A Refined Complexity Analysis of Degree Anonymization in Graphs☆

Sepp Hartung[a], André Nichterlein[a], Rolf Niedermeier[a], Ondřej Suchý[b]

[a] *Institut für Softwaretechnik und Theoretische Informatik, TU Berlin, Germany*
{`sepp.hartung, andre.nichterlein, rolf.niedermeier`}`@tu-berlin.de`
[b] *Faculty of Information Technology, Czech Technical University in Prague, Czech Republic*
`ondrej.suchy@fit.cvut.cz`

Abstract

Motivated by a strongly growing interest in graph anonymization, we study the NP-hard Degree Anonymity problem asking whether a graph can be made k-anonymous by adding at most a given number of edges. Herein, a graph is k-anonymous if for every vertex in the graph there are at least $k-1$ other vertices of the same degree. Our algorithmic results shed light on the performance quality of a popular heuristic due to Liu and Terzi [ACM SIGMOD 2008]; in particular, we show that the heuristic provides optimal solutions if "many" edges need to be added. Based on this, we develop a polynomial-time data reduction yielding a polynomial-size problem kernel for Degree Anonymity parameterized by the maximum vertex degree. In terms of parameterized complexity analysis, this result is in a sense tight since we also show that the problem is already NP-hard for H-index three, implying NP-hardness for smaller parameters such as average degree and degeneracy.

Keywords: Parameterized Complexity, Kernelization, Heuristics, f-Factors, Data Privacy

1. Introduction

For many scientific disciplines, including the understanding of the spread of diseases in a globalized world or power consumption habits with impact on fighting global warming, the availability of social network data becomes more and more important. To respect privacy issues, there is a strong demand to anonymize the associated data in a preprocessing phase [21]. In a landmark paper, Liu and Terzi [31] (also see Clarkson et al. [14] for an extended version) introduced the following simple graph-theoretic model for identity anonymization

☆A preliminary version appeared in the *Proceedings of the 40th International Colloquium on Automata, Languages, and Programming (ICALP 2013), Part II*, Riga, Latvia, July 2013, Volume 7966 of Lecture Notes in Computer Science, pages 594–606, Springer.

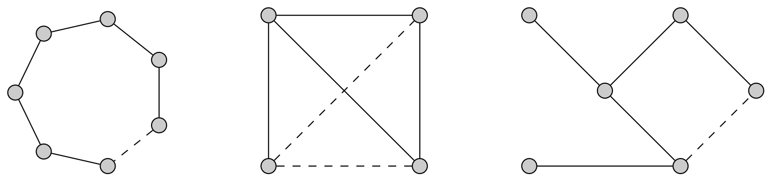

Figure 1: Three illustrating examples. The solid edges indicate the original graphs. Adding the dashed edges changes the graphs (from left to right) from being 2-anonymous to 7-anonymous, from 1-anonymous to 4-anonymous, and from 1-anonymous to 2-anonymous.

on (social) networks. Herein, they transferred the k-anonymity concept known for tabular data in databases [21, 41, 42, 43] to graphs (see Figure 1 for examples).

Degree Anonymity [31]
Input: An undirected graph $G = (V, E)$ and two positive integers k and s.
Question: Is there an edge set E' over V with $|E'| \leq s$ such that $G' = (V, E \cup E')$ is *k-anonymous*, that is, for every vertex $v \in V$ there are at least $k-1$ other vertices in G' having the same degree?

Liu and Terzi [31] assume in this model that an adversary (who wants to de-anonymize the network) knows only the degree of the vertex of a target individual; this is a modest adversarial model. Clearly, there are stronger adversarial models which (in many cases very realistically) assume that the adversary has more knowledge, making it possible to breach privacy provided by a "k-anonymized graph" [36]. Moreover, it has been argued that graph anonymization has fundamental theoretical barriers which prevent a fully effective solution [1]. Degree Anonymity, however, provides the perhaps most basic and still practically relevant model for graph anonymization; it is the subject of active research [10, 11, 13, 33].

Graph anonymization problems are typically NP-hard. Thus, almost all algorithms proposed in this field are heuristic in nature, this also being true for algorithms for Degree Anonymity [25, 31, 33]. Indeed, as the field of graph anonymization is young and under strong development, there is very little research on its theoretical foundations, particularly concerning computational complexity and algorithms with provable performance guarantees [11].

Our contributions. Our central result is to show that Degree Anonymity has a polynomial-size problem kernel when parameterized by the maximum vertex degree Δ of the input graph. In other words, we prove that there is a polynomial-time algorithm that transforms any input instance of Degree Anonymity into an equivalent instance with at most $O(\Delta^7)$ vertices. Indeed, we encounter a "win-win" situation when proving this result: We show that Liu and Terzi's heuristic strategy [31] finds an optimal solution when the size s of a minimum solution is larger than $2\Delta^4$. As a consequence, we can bound s in $O(\Delta^4)$ and, hence, a polynomial kernel we provide for the combined parameter (Δ, s) actually is also a polynomial kernel only for Δ. Furthermore, our kernelization has the useful property (for instance when combining it with approximation algorithms) that each solution derived for the kernel instance one-to-one corresponds to a solution of the original instance. While this kernelization directly implies fixed-parameter tractability for DEGREE ANONYMITY parameterized by Δ, we also develop a further fixed-parameter algorithm with an improved worst-case running time.

In addition, we prove that DEGREE ANONYMITY becomes NP-hard on graphs with H-index[1] three. The same proof also yields NP-hardness in 3-colorable graphs. Further, adopting the viewpoint of "standard parameterization", we show that DEGREE ANONYMITY is W[1]-hard when parameterized by the solution size s (the number of added edges), even when $k = 2$. In other words, there is no hope for fixed-parameter tractability even when the level k of anonymity is low and the graph needs only few edge additions (meaning little perturbation) to achieve k-anonymity.

Why is the parameter "maximum vertex degree Δ" of specific interest? First, note that from a parameterized complexity perspective it seems to be a "tight" parameterization in the sense that for the only little "stronger" (that is, provably smaller) parameter H-index our results already show NP-hardness for H-index three (also implying hardness e.g. for the parameters degeneracy and average degree). Social networks typically have few vertices with high degree and many vertices of small degree. Leskovec and Horvitz [30] studied a huge instant-messaging network (180 million vertices) with maximum degree 600. For the DBLP co-author graph[2] generated in February 2012 and containing more than 715,000 vertices we measured a maximum degree of 804 and an H-index of 208, so there are not more than 208 vertices with degree larger than 208. Thus, a plausible strategy might be to only anonymize vertices of "small" degree and to remove high-degree vertices for the anonymization process because it might be overly expensive to anonymize these high-degree vertices and since they might be well-known (that is, not anonymous) anyway. Indeed, high-degree vertices can be interpreted as outliers [2], potentially making their removal plausible.

Related work. The most important reference is Liu and Terzi's work [31] where the basic model of graph anonymization was introduced, sophisticated (heuristic) algorithms (also using algorithms to determine the realizability of degree sequences) have been developed and validated on experimental data. Somewhat more general models have been considered by Zhou and Pei [44] (studying the neighborhood of vertices instead of only the degree) and by Chester et al. [11] (anonymizing a *subset* of the vertices of the input). Chester et al. [13] investigated

[1]The H-index of a graph G is the maximum integer h such that G has at least h vertices with degree at least h. As a consequence, if G has H-index h, then it has at most h vertices of degree larger than h.

[2]In this graph the vertices represent the authors and an edge indicates that the two corresponding authors are co-authors of at least one paper.

the variant of adding vertices instead of edges; Bredereck et al. [8] provided first parameterized complexity results in this direction. Recently, building on Liu and Terzi's work, we enhanced their heuristic approach with the focus on improving lower and upper bounds on the solution size [25]. Lu et al. [33] and Casas-Roma et al. [10] proposed enhanced algorithms for Degree Anonymity. Again, these algorithms are heuristic in nature. Today, the field of graph anonymization has grown tremendously with numerous surveys and research directions. We only mention some directly related work.

Chester et al. [12] are among the few having performed formal computational complexity studies of Degree Anonymity and edge-labeled variants. On the positive side, they showed a polynomial-time algorithm for the unlabeled case on bipartite graphs. In particular, they ask for effective approximation algorithms (for the optimization versions of the underlying decision problems) for NP-hard problems and complain the lack of complexity investigations and theoretical research. In a sense, with our main result on polynomial-time provably effective preprocessing, we provide a "stricter" result than polynomial-time approximation since our reduced instance still allows for finding optimal solutions. Bredereck et al. [7] studied the variant of Degree Anonymity where one uses at most s vertex deletions instead of edge additions to make the given graph k-anonymous. They showed that this variant is NP-hard even on several restricted graph classes (e. g. on trees and on graphs with maximum degree three), but becomes fixed-parameter tractable with respect to each of the combined parameters (s, k) and (s, Δ). Bazgan and Nichterlein [3] studied the (parameterized) approximability of Degree Anonymity and its vertex deletion variant; they achieved several inapproximability results.

There are many other, often more complicated models for graph anonymization. Weaknesses of Degree Anonymity (mainly depending on the assumed adversary model where for many practical situations the adversary may e.g. have an auxiliary network that helps in de-anonymizing) and other models have been pointed out [1, 36, 40]. Notably, differential privacy, a "noise addition" framework, which is successfully used for anonymizing tabular data (instead of more combinatorial models such as k-anonymity), incurs other difficulties applied to anonymizing graphs [36, 40]. In conclusion, given the generality of background knowledge an adversary may or may not have, graph anonymization remains a chimerical target [33] and, thus, a universally best model is not available.

From a (parameterized) computational complexity perspective, the closest work we are aware of in terms of graph problems is due to Mathieson and Szeider [34] who provided a study on editing graphs to satisfy degree constraints. In their basic model, each vertex is equipped with a degree list and the task is to edit the graph such that each vertex achieves a degree contained in its degree list. They studied the editing operations edge addition, edge deletion, and vertex deletion, and achieved numerous parameterized tractability and intractability results. Interestingly, on the technical side they also relied on the computation of general factors in graphs (as we do) and they also studied kernelization, where they left as most challenging open problem to extend their kernelization results to cases that include vertex deletion and edge addition, emphasizing that the presence of edge additions is making their approach inapplicable. Earlier work by Moser and Thilikos [35] studied the case of obtaining an r-regular graph by vertex deletions.

Finally, we mention in passing that there is recent work on studying the parameterized complexity of k-Anonymity on tabular data with numerous tractability and intractability results [5, 6, 9, 17].

2. Preliminaries

Parameterized complexity. A parameterized problem is called *fixed-parameter tractable* if there is an algorithm that decides any instance (I, k), consisting of the "classical" instance I and a parameter $k \in \mathbb{N}_0$, in $f(k) \cdot |I|^{O(1)}$ time, for some computable function f solely depending on k.

A core tool in the development of fixed-parameter algorithms is polynomial-time preprocessing by data reduction, called *kernelization*[3] [4, 24, 29]. Here, the goal is to transform a given problem instance (I, k) in polynomial time into an equivalent instance (I', k') whose size is upper-bounded by a function of k. That is, (I, k) is a yes-instance if and only if (I', k'), $k' \leq g(k)$, and $|I'| \leq g(k)$ for some function g. Thus, such a transformation is a polynomial-time self-reduction with the constraint that the reduced instance is "small" (measured by $g(k)$). In case that such a transformation exists, I' is called *kernel* of size $g(k)$. Furthermore, if g is a polynomial, then it I' is called a *polynomial kernel*.

Parameterized complexity theory also a hardness program (most prominently, W[1]-hardness) that allows to show presumable fixed-parameter *intractability*. One can show W[1]-hardness of a parameterized problem L by providing a *parameterized reduction* from a W[1]-hard problem (e.g. Independent Set parameterized with the solution size) 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') (with $k' \leq g(k)$) such that $(I, k) \in L \Leftrightarrow (I', k') \in L'$. A W[1]-hard problem does not admit a fixed-parameter algorithm, unless the widely believed conjecture FPT $\neq$ W[1] fails. We refer to the monographs [16, 20, 37] for more details on parameterized complexity.

Graphs and k-anonymity. We use standard graph-theoretic notation. All graphs studied in this paper are undirected and simple, that is, there are no self-loops and no multi-edges. For a given graph $G = (V, E)$ with vertex set V and edge set E we set $n := |V|$ and $m := |E|$. Furthermore, by $\deg_G(v)$ we denote the degree of a vertex $v \in V$ in G and Δ_G denotes the maximum occurring vertex degree in G. For $0 \leq d \leq \Delta_G$ let $D_G(d) := \{v \in V \mid \deg_G(v) = d\}$ be the *block* of degree d, that is, the set of all vertices with degree d in G. Thus, being k-anonymous is equivalent to each block being of size either zero or at least k.

[3] It is well-known that a parameterized problem is fixed-parameter tractable if and only if it has a kernelization.

The complement graph of G is denoted by $\overline{G} := (V, \overline{E})$, $\overline{E} := \{\{u, v\} \mid u, v \in V, \{u, v\} \notin E\}$. The subgraph of G induced by a vertex subset $V' \subseteq V$ is denoted by $G[V']$. For an edge subset $E' \subseteq E$, $V(E')$ denotes the set of all endpoints of edges in E' and $G[E'] := (V(E'), E')$. For a set S of edges with endpoints in a graph G, we denote by $G + S$ the graph that results by inserting all edges in S into G and we call S an *edge insertion set* for G. Thus, Degree Anonymity is the question whether there is an edge insertion set S of size at most s such that $G + S$ is k-anonymous. In this case S is called *k-insertion set* for G. We omit subscripts if the graph is clear from the context.

3. Computational Hardness

In this section we provide two polynomial-time many-to-one reductions yielding three (parameterized) hardness results: Degree Anonymity is

- NP-hard on 3-colorable graphs,
- NP-hard on graphs with H-index three (that is on graphs with at most three vertices of degree more than three), and
- W[1]-hard with respect to s even if $k = 2$.

The problem is fixed-parameter tractable with respect to the maximum degree, showing that a small maximum degree makes the problem easy. Interestingly, the reduction given in the next proof contains exactly one vertex with degree more than three, showing that one high-degree vertex is sufficient to make the problem hard.

Theorem 1. Degree Anonymity *is NP-hard on 3-colorable graphs and on graphs with H-index three.*

Proof. We give a reduction from the NP-hard Independent Set problem.

Independent Set [23, GT20]
Input: A graph $G = (V, E)$ and a positive integer h.
Question: Is there an independent set $V' \subseteq V$ of size $|V'| = h$, that is, a vertex subset of pairwise nonadjacent vertices?

We assume without loss of generality that in the given Independent Set instance (G, h) it holds that $|V| \geq 2h + 1$. We construct an equivalent instance $(G' = (V', E'), k, s)$ for Degree Anonymity as follows. We start with a copy G' of G, denoting with $v' \in V'$ the copy of the vertex $v \in V$. Then, for each vertex $v \in V$ we add to G' degree-one vertices adjacent to v' such that v' has degree Δ_G in G'. Finally, we add a star with $\Delta_G + h - 1$ leaves and denote its central vertex c. We conclude the construction by setting $k := h + 1$ and $s := \binom{h}{2}$.

We prove the correctness of the reduction by showing that (G, h) is a yes-instance of Independent Set if and only if (G', k, s) is a yes-instance of Degree Anonymity.

"⇒:" Let $I \subseteq V$ be an independent set in G with $|I| = h$. We show that the edge set $\binom{I}{2}$ is a solution for (G', k, s): Since I is an independent set, none of the edges in $\binom{I}{2}$ is contained in G'. Furthermore, observe that $G' + \binom{I}{2}$ is k-anonymous: There are three different degrees in the degree distribution of G': 1, Δ_G, and $\Delta_G + h - 1$. Obviously, there are at least k degree-one and degree-Δ_G vertices. Furthermore, the vertices with degree $\Delta_G + h - 1$ are all in $I \cup \{c\}$. Thus, there are $|I| + 1 = k$ vertices with degree $\Delta_G + h - 1$. Finally, observe that $|\binom{I}{2}| = \binom{h}{2}$.

"⇐:" Let $E_s \subseteq \binom{V'}{2}$ be a solution to (G', k, s) with $|E_s| = s$. The following degrees occur in G': 1, Δ_G, and $\Delta_G + h - 1$. Furthermore, observe that there is exactly one vertex with degree $\Delta_G + h - 1$ in G'. In $G' + E_s$ there must be at least $k - 1 = h$ further vertices of degree at least $\Delta_G + h - 1$ and, hence, each of them has to have at least $h - 1$ incident edges in E_s. Thus there are exactly h such vertices, each incident to exactly $h - 1$ edges in E_s. These h vertices form an independent set of size h in G' and, by construction, the h corresponding vertices form an independent set of size h in G. This completes the proof of the correctness of the reduction.

Independent Set is NP-hard on 3-colorable graphs [39, Lemma 6] and on graphs with maximum degree three [23, GT20]. Clearly, if G is 3-colorable, then G' is 3-colorable as well. Furthermore, if G has maximum degree three, then only the central vertex c has degree larger than three, implying that the H-index of G' is three. □

The NP-hardness for constant H-index directly implies analogous NP-hardness results for the prominent parameters average degree and degeneracy[4]. We next prove W[1]-hardness for the "standard parameterization", that is, the number of edges s that are may be added.

Theorem 2. Degree Anonymity *is W[1]-hard parameterized by the number of inserted edges s, even if $k = 2$.*

Proof. We give a parameterized reduction from the Multicolored Independent Set problem.

Multicolored Independent Set
Input: A graph $G = (V, E)$, an integer h, and a vertex coloring $\mathrm{col}\colon V \to \{1, \ldots, h\}$.
Question: Is there a multicolored independent set $V' \subseteq V$ of size $|V'| = h$, that is, for every pair of vertices $u, v \in V'$ it holds that $\mathrm{col}(u) \neq \mathrm{col}(v)$ and $\{u, v\} \notin E$?

The W[1]-hardness of Multicolored Independent Set directly follows from the W[1]-hardness of the Multicolored Clique problem [18]. We assume without loss of generality that each color class contains at least three vertices.

[4]A graph G has degeneracy d if every subgraph of G (including G) contains a vertex of degree at most d.

Given a Multicolored Independent Set instance (G, h, col), we construct an equivalent instance $(G' = (V', E'), k, s)$ for Degree Anonymity as follows. We start with copying the graph G to G'. Then, for each vertex $v \in V$ we add in G' degree-one vertices adjacent to v until v has degree $h^3 \cdot \mathrm{col}(v) + \Delta_G$ in G'. Next we add h disjoint stars to G'—one for each color in $\{1, \ldots, h\}$. The star for color i has $h^3 i + \Delta_G + h - 1$ leaves and its central vertex is denoted by w_i. We conclude the construction by setting $k := 2$ and $s := \binom{h}{2}$.

We prove the correctness of the reduction by showing that (G, h, col) is a yes-instance of Multicolored Independent Set if and only if $(G', 2, s)$ is a yes-instance of Degree Anonymity.

"$\Rightarrow$:" Let $I \subseteq V$ be a multicolored independent set in G with $|I| = h$. It is easy to verify that $E_s = \binom{I}{2}$ is a k-insertion set for G' of size $|E_s| = s = \binom{h}{2}$.

"$\Leftarrow$:" Let S be a k-insertion set for G' with $|S| \leq s$. Observe that G' contains h blocks $D_{G'}(h^3 i + \Delta_G + h - 1)$ for $i \in \{1, \ldots, h\}$ of size exactly one. Since $k = 2$, this implies that $|V(S)| \geq h$. Since for $i \in \{1, \ldots, h\}$ there is no vertex in G' with degree $h^3 i + \Delta_G + h - 1 + j$ for any $j \in \{1, \ldots, h^2\}$ and $s = \binom{h}{2} < h^2$, it follows that in order to get a vertex of the same degree as w_i, the set S must increase the degree of at least one vertex by at least $h - 1$. It follows that such a vertex must have degree $h^3 i + \Delta_G$ in G', there is one such vertex for each $i \in \{1, \ldots, h\}$ in $V(S)$, and each of them is incident to exactly $h - 1$ edges of S. Due to the size of S this implies that $|V(S)| = h$ and $V(S)$ is an independent set in G' and, by construction, also in G. Furthermore, by construction, $V(S)$ is multicolored. $\square$

4. Polynomial Kernel for the Parameter Maximum Degree

In this main section of our work we provide a polynomial kernel with respect to the parameter maximum degree Δ (Theorem 4). To this end, we first analyze the heuristic Liu and Terzi [31] proposed to solve Degree Anonymity. Basically, this heuristic runs in three steps as follows (see Figure 2 for an example and Section 4.2 for the technical details):

1. Compute the degree sequence of the given graph G.
2. k-anonymize the degree sequence.
3. Realize the k-anonymized degree sequence as a super-graph of G.

The heuristic may fail to find a solution if the anonymized degree sequence computed in Step 2 cannot be realized in Step 3. However, in Section 4.2 we show that if there is a "large" difference between the degree sequence and the anonymized degree sequence, then there is always a realization of the anonymized degree sequence. This leads, as the heuristic runs in polynomial time, to the following win-win situation: For a given instance of Degree Anonymity, one can either find a k-insertion set in polynomial time using the above approach, or the solution—if existing—is "small" (containing less than $(\Delta^2 + 4\Delta + 3)^2$ edges). This win-win situation enables us to show that a polynomial kernel with respect

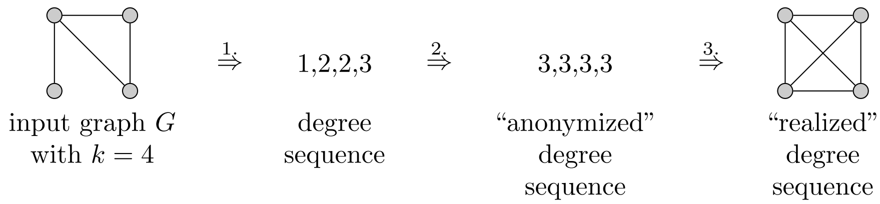

Figure 2: Example for the three basic steps in the heuristic of Liu and Terzi [31]. Step 1: Compute the degree sequence. Step 2. Anonymize the degree sequence (ignoring the graph), that is, increase its numbers such that each resulting number occurs at least k times. Step 3. Realize the anonymized degree sequence as super-graph of G.

to the combined parameter (Δ, s) provided in Section 4.3 is indeed polynomial only in Δ.

We begin, however, with presenting the main technical tool used in our work, the so-called f-FACTOR problem.

4.1. The f-Factor problem

DEGREE ANONYMITY has a close connection to the polynomial-time solvable f-FACTOR problem [32, Chapter 10]:

f-FACTOR
Input: A graph $G = (V, E)$ and a function $f : V \to \mathbb{N}_0$.
Question: Is there an *f-factor*, that is, a subgraph $G' = (V, E')$ of G such that $\deg_{G'}(v) = f(v)$ for all $v \in V$?

The f-FACTOR problem can be solved in $O(\sqrt{\sum_{v \in V} f(v)}|E|)$ time [22]. Using f-FACTOR, one can reformulate DEGREE ANONYMITY as follows: Given an instance (G, k, s), the question is whether there is a function $f : V \to \mathbb{N}_0$ such that the complement graph $\overline{G}$ contains an f-factor, $\sum_{v \in V} f(v) \leq 2s$ (every edge is counted twice in the sum of degrees), and for all $v \in V$ it holds that $|\{u \in V \mid \deg_G(u) + f(u) = \deg_G(v) + f(v)\}| \geq k$ (the k-anonymity requirement). As a warm-up, we use this formulation to make the following observation.

Observation 1. *If $k > n/2$, then* DEGREE ANONYMITY *can be solved in $O(n^4)$ time.*

Proof. Observe that if $k > n/2$, then all vertices in the k-anonymous graph have the same degree. Our polynomial-time algorithm is as follows: Branch in the at most n possibilities for the degree $d \geq \Delta$ in the k-anonymous graph. Then compute for each $v \in V$ the value $f(v) = d - \deg_G(v)$. If $1/2 \cdot \sum_{v \in V} f(v) > s$, then return no. Otherwise determine whether there is an f-factor in $\overline{G}$. If there is an f-factor, then return yes (and return the set of edges in the f-factor as solution set), otherwise return no.

As to the running time, observe that we solve at most n f-FACTOR instances. Each instance can be solved in $O(\sqrt{\sum_{v \in V} f(v)}|E|) = O(n^3)$ time. Summing up, the running time is bounded by $O(n^4)$. $\square$

In the above reformulation of Degree Anonymity one looks for an f-factor in the complement graph. Step 3 in Liu and Terzi's heuristic [31] (see Figure 2) can also be formulated as an f-Factor problem in the complement graph: Realizing the k-anonymized degree sequence as super-graph of G is equivalent to finding an f-factor in $\overline{G}$, where $f(v)$ captures the difference between the degree of v in G and the corresponding number in the k-anonymized degree sequence.

As mentioned in the introduction of Section 4, we prove that under certain conditions there exists a realization of the anonymized degree sequence (Step 3). These conditions come from the following lemma guaranteeing the existence of an f-factor.

Lemma 1 (Katerinis and Tsikopoulos [27])**.** *Let $G = (V, E)$ be a graph with minimum vertex degree δ and let $a \leq b$ be two positive integers. Suppose further that*

$$\delta \geq \frac{b}{a+b}|V| \text{ and } |V| > \frac{a+b}{a}(a+b-3).$$

Then, for any function $f\colon V \to \{a, a+1, ..., b\}$ where $\sum_{v \in V} f(v)$ is even, G has an f-factor.

As we are interested in an f-factor in the complement graph of our input graph G, we use Lemma 1 with minimum degree $\delta \geq n - \Delta - 1$, $a = 1$, and $b = \Delta + 2$. Using the next corollary, we will later show that for a minimal k-insertion set S with $|V(S)| > \Delta^2 + 4\Delta + 3$, the maximum degree in $G + S$ is at most $\Delta + 2$ (Lemma 2). This is the reason for setting b to $\Delta + 2$.

Corollary 1. *Let $G = (V, E)$ be a graph with n vertices, minimum degree $n - \Delta - 1$, $\Delta \geq 1$, and let $f\colon V \to \{1, \ldots, \Delta + 2\}$ be a function such that $\sum_{v \in V} f(v)$ is even. If $n \geq \Delta^2 + 4\Delta + 3$, then G has an f-factor.*

Proof. Set $a := 1$ and $b := \Delta + 2$. Since $n \geq \Delta^2 + 4\Delta + 3$ it follows that:

$$\frac{b}{a+b}n = \frac{\Delta+2}{\Delta+3}n \leq n - \Delta - 1.$$

Furthermore,

$$\frac{a+b}{a}(b+a-3) = (\Delta+3)\Delta = \Delta^2 + 3\Delta < \Delta^2 + 4\Delta + 3 = n$$

and, thus, all conditions of Lemma 1 are fulfilled. □

4.2. A polynomial-time algorithm for "large"-solution instances

In this subsection we give an algorithm based on the approach of Liu and Terzi [31] (see Figure 2) that, if a minimum-size k-insertion set S is "large" compared to Δ, solves the given instance in polynomial time (Lemma 5). The key point is to prove that in Step 3 there exists a realization of the anonymized degree sequence, that is, the corresponding f-factor in the complement graph exists (see previous subsection). To this end, we use Corollary 1 and therefore have to ensure that its conditions are fulfilled, namely:

1. The maximum function value is $\Delta + 2$.
2. There are at least $\Delta^2 + 4\Delta + 3$ "affected" vertices, that is, vertices $v \in V$ such that $f(v) > 0$.

In the next lemma we show that a "large" minimum-size k-insertion set increases the maximum degree by at most two implying the first condition. This further implies that if a minimum-size k-insertion set contains more than $(\Delta^2+4\Delta+3)^2$ edges, also the second condition is satisfied.

Lemma 2. *Let $G = (V, E)$ be a graph and let S be a minimum-size k-insertion set. If $|V(S)| \geq \Delta_G^2 + 4\Delta_G + 3$, then the maximum degree in $G + S$ is at most $\Delta_G + 2$.*

Proof. Let G be a graph with maximum degree Δ_G and k be an integer. Let S be a minimum-size edge set such that $G + S$ is k-anonymous and suppose that $|V(S)| \geq \Delta^2 + 4\Delta + 3$. Now assume towards a contradiction that the maximum degree in $G+S$ is at least Δ_G+3. We show that there exists an edge set S' such that $G + S'$ is k-anonymous, $|S'| < |S|$, and $G + S'$ has maximum degree at most $\Delta_G + 2$, contradicting the minimality of S.

First we introduce some notation. Let f be a function $f\colon V \to \mathbb{N}_0$ defined as $f(v) := \deg_{G+S}(v) - \deg_G(v)$ for all $v \in V$. Furthermore, denote with X the set of all vertices having degree more than $\Delta_G + 2$ in $G + S$, that is,

$$X := \{v \in V \mid f(v) + \deg_G(v) \geq \Delta_G + 3\}.$$

Observe that (V, S) is an f-factor of the complement graph $\overline{G}$ and $2|S| = \sum_{v \in V} f(v)$. We now define a new function $f'\colon V \to \mathbb{N}_0$ such that $\overline{G}$ contains an f'-factor denoted by $G' = (V, S')$ such that $G + S'$ is k-anonymous, $|S'| < |S|$, and $G + S'$ has maximum degree at most $\Delta_G + 2$.

We define f' for all $v \in V$ as follows:

$$f'(v) := \begin{cases} f(v) & \text{if } v \notin X, \\ \Delta_G - \deg_G(v) + 1 & \text{if } v \in X \text{ and } f(v) + \deg_G(v) - \Delta_G - 1 \text{ is even,} \\ \Delta_G - \deg_G(v) + 2 & \text{otherwise.} \end{cases}$$

First observe that $\deg_G(v) + f'(v) \leq \Delta_G + 2$ for all $v \in V$. Furthermore, observe that $f'(v) = f(v)$ for all $v \in V \setminus X$ and for all $v \in X$ it holds that $f'(v) < f(v)$ and $f(v) - f'(v)$ is even. Thus, $\sum_{v \in V} f(v) > \sum_{v \in V} f'(v)$ and $\sum_{v \in V} f'(v)$ is even. It remains to show that

(i) $\overline{G}$ contains an f'-factor $G' = (V, S')$ and

(ii) $G + S'$ is k-anonymous.

To prove (i) let $\widetilde{V} := \{v \in V \mid f'(v) > 0\}$. Next, observe that from the definition of X and f' it follows $f(v) > 0$ if and only if $f'(v) > 0$ and hence $\widetilde{V} = V(S)$. Furthermore, let $\widetilde{G} := \overline{G[\widetilde{V}]}$. Observe that $\widetilde{G}$ has minimum degree $|\widetilde{V}| - \Delta_G - 1$ and $|\widetilde{V}| = |V(S)| \geq \Delta^2 + 4\Delta + 3$. Thus, the conditions of Corollary 1 are satisfied and hence $\widetilde{G}$ contains an $f'|_{\widetilde{V}}$-factor $\widetilde{G}' = (\widetilde{V}, S')$. Here, $f'|_{\widetilde{V}}$ denotes f restricted to the domain $\widetilde{V}$. By definition of $\widetilde{V}$ it follows that $G' = (V, S')$ is an f'-factor of $\overline{G}$.

To show (ii), assume towards a contradiction that $G + S'$ is not k-anonymous, that is, there exists some vertex $v \in V$ such that $1 \leq |D_{G+S'}(\deg_{G+S'}(v))| < k$. Let $d := \deg_{G+S}(v)$ and $d' := \deg_{G+S'}(v)$. Observe that $d' = \deg_G(v) + f'(v)$. Thus, if $v \notin X$, then by definition of f' it holds that $d' = \deg_G(v) + f(v) = d \leq \Delta_G + 2$. Hence, for all vertices $u \in D_{G+S}(d')$ it follows that $u \notin X$. Thus, $D_{G+S}(d') \subseteq D_{G+S'}(d')$ and since $G+S$ is k-anonymous we have $|D_{G+S'}(d')| \geq k$, a contradiction. If $v \in X$, that is, $d > \Delta_G + 2$, then $|D_{G+S}(d)| \geq k$ since $G+S$ is k-anonymous. Furthermore, by the definitions of $D_{G+S}(d)$, f, and X we have for all $u \in D_{G+S}(d)$ that $\deg_G(u) + f(u) = d$, $u \in X$, and, thus, $f'(u) + \deg_G(u) = d'$. Therefore, $D_{G+S}(d) \subseteq D_{G+S'}(d')$ and $|D_{G+S'}(d')| \geq k$, a contradiction. $\square$

Note that the bound provided in Lemma 2 is tight: Consider a cycle with $2\ell+1$ vertices plus two additional adjacent vertices with degree one. By setting $k := |V|$ we ensure that the k-anonymized graph is regular. Observe that adding any k-insertion set ends up with a graph of maximum degree at least four.

Next, we formalize the anonymization of degree sequences. A multiset of positive integers $\mathcal{D} = \{d_1, \ldots, d_n\}$, that corresponds to the degrees of all vertices in a graph is called *degree sequence*. A degree sequence $\mathcal{D}$ is k-anonymous if each number in $\mathcal{D}$ occurs at least k times in $\mathcal{D}$. Clearly, the degree sequence of a k-anonymous graph G is k-anonymous. Moreover, if a graph G can be transformed by at most s edge insertions into a k-anonymous graph, then the degree sequence of G can be transformed into a k-anonymous degree sequence by increasing the integers by no more than $2s$ in total (clearly, in the other direction this fails in general because of the graph structure). As we are only interested in a degree sequence corresponding to a graph of a Degree Anonymity instance where s is large, by Lemma 2 we can require the integers in a k-anonymous degree sequence to be upper-bounded by $\Delta + 2$.

k-Degree Sequence Anonymity (k-DSA)
Input: Two positive integers k and s, and a degree sequence $\mathcal{D} = \{d_1, \ldots, d_n\}$ with $d_1 \leq d_2 \leq \ldots \leq d_n$ and $\Delta = d_n$.
Question: Is there a k-anonymous degree sequence $\mathcal{D}' = \{d'_1, \ldots, d'_n\}$ with $d_i \leq d'_i$ and $\max_{1 \leq i \leq n} d'_i \leq \Delta + 2$ such that $\sum_{i=1}^{n} d'_i - d_i = 2s$?

Observe that we require that the "cost" of anonymizing the degree sequence $\mathcal{D}$ is exactly $2s$ and not at most $2s$. This is due to the fact that we only can transfer "large" solutions of k-Degree Sequence Anonymity to Degree Anonymity, as we will show later. In particular, if we allowed the cost of the solution to be at most $2s$, then we could always get "small" solutions to k-Degree Sequence Anonymity, which actually might not be realized in the graph. Note that, due to the degree upper bound of $\Delta + 2$ and the required cost of exactly $2s$, k-Degree Sequence Anonymity is a modified variant compared to the original degree anonymization problem used in Liu and Terzi [31]. Hence, we need to slightly modify their dynamic programming-based approach to prove that k-Degree Sequence Anonymity is polynomial-time solvable.

Lemma 3. k-Degree Sequence Anonymity *can be solved in* $O(nsk\Delta)$ *time.*

Proof. We slightly adapt a dynamic programming algorithm provided by Liu and Terzi [31, Section 4] and Chester et al. [12, Section 6.2.2].

The dynamic programming uses a single table T with a boolean entry $T[i,j]$ for every $i \in \{1,\dots,n\}$ and $j \in \{0,\dots,2s\}$. The entry $T[i,j]$ is `true` if and only if there is a k-anonymous sequence $d'_1,\dots,d'_i$ with $d'_t \geq d_t$ for all $t \in \{1,\dots,i\}$ and the cost $\sum_{t=1}^{i} d'_t - d_t$ of the anonymization is exactly j. Thus, $T[n,2s]$ stores the answer to the k-Degree Sequence Anonymity problem.

Obviously, for $i < k$ we have $T[i,j] :=$ `false` for all j as there is no k-anonymous sequence with less than k numbers. To fill the rest of the table with increasing i, we use for $1 \leq a \leq b \leq n$ and a positive integer d the function $\mathrm{cost}(a,b,d) := \sum_{t=a}^{b} d - d_t$ (the cost of increasing $d_a,\dots,d_b$ up to d).

For $k \leq i < 2k$ we set $T[i,j]$ to `true` if and only if there is a $d \in \{d_i,\dots,\Delta+2\}$ such that $j = \mathrm{cost}(1,i,d)$. We next prove the correctness of this assignment: Clearly, the corresponding sequence $d'_1 = \dots = d'_i = d$ is k-anonymous. In the reverse direction, from $i < 2k$ it follows that $d'_1 = \dots = d'_i$ for each k-anonymous sequence $d'_1,\dots,d'_i$. Hence, the entry $T[i,j]$ is computed correctly in this case.

For $i \geq 2k$ we set $T[i,j]$ to `true` if and only if there are $\ell \in \{k,\dots,2k-1\}$ and $d \in \{d_i,\dots,\Delta+2\}$ such that $T[i-\ell, j-\mathrm{cost}(i-\ell+1,i,d)] =$ `true`. We next prove that this assignment is correct. In the first direction, corresponding to $T[i-\ell, j-\mathrm{cost}(i-\ell+1,i,d)]$ let $d'_1,\dots,d'_{i-\ell}$ be a k-anonymous sequence for $d_1,\dots,d_{i-\ell}$ with anonymization cost $j-\mathrm{cost}(i-\ell+1,i,d)$. Then, since $d \geq d_i$ the sequence

$$d'_1,\dots,d'_{i-\ell},\underbrace{d,\dots,d}_{\ell}$$

is a k-anonymous sequence of cost j for $d_1,\dots,d_i$. In the other direction, let $d'_1,\dots,d'_i$ be a k-anonymous sequence for $d_1,\dots,d_i$ with anonymization cost j. Denote by ℓ the largest integer such that $d'_{i-\ell} = \dots = d'_i$. Since the sequence is k-anonymous ℓ is at least k and if $\ell \geq 2k$, then set $\ell := k$. It follows that the sequence $d'_1,\dots,d'_{i-\ell}$ is k-anonymous and hence $T[i-\ell, j-\mathrm{cost}(i-\ell+1,i,d'_i)] =$ `true`. From this and since $\Delta+2 \geq d'_i \geq d_i$ it follows that the entry $T[i,j]$ is computed correctly.

As each of the recurrences only depends on at most $k \cdot (\Delta+2)$ other entries of the table and the table has $n(2s+1)$ entries, the algorithm runs in $O(nsk\Delta)$ time. It is easy to modify the algorithm to output the appropriate k-anonymous sequence in the same running time. $\square$

We now have all ingredients to solve Degree Anonymity in polynomial time in case it has a "large" minimum-size k-insertion set. The basic process is as follows (see Algorithm 1 for the pseudocode): Given an instance (G,k,s) of Degree Anonymity first compute the degree sequence $\mathcal{D}$ of G. Then, search a "large" solution for $(\mathcal{D},k,s)$, that is a solution of size i, $(\Delta^2+4\Delta+3)^2 \leq i \leq s$.

Algorithm 1 Pseudocode of an algorithm that, given an instance (G, k, s) of Degree Anonymity, either finds a k-insertion set of size at most s for G or decides that the size of a minimum k-insertion set for G is not between $(\Delta^2 + 4\Delta + 3)^2$ and s.

1: **procedure** searchForLarge$k$InsertionSet($G = (V, E)$, k, s)
2: $\quad \mathcal{D} \leftarrow$ degree sequence of G
3: $\quad j \leftarrow -1$
4: $\quad i \leftarrow (\Delta^2 + 4\Delta + 3)^2$
5: $\quad$ **while** $j = -1$ **and** $i \leq s$ **do**
// find minimum j s.t. $(\mathcal{D}, k, j)$ is a yes-instance of k-DSA
6: $\quad\quad$ **if** $(\mathcal{D}, k, i)$ is a yes-instance of k-DSA **then** // see Lemma 3
7: $\quad\quad\quad j \leftarrow i$
8: $\quad\quad\quad \mathcal{D}' \leftarrow$ solution for $(\mathcal{D}, k, i)$
9: $\quad\quad$ **else**
10: $\quad\quad\quad i \leftarrow i + 1$
11: $\quad$ **if** $j = -1$ **then** // no k-insertion set of size between $(\Delta^2 + 4\Delta + 3)^2$ and s
12: $\quad\quad$ **return** 'NO'
13: $\quad$ **else** // (G, k, s) is a yes-instance; the algorithm now computes a solution
14: $\quad\quad$ **for all** $v_i \in V$ **do**
15: $\quad\quad\quad f(v_i) \leftarrow d'_i - \deg_G(v_i)$ // $f(v_i) \mathrel{\widehat{=}}$ number of new incident edges
16: $\quad\quad G' = (V, S) \leftarrow f$-factor of $\overline{G}$
17: $\quad\quad$ **return** S

If there is such a large solution for the k-Degree Sequence Anonymity instance, then the next lemma states that this solution can be transferred to the Degree Anonymity instance.

Lemma 4. *Let (G, k, s) be an instance of* Degree Anonymity. *If the size of a minimum-size k-insertion set is at least $(\Delta^2 + 4\Delta + 3)^2$, then Algorithm 1 decides (G, k, s) in polynomial time. Furthermore, if Algorithm 1 returns an edge set S, then S is a k-insertion set of size $|S| \leq s$.*

Proof. We first show that if Algorithm 1 returns an edge set S, then S is a k-insertion set. Let S be an edge insertion set returned by the algorithm. First, observe that $|S| \leq s$ due to the while loop in Line 5. Since in Line 16 the algorithm determines an f-factor in $\overline{G}$ it follows that $S \cap E = \emptyset$. Furthermore, $(\mathcal{D}, k, |S|)$ is a yes-instance. Thus, by construction of f, it follows that $G + S$ is k-anonymous. Putting all this together implies that S is a k-insertion set of size $|S| \leq s$ and, hence, (G, k, s) is a yes-instance.

Now, let S be the minimum k-insertion set of size $|S| \geq (\Delta^2 + 4\Delta + 3)^2$ and $|S| \leq s$. We show that Algorithm 1 returns a k-insertion set. Observe that for any edge set S of size at least $(\Delta^2 + 4\Delta + 3)^2$ it holds that $|V(S)| > \sqrt{(\Delta^2 + 4\Delta + 3)^2} = \Delta^2 + 4\Delta + 3$. Thus, by Lemma 2, since S is minimum, $G + S$ has maximum degree $\Delta + 2$. Let $\mathcal{D}$ be the degree sequence of G. As already discussed before, the degree sequence $\mathcal{D}'$ of $G + S$ is a solution for k-Degree Sequence Anonymity. Thus, $(\mathcal{D}, k, |S|)$ is a yes-instance of k-Degree Sequence Anonymity. Hence, after leaving the while-loop in Line 5 it holds that $j \leq |S|$ and $\mathcal{D}'$ is the corresponding k-anonymous degree sequence. By definition, $\mathcal{D}'$ has a maximum degree of at most $\Delta + 2$. Hence, there are at least $(\Delta^2 + 4\Delta + 3)^2/(\Delta + 2) > \Delta^2 + 4\Delta + 3$ integers in $\mathcal{D}$ that have been increased to get $\mathcal{D}'$. Thus, for the function f computed in Line 15 it holds that $|\{v \in V \mid f(v) > 1\}| > \Delta^2+4\Delta+3$. Since $\overline{G}$ has minimum degree $|V|-\Delta-1$, if follows from Corollary 1 that $\overline{G}$ contains an f-factor. Thus, in Line 16 an f-factor $G' = (V, S)$ is found and the algorithm returns a k-insertion set. $\square$

Recall that f-Factor can be solved in $O(\sqrt{\sum_{v \in V} f(v)}|E|)$ time [22]. Together with Lemma 3, this implies that Algorithm 1 runs in polynomial time. Hence, Lemma 4 essentially shows that Degree Anonymity can be decided in polynomial time when a minimum-size k-insertion sets is large. If a minimum-size k-insertion set is not large, then, since any k-insertion set for G of size $j \leq s$ directly implies that $(\mathcal{D}, k, j)$ is a yes-instance for k-Degree Sequence Anonymity, it follows that we can bound the parameter s by a function in Δ, as stated in the next lemma stating the mentioned win-win situation.

Lemma 5. *There is an algorithm running in $O(ns^2k\Delta)$ time that given an instance (G, k, s) of* Degree Anonymity *returns 'YES' or 'NO'. If it answers 'YES', then (G, k, s) is a yes-instance. If it returns 'NO', then (G, k, s) is a yes-instance if and only if $(G, k, \min\{(\Delta^2 + 4\Delta + 3)^2, s\})$ is a yes-instance.*

Proof. The algorithm is obtained by replacing the Lines 14 to 17 of Algorithm 1 with "return 'YES' ". If the algorithm returns 'YES', then, by Lemma 4, the input instance (G, k, s) is a yes-instance. If the algorithm returns 'NO', then consider the following two cases. Let S be a minimum-size k-insertion set.

Case 1 $|S| \geq (\Delta^2 + 4\Delta + 3)^2$: As the algorithm returns 'NO', it follows from Lemma 4 that the given instance (G, k, s) is a no-instance (thus $s < |S|$). Hence, also $(G, k, \min\{(\Delta^2 + 4\Delta + 3)^2, s\})$ is a no-instance.

Case 2 $|S| < (\Delta^2 + 4\Delta + 3)^2$: If $s < |S|$, then $s < (\Delta^2 + 4\Delta + 3)^2$ and, thus, (G, k, s) as well as $(G, k, \min\{(\Delta^2 + 4\Delta + 3)^2, s\})$ are no-instances. Conversely, if $s > |S|$, then $\min\{(\Delta^2 + 4\Delta + 3)^2, s\} > |S|$ and, hence, (G, k, s) as well as $(G, k, \min\{(\Delta^2 + 4\Delta + 3)^2, s\})$ are yes-instances. Hence, it holds that (G, k, s) is a yes-instance if and only if $(G, k, \min\{(\Delta^2 + 4\Delta + 3)^2, s\})$ is a yes-instance.

As to the running time, observe that the algorithm runs in $O(ns^2k\Delta)$ time: The algorithm basically solves at most s instances of k-Degree Sequence Anonymity which requires $O(ns^2k\Delta)$ time, see Lemma 3, and then returns 'YES' or 'NO'. $\square$

We remark that Algorithm 1, constructing a solution if found, runs in $O(n^3 + ns^2k\Delta)$ time: The first part of deciding whether there exists a large solution runs in $O(ns^2k\Delta)$, see Lemma 5. Then, computing an f-factor in $\overline{G}$ is doable in $O(\sqrt{\sum_{v \in V} f(v)}|E|)$ time [22], that is, $O(n^2\sqrt{n^2}) = O(n^3)$ time.

4.3. Polynomial kernel

In this subsection we first show a kernel with respect to the combined parameter (Δ, s) and then use Lemma 5 to show that this kernel is polynomial only in Δ. Our kernelization algorithm is based on the following observation. For a given graph G, consider for some $1 \leq i \leq \Delta$ the block $D_G(i)$, that is, the set of all vertices of degree i. If $D_G(i)$ contains many vertices, then the vertices are "interchangeable":

Observation 2. *Let (G, k, s) with graph $G = (V, E)$ be an instance, let S be a k-insertion set for G with $|S| \leq s$, and let $v \in V(S) \cap D_G(i)$ be a vertex such that $|D_G(i)| > (\Delta + 2)s$. Then there exists a vertex $u \in D_G(i) \setminus V(S)$ such that replacing in S every edge $\{v, w\}$ by $\{u, w\}$ results in a k-insertion set for G.*

Proof. Since $|S| \leq s$, the vertex v can be incident to at most s edges in S. Denoting the set of these edges by S^v, one obviously can replace v by $u \in D_G(i)$ if u is non-adjacent to all vertices in $V(S^v) \setminus \{v\}$ (this allows to insert all edges) and $u \notin V(S)$ (no block in $G + S$ does change its size). However, as $V(S)$ contains at most $2s$ vertices from $D_G(i)$ and each of the at most s vertices in $V(S^v) \setminus \{v\}$ has at most Δ neighbors in G, it follows that such a vertex $u \in D_G(i)$ exists if $|D_G(i)| > (\Delta + 2)s$. □

By Observation 2, in our kernel we only need to keep at most $(\Delta+2)s$ vertices in each block: If in an optimal k-insertion set S there is a vertex $v \in V(S)$ that we did not keep, then by Observation 2 we can replace v by some vertex we kept. There are two major problems that need to be fixed to obtain a kernel: First, when removing vertices from the graph, the degrees of the remaining vertices change. Second, k might be "large" and, thus, removing vertices (during kernelization) in one block may breach the k-anonymity constraint. To overcome the first problem we insert some "dummy-vertices" which are guaranteed not to be contained in any k-insertion set. To solve the second problem, however, we need to adjust the parameter k as well as the number of vertices that we keep from each block.

Details of the Kernelization Algorithm. We now explain the kernelization algorithm in detail (see Algorithm 2 for the pseudocode). Let (G, s, k) be an instance of Degree Anonymity. For brevity we set $\beta := (\Delta + 4)s + 1$. We compute in polynomial time an equivalent instance (G', k', s) with at most $O(\Delta^3 s)$ vertices: First set $k' := \min\{k, \beta\}$ (Line 4). We arbitrarily select from each block $D_G(i)$ a certain number x of vertices and collect all these vertices into the set A (Line 14). To cope with the above mentioned second problem, the "certain number" is defined in a case distinction on the value of k (see Lines 5 to 14). Intuitively, if k is large then we distinguish between "small" blocks of size at most $2s$ and "large" blocks of size at least $k - 2s$. Obviously, if there is a block which is neither small nor large, then the instance is a no-instance (see Line 7). Thus, in the kernel we keep for small blocks the "distance to size zero" and for large blocks the "distance to size k". Furthermore, in order to distinguish between small and large blocks it is sufficient that $k' > 4s$. However, to guarantee that Observation 2

Algorithm 2 The pseudocode or the algorithm computing a polynomial kernel with respect to (Δ, s).

1: **procedure** PRODUCEPOLYKERNEL($G = (V, E)$, k, s)
2:     **if** $|V| \leq \Delta(\beta + 4s)$ **then** // β is defined as $\beta := (\Delta + 4)s + 1$
3:         **return** (G, k, s)
4:     $k' \leftarrow \min\{k, \beta\}$; $A \leftarrow \emptyset$
5:     **for** $i \leftarrow 1$ **to** Δ **do**
6:         **if** $2s < |D_G(i)| < k - 2s$ **then**
7:             **return** trivial no-instance // insufficient budget for $D_G(i)$
8:         **if** $k \leq \beta$ **then** // determine retained vertices
9:             $x \leftarrow \min\{|D_G(i)|, \beta + 4s\}$ // keep at most $\beta + 4$ vertices
10:         **else if** $|D_G(i)| \leq 2s$ **then** // "small" block
11:             $x \leftarrow |D_G(i)|$ // keep all vertices ("distance to size zero")
12:         **else** // "large" block and $k' = \beta$
13:             $x \leftarrow k' + \min\{4s, (|D_G(i)| - k)\}$ // keep "distance to size k".
14:         add x vertices from $D_G(i)$ to A
15:     $G' := G[A]$
16:     **for each** $v \in A$ **do** // add vertices to preserve degree of retained vertices
17:         add to G' $\deg_G(v) - \deg_{G'}(v)$ many degree-one vertices adjacent to v
18:     denote with P the set of vertices added in Line 17
19:     by adding matched pairs of vertices, ensure that $|P| \geq \max\{4\Delta + 4s + 4, k'\}$
20:     **if** $\Delta + s + 1$ is even **then**
21:         $G^F = (P, E^F) \leftarrow (\Delta + s + 1)$-factor in $\overline{G'[P]}$
22:     **else**
23:         $G^F = (P, E^F) \leftarrow (\Delta + s + 2)$-factor in $\overline{G'[P]}$
24:     $G' \leftarrow G' + E^F$
25:     **return** (G', k', s)

is applicable, the case distinction is a little bit more complicated, see Lines 5 to 14. The idea is to take enough vertices from each block into A such that we can guarantee that any solution on G can be transformed to G' and vice versa. Intuitively, for this it is enough to select $2s$ vertices from each block, as no solution can "affect" more vertices.

In Line 15 we start building G' by first copying $G[A]$ into it. Next, adding a pendant vertex to v means that we add a new vertex to G' and make it adjacent to v. For each $v \in A$ we add pendant vertices to v to ensure that $\deg_{G'}(v) = \deg_G(v)$ (Line 17). The vertices of A stay untouched in the following. Denote the set of all pendant vertices by P. Next, we add enough pairwise adjacent vertices to P to ensure that $|P| \geq \max\{k', 4\Delta + 4s + 4\}$ (Line 19). Hence, $|P| \leq \max\{|A| \cdot \Delta, k', 4\Delta + 4s + 4\} + 1$. To avoid that vertices in P help to anonymize the vertices in A we "shift" the degree of the vertices in P (see Lines 20 to 24): We add edges between the vertices in P to ensure that the degree of all vertices in P is $\Delta + s + 2$ (when $\Delta + s + 1$ is even) or $\Delta + s + 3$ (when $\Delta + s + 2$ is even). For the ease of notation let χ denote the new degree of the vertices in P. Observe that before adding edges all vertices in P have degree one in G'. Thus, the minimum degree in $\overline{G'[P]}$ is $|P| - 2$. Furthermore, for each $v \in P$ we denote by $f(v)$ the number of incident edges v requires to have the described degree. It follows that $f(v)$ is even and hence $\sum_{v \in P} f(v)$ is even. Hence setting $a = b := \chi$ fulfills all conditions of Lemma 1. Thus, the required f-factor exists and can be found in $O(|P|^2 \sqrt{|P|(\Delta + s)})$ time [22]. This completes the description of the kernelization algorithm.

The key point of the correctness of the kernelization is to show that without loss of generality, no k-insertion set S for G' of size $|S| \leq s$ affects any vertex in P. This is ensured by "shifting" the degree of all vertices in P by $s + 1$ (or $s + 2$), implying that none of the vertices in A can "reach" the degree of any vertex in P by adding at most s edges. Hence each block either is a subset of A or of P. We now prove that we may assume that an edge insertion set does not affect any vertex in P. All what we need to prove this is the fact that A contains at least $\beta + 4s$ vertices from at least one block in G. Observe that this is ensured by the condition in Line 2.

Lemma 6. *Let (G, k, s) be an instance of* Degree Anonymity *and let (G', k', s) be the instance computed by Algorithm 2. If there is a k-insertion set S for G' with $|S| \leq s$, then there is also a k-insertion set S' for G' with $|S'| = |S|$ such that $V(S') \cap P = \emptyset$.*

Before proving Lemma 6, we introduce the term "co-matching" and prove an observation concerning its existence. A graph $G = (V, E)$ contains a *co-matching* of size ℓ if the complement graph $\overline{G}$ contains a matching of size ℓ, that is, a subset of ℓ non-overlapping edges of $\overline{G}$. A *perfect* co-matching of G is a co-matching of size $|V|/2$. We prove the following observation that shows sufficient conditions for the existence of co-matchings.

Observation 3. *Let $G = (V, E)$ be a graph and let $V' \subseteq V$ be a vertex subset such that $|V'| \geq 2\Delta + 1$ and $|V'|$ is even. Then, $G[V']$ contains a perfect co-matching.*

Proof. Since $|V'| \geq 2\Delta + 1$, it follows that in $\overline{G[V']}$ every vertex has degree at least $|V'| - \Delta \geq |V'|/2$. Hence, using Dirac's Theorem [15], it follows that $\overline{G[V']}$ contains a Hamiltonian cycle C. If $|V'|$ is even, then taking every second edge of C results in a perfect matching. $\square$

We now can prove Lemma 6.

Proof of Lemma 6. Let S be a k-insertion set S for G' with $|S| \leq s$ and $V(S) \cap P \neq \emptyset$. As each block in $G' + S$ is either a subset of A or of P, it follows from $V(S) \cap P \neq \emptyset$ that $|V(S) \cap P| \geq k$. Additionally, as S can affect at most $2s$ vertices and A contains at least $\beta + 4s$ vertices from at least one block, say $D_G(i)$. It follows that block $D_{G'+S}(i)$ contains at least $\beta + 2s$ unaffected vertices.

We next restructure S in order to get a k-insertion set fulfilling the claimed properties. For this, one has to exchange all edges in S containing at least one endpoint from P. We start with those edges in S having only one endpoint in P.

Let $A^P \subseteq V(S) \cap A$ be all vertices in A that are incident to some edge in S with the second endpoint in P. For each $v \in A^P$ we select $|(N_{G'+S}(v) \setminus N_{G'}(v)) \cap P|$ vertices among the unaffected vertices from $D_{G'+S}(i)$ and replace each edge in S from v to some vertex in P (there are exactly $|(N_{G'+S}(v) \setminus N_{G'}(v)) \cap P|$ many) by an edge from v to one of the selected vertices (each unaffected vertex in $D_{G'+S}(i)$ is only used once). Note that this is always possible since each vertex v has at most Δ neighbors among the unaffected vertices in $D_{G'+S}(i)$, since there are at least $s + \Delta + 1$ unaffected vertices in $D_{G'+S}(i)$, and since there can be at most s edges in S that are replaced in this way.

Note that, after having exchanged all edges in S with one endpoint in P, $D_{G'+S}(i)$ contains still at least $\beta + 2s > 2\Delta + 2s$ unaffected vertices. Thus, by Observation 3, there exists a co-matching of size exactly $|S| \leq s$ among the unaffected vertices in $D_{G'+S}(i)$. Exchanging each edge in S with two endpoints in P by an edge in this matching yields the following: All vertices in P are unaffected. Hence, the block containing all vertices from P is of size at least k. Additionally, we increased for at least k vertices the degree from i to $i + 1$, thus $|D_{G'+S}(i + 1)| \geq k$. As the block $D_{G'+S}(i)$ still contains at least k vertices after restructuring, it follows that $G' + S$ is k-anonymous. □

Based on Lemma 6 we now prove the correctness of our kernelization algorithm.

Lemma 7. *If the instance (G', k', s) constructed by Algorithm 2 is a yes-instance, then (G, k, s) is a yes-instance.*

Proof. First, observe that if $k \leq \beta$, then $k' = k$ and each edge insertion set that makes G' k-anonymous also makes G k-anonymous as all blocks with less than $\beta + 4s$ vertices remain unchanged. Hence, assume that $k > \beta$ and, thus, $k' = \beta < k$.

Let S' be an edge insertion set with $|S'| \leq s$ such that $G' + S'$ is k-anonymous and $S' \cap P = \emptyset$ (see Lemma 6). To prove that $G + S'$ is also k-anonymous, assume towards a contradiction that there is a block $D_{G+S'}(j)$ with $0 < |D_{G+S'}(j)| < k$. We associate two numbers $d_i^G(j), d_o^G(j)$ to S' with respect to G where $d_i^G(j)$ is the number of vertices in $D_{G+S'}(j)$ but not in $D_G(j)$ and $d_o^G(j)$ is the number of vertices in $D_G(j)$ but not in $D_{G+S'}(j)$. Defining the numbers analogously for G', it holds that $d_i^G(j) = d_i^{G'}(j)$ and $d_o^G(j) = d_o^{G'}(j)$.

If $|D_{G'+S'}(j)| = 0$, then $d_o^{G'}(j) = |D_{G'}(j)| \leq 2s$ and $d_i^{G'}(j) = 0$. By Line 11 this implies $D_{G+S'}(j) = \emptyset$. Consider the remaining case, that is, $|D_{G'+S'}(j)| \geq k'$. If $|D_G(j)| \geq k + 2s$, then $|D_{G+S'}|(j) \geq k$. Otherwise $|D_{G'}(j)| = k' + |D_G(j)| - k$ by Line 13. But then we have

$$\begin{aligned} 0 \leq |D_{G'+S'}(j)| - k' &= |D_{G'}(j)| + d_i^{G'}(j) - d_o^{G'}(j) - k' = \\ &= |D_G(j)| + d_i^G(j) - d_o^G(j) - k = |D_{G+S'}(j)| - k. \end{aligned}$$

and, hence, $|D_{G+S'}(j)| \geq k$. □

Lemma 8. *If (G, k, s) is a yes-instance, then the instance (G', k', s) constructed by Algorithm 2 is a yes-instance.*

Proof. Recall that $k' = \min\{k, \beta\} = \min\{k, (\Delta+4)s+1\}$. Let S be a k-insertion set for G of size at most s. We now show how to construct a k'-insertion set S' for G' of size at most s. If $V(S) \setminus A \neq \emptyset$, then we do the following to ensure $V(S) \subseteq A$. We initialize $S^1 := S$. Observe that for each vertex $v \in V(S) \setminus A$ it holds that $|D_G(\deg_G(v)) \cap A| \geq \beta - 2s > (\Delta + 2)s$. Hence, by Observation 2, there exists a vertex $u \in D_G(\deg_G(v)) \cap A$ such that the set S^2 resulting from S^1 by replacing v with u, formally, $S^2 := S^1 \cup \{\{u, w\} \mid \{v, w\} \in S^1\} \setminus \{\{v, w\} \mid \{v, w\} \in S^1\}$, is also a k-insertion set for G. Note that $V(S^2)$ has larger overlap with A as $V(S^1)$, more precisely, $|V(S^2) \cap A| = |V(S^1) \cap A| + 1$. By iteratively applying this procedure we end up with a k-insertion set S' for G with $V(S') \subseteq A$.

We next show that $G'+S'$ is k'-anonymous. Observe that if $k \leq \beta$, then $k = k'$ and all blocks in G' with less than $\beta + 4s > k + 2s$ vertices remained unchanged during the kernelization (see Line 9). Hence, all these blocks fulfill the k-anonymity requirement in $G' + S'$. Furthermore, all blocks with more than $k+2s$ vertices in G also contain more than $k + 2s$ vertices in G' and more than k vertices in $G' + S'$. Thus, $G' + S'$ is k'-anonymous.

Now assume that $k > \beta$ and, thus, $k' = \beta$. Assume towards a contradiction that there is a block with $0 < |D_{G'+S'}(i)| < k'$. Observe that if $|D_{G'}(i)| \leq 2s$, then also $|D_G(i)| \leq 2s$, thus $D_{G'}(i) = D_G(i)$ (see Line 11) and $D_{G'+S'}(i) = D_{G+S'}(i)$, a contradiction to the assumption that $G + S'$ is k-anonymous. Hence, consider the case $|D_{G'}(i)| \geq 2s$ and, thus, $|D_G(i)| \geq k - 2s$ and $|D_{G'}(i)| = \beta + \min\{4s, (|D_G(i)| - k)\}$ (see Line 13). Observe that $|D_{G'+S'}(i)| - |D_{G'}(i)| = |D_{G+S'}(i)| - |D_G(i)|$ and, thus,

$$|D_{G'+S'}(i)| = (|D_{G+S'}(i)| - |D_G(i)|) + |D_{G'}(i)|. \tag{1}$$

Furthermore, observe that $|D_{G+S'}(i)| - |D_G(i)| \geq -2s$ and $|D_{G+S'}(i)| \geq k$. We now distinguish the two cases $|D_G(i)| - k \geq 4s$ and $|D_G(i)| - k < 4s$. In the first case it follows that $|D_{G'}(i)| = \beta + 4s$ and, hence, from Equation (1) it follows

$$|D_{G'+S'}(i)| \geq -2s + \beta + 4s > k',$$

a contradiction. In the second case it follows that $|D_{G'}(i)| = \beta + |D_G(i)| - k$ (see Line 13), and from Equation (1) we conclude that

$$|D_{G'+S'}(i)| \geq k - |D_G(i)| + \beta + |D_G(i)| - k = \beta = k',$$

a contradiction. □

From Lemma 8 and Lemma 7 it follows that the kernelization algorithm is correct. It is not hard to see that the size of the computed instances is bounded by a polynomial in Δ and s, leading to the following.

Theorem 3. Degree Anonymity *admits a kernel with $O(\Delta^3 s)$ vertices. The kernelization runs in $O(\Delta^8 s^3 + \Delta^2 sn)$ time.*

Proof. The kernel is computed by Algorithm 2. The correctness of the kernelization algorithm follows from Lemma 8 and Lemma 7. Observe that each block in A

has size at most $\beta + 4s$ (see Lines 9, 11 and 13). Thus, $|A| = O(\Delta\beta) = O(\Delta^2 s)$. Furthermore, the set P contains at most $\max\{\Delta|A|, k', 4s + 4\Delta + 1\}$ vertices (see Lines 17 to 19). Thus, $|P| = O(\Delta^3 s)$ and, hence, the reduced instance contains $O(\Delta^3 s)$ vertices.

It remains to show the running time. To this end, using bucket sort, one can sort the n vertices by degree in $O(n)$ time. Furthermore, in the same time one can create Δ lists—each list containing the vertices of some degree i, $1 \leq i \leq \Delta$. Then, the selection of the $O(\Delta^2 s)$ vertices of A can be done in $O(\Delta^2 sn)$ time. Clearly, adding the vertices in P can be done in $O(\Delta^3 s)$ time. Finally, as P contains $O(\Delta^3 s)$ vertices and an $\Delta + s + 1$-factor in $\overline{G[P]}$ can be found in $O(|P|^2 \sqrt{|P|(\Delta + s)})$ time [22], Algorithm 2 runs in $O(\Delta^6 s^2 \sqrt{\Delta^3 s(\Delta + s)} + \Delta^2 sn) = O(\Delta^8 s^3 + \Delta^2 sn)$ time. □

By Lemma 5 it follows that in $O(ns^2 k\Delta)$ time we can either decide the instance or we have $s \leq (\Delta^2 + 4\Delta + 3)^2$. By Theorem 3 this implies our main result—a polynomial kernel with respect to the maximum degree.

Theorem 4. Degree Anonymity *admits an $O(\Delta^7)$-vertex kernel. The kernelization runs in $O(\Delta^8 s^3 + (sk + \Delta)\Delta sn)$ time.*

5. Fixed-Parameter Algorithm for the Parameter Maximum Degree

Theorem 4 already implies that Degree Anonymity is fixed-parameter-tractable with respect to the parameter maximum degree. In this section, however, we provide a faster, direct combinatorial algorithm for the combined parameter (Δ, s) and, by Lemma 5, also for the parameter Δ.

Roughly speaking, for fixed k-insertion set S the algorithm branches into all suitable structures of $G[S]$, that is, graphs of at most $2s$ vertices with vertex labels from $\{1, \ldots, \Delta\}$. Then the algorithm checks whether the respective structure occurs as a subgraph in $\overline{G}$ such that the labels on the vertices match the degree of the corresponding vertex in G.

Theorem 5. Degree Anonymity *can be solved in $s(6s^2\Delta^2)^{2s} \cdot n^{O(1)}$ time.*

Proof. Let (G, k, s) be an instance of Degree Anonymity. Let S be a k-insertion set S of size at most s and consider the graph $G[S]$ that is induced by the edges in S. Clearly, $G[S]$ contains at most $2s$ vertices and we label each vertex with its initial degree (some vertices might have the same label). Roughly speaking, we branch into all possibilities for the structure (label of vertices and which "labels" are connected by an edge) of the graph $G[S]$ and then try to find the structure as a subgraph in $\overline{G}$.

More specifically, we first branch into all possibilities to first choose the right number of edges and vertices in $G[S]$. We then branch into all possibilities to choose for each vertex its label, that is, its degree in G. Note that there are at most Δ^{2s} possibilities. Finally, we branch into the at most $\binom{2s}{2}^s \leq 4^s s^{2s}$ possibilities to choose pairs of vertices that are connected by an edge from S. Denote the guessed graph by G^S. Clearly, if G^S corresponds to $G[S]$, then $\overline{G}$ contains G^S.

We now give an algorithm that finds the subgraph G^S in $\overline{G}$ if it exists. First, note that there are at most $2s$ vertices in G^S and each of them has degree at most Δ in G. Hence, if a block $D_G(d)$ has size at least $(2s-1)\Delta + 2s$, then it is always possible to choose a vertex from $D_G(d)$ that is non-adjacent to all vertices in a size-at-most-$(2s-1)$ vertex subset where at most s edges have been added. Thus we first can ignore vertices in G^s labeled with d where $|D_G(d)| \geq 3s\Delta$. For all other vertices we branch again into the at most $\binom{3s\Delta}{2s} \leq (3s\Delta)^{2s}$ possibilities to choose them from the "small" blocks. Afterwards we greedily add the required vertices from the blocks of size at least $3s\Delta$ such that they are non-adjacent to the vertices chosen before. As this can be done in polynomial time, the algorithm runs overall in $s \cdot \Delta^{2s} \cdot 4^s s^{2s} \cdot (3s\Delta)^{2s} \cdot n^{O(1)} = s(6s^2\Delta^2)^{2s} \cdot n^{O(1)}$ time. The correctness of the algorithm follows from the exhaustive search. $\square$

Note that due to the upper bound $s < (\Delta^2 + 4\Delta + 3)^2$ (see Lemma 5) and the polynomial kernel for the parameter Δ (see Theorem 4), Theorem 5 also provides the following.

Corollary 2. Degree Anonymity *can be solved in* $\Delta^{O(\Delta^4)} + n^{O(1)}$ *time.*

6. Conclusion

One of the grand challenges of theoretical research on computationally hard problems is to gain a better understanding of when and why heuristic algorithms work [26]. In this theoretical study, we contributed to a better theoretical understanding of a basic problem in graph anonymization, on the one side partially explaining the quality of a successful heuristic approach [31] and on the other side providing a first step towards a provably efficient algorithm for relevant special cases (bounded-degree graphs). Our work just being one of the first steps in the so far underdeveloped field of studying the computational complexity of graph anonymization [12], there are numerous challenges for future research. First, our focus was on classification results rather than engineering the upper bounds, a natural next step to do. Notably, some algorithm engineering efforts based on our theoretical work showed recently some promising (partially heuristic) results [25]. Second, it would be interesting to perform a data-driven analysis of parameter values on real-world networks in order to gain parameterizations that can be exploited in a broad-band multivariate complexity analysis [19, 28, 38] of Degree Anonymity. Finally, with Degree Anonymity we focused on a very basic problem of graph anonymization; there are numerous other models (partially mentioned in the introductory section) that ask for similar studies.

References

[1] C. C. Aggarwal, Y. Li, and P. S. Yu. On the hardness of graph anonymization. In *Proceedings of the 11th IEEE International Conference on Data Mining (ICDM '11)*, pages 1002–1007. IEEE, 2011.

[2] G. Aggarwal, T. Feder, K. Kenthapadi, S. Khuller, R. Panigrahy, D. Thomas, and A. Zhu. Achieving anonymity via clustering. *ACM Transactions on Algorithms*, 6(3):1–19, 2010.

[3] C. Bazgan and A. Nichterlein. Parameterized inapproximability of degree anonymization. In *Proceedings of the 9th International Symposium on Parameterized and Exact Computation (IPEC '14)*, LNCS. Springer, 2014. In press.

[4] H. L. Bodlaender. Kernelization: New upper and lower bound techniques. In *Proceedings of the 4th International Workshop on Parameterized and Exact Computation (IWPEC '09)*, volume 5917 of *LNCS*, pages 17–37. Springer, 2009.

[5] P. Bonizzoni, G. Della Vedova, and R. Dondi. Anonymizing binary and small tables is hard to approximate. *Journal of Combinatorial Optimization*, 22(1):97–119, 2011.

[6] P. Bonizzoni, G. Della Vedova, R. Dondi, and Y. Pirola. Parameterized complexity of k-anonymity: hardness and tractability. *Journal of Combinatorial Optimization*, 26(1):19–43, 2013.

[7] R. Bredereck, S. Hartung, A. Nichterlein, and G. J. Woeginger. The complexity of finding a large subgraph under anonymity constraints. In *Proceedings of the 24th International Symposium on Algorithms and Computation (ISAAC '13)*, volume 8283 of *LNCS*, pages 152–162. Springer, 2013.

[8] R. Bredereck, V. Froese, S. Hartung, A. Nichterlein, R. Niedermeier, and N. Talmon. The complexity of degree anonymization by vertex addition. In *Proceedings of the International Conference on Algorithmic Aspects of Information and Management (AAIM '14)*, volume 8546 of *LNCS*, pages 44–55. Springer, 2014.

[9] R. Bredereck, A. Nichterlein, R. Niedermeier, and G. Philip. The effect of homogeneity on the computational complexity of combinatorial data anonymization. *Data Mining and Knowledge Discovery*, 28(1):65–91, 2014.

[10] J. Casas-Roma, J. Herrera-Joancomartí, and V. Torra. An algorithm for k-degree anonymity on large networks. In *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining (ASONAM '13)*, pages 671–675. ACM Press, 2013.

[11] S. Chester, J. Gaertner, U. Stege, and S. Venkatesh. Anonymizing subsets of social networks with degree constrained subgraphs. In *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining (ASONAM '12)*, pages 418–422. IEEE Computer Society, 2012.

[12] S. Chester, B. Kapron, G. Srivastava, and S. Venkatesh. Complexity of social network anonymization. *Social Network Analysis and Mining*, 3(2): 151–166, 2013.

[13] 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. 3(3):381–399, 2013.

[14] 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.

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

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

[17] P. A. Evans, T. Wareham, and R. Chaytor. Fixed-parameter tractability of anonymizing data by suppressing entries. *Journal of Combinatorial Optimization*, 18(4):362–375, 2009.

[18] M. R. Fellows, D. Hermelin, F. A. Rosamond, and S. Vialette. On the parameterized complexity of multiple-interval graph problems. *Theoretical Computer Science*, 410(1):53–61, 2009.

[19] M. R. Fellows, B. M. P. Jansen, and F. A. Rosamond. Towards fully multivariate algorithmics: Parameter ecology and the deconstruction of computational complexity. *European Journal of Combinatorics*, 34(3):541–566, 2013.

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

[21] B. C. M. Fung, K. Wang, R. Chen, and P. S. Yu. Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys*, 42 (4):14:1–14:53, 2010.

[22] H. N. Gabow. An efficient reduction technique for degree-constrained subgraph and bidirected network flow problems. In *Proceedings of the 15th Annual ACM Symposium on Theory of Computing (STOC '83)*, pages 448–456. ACM, 1983.

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

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

[25] S. Hartung, C. Hoffmann, and A. Nichterlein. Improved upper and lower bound heuristics for degree anonymization in social networks. In *Proceedings of the 13th International Symposium on Experimental Algorithms (SEA '14)*, volume 8504 of *LNCS*, pages 376–387. Springer, 2014.

[26] R. M. Karp. Heuristic algorithms in computational molecular biology. *Journal of Computer and System Sciences*, 77(1):122–128, 2011.

[27] P. Katerinis and N. Tsikopoulos. Minimum degree and f-factors in graphs. *New Zealand Journal of Mathematics*, 29(1):33–40, 2000.

[28] C. Komusiewicz and R. Niedermeier. New races in parameterized algorithmics. In *Proceedings of the 37th International Symposium on Mathematical Foundations of Computer Science (MFCS '12)*, volume 7464 of *LNCS*, pages 19–30. Springer, 2012.

[29] S. Kratsch. Recent developments in kernelization: A survey. *Bulletin of EATCS*, 113:58–97, 2014.

[30] J. Leskovec and E. Horvitz. Planetary-scale views on a large instant-messaging network. In *Proceedings of the 17th International World Wide Web Conference (WWW '08)*, pages 915–924. ACM, 2008.

[31] K. Liu and E. Terzi. Towards identity anonymization on graphs. In *Proceedings of the ACM SIGMOD International Conference on Management*

of Data (SIGMOD '08), pages 93–106. ACM, 2008.
[32] L. Lovász and M. D. Plummer. *Matching Theory*, volume 29 of *Annals of Discrete Mathematics*. North-Holland, 1986.
[33] X. Lu, Y. Song, and S. Bressan. Fast identity anonymization on graphs. In *Proceedings of the 23rd International Conference on Database and Expert Systems Applications (DEXA 2012), Part I*, volume 7446 of *LNCS*, pages 281–295. Springer, 2012.
[34] L. Mathieson and S. Szeider. Editing graphs to satisfy degree constraints: A parameterized approach. *Journal of Computer and System Sciences*, 78 (1):179–191, 2012.
[35] H. Moser and D. M. Thilikos. Parameterized complexity of finding regular induced subgraphs. *Journal of Discrete Algorithms*, 7(2):181–190, 2009.
[36] A. Narayanan and V. Shmatikov. De-anonymizing social networks. In *Proceedings of the 30th IEEE Symposium on Security and Privacy (SP '09)*, pages 173–187. IEEE, 2009.
[37] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms.* Oxford University Press, 2006.
[38] R. Niedermeier. Reflections on multivariate algorithmics and problem parameterization. In *Proceedings of the 27th International Symposium on Theoretical Aspects of Computer Science (STACS '10)*, volume 5 of *LIPIcs*, pages 17–32. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2010.
[39] C. Phillips and T. J. Warnow. The asymmetric median tree—a new model for building consensus trees. *Discrete Applied Mathematics*, 71(1–3):311–335, 1996.
[40] A. Sala, X. Zhao, C. Wilson, H. Zheng, and B. Y. Zhao. Sharing graphs using differentially private graph models. In *Proceedings of the 11th ACM SIGCOMM Conference on Internet Measurement (SIGCOMM '11)*, pages 81–98. ACM, 2011.
[41] P. Samarati. Protecting respondents identities in microdata release. *IEEE Transactions on Knowledge and Data Engineering*, 13(6):1010–1027, 2001.
[42] P. Samarati and L. Sweeney. Generalizing data to provide anonymity when disclosing information. In *Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS '98)*, pages 188–188. ACM, 1998.
[43] L. Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5): 557–570, 2002.
[44] B. Zhou and J. Pei. The k-anonymity and l-diversity approaches for privacy preservation in social networks against neighborhood attacks. *Knowledge and Information Systems*, 28(1):47–77, 2011.