Essentially Tight Kernels For (Weakly) Closed Graphs

Tomohiro Koana

Algorithmics and Computational Complexity, Technische Universität Berlin, Germany

Christian Komusiewicz

Fachbereich Mathematik und Informatik, Philipps-Universität Marburg, Germany

Frank Sommer

Fachbereich Mathematik und Informatik, Philipps-Universität Marburg, Germany

Abstract

We study kernelization of classic hard graph problems when the input graphs fulfill triadic closure properties. More precisely, we consider the recently introduced parameters closure number c and weak closure number γ [Fox et al., SICOMP 2020] in addition to the standard parameter solution size k. The weak closure number γ of a graph is upper-bounded by the minimum of its closure number c and its degeneracy d. For Capacitated Vertex Cover, Connected Vertex Cover, and Induced Matching we obtain the first kernels of size $k^{\mathcal{O}(\gamma)}$, $k^{\mathcal{O}(\gamma)}$, and $(\gamma k)^{\mathcal{O}(\gamma)}$, respectively. This extends previous results on the kernelization of these problems on degenerate graphs. These kernels are essentially tight as these problems are unlikely to admit kernels of size $k^{o(\gamma)}$ by previous results on their kernelization complexity in degenerate graphs [Cygan et al., ACM TALG 2017]. For Capacitated Vertex Cover, we show that even a kernel of size $k^{o(c)}$ is unlikely. In contrast, for Connected Vertex Cover, we obtain a problem kernel with $\mathcal{O}(ck^2)$ vertices. Moreover, we prove that searching for an induced subgraph of order at least k belonging to a hereditary graph class $\mathcal{G}$ admits a kernel of size $k^{\mathcal{O}(\gamma)}$ when $\mathcal{G}$ contains all complete and all edgeless graphs. Finally, we provide lower bounds for the kernelization of Independent Set on graphs with constant closure number c and kernels for Dominating Set on weakly closed split graphs and weakly closed bipartite graphs.

2012 ACM Subject Classification Theory of computation → Parameterized complexity and exact algorithms; Theory of computation → Graph algorithms analysis

Keywords and phrases Fixed-parameter tractability, kernelization, c-closure, weak γ-closure, Independent Set, Induced Matching, Connected Vertex Cover, Ramsey numbers, Dominating Set

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2021.35

Related Version *A continously updated version of the paper is available at*: `https://arxiv.org/abs/2103.03914`

Funding *Tomohiro Koana*: Supported by the Deutsche Forschungsgemeinschaft (DFG), project DiPa, NI 369/21.
Frank Sommer: Supported by the Deutsche Forschungsgemeinschaft (DFG), project EAGR, KO 3669/6.

Acknowledgements We would like to thank the anonymous reviewers of ISAAC'21 for their many helpful remarks that have substantially improved the presentation of the results in this paper.

1 Introduction

A main tool for coping with hard computational problems is to shrink large input data to a computationally hard core by removing easy parts of the instance in polynomial time. Parameterized algorithmics provides the framework of *kernelization* for analyzing the power and limits of polynomial-time data reduction algorithms. In addition to the input instance I, a parameterized problem comes equipped with a parameter k which describes the structure of the input or is a bound on the solution size. A *kernelization* for a parameterized problem L is an algorithm that replaces every input instance (I, k) of L in polynomial time by an equivalent instance (I', k') of L (the kernel) whose size depends only on the parameter k, that is, $|I'| + k' \leq g(k)$ for some computable function g. The kernel is guaranteed to be small if k is small and g grows only modestly. A particularly important special case is thus a kernelization where g is a polynomial function. Such kernels are referred to as *polynomial kernelizations*.

Many problems do not admit a kernel simply because they are believed to be not *fixed-parameter tractable*. That is, it is assumed that they are not solvable in $f(k) \cdot |I|^{\mathcal{O}(1)}$ time. A classic example is Dominating Set parameterized by the solution size k. Moreover, even problems that do admit kernels are known to not admit polynomial kernels [2, 7, 21, 22][1]; a classic example is Connected Vertex Cover parameterized by the solution size k [7].

To devise kernelization algorithms for such problems, one considers either additional parameters or restricted classes of input graphs. One example for this approach is kernelization in degenerate graphs [4, 5, 24]. A graph G is d-degenerate if every subgraph of G contains at least one vertex with degree at most d. Dominating Set, for example, admits a kernel of size $k^{\mathcal{O}(d^2)}$ where d is the degeneracy of the input graph [24]. Thus, the exponent of the kernel size depends on d; we will say that Dominating Set admits a polynomial kernel on d-degenerate graphs. This kernelization was shown to be tight in the sense that there is no kernel of size $k^{o(d^2)}$ [4]. The situation is different for Independent Set which admits a trivial problem kernel with $\mathcal{O}(dk)$ vertices: here the kernel size is polynomial in $d + k$.

Real-world networks have small degeneracy d, making d an interesting parameter from an application point of view. Moreover, bounded degeneracy imposes combinatorial structure that can be exploited algorithmically as evidenced by the discussion above. Recently, Fox et al. [13] discovered two new parameters that share these two features; they are well-motivated from a practical standpoint and describe interesting and useful combinatorial features of graphs. The first parameter is the *closure* of a graph, defined as follows.

▶ **Definition 1.1** ([13, Definition 1.1]). *Let* $\mathsf{cl}_G(v) := \max_{w \in V(G) \setminus N[v]} \{|N(v) \cap N(w)|, 0\}$ *denote the* closure number *of a vertex v in a graph G. A graph G is c-*closed *if* $\mathsf{cl}_G(v) < c$ *for all $v \in V(G)$.*

In other words, a graph is c-closed if every pair of nonadjacent vertices has at most $c - 1$ common neighbors. The parameter models triadic closure in social networks, the observation that people with many common acquaintances are likely to know each other. Fox et al. [13] devised a another parameter, the *weak closure* which relates to c-closure as degeneracy relates to maximum degree: instead of demanding a bounded closure number for every vertex, one demands that every induced subgraph has some vertex with bounded closure number.

▶ **Definition 1.2** ([13, Definition 1.3]). *A graph G is weakly γ-closed if*

- *there exists a* weak closure ordering *$\sigma := v_1, \ldots, v_n$ of the vertices of G such that $\mathsf{cl}_{G_i}(v_i) < \gamma$ for all $i \in [n]$ where $G_i := G[\{v_i, \ldots, v_n\}]$, or, equivalently, if*
- *every induced subgraph G' of G has a vertex $v \in V(G')$ such that $\mathsf{cl}_{G'}(v) < \gamma$.*

The weak closure number *of a graph G is the minimum integer γ such that the graph G is weakly γ-closed.*

[1] All kernelization lower bounds mentioned in this work are based on the assumption coNP $\not\subseteq$ NP/poly.

Let G be a graph and let d, c, and γ be the degeneracy, the closure number and the weak closure number of G. The three parameters d, c, and γ are related as follows:

1. The weak closure number γ is at most $\max(d+1, c)$.
2. The weak closure number γ can be arbitrarily smaller than d as witnessed by large complete graphs.
3. The degeneracy d and the closure number c are incomparable as witnessed by large complete graphs (these have large degeneracy and are 1-closed) and large complete bipartite graphs where one part has size two (these are 2-degenerate and have large closure number).
4. The latter example also shows that γ can be much smaller than the closure number c which is very often the case in real-world data [13, 19].

Akin to degeneracy, c-closure and weak γ-closure have proven to be very useful parameters. In particular, all maximal cliques of a graph can be enumerated in $3^{\gamma/3} \cdot n^{\mathcal{O}(1)}$ time [13]. By the above discussion on the relation of γ and d, this result thus extends the range of tractable clique enumeration instances from the class of bounded-degeneracy graphs [9] to the larger class of graphs with bounded weak closure. The clique enumeration algorithm for weak closed graphs [13] has also been extended to the enumeration of other clique-like subgraphs [16, 19]. Concerning kernelization, in previous work, we showed that Independent Set and Induced Matching admit polynomial kernels with respect to the parameter $k + c$ and that Dominating Set admits a polynomial kernel on c-closed graphs [20]. Later, we extended the kernelization result for Independent Set to parameterization by weak closure. More generally, we showed that $\mathcal{G}$-Subgraph, where one wants to find a subgraph on at least k vertices belonging to $\mathcal{G}$ admits a kernel with $\mathcal{O}(\gamma k^2)$ vertices if $\mathcal{G}$ is closed under taking subgraphs [19]. To the best of our knowledge, this is is the only known kernelization result for the weak closure parameterization.

In this work we study the kernelization of several further hard graph problems on weakly closed graphs. In a nutshell, we provide kernels for a range of problems that have not been considered on weakly closed graphs so far. Our kernels are based on several combinatorial observations on the structure of weakly closed graphs that might be of more general interest.

Our Results. Building on a combinatorial lemma of Frankl and Wilson [14], we obtain a general lemma (Lemma 2.2) which can be used to bound the size of graphs in terms of their vertex cover number and weak closure number. More precisely, we show that in a graph G with vertex cover S of size k and weak closure γ, the number of different neighborhoods in the independent set $I := V(G) \setminus S$ is $k^{\mathcal{O}(\gamma)}$. Lemma 2.2 gives a general strategy for obtaining kernels in weakly closed graphs: Devise reduction rules that 1) bound the size of the vertex cover and 2) decrease the size of neighborhood classes. We also show that Lemma 2.2 can be extended to a more general notion of neighborhood types (Lemma 2.4).

We then show that this strategy helps in obtaining kernels on weakly closed graphs for Capacitated Vertex Cover, Connected Vertex Cover, Connected ℓ-Component Order Connectivity (Connected ℓ-COC), and Induced Matching all parameterized by the natural parameter solution size k. For these problems, polynomial kernels in degenerate graphs are known [4, 5, 11, 17]. Our results thus extend the class of graphs for which polynomial kernels are known for these problems. The kernels have size $k^{\mathcal{O}(\gamma)}$ and $(\gamma k)^{\mathcal{O}(\gamma)}$, respectively, and by previous results on degenerate graphs the dependence on γ in the exponent cannot be avoided [4, 5]. We complement these findings with a study of Capacitated Vertex Cover and Connected Vertex Cover on c-closed graphs. Interestingly, the kernelization complexity of the problems differs: Capacitated Vertex Cover does not admit a kernel of size $\mathcal{O}(k^{\frac{c-1}{2}-\epsilon})$ for all $\epsilon > 0$ whereas Connected Vertex Cover admits a kernel with $\mathcal{O}(ck^2)$ vertices.

Next, we study the kernelization complexity of INDEPENDENT SET on c-closed graphs. We show that INDEPENDENT SET does not admit a kernel of size $\mathcal{O}(k^{2-\epsilon})$ on c-closed graphs for constant c. This complements previous kernels of size $\mathcal{O}(c^2k^3)$ [20] and $\mathcal{O}(\gamma^2k^3)$ [19], narrowing the gap between upper and lower bound for the achievable kernel size on (weakly) closed graphs. We also obtain a lower bound of $\Omega(k^{4/3-\epsilon})$ on the number of *vertices* in the graph in case of constant c and show that at least a linear dependence on c is necessary in any kernelization of INDEPENDENT SET in c-closed graphs: under standard assumptions, there is no kernel of size $c^{(1-\epsilon)} \cdot k^{\mathcal{O}(1)}$. Some of our results also hold for Ramsey-type problems where one wants to find a large subgraph belonging to a class $\mathcal{G}$ containing all complete and all edgeless graphs. In this context, we observe that weakly γ-closed graphs fulfill the Erdős–Hajnal property [10] with a linear dependence on γ: There is a constant q such that every weakly γ-closed graph on $k^{q\gamma}$ vertices has either a clique of size k or an independent set of size k. We believe that this observation is of independent interest and that it will be useful in the further study of weakly γ-closed graphs.

Finally, we consider DOMINATING SET which admits a kernel of size $k^{\mathcal{O}(c)}$ [20]. It is open whether DOMINATING SET admits a kernel of size $k^{f(\gamma)}$ for some function f, which would extend the class of kernelizable input graphs from degenerate to weakly closed. We make partial progress towards answering this question by showing that there is a kernel of size $k^{\mathcal{O}(\gamma^2)}$ on graphs with constant clique number (such as bipartite graphs) and a kernel of size $(\gamma k)^{\mathcal{O}(\gamma)}$ in split graphs. In both cases these bounds are tight in the sense that kernels of size $k^{o(d^2)}$ and of size $k^{o(c)}$ are unlikely to exist [4, 19].

Due to lack of space, several proofs (marked with (*)) and all results for DOMINATING SET on bipartite and split graphs are deferred to the full version of this article.

Preliminaries. By $[n]$ we denote the set $\{1, \ldots, n\}$ for some $n \in \mathbb{N}$. For a graph G, let $V(G)$ denote its *vertex set*, $E(G)$ its *edge set*, and $n := |V(G)|$ the number of vertices. Let $X \subseteq V(G)$ be a vertex set. By $G[X]$ we denote the *subgraph induced* by X and by $G - X := G[V(G) \setminus X]$ we denote the graph obtained by removing the vertices of X. If the vertices of X are pairwise adjacent (nonadjacent), then X is a *clique* (an *independent set*, respectively). We denote by $N_G(X) := \{y \in V(G) \setminus X \mid xy \in E(G), x \in X\}$ the *open neighborhood* of X and by $N_G[X] := N_G(X) \cup X$ the *closed neighborhood* of X. The maximum degree of G is $\Delta_G := \max_{v \in V(G)} \deg_G(v)$. In the remainder of this paper we fix a weak closure ordering σ. Note that such an ordering can be computed in polynomial time [13]. We define $P^\sigma_G(v) := \{u \in N_G(v) \mid u \text{ appears before } v \text{ in } \sigma\}$ and $Q^\sigma_G(v) := \{u \in N_G(v) \mid u \text{ appears after } v \text{ in } \sigma\}$. We say that $P^\sigma_G(v)$ are *prior neighbors* of v and $Q^\sigma_G(v)$ are *posterior neighbors* of v. A *matching* M is a set of vertex-disjoint edges. By $V(M)$ we denote the union of all endpoints of edges in M. We omit the superscripts and subscripts when they are clear from the context. The following observation follows from the definition of weak closure.

▶ **Observation 1.3.** *For nonadjacent vertices $u, v \in V(G)$, it holds that $|Q(u) \cap Q(v)| \leq |Q(u) \cap N(v)| \leq \gamma - 1$.*

Proof. Let G_u and G_v be the graph induced by the vertices that appear after u and v, respectively. We have two cases based on whether u or v appears first in the weak closure ordering σ.

- u precedes v. By the definition of Q-neighbors, we have $Q_G(u) \cap Q_G(v) \subseteq Q_G(u) \cap N_G(v)$ and $Q_G(u) \cap N_G(v) = N_{G_u}(u) \cap N_{G_u}(v)$. Since $|N_{G_u}(u) \cap N_{G_u}(v)| \leq \mathsf{cl}_{G_u}(u)$, it follows from the definition of weak closure that $|N_{G_u}(u) \cap N_{G_u}(v)| \leq \gamma - 1$.

- v precedes u. Clearly, $Q_G(u) \cap P_G(v) = \emptyset$. We then have $Q_G(u) \cap N_G(v) = Q_G(u) \cap (P_G(v) \cup Q_G(v)) = Q_G(u) \cap Q_G(v)$. It follows from the definition of Q-neighbors that $|Q_G(u) \cap Q_G(v)| = |N_{G_u}(u) \cap N_{G_v}(v)| \le |N_{G_v}(u) \cap N_{G_v}(v)| \le \mathsf{cl}_{G_v}(v) \le \gamma - 1$.

We have shown that $|Q(u) \cap Q(v)| \le |Q(u) \cap N(v)| \le \gamma - 1$ for both cases. ◀

A parameterized problem is *fixed-parameter tractable* if every instance (I, k) can be solved in $f(k) \cdot |I|^{\mathcal{O}(1)}$ time for some computable function f. An algorithm with such a running time is an *FPT algorithm*. A *kernelization* is a polynomial-time algorithm which transforms every instance (I, k) of a parameterized language Q into an equivalent instance (I', k') of Q such that $|I'| + k' \le g(k)$ for some computable function g. A *compression* of a parameterized language Q into a language R is an algorithm that takes as input an instance $(x, k) \in \Sigma^* \times \mathbb{N}$ and returns a string y in time polynomial in $|x| + k$ such that $|y|$ is bounded by some polynomial in k, and $y \in R$ if and only if $(x, k) \in Q$. It is widely believed that W[t]-hard problems ($t \in \mathbb{N}$) do not admit an FPT algorithm. For more details on parameterized complexity, we refer to the standard monographs [3, 8].

2 Bounding the Size of Weakly Closed Graphs with Small Twin Sets

Frankl and Wilson [14] proved the following bound on the size of set systems where the number of different intersection sizes is bounded.

▶ **Proposition 2.1** ([14, Theorem 11]). *Let $\mathcal{F}$ be a collection of pairwise distinct subsets of $[n]$ and let $L \subseteq \{0\} \cup [n]$ be some subset. If $|S \cap S'| \in L$ for all distinct $S, S' \in \mathcal{F}$, then $|\mathcal{F}| \in \mathcal{O}(n^{|L|})$.*

We now use this proposition to achieve a bound on the size of weakly closed graphs when every vertex has few false twins and the size of the vertex cover is small. Herein, two vertices u and v are *false twins* if $N(u) = N(v)$.

▶ **Lemma 2.2.** *Let G be a weakly γ-closed graph and let I be an independent set of G. Suppose that each vertex $v \in I$ has at most $t - 1$ false twins. Then, $|I| \in t \cdot \mathcal{O}(3^{\gamma/3} \cdot k^{2\gamma+3})$, where $k := n - |I|$.*

Proof. We say that two vertices $v, v' \in I$ are *P-equivalent*, *Q-equivalent*, and *N-equivalent* if $P(v) = P(v')$, $Q(v) = Q(v')$, and $N(v) = N(v')$, respectively. Let $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{N}$ denote the collection of P-equivalence, Q-equivalence, and N-equivalence classes, respectively. We extend the notation of P, Q, and N to an equivalence classes A by defining $P(A) := P(v)$, $Q(A) := Q(v)$, and $N(A) := N(v)$ for some $v \in A$. Since there is at most one N-equivalence class for every pair of P-equivalent and Q-equivalent classes, we have $|\mathcal{N}| \le |\mathcal{P}| \cdot |\mathcal{Q}|$. By the assumption that there are at most t vertices in each N-equivalence class, we also have $|I| \le t \cdot |\mathcal{N}|$. Thus, it suffices to show suitable bounds on $|\mathcal{P}|$ and $|\mathcal{Q}|$.

First, we prove that $|\mathcal{Q}| \in \mathcal{O}(k^\gamma)$, using the result of Frankl and Wilson (Proposition 2.1 [14]). Since $I \supseteq A$ is an independent set, $Q(A) \subseteq S := V(G) \setminus I$. Moreover, for two distinct Q-equivalence classes A and A', we have $|Q(A) \cap Q(A')| < \gamma$ by Observation 1.3, and equivalently, $|Q(A) \cap Q(A')| \in L$ for $L := \{0\} \cup [\gamma - 1]$. By Proposition 2.1 we obtain $|\mathcal{Q}| \in \mathcal{O}(|S|^{|L|}) = \mathcal{O}(k^\gamma)$.

Next, we bound the size of $\mathcal{P}$. Let $I_0 := \{v \in I \mid \exists u, w \in P(v) \colon uw \notin E(G)\}$ be the set of vertices in I with nonadjacent prior neighbors. By the definition of weak γ-closure, there are at most $\gamma - 1$ vertices of I_0 for every pair of nonadjacent vertices in S. Thus, we have $|I_0| < \gamma\binom{|S|}{2} \in \mathcal{O}(\gamma k^2)$.

Let $I_1 := I \setminus I_0$ and let $\mathcal{P}_1$ be the collection of P-equivalence classes in I_1. Note that for every $A \in \mathcal{P}_1$, its neighborhood $P(A)$ is a clique. Since a weakly γ-closed graph on n vertices has $\mathcal{O}(3^{\gamma/3}n^2)$ maximal cliques [13], there are $\mathcal{O}(3^{\gamma/3}k^2)$ equivalence classes A such that $P(A)$ constitutes a maximal clique in $G[S]$. Consider an equivalence class A such that $P(A) \subset C$ for some maximal clique C in $G[S]$. We will show that there are $k^{\mathcal{O}(\gamma)}$ such equivalence classes. Let u be the first vertex of $C \setminus P(A)$ in the weak closure ordering σ. Since $P(A) \subset C \subseteq N(u) = P(u) \cup Q(u)$, we have $P(A) = (P(A) \cap P(u)) \cup (P(A) \cap Q(u))$. As $P(A) \cap P(u) = C \cap P(u)$ by the choice of u, we can rewrite $P(A) = (C \cap P(u)) \cup B$, where $B := P(A) \cap Q(u)$. Thus, there is at most one equivalence class of $\mathcal{P}_1$ for every maximal clique C in $G[S]$, vertex $u \in S$, and vertex subset $B \subseteq S$, and thereby, we have $|\mathcal{P}_1| \in \mathcal{O}(3^{\gamma/3}k^2 \cdot k \cdot b)$, where b denotes the number of choices for B. Observe that $P(A) = P(v)$ for some vertex $v \in I_1$ and thus that $B = Q(u) \cap P(v) \subseteq Q(u) \cap N(v)$. Recall that u and v are not adjacent by the choice of u. It follows that $|B| \leq |Q(u) \cap N(v)| \leq \gamma - 1$ by Observation 1.3, and hence $b \in \mathcal{O}(k^\gamma)$ and $|\mathcal{P}_1| \in \mathcal{O}(3^{\gamma/3} \cdot k^3 \cdot k^\gamma) = \mathcal{O}(3^{\gamma/3} \cdot k^{\gamma+3})$. Overall, we have $|\mathcal{P}| \leq (|I_0| + |\mathcal{P}_1|) \in \mathcal{O}(3^{\gamma/3} \cdot k^{\gamma+3})$. The total number of N-equivalence classes is thus at most $|\mathcal{Q}| \cdot |\mathcal{P}| \in \mathcal{O}(3^{\gamma/3} \cdot k^{2\gamma+3})$. ◀

We now show that Proposition 2.1 can be also applied to bound the graph size in terms of the ℓ-COC number, which is the smallest size of a vertex set S such that every connected component in $G - S$ has size at most ℓ, where ℓ is a fixed constant. The 1-COC number is the vertex cover number. To obtain this generalization, we extend the notion of twins.

▶ **Definition 2.3.** *Let $G = (V, E)$ be a graph and let $A, B \subseteq V(G)$ such that $|A| = |B| = \ell$. The sets A and B are* ℓ-twins *if there exists an ordering $a_1, \dots, a_\ell$ of A and an ordering $b_1, \dots, b_\ell$ of B such that $N(a_i) \setminus A = N(b_i) \setminus B$ for each $i \in [\ell]$.*

Note that u and v are false twins if and only if $\{u\}$ and $\{v\}$ are 1-twins.

▶ **Lemma 2.4 (*).** *Let G be a graph and let $D \subseteq V(G)$ be such that each connected component in $G[D]$ has size at most ℓ. Suppose that for every connected component Z in $G[D]$, there are at most $t-1$ other connected components Z' in $G[D] - Z$ such that Z and Z' are $|Z|$-twins. Then, $|D| \in \mathcal{O}(t \cdot k^{\mathcal{O}(\gamma)})$, where $k = n - |D|$.*

3 Applications of our Framework

We now apply Lemma 2.2 and Lemma 2.4 to obtain kernels for several well-known problems.

3.1 Capacitated Vertex Cover

The first problem to which we apply Lemma 2.2 is Capacitated Vertex Cover.

Capacitated Vertex Cover
Input: A graph G, a capacity function $\mathsf{cap}\colon V(G) \to \mathbb{N}$, and $k \in \mathbb{N}$.
Question: Is there a set S of at most k vertices and a function f mapping each edge of $E(G)$ to one of its endpoints in S such that $|\{e \in E(G) \mid f(e) = v\}| \leq \mathsf{cap}(v)$ for all $v \in S$?

Capacitated Vertex Cover admits a kernel with $\mathcal{O}(k^{d+1})$ vertices. Furthermore, this kernel is essentially tight: a kernel with $\mathcal{O}(k^{d-\epsilon})$ vertices would imply coNP $\subseteq$ NP/poly [4]. We will show that the reduction rule used to obtain a kernel in degenerate graphs also leads to a kernel in graphs with bounded weak closure. One may view this result as a way of showing that the rules are more powerful than what was previously known. The kernel uses the following rule.

▶ **Reduction Rule 3.1** ([4, Rule 3]). *If $S \subseteq V(G)$ is a subset of false twin vertices with a common neighborhood $N(S)$ such that $|S| = k + 2 \geq |N(S)|$, then remove a vertex with minimum capacity in S from G, and decrease all the capacities of vertices in $N(S)$ by one.*

We omit the proof for the correctness of Reduction Rule 3.1, referring to Cygan et al. [4, Lemma 20]. One can easily verify that it does not increase the weak γ-closure. In the following theorem, we show that Reduction Rule 3.1 indeed gives us a kernel with $k^{\mathcal{O}(\gamma)}$ vertices.

▶ **Theorem 3.2.** Capacitated Vertex Cover *has a kernel of size $k^{\mathcal{O}(\gamma)}$.*

Proof. We show that a Yes-instance which is reduced with respect to Reduction Rule 3.1 has size $k^{\mathcal{O}(\gamma)}$. Let S be a capacitated vertex cover of size at most k of (G, cap). Let $I := V(G) \setminus S$. By definition, I is an independent set and $N(v) \subseteq S$ for all $v \in I$. Moreover, since (G, cap) is reduced with respect to Reduction Rule 3.1 there is no set of $k + 2$ vertices in I that have the same neighborhood. Hence, I fulfills the condition of Lemma 2.2 with $t = k + 2$. Thus, $|I| \in k \cdot k^{\mathcal{O}(\gamma)}$ which implies $|V(G)| = |S| + |I| \in k^{\mathcal{O}(\gamma)}$. ◀

We also show that this kernel is essentially tight even if γ is replaced by c.

▶ **Theorem 3.3 (*).** *For $c \geq 4$,* Capacitated Vertex Cover *has no kernel of size $\mathcal{O}(k^{\frac{c-1}{2}-\epsilon})$ unless* coNP $\subseteq$ NP/poly.

3.2 Connected Vertex Cover

We now provide kernels for Connected Vertex Cover, a well-studied variant of Vertex Cover which is notoriously hard and does not admit a polynomial kernel when parameterized k [7].

Connected Vertex Cover
Input: A graph G and $k \in \mathbb{N}$.
Question: Is there a vertex cover S of size at most k in G such that $G[S]$ is connected?

We will show that by applying Lemma 2.2 we obtain a kernel of size $k^{\mathcal{O}(\gamma)}$. We may use the following known rule.

▶ **Reduction Rule 3.4** ([4, Rule 2]). *If $S \subseteq V(G)$ is a set of at least two twin vertices with a common neighborhood $N(S)$ such that $|S| > |N(S)|$, then remove one vertex v of S from G.*

After exhaustive application of Reduction Rule 3.4 we have, again by Lemma 2.2, that every Yes-instance has size $k^{\mathcal{O}(\gamma)}$. The proof is completely analogous to that of Theorem 3.2.

▶ **Theorem 3.5.** Connected Vertex Cover *admits a kernel of size $k^{\mathcal{O}(\gamma)}$.*

This kernel is essentially tight, because there is no kernel of size $k^{o(d)}$ [4]. We now show a polynomial kernel for $k + c$.

▶ **Theorem 3.6 (*).** Connected Vertex Cover *has a kernel with $\mathcal{O}(ck^2)$ vertices.*

This result stands in contrast to Capacitated Vertex Cover, which has no kernel of size $k^{o(c)}$ unless coNP $\subseteq$ NP/poly (Theorem 3.3).

An Extension to Connected ℓ-COC. In Connected ℓ-COC the task is to find a set S of at most k vertices such that $G[S]$ is connected and every connected component of $G - S$ has size at most ℓ, where ℓ is a fixed constant. We show that this problem also admits a kernel of size $k^{\mathcal{O}(\gamma)}$. The main idea lies in the extension of Reduction Rule 3.4:

▶ **Reduction Rule 3.7.** *Let $T_1, \ldots, T_x \subseteq V(G)$ be a set of x many r-twins, for some $r \in [\ell]$. If $x \geq k + \ell + 2$, then remove all vertices in T_x from G.*

Note that Reduction Rule 3.7 can be exhaustively performed in polynomial time since the r-twin relation can be computed in $n^{2r+\mathcal{O}(1)}$ time. We then obtain the following theorem from Lemma 2.4.

▶ **Theorem 3.8 (*).** Connected ℓ-COC *has a kernel of size $k^{\mathcal{O}(\gamma)}$ for constant ℓ.*

3.3 Induced Matching

In this section, we provide a kernel of size $(\gamma k)^{\mathcal{O}(\gamma)}$ for Induced Matching:

Induced Matching
Input: A graph G and $k \in \mathbb{N}$.
Question: Is there a set M of at least k edges such that the endpoints of distinct edges are pairwise nonadjacent?

Induced Matching is W[1]-hard for the parameter k on general graphs. For c-closed graphs, we developed a kernel with $\mathcal{O}(c^7 k^8)$ vertices [20]. For d-degenerate graphs, Kanj et al. [17] and Erman et al. [11] independently presented kernels of size $k^{\mathcal{O}(d)}$. Later, Cygan et al. [4] provided a matching lower bound $k^{o(d)}$ on the kernel size. Note that this also implies the nonexistence of $k^{o(\gamma)}$-size kernels unless coNP $\subseteq$ NP/poly.

It turns out that Lemma 2.2 is again helpful in designing a $k^{\mathcal{O}(\gamma)}$-size kernel for Induced Matching. In a nutshell, we show that the application of a series of reduction rules results in a graph with a $(\gamma k)^{\mathcal{O}(1)}$-size vertex cover. We do so by combining the kernelization of Erman et al. [11] for degenerate graphs with our previous one for c-closed graphs [20]. Lemma 2.2 and the reduction rule which removes twin vertices then give us a kernel of size $(\gamma k)^{\mathcal{O}(\gamma)}$.

Erman et al. [11] use the following observation for degenerate graphs.

▶ **Lemma 3.9** ([11, Proof of Theorem 2.10])**.** *Any graph G with a matching M has an induced matching of size $|M|/(4d_G + 1)$.*

Ideally, we would like to prove a lemma analogous to Lemma 3.9 on weakly γ-closed graphs. Note, however, that a complete graph on n vertices (which is weakly 1-closed) has no induced matching of size 2, although it contains a matching of size $\lfloor n/2 \rfloor$. So we follow a different route, and prove an analogous lemma on weakly γ-closed bipartite graphs (there exist bipartite 2-closed graphs whose degeneracy is unbounded; see e.g. Eschen et al. [12]). As we shall see, this serves our purposes.

▶ **Lemma 3.10.** *Suppose that G is a bipartite graph with a bipartition (A, B). If G has a matching M of size $f_\gamma(k) := 4\gamma k^2 + 3k$, then G has an induced matching of size k.*

Proof. Recall that $Q^\sigma(v) := \{u \in N(v) \mid u \text{ appears after } v \text{ in } \sigma\}$. Let $S \subseteq V(G)$ be the set of vertices v such that $|Q(v)| \geq \gamma k$. Suppose that $|S| \geq 2k$. Then, we may assume that $|A \cap S| \geq k$. Let $A' \subseteq A \cap S$ be an arbitrary vertex set of size exactly k and consider some vertex $v \in A'$. Since $|Q(v) \cap N(v')| < \gamma$ for every $v' \in A' \setminus \{v\}$ by Observation 1.3, we have $|Q(v) \setminus \bigcup_{v' \in A' \setminus \{v\}} N(v')| > 0$ for each $v \in A'$. Consequently, there is at least one vertex $q_v \in Q(v) \setminus \bigcup_{v' \in A' \setminus \{v\}} N(v')$. Then, the edge set $\{vq_v \mid v \in A'\}$ forms an induced matching of size k in G.

Now, consider the case $|S| < 2k$. By the definition of S, it holds that $|Q_{G-S}(v)| \leq |Q_G(v)| \leq \gamma k$ for each vertex $v \in V(G) \setminus S$. Hence, the degeneracy of $G - S$ is at most γk. Since $G - S$ has a matching M_{G-S} of size at least $|M| - |S| \geq f_\gamma(k) - 2k = 4\gamma k^2 + k$, Lemma 3.9 yields an induced matching of size $|M_{G-S}|/(4d_{G-S} + 1) \geq k$. ◀

We use the following reduction rule to sparsify the graph G so that every sufficiently large vertex set contains a large independent set (see Lemma 3.13).

▶ **Reduction Rule 3.11.** *If for some vertex $v \in V(G)$, there is a maximum matching M_v of size at least $2\gamma k$ in $G[Q(v)]$, then delete v.*

▶ **Lemma 3.12.** *Reduction Rule 3.11 is correct.*

Proof. Let $G' := G - v$. Suppose that G has an induced matching M of size k. If $v \notin V(M)$, then M is also an induced matching in G'. So assume that $vv' \in V(M)$ for some vertex $v' \in V(G)$. Then, we have $|N(u) \cap Q(v)| < \gamma$ for any vertex $u \in V(M \setminus \{vv'\})$ by Observation 1.3 and thus $|V(M_v) \setminus \bigcup_{u \in V(M \setminus \{vv'\})} N(u)| \geq 2|M_v| - (\gamma - 1)(2k - 2) > |M_v|$. By the pigeon-hole principle, this implies that there is an edge $e \in M_v$ not incident with any vertex in $V(M)$ and no endpoint of e is adjacent to any vertex in $V(M \setminus \{vv'\})$. Then, $(M \setminus \{vv'\}) \cup \{e\}$ is an induced matching of size k in G'. ◀

▶ **Lemma 3.13.** *Suppose that G is a graph in which Reduction Rule 3.11 is applied on every vertex. Then, every vertex set $S \subseteq V(G)$ of size at least $g_\gamma(k) := 4\gamma k^2 + k^2$ contains an independent set $I \subseteq S$ of size k.*

Proof. Suppose that there is no independent set of size k in $G' := G[S]$ for some vertex set S of size $g_\gamma(k)$. For every vertex $v \in S$, let M_v be a maximum matching in $Q_{G'}(v)$ and let $I_v := Q_{G'}(v) \setminus V(M_v)$. By Reduction Rule 3.11, we have $|V(M_v)| = 2|M_v| \leq 4\gamma k$. Since I_v is an independent set, we then have $|Q_{G'}(v)| = |M_v| + |I_v| < 4\gamma k + k$ for every vertex $v \in S$, and thus $d_{G'} < 4\gamma k + k$. Note, however, that G' has an independent set of size $|S|/(d_{G'} + 1) \geq k$, which is a contradiction. ◀

To identify a part of the graph with a sufficiently large induced matching, we rely on the LP relaxation of Vertex Cover, following our approach [20] to obtain a polynomial kernel on c-closed graphs. Recall that Vertex Cover can be formulated as an integer linear program as follows, using a variable x_v for each $v \in V(G)$:

$$\min \sum_{v \in V(G)} x_v \qquad \text{subject to} \quad x_u + x_v \geq 1 \quad \forall uv \in E(G), \qquad x_v \in \{0, 1\} \quad \forall v \in V(G).$$

We will refer to the LP relaxation of Vertex Cover as VCLP. We use the well-known facts that VCLP always admits an optimal solution in which $x_v \in \{0, 1/2, 1\}$ for each $v \in V(G)$ and that such a solution can be found in polynomial time. Suppose that we have such an optimal solution $(x_v)_{v \in V(G)}$. Let $V_0 := \{v \in V(G) \mid x_v = 0\}$, $V_1 := \{v \in V(G) \mid x_v = 1\}$, and $V_{1/2} := \{v \in V(G) \mid x_v = 1/2\}$. Also, let $\mathsf{opt}(G)$ be the optimum of VCLP. We show that we can immediately return Yes, whenever $\mathsf{opt}(G)$ is sufficiently large:

▶ **Reduction Rule 3.14.** *If $\mathsf{opt}(G) \geq 2g_\gamma(g_\gamma(f_\gamma(k)))$, then return Yes.*

Here, the functions f_γ and g_γ are as specified in Lemmas 3.10 and 3.13, respectively.

▶ **Lemma 3.15.** *Reduction Rule 3.14 is correct.*

Proof. We show that G has an induced matching of size k whenever $\mathsf{opt}(G) \geq 2g_\gamma(g_\gamma(f_\gamma(k)))$. Let M be an arbitrary maximal matching in G. Since $V(M)$ is a vertex cover, we have $\mathsf{opt}(G) \leq |V(M)| = 2|M|$, and hence $|M| \geq \mathsf{opt}(G)/2 \geq g_\gamma(g_\gamma(f_\gamma(k)))$. Let $M := \{a_1b_1, \ldots, a_{|M|}b_{|M|}\}$ and let $A := \{a_1, \ldots, a_{|M|}\}$ and $B := \{b_1, \ldots, b_{|M|}\}$. By Lemma 3.13, there exists an independent set $A' \subseteq A$ of size $s' := g_\gamma(f_\gamma(k))$. Without loss of generality, suppose that $A' = \{a'_1, \ldots, a'_{s'}\}$ and let $B' := \{b'_1, \ldots, b'_{s'}\}$ be the set of vertices matched to A' in M. Again by Lemma 3.13, we obtain an independent set $B'' \subseteq B'$ of size $s'' := f_\gamma(k)$. We assume without loss of generality that $B'' = \{b''_1, \ldots, b''_{s''}\}$. Let $A'' := \{a''_1, \ldots, a''_{s''}\}$ be the set of vertices matched to B'' in M. Then, $G[A'' \cup B'']$ is a bipartite graph with a matching of size at least $s'' = f_\gamma(k)$. By Lemma 3.10, $G[A'' \cup B'']$ has an induced matching of size k. ◀

Since $\mathsf{opt}(G) = |V_{1/2}|/2 + |V_1|$, it holds that $|V_{1/2}|/2 + |V_1| \leq 2g_\gamma(g_\gamma(f_\gamma(k))) \in \mathcal{O}(\gamma^7 k^8)$ after the application of Reduction Rule 3.14. Hence, it remains to bound the size of V_0. To do so, it suffices to remove twins:

▶ **Reduction Rule 3.16.** *If $N(u) = N(v)$ for some vertices $u, v \in V(G)$, then delete v.*

Since an induced matching contains at most one of u and v, the rule is obviously correct. We are finally ready to utilize Lemma 2.2 to derive an upper bound on V_0: Since V_0 is an independent set, Lemma 2.2 gives us $|V_0| \in |V_{1/2} \cup V_1|^{\mathcal{O}(\gamma)} \in (\gamma k)^{\mathcal{O}(\gamma)}$. Thus, we have the following result.

▶ **Theorem 3.17.** Induced Matching *has a kernel of size $(\gamma k)^{\mathcal{O}(\gamma)}$.*

4 Independent Set and Ramsey-Type Problems

We now investigate the kernelization complexity of Independent Set, where we are given a graph G and an integer k, and ask whether G has an independent set of size k. Independent Set admits a kernel with $\mathcal{O}(ck^2)$ vertices and $\mathcal{O}(c^2k^3)$ edges [20] and a kernel with $\mathcal{O}(\gamma k^2)$ and $\mathcal{O}(\gamma^2 k^3)$ edges [19]. We show a lower bound for these parameterizations: unless coNP $\subseteq$ NP/poly, Independent Set admits no kernel of size $k^{2-\epsilon}$ and no kernel with $k^{4/3-\varepsilon}$ vertices even if the c-closure is constant. We also show a kernel lower bound of size $c^{1-\varepsilon}k^{\mathcal{O}(1)}$ for any ε. We also consider the following related problem where $\mathcal{G}$ is a hereditary graph class containing all complete graphs and edgeless graphs.

$\mathcal{G}$-Subgraph
Input: A graph G and $k \in \mathbb{N}$.
Question: Is there a set S of at least k vertices such that $G[S] \in \mathcal{G}$?

Khot and Raman [18] showed that $\mathcal{G}$-Subgraph is FPT when parameterized by k, using Ramsey's theorem: for any $k \in \mathbb{N}$, any graph G on at least $R(k) \in 2^{\mathcal{O}(k)}$ vertices contains a clique of size k or an independent set of size k. Ramsey, the special case where $\mathcal{G}$ is the family of all complete and edgeless graphs, admits no polynomial kernel unless coNP $\subseteq$ NP/poly [21]. Similarly, $\mathcal{G}$-Subgraph admits no polynomial kernel for several graph classes $\mathcal{G}$, such as cluster graphs [23]. Our contribution for $\mathcal{G}$-Subgraph is two-fold: First, we observe that the lower bounds for Independent Set on graphs with constant c-closure also

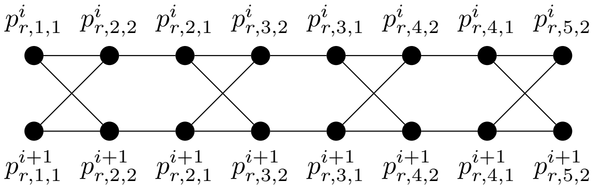

Figure 1 An illustration of P^i_r and P^{i+1}_r for $t = 5$.

hold for Ramsey. This complements a kernel for $\mathcal{G}$-Subgraph with $\mathcal{O}(ck^2)$ vertices [20].[2] Second, we provide a kernel of size $k^{\mathcal{O}(\gamma)}$. To show our kernel lower bounds, we will use *weak q-compositions*. Weak q-compositions exclude kernels of size $\mathcal{O}(k^{q-\varepsilon})$ for $\varepsilon > 0$.

▶ **Definition 4.1** ([6, 15]). *Let $q \geq 1$ be an integer, let $L_1 \subseteq \{0,1\}^*$ be a decision problem, and let $L_2 \subseteq \{0,1\}^* \times \mathbb{N}$ be a parameterized problem. A* weak q-composition *from L_1 to L_2 is a polynomial-time algorithm that on input $x_1, \dots, x_{t^q} \in \{0,1\}^n$ outputs an instance $(y, k') \in \{0,1\}^* \times \mathbb{N}$ such that:*

- *$(y, k') \in L_2 \Leftrightarrow x_i \in L_1$ for some $i \in [t^q]$, and*
- *$k' \leq t \cdot n^{\mathcal{O}(1)}$.*

▶ **Lemma 4.2** ([3, 6, 15]). *Let $q \geq 1$ be an integer, let $L_1 \subseteq \{0,1\}^*$ be an NP-hard problem, and let $L_2 \subseteq \{0,1\}^* \times \mathbb{N}$ be a parameterized problem. If there is a weak q-composition from L_1 to L_2, then L_2 has no compression of size $\mathcal{O}(k^{q-\epsilon})$ for any $\epsilon > 0$, unless* coNP $\subseteq$ NP/poly.

Weak Composition. We give a weak composition from the following problem:

Multicolored Independent Set
Input: A graph G and a partition $(V_1, \dots, V_k)$ of $V(G)$ into k cliques.
Question: Is there an independent set of size exactly k?

A standard reduction from a restricted variant of 3-SAT (for instance, each literal appears exactly twice [1]) shows that Multicolored Independent Set is NP-hard even when $\Delta_G \in \mathcal{O}(1)$ and $|V_i| \in \mathcal{O}(1)$ for all $i \in [k]$. Let $[t]^q$ be the set of q-dimensional vectors whose entries are in $[t]$. Suppose that $q \geq 2$ is a constant and that we are given t^q instances $\mathcal{I}_x = (G_x, (V^1_x, \dots, V^k_x))$ for $x \in [t]^q$, where $\Delta_{G_x} \in \mathcal{O}(1)$ and $|V^i_x| \in \mathcal{O}(1)$ for all $x \in [t]^q$ and $i \in [k]$. We construct an Independent Set instance (H, k'). The kernel lower bound of size $k^{2-\epsilon}$ will be based on the special case $q = 2$. To obtain the lower bound of $c^{1-\epsilon}k^{\mathcal{O}(1)}$, however, we need the composition to work for all $q \in \mathbb{N}$. Hence, we give a generic description in the following. First, we construct a graph H_i as follows for every $i \in [k]$:

- For every $x \in [t]^q$, include V^i_x into $V(H_i)$.
- For every $r \in [q]$, introduce a path P^i_r on $2t - 2$ vertices. We label the $(2j-1)$-th vertex as $p^i_{r,j,1}$ and the $2j$-th vertex as $p^i_{r,j+1,2}$ (see Figure 1 for an illustration). Note that $V(P^i_r) = \{p^i_{r,j,1}, p^i_{r,j+1,2} \mid j \in [t-1]\}$. For every $j \in [t]$, we new define the set $P^i_{r,j}$: let $P^i_{r,1} = \{p^i_{r,1,1}\}$, $P^i_{r,t} = \{p^i_{r,t,2}\}$, and $P^i_{r,j} = \{p^i_{r,j,1}, p^i_{r,j,2}\}$ for $j \in [2, t-2]$.
- For every $r \in [q]$ and $j \in [t]$, add edges such that $P^i_{r,j} \cup \bigcup_{x \in [t]^q, x_r = j} V^i_x$ forms a clique (see Figure 1 for an illustration).

[2] Any c-closed n-vertex graph contains a clique or an independent set of size $\Omega(\sqrt{n/c})$ [20].

Now, construct H by taking the disjoint union of the H_i, $i \in [k]$, and adding the following:

- For every $x \in [t]^q$, add edges such that $H[V(G_x)] = G_x$.
- For every $i \in [k-1]$, $r \in [q]$, and $j \in [t-1]$, add edges $p^i_{r,j,1}p^{i+1}_{r,j+1,2}$ and $p^{i+1}_{r,j,1}p^i_{r,j+1,2}$.

This concludes the construction of H. Let $k' := qkt - qk + k$.

We call the vertices of $\bigcup_{x\in[t]^q, i\in[k]} V^i_x$ the *instance vertices*. The other vertices, which are on P^i_r for some $i \in [k]$ and $r \in [q]$, serve as *instance selectors*: As we shall see later, any independent set J of size k' in H contains exactly $t-1$ vertices of P^i_r for every $i \in [k]$ and $r \in [q]$. In fact, there is exactly one $j \in [t]$ such that $J \cap P^i_{r,j} = \emptyset$ and $|J \cap P^i_{r,j'}| = 1$ for all $j' \in [t] \setminus \{j\}$. Consequently, J contains no instance vertex in V^i_x for $x_r \neq j$, and thereby, j is *selected* for the r-th dimension. We bound the c-closure of H and prove the correctness.

▶ **Lemma 4.3 (*).** *It holds that* $\mathsf{cl}_H \in \mathcal{O}(t^{q-2})$.

▶ **Lemma 4.4 (*).** *The graph G_x has a multicolored independent set of size k for some $x \in [t]^q$ if and only if the graph H has an independent set I of size k'.*

For $q = 2$, we have a weak 2-composition from Multicolored Independent Set to Independent Set on $\mathcal{O}(t^{q-2}) = \mathcal{O}(1)$-closed graphs by Lemmas 4.3 and 4.4. Since the constructed graph H has no clique of size k', the construction also constitutes a weak 2-composition to Ramsey on $\mathcal{O}(1)$-closed graphs. Thus, Lemma 4.2 implies the following:

▶ **Theorem 4.5.** *For any $\varepsilon > 0$, neither* Independent Set *nor* Ramsey *has a kernel of size $k^{2-\varepsilon}$ on graphs of constant c-closure, unless* coNP $\subseteq$ NP/poly.

By Theorem 4.5, neither Independent Set nor Ramsey admit a kernel of $k^{1-\varepsilon}$ vertices. We improve this bound on the number of vertices, taking advantage of the fact that any n-vertex c-closed graph can be encoded using $\mathcal{O}(cn^{1.5}\log n)$ bits in polynomial time [12]. Assume for a contradiction that Independent Set or Ramsey admit a kernel of $k^{4/3-\varepsilon'}$ vertices for constant c. Using the above-mentioned encoding, we obtain a string with $\mathcal{O}(k^{(4/3-\varepsilon')1.5}\log k) = \mathcal{O}(k^{2-\varepsilon})$ bits. So a kernel of $k^{4/3-\varepsilon'}$ vertices implies that there is a compression of Independent Set or Ramsey with bitsize $\mathcal{O}(k^{2-\varepsilon})$, a contradiction. Thus, we have the following:

▶ **Theorem 4.6.** *For any $\varepsilon > 0$, neither* Independent Set *nor* Ramsey *has a compression of $k^{4/3-\varepsilon}$ vertices on graphs of constant c-closure, unless* coNP $\subseteq$ NP/poly.

We also obtain another kernel lower bound for Independent Set; this bound excludes the existence of polynomial kernels (in terms of $c + k$) whose dependence on c is sublinear.

▶ **Theorem 4.7.** *For any $\varepsilon > 0$,* Independent Set *has no kernel of size $c^{1-\varepsilon}k^{\mathcal{O}(1)}$ unless* coNP $\subseteq$ NP/poly.

Proof. We show that Independent Set admits no kernel of size $c^{1-\varepsilon}k^i$ for any $\varepsilon, i > 0$, unless coNP $\subseteq$ NP/poly. Let q be a sufficiently large integer with $\frac{q-\varepsilon-i}{q-2} > 1 - \varepsilon$ (that is, $q > \frac{i+3\varepsilon-2}{\varepsilon}$). Recall, that in the constructed instance we have $c \in \mathcal{O}(t^{q-2})$ and $k' := qkt - qk + k$. A straightforward calculation shows that $\ell := c^{\frac{1}{q-2}\left(1-\frac{i}{q-\varepsilon}\right)}k'^{\frac{i}{q-\varepsilon}} \in \mathcal{O}(t)$, and hence Independent Set admits a weak q-decomposition for the parameterization ℓ. Thus, Lemma 4.2 implies that there is no kernel of size $\ell^{q-\varepsilon} = c^{\frac{q-\varepsilon-i}{q-2}}k'^i > c^{1-\varepsilon}k'^i$. ◀

Finally, we show that $\mathcal{G}$-Subgraph has a kernel of size $k^{\mathcal{O}(\gamma)}$ for any graph class $\mathcal{G}$ containing all complete graphs and empty graphs.

▶ **Proposition 4.8.** *Any graph G on at least $R_\gamma(a,b) \in (a \cdot b)^{\gamma+\mathcal{O}(1)}$ vertices has a clique of size a or an independent set of size b.*

Proof. The Ramsey number $R(a,b)$ denotes the smallest number such that every graph on $R(a,b)$ vertices contains a clique of size a or an independent set of size b. It is known that $R(a,b) \leq \binom{a+b}{b}$. So whenever $a \leq \gamma$ or $b \leq \gamma$, we have $R_\gamma(a,b) \leq \binom{a+b}{\gamma} \in \mathcal{O}((a+b)^\gamma)$. For $a,b > \gamma$, let $R_\gamma(a,b)$ be some number greater than $ab\binom{b}{\gamma} + b\binom{a}{\gamma}\sum_{b' \in [b]}\binom{b'-1}{\gamma}$.

Let $n = |V(G)|$ and let $v_1,\dots,v_n$ be a weak closure ordering σ of G. Divide $V(G)$ into b subsets $V_1,\dots,V_b$ of equal size[3]: let $V_i = \{v_{((b-i)n/b)+1},\dots,v_{(b-i+1)n/b}\}$ for each $i \in [b]$. Notably, V_1 is the set of n/b vertices occurring last in σ and V_b is the set of n/b vertices occurring first in σ. Moreover, let $G_i := G[\{v_{(b-i)n/b+1},\dots,v_n\}]$ be the subgraph induced by $\bigcup_{i' \in [i]} V_{i'}$ for each $i \in [b]$. Suppose that G contains no clique of size a. We show that G contains an independent set of size b. More precisely, we prove by induction that G_i contains an independent set of size i for each $i \in [b]$.

This clearly holds for $i = 1$. For $i > 1$, assume that there is an independent set I of size $i-1$ in G_{i-1} by the induction hypothesis. In the following, we consider subsets X of size at most γ of I to obtain an independent set I' of size at least i.

First, consider vertex sets $X \subseteq I$ of size γ and $V_X := \{v \in V_i \mid N_G(v) \supseteq X\}$. Note that $X \subseteq Q(v)$ for each $v \in V_X$. Hence, since G is weakly γ-closed, V_X is a clique. It follows that $|V_X| < a$. Therefore, less than $a\binom{b}{\gamma}$ vertices of V_i are adjacent to at least γ vertices in I.

Second, consider vertex sets $X \subseteq I$ with $X \neq \emptyset$ and $|X| < \gamma$. Furthermore, let $V'_X = \{v \in V_i \mid N_G(v) \cap I = X\}$. Since $n > R_\gamma(a,b)$, there exists $X \subseteq I$ of size at most $\gamma - 1$ such that $|V'_X| > R(a,\gamma)$. By Ramsey's theorem, we then find an independent set $I' \subseteq V'_X$ of size γ (recall that G has no clique of size a). It follows that $(I \setminus X) \cup I'$ is an independent set of size at least i in G_i. ◀

Now, we directly obtain to a kernel for $\mathcal{G}$-Subgraph where $\mathcal{G}$ contains all cliques and all independent sets since each graph on $k^{\mathcal{O}(\gamma)}$ vertices contains either a clique or an independent set of size at least k by the bound on the Ramsey number in weakly closed graphs shown in Proposition 4.8.

▶ **Corollary 4.9.** *Let $\mathcal{G}$ be a class of graphs containing all cliques and independent sets. $\mathcal{G}$-Subgraph has a kernel of size $k^{\mathcal{O}(\gamma)}$.*

5 Conclusion

We have provided several kernelization algorithms and kernelization lower bounds for classic graph problems on (weakly) closed graphs. How far can our results for Connected Vertex Cover and Capacitated Vertex Cover be extended to other cases of connected or capacitated vertex deletion problems? We did show that Connected ℓ-COC admits a kernel of size $k^{\mathcal{O}(\gamma)}$. In contrast, Connected Feedback Vertex Set does not admit a polynomial kernel for the solution size k even in 2-closed graphs [5]. Drawing a borderline between those desired graph properties where connected and capacitated vertex deletion problems do admit a kernel on (weakly) closed graphs and those where they do not would improve our understanding when (weak) closure can be exploited algorithmically. It is also open whether the Ramsey number of weakly closed graphs can be bounded by $(a+b+\gamma)^{\mathcal{O}(1)}$, such a bound would immediately improve some of our kernels. The most important open

[3] For ease of presentation we assume that $|V(G)|$ is divisible by b.

problem is arguably whether Dominating Set parameterized by the solution size k admits a polynomial kernel on weakly closed graphs. We made partial progress by showing that Dominating Set admits a kernel on weakly closed bipartite graphs and on weakly closed split graphs. Answering this question positively would need further insights into the structure of weakly closed graphs, however.

References

1 Piotr Berman, Marek Karpinski, and Alex D. Scott. Approximation hardness of short symmetric instances of MAX-3SAT. *Electronic Colloquium on Computational Complexity (ECCC)*, 049, 2003.

2 Hans L. Bodlaender, Rodney G. Downey, Michael R. Fellows, and Danny Hermelin. On problems without polynomial kernels. *Journal of Computer and System Sciences*, 75(8):423–434, 2009.

3 Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015.

4 Marek Cygan, Fabrizio Grandoni, and Danny Hermelin. Tight kernel bounds for problems on graphs with small degeneracy. *ACM Transactions on Algorithms*, 13(3):43:1–43:22, 2017.

5 Marek Cygan, Marcin Pilipczuk, Michal Pilipczuk, and Jakub Onufry Wojtaszczyk. Kernelization hardness of connectivity problems in d-degenerate graphs. *Discrete Applied Mathematics*, 160(15):2131–2141, 2012.

6 Holger Dell and Dániel Marx. Kernelization of packing problems. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '12)*, pages 68–81. SIAM, 2012.

7 Michael Dom, Daniel Lokshtanov, and Saket Saurabh. Kernelization lower bounds through colors and IDs. *ACM Transactions on Algorithms*, 11(2):13:1–13:20, 2014.

8 Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013.

9 David Eppstein, Maarten Löffler, and Darren Strash. Listing all maximal cliques in large sparse real-world graphs. *ACM Journal of Experimental Algorithmics*, 18, 2013.

10 Paul Erdös and András Hajnal. Ramsey-type theorems. *Discrete Applied Mathematics*, 25(1-2):37–52, 1989.

11 Rok Erman, Łukasz Kowalik, Matjaž Krnc, and Tomasz Waleń. Improved induced matchings in sparse graphs. *Discrete Applied Mathematics*, 158(18):1994–2003, 2010.

12 Elaine M. Eschen, Chính T. Hoàng, Jeremy P. Spinrad, and R. Sritharan. On graphs without a C_4 or a diamond. *Discrete Applied Mathematics*, 159(7):581–587, 2011.

13 Jacob Fox, Tim Roughgarden, C. Seshadhri, Fan Wei, and Nicole Wein. Finding cliques in social networks: A new distribution-free model. *SIAM Journal on Computing*, 49(2):448–464, 2020.

14 Peter Frankl and Richard M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1(4):357–368, 1981. doi:10.1007/BF02579457.

15 Danny Hermelin and Xi 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 '12)*, pages 104–113. SIAM, 2012.

16 Edin Husic and Tim Roughgarden. FPT algorithms for finding dense subgraphs in c-closed graphs. *CoRR*, abs/2007.09768, 2020. arXiv:2007.09768.

17 Iyad A. Kanj, Michael J. Pelsmajer, Marcus Schaefer, and Ge Xia. On the induced matching problem. *Journal of Computer and System Sciences*, 77(6):1058–1070, 2011.

18 Subhash Khot and Venkatesh Raman. Parameterized complexity of finding subgraphs with hereditary properties. *Theoretical Computer Science*, 289(2):997–1008, 2002.

19 Tomohiro Koana, Christian Komusiewicz, and Frank Sommer. Computing dense and sparse subgraphs of weakly closed graphs. In *Proceedings of the 31st International Symposium on Algorithms and Computation, (ISAAC '20)*, volume 181 of *LIPIcs*, pages 20:1–20:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020.

20 Tomohiro Koana, Christian Komusiewicz, and Frank Sommer. Exploiting c-closure in kernelization algorithms for graph problems. In *Proceedings of the 28th Annual European Symposium on Algorithms (ESA '20)*, volume 173 of *LIPIcs*, pages 65:1–65:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020.

21 Stefan Kratsch. Co-nondeterminism in compositions: A kernelization lower bound for a Ramsey-type problem. *ACM Transactions on Algorithms*, 10(4):19:1–19:16, 2014.

22 Stefan Kratsch. Recent developments in kernelization: A survey. *Bulletin of the EATCS*, 113, 2014.

23 Stefan Kratsch, Marcin Pilipczuk, Ashutosh Rai, and Venkatesh Raman. Kernel lower bounds using co-nondeterminism: Finding induced hereditary subgraphs. *ACM Transactions on Computation Theory*, 7(1):4:1–4:18, 2014.

24 Geevarghese Philip, Venkatesh Raman, and Somnath Sikdar. Polynomial kernels for dominating set in graphs of bounded degeneracy and beyond. *ACM Transactions on Algorithms*, 9(1):11:1–11:23, 2012.