Kernelization Hardness of Connectivity Problems in d-Degenerate Graphs*

Marek Cygan, Marcin Pilipczuk,
Michał Pilipczuk, and Jakub Onufry Wojtaszczyk

Faculty of Mathematics, Computer Science and Mechanics,
University of Warsaw, Poland
{cygan@,malcin@,michal.pilipczuk@students.,onufry@}mimuw.edu.pl

Abstract. A graph is d-degenerate if its every subgraph contains a vertex of degree at most d. For instance, planar graphs are 5-degenerate. Inspired by recent work by Philip, Raman and Sikdar, who have shown the existence of a polynomial kernel for DOMINATING SET in d-degenerate graphs, we investigate kernelization hardness of problems that include connectivity requirement in this class of graphs.

Our main contribution is the proof that CONNECTED DOMINATING SET does not admit a polynomial kernel in d-degenerate graphs for $d \geq 2$ unless the polynomial hierarchy collapses up to the third level. We prove this using a problem originated from bioinformatics — COLOURFUL GRAPH MOTIF — analyzed and proved to be NP-hard by Fellows et al. This problem nicely encapsulates the hardness of the connectivity requirement in kernelization. Our technique yields also an alternative proof that, under the same complexity assumption, STEINER TREE does not admit a polynomial kernel. The original proof, via reduction from SET COVER, is due to Dom, Lokshtanov and Saurabh.

We extend our analysis by showing that, unless $PH = \Sigma_p^3$, there do not exist polynomial kernels for STEINER TREE, CONNECTED FEEDBACK VERTEX SET and CONNECTED ODD CYCLE TRANSVERSAL in d-degenerate graphs for $d \geq 2$. On the other hand, we show a polynomial kernel for CONNECTED VERTEX COVER in graphs that do not contain the biclique $K_{i,j}$ as a subgraph.

1 Introduction

In the parameterized complexity setting, an instance comes with an integer parameter k — formally, a parameterized problem Q is a subset of $\Sigma^* \times \mathbb{N}$ for some finite alphabet Σ. We say that the problem is *fixed parameter tractable* (*FPT*) if there exists an algorithm solving any instance (x, k) in time $f(k)\text{poly}(|x|)$ for some (usually exponential) computable function f. It is known that a problem is FPT iff it is kernelizable: a kernelization algorithm for a problem Q takes an instance (x, k) and in time polynomial in $|x| + k$ produces an equivalent instance (x', k') (i.e., $(x, k) \in Q$ iff $(x', k') \in Q$) such that $|x'| + k' \leq g(k)$ for some computable function g. The function g is the *size of the kernel* and if it is polynomial, we say that Q admits a polynomial kernel.

* The first two authors were partially supported by Polish Ministry of Science grant no. N206 355636.

D.M. Thilikos (Ed.): WG 2010, LNCS 6410, pp. 147–158, 2010.

Kernelization techniques can be viewed as polynomial time preprocessing routines for tackling NP-hard problems. Parameterized complexity provides a formal framework for the analysis of such algorithms. In particular small (i.e. polynomial) kernels play an important role, and there are numerous positive results showing small kernels for various problems, including VERTEX COVER [5] and FEEDBACK VERTEX SET [19]. Recently, Bodlaender et al. [2] and Fortnow and Santhanam [12] came up with a technique that allows to prove negative results in this field: their tools provide a way to show that a parameterized problem does not admit a polynomial kernel unless the polynomial hierarchy collapses up to the third level. Up to this day, the list of non-polynomially-kernelizable (unless $PH = \Sigma_p^3$) FPT problems includes LONGEST PATH, LONGEST CYCLE [2], DIRECTED MAX LEAF OUT-BRANCHING [10], DISJOINT PATHS, DISJOINT CYCLES [4], RED-BLUE DOMINATING SET (aka SET COVER), STEINER TREE, CONNECTED VERTEX COVER [6] and CONNECTED FEEDBACK VERTEX SET [15].

On the other hand, many problems which are hard in general graphs — i.e. without a polynomial kernel or even not FPT — have small kernels in restricted graph classes, such as planar graphs, bounded genus graphs, apex-minor-free graphs or H-minor-free graphs. Recent results include linear kernels for DOMINATING SET and CONNECTED DOMINATING SET in apex-minor-free graphs and linear kernels for FEEDBACK VERTEX SET and CONNECTED VERTEX COVER in H-minor-free graphs [11].

The aforementioned results use the topological structure of the considered graph classes. However, sometimes an even weaker assumption on the graph class leads to significantly better algorithms and kernels than in the general case. One may, for instance, consider the class of d-degenerate graphs. A graph is called d-degenerate if its every induced subgraph contains a vertex of degree at most d. For instance, the class of 1-degenerate graphs is the class of forests, and all planar graphs are 5-degenerate. Moreover, every H-minor-free graph is d-degenerate, where the constant d depends on the minor [14,17,18]. Alon and Gutner [1] followed by Golovach and Villanger [13] proved that DOMINATING SET and CONNECTED DOMINATING SET, which are $W[2]$-hard in general graphs [7], become FPT when the input graph is d-degenerate. Very recently, Philip et al. [16] proved that DOMINATING SET is FPT and admits a polynomial kernel in a larger class of graphs: graphs excluding the biclique $K_{i,j}$ as a subgraph (note that a d-degenerate graph cannot contain $K_{d+1,d+1}$ as a subgraph).

A natural question arises: does the bounded degeneracy assumption help in the kernelization of other problems? In particular, the question of finding a polynomial kernel for CONNECTED DOMINATING SET in d-degenerated graphs was posted on the 1st Workshop on Kernels (WORKER'09, Bergen, Norway). In this paper we provide mostly negative answers to questions of existence of polynomial kernels for connectivity problems in graphs of bounded degeneracy. Note that this is in sharp contrast with the existence of the linear kernel for CONNECTED DOMINATING SET in apex-minor-free graphs [11].

The main contribution of this paper is the idea to use the COLOURFUL GRAPH MOTIF problem, which, intuitively, encapsulates the hardness of the connectivity requirement.

COLOURFUL GRAPH MOTIF	**Parameter:** k.
Input: A graph $G = (V, E)$, an integer k and a function $f : V \to \{1, 2, \dots, k\}$	
Question: Does there exist a connected set $S \subset V$ of cardinality k, such that $f|_S$ is bijective?	

We think of the function f to be a colouring of V — each number from $\{1, 2, \dots, k\}$ corresponds to a single colour — and we ask whether it is possible to choose a connected set containing exactly one vertex of each colour.

Fellows et al. [9] have shown that, surprisingly, this problem is NP-hard even in the class of trees of maximum degree 3. We use this fact to prove that COLOURFUL GRAPH MOTIF does not admit a polynomial kernel in 1-degenerate graphs (forests) unless $PH = \Sigma_p^3$.[1] This problem is simple enough to admit a reduction to CONNECTED DOMINATING SET in 2-degenerate graphs. As a by-product of this analysis, we obtain an alternative proof that STEINER TREE does not admit a polynomial kernel in arbitrary graphs. The original proof, via reduction from RED BLUE DOMINATING SET (aka SET COVER) is due to Dom et al. [6]. We analyze COLOURFUL GRAPH MOTIF in Section 4 and apply it to CONNECTED DOMINATING SET to show that CONNECTED DOMINATING SET does not admit a polynomial kernel in 2-degenerate graphs. In Section 4 we also show the reduction from COLOURFUL GRAPH MOTIF to STEINER TREE.

On the positive side (in Section 5) we provide a $O(k^2 + (i + j)k^{\min(i,j)})$-vertex kernel for CONNECTED VERTEX COVER in $K_{i,j}$-free graphs. In the analysis we use arguments similar to those developed by Philip et al. [16].

Preliminaries and notation are given in Section 2. As a warmup, in Section 3 by easy reductions and using already known results we show that STEINER TREE, CONNECTED FEEDBACK VERTEX SET and CONNECTED ODD CYCLE TRANSVERSAL do not admit polynomial kernels in 2-degenerate graphs. All discussed problems are parameterized by the solution size, except for STEINER TREE, which is parameterized both by the solution size and the size of the terminal set. Precise definitions of considered problems can be found in appropriate sections.

2 Preliminaries and Notation

Before we start, let us introduce some notation. All problems are considered on an undirected graph $G = (V, E)$. The set $N(v) = \{u : uv \in E\}$ is the neighbourhood of v and $N[v] = N(v) \cup \{v\}$ is the closed neighbourhood of v. We extend this notation to all subsets $A \subset V$: $N[A] = \bigcup_{v \in A} N[v]$ and $N(A) = N[A] \setminus A$. We say that a vertex v is dominated by a vertex set A if $v \in N[A]$; a vertex set A is dominating if $N[A] = V$. Whenever we speak of a parameterized problem Q, by d-deg-Q we denote the problem Q, where the class of input graphs is restricted to d-degenerate graphs.

In this section we recall all the required definitions about kernels, and ways to prove the non-existence of a polynomial kernel. In general, we follow the notation from [6]. Given a parameterized problem $Q \subset \Sigma^* \times \mathbb{N}$, its unparameterized version is a language

[1] In the full version of this paper we show NP-hardness and nonexistence of a polynomial kernel for COLOURFUL GRAPH MOTIF in comb graphs. A graph is called a comb graph if it is a tree, all vertices are of degree at most 3 and all the vertices of degree 3 lie on a single simple path.

$\tilde{Q} = \{x\#1^k : (x, k) \in Q\}$, i.e., we append the parameter written in unary. Let us now cite the main result of Bodlaender et al. [2] and Fortnow and Santhanam [12].

Definition 1 (Composition [2,6]). *A* composition algorithm *for a parameterized problem $Q \subset \Sigma^* \times \mathbb{N}$ is an algorithm that receives as input a sequence (x_1, k), (x_2, k), …, (x_t, k) with $(x_i, k) \in \Sigma^* \times \mathbb{N}$ for each $1 \le i \le t$, uses polynomial time in $\sum_{i=1}^{t} |x_i| + k$, and outputs $(y, k') \in \Sigma^* \times \mathbb{N}$ with $(y, k') \in Q$ iff $\exists_{1 \le i \le t} (x_i, k) \in Q$ and k' is polynomial in k. A parameterized problem is called* compositional *if there is a composition algorithm for it.*

Theorem 1 ([2,12]). *Let Q be a compositional parameterized problem whose unparameterized version $\tilde{Q}$ is NP-complete. Then, unless $PH = \Sigma_p^3$, there is no polynomial kernel for Q.*

To prove the non-existence of a polynomial kernel for some parameterized problem, it is not necessary to go through Theorem 1. As in the case of NP-complete problems, we can use reductions instead.

Definition 2 ([4,6]). *Let P and Q be parameterized problems. We say that P is polynomial parameter reducible to Q, written $P \le_{Ptp} Q$, if there exists a polynomial time computable function $f : \Sigma^* \times \mathbb{N} \to \Sigma^* \times \mathbb{N}$ and a polynomial p, such that for all $(x, k) \in \Sigma^* \times \mathbb{N}$ the following holds: $(x, k) \in P$ iff $(x', k') = f(x, k) \in Q$ and $k' \le p(k)$. The function f is called a* polynomial parameter transformation.

Theorem 2 ([4,6]). *Let P and Q be parameterized problems and $\tilde{P}$ and $\tilde{Q}$ be the unparameterized versions of P and Q respectively. Suppose that $\tilde{P}$ is NP-hard and $\tilde{Q}$ is in NP. Assume there is a polynomial parameter transformation from P to Q. Then if Q admits a polynomial kernel, so does P.*

3 Easy Cases: Steiner Tree, Connected Feedback Vertex Set and Connected Odd Cycle Transversal

We shall begin by showing that, unless $PH = \Sigma_p^3$, no polynomial kernel exists in the connected case even for 2-degenerate graphs for three problems: Steiner Tree, Connected Feedback Vertex Set and Connected Odd Cycle Transversal. We reduce them through Theorem 2 to the problems shown by other authors not to admit a polynomial kernel. We use the results of Dom et al. [6], where the authors show Steiner Tree and Connected Vertex Cover do not admit a polynomial kernel in the class of all graphs. Presented constructions are adjustments of reductions made for Connected Feedback Vertex Set [15].

Steiner Tree (ST) **Parameter:** $t := |T|$ and k.
Input: A graph $G = (V, E)$, a set of terminals $T \subset V$ and an integer k.
Question: Does there exist $S \subset V$, such that $G[S \cup T]$ is connected and $|S| \le k$?

Connected Feedback Vertex Set (CFVS) **Parameter:** k.
Input: A graph $G = (V, E)$ and an integer k.
Question: Does there exist a set $S \subset V$ of cardinality at most k, such that $G[S]$ is connected and $G[V \setminus S]$ contains no cycles?

CONNECTED ODD CYCLE TRANSVERSAL	**Parameter:** k.
Input: A graph $G = (V, E)$ and an integer k.	
Question: Does there exist a set $S \subset V$ of cardinality at most k, such that $G[S]$ is connected and $G[V \setminus S]$ is bipartite (that is, contains no cycles of odd length)?	

CONNECTED VERTEX COVER	**Parameter:** k.
Input: A graph $G = (V, E)$ and an integer k.	
Question: Does there exist a set $S \subset V$ of cardinality at most k, such that $G[S]$ is connected and every edge $e \in E$ has at least one endpoint in S?	

Now let us note the following simple observation.

Lemma 1. *Assume that in a graph G every edge has an endpoint of degree at most 2. Then G is 2-degenerate.*

We now show reductions to each of the three aforementioned problems.

Proposition 1. CONNECTED VERTEX COVER $\leq_{Ptp}$ 2*–deg–*CONNECTED FEEDBACK VERTEX SET.

Proof. Consider any instance (G, k) of CONNECTED VERTEX COVER. We create a graph $G' = (V', E')$. We take $V' = V \cup E_1 \cup E_2$ — the vertices of G' are the vertices of G plus two new vertices e_1, e_2 for each edge e of G. For each edge $e = uv \in E$ we add four edges to E': $\{u, e_1\}, \{u, e_2\}, \{v, e_1\}$ and $\{v, e_2\}$. This means we transform each edge of G into a cycle of length 4, where the original vertices are on opposite points of the cycle. Lemma 1 implies that G' is 2-degenerate.

We now prove that the answer to CONNECTED FEEDBACK VERTEX SET for $(G', 2k-1)$ is the same as the answer to CONNECTED VERTEX COVER for (G, k). First assume we have a positive answer for (G, k). This means that there exists a connected vertex cover S of G of size at most k. As S is connected, we can create a spanning tree in $G[S]$, this consists of at most $k - 1$ edges $E_S \subset E$. Let $E'_S = \{e_1 : e \in E_S\}$ — that is, for each edge $e \in E_S$ we take into E'_S one of the two vertices in V' corresponding to e. We claim $G'[V' \setminus (S \cup E_S)]$ contains no cycles. Assume C is a cycle in $G'[V' \setminus (S \cup E_S)]$. C cannot consist only of elements of V (since V is independent in G'), thus C contains some element e_i. As $\deg e_i = 2$, C also has to contain both vertices from V which the corresponding edge $e \in E$ connects. This, however, means in particular that neither of these vertices was in S, which is a contradiction with the assumption that S was a vertex cover of G, as the edge e is not covered.

Now assume we have a connected feedback vertex set $S \subset V'$ in G' of cardinality at most $2k - 1$. Assume $|S| \geq 2$ (the case $|S| = 1$ is trivial). Notice that $|S \cap V| \leq k$ — if we have more than k vertices from V, they form at least $k + 1$ connected components, and each vertex from E' connects at most two of them — thus S would not be connected. We claim $S \cap V$ forms a connected vertex cover of G. Consider any edge $e = uv$ in E and the corresponding cycle (u, e_1, v, e_2) in G'. As S is a feedback vertex set in G', at least one of these four vertices must belong to S. As $|S| \geq 2$ and S is connected, at least one of u, v is in S — and thus e is covered in G by $S \cap V$.

Proposition 2. CONNECTED VERTEX COVER $\leq_{Ptp}$ 2*–deg–*CONNECTED ODD CYCLE TRANSVERSAL.

Proof. We proceed as above, except we transform each edge into a cycle of length five.

Corollary 1. *The problems 2-deg-*CONNECTED ODD CYCLE TRANSVERSAL *and 2-deg-*CONNECTED FEEDBACK VERTEX SET *do not admit a polynomial kernel unless* $PH = \Sigma_p^3$.

The last reduction to degenerate graphs from previously known results is for 2-deg-STEINER TREE. The alternative proof of the kernelization hardness of 2-deg-STEINER TREE, via reduction from COLOURFUL GRAPH MOTIF, can be found in Section 4.

Proposition 3. STEINER TREE $\leq_{Ptp}$ 2*-deg-*STEINER TREE *and* 2*-deg-*STEINER TREE *does not admit a polynomial kernel unless* $PH = \Sigma_p^3$.

Proof. Take a general instance (G, k, T) of STEINER TREE. Create a new graph G' by subdividing each edge — formally, let $V' = V \cup E$ and $ve \in E'$ if v is an endpoint of e in G. The graph G' is 2–degenerate by Lemma 1.

We claim that the answer for (G, k, T) is the same as the answer for $(G', 2k + |T| - 1, T)$. Assume we have a solution S of (G, k, T). Then $G[S \cup T]$ is connected. Take any spanning tree of $G[S \cup T]$, let F be the set of its edges, we have $|F| \leq k + |T| - 1$. Now $F \cup S$ is a solution in $(G', 2k + |T| - 1, T)$. In the other direction, if we have a solution S' in $(G', 2k + |T| - 1, T)$, we consider $S = S' \cap V$. Note that $S \cup T$ has cardinality at most $k + |T|$ — since $|S' \cup T| \leq 2k + 2|T| - 1$, $S \cup T$ is isolated in G', and adding a single vertex from E connects at most two components of the set. Thus S has a cardinality at most k, and $G[S \cup T]$ is connected (for otherwise $S' \cup T$ could not be connected in G').

4 From COLOURFUL GRAPH MOTIF to CONNECTED DOMINATING SET

4.1 COLOURFUL GRAPH MOTIF

The CONNECTED VERTEX COVER problem is, in a number of cases, too specific to allow easy reductions. The COLOURFUL GRAPH MOTIF problem appeared to be very handy in our case.

We show the problem has no polynomial kernel in the class of forests of maximum degree 3. Fellows et al. [9] have shown that COLOURFUL GRAPH MOTIF in this class of graphs is already NP-complete. Since trees are 1-degenerate, we use Theorem 1 and take the disjoint union of graphs and the union of functions as the composition algorithm. Note that any feasible solution is required to induce a connected subgraph and therefore it needs to be contained in one connected component of the input graph. This yields the following theorem:

Theorem 3. *The* COLOURFUL GRAPH MOTIF *problem in the class of* 1*-degenerate graphs (forests) of maximum degree* 3 *does not admit a polynomial kernel unless* $PH = \Sigma_p^3$.

4.2 Reductions

We propose COLOURFUL GRAPH MOTIF as a simpler tool to prove that various other problems do not admit a polynomial kernel unless $PH = \Sigma_p^3$. Firstly, to give some intuition on COLOURFUL GRAPH MOTIF, let us note that COLOURFUL GRAPH MOTIF is a special case of GROUP STEINER TREE.

> GROUP STEINER TREE **Parameter:** k.
> **Input:** A graph $G = (V, E)$, sets of vertices $T_1, \ldots, T_k \subset V$ and an integer p.
> **Question:** Does there exist $S \subset V$, such that $G[S]$ is connected, $|S| = p$ and $S \cap T_i \neq \emptyset$ for $i = 1, \ldots, k$?

Proposition 4. *d–deg–*COLOURFUL GRAPH MOTIF $\leq_{Ptp}$ *d–deg–*GROUP STEINER TREE.

Proof. Assume we have an instance (G, k, f) of d–deg–COLOURFUL GRAPH MOTIF. We create an instance of d–deg–GROUP STEINER TREE as follows: we keep the graph G, we put $p = k$ and take $T_i = f^{-1}(i)$. Now the the problem GROUP STEINER TREE asks whether there exists a connected set S of cardinality $p = k$ which has a non–empty intersection with each T_i. As $p = k$, this means that the intersection with each T_i is to contain exactly one element. This is exactly the question in COLOURFUL GRAPH MOTIF, thus the answer to COLOURFUL GRAPH MOTIF in (G, f, k) is the same as the answer to GROUP STEINER TREE in $(G, \{T_i\}, p, k)$.

Corollary 2. COLOURFUL GRAPH MOTIF *can be solved in $2^k n^{O(1)}$ time and polynomial space.*

Proof. We reduce COLOURFUL GRAPH MOTIF to GROUP STEINER TREE as in the proof of Proposition 4 and use $2^k n^{O(1)}$-time algorithm described in [15].

Our original motivation for analyzing COLOURFUL GRAPH MOTIF was the CONNECTED DOMINATING SET problem.

> CONNECTED DOMINATING SET **Parameter:** k.
> **Input:** A graph $G = (V, E)$ and an integer k
> **Question:** Does there exist a set $S \subset V$ of cardinality at most k, such that $G[S]$ is connected and every vertex $v \in V$ is adjacent or equal to some vertex $u \in S$?

Proposition 5. *d–deg–*COLOURFUL GRAPH MOTIF $\leq_{Ptp}$ $(d+1)$*–deg–*CONNECTED DOMINATING SET*, and* 2*–deg–*CONNECTED DOMINATING SET *admits no polynomial kernel unless $PH = \Sigma_p^3$.*

Proof. We begin with an instance (G, k, f) of d–deg–COLOURFUL GRAPH MOTIF. Due to Corollary 2, we may assume $k \geq 2$, otherwise we can solve the input instance in polynomial time. We create a graph $G' = (V', E')$ as follows:

- $V \subset V', E \subset E'$;
- for each colour $l \in \{1, 2, \ldots, k\}$ we add two vertices v_l and v'_l to V';

– for each colour $l \in \{1, 2, \ldots, k\}$ we add an edge $v_l v'_l$ to E';
– for each vertex $v \in V$ we add an edge $vv_{f(v)}$ to E'.

Firstly, we prove G' is $(d+1)$–degenerate. Consider any $S \subset V'$. Then either $S \subset V' \setminus V$ (but then every vertex in $G'[S]$ is of degree at most 1) or $S \cap V$ is non–empty. Then $G[S \cap V]$ contains a vertex v, which had degree at most d in G, so it has degree at most $d+1$ in G' (as we added one edge to each vertex of V).

Now we prove the answer to COLOURFUL GRAPH MOTIF for (G, k, f) is the same as the answer to CONNECTED DOMINATING SET for $(G', 2k)$. Assume $k > 1$. If we have a solution S of COLOURFUL GRAPH MOTIF in G, we create a solution of CONNECTED DOMINATING SET by putting $S' = S \cup \{v_1, v_2, \ldots, v_k\}$. The vertices v'_l are neighbours of v_ls, any vertex $v \in V$ is a neighbour of $v_{f(v)}$, which is in S', and S' is connected, for S was connected and each v_l is adjacent to the vertex of colour l in S. On the other hand, any solution S' to CONNECTED DOMINATING SET in G' has to contain all the vertices v_l (there are two ways to dominate v'_l — either we take v_l, or we take v'_l, but in the second case we have to take v_l anyway for connectedness). To ensure connectedness we have to take at least one neighbour u_l of each v_l ($u_l \neq v'_l$). As the sets of neighbours of v_ls are disjoint and $|S'| \leq 2k$, this means exactly one neighbour of each v_l is in S'. In $G'[S']$ the vertices v_l are of degree 1 (they are not adjacent to each other, and are not adjacent to u_j for $j \neq l$), thus $G'[S' \setminus \{v_1, v_2, \ldots, v_k\}]$ is connected as $G'[S']$ is connected. This means $S' \setminus \{v_1, v_2, \ldots, v_k\}$ is a solution to COLOURFUL GRAPH MOTIF in G.

As a final example of the technique we show how to prove that the STEINER TREE problem admits no polynomial kernel in 2-degenerate graphs. The problem was studied in [6], where STEINER TREE was shown to admit no polynomial kernel in general graphs, and a simple reduction to 2-degenerate graphs was shown in Section 3. We now show a self–contained proof to demonstrate again the applicability of COLOURFUL GRAPH MOTIF.

Proposition 6. *d–deg–*COLOURFUL GRAPH MOTIF $\leq_{Ptp}$ *$(d+1)$–deg–*STEINER TREE *and* 2*–deg–*STEINER TREE *admits no polynomial kernel unless $PH = \Sigma_p^3$.*

Proof. Assume we have an instance (G, k, f) of d-deg-COLOURFUL GRAPH MOTIF. We create an instance (G', T, k) of $(d+1)$-deg-STEINER TREE as follows: we keep the graph G as the set of non–terminals $V \setminus T$. Additionally for each colour $i \in \{1, 2, \ldots, k\}$ we add a vertex $t_i \in T$ and edges vt_i for all $v \in f^{-1}(i)$. We ask for a Steiner tree of cardinality k in $T \cup V$ connecting all vertices from T.

First note G' is $(d+1)$-degenerate. Similarly like in the previous proof, the terminals T form an independent set, while to each non–terminal from G we added exactly one edge. Let S be the solution to STEINER TREE in G'. Note that S has to contain exactly one vertex of each colour — if some colour was excluded, the corresponding terminal could not be connected, and the number of colours is at least $|S|$. Moreover, S has to be connected in $G'[V] = G$, as there is only one vertex of each colour, each terminal is a leaf in the solution, so removing a terminal does not change the connectedness of the solution. On the other hand, it can be easily seen that any solution of COLOURFUL GRAPH MOTIF in G gives a solution of STEINER TREE in G'.

5 On the Positive Side: Polynomial Kernel for CONNECTED VERTEX COVER

As a counterpoint to the results above we show that CONNECTED VERTEX COVER in 2–degenerate graphs *does* admit a polynomial kernel. To show the problem is non–trivial we have to begin by proving CONNECTED VERTEX COVER is NP–hard in this class (otherwise finding a polynomial kernel would not be much of an achievement). This is not surprising — the CONNECTED VERTEX COVER problem was studied extensively and shown, for instance, to be NP–hard in graphs with maximum degree 4 (although it is in P for graphs of maximum degree 3, see [8]).

Proposition 7. *The unparameterized version of* 2*–deg–*CONNECTED VERTEX COVER *is NP–hard.*

Proof. We show a reduction of CNF–SAT to 2–deg–CONNECTED VERTEX COVER. Consider an instance $C_1 \wedge C_2 \wedge \ldots \wedge C_m$ with variables $x_1, \ldots, x_n$ of CNF–SAT. Let M be a total number of all literals in all clauses in this formula. We create a graph G as follows:

- we create two vertices v and v' and an edge vv';
- for each variable x we create vertices x^t and x^f, an edge $x^t x^f$ and edges vx^t and vx^f;
- for each clause C_j we create vertices C_j and C'_j and an edge $C_j C'_j$;
- for each clause C_j if x is a literal in C_j we create vertices L_{xj} and L'_{xj} and edges $L_{xj}L'_{xj}, L_{xj}C_j$ and $L_{xj}x^t$. If $\neg x$ is a literal in C_j we create the same vertices and edges, with the exception of the last edge being $L_{xj}x^f$;

First let us check the graph above is indeed 2–degenerate. Assume we have such a set $S \subset V$ that $G[S]$ does not contain a vertex of degree 2. The vertices v', L'_{xj} and C'_j are of degree 1 in G, so they cannot be contained in S. The vertices L_{xj} are of degree 2 in $G[V \setminus \{L'_{xj}\}]$, so they cannot be contained in S. After removing all L_{xj}s and C'_js the C_js become isolated, and the degree of x^ts and x^fs drops to 2, so S cannot contain any of them either. We are left with a single vertex v, which is isolated in $G[\{v\}]$, so S is empty.

We now claim that a solution to CNF–SAT in $C_1 \wedge \ldots \wedge C_m$ exists iff a solution to CONNECTED VERTEX COVER exists for $(G, n + m + M + 1)$. Assume we have a solution ϕ to CNF–SAT in $C_1 \wedge \ldots \wedge C_m$. We choose a set $S \subset V$ as follows:

- the vertices v, C_j for all j and L_{xj} for all x, j for which they exist are in S;
- the vertex x^t is in S if $\phi(x)$ is true, otherwise x^f is in S.

It is easy to see that the set above does indeed cover all the edges of G. It is also connected — all the x^ts and x^fs are connected with v, for each j the vertices C_j and L_{xj} are all connected, and as at least one literal of the clause C_j is set to be true by ϕ, at least one of the L_{xj}s for each j is connected to a $x^? \in S$. The solution given is of cardinality exactly $n + m + M + 1$.

On the other hand consider any solution S of CONNECTED VERTEX COVER in G. It has to contain v — to cover the edge vv' one of v, v' has to be in S, and if $v' \in S$,

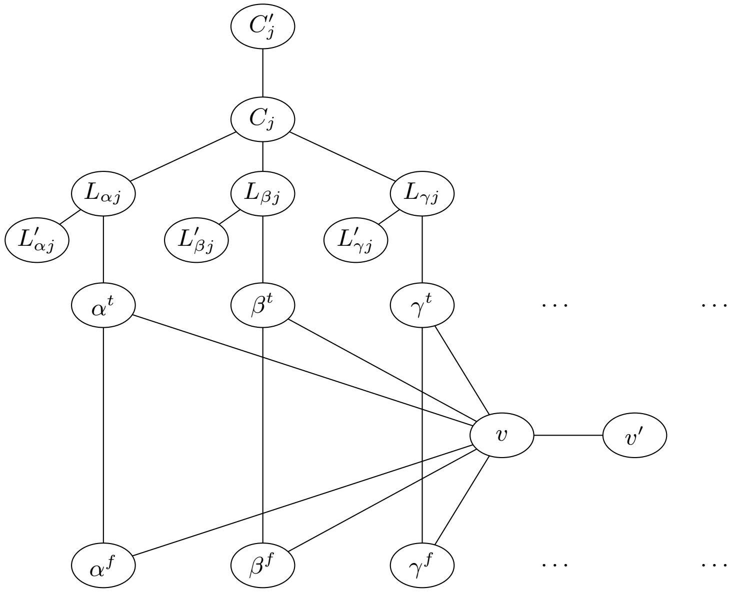

Fig. 1. Part of the constructed graph illustrating vertices added for clause $C_j = (\alpha \vee \beta \vee \gamma)$ and their interaction with vertices added for variables

then $v \in S$ to assure connectedness. For identical reasons $C_j \in S$ and $L_{xj} \in S$. We already have $m + M + 1$ vertices in S, so we can use at most n to cover the remaining edges. However the edges $x^f x^t$ form a matching of cardinality n in the remaining set, thus we have to take exactly one of $\{x^t, x^f\}$ to belong to S.

Consider a function $\phi : \{x_1, \ldots, x_n\} \to \{TRUE, FALSE\}$, setting $\phi(x) = TRUE$ iff $x^t \in S$. If for some clause C_j all of its literals were set to false by ϕ, then removing the vertices $x^?$ corresponding to these literals would split G into two connected components, one containing v, the other containing C_j, which would show S cannot be connected. As S is a solution to CONNECTED VERTEX COVER, this cannot happen, thus ϕ is a solution to our CNF–SAT instance.

Proposition 8. CONNECTED VERTEX COVER *in the class of $K_{i,j}$–free graphs admits a polynomial kernel for any i, j.*

Proof. Consider a graph $G = (V, E)$. We try to solve CONNECTED VERTEX COVER for (G, k). First, let $X = \{v \in V : \deg v > k\}$. Note that if there is a solution S for CONNECTED VERTEX COVER in (G, k), then $X \subset S$ — since if some $v \in X \setminus S$, then all the neighbours of v would have to be in S, but there are more than k of them. In particular if $|X| > k$, the answer is NO. Note we cannot remove X from G and analyse the remaining graph, for we could lose connectedness.

Consider the set F of edges which are not incident to X. If $|F| > k^2$, the answer is NO (for each vertex from V covers at most k such edges), let Z be the set of the endpoints of these edges, and let $Y = V \setminus (Z \cup X)$. Now for each vertex $x \in X$ add a vertex x' and an edge xx' (this is intended to assure that x is a part of any connected vertex cover not only of G, but also of the graphs we reduce G to, where the degree of x could drop). Denote the set of all vertices x' by X'.

If we do not know the answer yet, we have $|X| = |X'| \leq k$ and $|Z| \leq 2k^2$ (there are at most k^2 edges of which the vertices of Z are endpoints). Moreover $N(y) \subset X$ for any $y \in Y$.

Now if we have any two vertices $y_1, y_2 \in Y$ such that $N(y_1) \subset N(y_2)$, then the answer for G is the same as the answer for $G[V \setminus \{y_1\}]$. Indeed — if S is a solution in $G[V \setminus \{y_1\}]$, it is also a solution in G, since $N(y_1) \subset X \subset S$, as the edges xx' have to be covered, and thus all the edges incident to y_1 are covered by S, and of course $G[S]$ stays connected. On the other hand, if S is a solution in G, then either $y_1 \notin S$ (and then S is a solution in $G[V \setminus \{y_1\}]$) or $y_1 \in S$, and then $(S \cup \{y_2\}) \setminus \{y_1\}$ is a solution in $G[V \setminus \{y_1\}]$ (since y_2 connects everything that y_1 connected). Thus as long as a pair of vertices y_1, y_2 as above exists, we reduce G by removing y_1.

To simplify notation assume $i \leq j$. Now we show that after these reductions $|Y| \leq (i + j)k^i$. Consider any set $T \subset X$. There is at most one element $y \in Y$ such that $N(y) = T$ after the reductions. Moreover, if $|T| \geq i$, then there are at most $j - 1$ elements $y_1, \ldots, y_{j-1}$ of Y such that $N(y_l) \supset T$ — otherwise T and the y_ls would form a $K_{i,j}$ subgraph in G. For any element $y \in Y$ let $f(y) = N(y)$ if $|N(y)| < i$ and $f(y)$ be any $J \subset N(y)$, $|J| = i$ if $|N(y)| \geq i$. There are at most $\sum_{l=0}^{i-1} \binom{k}{l} \leq ik^i$ vertices y of the first type (as for such vertices each set appears at most once as the image of f) and at most $(j-1)\binom{k}{i} \leq jk^i$ vertices of the second type (as for such vertices each set appears at most $j - 1$ times as the image of f). Thus after the reductions we have $|V| = |X| + |X'| + |Z| + |Y| \leq k + k + 2k^2 + (i + j)k^i$, which is a polynomial of k.

The $K_{i,j}$–free graphs form a wider class than $(\min\{i, j\} - 1)$–degenerate graphs, thus CONNECTED VERTEX COVER admits a polynomial kernel in d–degenerate graphs for any d.

6 Conclusions and Open Problems

In this paper we investigated kernelization hardness in d-degenerate graphs for a number of problems that included the connectivity requirement. Generally, we proved that the bounded degeneracy assumption does not help much in existence of polynomial kernels. The question arises: does there exist a natural class larger than H-minor-free graphs or apex-minor-free graphs, for which CONNECTED DOMINATING SET, CONNECTED FEEDBACK VERTEX SET or CONNECTED ODD CYCLE TRANSVERSAL admit a small kernel?

Secondly, COLOURFUL GRAPH MOTIF appeared as a handy tool for proving kernelization hardness for 2–deg–CONNECTED DOMINATING SET and 2–deg–STEINER TREE. We believe this idea can inspire more negative results in the field of kernelization. In particular, such techniques may lead to a negative result for the question of existence of a polynomial kernel for PLANAR STEINER TREE, which today is a major open problem in kernelization and is not covered by meta-kernelization theorems of Bodlaender et al [3].

References

1. Alon, N., Gutner, S.: Linear time algorithms for finding a dominating set of fixed size in degenerated graphs. In: Lin, G. (ed.) COCOON 2007. LNCS, vol. 4598, pp. 394–405. Springer, Heidelberg (2007)
2. Bodlaender, H.L., Downey, R.G., Fellows, M.R., Hermelin, D.: On problems without polynomial kernels (extended abstract). In: Aceto, L., Damgård, I., Goldberg, L.A., Halldórsson, M.M., Ingólfsdóttir, A., Walukiewicz, I. (eds.) ICALP 2008, Part I. LNCS, vol. 5125, pp. 563–574. Springer, Heidelberg (2008)
3. Bodlaender, H.L., Fomin, F.V., Lokshtanov, D., Penninkx, E., Saurabh, S., Thilikos, D.M. (Meta) kernelization. In: Proc. of FOCS 2009, pp. 629–638 (2009)
4. Bodlaender, H.L., Thomasse, S., Yeo, A.: Analysis of data reduction: Transformations give evidence for non-existence of polynomial kernels, technical Report UU-CS-2008-030, Institute of Information and Computing Sciences, Utrecht University, Netherlands (2008)
5. Chen, J., Kanj, I.A., Jia, W.: Vertex cover: Further observations and further improvements. J. Algorithms 41(2), 280–301 (2001)
6. Dom, M., Lokshtanov, D., Saurabh, S.: Incompressibility through colors and IDs. In: Proc. of ICALP 2009, pp. 378–389 (2009)
7. Downey, R.G., Fellows, M.R.: Parameterized Complexity. Springer, Heidelberg (1999), `http://citeseer.ist.psu.edu/downey98parameterized.html`
8. Escoffier, B., Gourvès, L., Monnot, J.: Complexity and approximation results for the connected vertex cover problem. In: Brandstädt, A., Kratsch, D., Müller, H. (eds.) WG 2007. LNCS, vol. 4769, pp. 202–213. Springer, Heidelberg (2007)
9. Fellows, M.R., Fertin, G., Hermelin, D., Vialette, S.: Sharp tractability borderlines for finding connected motifs in vertex-colored graphs. In: Arge, L., Cachin, C., Jurdziński, T., Tarlecki, A. (eds.) ICALP 2007. LNCS, vol. 4596, pp. 340–351. Springer, Heidelberg (2007)
10. Fernau, H., Fomin, F.V., Lokshtanov, D., Raible, D., Saurabh, S., Villanger, Y.: Kernel(s) for problems with no kernel: On out-trees with many leaves. In: Proc. of STACS 2009, pp. 421–432 (2009)
11. Fomin, F., Lokshtanov, D., Saurabh, S., Thilikos, D.M.: Bidimensionality and kernels. In: Proc. of SODA 2010, pp. 503–510 (2010)
12. Fortnow, L., Santhanam, R.: Infeasibility of instance compression and succinct PCPs for NP. In: Proc. of STOC 2008, pp. 133–142 (2008)
13. Golovach, P.A., Villanger, Y.: Parameterized complexity for domination problems on degenerate graphs. In: Broersma, H., Erlebach, T., Friedetzky, T., Paulusma, D. (eds.) WG 2008. LNCS, vol. 5344, pp. 195–205. Springer, Heidelberg (2008)
14. Kostochka, A.V.: Lower bound of the hadwiger number of graphs by their average degree. Combinatorica 4(4), 307–316 (1984)
15. Misra, N., Philip, G., Raman, V., Saurabh, S., Sikdar, S.: FPT Algorithms for Connected Feedback Vertex Set. In: Rahman, M. S., Fujita, S. (eds.) WALCOM 2010. LNCS, vol. 5942, pp. 269–280. Springer, Heidelberg (2010)
16. Philip, G., Raman, V., Sikdar, S.: Solving dominating set in larger classes of graphs: Fpt algorithms and polynomial kernels. In: Fiat, A., Sanders, P. (eds.) ESA 2009. LNCS, vol. 5757, pp. 694–705. Springer, Heidelberg (2009)
17. Thomason, A.: An extremal function for contractions of graphs. Math. Proc. Cambridge Philos. Soc. 95(2), 261–265 (1984)
18. Thomason, A.: The extremal function for complete minors. J. Comb. Theory, Ser. B 81(2), 318–338 (2001)
19. Thomassé, S.: A quadratic kernel for feedback vertex set. In: Proc. of SODA 2009, pp. 115–119 (2009)