The Complexity of Degree Anonymization by Graph Contractions*

Sepp Hartung and Nimrod Talmon

Institut für Softwaretechnik und Theoretische Informatik, TU Berlin, Germany

sepp.hartung@tu-berlin.de, nimrodtalmon77@gmail.com

Abstract

We study the computational complexity of k-anonymizing a given graph by as few graph contractions as possible. A graph is said to be k-anonymous if for every vertex in it, there are at least $k-1$ other vertices with exactly the same degree. The general degree anonymization problem is motivated by applications in privacy-preserving data publishing, and was studied to some extent for various graph operations (most notable operations being edge addition, vertex addition, and vertex deletion). We complement this line of research with studying several variants of graph contraction, which are operations of interest, for example, in the contexts of social networks and clustering algorithms. We show that the problem of degree anonymization by graph contractions is NP-hard even for some very restricted inputs, and identify some fixed-parameter tractable cases.

1 Introduction

Motivated by concerns of data privacy in social networks, Clarkson et al. [9] introduced the general degree anonymization problem, defined as follows: given an input graph G and an allowed operation O, transform G into a k-anonymous graph by performing as few O operations as possible; a graph is said to be k-anonymous if for every vertex degree d in it, there are at least k vertices with the same degree d. This problem has been studied both theoretically and practically, for several graph modification operations such as edge addition [9, 17, 20], edge addition and edge deletion [7], vertex addition [8, 4], and vertex deletion [3]. From the perspective of degree anonymization, this paper can be seen as complementing this line of research by considering graph contractions, as a natural graph modification operation. Specifically, studying the (parameterized) complexity of this degree anonymization problem.

*NT was supported by DFG Research Training Group MDS (GRK 1408).

This paper also complements research done on the following problem: given an input graph G and a family $\mathcal{F}$ of graphs, find a set of edges E' of minimum size, such that after contracting the edges in E', G would be in the family $\mathcal{F}$. Asano and Hirata [1] defined a set of conditions on $\mathcal{F}$, which is sufficient for NP-hardness of this problem. Others studied specific graph classes (as $\mathcal{F}$), such as planar graphs [15], bipartite graphs [18], paths [18], trees [16], and d-regular graphs Belmonte et al. [2]. This last work is of particular interest, as the concept of k-anonymity is a generalization of the notion of regularity (in particular, a graph is n-anonymous if and only if it is regular).

Studying graph contractions in the context of degree anonymization is interesting for several reasons. First, some variants of contractions can preserve original properties of the input graph (for example, connectivity). Second, vertex contraction (where also non-adjacent vertices can be contracted), is the inverse operation of vertex cleaving (as defined by Oxley [22, Chapter 3]), which was studied in the context of degree anonymization by Bredereck et al. [4]. We note also the relation of graph contractions to communities in social networks and to clustering (see, for example, Delling et al. [10]).

2 Preliminaries

We assume familiarity with standard notions regarding algorithms, computational complexity theory, and graph theory. For a non-negative integer z, we write $[z]$ to mean $\{1, \ldots, z\}$.

2.1 Parameterized Complexity

An instance (I, k) of a parameterized problem consists of the "classical" problem instance I and an integer k being the *parameter* [12, 13, 21]. A parameterized problem is called *fixed-parameter tractable* (in FPT) if there is an algorithm solving it in $f(k) \cdot |I|^{O(1)}$ time, for an arbitrary computable function f only depending on parameter k. In difference to that, algorithms running in $|I|^{f(k)}$ time prove membership in the class XP (clearly, FPT $\subseteq$ XP). One can show that a parameterized problem L is (presumably) not fixed-parameter tractable by devising a *parameterized reduction* from a W[1]-hard or a W[2]-hard problem to L. A parameterized reduction from a parameterized problem L to another parameterized problem L' is a function that, given an instance (I, k), computes in $f(k) \cdot |I|^{O(1)}$ time an instance (I', k') such that $k' \leq g(k)$ and $(I, k) \in L \Leftrightarrow (I', k') \in L'$. A parameterized problem which is NP-hard even for instances for which the parameter is a constant is said to be Para-NP-hard.

2.2 Graph Theory and Contractions

Given a graph $G = (V, E)$, which may have self-loops and parallel edges, we denote the degree of a vertex $v \in V$ by $\deg(v)$, and $B_d = \{v \in V : \deg(v) = d\}$ is the set of vertices of degree d. As usual, we define the degree of a vertex v

with x neighbors and y self-loops to be $x+2y$ (in particular, we count a self-loop twice). We define a *path-star* of degree d and length l to be a graph consisting of one center vertex, connected to d paths of length l (indeed, this is a spider graph with equal-length legs). A *caterpillar-tree* is a tree for which removing the leaves and their incident edges leaves a path graph.

Given an undirected graph $G = (V, E)$ and two adjacent vertices, u and v, contracting the vertices u and v (usually referred to as contracting the edge $e = \{u, v\}$), means removing u and v from V and replacing them by one new vertex (denoted by $u \oplus v$), which is adjacent to exactly those vertices that were adjacent to u, v, or both. From this definition, it follows that edge contractions are symmetric and associative. The resulting graph is denoted by G/e, and given a set of edges $E_1 \subseteq E$, we denote by G/E_1 the graph obtained from G after contracting the edges of E_1. A graph $G = (V, E)$ is said to be k-contractible to a graph $G' = (V', E')$ if there is a set of edges $E_1 \subseteq E$ of size at most k, such that $G/E_1 = G'$. It follows that $G = (V, E)$ is k-contractible to $G' = (V', E')$ if and only if there exists a *witness structure* $V = V_1 \cup V_2 \cup \ldots \cup V_{|V'|}$, each V_i is called a *witness set*, such that for each $1 \leq i \leq |V'|$ the subgraph of G induced by each V_i is connected and for each pair of witness sets, V_i and V_j $(1 \leq i \neq j \leq |V'|)$, we have that $\{V_i, V_j\} \in E' \iff \exists v_i \in V_i, v_j \in V_j : \{v_i, v_j\} \in E$ (indeed, the vertices in each part V_i are contracted to form a single vertex). We denote by $\deg(V_i)$ the resulting degree of the vertex corresponding to the contraction of the witness set and we call graph G' the witness graph.

We also define the closely related operation of vertex contraction, which is defined similarly to edge contraction, with the only difference that it is allowed to contract non-adjacent vertices as well. It is clear that a graph contraction operation can sometimes introduce self-loops and parallel edges. We define three variants of edge and vertex contraction, differing by how these self-loops and parallel edges are treated:

- *Simple Contraction:* Both self-loops and parallel edges are removed.
- *Hybrid Contraction:* Only self-loops are removed.
- *Non-Simple Contraction:* Nothing is removed.

For the Hybrid and Non-Simple variants, we allow the input graph to be non-simple. See Figure 1 for some examples.

2.3 Main Problem

Given an undirected input graph G, we are interested in k-anonymizing it by performing at most c edge contractions, where a graph is said to be k-anonymous if every vertex degree in it occurs at least k times (equivalently, if $\forall i \in [n] : |B_i| = 0 \vee |B_i| \geq k$).

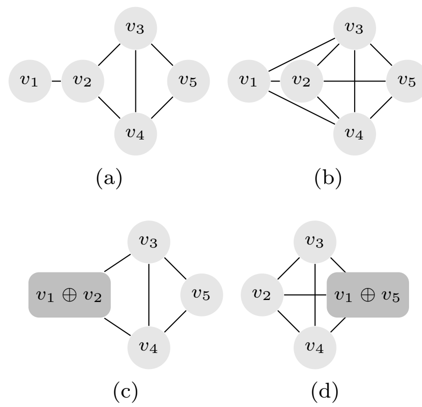

Figure 1: Example of 2-anonymizing an input graph. The input graph is depicted in (a), an optimal 2-anonymized graph with respect to edge addition is depicted in (b), an optimal 2-anonymized graph with respect to simple edge contraction or hybrid edge contraction is depicted in (c) (by contracting v_1 and v_2), and an optimal 2-anonymized graph with respect to non-simple vertex contraction is depicted in (d) (by contracting v_2 and v_5). Notice that there is no solution with respect to non-simple edge contraction, and the solution with respect to to edge addition is less efficient than the edge contractions solutions.

Degree Anonymization by Graph Contractions (DAGC)
Input: An undirected graph $G = (V, E)$, a budget $c \in \mathbb{N}$, and an anonymization level $k \in \mathbb{N}$.
Question: Can G be made k-anonymous by performing at most c contractions?

When the contraction operation is a simple (hybrid, non-simple) edge contraction operation, we denote the corresponding degree anonymization problem as SEC-A (respectively: HEC-A, NEC-A). Similarly, when the contraction operation is a simple (hybrid, non-simple) vertex contraction operation, we denote the corresponding degree anonymization problem as SVC-A (respectively: HVC-A, NVC-A).

We notice that sometimes it is not possible to anonymize a graph by simple/hybrid/non-simple edge/vertex contractions. As an example, consider n-anonymizing a complete graph with one missing edge: as the input graph is not n-anonymized, at least one edge needs to be contracted, but then the number of remaining vertices will be strictly less than n, rendering the graph not n-anonymous. This phenomenon stands in contrast to anonymization by edge addition, as completing any graph by adding all missing edges to it makes it n-anonymized. Interestingly, sometimes a graph can be made anonymous more efficiently by using edge contraction rather then edge addition (a very simple example is shown in Figure 1).

2.4 Overview

We study the parameterized complexity of degree anonymization by graph contractions, considering the solution size c, the anonymity level k, and the maximum degree Δ, as the most natural parameters. Table 1 gives an overview of our results.

We consider mainly SEC-A and HEC-A as these are the most common variants (see, for example, Diestel [11, Chapter 1.7] and Wolle and Bodlaender [24]). We mention that some of the results easily transfer to some other variants, while other results require additional research to transfer to other variants, and we leave this study for future research. We discuss these open questions further in Section 6. Our main results can be summarized as follows.

	solution size c	anonymization level k	maximum degree Δ
c	W-h[a] (Th. 3) XP (Obs. 1)	W-h[a] (Th. 3) XP (Obs. 1)	FPT (Th. 5)
k		Para-NP-h[a] (Th. 3)	FPT[b] (Cor. 1)
Δ			Para-NP-h[a] (Th. 4)

[a]Only for SEC-A and HEC-A.
[b]Only for NVC-A.

Table 1: Parameterized complexity landscape of Degree Anonymization by Graph Contractions. Rows and columns correspond to parameters, such that each cell corresponds to the combination of the corresponding parameters.

- Contrary to degree anonymization by some other graph operations (for example, by edge addition), here, even the underlying number problem (NVC-A) is NP-hard. Moreover, most variants considered here (SEC-A, HEC-A, and NVC-A) are NP-hard even on trees (in fact, even on caterpillar trees)
- Parameterizing by the solution size c, the maximum degree Δ, or the anonymity level k alone, does not help for tractability. However, combining Δ with c does help for tractability
- While we could show that combining the maximum degree Δ with the anonymity level k helps for tractability for some variants of the problem, we could show some evident against this claim for other variants.

3 NP-hardness

We begin by considering NVC-A, the main reason for considering this variant being the observation that for NVC-A, the structure of the graph does not matter, but only the degrees. This holds because any two vertices can be contracted, and the resulting graph after performing a contraction only depends on the original degrees of the contracted vertices. It follows that NVC-A is equivalent to the following number problem:

> An equivalent formulation of NVC-A
> **Input:** A set $V = \{d_1, \ldots, d_n\}$ of n integers ($\forall i : 0 \leq d_i \leq \Delta$) and two integers $k, c \in \mathbb{N}$.
> **Question:** Is there a partition $V = \bigcup_{i \in [z]} V_i$ (where $V_i \cap V_j = \emptyset$ for $i \neq j$) such that the set $S = \{\sum_{v \in V_i} v : i \in [z]\}$ is k-anonymized and $\sum_{i \in [z]}(|V_i| - 1) \leq c$?

Informally, this number problem is in the heart of the anonymization problem, and for this reason we call it the underlying number problem. Interestingly, contrary to the situation for edge addition and other operations, this underlying number problem is intractable (notice that, in order to be formally correct, we shall define the input to this number problem to be in unary; this is fine, as we next prove a reduction from a *strongly* NP-hard problem).

Theorem 1. NVC-A *is* NP*-hard even on caterpillar trees.*

Proof. We provide a reduction from the following strongly NP-hard problem [14]:

> Strictly Three Partition
> **Input:** A set of numbers $S = \{a_1, \ldots, a_{3m}\}$ such that $\sum_{a_i \in S} a_i = mB$ and $\forall i : B/4 < a_i < B/2$.
> **Question:** Are there m disjoint sets $S_1, \ldots, S_m$ such that $|S_i| = 3$ and $\forall j : \sum_{a_i \in S_j} a_i = B$?

Given an instance for Strictly Three Partition, we create an instance for NVC-A. Intuitively, the idea is to create a set of $3m$ vertices, such that each number a_i would have a corresponding vertex whose degree is proportional to a_i. Then, we will add a distinguished vertex with degree proportional to B, and we will make sure that the only way of anonymizing the block containing this distinguished vertex is by contracting m triplets of vertices corresponding to a triplet of numbers whose sum is m, that is, to a three partitioning. Details follow.

We scale the input numbers, specifically defining $a'_i = a_i \cdot mB$ and $B' = B \cdot mB$. For each number a'_i, we create a node $v_{a'_i}$ and connect it to another a'_i paths of length c consisting of new vertices, such that $\deg(v_{a'_i}) = a'_i$ holds for each i. We add a path-star of degree B' and length c. We set $k := m + 1$ and $c := 2m$ (indeed, G is a forest; it can be easily transformed into a caterpillar tree by placing all $v_{a'_i}$'s on a path together with the path-star, and adjusting the number of additional new vertices connected to each $v_{a'_i}$ accordingly).

Given a partitioning of S into triplets, it is possible to anonymize the graph by contracting, the three vertices $v_{a'_i}$ which correspond to each triplet, into a single vertex. Notice that we need two contractions to contract each triplet, and that the resulting graph is k-anonymized, as it has m new vertices of degree B'.

For the other direction, notice that due to the strictness constraint (that is, $\forall i : B/4 < a_i < B/2$), it follows that any witness set of size other than three will have degree which is far away from B'. Combining this observation with the fact that we multiplied each a_i by mB, it holds that if there is no partitioning of S to triplets, then in any partitioning, the degree of at least one triplet is far away from B' by at least mB, therefore the block containing the distinguished path-star cannot be anonymized in this way, Moreover, contracting the path-star itself does not help, as it can decrease its degree by at most c, which is not enough for it to fall on any other block. Therefore, any solution must introduce at least m new vertices of degree B', each corresponding to a triplet, therefore a solution must correspond into a partitioning of S into triplets. □

Informally, at least part of the hardness of SEC-A and HEC-A is due to the hardness of NVC-A, as it is somehow "in the core" of these problems as well. For SEC-A and HEC-A, we can show NP-hardness, even on a very restricted class of graphs.

Theorem 2. SEC-A *and* HEC-A *are both* NP*-hard even on caterpillar trees.*

Proof. We provide a reduction from the following strongly NP-hard problem [14]:

> Numerical Matching with Target Sums
> **Input:** Three sets of integers $A = \{a_1, \ldots, a_n\}$, $B = \{b_1, \ldots, b_n\}$, and $C = \{c_1, \ldots, c_n\}$.
> **Question:** Can the elements of A and B be paired such that c_i is the sum of the ith pair?

The variant where all $3n$ input integers are distinct is known to be also NP-hard in the strong-sense [19]. Without loss of generality, we also assume that all input integers are greater than 3. Given an instance for Numerical Matching with Target Sums, we create an instance for SEC-A and HEC-A. Intuitively, the idea is to create a set of $k - 1$ vertices for each c_i and a pair of vertices for each pair of a_i and b_j, such that the only possibility of anonymizing the vertices corresponding to the c_i's is to contract the correct pairs of a_i's and b_j's together. Details follows.

We set $k := n - 1$ and $c := n$. We construct some *c-gadgets:* for each c_i, we create $k - 1$ path-stars of degree $c_i - 2$ and length $c + 1$. We construct some *ab-gadgets:* for each pair of $i \in [n]$ and $j \in [n]$ we create two path-stars, one of degree a_i and length $c + 1$, and another of degree b_j and length $c + 1$, and connect them by an edge (indeed, the construction as such is a forest, but we can transform it into a tree by connecting arbitrarily each pair of disconnected components by a path of length $c + 1$). See Figure 2 for a visualization.

Correctness: Given a correct pairing of A and B, we can anonymize the input graph by contracting the corresponding *ab*-gadgets. For the other direction,

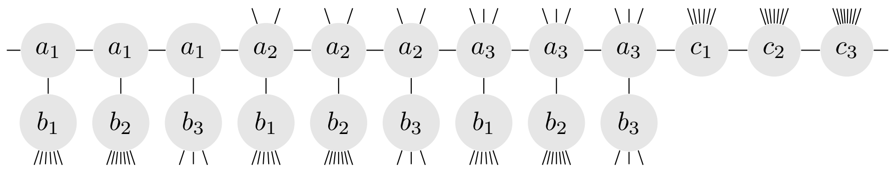

Figure 2: Example for the reduction used in the proof of Theorem 2. Specifically, the reduction is shown for the following instance for NUMERICAL MATCHING WITH TARGET SUMS: $A = \{3, 5, 6\}$, $B = \{7, 8, 4\}$, and $C = \{10, 11, 12\}$. All drawn edges, except for the edges from some a_i to some b_j, should be understood as paths of length 3.

notice that all of the c-gadgets must be anonymized in a solution, and because contracting c-gadgets together is not possible, they can be anonymized only by contracting some ab-gadgets. Notice also that it is not possible to contract two ab-gadgets sharing the same a_i or the same b_j, because this will introduce a new non-anonymized block, namely, the block of degree a_i. Therefore, a solution must correspond to a correct pairing. □

4 Non-structural parameters

In our quest for tractability, we go on to consider some non-structural parameters, beginning with the solution size c. It is easy to see that for constant c, we can simply use brute-force.

Observation 1. DAGC *is* XP *with respect to c.*

However, there is no hope for fixed-parameter tractability with respect to the solution size c, and even combining the solution size c with the anonymity level k does not help for tractability.

Theorem 3. *Both* SEC-A *and* HEC-A *are* NP*-hard and* W*-hard with respect to c, even if $k = 2$.*

Proof. For SEC-A, we provide a reduction from the following W[2]-hard problem, parameterized by the solution size [12]:

> SET COVER
> **Input:** Sets $S_1, \ldots, S_m$ containing elements from $x_1, \ldots, x_n$ and $h \in \mathbb{N}$.
> **Question:** Is there a set of at most h sets that covers all elements?

Given an instance for SET COVER, we create an instance for SEC-A. For each x_i we create a new vertex x'_i. For each S_j, we create two new vertices, S'_j and S''_j, and we connect them by an edge. Each S'_j and S''_j, corresponding to a set

S_j, are connected to all x'_i's corresponding to all elements $x_i \in S_j$. We add several paths of length h to each x'_i such that the degree of each x'_i will be $f(i) = ic + 2$. Similarly, we add several paths of length h to each S'_j and S''_j, such that the degree of each S'_j and S''_j will be $f(n+1)$. For every $i \in [n]$ and $z \in [h]$, we add a path-star of degree $f(i) - z$ and length c. We add k path-stars of degree $f(i+1)$ and length c in order to anonymize the vertices corresponding to the sets. We define $k := 2$ and $c := h$.

Given a set cover, by contracting together each pair of S'_j and S''_j corresponding to a set in the cover. the degrees of each x'_i will decrease by one, therefore the graph would be anonymized. For the other direction, notice that each x_i needs to be anonymized, and we can only decrease their degree (and not increase). Therefore, following a simple exchange argument, this must correspond to a set cover.

For HEC-A, we provide a reduction from the following W[1]-hard problem, parameterized by the solution size h [12] (an h-coloring is a function *color* : $V \to [h]$, assigning to each vertex v a color $\mathrm{color}(v) \in [h]$):

> Multi-Colored Clique
> **Input:** An undirected graph $G = (V, E)$ and an h-coloring of its vertices.
> **Question:** Is there a clique of size h including vertices of all h colors?

We notice that Cai [5] showed that Multi-Colored Clique remains hard even on regular graphs. We assume also, without loss of generality, that there are no monochromatic edges. Given an instance for Multi-Colored Clique, we create an instance for HEC-A. We define the following function, $f(i) = 2^i$, whose domain is the set of colors (that is, $i \in [h]$).

For every vertex v, we add $f(\mathrm{color}(v)) - \deg(v)$ paths of length c such that the degree of each vertex colored in color $i \in [h]$ is $f(i)$. We construct $k+1$ copies of this modified graph. We add $k - 1$ path-stars of degree $\sum_{i \in [h]} f(i) - 2\binom{h}{2}$ and length c. We set $k := 2$ and $c := h - 1$.

Given a multi-colored clique of size h, we contract the vertices of the clique into one vertex. The degree of the new vertex is equal to the degree of the path-stars, and the graph is anonymized, due to the $k+1$ copies.

For the other direction, notice that contracting the path-star does not change its degree. Moreover, as there are no monochromatic edges, we can only contract edges of different colors. Due to the way we defined $f(i)$, the only possible way of reaching the degree of the path-star (that is, $\sum_{i \in [h]} f(i) - 2\binom{h}{2}$), is by contracting a multi-colored clique, because all colors are needed for the first part (that is, $\sum_{i \in [h]} f(i)$), and all edges between the colors are needed for the second part (that is, $2\binom{h}{2}$). $\square$

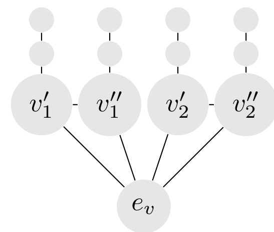

Figure 3: Gadget used for the reduction in the proof of Theorem 4 (for SEC-A). Specifically, the construction for two vertices, v_1 and v_2, connected by an edge $e = \{u, v\}$, is shown (indeed, as the input graph for the reduction is cubic, the vertices $v_{v_1'}$,$v_{v_1''}$,$v_{v_2'}$,$v_{v_2''}$ all have degree five, and not three, as is depicted here). The main crux is the fact that contracting $v_{v_1'}$ and v''_{v_1} together reduces the degree of e_v.

5 Structural parameters

We go on to consider the maximum degree Δ, as a natural structural parameter. It turns out that this parameter alone does not help for tractability.

Theorem 4. *Both* SEC-A *and* HEC-A *are para-*NP*-hard with respect to* Δ*.*

Proof. For SEC-A, we provide a reduction from the following NP-hard problem [14] (where a graph is said to be *cubic* if it is 3-regular):

> Vertex Cover on Cubic Graphs
> **Input:** An simple undirected cubic graph $G = (V, E)$ and $h \in \mathbb{N}$.
> **Question:** Is there a set of at most h vertices that covers all edges?

Given an instance for Vertex Cover on Cubic Graphs, we create an instance for SEC-A. For every edge $e \in E$, we create a new vertex v_e. For every vertex $v \in V$, we create a pair of new vertices v'_v and v''_v, and connect each such pair by an edge. For every edge $e = \{u, v\}$, we connect v_e to the four vertices v'_u, v''_u, v'_v, and v''_v. We also connect each v'_v and v''_v to a path of length c each. We add k path-stars of degrees 1, 2, 3, and 5, all with length c. We set $k := |E| + 1$ and $c := h$. See Figure 3 for a visualization.

The idea is that, given a vertex v, contracting the two vertices v'_v and v''_v corresponding to it, together, would decrease the degrees of the vertices v_e corresponding to all incident edges e of v, therefore, a vertex cover would correspond to decreasing the degrees of all the vertices corresponding to edges by at least one, thus anonymizing the block of degree four. More formally, given a vertex cover, we contract each pair of v'_v and v''_v that corresponds to a vertex v in the vertex cover. As a result, the degree of each vertex v_e is decreased from 4 to either 3 or 2, and the graph is k-anonymized.

For the other direction, notice that the block of degree four needs to be anonymized. However, there is no way of increasing the degree of the vertices in this block since any contraction would decrease their degree, therefore, their degree must decrease, in any solution. For an edge $e = \{u, v\}$, the only possibility of decreasing the degree of v_e is by contracting either the pair v'_u and v''_u or the pair v'_v and v''_v. By a simple exchange argument, this must correspond to a vertex cover.

For HEC-A, we reduce from the following problem, which was shown by van Rooij et al. [23] to be NP-hard even on 4-regular graphs:

> PARTITION INTO TRIANGLES
> **Input:** An undirected graph $G = (V, E)$
> **Question:** Can V be partitioned into sets $S_1, \ldots, S_{|V|/3}$ of size 3 each such that each S_i forms a triangle in G?

Given an 4-regular input graph G for PARTITION INTO TRIANGLES on 4-regular graphs, we create an input graph G' for HEC-A. We initialize G' by G, and add a path of length c to each vertex, consisting of new vertices, such that the G' is a 5-regular graph. We create a path-star of degree 9 and length c. We set $k := n/3 + 1$ and $c := 2n/3$ (we assume, without loss of generality, that $n \bmod 3 = 0$).

Given a partition of G into triangles, we contract the vertices of each triangle together. The degree of each such witness set is nine, therefore the graph is k-anonymized. For the other direction, notice that the path-star of degree nine needs to be anonymized, its degree cannot decrease or increase, and the only way of having other vertices of degree nine is by contracting some triangles. Finally, as we need $n/3$ of these triangles, this must correspond to a partitioning of G into triangles. □

Contrary to the above hardness results, combining the maximum degree with the solution size does help for tractability, for all variants of DEGREE ANONYMIZATION BY GRAPH CONTRACTIONS.

Theorem 5. DAGC *is* FPT *with respect to* (Δ, c).

Proof. Consider a yes-instance for DAGC. Then, there exists a set E' of at most c edges such that contracting them results in a k-anonymous graph. Consider a set V' of vertices, containing all endpoints of the edges in E', including all their neighbors. As each edge has two endpoints and each vertex has at most Δ neighbors, it follows that $|V'| \leq 2c(\Delta + 1)$. Consider the set V'' containing all vertices whose degree will be changed due to contracting the edges of E'. As it holds that $V'' \subseteq V'$, it is enough to find the subgraph induced by V'.

To this end, we consider all possible graphs H containing at most $2c(\Delta + 1)$ vertices. For each such graph H, we consider all possible sets C of at most c edges to be contracted. For each such pair of a graph H and a set C, we compute the degree changes in H incurred by contracting the edges in C. If these degree changes make the graph k-anonymous, then we try to find this graph H as a subgraph in G. This step can be performed using, for example, the result by Cai et al. [6]. □

We consider now the combined parameter Δ and k. First, for NVC-A, we can bound c in these parameters and thus the FPT-algorithm with respect to (Δ, k), indeed, also proves fixed-parameter tractability with respect to (Δ, c).

Lemma 1. *For any yes-instance* (V, k, c) *of* NVC-A *it holds that* (V, k, c') *with* $c' = k \cdot (\Delta \cdot \Delta!)^{\Delta}$ *is also a yes-instance.*

Proof. Let (V, k, c) be a yes-instance of NVC-A and denote by $c_{\text{opt}} \leq c$ the smallest number such that (V, k, c_{opt}) is still a yes-instance. Moreover, let the partition $P = \{V_1, \ldots, V_i\}$ of V be a solution which corresponds to c_{opt} (that is, P is the witness structure corresponding to a solution of (V, k, c_{opt})). In the following we define two operations on P, with the property that applying each of them, when it is applicable, results in another solution with less than c_{opt} contractions. Since we show that at least one of them is applicable in case of $c_{\text{opt}} > k \cdot (\Delta \cdot \Delta!)^{\Delta}$, this proves Lemma 1.

To formally describe our operations, we associate with each witness set V_i a witness vector $\vec{v_i} \in \mathbb{N}^{\Delta}$ with $\vec{v_i}[j]$ being equal to the number of vertices of degree j in the witness set V_i. The *degree* of a witness set is defined to be the sum of the degrees of the vertices in the witness set (that is, the degree of the vertex corresponding to contracting all of the vertices in the witness set).

Operation 1: This operation is applicable to P if there are at least k witness sets in P of equal degree and such that in each of them, say V_i, there is at least one j with $\vec{v_i}[j] \geq \Delta!$. If there exist such a collection of witness sets, then consider such a collection P which is maximal with respect to containment, and do the following. For each witness set V_i in this collection, let j be an integer with $\vec{v_i}[j] \geq \Delta!$. remove $(\Delta!/j)$-many vertices of degree j from V_i (notice that $\Delta!/j$ is always an integer), and form a new witness set containing these vertices.

We introduced at least k new witness sets all being of degree exactly $\Delta!$ and we decreased the degree of each of the initial witness sets in by the same number $\Delta!$. Since there are at least k of such witness sets, it follows that performing this operation results in a partition of V that is still a solution for (V, k, c_{opt}), but requires less edge contractions than P.

Operation 2: This operation is applicable to P if there is a collection of at least k witness sets in P, such that the witness sets in the collection all have the same witness vector, and this same witness vector is of hamming weight of at least 2 (that is, these are not singletons). If such a collection exists, then choose an arbitrary integer j occurring in this same witness vector. Then, for each witness set V_i in this collection, remove one vertex of degree j from V_i, and form a new witness set containing only this vertex of degree j (that is, form a new singleton witness set).

Since there are at least k witness sets where a vertex of same degree j is cut out from them, the resulting partition is a solution for V which requires less edge contractions than P.

Applicability: It remains to argue that in case of $c_{\text{opt}} > k \cdot (\Delta \cdot \Delta!)^{\Delta}$, at least one of the two operations described above is applicable. First, assume that P contains a witness set V_i of degree at least $(\Delta \cdot \Delta!)$. Then, since P is k-anonymous there are at least k witness set of the same degree, which is at least $(\Delta \cdot \Delta!)$. It follows that each of these witness sets must contain at least one integer j which occurs at least $\Delta!$ times in it. Thus, Operation 1 is applicable.

So, let assume now that the degree of each witness set in P is at most $(\Delta \cdot \Delta!)$. Then, since a witness set cannot contain vertices of degree 0, we have that there are at most $(\Delta \cdot \Delta!)^{\Delta-1}$ different witness vectors such that none of them has degree greater or equal to $(\Delta \cdot \Delta!)$. Hence, if P contains at least $k \cdot (\Delta \cdot \Delta!)^{\Delta-1}$ witness sets of size at least two, then Operation 2 is applicable.

Finally, a solution for which $c_{\text{opt}} > k \cdot (\Delta \cdot \Delta!)^{\Delta}$ edge contractions have been performed either contains a set of size at least $(\Delta \cdot \Delta!)$ or it contains at least

$$\frac{k \cdot (\Delta \cdot \Delta!)^{\Delta}}{(\Delta \cdot \Delta!)} = k \cdot (\Delta \cdot \Delta!)^{\Delta-1}$$

witness sets of size at least two. □

Using this Lemma, we can show the following.

Corollary 1. NVC-A *is* FPT *with respect to* (Δ, k).

Proof. For a given instance (V, k, c) of NVC-A we decide the instance $(V, k, \min\{c, k \cdot (\Delta \cdot \Delta!)^{\Delta}\})$ using the FPT-algorithm with respect to (Δ, c). By Lemma 1 these two instances are equivalent and the corresponding running time proves fixed-parameter tractability with respect to (Δ, k). □

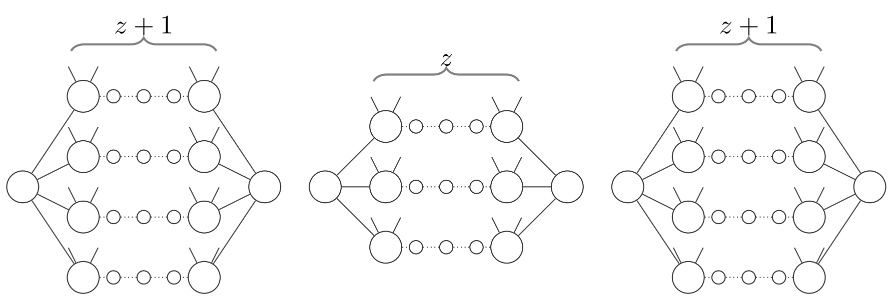

Figure 4: An input for SEC-A with small maximum degree Δ and small anonymity level k ($k := 3$), but with unbounded (in Δ and k) solution size c and output maximum degree Δ'. The edges going out of the nodes but reaching nothing are actually paths of length $z + 2$. The small nodes also have two paths of length $z+2$ going out of each of them, but these are omitted for picture clarity. For NEC-A, the same construction works, but we do not need the paths of length $z + 2$.

Consider the relation between k, Δ, Δ', and c, where Δ' denotes the maximum degree in the anonymized solution graph. While for NVC-A we could upper-bound both Δ' and c in k and Δ, as shown in the proof of Lemma 1, for SEC-A and NEC-A this is not always true, as the graph depicted in Figure 4 shows: for this graph we have $\Delta = 4$. If we additionaly set $k := 3$, then the best solution for SEC-A will require $3z + 4$ edge contractions. Therefore, in this example, we have that $c_{Opt} = \Omega(n)$ and also $\Delta' = \Omega(n)$. This means that either a completely different proof technique is needed for these variants, or, as we conjecture, that these variants are not fixed-parameter tractable for this combined parameter.

6 Conclusion

We investigated the parameterized complexity of degree anonymization by several variants of graph contractions. We showed that most of the variants are intractable, even on very restricted graph classes, and that even the underlying number problem is NP-hard, contrary to degree anonymization by other studied operations. However, we could find some fixed-parameter tractable cases.

For further research, one could consider some related graph operations, such as *edge splitting* (removing an edge and introducing a new vertex, connecting it to the endpoints of the removed edge), *structure contraction* (contracting a whole subgraph at the cost of one operation), and edge twisting or vertex dissolution (both defined, for example, in [22, Chapter 3]). Another research direction could be different notions of approximations, that is, either anonymizing the input graph by using more budget than is allowed, or only partially anonymizing it.

References

[1] T. Asano and T. Hirata. Edge-contraction problems. *J. Comput. Syst. Sci.*, 26(2):197–208, 1983.

[2] R. Belmonte, P. A. Golovach, P. Hof, and D. Paulusma. Parameterized complexity of three edge contraction problems with degree constraints.

[3] R. Bredereck, S. Hartung, A. Nichterlein, and G. J. Woeginger. The complexity of finding a large subgraph under anonymity constraints. In *Proc. 24th ISAAC*, volume 8283 of *LNCS*, pages 152–162. Springer, 2013.

[4] R. Bredereck, V. Froese, S. Hartung, A. Nichterlein, R. Niedermeier, and N. Talmon. The complexity of degree anonymization by vertex addition. In *Proceedings of AAIM 2014*. Springer, 2014.

[5] L. Cai. Parameterized complexity of cardinality constrained optimization problems. *The Computer Journal*, 51(1):102–121, 2008.

[6] L. Cai, S. M. Chan, and S. O. Chan. Random separation: A new method for solving fixed-cardinality optimization problems. In *Parameterized and Exact Computation*, pages 239–250. Springer, 2006.

[7] J. Casas-Roma, J. Herrera-Joancomartí, and V. Torra. An algorithm for k-degree anonymity on large networks. In *Proc. ASONAM'13*, pages 671–675. ACM Press, 2013.

[8] S. Chester, B. M. Kapron, G. Ramesh, G. Srivastava, A. Thomo, and S. Venkatesh. Why Waldo befriended the dummy? k-anonymization of social networks with pseudo-nodes. *Social Netw. Analys. Mining*, 3(3): 381–399, 2013. Available online.

[9] K. L. Clarkson, K. Liu, and E. Terzi. Towards identity anonymization in social networks. In *Link Mining: Models, Algorithms, and Applications*, pages 359–385. Springer, 2010.

[10] D. Delling, R. Görke, C. Schulz, and D. Wagner. Orca reduction and contraction graph clustering. In *Algorithmic Aspects in Information and Management*, pages 152–165. Springer, 2009.

[11] R. Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, 4th edition, 2010.

[12] R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity*. Springer, 2013.

[13] J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer, 2006.

[14] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, 1979.

[15] P. A. Golovach, P. Hof, and D. Paulusma. Obtaining planarity by contracting few edges. In *MFCS 2012*, pages 455–466. Springer, 2012.

[16] S. Guillemot and D. Marx. A faster FPT algorithm for bipartite contraction. *Information Processing Letters*, 113(22):906–912, 2013.

[17] S. Hartung, A. Nichterlein, R. Niedermeier, and O. Suchý. A refined complexity analysis of degree anonymization in graphs. In *Proc. 40th ICALP*, volume 7966 of *LNCS*, pages 594–606. Springer, 2013.

[18] P. Heggernes, P. Hof, B. Lévêque, D. Lokshtanov, and C. Paul. Contracting graphs to paths and trees. *Algorithmica*, 68(1):109–132, 2014.

[19] H. Hulett, T. G. Will, and G. J. Woeginger. Multigraph realizations of degree sequences: Maximization is easy, minimization is hard. *Operations Research Letters*, 36(5):594–596, 2008.

[20] X. Lu, Y. Song, and S. Bressan. Fast identity anonymization on graphs. In *Proc. 23rd DEXA*, volume 7446 of *LNCS*, pages 281–295. Springer, 2012.

[21] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford University Press, 2006.

[22] J. G. Oxley. *Matroid theory*, volume 3. Oxford university press, 2006.

[23] J. M. M. van Rooij, M. E. van Kooten Niekerk, and H. L. Bodlaender. Partition into triangles on bounded degree graphs. *Theory Comput. Syst.*, 52(4):687–718, 2013.

[24] T. Wolle and H. L. Bodlaender. A note on edge contraction. Technical report, Technical Report UU-CS-2004, 2004.