighborhood sets which is very useful for designing kernelization algorithms.

Input: G with vertex bipartition (A, S).

Step 0: $\widehat{A} \leftarrow A$, $\widehat{S} \leftarrow S$

Step 1: Let H be a capacitated acyclic digraph H with vertex set $V(H) = \widehat{A} \cup \widehat{S} \cup \{s, t\}$ and arc set $E(H) = \{sx : x \in \widehat{S}\} \cup \{xy \in E(G) : x \in \widehat{S}, y \in \widehat{A}\} \cup \{yt : y \in \widehat{A}\}$, and a capacity function c of $E(H)$ as follows: $c(s, x) = \ell$ for every $x \in \widehat{S}$; $c(y, t) = 1$ for every $y \in \widehat{A}$, and $c(x, y) = 1$ for every $xy \in E(G)$.

Find a min-cut C of H separating H into $\{s\} \cup S_1 \cup A_1$ and $S_2 \cup A_2 \cup \{t\}$ with $S_1 \cup S_2 = \widehat{S}$ and $A_1 \cup A_2 = \widehat{A}$. If the size of C is less than $\ell\widehat{S}$, then set $\widehat{S} = S_2$ and $\widehat{A} = \widehat{A} \setminus (N_H(S_1) \cap \widehat{A})$ and repeat Step 1. Otherwise, stop the algorithm.

Output: $\widehat{A}, \widehat{S}$.

Figure 1: Algorithm to compute $\widehat{A}$ and $\widehat{S}$ Lemma 3.2.

Lemma 3.1 ([33, 9], Expansion Lemma)**.** *Let G be a bipartite graph with bipartition (A, S) such that there are no isolated vertices in A. Let ℓ be a positive integer such that $|A| \geq \ell|S|$. Then, there are non-empty subsets $X \subseteq S$ and $Y \subseteq A$ such that*

- *there is a ℓ-expansion from X into Y,*
- *and there is no vertex in Y that has a neighbor in $S \setminus X$, i.e. $N_G(Y) = X$.*

Further, the sets X and Y can be computed in polynomial time.

An alternate but an equivalent view on expansion properties is as follows. Let $\ell > 0$ and G be a bipartite graph with vertex sets A, S, and $\widehat{S} \subseteq S$ and $\widehat{A} \subseteq A$. We say that $\widehat{S}$ has an *ℓ-expansion* into $\widehat{A}$ in G if $|N_G(Y) \cap \widehat{A}| \geq \ell|Y|$ for every $Y \subseteq \widehat{S}$. We call this **(ex1)** property. Using the classical Hall's Theorem one can observe that this condition is equivalent to having for each $u \in S$, an ℓ-star, and that all of the ℓ-stars are pairwise vertex disjoint. In the next two lemmas, and in Sections 5 and 7 we will use this definition of expansion, while in the rest of the paper, we will use the classical definition of expansion.

Lemma 3.2 (New Expansion Lemma)**.** *Let ℓ be a positive integer and G be a bipartite graph with bipartition (A, S). Then, there exist $\widehat{S} \subseteq S$ and $\widehat{A} \subseteq A$ such that $\widehat{S}$ has an ℓ-expansion into $\widehat{A}$ in G, $N_G(\widehat{A}) \subseteq \widehat{S}$ and $|A \setminus \widehat{A}| \leq \ell|S \setminus \widehat{S}|$. Moreover, the sets $\widehat{S}$ and $\widehat{A}$ can be computed in polynomial time.*

The property that $N_G(\widehat{A}) \subseteq \widehat{S}$ will be called **(ex2)**. Lemma 3.2 is slightly different from Lemma 3.1, as it does not require $|A| \geq \ell|S|$ and that there is no isolated vertex in A, and thus $\widehat{A}$ and $\widehat{S}$ may be *empty*. However, we still have the bound on the number of removed vertices. That is, $|A \setminus \widehat{A}| < \ell|S \setminus \widehat{S}|$, and hence, if $|A| \geq \ell|S|$, then $\widehat{A}$ is nonempty. The difference between Lemmas 3.1 and Lemma 3.2 indeed comes from they viewpoint: in Lemmas 3.1, we obtain Y by only keeping "desired" vertices in A, while in Lemmas 3.1 we obtain $\widehat{A}$ by only removing "undesired" vertices from A. Thus, in Lemmas 3.1, if X is empty then Y is empty, while in Lemmas 3.1, it is possible that $\widehat{S}$ is empty but $\widehat{A}$ is large.

Proof. We first give the formal description of our algorithm in Figure 1. We now analyze one call to **Step 1**. Clearly, the weight of C is $|N_G(S_1) \cap A_2| + \ell|S_2| + |A_1| + |N_G(S_2) \cap A_1|$. If the weight of C is less than $\ell|\widehat{S}|$, then

$$\begin{aligned}
&|N_G(S_1) \cap A_2| + \ell|S_2| + |A_1| + |N_G(S_2) \cap A_1| < \ell|\widehat{S}| \\
\implies &|N_G(S_1) \cap A_2| + \ell|S_2| + |A_1| < \ell|\widehat{S}| \\
\implies &|N_G(S_1) \cap A_2| + |A_1| < \ell|\widehat{S}| - \ell|S_2| \\
\implies &|N_G(S_1) \cap (A_1 \cup A_2)| < \ell|S_1| \\
\implies &|N_G(S_1) \cap \widehat{A}| < \ell|S_1|,
\end{aligned}$$

which implies $|S_1| \geq 1$. Hence, $|\widehat{A} \cup \widehat{S}|$ is reduced after every call to **Step 1**, except the last call. The inequality above also implies that after each step, the number of vertices removed from $\widehat{A}$ is less than ℓ times the number of vertices removed from $\widehat{S}$. Note that in the last step, we do not remove anything. Thus, for the output $\widehat{A}, \widehat{S}$, we have $|A \setminus \widehat{A}| \leq \ell|S \setminus \widehat{S}|$ (where the equality is achieved only when $|A \setminus \widehat{A}| = \ell|S \setminus \widehat{S}| = 0$, i.e. the algorithm stops after just one call of Step 1).

If the size of C is at least $\ell|\widehat{S}|$, then the algorithm stops. Note that by Max-Flow-Min-Cut Theorem [20], the max-flow, which is equal to the min-cut size, is at most $\ell|\widehat{S}|$, which is the total capacity of arcs incident with s. Hence, there is a max-flow in H with optimal capacity $\ell|\widehat{S}|$. Then for every $Y \subseteq \widehat{S}$, the amount of flow passing from s through Y to t in H must be $\ell|Y|$. This gives $|N_H(Y)| \geq \ell|Y|$, and since $V(H) \cap A = \widehat{A}$, so $|N_G(Y) \cap \widehat{A}| \geq \ell|Y|$. Therefore, output $\widehat{A}, \widehat{S}$ satisfy Property **(ex1)**.

Throughout the algorithm we maintain an invariant that $N_G(\widehat{A}) \subseteq \widehat{S}$. Clearly, this holds at the beginning of the algorithm. Since every time we remove a set B_1, we remove all its neighbors in $\widehat{A}$, and so $N_G(\widehat{A}) \subseteq \widehat{S}$ holds at the end of every step, and so Property **(ex2)** is maintained during the algorithm. Thus Property **(ex2)** holds for the output $\widehat{A}, \widehat{S}$.

The algorithm runs in polynomial time, since $\widehat{A} \cup \widehat{S}$ is reduced each time we call **Step 1**, except the last call, and that each call to **Step 1** runs in polynomial time. □

As we discussed before, the viewpoint in the New Expansion Lemma is how many vertices are removed, rather than how many vertices remains as in the classical Expansion Lemma. This viewpoint enables us to generalize the Expansion Lemma to the Double expansion Lemma, where we can simultaneously achieve expansions in many graphs. In the following lemma, we consider a scheme where we have a "global" bipartite graph and d vertex-disjoint "local" bipartite graphs and we would like to achieve the expansion in each of them simultaneously.

Lemma 3.3 (Double Expansion Lemma)**.** *Let ℓ be a positive integer, and $G, H_1, \ldots, H_d$ be bipartite graphs with bipartition $(A, S), (A_1, S_1), \ldots, (A_d, S_d)$, respectively, such that $A_i \cap A_j = \emptyset, S_i \cap S_j = \emptyset$ for every $i \neq j$, and $\bigcup_{i=1}^d A_i = A, \bigcup_{i=1}^d S_i \subseteq S$. We can in polynomial time find $\widehat{A} \subseteq A, \widehat{S} \subseteq S, \widehat{A}_i \subseteq A_i, \widehat{S}_i \subseteq S_i$ for every i, satisfying the following*

- $\widehat{A} = \bigcup_{i=1}^d \widehat{A}_i$.
- $|A \setminus \widehat{A}| \leq 2\ell|S|$.
- *$\widehat{S}$ has an ℓ-expansion into $\widehat{A}$ in G, and for every i, $\widehat{S}_i$ has an ℓ-expansion into $\widehat{A}_i$ in H_i.*
- *$N_G(\widehat{A}) \subseteq \widehat{S}$, and for all $1 \leq i \leq d$, $N_{H_i}(\widehat{A}_i) \subseteq \widehat{S}_i$.*

Roughly speaking, the lemma asserts that we can find a set $\widehat{A}$ such that $\widehat{A}$ is the "image" of an expansion in the global graph, and the set of vertices $\widehat{A}_i$ in every local graph is the image of another expansion in that local graph. Since, $\widehat{A} = \bigcup_{i=1}^d \widehat{A}_i$, we achieve simultaneous expansion. Since $|A \setminus \widehat{A}| \leq 2\ell|S|$, we again have the property that if $|A| \geq 2\ell|S|$, then $\widehat{A}$ is non-empty. Note that, unlike $\widehat{A}$ and $\widehat{A}_i$, we do not have $\widehat{S} = \bigcup_{i=1}^d \widehat{S}_i$, or even $\widehat{S}_i \subseteq \widehat{S}$.

To prove Lemma 3.3, we repeatedly apply Lemma 3.2, alternately to the global graph and then to local graphs, and refine $\widehat{A}$ and $\bigcup_{i=1}^d \widehat{A}_i$ until they are equal.

Input: G, H_i for every i.

Step 0: Initialize $\widehat{A} \leftarrow A, \widehat{S} \leftarrow S, \widehat{A}_i \leftarrow A_i, \widehat{S}_i \leftarrow S_i$ for every i.

Stage 1: It consists of the following two steps.

> **Step 1:** Apply the New Expansion Lemma (Lemma 3.2) on $G[\widehat{A} \cup \widehat{S}]$ and get $S^\star \subseteq \widehat{S}$ and $A^\star \subseteq \widehat{A}$ satisfying the Expansion Lemma. Set $\widehat{S} \leftarrow S^\star, \widehat{A} \leftarrow A^\star$ and $\widehat{A}_i \leftarrow A^\star \cap \widehat{A}_i$ for every i (we do not update $\widehat{S}_i$). [*This is to ensure that* **(ex2)** *still holds.*]
>
> **Step 2:** For every i, apply the New Expansion Lemma (Lemma 3.2) on $H_i[\widehat{A}_i \cup \widehat{S}_i]$ and get $S_i^\star \subseteq \widehat{S}_i$ and $A_i^\star \subseteq \widehat{A}_i$ satisfying the New Expansion Lemma (Lemma 3.2). Set $\widehat{S}_i \leftarrow S_i^\star, \widehat{A}_i \leftarrow A_i^\star$ for every i, and $\widehat{A} \leftarrow \bigcup_i A_i^\star$ (we do not update $\widehat{S}$). [*Similarly, this is to ensure that* **(ex2)** *still holds.*]
>
> If at least one of $\widehat{A}, \widehat{S}, \widehat{A}_i, \widehat{S}_i$ changes, repeat **Stage 1**. Otherwise, stop the algorithm.

Output: $\widehat{A}, \widehat{S}, \widehat{A}_i, \widehat{S}_i$ for every i.

Figure 2: Algorithm to compute $\widehat{A}, \widehat{S}, \widehat{A}_i, \widehat{S}_i$ for every i in Lemma 3.3.

Proof. We first give the formal description of our algorithm in Figure 2 and an illustration in Figure 3.

Observe that each call of **Stage 1** runs in polynomial time. Towards this, note that the size of at least one of $\widehat{A}, \widehat{S}, \widehat{A}_i, \widehat{S}_i$ reduces after each call of **Stage 1** (except the last call), and since each step itself can be carried out in polynomial time, the algorithm itself runs in polynomial time. We will show that the output satisfies all the properties stated in the lemma.

The first property, $\widehat{A} = \bigcup_{i=1}^{d} \widehat{A}_i$ is vacous, since it is always maintained as an invariant during the algorithm. To prove the second property, observe that each time we call **Step 1**, we remove some vertices from $\widehat{S}$ and $\widehat{A}$. The number of vertices removed from $\widehat{A}$ at **Step 1** is at most ℓ times the number of vertices removed from $\widehat{S}$ at the same step (guaranteed by Lemma 3.2). Besides, initially $|\widehat{S}| = |S|$, so there are at most $|S|$ vertices removed from $\widehat{S}$ in all calls to **Step 1**. This implies that there are at most $\ell|S|$ vertices removed from $\widehat{A}$ in all calls to **Step 1**. Similarly, each time we call **Step 2**, we remove some vertices from $\bigcup_i \widehat{S}_i$ and $\bigcup_i \widehat{A}_i$. The number of vertices removed from $\bigcup_i \widehat{A}_i$ at **Step 2** is at most ℓ times the number of vertices removed from $\bigcup_i \widehat{S}_i$ at the same step. Besides, initially $\left|\bigcup_i \widehat{S}_i\right| \le |S|$, so there are at most $|S|$ vertices removed from $\bigcup_i \widehat{S}_i$ in all calls to **Step 2**. This implies that there are at most $\ell|S|$ vertices removed from $\bigcup_i \widehat{A}_i$ in all calls to **Step 2**, which is also exactly the number of vertices removed from $\widehat{A}$ in all calls to **Step 2**. In conclusion, there are at most $2\ell|S|$ vertices removed from $\widehat{A}$ during the algorithm, and so $|A \setminus \widehat{A}| < 2\ell|S|$.

To prove that $\widehat{S}$ and $\widehat{A}$ satisfies **(ex1)** and **(ex2)**, we first observe that $\widehat{S}$ and $\widehat{A}$ satisfies **(ex1)** after every **Step 1** of the algorithm, and so $\widehat{S}$ and $\widehat{A}$ satisfies **(ex1)** after **Step 2** if no vertex of $\widehat{A}$ is removed in that step. This means that the output $\widehat{S}$ and $\widehat{A}$ satisfies **(ex1)** since $\widehat{S}$ and $\widehat{A}$ are unchanged in the last stage. It remains to show that output $\widehat{S}$ and $\widehat{A}$ satisfies **(ex2)**. To do so, we prove by induction that $N_G(\widehat{A}) \subseteq \widehat{S}$ at the end of every stage. Clearly it is true at the beginning of the algorithm. Suppose that it is true at after **Stage** j (i.e. the j^{th} call of **Stage 1**), then there are no edge between $\widehat{A}$ and $S \setminus \widehat{S}$ in G. At **Step 1** of **Stage** $j+1$, we apply Lemma 3.2 on $G[\widehat{A} \cup \widehat{S}]$ and get $S^\star$ and $A^\star$ such that $N_{G[\widehat{A} \cup \widehat{S}]}(A^\star) \subseteq S^\star$, then there is no edge between $A^\star$ and $\widehat{S} \setminus S^\star$ in G. Thus there is no edge between $A^\star$ and $(S \setminus \widehat{S}) \cup (\widehat{S} \setminus S^\star)$ in G, i.e., $N_G(A^\star) \subseteq S^\star$. We then set $\widehat{A} \leftarrow A^\star, \widehat{S} \leftarrow S^\star$, and so $N_G(\widehat{A}) \subseteq \widehat{S}$ holds at the end of **Step 1** of **Stage** $j+1$. At **Step 2** of **Stage** $j+1$, some vertices are removed from $\widehat{A}$ while $\widehat{S}$ is unchanged, and hence $N_G(\widehat{A}) \subseteq \widehat{S}$ holds at the end of **Stage** $j+1$. This means that output

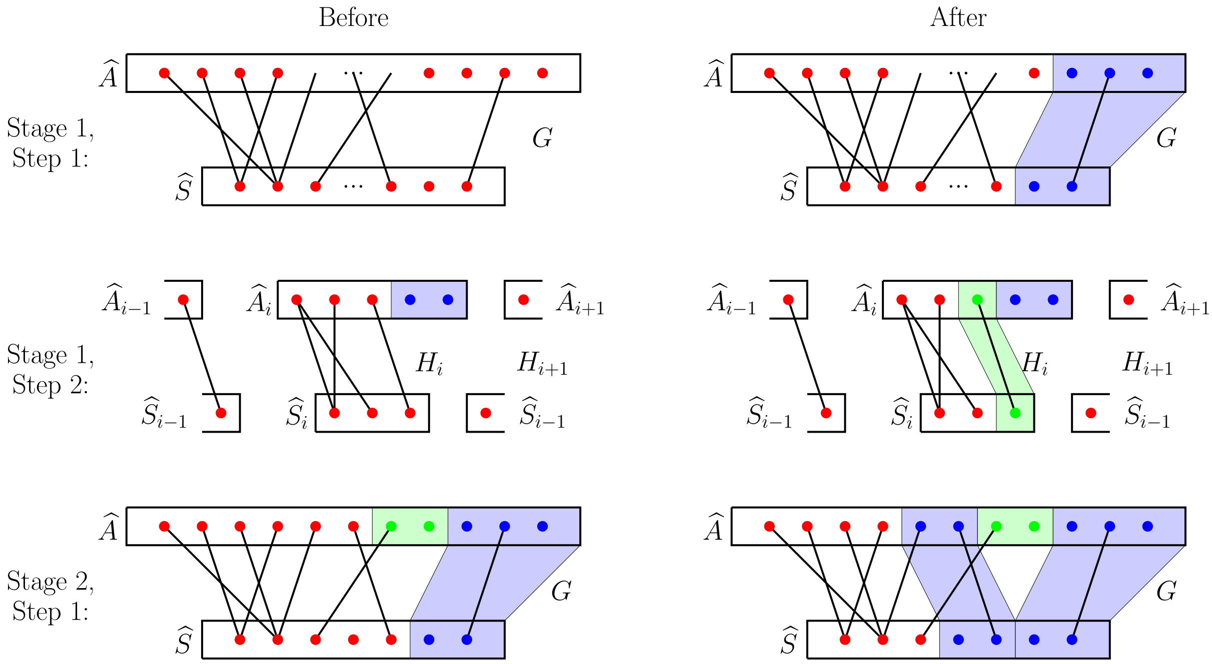

Figure 3: Illustration of the Double Expansion Algorithm – original vertices are red; a vertex turns blue if it is removed by a call of Step 1, and turns green if it is removed by a call of Step 2. Note that changing colors in S of G does not affect colors in S_i of H_i and vice versa. (Vertices in this figure are "well-ordered" to illustrate the algorithm).

$\widehat{S}$ and $\widehat{A}$ satisfies **(ex2)**.

Fix and integer $i \leq d$. To prove that $\widehat{S}_i$ and $\widehat{A}_i$ satisfies **(ex1)** and **(ex2)**, we first observe that $\widehat{S}_i$ and $\widehat{A}_i$ satisfies **(ex1)** after every execution of **Step 2** of the algorithm, and so the output $\widehat{S}_i$ and $\widehat{A}_i$ satisfies **(ex1)**. It remains to show that output $\widehat{S}_i$ and $\widehat{A}_i$ satisfies **(ex2)**. To do so, we prove by induction that $N_{H_i}(\widehat{A}_i) \subseteq \widehat{S}_i$ at the end of every stage. Clearly it is true at the beginning of the algorithm. Suppose that it is true after **Stage** j, then there are no edge between $\widehat{A}_i$ and $S_i \setminus \widehat{S}_i$ in H_i. At **Step 1** of **Stage** $j+1$, some vertices are removed from $\widehat{A}_i$ while $\widehat{S}_i$ is unchanged, then obviously $N_{H_i}(\widehat{A}_i) \subseteq \widehat{S}_i$ holds at the end of **Step 1** of **Stage** $j+1$. At **Step 2** of **Stage** $j+1$, we apply Lemma 3.2 on $H[\widehat{A}_i \cup \widehat{S}_i]$ and get $S_i^\star$ and $A_i^\star$ such that $N_{H_i[\widehat{A}_i \cup \widehat{S}_i]}(A_i^\star) \subseteq S_i^\star$, then there is no edge between $A_i^\star$ and $\widehat{S}_i \setminus S_i^\star$ in H_i. Thus there are no edge between $A_i^\star$ and $(S_i \setminus \widehat{S}_i) \cup (\widehat{S}_i \setminus S_i^\star)$ in H_i, i.e., $N_{H_i}(A_i^\star) \subseteq S_i^\star$. We then set $\widehat{A}_i \leftarrow A_i^\star, \widehat{S}_i \leftarrow S_i^\star$, and so $N_{H_i}(\widehat{A}_i) \subseteq \widehat{S}_i$ holds at the end of **Stage** $j+1$. This means that output $\widehat{S}$ and $\widehat{A}$ satisfies **(ex2)**. This concludes the proof of the lemma. □

We would like to remark that the Double Expansion Lemma can be generalized to the Triple Expansion Lemma (or η-levels Expansion Lemma), where the system contains a global bipartite graph G_i, local bipartite graphs H_i, and super-local bipartite graphs $H_{i,j}$. The proofs of these generalized version are similar to that of the Double Expansion Lemma. The idea of the Double Expansion Lemma (or its generlizations) is that one tries to capture different properties using different bipartite graphs at the same time.

4 Kernel for Cluster Vertex Deletion

In this section, we prove the following theorem.

Theorem 1. CVD *admits a kernel with* $\mathcal{O}(k^{1\frac{2}{3}})$ *vertices.*

Let (G, k) be an instance of CVD. Recall that CVD admits a polynomial-time 3-approximation algorithm. We call this algorithm with G as input, and thus we obtain a 3-approximate solution S. If $|S| > 3k$, then we conclude that (G, k) is a No-instance. Thus, we next assume that $|S| \leq 3k$. Notice that $G \setminus S$ is a collection of cliques, which we denote by $\mathcal{C}$.

In what follows, we denote α =2, β =1, γ =10, δ =3, λ =1 and η =1, so that $(1 - \frac{1}{\delta})\gamma \geq 2\eta$ (used in the proof of Lemma 4.11), $(\frac{1}{2} - \frac{1}{\delta})\gamma > (\frac{1}{(\alpha-1)\beta} + \lambda)$ (used in the proof of Lemma 4.13), and $\gamma \geq \frac{\delta}{\delta-1}(\frac{1}{(\alpha-1)\beta} + \lambda)$ (used in the proof of 4.14).

4.1 Bounding the Number of Cliques

First, we have the following simple rule, whose safeness is obvious.

Reduction Rule 4.1. *If there exists $C \in \mathcal{C}$ such that no vertex in C has a neighbor in S, then remove C from G. The new instance is $(G \setminus C, k)$.*

Now, we define the bipartite graph B by setting one side of the bipartition to be S and the other side to be $\mathcal{C}$,[5] such that there exists an edge between $s \in S$ and $C \in \mathcal{C}$ if and only if s is adjacent to at least one vertex in C. Note that by Reduction Rule 4.1, no clique in $\mathcal{C}$ is an isolated vertex in B. We thus proceed by presenting the following rule, where we rely on the Expansion Lemma (Lemma 3.1). It should be clear that the conditions required to apply the algorithm provided by this lemma are satisfied.

Reduction Rule 4.2. *If $|\mathcal{C}| \geq 2|S|$, then call the algorithm provided by Lemma 3.1 to compute sets $X \subseteq S$ and $\mathcal{Y} \subseteq \mathcal{C}$ such that X has a 2-expansion into $\mathcal{Y}$ in B and $N_B(\mathcal{Y}) \subseteq X$. The new instance is $(G \setminus X, k - |X|)$.*

We now argue that this rule is safe.

Lemma 4.1. *Reduction Rule 4.2 is safe.*

Proof. In one direction, it is clear that if $S^\star$ is a solution to $(G \setminus X, k - |X|)$, then $S^\star \cup X$ is a solution to (G, k). For the other direction, let $S^\star$ be a solution to (G, k). We denote $S' = (S^\star \setminus V(\mathcal{Y})) \cup X$. Notice that for all $s \in X$, there exists an induced P_3 in G of the form $u - s - v$ where u is any vertex in one clique associated to s by the 2-expansion that is adjacent to s and v is any vertex in the other clique associated to s by the 2-expansion that is adjacent to v. The existence of such u and v is implied by the definition of the edges of B. Thus, as $S^\star$ is a solution to (G, k), we have that $|X \setminus S^\star| \leq |S^\star \cap V(\mathcal{Y})|$, and hence $|S'| \leq |S^\star| \leq k$. Note that $G \setminus S'$ is a collection of isolated cliques together with a subgraph of $G \setminus S^\star$. Thus, as $G \setminus S^\star$ does not contain any induced P_3, we derive that $G \setminus S'$ also does not contain any induced P_3. We conclude that S' is a solution to (G, k), and as $X \subseteq S'$, we have that $S' \setminus X$ is a solution to $(G \setminus X, k - |X|)$. Thus, $(G \setminus X, k - |X|)$ is a Yes-instance. □

Due to Reduction Rule 4.2, from now on $|\mathcal{C}| \leq 6k$.

4.2 The Specification of the Marking Procedure

We proceed by presenting a procedure called Mark. Clearly, every vertex in S that has both a neighbor and a non-neighbor in a clique in $\mathcal{C}$ is a vertex due to which that clique in $\mathcal{C}$ is not a module. The procedure Mark accordingly associates vertices $s \in S$ with sets $\mathsf{mark}(s)$ of edges that belong to cliques in $\mathcal{C}$. In particular, we would ensure that for all $s \in S$, there would not exist two distinct edges $e, e' \in \mathsf{mark}(s)$ that have a common endpoint, as well as that for all distinct $s, s' \in S$, there would not exist two distinct edges $e \in \mathsf{mark}(s), e' \in \mathsf{mark}(s')$ that have a common endpoint.

[5] Here, we slightly abuse notation. Specifically, we mean that each clique in $\mathcal{C}$ is represented by a unique vertex in $V(B)$, and we refer to both the clique and the corresponding vertex identically.

Specification. The procedure Mark first initializes $M \Leftarrow \emptyset$, $T \Leftarrow S$, and for all $s \in S$, $\mathsf{mark}(s) \Leftarrow \emptyset$. At each *stage* i, $i = 1, 2, \ldots, k+1$, Mark executes the following process. For each $s \in T$, if there exist $C \in \mathcal{C}$ and $\{u,v\} \in E(C)$ such that $\{s,u\} \in E(G)$ but $\{s,v\} \notin E(G)$ and $\{u,v\} \cap M = \emptyset$, then insert u,v into M and $\{u,v\}$ into $\mathsf{mark}(s)$, and otherwise remove s from T. The order in which the process examines the vertices in T is immaterial given that it examines each vertex in T exactly once. Moreover, if $i = \lceil \beta k^{2/3} \rceil$, then the process sets U to be equal to T. If T is updated in subsequent stages, U is *not* updated as well.

We say that Mark *succeeded* if $|U| \geq \lceil \alpha k^{2/3} \rceil$, and otherwise we say that Mark failed. Moreover, if there exists $s \in S$ such that $|\mathsf{mark}(s)| \geq k+1$, then we say that Mark was *lucky*. Let us begin the analysis of Mark with the following simple lemma.

Lemma 4.2. *For any solution $S^\star$ to (G,k) and vertex $s \in S \setminus S^\star$, it holds that $S^\star \cap \{u,v\} \neq \emptyset$ for all $\{u,v\} \in \mathsf{mark}(s)$.*

Proof. Let $S^\star$ be a solution to (G,k). Consider some vertex $s \in S$ and edge $\{u,v\} \in \mathsf{mark}(s)$. Note that $\{s,u,v\}$ is the vertex set of an induced P_3 in G. Therefore, $S^\star \cap \{s,u,v\} \neq \emptyset$. We thus have that if $s \notin S^\star$, then $S^\star \cap \{u,v\} \neq \emptyset$. □

In light of Lemma 4.2, we employ the following rule.

Reduction Rule 4.3. *If there exists $s \in S$ such that $|\mathsf{mark}(s)| \geq k+1$ (i.e., Mark was lucky), then remove s from G and decrement k by 1. The new instance is $(G \setminus s, k-1)$.*

Lemma 4.3. *Reduction Rule 4.3 is safe.*

Proof. In one direction, it is clear that if $S^\star$ is a solution to $(G \setminus s, k-1)$, then $S^\star \cup \{s\}$ is a solution to (G,k). For the other direction, let $S^\star$ be a solution to (G,k). Observe that for all $s' \in S$ and $\{u,v\}, \{u',v'\} \in \mathsf{mark}(s')$, it holds that $\{u,v\} \cap \{u',v'\} = \emptyset$. Thus, by Lemma 4.2 and since $|\mathsf{mark}(s)| \geq k+1$, if $s \notin S^\star$ then $|S^\star| \geq k+1$, which is not possible as $|S^\star| \leq k$. We derive that $s \in S^\star$, and therefore $S^\star \setminus \{s\}$ is a solution to $(G \setminus s, k-1)$. □

The main purpose of Mark is to derive information on (G,k) also when it is not coincidentally lucky. More precisely, we have the following simple but useful lemma.

Lemma 4.4. *For any solution $S^\star$ to (G,k), $|U \setminus S^\star| \leq \frac{1}{\beta} k^{1/3}$.*

Proof. Let $S^\star$ be a solution to (G,k). Again, observe that for all $s \in S$ and $\{u,v\}, \{u',v'\} \in \mathsf{mark}(s)$, it holds that $\{u,v\} \cap \{u',v'\} = \emptyset$. In addition, observe that for all $s, s' \in S$, $\{u,v\} \in \mathsf{mark}(s)$ and $\{u',v'\} \in \mathsf{mark}(s')$, it holds that $\{u,v\} \cap \{u',v'\} = \emptyset$. Thus, by Lemma 4.2, $|S^\star| \geq \sum_{s \in U \setminus S^\star} |\mathsf{mark}(s)| \geq \lceil \beta k^{2/3} \rceil |U \setminus S^\star|$. Since $|S^\star| \leq k$, we conclude that $|U \setminus S^\star| \leq \frac{1}{\beta} k^{1/3}$. □

We also need to derive an upper bound on the number of marked vertices, namely $|M|$.

Lemma 4.5. *If Mark was neither lucky nor successful, then $|M| \leq 6(\alpha+\beta)k^{1\frac{2}{3}}$.*

Proof. Since Mark was unlucky, $|\mathsf{mark}(s)| \leq k$ for all $s \in S$. Thus, $|M| \leq 2|U|k + 2|S \setminus U|(\lceil \beta k^{2/3} \rceil - 1)$. Since Mark failed, we further have that $|M| \leq 2(\lceil \alpha k^{2/3} \rceil - 1)k + 6k(\lceil \beta k^{2/3} \rceil - 1) \leq 6(\alpha+\beta)k^{1\frac{2}{3}}$. □

4.3 Multiple Calls to the Marking Procedure

Let us now explain how we employ Mark. We initialize $\widehat{U} = \emptyset$ and $\widehat{G} = G$. Then, we call Mark with $(\widehat{G}, k)$ as input. If Mark was lucky, then we execute Reduction Rule 4.3 and restart the entire process (including the initialization phase). Else, if Mark succeeded, then for the set U computed by the current call, we update $\widehat{U} \Leftarrow \widehat{U} \cup U$ and $\widehat{G} \Leftarrow \widehat{G} \setminus U$, and then we proceed to execute another call. Otherwise, Mark was unlucky and also failed, and we let M denote the same set $M \subseteq V(G) \setminus S$ as computed by the *current call* to Mark, after which we terminate the process. Note that after each call to Mark, either Reduction Rule 4.3 is executed or the size of $\widehat{U}$ increases, and therefore it is clear that the process eventually terminates. We denote $L = S \setminus \widehat{U}$.

By relying on Lemma 4.4, we have the following lemma.

Lemma 4.6. *Let i be the number of calls to Mark that succeeded but were unlucky. For any solution $S^\star$ to (G, k), $|\widehat{U} \setminus S^\star| \leq i \cdot \frac{1}{\beta} k^{1/3}$ and $|S^\star \cap \widehat{U}| \geq i \cdot (\alpha \lceil k^{2/3} \rceil - \frac{1}{\beta} k^{1/3})$.*

Proof. First, note that $|S^\star \cap \widehat{U}| \geq i \cdot \alpha \lceil k^{2/3} \rceil - |\widehat{U} \setminus S^\star|$ as the sets U computed at distinct iterations are pairwise disjoint and the size of each one of them is at least $\alpha \lceil k^{2/3} \rceil$. Thus, it is sufficient to prove that $|\widehat{U} \setminus S^\star| \leq i \cdot \frac{1}{\beta} k^{1/3}$. However, this inequality follows from Lemma 4.4. $\square$

As a consequence of the two bounds in Lemma 4.6, we have the following corollary.

Corollary 4.1. *For any solution $S^\star$ to (G, k), $|\widehat{U} \setminus S^\star| \leq \frac{1}{(\alpha-1)\beta} k^{2/3}$.*

Proof. First, note that $k \geq |S^\star \cap \widehat{U}|$. Thus, by the second bound in Lemma 4.6, $k \geq i \cdot (\alpha \lceil k^{2/3} \rceil - \frac{1}{\beta} k^{1/3}) \geq i \cdot (\alpha k^{2/3} - \frac{1}{\beta} k^{1/3})$, which implies that $i \leq \dfrac{k}{\alpha k^{2/3} - \frac{1}{\beta} k^{1/3}} = \dfrac{k^{2/3}}{\alpha k^{1/3} - \frac{1}{\beta}} \leq \dfrac{1}{\alpha - 1} k^{1/3}$. By the first bound in Lemma 4.6, we thus derive that indeed $|\widehat{U} \setminus S^\star| \leq \frac{1}{(\alpha-1)\beta} k^{2/3}$. $\square$

The usefulness of Corollary 4.1 stems from the observation that it implies that we have found a (possibly large) set $\widehat{U} \subseteq S$ such that not only any $S^\star$ to (G, k) contains *almost all* the vertices in $\widehat{U}$, but also that the removal of $\widehat{U}$ from G significantly simplifies G as described by the following lemma.

Lemma 4.7. *For every clique $C \in \mathcal{C}$, $C[V(C) \setminus M]$ is a module in $G \setminus \widehat{U}$.*

Proof. Let C be a clique in $\mathcal{C}$. By the specification of Mark, for every vertex $s \in L$, it holds that there do not exist $u, v \in V(C) \setminus M$ such that $u \in N_G(s)$ and $v \notin N_G(s)$ (since $\{u, v\} \notin \mathsf{mark}(s)$). Furthermore, every vertex in C is adjacent to both u and v, and every vertex in a clique in $\mathcal{C} \setminus \{C\}$ is adjacent to neither u nor v. Thus, $C[V(C) \setminus M]$ is indeed a module in $G \setminus \widehat{U}$. $\square$

4.4 Sieving Bad Cliques

We sieve cliques based on three classifications. First, we say that a clique $C \in \mathcal{C}$ is *big* if $|V(C)| > \gamma k^{2/3}$, and otherwise it is small. Furthermore, we say that a clique $C \in \mathcal{C}$ is *huge* if $|V(C)| > 3k$. Recall that by Reduction Rule 4.2, $|\mathcal{C}| \leq 6k$. Thus, we directly have the following observation.

Observation 4.1. *The total number of vertices in small cliques in $\mathcal{C}$ is upper bounded by $6\gamma k^{1\frac{2}{3}}$.*

Second, we say that a clique $C \in \mathcal{C}$ is *heavy* if $|V(C) \cap M| \geq \frac{1}{\delta} |V(C)|$, and otherwise it is *light*. It is clear that the total number of vertices in heavy cliques in $\mathcal{C}$ is upper bounded by $\delta |M|$. Thus, by Lemma 4.5, we have the following observation.

Observation 4.2. *The total number of vertices in heavy cliques in $\mathcal{C}$ is upper bounded by $6\delta(\alpha+\beta)k^{1\frac{2}{3}}$.*

Third, for a clique $C \in \mathcal{C}$ and a vertex $s \in S$, we say that C is *visible* to s if $|N_G(s) \cap V(C)| \geq 2\eta k^{2/3}$, and otherwise we say that C is *hidden* from s. For a clique $C \in \mathcal{C}$, we let $\mathsf{vis}(C)$ denote that set of vertices in S to which C is visible. Moreover, we say that a clique $C \in \mathcal{C}$ is *visible* if $|\mathsf{vis}(C)| \geq \lambda k^{2/3}$, and otherwise we say that it is *hidden*. To bound the number of visible cliques, we need the following rule.

Reduction Rule 4.4. *If there exists a vertex $s \in S$ with at least $\frac{1}{2\eta}k^{1/3}+2$ cliques in $\mathcal{C}$ visible to s, then remove s from G and decrement k by 1. The new instance is $(G \setminus s, k-1)$.*

Lemma 4.8. *Reduction Rule 4.4 is safe.*

Proof. In one direction, it is clear that if $S^\star$ is a solution to $(G \setminus s, k-1)$, then $S^\star \cup \{s\}$ is a solution to (G,k). For the other direction, let $S^\star$ be a solution to (G,k). Let $\mathcal{A}$ denote the set of cliques in $\mathcal{C}$ that are visible to s. Since $|S^\star| \leq k$, $|\mathcal{A}| \geq \frac{1}{2\eta}k^{1/3}+2$ and by the definition of visibility, we have that there necessarily exist two distinct cliques $A, A' \in \mathcal{A}$ such that each clique among A, A' has a vertex that is a neighbor of s and does not belong to $S^\star$. Since these two vertices together with s form an induced P_3 in G, we derive that necessarily $s \in S^\star$. Therefore, $S^\star \setminus \{s\}$ is a solution to $(G \setminus s, k-1)$. □

After we exhaustively apply Reduction Rule 4.4, for every vertex $s \in S$ there exist at most $\frac{1}{2\eta}k^{1/3}+1 \leq \frac{1}{\eta}k^{1/3}$ cliques in $\mathcal{C}$ visible to s. Since $|S| \leq 3k$, we derive that there are at most $\dfrac{|S|\frac{1}{\eta}k^{1/3}}{\lambda k^{2/3}} = \dfrac{3}{\lambda\eta}k^{2/3}$ visible cliques. Thus, we have the following observation.

Observation 4.3. *The total number of vertices in non-huge visible cliques in $\mathcal{C}$ is upper bounded by $\frac{9}{\lambda\eta}k^{1\frac{2}{3}}$.*

Altogether, we say that a clique $C \in \mathcal{C}$ is *good* if it is **(i)** big, **(ii)** light and **(iii)** hidden or huge (or both), and otherwise we say that it is *bad*. We denote the set of all good cliques in $\mathcal{C}$ by $\mathcal{D}$. By Observations 4.1, 4.2 and 4.3, we derive the following lemma.

Lemma 4.9. *The total number of vertices in bad cliques in $\mathcal{C}$ is upper bounded by $9(\gamma+\delta(\alpha+\beta)+\frac{1}{\lambda\eta})k^{1\frac{2}{3}}$.*

4.5 Properties of Clique Sides

For all $C \in \mathcal{C}$ and $s \in S$, denote $N_C(s) = N_G(s) \cap V(C)$ and $\overline{N}_C(s) = V(C) \setminus N_C(s)$. Notice that for all $C \in \mathcal{C}$, $s \in S$, $u \in N_C(s)$ and $v \in \overline{N}_C(s)$, it holds that $s-u-v$ is an induced P_3 in G. Thus, we have the following observation.

Observation 4.4. *Let $S^\star$ be a solution to (G,k). Then, for all $C \in \mathcal{C}$ and $s \in S$, at least one of the following three condition holds:* **(i)** *$s \in S^\star$;* **(ii)** *$N_C(s) \subseteq S^\star$;* **(iii)** *$\overline{N}_C(s) \subseteq S^\star$.*

For all $C \in \mathcal{C}$ and $s \in S$, let $M_C(s)$ denote the set of minimum size among $N_C(s)$ and $\overline{N}_C(s)$ (if they have equal sizes, the choice is arbitrary). We first need to apply the following simple rule.

Reduction Rule 4.5. *If there exist $C \in \mathcal{C}$ and $s \in S$ such that $|M_C(s)| > k$, then remove s from G and decrement k by 1. The new instance is $(G \setminus s, k-1)$.*

Lemma 4.10. *Reduction Rule 4.5 is safe.*

Proof. In one direction, it is clear that if $S^\star$ is a solution to $(G \setminus s, k-1)$, then $S^\star \cup \{s\}$ is a solution to (G, k). For the other direction, let $S^\star$ be a solution to (G, k). Since $|S^\star| \leq k$ and $|M_C(s)| > k$, we have that both $N_C(s) \setminus S^\star \neq \emptyset$ and $\overline{N}_C(s) \setminus S^\star \neq \emptyset$. Thus, by Observation 4.4, we have that necessarily $s \in S^\star$. Therefore, $S^\star \setminus \{s\}$ is a solution to $(G \setminus s, k-1)$. ☐

Specifically, since for every $s \in S$ and huge clique $C \in \mathcal{C}$, $|M_C(s)| \leq k$, we have the following corollary, which exhibits a "vertex cover-like" interaction between S and huge cliques.

Observation 4.5. *Let $S^\star$ be a solution to (G, k). Then, for every $s \in S$ and huge clique $C \in \mathcal{C}$, at least one of the following two conditions holds:* **(i)** *$s \in S^\star$;* **(ii)** *$M_C(s) \subseteq S^\star$.*

Next, we prove that a similar result holds also for non-huge cliques given that they are good. To this end, we first prove the following simple lemma.

Lemma 4.11. *For all $s \in L$ and $C \in \mathcal{D}$ such that $N_G(s) \cap (V(C) \setminus M) \neq \emptyset$, it holds that C is visible to s.*

Proof. Let $s \in L$ and $C \in \mathcal{D}$ such that $N_G(s) \cap (V(C) \setminus M) \neq \emptyset$. Then, by Lemma 4.7, we have that $V(C) \setminus M \subseteq N_G(s)$. Thus, to prove that C is visible to s, it is sufficient to show that $|V(C) \setminus M| \geq 2\eta k^{2/3}$. Since $C \in \mathcal{D}$, we have that C is light, and therefore $|V(C) \setminus M| > (1 - \frac{1}{\delta})|V(C)|$. Moreover, since C is big, $|V(C)| > \gamma k^{2/3}$, and hence $|V(C) \setminus M| > (1 - \frac{1}{\delta})\gamma k^{2/3}$. Since $(1 - \frac{1}{\delta})\gamma \geq 2\eta$, the proof is completed. ☐

Lemma 4.12. *Let $S^\star$ be a solution to (G, k) of minimum size. Then, for every non-huge clique $C \in \mathcal{D}$, it holds that $|V(C) \cap S^\star| \leq |V(C) \cap M| + (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$.*

Proof. Let $C \in \mathcal{D}$ be a non-huge clique. Suppose, by way of contradiction, that $|V(C) \cap S^\star| > |V(C) \cap M| + (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$. Define $S' = (S^\star \setminus V(C)) \cup \widehat{U} \cup (V(M) \cap V(C)) \cup \mathsf{vis}(C)$. By Corollary 4.1 and since C is a non-huge clique in $\mathcal{D}$, $|\widehat{U} \setminus S^\star| \leq \frac{1}{(\alpha-1)\beta}k^{2/3}$ and $|\mathsf{vis}(C)| \leq \lambda k^{2/3}$. Thus, $|S'| < |S^\star| \leq k$. Next we show that $(V(C)) \setminus S'$ is an isolated clique. Towards this we will show that it has no neighbor in the approximate solution S. The only possible neighbors of $(V(C)) \setminus S'$ in S are in L. However, if there exists a vertex $s \in L$ such that $N_G(s) \cap (V(C) \setminus M) \neq \emptyset$, then by Lemma 4.11, it holds that C is visible to s. This implies that a vertex $s \in L$ is either in $\mathsf{vis}(C)$ or $N(s) \cap V(C) \subseteq M \cap V(C)$. Since, S' contains $M \cap V(C)) \cup \mathsf{vis}(C)$ we have that $(V(C)) \setminus S'$ is an isolated clique. Thus, by Lemma 4.11, the graph $G \setminus S'$ consists of an isolated clique on the vertex set $(V(C)) \setminus S'$ and a subgraph of $G \setminus S^\star$. Therefore, as $G \setminus S^\star$ does not contain any induced P_3, so does $G \setminus S'$. This implies that S' is a solution to (G, k), but since $|S'| < |S^\star|$, we obtain a contradiction to the choice of $S^\star$. ☐

Lemma 4.13. *Let $S^\star$ be a solution to (G, k) of minimum size. Then, for every $s \in S$ and non-huge clique $C \in \mathcal{D}$, at least one of the following two conditions holds:* **(i)** *$s \in S^\star$;* **(ii)** *$M_C(s) \subseteq S^\star$.*

Proof. Let s be a vertex in S, and let $C \in \mathcal{D}$ be a non-huge clique. Suppose, by way of contradiction, that neither $s \in S^\star$ nor $M_C(s) \subseteq S^\star$. By Observation 4.4, we necessarily have that $V(C) \setminus M_C(s) \subseteq S^\star$. Thus, $|V(C) \cap S^\star| \geq |V(C) \setminus M_C(s)| \geq \frac{1}{2}|V(C)|$. Therefore, to obtain a contradiction to Lemma 4.12, it is sufficient to show that $\frac{1}{2}|V(C)| > |V(C) \cap M| + (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$. Since C is light, we have that $|V(C) \cap M| < \frac{1}{\delta}|V(C)|$, and therefore it remains to show that $(\frac{1}{2} - \frac{1}{\delta})|V(C)| > (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$. Since C is big, $|V(C)| > \gamma k^{2/3}$. Thus, we only need to show that $(\frac{1}{2} - \frac{1}{\delta})\gamma > (\frac{1}{(\alpha-1)\beta} + \lambda)$, which follows from the definition of $\alpha, \beta, \gamma, \delta$ and λ. ☐

4.6 Expansion with Respect to Clique Sides

We construct the bipartite graph B' by setting one side of the bipartition to be S and the other side Q' to be the set of vertices in good cliques (i.e., $Q' = \bigcup_{C \in \mathcal{D}} V(C)$), such that there exists an edge between $s \in S$ and $v \in Q'$ if and only if $v \in M_D(s)$ where D is the clique in $\mathcal{D}$ containing v. Let I denote the set of isolated vertices in B' that belong to Q', and denote $Q = Q' \setminus I$. Moreover, define $B = B' \setminus I$. Clearly, no clique in Q is an isolated vertex in B. We thus proceed by presenting the following rule, where we rely on the Expansion Lemma (Lemma 3.1). It should be clear that the conditions required to apply the algorithm provided by this lemma are satisfied.

Reduction Rule 4.6. *If $|Q| \geq (\frac{1}{(\alpha-1)\beta}k^{2/3}+1)|S|$, then call the algorithm provided by Lemma 3.1 to compute sets $X \subseteq S$ and $Y \subseteq Q$ such that X has a $(\frac{1}{(\alpha-1)\beta}k^{2/3}+1)$-expansion into Y in B and $N_B(Y) \subseteq X$. The new instance is $(G \setminus X, k-|X|)$.*

We now argue that this rule is safe.

Lemma 4.14. *Reduction Rule 4.6 is safe.*

Proof. In one direction, it is clear that if $S^\star$ is a solution to $(G \setminus X, k-|X|)$, then $S^\star \cup X$ is a solution to (G,k). For the other direction, let $S^\star$ be a solution to (G,k) of minimum size. Define $S' = (S^\star \setminus Y) \cup X \cup \widehat{U}$. First, due to Corollary 4.1, note that $|\widehat{U} \setminus U^\star| \leq \frac{1}{(\alpha-1)\beta}k^{2/3}$. Moreover, by Observation 4.5 and Lemma 4.13, for every vertex $s \in S \setminus S^\star$, it holds that $N_B(s) \subseteq S^\star$. Thus, since X has a $(\frac{1}{(\alpha-1)\beta}k^{2/3}+1)$-expansion into Y in B, we have that $|Y \setminus S^\star| \leq (\frac{1}{(\alpha-1)\beta}k^{2/3}+1)|X \setminus S^\star|$. This implies that $|S'| \leq |S^\star| \leq k$.

Notice that if $G \setminus S'$ does not contain any induced P_3, then since $X \subseteq S^\star$ and we have shown that $|S'| \leq k$, this would imply that S' is a solution to $(G \setminus X, k-|X|)$. Suppose, by way of contradiction, that $G \setminus S'$ contains some induced P_3, which we denote by W. Note that $V(W) \cap (X \cup \widehat{U}) = \emptyset$. Since $G \setminus S^\star$ does not contain any induced P_3 and since S is an approximate solution, we also derive that $V(W) \cap Y \neq \emptyset$ and $V(W) \cap S \neq \emptyset$. Accordingly, the following case analysis is exhaustive.

- **Case 1:** $W = s - u - v$ where $s \in S \setminus (X \cup \widehat{U})$, $u, v \in V(C)$ for some $C \in \mathcal{D}$ and $\{u,v\} \cap Y \neq \emptyset$. In this case, let $y \in \{u,v\}$ denote some vertex in $\{u,v\} \cap Y$ and let x denote the other vertex in $\{u,v\}$ (which might also be in Y). Since $N_B(Y) \subseteq X$ and $s \notin X$, we have that $y \notin N_B(s)$. Since $u \in N_C(s)$ and $v \in \overline{N}_C(s)$, we have that $|M_C(s) \cap \{u,v\}| = 1$. Since $y \notin N_B(s)$, we have that $y \notin M_C(s)$ and $x \in M_C(s)$. In particular, as $N_B(Y) \subseteq X$ and $s \notin X$, we have that $x \in N_B(s) \setminus Y$ (in fact, $N_B(s) \cap Y = \emptyset$). By Observation 4.5 and Lemma 4.13, we derive that $S^\star \cap \{s,x\} \neq \emptyset$. However, as $x \notin Y$, this implies that $S' \cap \{s,x\} \neq \emptyset$, which is a contradiction.

- **Case 2:** $W = s - v - s'$ where $s, s' \in S \setminus (X \cup \widehat{U})$, and $v \in V(C) \cap Y$ for some $C \in \mathcal{D}$. Since $N_B(Y) \subseteq X$ and $s, s' \notin X$, we have that $v \notin M_C(s) \cup M_C(s')$, which means that $N_C(s) = V(C) \setminus M_C(s)$ and $N_C(s') = V(C) \setminus M_C(s')$. Therefore, $|N_C(s) \cap V(C)|, |N_C(s') \cap V(C)| \geq \frac{1}{2}|V(C)|$. Thus, since $|M \cap V(C)| < \frac{1}{\delta}|V(C)| < \frac{1}{2}|V(C)|$ (because $C \in \mathcal{D}$), we have that there exist $w \in N_C(s) \setminus M$ and $w' \in N_C(s') \setminus M$. By Lemma 4.7 and since $s, s' \notin \widehat{U}$, we derive that C is huge or both $V(C) \setminus M \subseteq N_C(s)$ and $V(C) \setminus M \subseteq N_C(s')$.

  Let us first consider the subcase where C is huge. Due to Reduction Rule 4.5, we have that $|\overline{N}_C(s)|, |\overline{N}_C(s')| \leq k$. Since $|V(C)| > 3k$, we derive that $|N_C(s) \cap N_C(s')| \geq k+1$. Note that any vertex $w \in N_C(s) \cap N_C(s')$, along with s and s', forms the induced P_3 in G that is $s - w - s'$. Note that $s, s' \notin S^\star$, as otherwise $\{s,s'\} \cap S' \neq \emptyset$, which contradicts the choice of W. Thus, as $S^\star$ is a solution to (G,k), it must hold that $N_C(s) \cap N_C(s') \subseteq S^\star$, but as $|N_C(s) \cap N_C(s')| \geq k+1$, this is a contradiction.

Let us now consider the subcase where C is not huge, and in particular $V(C) \setminus M \subseteq N_C(s)$ and $V(C) \setminus M \subseteq N_C(s')$. Then, any vertex $w \in V(C) \setminus M$, along with s and s', forms the induced P_3 in G that is $s - w - s'$. Again, note that $s, s' \notin S^\star$. Thus, since $S^\star$ is a solution to (G,k), we have that $V(C) \setminus M \subseteq S^\star$. Now, recall that by Lemma 4.12 and since $C \in \mathcal{D}$ is not huge, $|V(C) \cap S^\star| \leq |V(C) \cap M| + (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$. Hence, $|V(C) \setminus M| \leq (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$. As $|V(C) \cap M| < \frac{1}{\delta}|V(C)|$ (because $C \in \mathcal{D}$), we have that $(1 - \frac{1}{\delta})|V(C)| < (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$, and hence $|V(C)| < \frac{\delta}{\delta-1}(\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$, which is a contradiction, since $C \in \mathcal{D}$ implies that C is in particular big and $\gamma \geq \frac{\delta}{\delta-1}(\frac{1}{(\alpha-1)\beta} + \lambda)$.

- **Case 3:** $W = v - s - s'$ where $s, s' \in S \setminus (X \cup \widehat{U})$, and $v \in V(C) \cap Y$ for some $C \in \mathcal{D}$. The analysis of this case is similar to the one of the previous case, and is only given for completeness. Since $N_B(Y) \subseteq X$ and $s, s' \notin X$, we have that $v \notin M_C(s) \cup M_C(s')$, which means that $N_C(s) = V(C) \setminus M_C(s)$ and $\overline{N}_C(s') = V(C) \setminus M_C(s')$. Therefore, $|N_C(s) \cap V(C)|, |\overline{N}_C(s') \cap V(C)| \geq \frac{1}{2}|V(C)|$. Thus, since $|M \cap V(C)| < \frac{1}{\delta}|V(C)| < \frac{1}{2}|V(C)|$ (because $C \in \mathcal{D}$), we have that there exist $w \in N_C(s) \setminus M$ and $w' \in \overline{N}_C(s') \setminus M$. By Lemma 4.7 and since $s, s' \notin \widehat{U}$, we derive that C is huge or both $V(C) \setminus M \subseteq N_C(s)$ and $V(C) \setminus M \subseteq \overline{N}_C(s')$.

  Let us first consider the subcase where C is huge. Due to Reduction Rule 4.5, we have that $|\overline{N}_C(s)|, |N_C(s')| \leq k$. Since $|V(C)| > 3k$, we derive that $|N_C(s) \cap \overline{N}_C(s')| \geq k+1$. Note that any vertex $w \in N_C(s) \cap \overline{N}_C(s')$, along with s and s', forms the induced P_3 in G that is $w - s - s'$. Note that $s, s' \notin S^\star$, as otherwise $\{s, s'\} \cap S' \neq \emptyset$, which contradicts the choice of W. Thus, as $S^\star$ is a solution to (G,k), it must hold that $N_C(s) \cap \overline{N}_C(s') \subseteq S^\star$, but as $|N_C(s) \cap \overline{N}_C(s')| \geq k+1$, this is a contradiction.

  Let us now consider the subcase where C is not huge, and in particular $V(C) \setminus M \subseteq N_C(s)$ and $V(C) \setminus M \subseteq \overline{N}_C(s')$. Then, any vertex $w \in V(C) \setminus M$, along with s and s', forms the induced P_3 in G that is $w - s - s'$. Again, note that $s, s' \notin S^\star$. Thus, since $S^\star$ is a solution to (G,k), we have that $V(C) \setminus M \subseteq S^\star$. Now, recall that by Lemma 4.12 and since $C \in \mathcal{D}$ is not huge, $|V(C) \cap S^\star| \leq |V(C) \cap M| + (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$. Hence, $|V(C) \setminus M| \leq (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$. As $|V(C) \cap M| < \frac{1}{\delta}|V(C)|$ (because $C \in \mathcal{D}$), we have that $(1 - \frac{1}{\delta})|V(C)| < (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$, and hence $|V(C)| < \frac{\delta}{\delta-1}(\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$, which is a contradiction, since $C \in \mathcal{D}$ implies that C is in particular big and $\gamma \geq \frac{\delta}{\delta-1}(\frac{1}{(\alpha-1)\beta} + \lambda)$.

- **Case 4:** $W = u - s - v$ where $s \in S \setminus (X \cup \widehat{U})$, $u \in V(C) \cap Y$ for some $C \in \mathcal{D}$, and $v \in V(C')$ for some $C' \in \mathcal{C} \setminus \{C\}$. Since $N_B(Y) \subseteq X$ and $s \notin X$, we have that $u \notin M_C(s)$, which means that $N_C(s) = V(C) \setminus M_C(s)$. Therefore, $|N_C(s) \cap V(C)| \geq \frac{1}{2}|V(C)|$. Thus, since $|M \cap V(C)| < \frac{1}{\delta}|V(C)| < \frac{1}{2}|V(C)|$ (because $C \in \mathcal{D}$), we have that there exists $w \in N_C(s) \setminus M$. By Lemma 4.7 and since $s \notin \widehat{U}$, we derive that C is huge or $V(C) \setminus M \subseteq N_C(s)$. Symmetrically, we derive that *if* $v \in Y$, then C' is huge or $V(C') \setminus M \subseteq N_{C'}(s)$.

  Note that for all $w \in N_C(s)$ and $w' \in N_{C'}(s)$, it holds that $w - s - w'$ is an induced P_3 in G. As $s \notin S^\star$ (as otherwise $s \in S'$), we have that $N_C(s) \subseteq S^\star$ or $N_{C'}(s) \subseteq S^\star$. Observe that if $v \notin Y$, then since $v \notin S'$, we have that $v \notin S^\star$ and therefore it clearly holds that $N_{C'}(s) \nsubseteq S^\star$. If $v \in Y$ (which means that $C' \in \mathcal{D}$), then the proof that $N_{C'}(s) \nsubseteq S^\star$ is symmetric to the proof that $N_C(s) \nsubseteq S^\star$. Therefore, in what follows, we only show that $N_C(s) \nsubseteq S^\star$.

  Let us first consider the subcase where C is huge. Due to Reduction Rule 4.5, we have that $|\overline{N}_C(s)| \leq k$. Since $|V(C)| > 3k$, we derive that $|N_C(s)| \geq 2k+1$. Since $|S^\star| \leq k$, it is then clear that $N_C(s) \not\subseteq S^\star$. Now, let us now consider the subcase where C is not huge, and in particular $V(C) \setminus M \subseteq N_C(s)$. Recall that by Lemma 4.12 and since $C \in \mathcal{D}$ is not huge, $|V(C) \cap S^\star| \leq |V(C) \cap M| + (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$. Suppose, by way of contradiction, that $N_C(s) \subseteq S^\star$. Then, $|V(C) \setminus M| \leq (\frac{1}{(\alpha-1)\beta} + \lambda)k^{2/3}$, which leads to a contradiction as in the previous two cases.

Since each case led to a contradiction, the proof is complete. □

4.7 Reduction of Almost Modules

At this point, it remains to bound the size of I. We first show that the sets of vertices in I, defined according to the cliques in $\mathcal{C}$, are modules also with respect to $\widehat{U}$. More precisely, we prove the following lemma.

Lemma 4.15. *For every clique $C \in \mathcal{D}$, $C[I \cap V(C)]$ is a module in G.*

Proof. Let C be a clique in $\mathcal{D}$. Consider two vertices $u, v \in I \cap V(C)$. Clearly, every vertex in C is adjacent to both u and v, and every vertex in a clique in $\mathcal{C} \setminus \{C\}$ is adjacent to neither u nor v. Thus, $C[I \cap V(C)]$ is indeed a module in $G \setminus S$. Now, consider some vertex $s \in S$. Then, as $u, v \in I \cap V(C)$, we have that $u, v \in V(C) \setminus M_C(s)$, because otherwise u or v would have been adjacent to s in the bipartite graph B of Section 4.6. Thus, we have that either both $u, v \in N_C(s)$ or both $u, v \in \overline{N}_C(s)$. As the choices of u, v and s were arbitrary, we conclude that $C[I \cap V(C)]$ is indeed a module in G. □

We now present a rule that concerns the set I.

Reduction Rule 4.7. *If there exists a visible clique $C \in \mathcal{D}$ such that $|I \cap V(C)| > k + 1$ or a hidden clique $C \in \mathcal{D}$ such that $|I \cap V(C)| > |M \cap V(C)| + \frac{1}{(\alpha-1)\beta}k^{2/3} + \lambda k^{2/3}$, then remove an arbitrarily chosen vertex $v \in V(C) \cap I$ from G. The new instance is $(G \setminus v, k)$.*

Lemma 4.16. *Reduction Rule 4.7 is safe.*

Proof. In one direction, it is clear that if (G, k) has a solution, so does $(G \setminus v, k)$. Now, let $S^\star$ be a solution to $(G \setminus v, k)$. If $S^\star$ is also a solution to (G, k), then the proof is complete. Therefore, we next assume that $S^\star$ is not a solution to (G, k). Then, there exists an induced P_3, denoted by W, in $G \setminus S^\star$. Since $S^\star$ is a solution to $(G \setminus v, k)$, $v \in V(W)$. Furthermore, since $v \in V(C) \cap I$ and $I \cap V(C)$ is a clique that is a module (by Lemma 4.15), for any vertex $u \in I \cap V(C)$, the vertex set $(V(W) \setminus \{v\}) \cup \{u\}$ induces a P_3 in $G \setminus v$. As $(V(W) \setminus \{v\}) \cap S^\star = \emptyset$ and $S^\star$ is a solution to $(G \setminus v, k)$, we deduce that $(I \cap V(C)) \setminus \{v\} \subseteq S^\star$.

In case $|I \cap V(C)| > k + 1$, the conclusion that $(I \cap V(C)) \setminus \{v\} \subseteq S^\star$ implies that $|S^\star| > k$, which is a contradiction. Now, suppose that C is a hidden clique in $\mathcal{D}$ such that $|I \cap V(C)| > |M \cap V(C)| + \frac{1}{(\alpha-1)\beta}k^{2/3} + \lambda k^{2/3}$. Let us denote $S' = (S^\star \setminus (I \cap V(C))) \cup (M \cap V(C)) \cup \widehat{U} \cup \mathsf{vis}(C)$. By Corollary 4.1 and since C is a hidden clique in $\mathcal{D}$, we have that $|S'| \leq |S^\star| - |I \cap V(C)| + |M \cap V(C)| + \frac{1}{(\alpha-1)\beta}k^{2/3} + \lambda k^{2/3} \leq |S^\star| \leq k$. Moreover, by Lemma 4.11, the graph $G \setminus S'$ consists of an isolated clique on the vertex set $V(C) \setminus S'$ (for a detail argument see the proof of Lemma 4.12) and a subgraph of $(G \setminus v) \setminus S^\star$. Therefore, as $(G \setminus v) \setminus S^\star$ does not contain any induced P_3, so does $G \setminus S'$. This implies that S' is a solution to (G, k), and therefore (G, k) is a Yes-instance. □

Finally, after the exhaustive application of Reduction Rule 4.7, we can bound the size of I.

Lemma 4.17. *After the exhaustive application of Reduction Rule 4.7, $|I| \leq 6(\frac{1}{\lambda\eta} + \alpha + \beta + \frac{1}{(\alpha-1)\beta} + \lambda)k^{1\frac{2}{3}}$.*

Proof. First, note that after the exhaustive application of Reduction Rule 4.7, every visible clique $C \in \mathcal{D}$ satisfies $|V(C) \cap I| \leq k+1$ and every hidden clique $C \in \mathcal{D}$ satisfies $|V(C) \cap I| \leq |M \cap V(C)| + \frac{1}{(\alpha-1)\beta}k^{2/3} + \lambda k^{2/3}$. Recalling that the number of visible cliques is upper bounded by $\frac{3}{\lambda\eta}k^{2/3}$, we have that the total number of vertices in I that belong to visible cliques in $\mathcal{D}$ is upper bounded by $\frac{3}{\lambda\eta}k^{2/3} \cdot (k+1)$. Now, recalling that $|\mathcal{C}| \leq 6k$, we also have that the total number of vertices in I that belong to hidden cliques in $\mathcal{D}$ is upper bounded by $|M| + 6(\frac{1}{(\alpha-1)\beta} + \lambda)k^{1\frac{2}{3}}$. By Lemma 4.5, $|M| \leq 6(\alpha+\beta)k^{1\frac{2}{3}}$. Thus, $\frac{3}{\lambda\eta}k^{2/3}(k+1) + |M| + 6(\frac{1}{(\alpha-1)\beta} + \lambda)k^{1\frac{2}{3}} \leq 6(\frac{1}{\lambda\eta} + \alpha + \beta + \frac{1}{(\alpha-1)\beta} + \lambda)k^{1\frac{2}{3}}$, which completes the proof. $\square$

4.8 Proof of Theorem 1

We are finally ready to present the proof of Theorem 1.

Proof of Theorem 1. Let (G,k) be an instance of CVD. Our kernelization algorithm simply applies (exhaustively) Reduction Rules 4.1 to 4.7. The output is the instance obtained once none of these rules is applicable. Let us observe that each rule among Reduction Rules 4.1 to 4.16 can be applied in polynomial time, it strictly decreases the size of G and it does not increase k. Thus, our kernelization algorithm runs in polynomial time.

For the sake of clarity, let us now abuse notation and denote the outputted instance by (G,k). Let us observe that $V(G)$ consists of the following vertices.

- Vertices in S, whose number is at most $3k$.
- Vertices in bad cliques, whose number is at most $9(\gamma + \delta(\alpha+\beta) + \frac{1}{\lambda\eta})k^{1\frac{2}{3}} = \mathcal{O}(k^{1\frac{2}{3}})$ (by Lemma 4.9).
- Vertices in good cliques that are not isolated in B', whose number is at most $(\frac{1}{(\alpha-1)\beta}k^{2/3} + 1)|S| = \mathcal{O}(k^{1\frac{2}{3}})$ (due to Reduction Rule 4.6).
- Vertices in the set I, whose number is at most $6(\frac{1}{\lambda\eta} + \alpha + \beta + \frac{1}{(\alpha-1)\beta} + \lambda)k^{1\frac{2}{3}} = \mathcal{O}(k^{1\frac{2}{3}})$ (by Lemma 4.17).

Thus, the total number of vertices is indeed $\mathcal{O}(k^{1\frac{2}{3}})$. This completes the proof. $\square$

5 Feedback Vertex Set in Tournaments

In this section, we prove the following theorem.

Theorem 2. FVST *admits a kernel with $\mathcal{O}(k^{3/2})$ vertices.*

To prove Theorem 2, we will also use the following folklore result.

Proposition 5.1. *Let T be a tournament. Then, the following conditions hold.*

1. *T has a directed cycle if and only if T has a directed triangle.*
2. *If T is acyclic then it has a unique topological ordering. That is, there exists a unique ordering $\prec$ of the vertices of T such that for every directed arc uv, we have $u \prec v$ (that is, u appears before v in the ordering $\prec$).*

Let (T,k) be an instance of FVST. By Proposition 5.1, to find a set S such that $T \setminus S$ is a directed acyclic graph, it is sufficient to find a set that intersects all the triangles of T. This immediately yields a simple polynomial-time 3-approximation algorithm for FVST. Indeed, start by greedily finding a maximal collection, say $\mathcal{S}$, of vertex-disjoint triangles in T and output $V(\mathcal{S})$. We call this algorithm with T as input, and obtain a 3-approximate solution S. If $|S| > 3k$, then we conclude that (T, k) is a No-instance. Hence, we assume that $|S| \leq 3k$. We call the vertex set S such that $G \setminus S$ does not have any directed cycle as *feedback vertex set*. Let $X = T \setminus S$. Note that since S is a feedback vertex set, X is a *transitive tournament*. Let (T, h) be an instance of FVST. We say that a feedback vertex set of of size at most h of T is a solution to the instance (T, h). For the sake of clarity of the analysis, we omit floor/ceiling signs and remainders whenever they are not crucial.

We have the following simple rule, whose safeness can be easily observed.

Reduction Rule 5.1. *If there exists $s \in S$ such that there are $k+1$ triangles intersecting pairwise (only) at s, then remove s from T. The new instance is $(T \setminus \{s\}, k-1)$.*

We apply Reduction Rule 5.1 exhaustively. Note that each application can be performed in polynomial time, as for any vertex $s \in S$, we can check whether that exist $k+1$ triangles intersecting pairwise (only) at s as follows: we construct a bipartite graph where one side of the bipartition is the set A of in-neighbors of s, the other side of the bipartition is the set B of out-neighbors of s, and there exists an edge between $a \in A$ and $b \in B$ if and only if a is an out-neighbor of b; then, there exist $k+1$ triangles intersecting pairwise (only) at s if and only if the size of a maximum matching in this bipartite graph is at least $k+1$ (which can be checked in polynomial time). Thus, from now onwards, we assume that Reduction Rule 5.1 is no longer applicable. Throughout this section, we work with the unique ordering $\prec$ of the vertices of X. For example, whenever we will use a phrase such as *the vertices are consecutive in* X, we mean that the vertices occur consecutively with respect to the ordering $\prec$. Similarly, we define the notion of the *smallest* and the *largest* vertex in X according to the ordering $\prec$.

5.1 Exploring the Vertex Cover Structure

Let us now define a notion of *vertex cover* for a set of arcs of T. Formally, for a subset of arcs $A \subseteq E(T)$, a subset $O \subseteq V(T)$ is called a *vertex cover* for A if for every arc $uv \in A$, either $u \in O$ or $v \in O$ (or both). An arc xy of T is called *strong* if **(i)** at least one vertex among x and y belongs to S, and **(ii)** there are at least $k+2$ vertices $z \in V(T)$ such that xyz is a triangle. Let F be the set of all the strong arcs of T, which can be easily found in polynomial time. We start our analysis with the following simple observation regarding the set F.

Observation 5.1. *If O is a solution to $(T, k+1)$, then O is a vertex cover of F.*

The proof is simple: if O does not hit $xy \in F$, then O contains all $z \in V(T)$ such that xyz is a triangle, i.e. $|O| \geq k+2$, which is a contradiction.

Recall that throughout our kernelization algorithm, we work with the unique topological ordering $\prec$ of X. Accordingly, we have that if xx' is an arc in $E(X)$, then $x \prec x'$. Furthermore, we need the following notion of distance.

Definition 5.1. *Let $x, x' \in X$ be two vertices such that $x \prec x'$, and let $d-1$ be the number of vertices y such that $x \prec y \prec x'$. Then, the* distance *between x and x' is d. Accordingly, $x' - x := d$ and $x - x' := -d$.*

In addition, we need the following definition which concerns the relations between the vertices in S and the vertices in X.

Definition 5.2. *For $s \in S$ and $x \in V(X)$, define $f_s^-(x) = \big|\{y \in V(X) : y \preceq x,\ sy \in E(T)\}\big|$, and $f_s^+(x) = \big|\{y \in V(X) : y \succ x,\ ys \in E(T)\}\big|$.*

Intuitively, the functions $f_s^-(x)$ and $f_s^+(x)$ measure how many arcs would have been in the "wrong direction" (with respect to the ordering $\prec$) if we inserted s into the position immediately after x in X. First, for every $s \in S$ we would like to find $x_s \in X$ such that $f_s^-(x_s)$ and $f_s^+(x_s)$ are almost equal.

Lemma 5.1. *For each $s \in S$, there exists $x_s \in V(X)$ such that $0 \leq f_s^-(x_s) - f_s^+(x_s) \leq 1$.*

Proof. Let x_m be the smallest vertex in $V(X)$, and x_M be the largest vertex in $V(X)$. Fix some $s \in S$. In what follows, we omit the subscript s. We have the following two inequalities:

- $f^-(x_M) - f^+(x_M) \geq 0$ (since $f^+(x_M) = 0$), and
- $f^-(x_m) - f^+(x_m) \leq 1$ (since $f^-(x_m) \leq 1$).

Let $x, x' \in V(X)$ where $x' = x + 1$. Then, $f^-(x') - f^+(x') = f^-(x) - f^+(x) + 1$. That is, the function $f^-(x) - f^+(x)$ increases by 1 whenever x increases by 1. Indeed, observe that if $sx' \in E(T)$, then $f^-(x') = f^-(x) + 1$ and $f^+(x') = f^+(x)$. Otherwise, $x's \in E(T)$, and so $f^-(x') = f^-(x)$ and $f^+(x') = f^+(x) - 1$. Thus, the two inequalities above, and the fact that the function $f^-(x) - f^+(x)$ increases by 1 whenever x increases by 1, together imply that there exists $x_s \in V(X)$ such that $0 \leq f^-(x_s) - f^+(x_s) \leq 1$. ☐

For the sake of clarity, we extract the implication of Lemma 5.1 to the following notation.

Definition 5.3. *For any $s \in S$, define $\varphi(s)$ as the* smallest *vertex $x_s \in V(X)$ satisfying the inequalities in Lemma 5.1.*

We now show that given Reduction Rule 5.1, neither $f_s^-(\varphi(s))$ nor $f_s^+(\varphi(s))$ can be too "large". Indeed, if there existed $s \in S$ such that $f_s^-(\varphi(s)) \geq k+2$, then $f_s^+(\varphi(s)) \geq k+1$, and we could have formed $k+1$ triangles, each consisting of s, a vertex from $\{x \in V(X) : x \preceq \varphi(s),\ sx \in E(T)\}$, and a vertex from $\{y \in V(X) : y \succ \varphi(s),\ ys \in E(T)\}$. In this case, Reduction Rule 5.1 is applicable. However, as we assumed that Reduction Rule 5.1 is no longer applicable, we have that for all $s \in S$, $f_s^-(\varphi(s)), f_s^+(\varphi(s)) \leq k+1$. By using this assumption, we have useful certificates for strong arcs as follows.

Lemma 5.2. *Let $x \in X$, and $s, s' \in S$. The following statements are true.*

1. *If $sx \in E(T)$ and $\varphi(s) - x \geq 2k+3$, then sx is strong.*
2. *If $xs \in E(T)$ and $x - \varphi(s) \geq 2k+3$, then xs is strong.*
3. *If $s's \in E(T)$ and $\varphi(s') - \varphi(s) \geq 3k+5$, then $s's$ is strong.*

Proof. We first prove $\boxed{1}$. As $\varphi(s) - x \geq 2k+3$, there are at least $2k+2$ vertices between x and $\varphi(s)$. Since $f_s^-(\varphi(s)) \leq k+1$, we have $\left|\{y : x \prec y \preceq \varphi(s), sy \in E(T)\}\right| \leq k+1$. Hence, the set $R = \{y : x \prec y \prec \varphi(s), ys \in E(T)\}$ has at least $k+2$ vertices. Note that sxy is a triangle for each $y \in R$ since $sx \in E(T)$. This shows that sx is strong. The proof of $\boxed{2}$ is similar.

To prove $\boxed{3}$, we note that $f_s^+(\varphi(s)) \leq k+1$ and $f_{s'}^-(\varphi(s')) \leq k+1$. That is, $\left|\{x : \varphi(s) \prec x \prec \varphi(s'),\ s'x \in E(T)\}\right| \leq k+1$, and $\left|\{x : \varphi(s) \prec x \prec \varphi(s'),\ xs \in E(T)\}\right| \leq k+1$. Since there are at least $3k+4$ vertices between $\varphi(s)$ and $\varphi(s')$, this implies that $\left|\{x : \varphi(s) \prec x \prec \varphi(s'),\ sx, xs' \in E(T)\}\right| \geq k+2$, i.e. $s's$ is strong. ☐

To proceed, we also need to introduce two terms concerning triangles.

Definition 5.4. *Let $x_1x_2x_3$ be a triangle of T, and $A = \{x_1, x_2, x_3\}$. The* span *of $x_1x_2x_3$ is the maximum distance between any two vertices in $(A \setminus S) \cup \varphi(A \cap S)$. Moreover, the triangle is called* local *if none of its arcs belongs to F.*

In the following lemma, we will show that a local triangle is indeed local in the sense that it must have a "short" span.

Lemma 5.3. *Let $x_1x_2x_3$ be a local triangle with at least one vertex from X. Then, its span is at most $6k+8$.*

Proof. For $1 \le i \le 3$, define

$$\varphi'(x_i) = \begin{cases} x_i, & \text{if } x_i \in V(X), \text{ and} \\ \varphi(x_i) & \text{otherwise.} \end{cases}$$

If the claim is false, then

$$\max\{|\varphi'(x_1) - \varphi'(x_2)|, |\varphi'(x_2) - \varphi'(x_3)|, |\varphi'(x_3) - \varphi'(x_1)|\} \ge 6k+9.$$

By symmetry, we may assume that $|\varphi'(x_1) - \varphi'(x_2)| \ge 6k+9$. We first claim that $(\star)$ *there is an index $i \in [3]$ such that* $\varphi'(x_i) - \varphi'(x_{i+1}) \ge 3k+5$ (where the calculation $i+1$ is modulo 3). Indeed, if $\varphi'(x_1) - \varphi'(x_2) \ge 6k+9$, then $(\star)$ is true. Therefore, next suppose that $\varphi'(x_2) - \varphi'(x_1) \ge 6k+9$. If $\varphi'(x_3) \succ \varphi'(x_2)$, then $\varphi'(x_3) - \varphi'(x_1) > 6k+9$, and then $(\star)$ is true. Moreover, if $\varphi'(x_3) \prec \varphi'(x_1)$, then $\varphi'(x_2) - \varphi'(x_3) > 6k+9$, and then $(\star)$ is true. Hence, we next suppose that $\varphi'(x_1) \prec \varphi'(x_3) \prec \varphi'(x_2)$. Then, as $\varphi'(x_3) - \varphi'(x_1) > 6k+9$, we have that either $\varphi'(x_3) - \varphi'(x_1) \ge 3k+5$ or $\varphi'(x_2) - \varphi'(x_3) \ge 3k+5$, so $(\star)$ is true. This proves $(\star)$.

Let $i \in [3]$ be an index satisfying $(\star)$, that is, $\varphi'(x_i) - \varphi'(x_{i+1}) \ge 3k+5$. If $x_i, x_{i+1} \in V(X)$, then since $x_ix_{i+1} \in E(T)$, we have that $x_i \prec x_{i+1}$, which contradicts that $\varphi'(x_i) - \varphi'(x_{i+1}) \ge 3k+5$ is positive. Thus, at least one vertex among x_i and x_{i+1} is in S. However, then Lemma 5.2 implies that x_ix_{i+1} is a strong arc, which contradicts the fact that $x_1x_2x_3$ is a local triangle. ☐

5.2 Applying the Double Expansion Lemma

In what follows, we denote $\alpha = 3$, $\beta = 20$, $\gamma = 7$, $\mu = 3$, $\delta = 2$, and $\ell = 3$ so that $\beta - 13 \ge \mu\delta$ (used in Observation 5.3), $\gamma\mu > 6\ell$ (used in Observation 5.5), $\frac{1}{\ell} + \frac{1}{\delta} < 1$ and $\ell - 1 \ge \delta$ (used in the proof of Lemma 5.4).

In order to proceed with our analysis, we need to classify "intervals" of vertices from X as either good or bad, depending on how many vertices from S are mapped into these intervals. Formally, we have the following definition.

Definition 5.5. *A set $Y \subseteq V(X)$ is an* interval *if it contains all the vertices in X that lie between the largest and smallest elements in Y (with respect to the ordering $\prec$ induced by X).*[6] *We refer to $|Y|$ as the* length *of Y. Moreover, Y is* good *if the size of $S_Y = \{s \in S \mid \varphi(s) \in Y\}$ is at most $\alpha\sqrt{k}$, and otherwise it is* bad.

Note that for two interval $Y, Y' \subseteq V(X)$, if $Y \cap Y' = \emptyset$, then $S_Y \cap S_{Y'} = \emptyset$ as well.

We partition $V(X)$ into disjoint intervals, each of length βk. That is, we follow the vertices of $V(X)$ from left to right in the ordering $\prec$, and partition them into disjoint intervals $Y_1^\star, \dots, Y_p^\star$ such that each $Y_j^\star$, $1 \le j \le p$, is of length βk. Note that among $Y_1^\star, \dots, Y_p^\star$, at most $\frac{3k}{\alpha\sqrt{k}} = \frac{3\sqrt{k}}{\alpha}$ intervals are bad; otherwise, $|S| > (\frac{3\sqrt{k}}{\alpha} + 1) \cdot \alpha\sqrt{k} > 3k$, which contradicts our assumption that $|S| \le 3k$. Thus, we have the following upper bound on the number of bad intervals among $Y_1^\star, \dots, Y_p^\star$.

Observation 5.2. *There are at most $\frac{3\sqrt{k}}{\alpha}$ bad intervals among $Y_1^\star, \dots, Y_p^\star$.*

[6] That is, the elements of Y are consecutive with respect to $\prec$.

Thus, if $p \geq \left(\frac{3}{\alpha} + \beta\right)\sqrt{k}$, there are at least $p - \frac{3\sqrt{k}}{\alpha} \geq \beta\sqrt{k}$ good intervals. Consider the first $\gamma\sqrt{k}$ good intervals among $Y_1^\star, \ldots, Y_p^\star$, and rename them as $Y_1, \ldots, Y_{\gamma\sqrt{k}}$ according to the order of the appearance (by $\prec$ of their vertices). The fact that the relative order of the intervals is preserved will be used later. For all $i \in [\gamma\sqrt{k}]$, denote $S_i = S_{Y_i}$ (recall that $S_{Y_i} = \{s \in S \mid \varphi(s) \in Y_i\}$), and let Y_i' be the sub-interval of Y_i excluding the $6k+9$ largest and the $6k+9$ smallest vertices of Y_i. The purpose of this exclusion is to ensure that the vertex set of any local triangle hit by Y_i' (that is, the triangle contains at least one vertex of Y_i') is completely contained in $Y_i \cup S_i$ (see Lemma 5.6).

Observation 5.3. *For all $i \in [\gamma\sqrt{k}]$, the length of Y_i' is at least $\beta k - 2(6k+9) > (\beta - 13)k \geq \mu\delta k$.*

We now ready apply the Double Expansion Lemma. One naive idea is to construct a bipartite graph G with vertex set $(\bigcup Y_i', S)$ and every H_i with vertex set (Y_i', S_i). However, this attempt does not work out mainly because we have too little information about the edge set of H_i to exploit. To overcome this, we chop down every Y_i' into sub-intervals $Y_{i,j}$'s, and we merge each $Y_{i,j}$ into a single "representative" vertex $a_{i,j}$, and we put an edge between $a_{i,j}$ and s in H_i if the arcs between s and $Y_{i,j}$ have different orientations. Precisely, the construction of G and H_i is as follows.

We first partition each Y_i' into $\mu\sqrt{k}$ sub-intervals, $Y_{i,1}, \ldots, Y_{i,\mu\sqrt{k}}$, each of length $\delta\sqrt{k}$ such that $x \prec x'$ for every $x \in Y_{i,j}, x' \in Y_{i,j'}$ with $j < j'$. We now construct the bipartite graphs $G, H_1, \ldots, H_{\gamma\sqrt{k}}$. To this end, for all $i, 1 \leq i \leq \gamma\sqrt{k}$, define $A_i = \{a_{i,1}, \ldots, a_{i,\mu\sqrt{k}}\}$, and $A = \bigcup_{i=1}^{\gamma\sqrt{k}} A_i$. Then, $|A| = \gamma\sqrt{k} \cdot \mu\sqrt{k} = \gamma\mu k$. It is useful to think of $a_{i,j}$ as the *representative* of the sub-interval $Y_{i,j}$ for every i, j. Let us now define the bipartite graphs (see Figure 4 for an illustration).

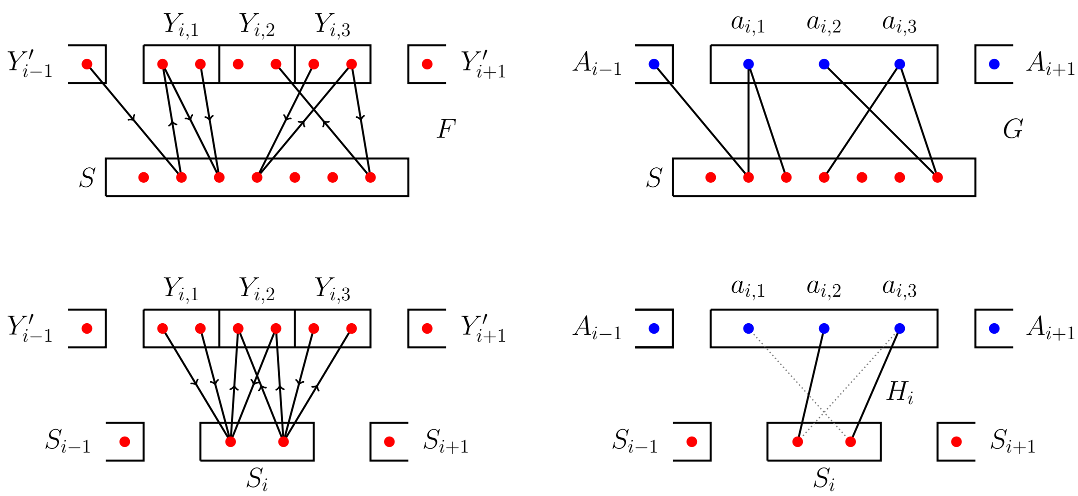

Figure 4: Construction of G from F (upper figure) and construction of H_i from $T[S_i \cup Y_i']$ (lower figure). Not all arcs of $T[S_i \cup Y_i']$ are shown; for arcs not shown in $T[S_i \cup Y_i']$, their corresponding (possibly) edges in H_i are dotted.

1. G**:** The (undirected) bipartite graph with vertex set (A, S) and edge set $E(G) = \{a_{i,j}s : \exists x \in Y_i'$ such that $\{xs, sx\} \cap F \neq \emptyset\}$. This is to take care of strong arcs.
2. H_i**:** The (undirected) bipartite graph with vertex set (A_i, S_i) and edge set $E(H_i) = \{a_{i,j}s : \exists x, x'$ such that $sx, xs' \in E(T)\}$. In other words, $a_{i,j}s \notin E(H_i)$ if and only if either $sx \in E(T)$ for every $x \in Y_{i,j}$ or $xs \in E(T)$ for every $x \in Y_{i,j}$. This is to take care of local triangles.

Before applying the Double Expansion Lemma, we mention here an observation for later use, which is the main purpose of our "merging" vertices into representative.

Observation 5.4. *If $sa_{i,j}, sa_{i,j'} \in E(H_i)$ for some $j < j'$, then there is a triangle sxx' with $x \in Y_{i,j}, x' \in Y_{i,j'}$.*

The proof is trivial: by definition of $E(H_i)$, there is $x \in Y_{i,j}$ such that $sx \in E(T)$, and there is $x' \in Y_{i,j'}$ such that $x's \in E(T)$. Since $i < j'$, $x \prec x'$, and so $xx' \in E(T)$. Thus, sxx' is a triangle.
By applying the Double Expansion Lemma 3.3, in polynomial time, we find $\widehat{A} \subseteq A$, $\widehat{S} \subseteq S$, as well as $\widehat{A}_i \subseteq A_i$ and $\widehat{S}_i \subseteq S_i$ for all $1 \leq i \leq \gamma\sqrt{k}$, such that

- $\widehat{A} = \bigcup_{i=1}^{\gamma\sqrt{k}} \widehat{A}_i$;
- $|A \setminus \widehat{A}| \leq 2\ell|S|$;
- $\widehat{S}$ has an ℓ-expansion into $\widehat{A}$ in G, and $N_G(\widehat{A}) \subseteq N_G(\widehat{S})$;
- $\widehat{S}_i$ has an ℓ-expansion into $\widehat{A}_i$ in H_i, and $N_{H_i}(\widehat{A}_i) \subseteq N_{H_i}(\widehat{S}_i)$.

Let $\widehat{Y} = \bigcup_{a_{i,j} \in \widehat{A}} Y_{i,j}$ and $\widehat{Y}_i = \bigcup_{a_{i,j} \in \widehat{A}_i} Y_{i,j}$. Since $\widehat{A} = \bigcup_{i=1}^{\gamma\sqrt{k}} \widehat{A}_i$, we have $\widehat{Y} = \bigcup_{i=1}^{\gamma\sqrt{k}} \widehat{Y}_i$.

Observation 5.5. *$\widehat{Y}$ is nonempty.*

Proof. Recall that $|S| \leq 3k$ and $|A| = \gamma\mu k$. Since $|A \setminus \widehat{A}| \leq 2\ell|S|$, we have $|\widehat{A}| \geq |A| - 2\ell|S| \geq \gamma\mu k - 2\ell \cdot 3k > 0$. Since $\widehat{A} \neq \emptyset$, there exists $a_{i,j} \in \widehat{A}$, and so $\widehat{Y} \supseteq Y_{ij} \neq \emptyset$. □

5.3 Using Expansion to Detect an Irrelevant Vertex

Let O be a solution to $(T, k+1)$, and define

$$O' = \left(O \setminus \widehat{Y}\right) \cup \left(\widehat{S} \cup \bigcup_{i=1}^{\gamma\sqrt{k}} S'_i\right), \text{ where}$$

$$S'_i = \begin{cases} \widehat{S}_i & \text{if } |O \cap \widehat{Y}_i| < \delta\sqrt{k}, \text{ and} \\ S_i & \text{otherwise.} \end{cases}$$

In the rest of this subsection, we show that if $O \cap \widehat{Y} \neq \emptyset$, then $|O'| < |O|$ and O' is a solution to $(T, k+1)$.

Lemma 5.4. *If $O \cap \widehat{Y} \neq \emptyset$, then $|O'| < |O|$.*

Proof. Observe that, to obtain O' from O, we remove $O \cap \widehat{Y}$, and add $\widehat{S} \setminus O$ and $\bigcup_{i=1}^{\gamma\sqrt{k}} (S'_i \setminus O)$.
We will prove that

$$\frac{|O \cap \widehat{Y}|}{\ell} \geq |\widehat{S} \setminus O|, \text{ and} \tag{1}$$

$$\frac{|O \cap \widehat{Y}|}{\delta} \geq \Big|\bigcup_{i=1}^{\gamma\sqrt{k}} (S'_i \setminus O)\Big|. \tag{2}$$

Combining (1) and (2) with $\frac{1}{\ell} + \frac{1}{\delta} < 1$ and the hypothesis of the lemma that $|O \cap \widehat{Y}| > 0$, we have

$$|O \cap \widehat{Y}| > |\widehat{S} \setminus O| + \Big|\bigcup_{i=1}^{\gamma\sqrt{k}} (\widehat{S}_i \setminus O)\Big|,$$

which implies that $|O'| < |O|$, proving the lemma.

To prove (1), recall that $\widehat{S}$ has an ℓ-expansion into $\widehat{A}$ in G, so $|N_G(\widehat{S} \setminus O) \cap \widehat{A}| \geq \ell|\widehat{S} \setminus O|$. Thus, it suffices to show that $|O \cap \widehat{Y}| \geq |N_G(\widehat{S} \setminus O) \cap \widehat{A}|$. Suppose for a contradiction that $|O \cap \widehat{Y}| < |N_G(\widehat{S} \setminus O) \cap \widehat{A}|$. Then

$$\sum_{a_{i,j} \in N_G(\widehat{S}\setminus O)\cap \widehat{A}} |O \cap Y_{i,j}| \leq \sum_{a_{i,j} \in \widehat{A}} |O \cap Y_{i,j}| = |O \cap \widehat{Y}| < |N_G(\widehat{S} \setminus O) \cap \widehat{A}|. \tag{3}$$

If $|O \cap Y_{i,j}| \geq 1$ for every $a_{i,j} \in N_G(\widehat{S}\setminus O)\cap\widehat{A}$, then $\sum_{a_{i,j}\in N_G(\widehat{S}\setminus O)\cap\widehat{A}} |O\cap Y_{i,j}| \geq |N_G(\widehat{S}\setminus O)\cap\widehat{A}|$, contradicting (3). Thus we conclude that there exists $a_{i,j} \in N_G(\widehat{S}\setminus O)\cap\widehat{A}$ such that $O\cap Y_{i,j} = \emptyset$. Let $s \in \widehat{S} \setminus O$ such that $sa_{i,j} \in E(G)$ (such a vertex s exists, since $a_{i,j} \in N_G(\widehat{S} \setminus O)$). By the definition of $E(G)$, there exists $x \in Y_{i,j}$ such that $sx \in F$. Note that $x \notin O$, since $O \cap Y_{i,j} = \emptyset$, and $s \notin O$, since $s \in \widehat{S} \setminus O$. As O is a solution to $(T, k + 1)$, and because of Observation 5.1, O must be a vertex cover of F. But $x, s \notin O$, which is a contradiction. From this we conclude that (1) holds.

To prove (2), note that

$$|O \cap \widehat{Y}| = \sum_{i=1}^{\gamma\sqrt{k}} |O \cap \widehat{Y}_i| \text{ and } \sum_{i=1}^{\gamma\sqrt{k}} |S'_i \setminus O| = \Big|\bigcup_{i=1}^{\gamma\sqrt{k}} (S'_i \setminus O)\Big|.$$

Thus, it suffices to show that $|O \cap \widehat{Y}_i| \geq \delta|S'_i \setminus O|$ for every i. If $S'_i = S_i$, then $|O \cap \widehat{Y}_i| \geq \delta\sqrt{k}$ by the definition of S'. Since Y_i is a good interval, $|S_i| \leq \sqrt{k}$. Hence, $|O \cap \widehat{Y}_i| \geq \delta\sqrt{k} \geq \delta|S_i| \geq \delta|S'_i \setminus O|$. Now suppose $S'_i = \widehat{S}_i$. Since $\widehat{S}_i$ has an ℓ-expansion into $\widehat{A}_i$ in H_i, we have $|N_{H_i}(\widehat{S}_i \setminus O) \cap \widehat{A}_i| \geq \ell|\widehat{S}_i \setminus O|$. Call $a_{i,j} \in N_{H_i}(\widehat{S}_i \setminus O) \cap \widehat{A}_i$ *pure*, if $Y_{i,j} \cap O = \emptyset$. Observe that if $s \in \widehat{S}_i \setminus O$ is adjacent to two pure vertices in H_i, say $a_{i,j}$ and $a_{i,j'}$ with $j < j'$, then by the definition of $E(H_i)$, then by Observation 5.4 there is a triangle sxx' with $x \in Y_{i,j}$ and $x' \in Y_{i,j'}$, and so sxx' is not hit by O by definition of purity, which contradicts the assumption that O is a feedback vertex set for T. Thus, each $s \in \widehat{S}_i \setminus O$ is adjacent to at most one pure vertex, i.e. there are at most $|\widehat{S}_i \setminus O|$ pure vertices. Thus, the number of non-pure vertices is at least (recall that $\ell - 1 \geq \delta$)

$$|N_{H_i}(\widehat{S}_i \setminus O) \cap \widehat{A}| - |\widehat{S}_i \setminus O| \geq (\ell - 1)|\widehat{S}_i \setminus O| \geq \delta|\widehat{S}_i \setminus O|.$$

For each non-pure $a_{i,j} \in N_{H_i}(\widehat{S}_i \setminus O) \cap \widehat{A}_i$, we have $|Y_{i,j} \cap O| \geq 1$, by the definition of purity. Recall that $|O \cap \widehat{Y}_i| = \sum_{a_{i,j}\in\widehat{A}_i} |O \cap Y_{i,j}|$. Thus, $|O \cap \widehat{Y}_i|$ is at least the number of non-pure vertices, i.e. $|O\cap\widehat{Y}_i| \geq \delta|\widehat{S}_i\setminus O|$. As $S'_i = \widehat{S}_i$, we have $|O\cap\widehat{Y}_i| \geq \delta|S'_i\setminus O|$, and this proves (2). □

It remains to show that O' is a solution to $(T, k + 1)$. To do so, we will prove that O' is a vertex cover of F and O' hits all local triangles.

Lemma 5.5. *O' is a vertex cover of F.*

Proof. By Observation 5.1, O is a vertex cover of F, so every $ss' \in F$ with $s, s' \in S$ is hit by O, and hence ss' is hit by O', since $O \cap S \subseteq O' \cap S$. Thus, we only need to show that every $xs \in F$ with $x \in V(X)$ and $s \in S$ is hit by O'. Suppose for contradiction that $xs \in F$ is not hit by O'. Then either $x \in O \setminus O'$ or $s \in O \setminus O'$. Note that since $O \cap S \subseteq O' \cap S$, $s \notin O \setminus O'$. So $x \in O \setminus O'$, which implies that $x \in O \cap \widehat{Y}$. Let $x \in Y_{i,j}$. Then $a_{i,j} \in \widehat{A}$, and since $xs \in F$, $a_{i,j}s \in E(G)$. Recall (from the list properties obtained after applying the Double Expansion Lemma) that $N_G(\widehat{A}) \subseteq \widehat{S}$, which implies that $s \in \widehat{S}$. However, $\widehat{S} \subseteq O'$ by the definition of O'. This implies that $s \in O'$, a contradiction to the assumption that xs is not hit by O'. This concludes the proof of the claim. □

Recall that a triangle is local if it has no strong arcs.

Lemma 5.6. *If xyz is a local triangle with $x \in O \cap \widehat{Y}_i$, then O' hits xyz.*

Proof. Suppose, for a contradiction, that O' does not hit xyz. By Lemma 5.3, xyz has span at most $6k+8$. Note that $x \in \widehat{Y}_i \subseteq Y_i'$, while Y_i' is obtained from Y_i by excluding $6k+9k$ smallest and $6k+9k$ largest vertices, so $(\{x,y,z\} \cap X) \cup \varphi(\{x,y,z\} \cap S)$ is a subset of Y_i. In other words, $x,y,z \in S_i \cup Y_i$. At least one of y,z belongs to S (otherwise, xyz is transitive), so at least one of y,z belongs to S_i. We consider two cases.

Case 1: $y \in S_i$. If $|O \cap \widehat{Y}_i| \geq \delta\sqrt{k}$, then $S_i = S_i' \subseteq O'$, and so $y \in O'$, a contradiction. We conclude that $|O \cap \widehat{Y}_i| < \delta\sqrt{k}$. If $y \in \widehat{S}_i$, then $y \in O'$, a contradiction again. Hence $y \in S_i \setminus \widehat{S}_i$. Let $x \in Y_{i,j}$, then $a_{i,j} \in \widehat{A}_i$. Recall that $N_{H_i}(\widehat{A}_i) \subseteq \widehat{S}_i$, and so $ya_{i,j} \notin E(H_i)$. Thus, by definition of $E(H_i)$, we have $x'y \in E(T)$ for every $x' \in Y_{i,j}$, since $xy \in E(T)$.

$$\text{If } z \notin O, \text{ then there is } x' \in Y_{i,j} \text{ such that } x'z \in E(T). \tag{4}$$

To prove (4), assume that $z \notin O$ and $zx' \in E(T)$ for every $x' \in Y_{i,j}$. This implies that $x'yz$ is a triangle for every $x' \in Y_{i,j}$. Since O' does not hit xyz, we have $y \notin O'$, and so $y \notin O$ (since $O \cap S \subseteq O' \cap S$). However, O is a solution to $(T, k+1)$, while $y,z \notin O$, so $x' \in O$ for every $x' \in Y_{i,j}$, i.e., $Y_{i,j} \subseteq O$. Since $Y_{i,j} \subseteq \widehat{Y}_i$, we have $|O \cap \widehat{Y}_i| \geq |Y_{i,j}| = \delta\sqrt{k}$, a contradiction to the observation $|O \cap \widehat{Y}_i| < \delta\sqrt{k}$ (made at the beginning of Case 1), which proves (4).

Note that $z \in S_i \cup Y_i$. We now consider all possibilities of z.

- If $z \in \widehat{S}_i$, then clearly $z \in O'$ since $\widehat{S}_i \subseteq O'$, a contradiction.
- If $z \in S_i \setminus \widehat{S}_i$, then recall that $N_{H_i}(\widehat{A}_i) \subseteq \widehat{S}_i$, and so $a_{i,j}z \notin E(H_i)$. Hence by definition of $E(H_i)$, we have $zx' \in E(T)$ for every $x' \in Y_{i,j}$ (since $zx \in E(T)$). If $z \in O$, then $z \in O'$ (since $O \cap S \subseteq O' \cap S$), a contradiction. Then $z \notin O$, which contradicts (4). We conclude that $z \notin S_i$, i.e. $z \in Y_i$.
- If $z \in Y_i$ and $z \notin O$, since $yz \in E(T)$ while $x'y \in E(T)$ for every $x' \in Y_{i,j}$, we have $z \notin Y_{i,j}$. Since $zx \in E(T)$, we have $z \prec x$, and so $z \prec x'$ for every $x' \in Y_{i,j}$. In other words, $zx' \in E(T)$ for every $x' \in Y_{i,j}$, a contradiction to (4).
- Otherwise, $z \in Y_i$ and $z \in O$. Then $z \in \widehat{Y}$ since $z \notin O'$. Let $z \in Y_{i,j'}$, then $a_{i,j'} \in \widehat{A}$. Observe that $j \neq j'$ since $x'y \in E(T)$ for every $x' \in Y_{i,j}$ while $zy \notin E(T)$. Since $y \in S \setminus \widehat{S}$, and recall that $N_{H_i}(\widehat{A}_i) \subseteq \widehat{S}_i$, we have $ya_{ij'} \notin E(H_i)$, and so $yz' \in E(T)$ for every $z' \in Y_{i,j'}$ (since $yz \in E(T)$). If there are $x' \in Y_{i,j}, z' \in Y_{i,j'}$ such that $x', z' \notin O$, then $x'yz'$ is not hit by O, a contradiction. Then either $Y_{i,j} \subseteq O$ or $Y_{i,j'} \subseteq O$, so

  $$|O \cap \widehat{Y}_i| \geq |O \cap (Y_{i,j} \cup Y_{i,j'})| \geq \min(|Y_{i,j}|, |Y_{i,j'}|) = \delta\sqrt{k},$$

  a contradiction to the observation $|O \cap \widehat{Y}_i| < \delta\sqrt{k}$ at the beginning of of Case 1.

We conclude that in all cases, O' always hits xyz.

Case 2: $z \in S_i$. The argument is similar as for Case 1. $\square$

Using Lemmas 5.3 to 5.6, we derive the following result.

Lemma 5.7. *O' is a solution to $(T, k+1)$.*

Proof. Suppose that O' is not a solution to $(T, k+1)$. Then there is a triangle xyz which is not hit by O'. Note that O is a solution to $(T, k+1)$, and $O \setminus O' \subseteq O \cap \widehat{Y}$, so at least one vertex of xyz belongs to $\emptyset \cap hy$, say x. Let $x \in \widehat{Y}_i$, i.e., $x \in O \cap \widehat{Y}_i$. If one of the arcs of xyz belongs to F, then O' hits xyz, by Lemma 5.5, which is a contradiction. So, xyz is local, and O' hits xyz, by Lemma 5.6, which again contradicts our assumption. $\square$

From Lemmas 5.4 to 5.7, we now conclude that if O is a solution to $(T, k+1)$, and $O \cap \widehat{Y} \neq \emptyset$, then there is another solution O' to $(T, k+1)$ with $|O'| \leq |O|-1$. Therefore, we have the following reduction rule to remove an irrelevant vertex.

Reduction Rule 5.2. *Let x be an arbitrary vertex in $\widehat{Y}$. Remove x from T. The new instance is $(T - \{x\}, k)$.*

Lemma 5.8. *Reduction rule 5.2 is safe.*

Proof. In one direction, it is clear that if $S^\star$ is a solution to (T, k), then $S^\star$ is a solution to $(T \setminus \{x\}, k)$. For the other direction, let $S^\star$ be a solution to $(T \setminus \{x\}, k)$. Then $O = S^\star \cup \{x\}$ is a solution to $(T, k+1)$. Since $x \in O \cap \widehat{Y}$, we have $O \cap \widehat{Y} \neq \emptyset$, and so there is a solution O' to $(T, k+1)$ with $|O'| \leq |O| - 1 = (|S^\star| + 1) - 1 \leq k$. Thus, O' is a solution to (T, k). $\square$

5.4 Proof of Theorem 2

We are finally ready to present the proof of Theorem 2.

Proof of Theorem 2. Let (T, k) be an instance of FVST. Our kernelization algorithm simply applies (exhaustively) Reduction Rules 5.1 and 5.2. The output is the instance obtained once none of these rules is applicable. Let us observe that each of Reduction Rules 5.1 and 5.2 can be applied in polynomial time, it strictly decreases the size of G and it does not increase k. Thus, our kernelization algorithm runs in polynomial time.

For the sake of clarity, let us now abuse notation and denote the outputted instance by (T, k). Let us observe that $V(T)$ consists of the following vertices.

- Vertices in S, whose number is at most $3k$.
- Vertices of X, whose number is at most $p\beta\sqrt{k} = \mathcal{O}(k^{3/2})$ since $p \leq \left(\frac{3}{\alpha} + \beta\right)\sqrt{k}$.

Thus, the total number of vertices is indeed $\mathcal{O}(k^{3/2})$. This complete the proof. $\square$

6 Kernel for Induced P_3-Packing

In this section, we prove the following theorem.

Theorem 3. Induced P_3-Packing *admits a kernel with $\mathcal{O}(k^{1\frac{2}{3}})$ vertices.*

Our kernel for Induced P_3-Packing is based on the kernel for CVD. In fact, several of the steps of both the kernelization algorithms are almost the same, but the subtle differences between them are crucial. Specifically, while in CVD we analyze properties that must be satisfied by *all* solutions, in Induced P_3-Packing we analyze properties such that there *exists* a solution that satisfies them (if there exists a solution at all). As we progress with the description of our kernelization algorithm for Induced P_3-Packing, the deviations from the kernelization algorithm for CVD become more palpable; in particular, the later proofs of both algorithms are completely different (for example, here we we do not even construct the bipartite graph B' as we did in Section 4.6).

Let (G, k) be an instance of Induced P_3-Packing. We start by greedily finding a maximal collection, say $\mathcal{S}$, of vertex-disjoint induced P_3's in G. Clearly, this greedy procedure can be run in polynomial time. If $|\mathcal{S}| \geq k$, then we conclude that (G, k) is a Yes-instance. Thus, we next suppose that $|\mathcal{S}| < k$. Let S be the set of vertices that belong to the induced P_3's in $\mathcal{S}$. Since $|\mathcal{S}| < k$, we have that $|S| \leq 3k$. Notice that $G \setminus S$ is a collection of cliques, which we denote by $\mathcal{C}$.

In what follows, we denote $\alpha = 2$, $\beta = 1$, $\gamma = 43$, $\mu = 26$, $\delta = 3$, $\lambda = 1$ and $\eta = 1$, so that $(1 - \frac{1}{\delta})\gamma \geq 6\eta$ (used in the proof of 6.11), $\frac{\delta-1}{\delta}\mu - \frac{14}{(\alpha-1)\beta} > 3$ (used in the proof of Lemma 6.13), $\frac{\delta-1}{\delta}\gamma \geq \frac{20}{(\alpha-1)\beta} + \lambda$ (used in the proof of Lemma 6.13), $\frac{\mu}{2} \geq 3$ (used in the proof of Lemma 6.14), and $\frac{6}{(\alpha-1)\beta} + \lambda + \frac{1}{\delta}\gamma \leq \frac{\gamma}{2}$ (used in the proof of Lemma 6.14).

6.1 Bounding the Number of Cliques

First, as in the case of CVD, we have the following simple rule, whose safeness is obvious.

Reduction Rule 6.1. *If there exists $C \in \mathcal{C}$ such that no vertex in C has a neighbor in S, then remove C from G. The new instance is $(G \setminus C, k)$.*

Now, also as in the case of CVD, we define the bipartite graph B by setting one side of the bipartition to be S and the other side to be $\mathcal{C}$, such that there exists an edge between $s \in S$ and $C \in \mathcal{C}$ if and only if s is adjacent to at least one vertex in C. Note that by Reduction Rule 6.1, no clique in $\mathcal{C}$ is an isolated vertex in B. We thus proceed by presenting the following rule (which is slightly different than Reduction Rule 4.2), where we rely on the Expansion Lemma (Lemma 3.1). It should be clear that the conditions required to apply the algorithm provided by this lemma are satisfied.

Reduction Rule 6.2. *If $|\mathcal{C}| \geq 2|S|$, then call the algorithm provided by Lemma 3.1 to compute sets $X \subseteq S$ and $\mathcal{Y} \subseteq \mathcal{C}$ such that X has a 2-expansion into $\mathcal{Y}$ in B and $N_B(\mathcal{Y}) \subseteq X$. The new instance is $(G \setminus (X \cup V(\mathcal{Y})), k - |X|)$. Here, $V(\mathcal{Y}) = \bigcup_{C \in \mathcal{Y}} V(C)$.*

We now argue that this rule is safe.

Lemma 6.1. *Reduction Rule 6.2 is safe.*

Proof. For every vertex $s \in X$, let C_s and C'_s be the two clique assigned to s by the 2-expansion. Notice that for all $s \in X$, there exists an induced P_3 in G of the form $u_s - s - v_s$, where u_s is any neighbor of s in C_s (as s and C_s are neighbors in B, at least one such vertex exists), and v_s is any neighbor of s in C'_s (again, at least one such vertex exists). Let this special collection of induced P_3's be denoted by $\mathcal{X}^\star$, that is $\mathcal{X}^\star = \{u_s - s - v_s : s \in X\}$. In one direction, it is clear that if $\mathcal{S}^\star$ is a solution to $(G \setminus (X \cup V(\mathcal{Y}), k - |X|)$, then $\mathcal{S}^\star \cup \mathcal{X}^\star$ is a solution to (G, k). For the other direction, let $\mathcal{S}^\star$ be a solution to (G, k). Let $\mathcal{W}$ denote the set of every induced P_3 in $\mathcal{S}^\star$ that contains at least one vertex from X. We denote $\mathcal{S}' = (\mathcal{S}^\star \setminus \mathcal{W}) \cup \mathcal{X}^\star$. Observe that since $N_B(\mathcal{Y}) \subseteq X$, we have that no induced P_3 in $\mathcal{S}^\star \setminus \mathcal{W}$ contains any vertex from $V(\mathcal{Y}) \cup X$. Thus, it holds that $\mathcal{S}'$ is a collection of induced P_3's in G. Since $|\mathcal{W}| \leq |X|$, we have that $|\mathcal{S}'| \geq k$. We conclude that $\mathcal{S}'$ is a solution to (G, k), and as $\mathcal{X}^\star \subseteq \mathcal{S}'$, we have that $\mathcal{S}' \setminus \mathcal{X}^\star$ is a solution to $(G \setminus (X \cup V(\mathcal{Y}), k - |X|)$. Thus, $(G \setminus (X \cup V(\mathcal{Y}), k - |X|)$ is a Yes-instance. ☐

Due to Reduction Rule 6.2, from now on $|\mathcal{C}| \leq 6k$.

6.2 The Specification of the Marking Procedure

We proceed by presenting a procedure called Mark. The specification of this procedure is similar to one presented in Section 4.2. In particular, let us emphasize one subtle difference: now we mark an additional set N, which will be a crucial component of latter rules and arguments.

Specification. The procedure Mark first initializes $M \Leftarrow \emptyset$, $T \Leftarrow S$, and for all $s \in S$, $\mathsf{mark}(s) \Leftarrow \emptyset$. At each *stage* i, $i = 1, 2, \ldots, 3k+1$, Mark executes the following process. For each $s \in T$, if there exist $C \in \mathcal{C}$ and $\{u,v\} \in E(C)$ such that $\{s,u\} \in E(G)$ but $\{s,v\} \notin E(G)$ and $\{u,v\} \cap M = \emptyset$, then insert u, v into M and $\{u,v\}$ into $\mathsf{mark}(s)$, and otherwise remove s from T. The order in which the process examines the vertices in T is immaterial given that it examines each vertex in T exactly once. Moreover, if $i = \lceil \beta k^{2/3} \rceil$, then the process sets U to T if $|T| \leq \lceil \alpha k^{2/3} \rceil$ and to an arbitrarily chosen subset of T of size $\lceil \alpha k^{2/3} \rceil$ otherwise, and it also sets N to be equal to $\bigcup_{s \in U} \mathsf{mark}(s)$. If T or M are updated in subsequent stages, U and N are *not* updated as well.

We say that Mark *succeeded* if $|U| = \lceil \alpha k^{2/3} \rceil$, and otherwise we say that Mark failed. Moreover, if there exists $s \in S$ such that $|\mathsf{mark}(s)| \geq 3k+1$, then we say that Mark was *lucky*. Let us begin the analysis of Mark with the following simple rule.

Reduction Rule 6.3. *If there exists $s \in S$ such that $|\mathsf{mark}(s)| \geq 3k+1$ (i.e., Mark was lucky), then remove s from G and decrement k by 1. The new instance is $(G \setminus s, k-1)$.*

Lemma 6.2. *Reduction Rule 6.3 is safe.*

Proof. If there exists $s \in S$ such that $|\mathsf{mark}(s)| \geq 3k+1$, then there exist $3k+1$ induced P_3's in the graph of the form $s - u_i - w_i$, $i \in \{1, \ldots, 3k+1\}$, that intersect only at s. That is, we have a "flower" whose core is s and whose petals are $\{u_i, w_i\}$. In one direction, let $\mathcal{S}^\star$ be a solution to $(G \setminus s, k-1)$. Note that $|V(\mathcal{S}^\star)| \leq 3(k-1)$. Thus, the number of induced paths of the form $s - u_i - w_i$ that intersect $V(\mathcal{S}^\star)$ is also upper bounded by $3(k-1)$. This implies that there exists an induced path $s - u_j - w_j$ that does not contain any vertex from $V(\mathcal{S}^\star)$. Then, $\mathcal{S}^\star \cup \{s - u_j - w_j\}$ is a solution to (G,k). For the other direction, let $\mathcal{S}^\star$ be a solution to (G,k). Observe that there is at most *one* induced P_3 in $\mathcal{S}^\star$ that contains the vertex s. Let $\mathcal{S}'$ be the set of induced P_3's obtained by deleting the induced P_3 in $\mathcal{S}^\star$ that contains s (if it exists). Then, $\mathcal{S}'$ is a solution to $(G \setminus s, k-1)$. $\square$

As in the case of CVD, the main purpose of Mark is to derive information on (G,k) when it is not coincidentally lucky. However, the information we require here is different than the one we require in the case of CVD. Not only do we analyze one solution rather than all solutions, we also need to state explicit relations between U and the set of vertices marked by U (that is, the set N).

Lemma 6.3. *For any induced P_3-packing $\mathcal{S}'$ of size at most k there exists an induced P_3-packing $\mathcal{S}^\star$ of size at least $|\mathcal{S}'|$ such that the following conditions hold.*

- *Let $\mathcal{P}'$ be the set of induced P_3's in $\mathcal{S}'$ that do not contain any vertex from U. Then, $\mathcal{P}' \subseteq \mathcal{S}^\star$.*
- *There exists a set $A \subseteq U$ of size at most $\frac{3}{\beta} k^{1/3}$ such that for all $s \in U \setminus A$, there exist $P \in \mathcal{S}^\star$ and $u, v \in N$ such that $P = s - u - v$.*

Proof. Let $\mathcal{S}'$ be an induced P_3-packing of G of size at most k. Observe that for all $s \in S$ and $\{u,v\}, \{u',v'\} \in \mathsf{mark}(s)$, it holds that $\{u,v\} \cap \{u',v'\} = \emptyset$. In addition, observe that for all $s, s' \in S$, $\{u,v\} \in \mathsf{mark}(s)$ and $\{u',v'\} \in \mathsf{mark}(s')$, it holds that $\{u,v\} \cap \{u',v'\} = \emptyset$. As $|V(\mathcal{S}')| \leq 3k$ and for all $s \in U$, $|\mathsf{mark}(s)| \geq \lceil \beta k^{2/3} \rceil$, we derive that there exist at most $3k / \lceil \beta k^{2/3} \rceil \leq \frac{3}{\beta} k^{1/3}$ vertices $s \in U$ such that for all $\{u,v\} \in \mathsf{mark}(s)$, $V(\mathcal{S}') \cap \{u,v\} \neq \emptyset$. Let A denote the set of these vertices in U. Moreover, let $\mathcal{P}^\star$ be the set of induced P_3's in $\mathcal{S}'$ that do not contain any vertex from $U \setminus A$. Notice that $\mathcal{P}' \subseteq \mathcal{P}^\star$. Moreover, notice that $|\mathcal{S}' \setminus \mathcal{P}^\star| \leq |U \setminus A|$. Now, define $\widehat{\mathcal{P}}$ as the P_3-packing obtained by selecting, for every vertex $s \in U \setminus A$, an induced P_3 that consists of s and an arbitrarily chosen edge $\{u,v\} \in \mathsf{mark}(s)$ such that $V(\mathcal{S}') \cap \{u,v\} \neq \emptyset$ (there exists at least one such edge). Then, $\mathcal{S}^\star = \mathcal{P}^\star \cup \widehat{\mathcal{P}}$ is an induced

P_3-packing. As $|\mathcal{S}' \setminus \mathcal{P}^\star| \leq |U \setminus A|$, we derive that $|\mathcal{S}^\star| \geq |\mathcal{S}'|$. Moreover, it is clear from its construction that $\mathcal{S}^\star$ satisfied the two properties in the statement of the lemma. This completes the proof. $\square$

We also need to derive an upper bound on the number of marked vertices, namely $|M|$.

Lemma 6.4. *If* Mark *was neither lucky nor successful, then* $|M| \leq 6(\alpha+\beta)k^{1\frac{2}{3}}$.

Proof. Since Mark was unlucky, $|\mathsf{mark}(s)| \leq 3k$ for all $s \in S$. Thus, $|M| \leq 2|U|3k + 2|S \setminus U|(\lceil \beta k^{2/3}\rceil - 1)$. Since Mark failed, we further have that $|M| \leq 6(\lceil \alpha k^{2/3}\rceil - 1)k + 6k(\lceil \beta k^{2/3}\rceil - 1) \leq 6(\alpha+\beta)k^{1\frac{2}{3}}$. $\square$

6.3 Multiple Calls to the Marking Procedure

We employ Mark exactly as in the case of CVD, with the exception that now we also compute a set $\widehat{M}$. For the sake of readability, let us repeat this short description (with the computation of $\widehat{M}$). We initialize $\widehat{U} = \emptyset$, $\widehat{M} = \emptyset$ and $\widehat{G} = G$. Then, we call Mark with $(\widehat{G}, k)$ as input. If Mark was lucky, then we execute Reduction Rule 6.3 and restart the entire process (including the initialization phase). Else, if Mark succeeded, then for the sets U and N computed by the current call, we update $\widehat{U} \Leftarrow \widehat{U} \cup U$, $\widehat{M} \Leftarrow \widehat{M} \cup N$ and $\widehat{G} \Leftarrow \widehat{G} \setminus U$, and then we proceed to execute another call. Otherwise, Mark was unlucky and also failed, and we let M denote the same set $M \subseteq V(G) \setminus S$ as computed by the *current call* to Mark, after which we terminate the process. (It may hold that $M \cap \widehat{M} \neq \emptyset$.) Note that after each call to Mark, either Reduction Rule 6.3 is executed or the size of $\widehat{U}$ increases, and therefore it is clear that the process eventually terminates. We denote $L = S \setminus \widehat{U}$.

By relying on Lemma 6.3, we have the following lemma.

Lemma 6.5. *Let i be the number of calls to* Mark *that succeeded but were unlucky. If (G,k) is a* Yes*-instance, then there exists a solution $\mathcal{S}^\star$ to (G,k) and a set $A \subseteq \widehat{U}$ of size at most $i \cdot \frac{3}{\beta}k^{1/3}$ such that for all $s \in \widehat{U} \setminus A$, there exists $P \in \mathcal{S}^\star$ and $u,v \in \widehat{M}$ such that $P = s - u - v$.*

Proof. Suppose that (G,k) is a Yes-instance, and let $\mathcal{S}'$ be a solution to (G,k) that minimizes the number of vertices $s \in \widehat{U}$ for which there do not exist $P \in \mathcal{S}'$ and $u,v \in \widehat{M}$ such that $P = s - u - v$. Let A denote the set of these vertices in $\widehat{U}$. Suppose, by way of contradiction, that $|A| > i \cdot \frac{3}{\beta}k^{1/3}$. Then, by the pigeonhole principle, there exists an iteration $j \in \{1, 2 \ldots, i\}$ such that $|A \cap U_j| > \frac{3}{\beta}k^{1/3}$, where U_j denotes the set U computed in iteration j. By Lemma 6.3, there exists a solution $\mathcal{S}^\star$ to (G,k) such that the following conditions hold.

- Let $\mathcal{P}'$ be the set of induced P_3's in $\mathcal{S}'$ that do not contain any vertex from U_j. Then, $\mathcal{P}' \subseteq \mathcal{S}^\star$.
- There exists a set $A^\star \subseteq U_j$ of size at most $\frac{3}{\beta}k^{1/3}$ such that for all $s \in U_j \setminus A^\star$, there exist $P \in \mathcal{S}^\star$ and $u,v \in \widehat{M}$ such that $P = s - u - v$. In fact, $u,v \in N$, the set computed in round j.

By the first condition, we deduce that for every $P \in \mathcal{S}'$ such that $P = s - u - v$ for some $s \in \widehat{U} \setminus U_j$ and $u,v \in \widehat{M}$, it also holds that $P \in \mathcal{S}^\star$. Furthermore, from the second condition we derive that $\mathcal{S}^\star$ has fewer vertices $s \in U_j$ than $\mathcal{S}'$ for which there do not exist $P \in \mathcal{S}'$ and $u,v \in \widehat{M}$ such that $P = s - u - v$. However, we thus conclude that $\mathcal{S}^\star$ has fewer vertices $s \in \widehat{U}$ than $\mathcal{S}'$ for which there do not exist $P \in \mathcal{S}'$ and $u,v \in \widehat{M}$ such that $P = s - u - v$. Since this contradicts the choice of $\mathcal{S}'$, we have that $|A| < i \cdot \frac{3}{\beta}k^{1/3}$. This completes the proof. $\square$

Before we proceed to proceed to present a consequence of Lemma 6.5, we need to present a new rule that is also necessary to upper bound $|\widehat{M}|$.

Reduction Rule 6.4. *Let i be the number of calls to* Mark *that succeeded but were unlucky. If $i \geq \frac{1}{\alpha-1}k^{1/3}$, then return a trivial* Yes*-instance.*

Lemma 6.6. *Reduction Rule 6.4 is safe.*

Proof. Let us consider the following simple procedure. Initialize $\mathcal{S}_0 = \emptyset$. Now, for $j = 1, 2, \ldots, i$, perform the following computation: Let $\mathcal{S}_j$ be the induced P_3-packing whose existence is guaranteed by Lemma 6.3 when applied with $\mathcal{S}' = \mathcal{S}_{j-1}$. (We implicitly assume that induced P_3's that are not of the form $s-u-v$, for $s \in \widehat{U}$ and $u, v \in \widehat{M}$, are discarded.) By the two properties of $\mathcal{S}^\star$ as specified by Lemma 6.3, we have that for all $j \in \{1, 2, \ldots, i\}$, $|\mathcal{S}_j| \geq |\mathcal{S}_{j-1}| + |U_j| - \frac{3}{\beta}k^{1/3}$, where U_j is the set U computed in iteration j. Since for all $j \in \{1, 2, \ldots, i\}$, $|U_j| = \lceil \alpha k^{2/3} \rceil$, we overall have that $|\mathcal{S}_i| \geq i \cdot (\alpha k^{2/3} - \frac{3}{\beta}k^{1/3})$. Observe that $\frac{k}{\alpha k^{2/3} - \frac{3}{\beta}k^{1/3}} = \frac{k^{2/3}}{\alpha k^{1/3} - \frac{3}{\beta}} \leq \frac{1}{\alpha-1}k^{1/3}$. Thus, if $i \geq \frac{1}{\alpha-1}k^{1/3}$, then $|\mathcal{S}_i| \geq k$, in which case $\mathcal{S}_i$ is a solution to (G, k). This implies that Reduction Rule 6.4 is indeed safe. $\square$

For the sake of clarity, let us formally define the solutions that we would like to analyze.

Definition 6.1. *We say that a pair $(\mathcal{S}^\star, A)$ is a* nice solution *to (G, k) if $\mathcal{S}^\star$ is a solution to (G, k) and $A \subseteq \widehat{U}$ is a set of size at most $\frac{3}{(\alpha-1)\beta}k^{2/3}$ such that for all $s \in \widehat{U} \setminus A$, there exist $P \in \mathcal{S}^\star$ and $u, v \in \widehat{M}$ such that $P = s - u - v$.*

Now, as a consequence of Lemma 6.5 and Reduction Rule 6.4, we have the following corollary.

Corollary 6.1. *If (G, k) is a* Yes*-instance, then there exists a nice solution to (G, k).*

Proof. Suppose that (G, k) is a Yes-instance. Let i be the number of calls to Mark that succeeded but were unlucky. By Lemma 6.5, there exists a solution $\mathcal{S}^\star$ to (G, k) and a set $A \subseteq \widehat{U}$ of size at most $i \cdot \frac{3}{\beta}k^{1/3}$ such that for all $s \in \widehat{U} \setminus A$, there exist $P \in \mathcal{S}^\star$ and $u, v \in \widehat{M}$ such that $P = s - u - v$. By Reduction Rule 6.4, we have that $i < \frac{1}{\alpha-1}k^{1/3}$. Therefore, we have that $|A| \leq \frac{1}{\alpha-1}k^{1/3} \cdot \frac{3}{\beta}k^{1/3} = \frac{3}{(\alpha-1)\beta}k^{2/3}$. We have thus obtained a nice solution $(\mathcal{S}^\star, A)$ to (G, k). $\square$

The usefulness of Corollary 6.1 stems from the observation that it implies that we have found a (possibly large) set $\widehat{U} \subseteq S$ such that not only there exists a solution that packs *almost all* the vertices in $\widehat{U}$ in induced P_3's with vertices in $\widehat{M}$, but also that the removal of $\widehat{U}$ from G significantly simplifies G as described by the following lemma. As the proof (and statement) of this lemma is identical to the proof of Lemma 4.7, it is omitted.

Lemma 6.7. *For every clique $C \in \mathcal{C}$, $C[V(C) \setminus M]$ is a module in $G \setminus \widehat{U}$.*

Before we proceed to sieve bad clique, let us upper bound $|\widehat{M}|$.

Lemma 6.8. $|\widehat{M}| \leq \frac{2\alpha\beta}{1-\alpha}k^{1\frac{2}{3}}$.

Proof. Due to each call to Mark, at most $2\lceil \alpha k^{2/3} \rceil \cdot \lceil \beta k^{2/3} \rceil$ new vertices are inserted into $\widehat{M}$. By Reduction Rule 6.4, Mark was called less than $\frac{1}{\alpha-1}k^{1/3}$ times. Thus, the total number of vertices inserted into $\widehat{M}$ is upper bounded by $2(\frac{1}{\alpha-1}k^{1/3} - 1) \cdot \lceil \alpha k^{2/3} \rceil \cdot \lceil \beta k^{2/3} \rceil \leq \frac{2\alpha\beta}{1-\alpha}k^{1\frac{2}{3}}$. $\square$

6.4 Sieving Bad Cliques

We sieve cliques based on three classifications, similarl to the case of CVD. First, we say that a clique $C \in \mathcal{C}$ is *big* if $|V(C)| > \gamma k^{2/3}$, and otherwise it is small. Furthermore, we say that a clique $C \in \mathcal{C}$ is *huge* if $|V(C)| > \mu k$. Recall that by Reduction Rule 6.2, $|\mathcal{C}| \leq 6k$. Thus, as in the case of CVD, we directly have the following observation.

Observation 6.1. *The total number of vertices in small cliques in $\mathcal{C}$ is upper bounded by $6\gamma k^{1\frac{2}{3}}$.*

Second, we say that a clique $C \in \mathcal{C}$ is *heavy* if $|V(C) \cap (M \cup \widehat{M})| \geq \frac{1}{\delta}|V(C)|$, and otherwise it is *light*. In particular, heaviness is now measured with respect to $M \cup \widehat{M}$, while in the case of CVD it was measure only with respect to M. It is clear that the total number of vertices in heavy cliques in $\mathcal{C}$ is upper bounded by $\delta|M \cup \widehat{M}|$. Thus, by Lemmata 6.4 and 6.8, we have the following observation.

Observation 6.2. *The total number of vertices in heavy cliques in $\mathcal{C}$ is upper bounded by $6\delta(\alpha+\beta+\frac{\alpha\beta}{1-\alpha})k^{1\frac{2}{3}}$.*

Third, as in the case of CVD (except that the constant 2 is replaced by 6), for a clique $C \in \mathcal{C}$ and a vertex $s \in S$, we say that C is *visible* to s if $|N_G(s) \cap V(C)| \geq 6\eta k^{2/3}$, and otherwise we say that C is *hidden* from s. For a clique $C \in \mathcal{C}$, we let $\mathsf{vis}(C)$ denote that set of vertices in S to which C is visible. Moreover, we say that a clique $C \in \mathcal{C}$ is *visible* if $|\mathsf{vis}(C)| \geq \lambda k^{2/3}$, and otherwise we say that it is *hidden*. To bound the number of visible cliques, we need the following rule.

Reduction Rule 6.5. *If there exists a vertex $s \in S$ with at least $\frac{1}{2\eta}k^{1/3}+2$ cliques in $\mathcal{C}$ visible to s, then remove s from G and decrement k by 1. The new instance is $(G \setminus s, k-1)$.*

Lemma 6.9. *Reduction Rule 6.5 is safe.*

Proof. In one direction, let $\mathcal{S}^\star$ be a solution to $(G \setminus s, k-1)$. Let $\mathcal{A}$ denote the set of cliques in $\mathcal{C}$ that are visible to s. Since $|V(\mathcal{S}^\star)| \leq 3(k-1)$, $|\mathcal{A}| \geq \frac{1}{2\eta}k^{1/3}+2$ and by the definition of visibility, we have that there necessarily exist two distinct cliques $A, A' \in \mathcal{A}$ such that each clique among A, A' has a vertex that is a neighbor of s and does not belong to $V(\mathcal{S}^\star)$. Since these two vertices together with s form an induced P_3 in G, called P, we derive that $\mathcal{S}^\star \cup \{P\}$ is a solution to (G,k). For the other direction, let $\mathcal{S}^\star$ be a solution to (G,k). Observe that there is at most *one* induced P_3 in $\mathcal{S}^\star$ that contains the vertex s. Let $\mathcal{S}'$ be the set of induced P_3's obtained by deleting the induced P_3 in $\mathcal{S}^\star$ that contains s (if it exists). Then, $\mathcal{S}'$ is a solution to $(G \setminus s, k-1)$. $\square$

After we exhaustively apply Reduction Rule 6.5, as in the case of CVD, for every vertex $s \in S$ there exist at most $\frac{1}{2\eta}k^{1/3}+1 \leq \frac{1}{\eta}k^{1/3}$ cliques in $\mathcal{C}$ visible to s. Since $|S| \leq 3k$, we derive that there are at most $\dfrac{|S|\frac{1}{\eta}k^{1/3}}{\lambda k^{2/3}} = \dfrac{3}{\lambda\eta}k^{2/3}$ visible cliques. Thus, we have the following observation.

Observation 6.3. *The total number of vertices in non-huge visible cliques in $\mathcal{C}$ is upper bounded by $\frac{3\mu}{\lambda\eta}k^{1\frac{2}{3}}$.*

Altogether, we say that a clique $C \in \mathcal{C}$ is *good* if it is **(i)** big, **(ii)** light and **(iii)** hidden or huge (or both), and otherwise we say that it is *bad*. We denote the set of all good cliques in $\mathcal{C}$ by $\mathcal{D}$. By Observations 6.1, 6.2 and 6.3, and that $\mu = 26$, we derive the following lemma.

Lemma 6.10. *The total number of vertices in bad cliques in $\mathcal{C}$ is upper bounded by $3\mu(\gamma+\delta(\alpha+\beta+\frac{\alpha\beta}{1-\alpha})+\frac{1}{\lambda\eta})k^{1\frac{2}{3}}$.*

6.5 Properties of Clique Sides

For all $C \in \mathcal{C}$ and $s \in S$, denote $N_C(s) = N_G(s) \cap V(C)$ and $\overline{N}_C(s) = V(C) \setminus N_C(s)$. Notice that for all $C \in \mathcal{C}$, $s \in S$, $u \in N_C(s)$ and $v \in \overline{N}_C(s)$, it holds that $s-u-v$ is an induced P_3 in G. Furthermore, for all $C \in \mathcal{C}$ and $s \in S$, let $M_C(s)$ denote the set of minimum size among $N_C(s)$ and $\overline{N}_C(s)$ (if they have equal sizes, the choice is arbitrary). Let us first verify that Lemma 4.11 also holds in the context of INDUCED P_3-PACKING.

Lemma 6.11. *For all $s \in L$ and $C \in \mathcal{D}$ such that $N_G(s) \cap (V(C) \setminus M) \neq \emptyset$, it holds that C is visible to s.*

Proof. Let $s \in L$ and $C \in \mathcal{D}$ such that $N_G(s) \cap (V(C) \setminus M) \neq \emptyset$. Then, by Lemma 6.7, we have that $V(C) \setminus M \subseteq N_G(s)$. Thus, to prove that C is visible to s, it is sufficient to show that $|V(C) \setminus M| \geq 6\eta k^{2/3}$. Since $C \in \mathcal{D}$, we have that C is light, and therefore $|V(C) \setminus M| > (1 - \frac{1}{\delta})|V(C)|$. Moreover, since C is big, $|V(C)| > \gamma k^{2/3}$, and hence $|V(C) \setminus M| > (1 - \frac{1}{\delta})\gamma k^{2/3}$. Since $(1 - \frac{1}{\delta})\gamma \geq 6\eta$, the proof is completed. □

Now, let us also explicitly state the following simple corollary to Lemma 6.11

Corollary 6.2. *For all non-huge $C \in \mathcal{D}$, the number of vertices $s \in L$ such that $N_G(s) \cap (V(C) \setminus M) \neq \emptyset$ is upper bounded by $\lambda k^{2/3}$.*

Proof. Let $C \in \mathcal{D}$ be a non-huge clique. By Lemma 6.11, C is visible to every vertex $s \in L$ such that $N_G(s) \cap (V(C) \setminus M) \neq \emptyset$. Thus, since C is hidden, the statement is true. □

Let us now argue that for any nice solution to (G, k), it holds that for every clique $C \in \mathcal{D}$, most of the clique C is "unused".

Lemma 6.12. *Let $(\mathcal{S}^\star, A)$ be a nice solution to (G, k). For all non-huge $C \in \mathcal{D}$, it holds that $|(V(C) \cap V(\mathcal{S}^\star)) \setminus (M \cup \widehat{M})| \leq (\frac{6}{(\alpha-1)\beta} + \lambda)k^{2/3}$.*

Proof. Let $C \in \mathcal{D}$ be a non-huge clique. Since C is a clique where $N_G(C) \subseteq S$, every induced P_3 in $\mathcal{S}^\star$ that contains at least one vertex from $V(C)$ must also contain at least one vertex from S. Because $(\mathcal{S}^\star, A)$ is a nice solution, every induced P_3 in $\mathcal{S}^\star$ that contains at least one vertex from $\widehat{U} \setminus A$ cannot contain any vertex from $V(C) \setminus \widehat{M}$. Furthermore, since $|A| \leq \frac{3}{(\alpha-1)\beta}k^{2/3}$, there exist at most $2|A| \leq \frac{6}{(\alpha-1)\beta}k^{2/3}$ vertices $v \in V(\mathcal{S}^\star)$ for which there exists an induced P_3 in $\mathcal{S}^\star$ that contains both v and at least one vertex from A. Now, let us denote the set of induced P_3's in $\mathcal{S}^\star$ that contain at least one vertex from $V(C) \setminus M$ and no vertex from $\widehat{U}$ by $\mathcal{P}$. Then, we note that every induced path $P \in \mathcal{P}$ must contain an edge $\{s, v\} \in E(G)$ for some $s \in L$ and $v \in V(C)$, and that $|V(P) \cap (V(C) \setminus M)| = 1$ (by Lemma 6.7). By Corollary 6.2, we derive that $|(V(C) \cap V(\mathcal{P})) \setminus M| \leq \lambda k^{2/3}$. This completes the proof. □

In order to proceed with our analysis, we need to refine Definition 6.1 with respect to a set of vertices.

Definition 6.2. *Let $T \subseteq V(\mathcal{D}) \setminus (M \cup \widehat{M})$. We say that a pair $(\mathcal{S}^\star, A)$ is a T-*nice solution *to (G, k) if $(\mathcal{S}^\star, A)$ is a nice solution, and for all $P \in \mathcal{S}^\star$ such that $V(P) \cap \widehat{U} = \emptyset$, it holds that $V(P) \cap T = \emptyset$.*

We now claim that that for any small enough set T, it is possible to focus on seeking nice solutions with respect to T. Formally, we prove the following lemma.

Lemma 6.13. *Let $T \subseteq V(\mathcal{D}) \setminus (M \cup \widehat{M})$ be a set of size at most $\frac{14}{(\alpha-1)\beta}k^{2/3}$. If (G, k) is a* Yes*-instance, then there exists a T-nice solution to (G, k).*

Proof. Suppose that (G, k) is a Yes-instance. Then, by Corollary 6.1, there exists a nice solution to (G, k). Let $(\mathcal{S}^\star, A)$ be a nice solution to (G, k) that minimizes the number of vertices $v \in T$ for which there exists $P \in \mathcal{S}^\star$ such that $V(P) \cap \widehat{U} = \emptyset$ and $v \in V(P)$. We claim that there do not exist $v \in T$ and $P \in \mathcal{S}^\star$ such that $V(P) \cap \widehat{U} = \emptyset$ and $v \in V(P)$. Suppose, by way of contradiction, that there exist $v \in T$ and $P \in \mathcal{S}^\star$ such that $V(P) \cap \widehat{U} = \emptyset$ and $v \in V(P)$. Let C denote the clique in $\mathcal{D}$ such that $v \in V(C)$. We first observe that due to Lemma 6.7 and because $v \notin M$ and $V(P) \cap \widehat{U} = \emptyset$, if we replace v in P by any other vertex in $V(C) \setminus M$, we obtain yet another induced P_3. Thus, by the choice of $(\mathcal{S}^\star, A)$, we derive that $V(C) \setminus (T \cup M \cup V(\mathcal{S}^\star)) = \emptyset$. In other words, $V(C) \setminus (T \cup M) \subseteq V(\mathcal{S}^\star)$. Hence, $|(V(C) \cap V(\mathcal{S}^\star)) \setminus (M \cup \widehat{M})| \geq |V(C) \setminus (T \cup M \cup \widehat{M})|$. Then, because $C \in \mathcal{D}$ and $|T| \leq \frac{14}{(\alpha-1)\beta}k^{2/3}$, we have that $|V(C) \setminus (T \cup M \cup \widehat{M})| \geq |V(C) \setminus (M \cup \widehat{M})| - \frac{14}{(\alpha-1)\beta}k^{2/3} > \frac{\delta-1}{\delta}|V(C)| - \frac{14}{(\alpha-1)\beta}k^{2/3}$. Thus, $\frac{\delta-1}{\delta}|V(C)| - \frac{14}{(\alpha-1)\beta}k^{2/3} < |(V(C) \cap V(\mathcal{S}^\star)) \setminus (M \cup \widehat{M})|$. Now, let us consider two cases, corresponding to whether or not C is huge.

- Suppose that C is huge. Then, $\frac{\delta-1}{\delta}|V(C)| - \frac{14}{(\alpha-1)\beta}k^{2/3} \geq \frac{\delta-1}{\delta}\mu k - \frac{14}{(\alpha-1)\beta}k^{2/3} > 3k$ (because $\frac{\delta-1}{\delta}\mu - \frac{14}{(\alpha-1)\beta} > 3$). However, $|(V(C) \cap V(\mathcal{S}^\star)) \setminus (M \cup \widehat{M})| \leq |V(\mathcal{S}^\star)| \leq 3k$. Thus, we have reached a contradiction.

- Suppose that C is not huge. Then, by Lemma 6.12, this means that $|(V(C) \cap V(\mathcal{S}^\star)) \setminus (M \cup \widehat{M})| \leq (\frac{6}{(\alpha-1)\beta} + \lambda)k^{2/3}$. Since $\frac{\delta-1}{\delta}|V(C)| - \frac{14}{(\alpha-1)\beta}k^{2/3} < |(V(C) \cap V(\mathcal{S}^\star)) \setminus (M \cup \widehat{M})|$ and $|V(C)| \geq \gamma k^{2/3}$, we have that $\frac{\delta-1}{\delta}k^{2/3} - \frac{14}{(\alpha-1)\beta}k^{2/3} < (\frac{6}{(\alpha-1)\beta} + \lambda)k^{2/3}$. However, since $\frac{20}{(\alpha-1)\beta} + \lambda \leq \frac{\delta-1}{\delta}\gamma$, we have reached a contradiction.

As both cases led to a contradiction, the proof is complete. $\square$

6.6 Assigning Sets of Vertices to Vertices in $\widehat{U}$

For every vertex $s \in \widehat{U}$, denote $Q'(s) = \bigcup_{C \in \mathcal{D}}(M_C(s) \setminus (M \cup \widehat{M}))$. Moreover, for every vertex $s \in \widehat{U}$, if $|Q'(s)| \leq \lfloor \frac{6}{(\alpha-1)\beta}k^{2/3} \rfloor$, then denote $Q(s) = Q'(s)$, and otherwise let $Q(s)$ be an arbitrarily chosen subset of $Q(s)$ of size exactly $\lfloor \frac{7}{(\alpha-1)\beta}k^{2/3} \rfloor$. Furthermore, we denote $\widehat{Q} = \bigcup_{s \in \widehat{U}} Q(s)$. Since $|S| \leq 3k$, the following observation is immediate.

Observation 6.4. $|\widehat{Q}| \leq \frac{21}{(\alpha-1)\beta}k^{1\frac{2}{3}}$.

Now, we proceed to apply the following rule, whose safeness is based on Lemma 6.13.

Reduction Rule 6.6. *If there exists a vertex $v \in V(\mathcal{D}) \setminus (M \cup \widehat{M} \cup \widehat{Q})$, then remove v from G. The new instance is $(G \setminus v, k)$.*

Lemma 6.14. *Reduction Rule 6.5 is safe.*

Proof. In one direction, it is clear that if $(G \setminus v, k)$ is a Yes-instance, then (G, k) is a Yes-instance. For the other direction, let us suppose that (G, k) is a Yes-instance. By Lemma 6.13 and since $v \notin M \cup \widehat{M}$, there exists a $\{v\}$-nice solution $(\mathcal{S}^\star, A)$. If $(\mathcal{S}^\star, A)$ is a solution to $(G \setminus v, k)$, then the proof is complete. Thus, we next suppose that $(\mathcal{S}^\star, A)$ is not a solution to $(G \setminus v, k)$. Because $(\mathcal{S}^\star, A)$ is a $\{v\}$-nice solution, this means that there exists $P^\star \in \mathcal{S}^\star$ such that $v \in V(P^\star)$ and $V(P^\star) \cap \widehat{U} \neq \emptyset$. Let $s^\star$ denote some vertex in $V(P^\star) \cap \widehat{U} \neq \emptyset$ (if there exist two vertices in $V(P^\star) \cap \widehat{U}$, we arbitrarily choose one of them). Now, observe that $\mathcal{S}^\star \setminus \{P\}$ is a solution to $(G \setminus \{v, s^\star\}, k-1)$, and therefore $(G \setminus \{v, s^\star\}, k-1)$ is a Yes-instance. Moreover, note that $|Q(s^\star)| \leq \frac{7}{(\alpha-1)\beta}k^{2/3} = \frac{7}{(\alpha-1)\beta}(k-1)^{2/3} \cdot \frac{k^{2/3}}{(k-1)^{2/3}} = \frac{7}{(\alpha-1)\beta}(k-1)^{2/3} \cdot \frac{1}{(1-\frac{1}{k})^{2/3}} \leq \frac{14}{(\alpha-1)\beta}(k-1)^{2/3}$. Then, by Lemma 6.13, there exists a $Q(s^\star)$-nice solution $(\mathcal{S}', A')$ to $(G \setminus \{v, s^\star\}, k-1)$.

Since $(\mathcal{S}', A')$ is a $Q(s^\star)$-nice solution, we have that for all $u \in V(\mathcal{S}') \cap Q(s^\star)$, there exists $P \in \mathcal{S}'$ such that $V(P) \cap \widehat{U} \neq \emptyset$. However, since $(\mathcal{S}', A')$ is a nice solution and $Q(s^\star) \cap (M \cup \widehat{M}) = \emptyset$, we further derive that for all $u \in V(\mathcal{S}') \cap Q(s^\star)$, there exists $P \in \mathcal{S}'$ such that $V(P) \cap A' \neq \emptyset$. Because $|A'| \leq \frac{3}{(\alpha-1)\beta}(k-1)^{2/3}$, we deduce that $|V(\mathcal{S}') \cap Q(s^\star)| \leq 2|A'| \leq \frac{6}{(\alpha-1)\beta}(k-1)^{2/3}$. However, since $Q(s^\star) = \lfloor \frac{7}{(\alpha-1)\beta}k^{2/3} \rfloor$ (because $v \in Q'(s^\star) \setminus Q(s^\star)$), we have that $Q(s^\star) \setminus V(\mathcal{S}') \neq \emptyset$. Let $v^\star$ denote some vertex in $Q(s^\star) \setminus V(\mathcal{S}')$ (by our previous argument, such a vertex exists), and let $C^\star$ denote the clique in $\mathcal{D}$ that contains $v^\star$. Then, by the definition of $Q(s^\star)$, we have that $v^\star \in M_{C^\star}(s^\star)$. Observe that any vertex in $V(C^\star) \setminus M_{C^\star}(s^\star)$ together with $s^\star$ and $v^\star$ forms an induced P_3. Hence, if $V(C^\star)\setminus(M_{C^\star}(s^\star)\cup V(\mathcal{S}')) \neq \emptyset$, then $\mathcal{S}'$ along with an induced P_3 consisting of some vertex in $V(C^\star)\setminus(M_{C^\star}(s^\star)\cup V(\mathcal{S}'))$, $s^\star$ and $v^\star$, forms a solution to (G,k), in which case the proof is complete. However, we claim that necessarily $V(C^\star) \setminus (M_{C^\star}(s^\star) \cup V(\mathcal{S}')) \neq \emptyset$. For this purpose, it is sufficient to prove that $|V(\mathcal{S}')| < |V(C^\star) \setminus M_{C^\star}(s^\star)|$. Let us first observe that $|V(C^\star) \setminus M_{C^\star}(s^\star)| \geq \frac{1}{2}|V(C^\star)|$. Hence, it is sufficient to prove that $|V(C^\star) \cap V(\mathcal{S}')| < \frac{1}{2}|V(C^\star)|$. To this end, we consider two cases, corresponding to whether or not $C^\star$ is huge.

- Suppose that $C^\star$ is huge. In this case, $\frac{1}{2}|V(C^\star)| \geq \frac{\mu}{2}k$. Since $|V(C^\star) \cap V(\mathcal{S}')| \leq |V(\mathcal{S}')| = 3(k-1)$ and $\frac{\mu}{2} \geq 3$, indeed $|V(\mathcal{S}')| < \frac{1}{2}|V(C^\star)|$.
- Suppose that $C^\star$ is not huge. In this case, by Lemma 6.12, $|(V(C^\star) \cap V(\mathcal{S}')) \setminus (M \cup \widehat{M})| \leq (\frac{6}{(\alpha-1)\beta} + \lambda)(k-1)^{2/3}$. Observe that since $C^\star \in \mathcal{D}$, we have that $|(V(C^\star) \cap V(\mathcal{S}')) \setminus (M \cup \widehat{M})| = |V(C^\star) \cap V(\mathcal{S}')| - |V(C^\star) \cap V(\mathcal{S}') \cap (M \cup \widehat{M})| \geq |V(C^\star) \cap V(\mathcal{S}')| - |V(C^\star) \cap (M \cup \widehat{M})| > |V(C^\star) \cap V(\mathcal{S}')| - \frac{1}{\delta}|V(C^\star)| > |V(C^\star) \cap V(\mathcal{S}')| - \frac{1}{\delta}\gamma k^{2/3}$. Thus, we derive that $|V(C^\star) \cap V(\mathcal{S}')| - \frac{1}{\delta}\gamma k^{2/3} \leq (\frac{6}{(\alpha-1)\beta} + \lambda)(k-1)^{2/3}$, and therefore $|V(C^\star) \cap V(\mathcal{S}')| < (\frac{6}{(\alpha-1)\beta} + \lambda + \frac{1}{\delta}\gamma)k^{2/3}$. However, $\frac{1}{2}|V(C^\star)| > \frac{\gamma}{2}k^{2/3}$. Since $\frac{6}{(\alpha-1)\beta} + \lambda + \frac{1}{\delta}\gamma \leq \frac{\gamma}{2}$, indeed $|V(\mathcal{S}')| < \frac{1}{2}|V(C^\star)|$.

As both cases led to the desired claim, the proof is complete. $\square$

6.7 Proof of Theorem 3

We are finally ready to present the proof of Theorem 3.

Proof of Theorem 3. Let (G,k) be an instance of Induced P_3-Packing. Our kernelization algorithm simply applies (exhaustively) Reduction Rules 6.1 to 6.6. The output is the instance obtained once none of these rules is applicable. Let us observe that each rule among Reduction Rules 4.1 to 4.16 can be applied in polynomial time, it strictly decreases the size of G and it does not increase k. Thus, our kernelization algorithm runs in polynomial time.

For the sake of clarity, let us now abuse notation and denote the outputted instance by (G,k). Let us observe that $V(G)$ consists of the following vertices.

- Vertices in S, whose number is at most $3k$.
- Vertices in bad cliques, whose number is at most $3\mu(\gamma + \delta(\alpha + \beta + \frac{\alpha\beta}{1-\alpha}) + \frac{1}{\lambda\eta})k^{1\frac{2}{3}}$ (by Lemma 6.10).
- Vertices in $M \cup \widehat{M}$, whose number is at most $(6(\alpha+\beta) + \frac{2\alpha\beta}{1-\alpha})k^{1\frac{2}{3}}$ (by Lemmata 6.4 and 6.8)
- Vertices in $\widehat{Q}$, whose number is at most $\frac{21}{(\alpha-1)\beta}k^{1\frac{2}{3}}$ (by Observation 6.4).

Thus, the total number of vertices is indeed $\mathcal{O}(k^{1\frac{2}{3}})$. This completes the proof. $\square$

7 Triangle Packing in Tournaments

In this section, we prove the following theorem.

Theorem 4. TPT *admits a kernel with $\mathcal{O}(k^{3/2})$ vertices.*

Let (T, k) be an instance of TPT. There is a simple polynomial-time $\frac{1}{3}$-approximation algorithm for TPT: greedily find a maximal collection, say $\mathcal{S}$, of vertex-disjoint triangles in T and output $\mathcal{S}$. Indeed, if there is a collection $\mathcal{S}^\star$ of of vertex-disjoint triangles in T with $|\mathcal{S}^\star| > 3|\mathcal{S}|$, then there is a triangle in $\mathcal{S}^\star$ not hit by $V(\mathcal{S})$, contradicting the assumption that $\mathcal{S}$ is maximal. If $|\mathcal{S}| < \frac{k}{3}$, then we conclude that (T, k) is a No-instance. If $|\mathcal{S}| \geq k$, then we conclude that (T, k) is a Yes-instance. Hence, we assume that $\frac{k}{3} \leq |\mathcal{S}| \leq k-1$. Let $S = V(\mathcal{S})$, then $|S| \leq 3k-3$. By maximality of $\mathcal{S}$, $T - S$ does not have any directed triangle, and so by Proposition 5.1, $T - S$ does not have any directed cycle. Hence, S is a feedback vertex set of T. Let $X = T - S$. Note that since S is a feedback vertex set, X is a *transitive tournament.*

Let (T, h) be an instance of TPT. We call a collection of at least h vertex-disjoint triangles of T as a solution to the instance (T, h). First, we have the following reduction rule.

Reduction Rule 7.1. *If there exists $s \in S$ such that there are $3k-2$ triangles pairwise intersecting only at s, then remove s from T. The new instance is $(T \setminus \{s\}, k-1)$.*

The safeness of rule 7.1 is simple. In one direction, if $\mathcal{S}^\star$ is a solution to (T, k), by removing the triangle (if any) containing s, we obtain a solution to $(T \setminus \{s\}, k-1)$. In the other direction, suppose that $\mathcal{S}^\star$ is a solution to $(T \setminus \{s\}, k-1)$. If $|\mathcal{S}^\star| \geq k$, then $\mathcal{S}^\star$ is a solution to (T, k). Hence $|\mathcal{S}^\star| = k-1$. If there is a triangle, say sxy that is not hit by $V(\mathcal{S}^\star)$, then $\mathcal{S}^\star \cup \{sxy\}$ is a solution to (T, k). Otherwise, $V(\mathcal{S}^\star)$ hits all $3k-2$ triangles pairwise intersecting only at s, and so $|V(\mathcal{S}^\star)| \geq 3k-2$, which contradicts $|\mathcal{S}^\star| = k-1$.

We apply Reduction Rule 7.1 exhaustively. By the same argument as for Reduction Rule 5.1, for any vertex $s \in S$, we can check whether that exist $3k$ triangles intersecting pairwise only at s in polynomial time. Thus, from now onwards, we assume that Reduction Rule 7.1 is no longer applicable.

In this section, we reuse the notation used in Section 5. Throughout this section we work with the unique ordering $\prec$ of vertices of X and use terms like consecutive vertices in X, smallest and largest vertex in X.

7.1 Exploring the Vertex Cover Structure

Recall the notion of *vertex cover* for a set of arcs of T. Formally, for a subset of arcs $A \subseteq E(T)$, a subset $O \subseteq V(T)$ is called a *vertex cover* for A if for every arc $uv \in A$, either $u \in O$ or $v \in O$ (or both). However, the definition of strong arc is slightly different from that in Section 5. An arc xy of T is called *strong* if **(i)** at least one vertex among x and y belongs to S, and **(ii)** there are at least $3k$ vertices $z \in V(T)$ such that xyz is a triangle. Let F be the set of all the strong arcs of T, which can be easily found in polynomial time.

Recall that throughout our kernelization algorithm, we work with the unique topological ordering $\prec$ of X. Accordingly, we have that if xx' is an arc in $E(X)$, then $x \prec x'$. Furthermore, we need the following notion of distance.

Definition 7.1. *Let $x, x' \in X$ be two vertices such that $x \prec x'$, and let $d-1$ be the number of vertices y such that $x \prec y \prec x'$. Then, the* distance *between x and x' is d. Accordingly, $x' - x := d$ and $x - x' := -d$.*

In addition, we need the following definition which concerns the relations between the vertices in S and the vertices in X.

Definition 7.2. *For $s \in S$ and $x \in V(X)$, define $f_s^-(x) = \big|\{y \in V(X) : y \preceq x,\ sy \in E(T)\}\big|$, and $f_s^+(x) = \big|\{y \in V(X) : y \succ x,\ ys \in E(T)\}\big|$.*

Similar to Lemma 5.1, we can prove the following.

Lemma 7.1. *For every $s \in S$, there is $x \in X$ such that $0 \leq f_s^-(x) - f_s^+(x) \leq 1$.*

As in Section 5 we have the following notation.

Definition 7.3. *For any $s \in S$, define $\varphi(s)$ as the* smallest *vertex $x_s \in V(X)$ satisfying the inequalities in Lemma 7.1.*

We now show that given Reduction Rule 7.1, neither $f_s^-(\varphi(s))$ nor $f_s^+(\varphi(s))$ can be too "large". Indeed, if there existed $s \in S$ such that $f_s^-(\varphi(s)) \geq 3k - 1$, then $f_s^+(\varphi(s)) \geq 3k - 2$, and we could have formed $3k - 2$ triangles, each consisting of s, a vertex from $\{x \in V(X) : x \preceq \varphi(s),\ sx \in E(T)\}$, and a vertex from $\{y \in V(X) : y \succ \varphi(s),\ ys \in E(T)\}$. In this case, Reduction Rule 7.1 is applicable. However, as we assumed that Reduction Rule 7.1 is no longer applicable, we have that for all $s \in S$, $f_s^-(\varphi(s)), f_s^+(\varphi(s)) \leq 3k - 2$. By using this assumption, we have useful certificates for strong arcs similar to the one in Lemma 5.2.

Lemma 7.2. *Let $x \in X$, and $s, s' \in S$. The following statements are true.*

1. *If $sx \in E(T)$ and $\varphi(s) - x \geq 6k - 2$, then sx is strong.*
2. *If $xs \in E(T)$ and $x - \varphi(s) \geq 6k - 2$, then xs is strong.*
3. *If $s's \in E(T)$ and $\varphi(s') - \varphi(s) \geq 9k - 4$, then ss' is strong.*

To proceed, as before, we also need to introduce two terms concerning triangles.

Definition 7.4. *Let $x_1x_2x_3$ be a triangle of T, and $A = \{x_1, x_2, x_3\}$. The* span *of $x_1x_2x_3$ is the maximum distance between any two vertices in $(A \setminus S) \cup \varphi(A \cap S)$. Moreover, the triangle is called* local *if none of its arcs belongs to F.*

In the following lemma, we will show that a local triangle is indeed local in the sense that it must have a "short" span. The proof of the following is identical to the one for Lemma 5.3

Lemma 7.3. *Let $x_1x_2x_3$ be a local triangle with at least one vertex from X. Then, its span is at most $18k - 8$.*

7.2 Applying the New Expansion Lemma

In what follows, we denote $\alpha = 3$, $\beta = 845$, $\gamma = 32$, $\mu = 9$, $\lambda = 25$, $\delta = 11$, and $\ell = 3$ so that $\beta - 36 - 3\ell \geq \lambda\gamma$ (used in Observation 7.8), $(\mu - 2)\ell > 4$ (used in the proof of Observation 7.10), $(\delta - 9)\ell > 4$ (used in the proof of Lemma 7.6), $\ell^2 > 4$ (used in Lemma 7.7), and $(\lambda - \frac{2\alpha\mu}{3} - \alpha)(\gamma - 2\delta) - 2\alpha\mu - \ell\alpha > 0$ (used in the proof of Lemma 7.5).

Next we give the definition of intervals.

Definition 7.5. *A set $Y \subseteq V(X)$ is an* interval *if it contains all the vertices in X that lie between the largest and smallest elements in Y (with respect to the ordering $\prec$ induced by X).*[7] *We refer to $|Y|$ as the* length *of Y.*

We partition $V(X)$ into disjoint intervals, each of length βk. That is, we follow the vertices of $V(X)$ from left to right in the ordering $\prec$, and partition them into disjoint intervals $X_1, \ldots, X_p$ such that each X_i, $1 \leq i < p$, is of length βk. Let $S_i := \{s \in S : \varphi(s) \in X_i\}$.

Definition 7.6. *Let X_i be an interval such that $|S_i| \geq \alpha\sqrt{k}$, then we call X_i* bad of Type 1.

Clearly, there are less than $\frac{3}{\alpha}\sqrt{k}$ bad intervals of Type 1, since $|S| < 3k$.

Observation 7.1. *There are at most $\frac{3\sqrt{k}}{\alpha}$ bad intervals of Type 1 among $X_1, \ldots, X_p$.*

[7] That is, the elements of Y are consecutive with respect to $\prec$.

For each i, we call a 3-approximation algorithm to TPT on the tournament $T[X_i \cup S_i]$. If the 3-approximation algorithm returns a solution of size least $\sqrt{k}$, we call X_i *bad* of Type 2. There are at most $\sqrt{k}$ bad interval of Type 2; otherwise, the (obviously vertex-disjoint) union of 3-approximate solutions of all these bad intervals has size at least k, and we conclude immediately that (T, k) is a Yes-instance.

Observation 7.2. *There are at most $\sqrt{k}$ bad intervals of Type 2 among $X_1, \ldots, X_p$.*

Observations 7.1 and 7.2 imply that there are at least $(p-1) - \frac{3}{\alpha}\sqrt{k} - \sqrt{k}$ *non-bad intervals* – we do not call them good yet, since we will introduce another type of bad intervals. For every non-bad interval X_i, let $Y_i^\star$ be the sub-interval of X_i excluding $18k$ smallest vertices and $18k$ largest vertices. Then every $Y_i^\star$ has length $(\beta - 36)k$. Recall that a triangle is local if it has no arcs in common with F. We give here two observations for later use.

Observation 7.3. *If xyz is a local triangle with $x \in Y_i^\star$ for some i, then $x, y, z \in X_i \cup S_i$.*

Proof. By Lemma 7.3, xyz has span at most $18k - 8$. Note that $x \in Y_i^\star$, while $Y_i^\star$ is obtained from X_i by excluding $18k$ smallest and $18k$ largest vertices, so $(\{x, y, z\} \cap X) \cup \varphi(\{x, y, z\} \cap S)$ is a subset of X_i. In other words, $x, y, z \in S_i \cup X_i$. □

Observation 7.4. *If X_i is a non-bad interval, then every collection of vertex-disjoint local triangles contains less than $6\sqrt{k}$ vertices in $Y_i^\star$.*

Proof. Suppose for a contradiction that there is a collection $\mathcal{O}$ of local triangles with at least $6\sqrt{k}$ vertices of $Y_i^\star$. Since each local triangle contains at most two vertices of $Y_i^\star$ ($Y_i^\star$ is transitive), the collection has at least $3\sqrt{k}$ local triangles. Let us consider a local triangle xyz with $x \in Y_i^\star$. By Observation 7.3, $x, y, z \in S_i \cup X_i$. Hence $\mathcal{O}$ contains at least $3\sqrt{k}$ triangles in $X_i \cup S_i$. This implies that a 3-approximation algorithm for TPT when run on $T[X_i \cup S_i]$ returns a solution of size at least $\sqrt{k}$. Therefore X_i is bad of Type 2, a contradiction. □

We remark that Observation 7.4 is very strong since it allows us to upper bound the number of vertex-disjoint triangles intersects a specific interval.

We now apply the New Expansion Lemma the first time to introduce the bad intervals of Type 3; later on, we will apply the New Expansion Lemma the second time to detect a relevant vertex. Let $Y^\star$ be the union of all $Y_i^\star$ such that the corresponding X_i is non-bad. Let us consider the (undirected) bipartite graph G with vertex bipartition $(S, Y^\star)$, and $E(G)$ consists of edges corresponding to those arcs in F which has one endpoint in S and another endpoint in $Y^\star$. By applying New Expansion Lemma (Lemma 3.2) on G, we obtain $\widehat{Y}^\star$ and $\widehat{S}$ satisfying the following.

Observation 7.5. *$\widehat{S}$ has an $\ell\sqrt{k}$-expansion to $\widehat{Y}^\star$ in G, $N_G(\widehat{Y}_i^\star) \subseteq \widehat{S}$ and $|Y^\star \setminus \widehat{Y}^\star| \leq \ell\sqrt{k}|S|$.*

We can now define the third type of bad intervals.

Definition 7.7. *For every i, if $|Y_i^\star \setminus \widehat{Y}^\star| \geq 3\ell k$, then X_i is called* bad *of Type 3.*

Then there are at most $\sqrt{k}$ bad intervals of Type 3 since $|Y^\star \setminus \widehat{Y}^\star| \leq \ell\sqrt{k}|S| < 3\ell k^{3/2}$.

Observation 7.6. *There are at most $\sqrt{k}$ bad intervals of Type 3 among $X_1, \ldots, X_p$.*

Finally, we are ready to define the notion of *good* interval.

Definition 7.8. *Let X_i be an interval such that it is not bad of Types 1, 2 or 3, then it is called* good.

Observations 7.1, 7.2 and 7.6 imply that there are at least $p - -\frac{3}{\alpha}\sqrt{k} - \sqrt{k} - \sqrt{k}$ good intervals. Then we have the following observation.

Observation 7.7. *If $p \geq \left(\frac{3}{\alpha}+3\right)\sqrt{k}$ then there are at least $\sqrt{k}$ good intervals.*

For every good X_i, let $Y_i = Y_i^\star \cap \widehat{Y}^\star$. Then $|Y_i| \geq |Y_i^\star| - |Y_i^\star \setminus \widehat{Y}^\star| \geq (\beta - 36 - 3\ell)k$. Since, $\beta - 36 - 3\ell \geq \lambda\gamma$, we have the following observation.

Observation 7.8. *$|Y_i| \geq \lambda\gamma k$ for every good X_i.*

Given $x, x' \in Y_i$ with $x \prec x'$, we say that x and x' are *consecutive* in Y_i if there is no $y \in Y_i$ such that $x \prec y \prec x'$. Note that Y_i is *not a sub-interval* of X_i, so two consecutive vertices in Y_i are not necessary consecutive in X_i. To avoid confusion, we do not introduce the distance notion between two vertices in Y_i; however, the order $\prec$ in Y_i is the restriction of $\prec$ on X. The following observation is immediate from Observation 7.4.

Observation 7.9. *If X_i is good, then every collection of vertex-disjoint local triangles contains less than $6\sqrt{k}$ vertices in Y_i.*

For a vertex $s \in S_i$, $\varphi(s)$ can be thought as a "balanced projection" of s on X. However, $\varphi(s)$ may not be a balanced projection of s on Y_i. Thus, we wish to find a balanced projection of s on Y_i. To do so, we repeat what we did before to find $\varphi(s)$ as follows. For every $s \in S_i$, let $R_s^-(x) = \{y \in Y_i : y \preceq x, sy \in E(T)\}$ and $R_s^+(x) = \{y \in Y_i : y \succ x, ys \in E(T)\}$. Note that $R_s^-(x)$ and $R_s^+(x)$ only count arcs between s and Y_i.

Lemma 7.4. *For every $s \in S_i$, there is $x \in Y_i$ such that $0 \leq |R_s^-(x)| - |R_s^+(x)| \leq 1$.*

The proof of Lemma 7.4 is similar to that of Lemma 5.1, where note that $|R_s^-(x')| - |R_s^+(x')| = |R_s^-(x)| - |R_s^+(x)| + 1$ for every $x \prec x'$ consecutive in Y_i.

Definition 7.9. *For any $s \in S_i$, define $\theta(s)$ to be the smallest vertex in Y_i satisfying the inequalities in Lemma 7.4.*

We denote $R_s^+ = R_s^+(\theta(s)), R_s^- = R_s^-(\theta(s))$ for short. We could not upper bound $|R_s^-|$ and $|R_s^+|$ as what we did for $\varphi(s)$; thus, we overcome this by introducing the notions of *heavy* and *light*.

Definition 7.10. *Given $s \in S_i$, if $|R_s^-| \geq \mu\sqrt{k}+1$, then we call s* heavy*; otherwise, we call s* light.

Thus, if s is light, $|R_s^-|, |R_s^+| \leq \mu\sqrt{k}$. Let

$$R_i = \bigcup_{\{s \in S_i | s \text{ is light}\}} \left(R_s^- \cup R_s^+\right).$$

Then $|R_i| \leq 2\mu\sqrt{k}|S_i| \leq 2\alpha\mu k$ since $|S_i| \leq \alpha\sqrt{k}$.

Recall from Observation 7.8 that $|Y_i| \geq \lambda\gamma$. We partition Y_i into subsets $Y_{i,1}, ..., Y_{i,\lambda\sqrt{k}}$ where $|Y_{i,j}| \geq \gamma\sqrt{k}$ for every $j \leq \lambda\sqrt{k}$, and $x \prec x'$ for every $x \in Y_{i,j}, x' \in Y_{i,j'}$ with $j < j'$ (it is useful to think that $Y_{i,j}$ is a "sub-interval" of Y_i; however, we would like to avoid that term since Y_i itself is not an interval).

Definition 7.11. *A set $Y_{i,j}$ is called* fit *if $|Y_{i,j} \cap R_i| < 3\sqrt{k}$ and $\theta(S_i) \cap Y_{i,j} = \emptyset$.*

Since $|R_i| \leq 2\alpha\mu k$, there are at most $\frac{2\alpha\mu}{3}\sqrt{k}$ sets $Y_{i,j}$ such that $|Y_{i,j} \cap R_i| \geq 3\sqrt{k}$. Since $|S_i| \leq \alpha\sqrt{k}$, there are at most $\alpha\sqrt{k}$ intervals $Y_{i,j}$ such that $Y_{i,j}$ contains $\theta(s)$ for some $s \in S_i$. Thus, there are at least $(\lambda - \frac{2\alpha\mu}{3} - \alpha)\sqrt{k}$ fit subset $Y_{i,j}$ of every Y_i. For each fit $Y_{i,j}$, let $Y_{i,j}^-, Y_{i,j}^+$ be the $\delta\sqrt{k}$ smallest vertices and $\delta\sqrt{k}$ largest vertices in $Y_{i,j}$, respectively, and $Y'_{i,j} \geq Y_{i,j} \setminus (Y^-_{i,j} \cup Y^+_{i,j})$. Then $Y'_{i,j} = (\gamma - 2\delta)\sqrt{k}$. Let

$$A_i = \left(\bigcup_{\{j|Y_{i,j} \text{ is fit}\}} Y'_{i,j} \right) \setminus R_i \quad \text{and} \quad A = \bigcup_{\{i|X_i \text{ is good}\}} A_i.$$

Then

$$|A_i| \geq \left(\lambda - \frac{2\alpha\mu}{3} - \alpha \right) \sqrt{k}(\gamma - 2\delta)10\sqrt{k} - |R_i| \geq \left(\left(\lambda - \frac{2\alpha\mu}{3} - \alpha \right) (\gamma - 2\delta) - 2\alpha\mu \right) k,$$

and by Observation 7.7, we have $|A| \geq \sqrt{k}((\lambda - \frac{2\alpha\mu}{3} - \alpha)(\gamma - 2\delta) - 2\alpha\mu)k$.

Now we apply the New Expansion Lemma (Lemma 3.2) the second time, this time on $G[A \cup \widehat{S}]$, to get $\widehat{A}$ and $\ddot{S}$ such that $\ddot{S}$ has an $\ell\sqrt{k}$-expansion into $\widehat{A}$ in $G[A \cup \widehat{S}]$, $N_{G[A \cup \widehat{S}]}(\widehat{A}) \subseteq \ddot{S}$ and $|A \setminus \widehat{A}| \leq \ell\sqrt{k}|\widehat{S}|$.

Lemma 7.5. *$\ddot{S}$ has an $\ell\sqrt{k}$-expansion into $\widehat{A}$ in G, $N_G(\widehat{A}) \subseteq \ddot{S}$ and $\widehat{A}$ is nonempty.*

Proof. Since $G[A \cup \widehat{S}]$ is a induced subgraph of G, then clearly $\ddot{S}$ has an $\ell\sqrt{k}$-expansion into $\widehat{A}$ in G. By Observation 7.5, $N_G(\widehat{Y}^\star) \subseteq \widehat{S}$, so there is no edge between $\widehat{Y}^\star$ and $S \setminus \widehat{S}$ in G. Thus, there is no edge between $\widehat{A}$ and $S \setminus \widehat{S}$ in G since $\widehat{A} \subseteq \widehat{Y}^\star$. Since $N_{G[A \cup \widehat{S}]}(\widehat{A}) \subseteq \ddot{S}$, there is no edge between $\widehat{A}$ and $\widehat{S} \setminus \ddot{S}$ in $G[A \cup \widehat{S}]$. Thus, there is no edge between $\widehat{A}$ and $\widehat{S} \setminus \ddot{S}$ in G. In other words, $N_G(\widehat{A}) \subseteq \ddot{S}$.

Observe that $|A \setminus \widehat{A}| \leq \ell\sqrt{k}|\widehat{S}| \leq \ell\sqrt{k}|S| \leq \ell\alpha k^{3/2}$. Hence, $\widehat{A} \geq |A| - |A \setminus \widehat{A}| \geq ((\lambda - \frac{2\alpha\mu}{3} - \alpha)(\gamma - 2\delta) - 2\alpha\mu)k^{3/2} - \ell\alpha k^{3/2} > 0$. This proves the lemma. □

7.3 Using Expansion to Detect an Irrelevant Vertex

Recall that a triangle is local if it contains no strong arc, i.e. it has no arcs in common with F. In the next lemma, we will show that given a mixed collection of local triangles and strong arcs, it is possible to exclude a particular vertex of $\widehat{A}$ from the collection.

Lemma 7.6. *Let $x \in \widehat{A}$ and assume that there is a vertex-disjoint collection $\mathcal{O}$ of local triangles and strong arcs such that $|\mathcal{O}| = k$ and x belongs to a local triangle of $\mathcal{O}$. Then there is a vertex-disjoint collection $\mathcal{O}'$ of local triangles and strong arcs such that $|\mathcal{O}'| = k$ and x does not belong to any local triangle or strong arc of $\mathcal{O}$.*

Proof. Let O be the vertex set of $\mathcal{O}$, then $|O| \leq 3k$. Assume that the statement of the lemma was false and let $xyz \in \mathcal{O}$ and $x \in Y_{i,j}$. By Observation 7.3, we have $y, z \in X_i \cup S_i$ since $Y_{ij} \subseteq Y_i^\star$. Thus, either $y \in S_i$ or $z \in S_i$; otherwise, xyz is transitive. We first prove the following observation.

Observation 7.10. *Neither y nor z is heavy.*

Proof. Suppose that y is heavy. If there are $v \in R_y^-, v' \in R_y^+$ such that $v, v' \notin O$, then $\mathcal{O}' := (\mathcal{O} \setminus \{xyz\}) \cup \{yvv'\}$ is a desired collection, a contradiction. Thus we conclude that either $R_y^- \subseteq O$ or $R_y^+ \subseteq O$.

If $R_y^+ \subseteq O$, we will show that we can exchange some strong arc of $\mathcal{O}$ with some strong arc outside to "free" one vertex of R_y^+ from $\mathcal{O}$. Since $R_y^+ \subseteq Y_i$, by Observation 7.9, at most $6\sqrt{k}$ vertices in R_y^+ belong to a local triangle in $\mathcal{O}$. Recall that $|R_y^+| \geq \mu\sqrt{k}$ by the definition of heaviness. Thus at least $(\mu - 6)\sqrt{k}$ vertices belong to some strong arcs of $\mathcal{O}$; we call that set Z. Then $\mathcal{O}$ contains a matching of strong arcs from Z into S (since a strong arc must contain at least one vertex in S). Let W be the set of vertices of S in that matching, then $|W| = |Z| \geq (\mu - 2)\sqrt{k}$. Note that $Z \subseteq \widehat{Y}^\star$ since $R_i \subseteq Y_i \subseteq \widehat{Y}^\star$. By Observation 7.5, we have $N_G(\widehat{Y}^\star) \subseteq \widehat{S}$, and so $W \subseteq N_G(Z) \subseteq \widehat{S}$. By Observation 7.5 again, we have $|N_G(W)| \geq \ell\sqrt{k}|W| \geq (\mu-2)\ell k > 4k = |O| + k$. We choose an arbitrary $u \in N_G(W)$ (among at least k candidates) such that $u \notin O$, let $w \in W$ be a neighbor of u in G (such w always exists since $u \in N_G(W)$). Suppose that $wv \in \mathcal{O}$, then $v \in Z \subseteq R_y^+$. Then remove wv from $\mathcal{O}$ and add wu to $\mathcal{O}$. We still call the new collection $\mathcal{O}$. In doing so, we free $v \in R_y^+$ from $\mathcal{O}$.

If $R_y^- \subseteq O$, by repeating the above argument, we can free some $v' \in R_y^-$ from $\mathcal{O}$. Note that since we have k candidates to choose to exchange strong arcs, we can avoid "recapturing" v into $\mathcal{O}$. Then $\mathcal{O}' := (\mathcal{O} \setminus \{xyz\}) \cup \{yvv'\}$ is a desired collection, a contradiction. Similarly, we can show that z is not heavy. $\square$

We have 3 cases:

Case 1: $y \in S_i$ and $z \in X_i$. Then y is light by Observation 7.10, and since $x \in \widehat{A} \subseteq A$ while $A \cap R_i = \emptyset$, we have $x \notin R_i$, and so $x \notin R_y^+ \cup R_y^-$. Combining with $xy \in E(T)$, we have $\theta(y) > x$. Recall that $Y_{i,j}^+$ is the set of $\delta\sqrt{k}$ largest vertices of $Y_{i,j}$, and hence $u \prec v$ for every $u \in Y'_{i,j}, v \in Y_{i,j}^+$. Since $x \in Y'_{i,j}$, we have $x \prec v$ for every $v \in Y_{i,j}^+$. Note that $\theta(y) \succ x$, and $\theta(y) \notin Y_{i,j}$ since $Y_{i,j}$ is fit. Thus $\theta(y) \succ v$ for every $v \in Y_{i,j}^+$. Since $R_y^+ \cup R_y^- \subseteq R_i$, we have $vy \in E(T)$ for every $v \in Y_{i,j}^+ \setminus R_i$. Besides, $|Y_{i,j}^+ \cap R_i| < 3\sqrt{k}$ since $Y_{i,j}$ is fit, and so $|Y_{i,j}^+ \setminus R_i| \geq |Y_{i,j}^+| - |Y_{i,j}^+ \cap R_i| \geq (\delta - 3)\sqrt{k}$.

Note that $z \in X_i$ and $zx \in E(T)$, so $z \prec x \prec v$ for every $v \in Y_{i,j}^+$. If there is $v \in Y_{i,j}^+ \setminus R_i$ such that $v \notin O$, then $\mathcal{O}' := (\mathcal{O} \setminus \{xyz\}) \cup \{vyz\}$ is a desired collection, a contradiction. Thus we conclude that there is no such v. In other words, $Y_{i,j}^+ \setminus R_i \subseteq O$. Since $Y_{i,j}^+ \subseteq Y_i$, by Observation 7.9, at most $6\sqrt{k}$ vertices in $Y_{i,j}^+ \setminus R_i$ belong to local triangles in $\mathcal{O}$, while we showed above that $|Y_{i,j}^+ \setminus R_i| \geq (\delta - 3)\sqrt{k}$. Thus at least $(\delta - 3 - 6)\sqrt{k} = (\delta - 9)\sqrt{k}$ vertices of $Y_{i,j}^+ \setminus R_i$ belong to some strong arcs of $\mathcal{O}$. Then by the same arguments as in the proof of Observation 7.10, combined with the assumption that $(\delta - 9)\ell > 4$, we can exchange strong arcs of $\mathcal{O}$ to free some $v \in Y_{i,j}^+ \setminus R_i$ from $\mathcal{O}$. Then $\mathcal{O}' := (\mathcal{O} \setminus \{xyz\}) \cup \{vyz\}$ is a desired collection, a contradiction.

Case 2: $y \in X_i$ and $z \in S_i$. This case is similar to Case 1, but we will consider $Y_{i,j}^-$ (instead of $Y_{i,j}^+$) to employ the fact that $v \prec x$ for every $v \in Y_{i,j}^-$.

Case 3: $y, z \in S_i$. Then by a similar argument as in Case 1, both y, z are light and $\theta(z) \prec x \prec \theta(y)$. Then we have $\theta(z) \prec v \prec \theta(y)$ for every $v \in Y_{i,j}^+$ since $\theta(y), \theta(z) \notin Y_{i,j}$. Note that $zv, vy \in E(T)$ for every $v \in Y_{i,j}^+ \setminus R_i$, then we repeat the argument in Case 1 to reach the contradiction. This concludes the proof. $\square$

We can now strengthen Lemma 7.6 by omitting the assumption that x belongs to a local triangle of $\mathcal{O}$.

Lemma 7.7. *Let $x \in \widehat{A}$ and suppose that there is a vertex-disjoint collection $\mathcal{O}$ of local triangles and strong arcs with $|\mathcal{O}| = k$. Then there is a vertex-disjoint collection $\mathcal{O}'$ of local triangles and strong arcs such that $|\mathcal{O}'| = k$ and x does not belong to any triangle or strong arc of $\mathcal{O}$.*

Proof. Let O be the vertex set of $\mathcal{O}$, then $|O| \leq 3k$ and $x \in O$ (otherwise, the lemma is obvious). If x belongs to a local triangle of $\mathcal{O}$, then we apply Lemma 7.6. Otherwise, x belongs to a strong arc of $\mathcal{O}$, say xy (note that in this proof, we do not consider the orientation of a strong arc, i.e. when we say uv is a strong arc, we mean either uv or vu is a strong arc).

By Lemma 7.5, $N_G(\widehat{A}) \subseteq \ddot{S}$, and so $y \in N_G(x) \subseteq \ddot{S}$, and $|N_G(y) \cap \widehat{A}| \geq \ell\sqrt{k}|\{y\}| = \ell\sqrt{k}$. Let $Z = N_G(y) \cap \widehat{A}$. If there is $v \in Z$ such that $v \notin O$, then $\mathcal{O}' := (\mathcal{O} \setminus \{xy\}) \cup \{vy\}$ is a desired collection. Thus, we conclude that $Z \subseteq O$.

Suppose that there is $v \in Z$ such that v belongs to a local triangle of $\mathcal{O}$. Since $v \in \widehat{A}$, we apply Lemma 7.6 to v and obtain a collection $\mathcal{O}''$ such that $|\mathcal{O}''| = k$ and v does not belong to any triangle or strong arc of $\mathcal{O}''$. Note also that according to the proof of Lemma 7.6, $\mathcal{O}''$ is obtained from $\mathcal{O}$ by exchanging some strong arcs and a local triangle. If x is freed by these exchange, then $\mathcal{O}''$ is a desired collection. Note that it is impossible that x is freed and then recaptured to $\mathcal{O}'$ in a strong arc, since we can always have k candidates of strong arcs, and so we can avoid recapturing x. x is freed and then recaptured to $\mathcal{O}'$ in a local triangle, then we just apply Lemma 7.6 again to x to find a desired collection. Thus, we concluded that x is "untouched" during the swapping procedure above, i.e. $xy \in \mathcal{O}''$. Then $\mathcal{O}' := (\mathcal{O}'' \setminus \{xy\}) \cup \{vy\}$ is a desired collection.

We conclude that every element of Z belongs to some strong arc of $\mathcal{O}$. Then $\mathcal{O}$ contains a matching of strong arcs from Z to S. Let W be the set of vertices of S in that matching, then $|W| = |Z| \geq \ell\sqrt{k}$. Note that $Z \subseteq \widehat{A}$. By Lemma 7.5, we have $N_G(\widehat{A}) \subseteq \ddot{S}$, and so $W \subseteq \ddot{S}$. By Lemma 7.5 again, we have $|N_G(W)| \geq \ell\sqrt{k}|W| \geq \ell\sqrt{k} \times \ell\sqrt{k} = \ell^2 k > 4k = |O| + 4$. Thus, there is $u \in N_G(W)$ such that $u \notin O$. Let $w \in W$ be a neighbor of u in G (such w always exists since $u \in N_G(W)$). Let $wv \in \mathcal{O}$. Then $\mathcal{O}' := (\mathcal{O} \setminus \{xy, wv\}) \cup \{vy, wu\}$ is a desired collection. $\square$

Finally, we are ready to state the reduction rule that removes an irrelevant vertex.

Reduction Rule 7.2. *Let x be an arbitrary vertex in $\widehat{A}$. Remove x from T. The new instance is $(T \setminus \{x\}, k)$.*

Lemma 7.8. *Reduction Rule 7.2 is safe.*

Proof. If is obvious that if $(T \setminus \{x\}, k)$ is a Yes-instance, then (T, k) is a Yes-instance. Conversely, suppose that (T, k) is a Yes-instance with some solution $\mathcal{O}^\star$, while $(T \setminus \{x\}, k)$ is a No-instance. Then $|\mathcal{O}^\star| = k$. Let $\mathcal{O}$ be the collection of local triangles and strong arcs obtained from $\mathcal{O}$ as follows. For every $uvw \in \mathcal{O}^\star$, if uvw is local, then $uvw \in \mathcal{O}$; otherwise, uvw must contain some strong arcs, then choose an arbitrary strong arc of uvw to be in $\mathcal{O}$. Then $|\mathcal{O}| = k$. Applying Lemma 7.7, we obtain a collection $\mathcal{O}'$ such that x does not belong to any local triangle and strong arc of $\mathcal{O}'$.

We now construct a solution to $(T \setminus \{x\}, k)$ by repeating the following argument sequentially. Pick an arbitrary strong arc of $\mathcal{O}'$, say yz, we choose a vertex w such that yzw is a triangle, $w \neq x$ and w does not belong to any element of $\mathcal{O}'$ and set $\mathcal{O}' := (\mathcal{O}' \setminus \{yz\}) \cup \{yzw\}$. It is clear that we can always proceed the exchange since the vertex set $\mathcal{O}' \setminus \{yz\}$ has at most $3k - 3$ vertices, while there are $3k$ possible choices for w since yz is strong, so we can always find the desired w. At the end of the process $\mathcal{O}'$ is a solution to $(T \setminus \{x\}, k)$. This concludes the proof. $\square$

7.4 Proof of Theorem 4

We are finally ready to present the proof of Theorem 2.

Proof of Theorem 2. Let (T, k) be an instance of TPT. Our kernelization algorithm simply applies (exhaustively) Reduction Rules 7.1 and 7.2. The output is the instance obtained once none of these rules is applicable. Let us observe that each of Reduction Rules 7.1 and 7.2 can be applied in polynomial time, it strictly decreases the size of G and it does not increase k. Thus, our kernelization algorithm runs in polynomial time.

For the sake of clarity, let us now abuse notation and denote the output instance by (T, k). Let us observe that $V(T)$ consists of the following vertices.

- Vertices in S, whose number is at most $3k$.
- Vertices of X, whose number is at most $p\beta\sqrt{k} = \mathcal{O}(k^{3/2})$ since $p \leq \left(\frac{3}{\alpha} + 3\right)\sqrt{k}$.

Thus, the total number of vertices is indeed $\mathcal{O}(k^{3/2})$. This complete the proof. $\square$

8 Conclusion

In this paper we designed first subquadratic vertex kernels for FVST, CVD, TPT, and INDUCED P_3-PACKING. All our kernels were based on the classical Expansion Lemma and the two new versions we proved in this article. We believe that our approach of designing kernels will be fruitful for similar implicit packing and covering problems. A most natural open question is whether these problems admit a kernel with $\mathcal{O}(k)$ vertices. Another interesting avenue is to find other problems where the methods developed in this paper can be applied.

References

[1] F. N. ABU-KHZAM, *An improved kernelization algorithm for r-set packing*, Inf. Process. Lett., 110 (2010), pp. 621–624. 2

[2] F. N. ABU-KHZAM, *A kernelization algorithm for d-Hitting Set*, J. Comput. Syst. Sci., 76 (2010), pp. 524–531. 1, 2

[3] S. BESSY, M. BOUGERET, AND J. THIEBAUT, *Triangle packing in (sparse) tournaments: approximation and kernelization*, in Proceedings of the 25th Annual European Symposium on Algorithms (ESA), To Appear, 2017. 5

[4] S. BESSY, F. V. FOMIN, S. GASPERS, C. PAUL, A. PEREZ, S. SAURABH, AND S. THOMASSÉ, *Kernels for feedback arc set in tournaments*, J. Comput. Syst. Sci., 77 (2011), pp. 1071–1078. 2

[5] H. L. BODLAENDER, R. G. DOWNEY, M. R. FELLOWS, AND D. HERMELIN, *On problems without polynomial kernels*, Journal of Computer and System Sciences, 75 (2009), pp. 423–434. 1

[6] H. L. BODLAENDER, F. V. FOMIN, AND S. SAURABH, *Open problems, worker 2010.*, Available at `http://fpt.wikidot.com/open-problems`, (2010). 2

[7] A. BORAL, M. CYGAN, T. KOCIUMAKA, AND M. PILIPCZUK, *A fast branching algorithm for cluster vertex deletion*, Theory Comput. Syst., 58 (2016), pp. 357–376. 2

[8] M. CAI, X. DENG, AND W. ZANG, *An approximation algorithm for feedback vertex sets in tournaments*, SIAM J. Comput., 30 (2000), pp. 1993–2007. 2

[9] M. CYGAN, F. V. FOMIN, Ł. KOWALIK, D. LOKSHTANOV, D. MARX, M. PILIPCZUK, M. PILIPCZUK, AND S. SAURABH, *Parameterized Algorithms*, Springer, 2015. 1, 3, 4, 7

[10] H. DELL AND D. MARX, *Kernelization of packing problems*, in Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, 2012, pp. 68–81. 1, 5

[11] H. DELL AND D. VAN MELKEBEEK, *Satisfiability allows no nontrivial sparsification unless the polynomial-time hierarchy collapses*, Journal of the ACM, 61 (2014), pp. 23:1–23:27. 1, 5

[12] M. DOM, J. GUO, F. HÜFFNER, R. NIEDERMEIER, AND A. TRUSS, *Fixed-parameter tractability results for feedback set problems in tournaments*, J. Discrete Algorithms, 8 (2010), pp. 76–86. 1, 2

[13] R. G. DOWNEY AND M. R. FELLOWS, *Fundamentals of Parameterized Complexity*, Texts in Computer Science, Springer, 2013. 1

[14] A. Drucker, *New limits to classical and quantum instance compression*, SIAM Journal on Computing, 44 (2015), pp. 1443–1479. 1

[15] M. R. Fellows, C. Knauer, N. Nishimura, P. Ragde, F. A. Rosamond, U. Stege, D. M. Thilikos, and S. Whitesides, *Faster fixed-parameter tractable algorithms for matching and packing problems*, Algorithmica, 52 (2008), pp. 167–176. 1

[16] S. Fiorini, G. Joret, and O. Schaudt, *Improved approximation algorithms for hitting 3-vertex paths*, in Integer Programming and Combinatorial Optimization - 18th International Conference, IPCO 2016, Liège, Belgium, June 1-3, 2016, Proceedings, vol. 9682 of Lecture Notes in Computer Science, 2016, pp. 238–249. 2

[17] J. Flum and M. Grohe, *Parameterized Complexity Theory*, Texts in Theoretical Computer Science. An EATCS Series, Springer-Verlag, Berlin, 2006. 1

[18] F. V. Fomin, S. Gaspers, D. Lokshtanov, and S. Saurabh, *Exact algorithms via monotone local search*, in Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016, 2016, pp. 764–775. 1, 2

[19] F. V. Fomin and S. Saurabh, *Kernelization methods for fixed-parameter tractability*, in Tractability, Cambridge Univ. Press, Cambridge, 2014, pp. 260–282. 1

[20] L. R. Ford and D. R. Fulkerson, *Maximal flow through a network*, Canadian journal of Mathematics, 8 (1956), pp. 399–404. 8

[21] L. Fortnow and R. Santhanam, *Infeasibility of instance compression and succinct PCPs for NP*, Journal of Computer and System Sciences, 77 (2011), pp. 91–106. 1

[22] J. Guo and R. Niedermeier, *Invitation to data reduction and problem kernelization*, SIGACT News, 38 (2007), pp. 31–45. 1

[23] D. Hermelin, S. Kratsch, K. Soltys, M. Wahlström, and X. Wu, *A completeness theory for polynomial (turing) kernelization*, Algorithmica, 71 (2015), pp. 702–730. 1

[24] D. Hermelin and X. Wu, *Weak compositions and their applications to polynomial lower bounds for kernelization*, in Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, 2012, pp. 104–113. 1

[25] F. Hüffner, C. Komusiewicz, H. Moser, and R. Niedermeier, *Fixed-parameter algorithms for cluster vertex deletion*, Theory Comput. Syst., 47 (2010), pp. 196–217. 1, 2

[26] B. M. P. Jansen and D. Marx, *Characterizing the easy-to-find subgraphs from the viewpoint of polynomial-time algorithms, kernels, and turing kernels*, in Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015, 2015, pp. 616–629. 5

[27] S. Kratsch, *Recent developments in kernelization: A survey*, Bulletin of the EATCS, 113 (2014). 1

[28] M. Kumar and D. Lokshtanov, *Faster exact and parameterized algorithm for feedback vertex set in tournaments*, in 33rd Symposium on Theoretical Aspects of Computer Science, STACS 2016, February 17-20, 2016, Orléans, France, vol. 47 of LIPIcs, 2016, pp. 49:1–49:13. 2

[29] D. Lokshtanov, N. Misra, and S. Saurabh, *Kernelization - preprocessing with a guarantee*, in The Multivariate Algorithmic Revolution and Beyond - Essays Dedicated to Michael R. Fellows on the Occasion of His 60th Birthday, 2012, pp. 129–161. 1

[30] M. Mnich, V. V. Williams, and L. A. Végh, *A 7/3-approximation for feedback vertex sets in tournaments*, in 24th Annual European Symposium on Algorithms, ESA 2016, August 22-24, 2016, Aarhus, Denmark, vol. 57 of LIPIcs, 2016, pp. 67:1–67:14. 2

[31] H. Moser, *A problem kernelization for graph packing*, in SOFSEM 2009: Theory and Practice of Computer Science, 35th Conference on Current Trends in Theory and Practice of Computer Science, Spindleruv Mlýn, Czech Republic, January 24-30, 2009. Proceedings, vol. 5404 of Lecture Notes in Computer Science, 2009, pp. 401–412. 5

[32] R. Niedermeier, *Invitation to fixed-parameter algorithms*, vol. 31 of Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2006. 1

[33] S. Thomassé, *A $4k^2$ kernel for feedback vertex set*, ACM Transactions on Algorithms, 6 (2010), pp. 32:1–32.8. 3, 7

[34] M. Wahlström, *Algorithms, measures and upper bounds for satisfiability and related problems*, PhD thesis, Linköping University, Sweden, 2007. 1

[35] J. You, J. Wang, and Y. Cao, *Approximate association via dissociation*, Discrete Applied Mathematics, 219 (2017), pp. 202–209. 2