Edge-Editing to a Dense and a Sparse Graph Class

Michal Kotrbčík[1], Rastislav Královič[2], and Sebastian Ordyniak[3(✉)]

[1] Faculty of Informatics, Masaryk University, Brno, Czech Republic
kotrbcik@fi.muni.cz

[2] Department of Computer Science, Comenius University, Bratislava, Slovakia
kralovic@dcs.fmph.uniba.sk

[3] Institute of Information Systems, TU Wien, Vienna, Austria
sordyniak@gmail.com

Abstract. We consider a graph edge-editing problem, where the goal is to transform a given graph G into a disjoint union of two graphs from a pair of given graph classes, investigating what properties of the classes make the problem fixed-parameter tractable. We focus on the case when the first class is dense, i.e. every such graph G has minimum degree at least $|V(G)| - \delta$ for a constant δ, and assume that the cost of editing to this class is fixed-parameter tractable parameterized by the cost. Under the assumptions that the second class either has bounded maximum degree, or is edge-monotone, can be defined in MSO_2, and has bounded treewidth, we prove that the problem is fixed-parameter tractable parameterized by the cost. We also show that the problem is fixed-parameter tractable parameterized by degeneracy if the second class consists of independent sets and Subgraph Isomorphism is fixed-parameter tractable for the input graphs. On the other hand, we prove that parameterization by degeneracy is in general W[1]-hard even for editing to cliques and independent sets.

Keywords: Graph modification problems · Clique-editing · Degeneracy · Parameterized complexity · Treewidth

1 Introduction

Graph editing problems ask for modifying an input graph to a graph with a given property using at most k operations, where the allowed operations are usually edge addition, edge deletion, vertex deletion, or their combinations. Variants of graph editing problems received significant attention ever since the seminal Yannakakis's paper [23], and have found use in many application areas including machine learning [2], and social networks [13,18]. Graph editing is similarly

M. Kotrbčík and S. Ordyniak—Research funded by the Employment of Newly Graduated Doctors of Science for Scientific Excellence (CZ.1.07/2.3.00/30.0009) and the Austrian Science Fund (FWF, project P26696).

E. Kranakis et al. (Eds.): LATIN 2016, LNCS 9644, pp. 562–575, 2016.
DOI: 10.1007/978-3-662-49529-2_42

interesting from the algorithm design point of view and a considerable effort was invested into determining the (parameterized) complexity of editing into particular graph classes, for example, Eulerian graphs [7,11], regular graphs [16,17], or hereditary classes [6]; for a recent survey we refer to Bodlaender et al. [3].

Recently, variants of graph editing involving two classes were considered. The goal in these problems is to change a given graph into a disjoint union of two graphs, one from each class, using as few edge insertions and deletions as possible. A particular instance of editing to two classes, the CliqueEditing problem, asks for editing into a disjoint union of a clique and an independent set and arises as a clustering problem in noisy data sets [8]. CliqueEditing is fixed-parameter tractable (and solvable in subexponential time) parameterized by the number of edge modifications [8]. Only later it was shown that CliqueEditing is NP-hard, both in the general case and in the class of bipartite graphs [15]. On the other hand, CliqueEditing is solvable in polynomial-time on planar graphs [15], admits a PTAS on bipartite graphs [15], and there is a 3-approximation algorithm on general graphs [14]. A different case of editing to two classes considers editing to a disjoint union of a balanced biclique ($K_{n,n}$) and an independent set [14] — the problem admits a kernelization scheme that for any positive ϵ yields a kernel of size ϵk, where k is the cost of editing.

Our Results. We consider the problem of $(\mathcal{C}_1, \mathcal{C}_2)$-Editing, where the goal is to transform an input graph G into a disjoint union of two graphs, one from $\mathcal{C}_1$ and the other from $\mathcal{C}_2$, using as few edge insertions and deletions as possible. Editing to two classes is a quite general problem, for example, one can imagine a scenario as follows. A bus network, represented by a graph, is going to be split and sold to two interested parties. The parties are interested only in networks with certain topology, and the network may need to be changed to accommodate the supported topologies. The cost of edge modifications then represents the inherent cost of changing the infrastructure accordingly. The cost of editing to two classes corresponds to the minimum cost of infrastructure changes required to transfer the network to parties (or one of them) while still serving all the nodes in the network.

In this paper we focus on the theoretical aspects of the problem and pursue the question *when is editing to two classes tractable.* In line with the existing research, we focus on the case where $\mathcal{C}_1$ is a dense and $\mathcal{C}_2$ is a sparse class. Technically, we achieve this by requiring all graphs G from $\mathcal{C}_1$ to have minimum degree at least $|V(G)| - \delta$ for some constant δ and requiring $\mathcal{C}_2$ to have bounded either maximum degree, or treewidth. Our main contribution is a general approach which allows us to treat many cases in a unified fashion, using very little class-specific structural information. First, we introduce novel notions of weakly-hereditary and weakly anti-hereditary classes. These properties guarantee that the cost of editing to a graph in $\mathcal{C}$ does not change enormously by omitting only one vertex from the graph, respectively by adding an isolated vertex to a graph, see Sect. 4 for details. Second, we make use of a separation property between dense weakly hereditary and weakly-anti hereditary classes, which implies that the degrees of the induced subgraph that is being edited to the dense class are large in terms of the size of the subgraph. Finally, to obtain a solution, at some point it is necessary to perform computations for individual classes. This is achieved by assuming that the cost of single-class editing is either decidable or fixed-parameter tractable, or by assuming that the class can be defined in MSO_2 logic. In particular, we show that for bounded degree sparse classes $(\mathcal{C}_1, \mathcal{C}_2)$-Editing is fixed-parameter tractable parameterized by the cost of the editing as long as $\mathrm{cost}_{\mathcal{C}_1}(G)$ and $\mathrm{cost}_{\mathcal{C}_2}(G)$ are fixed-parameter tractable. For bounded treewidth sparse classes, the problem is fixed-parameter tractable parameterized by the cost whenever $\mathrm{cost}_{\mathcal{C}_1}(G)$ is fixed-parameter tractable and $\mathcal{C}_2$ is an edge-monotone and MSO_2-definable class. Our assumptions, in particular on the sparse classes, are weak, since our results hold for example for regular graphs, acyclic graphs, bounded degree trees, and k-colorable or bounded genus graphs with bounded treewidth.

For parameterization by degeneracy we prove that the problem is fixed-parameter tractable assuming only computability of $\mathrm{cost}_{\mathcal{C}_1}(G)$, the tradeoff is that the input graphs have to belong to a class for which Subgraph Isomorphism is fixed-parameter tractable, and that $\mathcal{C}_2$ contains only independent sets. Since graphs with bounded expansion have bounded degeneracy and admit Subgraph Isomorphism in FPT, we obtain a linear-time algorithm for $(\mathcal{C}, \mathcal{I})$-Editing when the input graphs have bounded expansion, where $\mathcal{I}$ is the class of all independent sets, which is a great improvement over the existing $O(n^{11})$ algorithm for CliqueEditing on planar graphs. On the other hand, we prove that the parameterization by degeneracy is in general W[1]-hard already for CliqueEditing. Finally, we obtain a kernel if $\mathrm{cost}_{\mathcal{C}_1}(G)$ is computable in polynomial time and the class $\mathcal{C}_2$ contains all graphs with maximum degree Δ. Without the second condition, or a similarly strong assumption about the structure of the graphs in the second class, it would not be possible to guarantee that the instance constructed as a kernel has the desired cost.

2 Preliminaries

We assume that the reader is familiar with the basic concepts and definitions of parameterized complexity; for details we refer to the standard texts on parameterized complexity [9,10,20].

2.1 Graphs

We use standard graph-theoretic notation and terminology and consider only finite, undirected, and simple graphs. For disjoint sets of vertices A and B of a graph G, by $E_G(A, B)$ we denote the set of edges of G with one endpoint in A and the other in B; if the graph G is clear from the context we omit the subscript G. For a graph G, by $|G|$ we denote the number of vertices of G and by $\deg_G(v)$ the degree of a vertex v in G. For a set of vertices X and a set of edges Y of a graph G, we denote by $G[X]$ the subgraph of G induced by X, by $G \backslash X$ the subgraph of G induced by $V(G) \backslash X$, and by $G \backslash Y$ the subgraph of G with vertices $V(G)$ and edges $E(G) \backslash Y$. The *degeneracy* of a graph G is defined as the minimum integer r such that every subgraph of G has minimum degree at most r and is denoted by $\mathrm{degen}(G)$. Finally, the isomorphism relation between graphs is denoted by $\cong$.

2.2 Treewidth

A *tree-decomposition* $\mathcal{T}$ of a graph G is a pair (T, χ), where T is a tree and χ is a function that assigns each tree node t a set $\chi(t) \subseteq V(G)$ of vertices such that the following conditions hold: (T1) For every vertex $v \in V(G)$, there is a tree node t such that $v \in \chi(t)$, (T2) For every edge $\{u, v\} \in E(G)$ there is a tree node t such that $u, v \in \chi(t)$, (T3) For every vertex $v \in V(G)$, the set of tree nodes t with $v \in \chi(t)$ forms a subtree of T. The sets $\chi(t)$ are called *bags* of the decomposition $\mathcal{T}$ and $\chi(t)$ is the bag associated with the tree node t. The *width* of a tree-decomposition (T, χ) is the size of a largest bag minus 1. A tree-decomposition of minimum width is called *optimal*. The *treewidth* of a graph G, denoted by $\mathrm{tw}(G)$, is the width of an optimal tree decomposition of G.

Proposition 1. *Let G be a graph and let G' be a graph obtained from G by applying at most k edge-modifications to G. Then $tw(G) \leq tw(G') + k$.*

Proposition 2 [4]. *Let G be a graph and ω a natural number. Then the problems of deciding whether G has a tree-decomposition of width at most ω and if yes, constructing such a tree-decomposition, are linear time fixed-parameter tractable parameterized by ω.*

2.3 Monadic Second Order Logic and Monotone Classes

We consider *Monadic Second Order* (MSO_2) logic on graphs in terms of their incidence structure whose universe contains vertices and edges; the incidence between vertices and edges is represented by a binary relation. We assume an infinite supply of *individual variables* $x, x_1, x_2, \ldots$ and of *set variables* $X, X_1, X_2, \ldots$ The *atomic formulas* are Ixy ("vertex x is incident with edge y"), $x = y$ (equality), $x \neq y$ (inequality), and Xx ("vertex or edge x is an element of set X"). *MSO formulas* are built up from atomic formulas using the usual Boolean connectives $(\neg, \wedge, \vee, \rightarrow, \leftrightarrow)$, quantification over individual variables ($\forall x$, $\exists x$), and quantification over set variables ($\forall X$, $\exists X$).

Let $\Phi(X)$ be an MSO_2 formula with a free set variable X. For a graph G and a set $S \subseteq E(G)$ we write $G \models \Phi(S)$ if the formula Φ holds true on G whenever X is instantiated with S.

The following theorem shows that if G has bounded treewidth, then in linear-time we can verify whether there is an S with $G \models \Phi(S)$ and $|S| \leq \ell$.

Theorem 3 [1]. *Let $\Phi(X)$ be an MSO_2 formula with a free set variable X and let ω be a constant. Then there is a linear-time algorithm that, given a graph G*

of treewidth at most ω, and an integer ℓ, decides whether there is a set $S \subseteq E(G)$ with $|S| \leq \ell$ such that $G \models \Phi(S)$.

While the original version of Theorem 3 [1] requires a tree-decomposition of width at most w to be provided with the input, for a graph of treewidth at most w such a tree decomposition can be found in linear time (Proposition 2) and thus the assumption is not necessary.

To employ Theorem 3, we utilize MSO_2-definable graph classes. A class $\mathcal{C}$ is MSO_2-*definable* if there is an MSO_2 formula $\Phi_{\mathcal{C}}(X)$ such that for any graph G, the formula $\Phi_{\mathcal{C}}(X)$ is satisfiable on G if and only if G belongs to $\mathcal{C}$. The vast majority of well studied graph classes are MSO_2-definable, examples include all graphs with bounded treewidth or bounded genus, bipartite, chordal, and perfect graphs, r-degenerate graphs for any r, trees, cliques, and many others. For our approach we additionally need the graph class to be *monotone* (with respect to edge deletion), that is, if G belongs to the class, then every subgraph of G on $V(G)$ also belongs to the class. Monotone classes are well studied (see for example Rivest and Vuillemin [22]) and examples of monotone MSO_2-definable graph classes include acyclic graphs, bipartite graphs, r-degenerate graphs for any r, or the class of all graphs with genus at most g.

3 Problem Definition

We consider the problem of edge-editing a graph to a disjoint union of two graphs, one from $\mathcal{C}_1$ and the other from $\mathcal{C}_2$. More formally, for a graph class $\mathcal{C}$ and a graph G, by $\mathrm{cost}_{\mathcal{C}}(G)$ we denote the minimum number of edge additions and edge removals required to modify G to a graph in $\mathcal{C}$. In other words, $\mathrm{cost}_{\mathcal{C}}(G)$ is the minimum size of a set of edges F such that the graph with vertex set $V(G)$ and edge set $E(G) \triangle F$ belongs to $\mathcal{C}$, where $\triangle$ denotes the symmetric difference. Note that $\mathrm{cost}_{\mathcal{C}}(G) = 0$ if and only if G belongs to $\mathcal{C}$. If $\mathcal{C}$ does not contain a graph on $|V(G)|$ vertices, then we let $\mathrm{cost}_{\mathcal{C}}(G) = \infty$. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two graph classes. Let G be a graph and D and S two induced subgraphs of G whose vertex sets partition $V(G)$. We define $\mathrm{cost}^G_{\mathcal{C}_1,\mathcal{C}_2}(D, S)$ to be the minimum number of edge additions and edge removals required to modify D to a graph in $\mathcal{C}_1$, modify S to a graph in $\mathcal{C}_2$, and to remove all edges between D and S. Formally, $\mathrm{cost}^G_{\mathcal{C}_1,\mathcal{C}_2}(D, S) = \mathrm{cost}_{\mathcal{C}_1}(D) + \mathrm{cost}_{\mathcal{C}_2}(S) + |E_G(D, S)|$. For a graph G, any pair (D, S) of vertex-disjoint induced subgraphs of G such that $V(D) \cup V(S) = V(G)$ is called a *solution* of $(\mathcal{C}_1, \mathcal{C}_2)$-Editing. *The cost of editing* of G to $(\mathcal{C}_1, \mathcal{C}_2)$ is defined by $\mathrm{cost}_{\mathcal{C}_1,\mathcal{C}_2}(G) = \min_{(D,S)} \mathrm{cost}^G_{\mathcal{C}_1,\mathcal{C}_2}(D, S)$, where the minimum is taken over all solutions. A solution (D, S) of $(\mathcal{C}_1, \mathcal{C}_2)$-Editing is *optimum* if $\mathrm{cost}^G_{\mathcal{C}_1,\mathcal{C}_2}(D, S) = \mathrm{cost}_{\mathcal{C}_1,\mathcal{C}_2}(G)$. These definitions lead to the following problem.

$(\mathcal{C}_1, \mathcal{C}_2)$-Editing **Input:** Graph G and a natural number k. **Question:** Is $\mathrm{cost}_{\mathcal{C}_1,\mathcal{C}_2}(G) \leq k$?

We denote the class of all edgeless graphs by $\mathcal{I}$ and the class of all cliques by $\mathcal{K}$. If $(\mathcal{C}_1, \mathcal{C}_2) = (\mathcal{K}, \mathcal{I})$, then we call $(\mathcal{C}_1, \mathcal{C}_2)$-Editing problem CliqueEditing and omit $(\mathcal{C}_1, \mathcal{C}_2)$ from the cost functions.

Consider the special case of $(\mathcal{C}, \mathcal{I})$-Editing. Clearly, for any solution (D, S) we have $\mathrm{cost}^G_{\mathcal{C},\mathcal{I}}(D, S) = \mathrm{cost}_{\mathcal{C}}(D) + |E(D, S)| + |E(S)|$. Since $|E(D)| + |E(D, S)| + |E(S)| = |E(G)|$, we have

$$\mathrm{cost}^G_{\mathcal{C},\mathcal{I}}(D, S) = \mathrm{cost}_{\mathcal{C}}(D) + |E(G)| - |E(D)|. \tag{1}$$

For a class of graphs $\mathcal{C}$, the $\mathcal{C}$-Cost problem asks whether, given a graph G and a natural number k, it holds that $\mathrm{cost}_{\mathcal{C}}(G) \le k$. For the remainder of the paper, we implicitly assume that the problems $(\mathcal{C}_1, \mathcal{C}_2)$-Editing and $\mathcal{C}$-Cost are parameterized by the cost of editing unless stated otherwise.

4 Graph Classes

In general, we are concerned with editing a given graph to a disjoint union of a dense and a sparse graph; in this section we make the technical requirements on the classes precise.

A class of graphs $\mathcal{D}$ is called a $\mathcal{D}(d, \delta)$-class, or a *dense class*, denoted by $\mathcal{D} \in \mathcal{D}(d, \delta)$, if each graph G from $\mathcal{D}$ satisfies the following two conditions:

($\mathcal{D}$1) For each vertex v of G, the degree of v is at least $|V(G)| - \delta$;
($\mathcal{D}$2) For each vertex v of G we have $\mathrm{cost}_{\mathcal{D}}(G - v) \le d$.

In the absence of significant structural information about the class, the condition ($\mathcal{D}$2) or a similar one seems to be necessary for our approach. In particular, without condition ($\mathcal{D}$2), it would be possible that $\mathcal{C}$ contains a graph H with n vertices, but does not contain any graph on $n - 1$ vertices (or all graphs from $\mathcal{C}$ on $n - 1$ vertices are very far from H in terms of editing cost). In such cases, an optimum solution might be forced to include a costly vertex just to raise the number of vertices in the dense part to n. Consequently, it would not be possible to prove a separation property analogous to the one we will prove in Lemma 7, which is crucial for our results.

On the other hand, the condition ($\mathcal{D}$2) is not particularly restrictive. Indeed, all hereditary graph classes (and in particular all minor-closed graph classes) satisfy ($\mathcal{D}$2) with $d = 0$, which leads us to call classes satisfying ($\mathcal{D}$2) *weakly hereditary*. The common property shared by the sparse classes considered in this paper is the following property, which is in a sense complementary to ($\mathcal{D}$2). A class of graphs $\mathcal{S}$ is called *weakly anti-hereditary* if there is an integer s such that for each graph G from $\mathcal{S}$ and for a vertex v not in G we have $\mathrm{cost}_{\mathcal{S}}(G + v) \le s$, where $G + v$ is the disjoint union of G and the single-vertex graph $\{v\}$. We call a class $\mathcal{S}$ s-weakly anti-hereditary to indicate the smallest integer s for which $\mathcal{S}$ satisfies the definition of weakly anti-hereditary class. Again, being weakly anti-hereditary is not particularly restrictive, since many well studied graph classes are weakly anti-hereditary. Examples include connected, bipartite, r-regular for any r, k-colorable for any k, chordal, perfect, and bounded-genus graphs.

Informally, to guarantee that a weakly anti-hereditary class is indeed sparse, we additionally require either bounded maximum degree, or bounded treewidth (graphs with treewidth ω have at most ωn edges). Specific basic examples of weakly anti-hereditary classes of bounded degree Δ include independent sets ($\Delta = 0$, $s = 0$), matchings covering all but at most one vertex ($\Delta = 1$, $s = 1$), paths ($\Delta = 2$, $s = 1$), cycles ($\Delta = 2$, $s = 3$), disjoint unions of paths ($\Delta = 2$, $s = 0$), disjoint unions of cycles ($\Delta = 2$, $s = 3$), forests with bounded maximum degree Δ ($s = 0$), trees with bounded maximum degree Δ ($s = 1$), and r-regular graphs for any even r ($\Delta = r$, $s = 2r$).[1] Further examples of sparse classes may be obtained by considering any weakly anti-hereditary class and restricting it to graphs with maximum degree Δ for suitably chosen Δ. Similarly, examples of weakly anti-hereditary classes of bounded treewidth include acyclic graphs and any weakly anti-hereditary class restricted to graphs with treewidth at most ω. Examples of dense classes may be obtained by taking complements of bounded degree sparse classes (see the next section for the precise definition). While we postpone verification of the fact that these classes have all the required properties to Sect. 7, the preceding classes can be used by the reader as specific illustrative examples of classes covered by our main theorems in Sect. 6.

5 Editing to a Single Class

To design parameterized algorithms for editing to two graph classes, it is necessary to assume that the cost of single-class editing can be determined efficiently. In this section we introduce the notation for the complexity of single-class editing and present several tools assuring efficient computation of the cost of single-class editing. These results are then used in Sect. 7 to establish the required complexity of editing to the basic examples of graph classes from the previous section. Except for the introduction of our notation in the following paragraph, the reader interested only in the framework for editing into two classes may skip the remainder of this section.

To discriminate among the complexity classes of the cost of single-class editing, we introduce additional notation as follows. By $\mathcal{D}_{\mathbf{P}}(d,\delta)$ we denote the set of all $\mathcal{D}(d,\delta)$ classes $\mathcal{C}$ such that $\mathcal{C}$-Cost can be computed in polynomial time. Similarly, by $\mathcal{D}_{\mathbf{FPT}}(d,\delta)$ we denote the set of all $\mathcal{D}(d,\delta)$ classes $\mathcal{C}$ such that $\mathcal{C}$-Cost is fixed-parameter tractable. Finally, by $\mathcal{D}_{\mathbf{C}}(d,\delta)$ we denote the set of all $\mathcal{D}(d,\delta)$ classes $\mathcal{C}$ such that $\mathcal{C}$-Cost is computable. Observe that $\mathcal{D}_{\mathbf{P}}(d,\delta) \subseteq \mathcal{D}_{\mathbf{FPT}}(d,\delta) \subseteq \mathcal{D}_{\mathbf{C}}(d,\delta)$ and that if $\mathcal{C} \in \mathcal{D}_{\mathbf{FPT}}(d,\delta)$, then it is possible to decide the membership in $\mathcal{C}$ in polynomial time.

The following lemma allows us to considerably weaken the requirements on the complexity of computation of the single-class editing cost function, and thus it may be interesting on its own.

[1] Note that in the case of r-regular graphs the smallest graph in the class has $r + 1$ vertices.

Lemma 4. *Let $\mathcal{C}$ be a class of graphs with maximum degree at most Δ such that $\mathcal{C}$-*Cost *is fixed-parameter tractable when the input is restricted to graphs with maximum degree at most Δ. Then $\mathcal{C}$-*Cost *is fixed-parameter tractable.*

In particular, we use the preceding lemma to prove Theorem 16 that establishes the complexity of single-class editing of several of the examples given in Sect. 4. While the proof of Theorem 16 treats each class separately and it is rather difficult to envision a general proof, we are able to show in a uniform way that $\mathcal{S}$-Cost is fixed-parameter tractable for MSO_2-definable sparse classes of bounded treewidth

Theorem 5. *Let $\mathcal{S}$ be a monotone, MSO_2-definable class of bounded treewidth. Then $\mathcal{S}$-*Cost *is fixed-parameter tractable.*

We say that a class $\mathcal{C}_1$ is *complement* of a class $\mathcal{C}_2$ if a graph G lies in $\mathcal{C}_1$ if and only if the complement of G lies in $\mathcal{C}_2$.

Proposition 6. *Let $\mathcal{S}$ be a weakly hereditary class of bounded maximum degree. Then the complement $\mathcal{D}$ of $\mathcal{S}$ is a $\mathcal{D}(d, \delta)$ class for some d and δ. Furthermore, if $\mathcal{S}$-*Cost *is computable in polynomial time, fixed-parameter tractable, respectively computable, then so is $\mathcal{D}$-*Cost*.*

6 Editing to Two Classes

In this section we give our fixed-parameter (in-)tractability results for variants of the $(\mathcal{C}_1, \mathcal{C}_2)$-Editing problem parameterized by treewidth, the cost of editing, respectively the degeneracy of the input graph.

We make use of the following separation lemma, which provides a lower bound on the minimum degree in the dense part of any optimal solution. While the lemma may seem intuitively clear, it is not necessarily the case. For instance, it is not immediately obvious why the bound should not be a function of degrees in S. However, even for weakly anti-hereditary classes of bounded degree Δ the lemma would be false with the bound $\deg_D(v) \geq (|D| + \Delta - \delta - d - s)/2$.

Lemma 7. *Assume that $\mathcal{D} \in \mathcal{D}(d, \delta)$, $\mathcal{S}$ is an s-weakly anti-hereditary class, and let (D, S) be an optimum solution of $(\mathcal{D}, \mathcal{S})$-*Editing *on a graph G. Then $\deg_D(v) \geq (|D| - \delta - d - s)/2$ for each vertex v of D.*

6.1 Parameterization by Treewidth

Theorem 8. *Let $\mathcal{D} \in \mathcal{D}_{\mathbf{C}}(d, \delta)$ and let $\mathcal{S}$ be a weakly anti-hereditary class. If $\mathcal{S}$ is monotone and MSO_2-definable, then $(\mathcal{D}, \mathcal{S})$-*Editing *is fixed-parameter tractable parameterized by treewidth.*

Let (G, k) be an instance of $(\mathcal{D}, \mathcal{S})$-Editing satisfying the assumptions of Theorem 8 and let (D, S) be an optimum solution of (G, k). To prove the theorem we first use Lemma 7 and the fact that a graph with treewidth ω has degeneracy at most ω to show that $|D| \leq 2\omega + d + \delta + s$. Since $\mathcal{S}$ is monotone, no edge is added to S. Therefore, to prove the theorem it is sufficient to show that the problem of finding the dense part D with at most $2\omega + d + \delta + s$ vertices and the set of edges to be deleted from S is fixed-parameter tractable parameterized by treewidth. The proof is concluded by constructing an MSO_2-formula deciding the last problem and using Theorem 3.

6.2 Parameterization by Editing Cost

In this section we show that $(\mathcal{D}, \mathcal{S})$-Editing is fixed-parameter tractable parameterized by the cost of editing for every $\mathcal{D} \in \mathcal{D}_{\mathbf{FPT}}(d, \delta)$ provided that $\mathcal{S}$ is weakly anti-hereditary and either has bounded maximum degree, or is monotone, MSO_2-definable, and has bounded treewidth. We also obtain polynomial kernels for several cases if $\mathcal{D} \in \mathcal{D}_{\mathbf{P}}(d, \delta)$ and $\mathcal{S}$ contains all graphs with maximum degree at most Δ. Our starting point is the following result for sparse classes with bounded treewidth.

Theorem 9. *Let $\mathcal{D} \in \mathcal{D}_{\mathbf{FPT}}(d, \delta)$ and let $\mathcal{S}$ be a weakly anti-hereditary class. If $\mathcal{S}$ is monotone, MSO_2-definable, and has bounded treewidth, then $(\mathcal{D}, \mathcal{S})$-*Editing *is fixed-parameter tractable.*

The proof of Theorem 9 proceeds as follows. Let $\mathcal{D}$ and $\mathcal{S}$ be classes satisfying the assumptions of the theorem, let (G, k) be an instance of $(\mathcal{D}, \mathcal{S})$-Editing and let (D, S) be an optimum solution of (G, k). We separately treat two cases, namely $|D| \geq 2k+\delta+d+s+1$ and $|D| \leq 2k+\delta+d+s$. In the first case we construct the largest set of vertices X such that the minimum degree of the subgraph of G induced by X is more than k. We further show that $V(D) \subseteq X$ and that $|X \backslash V(D)|$ is constant (with respect to $\mathcal{D}$ and $\mathcal{S}$). It follows that there are only polynomially many choices for the dense part $\mathcal{D}$ and because $\mathcal{D} \in \mathcal{D}_{\mathbf{FPT}}(d, \delta)$ and Theorem 5, we can compute the cost for each of them in fpt-time.

In the second case, we show that the treewidth of any YES-instance with $|D| \leq 2k + \delta + d + s$ is at most $3k + \delta + d + s + \omega_{\mathcal{S}}$, where $\omega_{\mathcal{S}}$ is the bound on the treewidth of graphs in $\mathcal{S}$. To see this, let (D', S') be a pair of graphs on $V(D)$, respectively $V(S)$, such that there is a set of $\text{cost}_{\mathcal{D},\mathcal{S}}(G)$ edge modifications that turn G into disjoint union of D' and S'. Observe that treewidth of D' is at most $2k + \delta + d + s$, the treewidth of S' is at most $\omega_{\mathcal{S}}$, and G differs from the disjoint union of D' and S' by at most k edges. Using Proposition 1 we obtain that the treewidth of G is at most $3k + \delta + d + s + \omega_{\mathcal{S}}$ and we can solve $(\mathcal{D}, \mathcal{S})$-Editing with the help of Theorem 8.

Theorem 9 implies fixed-parameter tractability of $(\mathcal{D}, \mathcal{S})$-Editing for sparse classes such as acyclic graphs and series-parallel graphs, as well as for k-colorable graphs and bounded genus graphs of bounded treewidth.

The following theorem, which is our main result for sparse classes of bounded maximum degree, is proved directly by first decomposing the graph according to degrees, using Lemma 7 to restrict the dense part, and then employing the fact that the cost of editing to the sparse class is fixed-parameter tractable.

Theorem 10. *Assume that $\mathcal{D} \in \mathcal{D}_{\mathbf{FPT}}(d,\delta)$ and $\mathcal{S}$ is a weakly anti-hereditary class. If $\mathcal{S}$ has bounded maximum degree and $\mathcal{S}$-*Cost *is fixed-parameter tractable, then $(\mathcal{D},\ \mathcal{S})$-*Editing *is fixed-parameter tractable.*

In a more restricted setting, we are able to obtain polynomial kernels, as shown by the following theorem.

Theorem 11. *Let $\mathcal{D} \in \mathcal{D}_{\mathbf{P}}(d,\delta)$ and let $\mathcal{S}_\Delta$ be the class containing all graphs with maximum degree at most Δ. If $\Delta = 0$, then $(\mathcal{D},\ \mathcal{S}_\Delta)$-*Editing *admits a kernel with $O(k)$ vertices and $O(k^2)$ edges. If $\Delta \leq 1$ or $\Delta \geq \max\{\delta, 2d+1, 6\}$, then $(\mathcal{D},\ \mathcal{S}_\Delta)$-*Editing *admits a kernel with $O(k^2)$ vertices and $O(k^3)$ edges.*

6.3 Parameterization by Degeneracy

In this section we consider $(\mathcal{C},\mathcal{I})$-Editing parameterized by the degeneracy of the input graph. Before we can state our results, we need to introduce the Subgraph Isomorphism problem, where given two graphs G and H, one asks whether G contains a subgraph isomorphic to H. In the following we will assume that Subgraph Isomorphism is parameterized by the size of H. We will denote by SI-FPT the set of all classes of graphs such that Subgraph Isomorphism is fixed-parameter tractable whenever the graph G is restricted to come from some class in SI-FPT. We will show that $(\mathcal{C},\mathcal{I})$-Editing is fixed-parameter tractable parameterized by degeneracy if the input graph comes from some class in SI-FPT and also if the input graph is bipartite. We also show that already CliqueEditing is W[1]-hard parameterized by degeneracy on general graphs.

Theorem 12. *For any $\mathcal{D}_{\mathbf{C}}(d,\delta)$-class $\mathcal{D}$, $(\mathcal{D},\mathcal{I})$-*Editing *is fixed-parameter tractable parameterized by the degeneracy for input graphs restricted to a class in* SI-FPT.

We want to note here that the above theorem is obtained via a linear time fpt-reduction from $(\mathcal{D},\mathcal{I})$-Editing parameterized by degeneracy to Subgraph Isomorphism. This, in particular, implies that if the Subgraph Isomorphism is linear time fixed-parameter tractable, then so is the $(\mathcal{D},\mathcal{I})$-Editing problem.

Note that restricting the input graphs to a class in SI-FPT is not particularly restrictive. Indeed, it is known that the Subgraph Isomorphism problem is fpt-equivalent to the model checking problem of existential first-order logic parameterized by the length of the formula [5, Proposition 1]. Therefore, SI-FPT contains also all classes for which the first-order model checking problem parameterized by the length of the formula is fixed-parameter tractable and, in particular, the very general class of nowhere-dense graphs [12]. A particular example of a nowhere-dense class is formed by graphs with bounded expansion, which contain for example all planar graphs, all classes with bounded treewidth, and all classes defined by a finite set of forbidden minors. Since we do not need the precise definition of bounded expansion, which is rather lengthy and technical, we only collect the following two results about classes with bounded expansion: (i) they belong to SI-FPT, and (ii) they have bounded degeneracy.

We refer to a book by Nešetřil and Ossona de Mendez [19] for more details on classes with bounded expansion and related classes in the context of nowhere-dense/somewhere-dense dichotomy. In the light of the preceding discussion, it follows that if input graphs are restricted to any class with bounded expansion, both conditions in Theorem 12 are satisfied, and since Subgraph Isomorphism is linear-time fixed-parameter tractable on graphs of bounded expansion [12], we obtain that $(\mathcal{D}, \mathcal{I})$-Editing is solvable in linear time.

Corollary 13. *For any class $\mathcal{C}$ of graphs of bounded expansion and any $\mathcal{D} \in \mathcal{D}_{\mathbf{C}}(d, \delta)$, $(\mathcal{D}, \mathcal{I})$-*Editing *can be solved in linear time for input graphs restricted to $\mathcal{C}$.*

While on bipartite graphs CliqueEditing is still NP-hard, it appears that requiring the input graphs to be bipartite leads to a somewhat simpler problem, as there is a polynomial-time approximation scheme [15]. We observe a similar behavior also with respect to parameterization — we prove that while CliqueEditing parameterized by degeneracy is in general W[1]-hard, it is fixed-parameter tractable when the input graphs are bipartite.

Theorem 14. CliqueEditing *is linear-time fixed-parameter tractable parameterized by the degeneracy for input graphs restricted to being bipartite.*

Our final result, proved by a reduction from Clique, which is W[1]-hard [21], is that editing to two classes is W[1]-hard parameterized by the degeneracy even for the case of cliques and independent sets.

Theorem 15. CliqueEditing *is* W[1]*-hard parameterized by the degeneracy.*

Proof. To show W[1]-hardness for CliqueEditing parameterized the degeneracy of the input graph, we use a parameterized reduction from the Clique problem, which is well-known to be W[1]-complete [21].

Clique	**Parameter:** k
Input: A graph G and a natural number k.	
Question: Does G have a clique of size at least k?	

Given G and k we construct a graph H with $\mathrm{degen}(H) \leq 1 + 3\binom{k}{2} + k$ such that G has a clique of size at least k if and only if $\mathrm{cost}(H) \leq b$, where $b = |E(H)| - (3\binom{k}{2} + k - 1) - 8\binom{k}{2}$.

Let G' be the graph obtained from G after subdividing every edge of G three times, i.e., replacing every edge of G by a path consisting of three novel internal vertices. Then the graph H consists of a copy of G' together with a set I of $3\binom{k}{2} + k - 1$ vertices, which are completely connected to all vertices in the copy of G'. Then $\mathrm{degen}(H) \leq 1 + 3\binom{k}{2} + k$ and it remains to show that G has a clique of size at least k if and only if $\mathrm{cost}(H) \leq b$, the proof of which can be found in the full version of the paper.

7 Applications

In this section we outline applications of our results by showing that our examples of sparse and dense classes have the properties required to employ Theorems 8 and 10. Let $\mathcal{G}$ be the set of the following graph classes: independent sets, matchings covering all but at most one vertex, paths, cycles, disjoint unions of paths, disjoint unions of cycles, forests with bounded maximum degree, trees with bounded maximum degree, and r-regular graphs for any even r.

The cost function $\text{cost}_{\mathcal{C}}(G)$ is computable in polynomial time for the class of independent sets, the class of matchings covering all but at most one vertex, and the class of graphs with bounded maximum degree [24]. It is also known that computing $\text{cost}_{\mathcal{C}}(G)$ is fixed-parameter tractable if $\mathcal{C}$ is the class of all r-regular graphs [16]. To show that the same holds also for the remaining classes in $\mathcal{G}$, we employ Lemma 4.

Theorem 16. *Let $\mathcal{C}$ be a class in $\mathcal{G}$. Then $\mathcal{C}$-*Cost *is fixed-parameter tractable.*

Proof. It is easy to see that all the classes given in $\mathcal{G}$ have bounded maximum degree and are weakly anti-hereditary. It hence remains to show that for all these classes computing the cost of editing is fixed-parameter tractable w.r.t. the editing cost. Because of Lemma 4 it is sufficient to show that for all these classes (having maximum degree at most Δ) computing the cost of editing is fixed-parameter tractable w.r.t. the editing cost if the input graph has maximum degree at most Δ.

Here we only show that this is indeed the case if $\mathcal{C}$ is the class of all connected graphs of maximum degree at most $d = \Delta$. The proofs for other cases are similar and can be found in the full version of the paper. Let c_1 be the number of components of G with minimum degree less than Δ and c_2 be the number of components of G with minimum degree at least Δ. Since the maximum degree of G is Δ, each component that is contributing to c_2 is a Δ-regular graph. To edit a graph with c_1+c_2 components into a connected graph c_1+c_2-1 edge additions are necessary and sufficient, and to maintain maximum degree Δ it is necessary and sufficient to delete one edge from each component that is contributing to c_2. Therefore, the cost of editing into $\mathcal{C}$ is $\text{cost}_{\mathcal{C}}(G) = c_1 + c_2 - 1 + c_2 = c_1 + 2c_2 - 1$. □

To the best of our knowledge, the exact (parameterized) complexity status of computing the cost for the classes in $\mathcal{G}$ is not known. In particular, none of the classes in $\mathcal{G}$ fall into the general setting of classes characterized by finitely many forbidden induced subgraphs considered previously [6]. However, at least for the class of all paths and the class of all cycles there is a straightforward reduction from the Hamiltonian Path or Hamiltonian Cycle problem, respectively, which shows NP-completeness of computing the cost of editing to these classes.

Recall that a class $\mathcal{C}_1$ is the complement of a class $\mathcal{C}_2$ if a graph G lies in $\mathcal{C}_1$ if and only if the complement of G lies in $\mathcal{C}_2$. Let $\mathcal{G}'$ be the set $\mathcal{G}$ without the classes of regular graphs, cycles, and disjoint unions of cycles. Since all the examples of weakly anti-hereditary classes in $\mathcal{G}'$ are also weakly hereditary, Proposition 6 yields that their complements are examples of dense classes. For the complementary classes of $2r$-regular graphs, cycles, or disjoint unions of cycles to be weakly hereditary, they would need to contain at least one graph on n vertices for each integer n. It is easy to see that if we add to the class of $2r$-regular graphs any graph on n vertices for each $n \leq 2r$, the class becomes weakly hereditary and thus its complement is a dense class. Similarly, if we add to the classes of cycles and disjoint union of cycles a graph with one vertex and a graph with two vertices, then the resulting classes are weakly hereditary and its complement is a $\mathcal{D}(d, \delta)$ class.

Finally, since $\mathcal{S}$-Cost is fixed-parameter tractable for each class $\mathcal{S}$ in $\mathcal{G}$ by Theorem 16, Proposition 6 yields that $\mathcal{C}_{\mathcal{S}}$-Cost is fixed-parameter tractable for the complement $\mathcal{C}_{\mathcal{S}}$ of each such class $\mathcal{S}$. Therefore, Theorem 8 holds with any class from $\mathcal{G}$ in the role of sparse class and the complement of any class from $\mathcal{G}$' (resp. complements of the extended classes of $2r$-regular graphs, cycles, or disjoint unions of cycles) in the role of dense class. Similarly, Theorem 10 also holds for the same set of complements in the role of dense class.

8 Conclusion

This paper introduces the $(\mathcal{C}_1, \mathcal{C}_2)$-Editing problem, where the goal is to edit an input graph into a disjoint union of two graphs, one from each class. We investigate for which classes the problem is fixed-parameterized tractable using only limited class-specific structural information. This is achieved by focusing on novel relevant properties, weakly hereditary and weakly anti-hereditary, and a separation property between such classes. While as far as we know weakly hereditary classes were not considered before, they may prove to be an useful concept also for other editing problems, extending the results from hereditary classes. Our results allow us to prove fixed-parameter tractability for a number of interesting sparse classes such as acyclic graphs, bounded degree trees or forests, series-parallel graphs, or k-colorable graphs of bounded treewidth, for instance, bipartite graphs of bounded treewidth. Since this is the first attempt to solve the problem in general, there are many open problems, including other conceivable approaches to choosing the technical requirements on the dense class. Another particular problem left is extending our results for parameterization by the degeneracy beyond independent sets. Our approach uses the equality in Eq. (1) in an essential way, and it seems to be very difficult, and necessary, to calculate the cost of editing to the class exactly when the class is different from independent sets.

References

1. Arnborg, S., Lagergren, J., Seese, D.: Easy problems for tree-decomposable graphs. J. Algorithms **12**(2), 308–340 (1991)
2. Bansal, N., Blum, A., Chawla, S.: Correlation clustering. Mach. Learn. **56**(1–3), 89–113 (2004)

3. Bodlaender, H., Heggernes, P., Lokshtanov, D.: Graph modification problems (Dagstuhl seminar 14071). Dagstuhl Rep. **4**(2), 38–59 (2014)
4. Bodlaender, H.L.: A linear-time algorithm for finding tree-decompositions of small treewidth. SIAM J. Comput. **25**(6), 1305–1317 (1996)
5. Bova, S., Ganian, R., Szeider, S.: Model checking existential logic on partially ordered sets. In: CSL-LICS. ACM (2014)
6. Cai, L.: Fixed-parameter tractability of graph modification problems for hereditary properties. Inform. Process. Lett. **58**(4), 171–176 (1996)
7. Cygan, M., Marx, D., Pilipczuk, M., Pilipczuk, M., Schlotter, I.: Parameterized complexity of eulerian deletion problems. Algorithmica **68**(1), 41–61 (2014)
8. Damaschke, P., Mogren, O.: Editing simple graphs. J. Graph Algorithms Appl. **18**(4), 557–576 (2014)
9. Downey, R.G., Fellows, M.R.: Parameterized Complexity. Monographs in Computer Science. Springer, New York (1999)
10. Flum, J., Grohe, M.: Parameterized Complexity Theory. Texts in Theoretical Computer Science. An EATCS Series, vol. 41. Springer, Heidelberg (2006)
11. Fomin, F., Golovach, P.: Long circuits and large euler subgraphs. SIAM J. Discrete Math. **28**(2), 878–892 (2014)
12. Grohe, M., Kreutzer, S., Siebertz, S.: Deciding first-order properties of nowhere dense graphs. In: STOC. ACM (2014)
13. Hartung, S., Nichterlein, A., Niedermeier, R., Suchý, O.: A refined complexity analysis of degree anonymization in graphs. In: Fomin, F.V., Freivalds, R., Kwiatkowska, M., Peleg, D. (eds.) ICALP 2013, Part II. LNCS, vol. 7966, pp. 594–606. Springer, Heidelberg (2013)
14. Hüffner, F., Komusiewicz, C., Nichterlein, A.: Editing graphs into few cliques: complexity, approximation, and kernelization schemes. In: Dehne, F., Sack, J.-R., Stege, U. (eds.) WADS 2015. LNCS, vol. 9214, pp. 410–421. Springer, Heidelberg (2015)
15. Kováč, I., Selečéniová, I., Steinová, M.: On the clique editing problem. In: Csuhaj-Varjú, E., Dietzfelbinger, M., Ésik, Z. (eds.) MFCS 2014, Part II. LNCS, vol. 8635, pp. 469–480. Springer, Heidelberg (2014)
16. Mathieson, L., Szeider, S.: Editing graphs to satisfy degree constraints: a parameterized approach. J. Comput. Syst. Sci. **78**(1), 179–191 (2012)
17. Moser, H., Thilikos, D.M.: Parameterized complexity of finding regular induced subgraphs. J. Discrete Algorithms **7**(2), 181–190 (2009)
18. Nastos, J., Gao, Y.: Familial groups in social networks. Soc. Netw. **35**(3), 439–450 (2013)
19. Nešetřil, J., de Mendez, P.O.: Sparsity: Graphs, Structures, and Algorithms. Algorithms and Combinatorics, vol. 28. Springer, Heidelberg (2012)
20. Niedermeier, R.: Invitation to Fixed-Parameter Algorithms. Oxford University Press, Oxford (2006)
21. Pietrzak, K.: On the parameterized complexity of the fixed alphabet shortest common supersequence and longest common subsequence problems. J. Comput. Syst. Sci. **67**(4), 757–771 (2003)
22. Rivest, R.L., Vuillemin, J.: On recognizing graph properties from adjacency matrices. Theor. Comput. Sci. **3**(3), 371–384 (1976)
23. Yannakakis, M.: Node-and edge-deletion NP-complete problems. In: STOC, pp. 253–264. ACM (1978)
24. Yannakakis, M.: Edge-deletion problems. SIAM J. Comput. **10**(2), 297–309 (1981)