Win-Win Kernelization for Degree Sequence Completion Problems

Vincent Froese*, André Nichterlein, and Rolf Niedermeier

Institut für Softwaretechnik und Theoretische Informatik, TU Berlin, Germany
{vincent.froese, andre.nichterlein, rolf.niedermeier}@tu-berlin.de

Abstract. We study the kernelizability of a class of NP-hard graph modification problems based on vertex degree properties. Our main positive results refer to NP-hard graph completion (that is, edge addition) cases while we show that there is no hope to achieve analogous results for the corresponding vertex or edge deletion versions. Our algorithms are based on a method that transforms graph completion problems into efficiently solvable number problems and exploits f-factor computations for translating the results back into the graph setting. Indeed, our core observation is that we encounter a win-win situation in the sense that either the number of edge additions is small (and thus faster to find) or the problem is polynomial-time solvable. This approach helps in answering an open question by Mathieson and Szeider [JCSS 2012].

1 Introduction

In this work, we propose a general approach for achieving polynomial-size problem kernels for a class of graph completion problems where the goal graph has to fulfill certain degree properties. Thus, we explore and enlarge results on provably effective polynomial-time preprocessing for these NP-hard graph problems. To a large extent, the initial motivation for our work comes from studying the NP-hard graph modification problem Degree Constraint Editing(S) for non-empty subsets $S \subseteq \{\mathrm{v}^-, \mathrm{e}^+, \mathrm{e}^-\}$ of editing operations (v^-: "vertex deletion", e^+: "edge addition", e^-: "edge deletion") as introduced by Mathieson and Szeider [22].[1] The definition reads as follows.

Degree Constraint Editing(S) (DCE(S))
Input: An undirected graph $G = (V, E)$, two integers $k, r > 0$, and a "degree list function" $\tau\colon V \to 2^{\{0,\dots,r\}}$.
Question: Is it possible to obtain a graph $G' = (V', E')$ from G using at most k editing operations of type(s) as specified by S such that $\deg_{G'}(v) \in \tau(v)$ for all $v \in V'$?

* Supported by DFG, project DAMM (NI 369/13).
[1] Mathieson and Szeider [22] originally introduced a weighted version of the problem, where the vertices and edges can have positive integer weights incurring a cost for each editing operation. Here, we focus on the unweighted version.

To appear in *Proceedings of the 14th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT' 14), Copenhagen, Denmark, July 2014.*

In our work, the set S always consists of a single editing operation. Our studies focus on the two most natural parameters: the number k of editing operations and the maximum allowed degree r. We will show that, although all three variants are NP-hard, DCE(e^+) is amenable to a generic kernelization method we propose. This method is based on dynamic programming solving a corresponding number problem and f-factor computations. For DCE(e^-) and DCE(v^-), however, we show that there is little hope to achieve analogous results.

Previous Work. There are basically two fundamental starting points for our work. First, there is our previous theoretical work on degree anonymization in social networks [15] motivated and strongly inspired by a preceding heuristic approach due to Liu and Terzi [19]. Indeed, our previous work for degree anonymization very recently inspired empirical work with encouraging experimental results [16]. A fundamental contribution of this work now is to systematically reveal what the problem-specific parts (tailored towards degree anonymization) and what the "general" parts of that approach are. In this way, we develop this approach into a general method of significantly wider applicability for a large number of graph completion problems based on degree properties. The second fundamental starting point is Mathieson and Szeider's work [22] on DCE(S). They showed several parameterized preprocessing (also known as kernelization) results and left open whether it is possible to reduce DCE(e^+) in polynomial time to a problem kernel of size polynomial in r —we will affirmatively answer this question. Finally, Golovach [13] achieved a number of kernelization results for closely related graph editing problems; his methods, however, significantly differ from ours.

From a more general perspective, all these considerations fall into the category of "graph editing to fulfill degree constraints", which recently received increased interest in terms of parameterized complexity analysis [10, 13, 23].

Our Contributions. Answering an open question of Mathieson and Szeider [22], we present an $O(kr^2)$-vertex kernel for DCE(e^+) which we then transfer into an $O(r^5)$-vertex kernel using a strategy rooted in previous work [15, 19]. A further main contribution of our work in the spirit of meta kernelization [2] is to clearly separate problem-specific from problem-independent aspects of this strategy, thus making it accessible to a wider class of degree sequence completion problems. We observe that in case that the goal graph shall have "small" maximum degree r, then the actual graph structure is in a sense negligible and thus allows for a lot of freedom that can be algorithmically exploited. This paves the way to a *win-win situation* of either having guaranteed a small number of edge additions or the overall problem being solvable in polynomial-time anyway.

Besides our positive kernelization results, we exclude polynomial-size problem kernels for DCE(e^-) and DCE(v^-) subject to the assumption that NP $\not\subseteq$ coNP/poly, thereby showing that the exponential-size kernel results by Mathieson and Szeider [22] are essentially tight. In other words, this demonstrates that in our context edge completion is much more amenable to kernelization than edge deletion or vertex deletion are. We also prove NP-hardness of DCE(v^-) and DCE(e^+) for graphs of maximum degree three, implying that the maximum degree is not a useful parameter for kernelization purposes. Last but not least, we develop a general preprocessing approach for DEGREE SEQUENCE COMPLETION problems which yields a search space size that is polynomially bounded in the parameter. While this per se does not give polynomial kernels, we derive fixed-parameter tractability with respect to the combined parameter maximum degree and solution size. The usefulness of our method is illustrated by further example degree sequence completion problems.

Notation. All graphs in this paper are undirected, loopless, and simple (that is, without multiple edges). For a graph $G = (V, E)$, we set $n := |V|$ and $m := |E|$. The degree of a vertex $v \in V$ is denoted by $\deg_G(v)$, the maximum vertex degree by Δ_G, and the minimum vertex degree by δ_G. For a finite set U, we denote with $\binom{U}{2}$ the set of all size-two subsets of U. We denote by $\overline{G} := (V, \binom{V}{2} \setminus E)$ the complement graph of G. For a vertex subset $V' \subseteq V$, the subgraph induced by V' is denoted by $G[V']$. For an edge subset $E' \subseteq \binom{V}{2}$, $V(E')$ denotes the set of all endpoints of edges in E' and $G[E'] := (V(E'), E')$. For a set E' of edges with endpoints in a graph G, we denote by $G + E' := (V, E \cup E')$ the graph that results by inserting all edges in E' into G. Similarly, we define for a vertex set $V' \subseteq V$, the graph $G - V' := G[V \setminus V']$. For each vertex $v \in V$, we denote by $N_G(v)$ the open neighborhood of v in G and by $N_G[v] := N_G(v) \cup \{v\}$ the closed neighborhood. We omit subscripts if the corresponding graph is clear from the context. A vertex $v \in V$ with $\deg(v) \in \tau(v)$ is called *satisfied* (otherwise *unsatisfied*). We denote by $U \subseteq V$ the set of all unsatisfied vertices, formally $U := \{v \in V \mid \deg_G(v) \notin \tau(v)\}$.

Parameterized Complexity. This is a two-dimensional framework for studying computational complexity [8, 11, 24]. One dimension of a parameterized problem is the input size s, and the other one is the *parameter* (usually a positive integer). A parameterized problem is called *fixed-parameter tractable* (fpt) with respect to a parameter ℓ if it can be solved in $f(\ell) \cdot s^{O(1)}$ time, where f is a computable function only depending on ℓ. This definition also extends to *combined parameters*. Here, the parameter usually consists of a tuple of positive integers $(\ell_1, \ell_2, \ldots)$ and a parameterized problem is called fpt with respect to $(\ell_1, \ell_2, \ldots)$ if it can be solved in $f(\ell_1, \ell_2, \ldots) \cdot s^{O(1)}$ time.

A core tool in the development of fixed-parameter algorithms is polynomial-time preprocessing by *data reduction* [1, 14, 20]. Here, the goal is to transform a given problem instance I with parameter ℓ in polynomial time into an equivalent instance I' with parameter $\ell' \leq \ell$ such that the size of I' is upper-bounded by some function g only depending on ℓ. If this is the case, we call I' a (problem) *kernel* of size $g(\ell)$. If g is a polynomial, then we speak of a *polynomial kernel*. Usually, this is achieved by applying polynomial-time executable data reduction rules. We call a data reduction rule $\mathcal{R}$ *correct* if the new instance I' that results from applying $\mathcal{R}$ to I is a yes-instance if and only if I is a yes-instance. The whole process is called *kernelization*. It is well known that a parameterized problem is fixed-parameter tractable if and only if it has a problem kernel.

Due to a lack of space several proofs are deferred to a full version.[2]

[2] Available on arXiv:1404.5432.

2 Degree Constraint Editing

Mathieson and Szeider [22] showed fixed-parameter tractability of DCE(S) for all non-empty subsets $S \subseteq \{v^-, e^-, e^+\}$ with respect to the combined parameter (k, r) and W[1]-hardness with respect to the single parameter k. The fixed-parameter tractability is in a sense tight as Cornuéjols [7] proved that DCE(e$^-$) is NP-hard on planar graphs with maximum degree three and with $r = 3$ and thus presumably not fixed-parameter tractable with respect to r. We complement his result by showing that DCE(v$^-$) is NP-hard on cubic (that is three-regular) planar graphs, even if $r = 0$, and that DCE(e$^+$) is NP-hard on graphs with maximum degree three.

Theorem 1. DCE(v$^-$) *is NP-hard on cubic planar graphs, even if* $r = 0$.

Proof (Sketch). We provide a polynomial-time many-one reduction from the NP-hard Vertex Cover on cubic planar graphs [12]. Let $(G = (V, E), h)$ be a Vertex Cover instance with the cubic planar graph G. It is not hard to see that (G, h) is a yes-instance of Vertex Cover if and only if $(G, h, 0, \tau)$ with $\tau(v) = \{0\}$ for all $v \in V$ is a yes-instance of DCE(v$^-$). □

Theorem 2. DCE(e$^+$) *is NP-hard on planar graphs with maximum degree three.*

In contrast to DCE(e$^-$) and DCE(v$^-$), unless P = NP, DCE(e$^+$) cannot be NP-hard for constant values of r since we later show fixed-parameter tractability for DCE(e$^+$) with respect to the parameter r.

Excluding Polynomial Kernels. Mathieson and Szeider [22] gave exponential-size problem kernels for DCE(v$^-$) and DCE({v$^-$, e$^-$}) with respect to the combined parameter (k, r). We prove that these results are tight in the sense that, under standard complexity-theoretic assumptions, neither DCE(e$^-$) nor DCE(v$^-$) admits a polynomial-size problem kernel when parameterized by (k, r).

Theorem 3. DCE(e$^-$) *does not admit a polynomial-size problem kernel with respect to* (k, r) *unless NP* $\subseteq$ *coNP/poly.*

Theorem 4. DCE(v$^-$) *does not admit a polynomial-size problem kernel with respect to* (k, r) *unless NP* $\subseteq$ *coNP/poly.*

Having established these computational lower bounds, we now show that in contrast to DCE(e$^-$) and DCE(v$^-$), DCE(e$^+$) admits a polynomial kernel.

2.1 A Polynomial Kernel for DCE(e$^+$) with respect to (k, r)

In order to describe the kernelization, we need some further notation: For $i \in \{0, \ldots, r\}$, a vertex $v \in V$ is of *type* i if and only if $\deg(v) + i \in \tau(v)$, that is, v can be satisfied by adding i edges to it. The set of all vertices of type i is denoted by T_i. Observe that a vertex can be of multiple types, implying that for $i \neq j$ the vertex sets T_i and T_j are not necessarily disjoint. Furthermore, notice that the

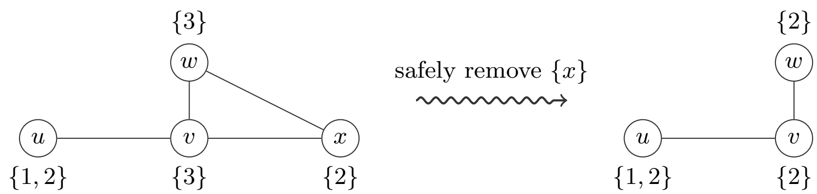

Fig. 1. An example for safely removing a vertex from a graph. The sets next to the vertices denote the degree lists defined by τ. Observe that in both graphs u is of type zero and of type one, v is of type zero, and w is of type one.

type-0 vertices are exactly the satisfied ones. We remark that there are instances for DCE(e$^+$) where we might have to add edges between two satisfied vertices (though this may seem counter-intuitive): Consider, for example, a three-vertex graph without any edges, the degree list function values are $\{2\}, \{0, 2\}, \{0, 2\}$, and $k = 3$. The two vertices with degree list $\{0, 2\}$ are satisfied. However, the only solution for this instance is to add *all* edges.

Now, we can describe our kernelization algorithm: The basic strategy is to keep the unsatisfied vertices U and "enough" arbitrary vertices of each type (from the satisfied vertices) and delete all other vertices. The idea behind the correctness is that the vertices in a solution are somehow "interchangeable". If an unsatisfied vertex needs an edge to a satisfied vertex of type i, then it is not important which satisfied type-i vertex is used. We only have to take care not to "reuse" the satisfied vertices to avoid the creation of multiple edges.

Next, we specify what we mean by "enough" vertices: The "magic number" is $\alpha := k(\varDelta_G + 2)$. This leads to the definition of *α-type sets*: An α-type set $C \subseteq V$ is a vertex subset containing all unsatisfied vertices U and $\min\{\alpha, |T_i \setminus U|\}$ type-i vertices from $T_i \setminus U$ for each $i \in \{1, \ldots, r\}$. We will soon show that for any fixed α-type set C, deleting all vertices in $V \setminus C$ results in an equivalent instance. However, deleting a vertex changes the degrees of its neighbors. Thus, we also have to adjust their degree lists. Formally, for a vertex subset $V' \subseteq V$, we define $\tau_{V'} : (V \setminus V') \to 2^{\{0,\ldots,r\}}$, where for each $u \in V \setminus V'$, we set $\tau_{V'}(u) := \{d \in \mathbb{N} \mid d + |N_G(u) \cap V'| \in \tau(u)\}$. Then, *safely removing* a vertex set $V' \subseteq V$ from the instance (G, k, r, τ) means to replace the instance with $(G - V', k, r, \tau_{V'})$, see Figure 1 for an example. With these definitions we can provide our reduction rules leading to a polynomial-size problem kernel.

Reduction Rule 1. Let $(G = (V, E), k, r, \tau)$ be an instance of DCE(e$^+$) and let $C \subseteq V$ be an α-type set in G. Then, safely remove all vertices in $V \setminus C$.

Lemma 1. *Reduction Rule 1 is correct and can be applied in linear time.*

As each α-type set contains at most α satisfied vertices of each vertex type, it follows that after one application of Reduction Rule 1 the graph contains at most $|C| = |U| + r\alpha$ vertices. The number of unsatisfied vertices in an α-type set can always be bounded by $|U| \leq 2k$ since we can increase the degrees of at most $2k$

vertices by adding k edges. If there are more unsatisfied vertices, then we return a trivial no-instance. Thus, we end up with $|C| \leq 2k + rk(\Delta_G + 2)$. To obtain a polynomial-size problem kernel with respect to the combined parameter (k, r), we need to bound Δ_G. However, this can easily be achieved: Since we only allow edge additions, for each vertex $v \in V$, we have $\deg(v) \leq \max \tau(v) \leq r$. Formalized as a data reduction rule, this reads as follows:

Reduction Rule 2. Let $(G = (V, E), k, r, \tau)$ be an instance of DCE(e^+). If G contains more than $2k$ unsatisfied vertices or if there exists a vertex $v \in V$ with $\deg(v) > \max \tau(v)$, then return a trivial no-instance.

Having applied Reduction Rules 1 and 2 once, it holds that $\Delta_G \leq r$ and thus the graph contains at most $2k + rk(r + 2)$ vertices. Lemma 1 ensures that we can apply Reduction Rule 1 in linear time. Note that linear time means $O(m + |\tau|)$ time, where $|\tau| \geq n$ denotes the encoding size of τ. Clearly, Reduction Rule 2 can be applied in linear time too. This leads to the following.

Theorem 5. DCE(e^+) *admits a problem kernel containing* $O(kr^2)$ *vertices computable in* $O(m + |\tau|)$ *time.*

2.2 A Polynomial Kernel for DCE(e^+) with respect to r

In this subsection, we show how to extend the polynomial-size problem kernel provided in Theorem 5 to a polynomial-size problem kernel for the single parameter r. To this end, among other things, we adapt some ideas of Hartung et al. [15] to show how to bound k in a polynomial of r. The general strategy, inspired by a heuristic of Liu and Terzi [19], will be as follows: First, remove the graph structure and solve the problem on the degree sequence of the input graph by using dynamic programming. The solution to this number problem will indicate the *demand* for each vertex, that is, the number of added edges incident to that vertex. Then, using a result of Katerinis and Tsikopoulos [17], we prove that either $k \leq r(r + 1)^2$ or we can find a set of edges satisfying the specified demands in polynomial time.

We start by formally defining the corresponding number problem and showing its polynomial-time solvability.

Number Constraint Editing (NCE)
Input: A function $\phi\colon \{1, \ldots, n\} \to 2^{\{0,\ldots,r\}}$ and positive integers $d_1, \ldots, d_n, k, r$.
Question: Are there n positive integers $d'_1, \ldots, d'_n$ such that $\sum_{i=1}^{n}(d'_i - d_i) = k$ and for all $i \in \{1, \ldots, n\}$ it holds that $d'_i \geq d_i$ and $d'_i \in \phi(i)$?

Lemma 2. NCE *is solvable in* $O(n \cdot k \cdot r)$ *time.*

Lemma 2 can be proved with a dynamic program that specifies the demand for each vertex, that is, the number of added edges incident to each vertex. Given these demands, the remaining problem is to decide whether there exists a set of edges that satisfy these demands and are not contained in the input graph G. This problem is closely related to the polynomial-time solvable f-Factor problem [21], a special case of DCE(e$^-$) where $|\tau(v)| = 1$ for all $v \in V$; it is formally defined as follows:

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 $v \in V$?

Observe that our problem of satisfying the demands of the vertices in G is essentially the question whether there is an f-factor in the complement graph $\overline{G}$ where the function f stores the demand of each vertex. Using a result of Katerinis and Tsikopoulos [17], we can show the following lemma about the existence of an f-factor:

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

We now have all ingredients to show that we can upper-bound k by $r(r + 1)^2$ or solve the given instance of DCE(e$^+$) in polynomial time. The main technical statement towards this is the following.

Lemma 4. *Let $I := (G = (V, E), k, r, \tau)$ be an instance of* DCE(e$^+$) *with $k \geq r(r + 1)^2$ and $V = \{v_1, \ldots, v_n\}$. If there exists a $k' \in \{r(r + 1)^2, \ldots, k\}$ such that $(\deg(v_1), \ldots, \deg(v_n), 2k', r, \phi)$ with $\phi(i) := \tau(v_i)$ is a yes-instance of* NCE*, then I is a yes-instance of* DCE(e$^+$).

Proof. Assume that $(\deg(v_1), \ldots, \deg(v_n), 2k', r, \phi)$ is a yes-instance of NCE. Let $d'_1, \ldots, d'_n$ be integers such that $d'_i \in \tau(v_i)$, $\sum_{i=1}^{n} d'_i - \deg(v_i) = 2k'$, and $d'_i \geq d_i$. Hence, we know that the degree constraints can numerically be satisfied, giving rise to a new target degree d'_i for each vertex v_i. Let $A := \{v_i \in V \mid d'_i > \deg(v_i)\}$ denote the set of *affected* vertices containing all vertices which require addition of some edges. It remains to show that the degree sequence of the affected vertices can in fact be realized by adding k' edges to $G[A]$. To this end, it is sufficient to prove the existence of an f-factor in the complement graph $\overline{G[A]}$ with $f(v_i) := d'_i - \deg(v_i) \in \{1, \ldots, r\}$ for all $v_i \in A$ since such an f-factor contains exactly the k' edges we want to add to G. Thus, it remains to check that all conditions of Lemma 3 are indeed satisfied to conclude the existence of the sought f-factor. First, note that $\delta_{\overline{G[A]}} \geq |A| - r - 1$ since $\Delta_{G[A]} \leq r$. Moreover, $\sum_{v_i \in A} (d'_i - \deg(v_i)) = 2k' \leq |A|r$, and thus $|A| \geq 2k'/r \geq 2(r + 1)^2$. Finally, $\sum_{v_i \in A} f(v_i) = 2k'$ is even and thus Lemma 3 applies. ☐

As NCE is polynomial-time solvable, Lemma 4 states a win-win situation: either the solution is bounded in size or can be found in polynomial time. From this and Theorem 5, we obtain the polynomial-size problem kernel.

Theorem 6. DCE(e$^+$) *admits a problem kernel containing $O(r^5)$ vertices computable in $O(k^2 \cdot r \cdot n + m + |\tau|)$ time.*

3 A General Approach for Degree Sequence Completion

In the previous section, we dealt with the problem DCE(e^+), where one only has to *locally* satisfy the degree of each vertex. In this section, we show how the presented ideas for DCE(e^+) can also be used to solve more *globally* defined problems where the degree sequence of the solution graph G' has to fulfill a given property. For example, consider the problem of adding a minimum number of edges to obtain a regular graph, that is, a graph where all vertices have the same degree. In this case the degree of a vertex in the solution is a priori not known but depends on the degrees of the other vertices.

The *degree sequence* of a graph G with n vertices is an n-tuple containing the vertex degrees. Then, for some tuple property Π, we consider the following problem:

Π-Degree Sequence Completion (Π-DSC)
Input: A graph $G = (V, E)$, an integer $k \in \mathbb{N}$.
Question: Is there a set of edges $E' \subseteq \binom{V}{2} \setminus E$ with $|E'| \leq k$ such that the degree sequence of $G + E'$ fulfills Π?

Note that Π-DSC is not a generalization of DCE(e^+) since in DCE(e^+) one can require for two vertices u and v of the same degree that u gets two more incident edges and v not. This cannot be expressed in Π-DSC. We remark that the results stated in this section can be extended to hold for a generalized version of Π-DSC where a "degree list function" τ is given as additional input and the vertices in the solution graph G' also have to satisfy τ, thus generalizing DCE(e^+). For simplicity, however, we stick to the easier problem definition as stated above.

3.1 Fixed-Parameter Tractability of Π-DSC

In this subsection, we first generalize the ideas behind Theorem 5 to show fixed-parameter tractability of Π-DSC with respect to the combined parameter (k, Δ_G). Then, we present an adjusted version of Lemma 4 and apply it to show fixed-parameter tractability for Π-DSC with respect to the parameter $\Delta_{G'}$. Clearly, a prerequisite for both these results is that the following problem has to be fixed-parameter tractable with respect to the parameter $\Delta_T := \max\{d_1, \ldots, d_n\}$.

Π-Decision
Input: An integer tuple $T = (d_1, \ldots, d_n)$.
Question: Does T fulfill Π?

For the next result, we need some definitions. 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. A subset $V' \subseteq V$ is an *α-block set* if V' contains for every $d \in \{1, \ldots, \Delta_G\}$ exactly $\min\{\alpha, |D_G(d)|\}$ vertices. Recall that $\alpha = (\Delta_G + 2)k$, see Section 2.1, and notice the similarity of α-block sets and α-type sets. This similarity is not a coincidence for we use ideas of Reduction Rule 1 and Lemma 1 to obtain the following lemma.

Lemma 5. *Let $I := (G = (V, E), k)$ be a yes-instance of Π-DSC and let $C \subseteq V$ be an α-block set. Then, there exists a set of edges $E' \subseteq \binom{C}{2} \setminus E$ with $|E'| \leq k$ such that the degree sequence of $G + E'$ fulfills Π.*

In the context of DCE(S), we introduced the notion of safely removing a vertex subset to obtain a problem kernel. On the contrary, in the context of Π-DSC, it seems impossible to remove vertices in general without further knowledge about the tuple property Π. Thus, Lemma 5 does not lead to a problem kernel but only to a reduced search space for a solution, namely any α-block set. Clearly, an α-block set C can be computed in polynomial time. Then, one can simply try out all possibilities to add edges with endpoints in C and check whether in one of the cases the degree sequence of the resulting graph satisfies Π. As $|C| \leq (\Delta_G + 2)k\Delta_G$, there are at most $O(2^{((\Delta_G+2)k\Delta_G)^2})$ possible subsets of edges to add. Overall, this leads to the following theorem.

Theorem 7. *Let Π be some tuple property. If Π-DECISION is fixed-parameter tractable with respect to Δ_T, then Π-DSC is fixed-parameter tractable with respect to (k, Δ_G).*

Bounding the Solution Size k in $\Delta_{G'}$. We now show how to extend the ideas of Section 2.2 to the context of Π-DSC in order to bound the solution size k by a polynomial in $\Delta_{G'}$. The general procedure still is the one inspired by Liu and Terzi [19]: Solve the number problem corresponding to Π-DSC on the degree sequence of the input graph and then try to "realize" the solution. To this end, we define the corresponding number problem as follows:

Π-NUMBER SEQUENCE COMPLETION (Π-NSC)
Input: Positive integers $d_1, \ldots, d_n, k, \Delta$.
Question: Are there n nonnegative integers $x_1, \ldots, x_n$ with $\sum_{i=1}^{n} x_i = k$ such that $(d_1 + x_1, \ldots, d_n + x_n)$ fulfills Π and $d_i + x_i \leq \Delta$?

With these problem definitions, we can now generalize Lemma 4.

Lemma 6. *Let $I := (G, k)$ be an instance of Π-DSC with $V = \{v_1, \ldots, v_n\}$ and $k \geq \Delta_{G'}(\Delta_{G'} + 1)^2$. If there exists a $k' \in \{\Delta_{G'}(\Delta_{G'} + 1)^2, \ldots, k\}$ such that the corresponding Π-NSC instance $I' := (\deg(v_1), \ldots, \deg(v_n), 2k', \Delta_{G'})$ is a yes-instance, then I is a yes-instance.*

Let function $g(|I|)$ denote the running time for solving the Π-NSC instance I. Clearly, if there is a solution for an instance of Π-DSC, then there also exists a solution for the corresponding Π-NSC instance. It follows that we can decide whether there is a large solution for Π-DSC (with at least $\Delta_{G'}(\Delta_{G'} + 1)^2$ edges) in $k \cdot g(n \log(n))$ time. Hence, we arrive at the following win-win situation:

Corollary 1. *Let $I := (G, k)$ be an instance of Π-DSC. Then, either one can decide in $k \cdot g(n \log(n))$ time that I is a yes-instance, or I is a yes-instance if and only if $(G, \min\{k, \Delta_{G'}(\Delta_{G'} + 1)^2\})$ is a yes-instance.*

Using Corollary 1, we can transfer fixed-parameter tractability of Π-NSC with respect to Δ to fixed-parameter tractability of Π-DSC with respect to $\Delta_{G'}$. Notice that $\Delta_{G'} \leq k + \Delta_G$, that is, $\Delta_{G'}$ is a smaller and thus "stronger" parameter [18]. Also, showing Π-NSC to be fixed-parameter tractable with respect to Δ is a significantly easier task than proving fixed-parameter tractability for Π-DSC with respect to $\Delta_{G'}$ directly since the graph structure can be completely ignored.

Theorem 8. *If Π-NSC is fixed-parameter tractable with respect to Δ, then Π-DSC is fixed-parameter tractable with respect to $\Delta_{G'}$.*

If Π-NSC can be solved in polynomial time, then Corollary 1 shows that we can assume that $k \leq \Delta_{G'}(\Delta_{G'} + 1)^2$. Thus, as in the DCE($e^+$) setting (Theorem 6), polynomial kernels with respect to (k, Δ_G) transfer to the parameter $\Delta_{G'}$, leading to the following.

Theorem 9. *If Π-NSC is polynomial-time solvable and Π-DSC admits a polynomial kernel with respect to (k, Δ_G), then Π-DSC also admits a polynomial kernel with respect to $\Delta_{G'}$.*

3.2 Applications

As our general approach is inspired by ideas of Hartung et al. [15], it is not surprising that it can be applied to "their" Degree Anonymity problem, where given an undirected graph $G = (V, E)$ and two positive integers k and s, one seeks an edge set E' over V of size at most s such that $G' := G + E'$ is k-*anonymous*, that is, for each vertex $v \in V$, there are at least $k - 1$ other vertices in G' having the same degree. The property Π of being k-anonymous clearly can be decided in polynomial time for a given degree sequence, and thus, by Theorem 7, we immediately get fixed-parameter tractability with respect to (s, Δ_G). Theorem 9 then basically yields the kernel results obtained by Hartung et al. [15]. There are more general versions of Degree Anonymity as proposed by Chester et al. [6]. For example, just a given subset of the vertices has to be anonymized or the vertices can have labels. As in each of these generalizations one can decide in polynomial time whether a given graph satisfies the particular anonymity requirement, Theorem 7 applies also in these scenarios. However, checking in which of these more general settings the conditions of Theorem 8 or Theorem 9 are fulfilled is future work.

Besides the graph anonymization setting, one could think of further, more general constraints on the degree sequence. For example, if $p_i(\mathcal{D})$ denotes how often degree i appears in a degree sequence $\mathcal{D}$, then being k-anonymous translates into $p_i(\mathcal{D}_{G'}) \geq k$ for all degrees i occurring in the degree sequence $\mathcal{D}_{G'}$ of the modified graph G'. Now, it is natural to consider not only a lower bound $k \leq p_i(\mathcal{D})$, but also an upper bound $p_i(\mathcal{D}) \leq u$ or maybe even a set of allowed frequencies $p_i(\mathcal{D}) \in F_i \subseteq \mathbb{N}$. Constraints like this allow to express properties not of individual degrees but of the whole distribution of the degrees in the resulting sequence. For example, in order to have some "balancedness" one can require that each occurring degree occurs exactly ℓ times for some $\ell \in \mathbb{N}$ [5]. To obtain some sort of "robustness" it might be useful to ask for an h-index of ℓ, that is, in the solution graph there are at least ℓ vertices with degree at least ℓ [9].

Another range of problems which fit naturally into our framework involves completion problems to a graph class that is completely characterized by degree sequences. For example, a graph is a *unigraph* if it is determined by its degree sequence up to isomorphism [4]. Given a degree sequence $\mathcal{D} = (d_1, \ldots, d_n)$, one can decide in linear time whether $\mathcal{D}$ defines a unigraph [3]. Thus, by Theorem 8, we conclude fixed-parameter tractability for the unigraph completion problem with respect to the parameter $\Delta_{G'}$.

4 Conclusion

We proposed a method for deriving efficient preprocessing algorithms for degree sequence completion problems. DCE(e^+) served as our main illustrating example. Roughly speaking, the core of the approach (as basically already used in previous work [15, 19]) consists of extracting the degree sequence from the input graph, efficiently solving a simpler number editing problem, and translating the obtained solution back into a solution for the graph problem using f-factors. While previous work [15, 19] was specifically tailored towards an application for graph anonymization, we generalized the approach by filtering out problem-specific parts and "universal" parts. Thus, whenever one can solve these problem-specific parts efficiently, we can automatically obtain efficient preprocessing and fixed-parameter tractability results.

Our approach seems promising for future empirical investigations concerning its practical usefulness, a very recent experimental work has been performed for Degree Anonymity [16]. Another line of future research could be to study polynomial-time approximation algorithms for the considered degree sequence completion problems. Perhaps parts of our preprocessing approach might find use here as well. A more specific open question concerning our work would be how to deal with additional connectivity requirements for the generated graphs.

Bibliography

[1] H. L. Bodlaender. Kernelization: New upper and lower bound techniques. In *Proc. 4th IWPEC*, volume 5917 of *LNCS*, pages 17–37. Springer, 2009. 3

[2] H. L. Bodlaender, F. V. Fomin, D. Lokshtanov, E. Penninkx, S. Saurabh, and D. M. Thilikos. (Meta) kernelization. In *Proc. 50th FOCS*, pages 629–638. IEEE, 2009. 2

[3] A. Borri, T. Calamoneri, and R. Petreschi. Recognition of unigraphs through superposition of graphs. *J. Graph Algorithms Appl.*, 15(3):323–343, 2011. 11

[4] A. Brandstädt, V. B. Le, and J. P. Spinrad. *Graph Classes: a Survey*, volume 3 of *SIAM Monographs on Discrete Mathematics and Applications*. SIAM, 1999. 11

[5] G. Chartrand, L. Lesniak, C. M. Mynhardt, and O. R. Oellermann. Degree uniform graphs. *Ann. N. Y. Acad. Sci.*, 555(1):122–132, 1989. 10

[6] S. Chester, B. Kapron, G. Srivastava, and S. Venkatesh. Complexity of social network anonymization. *Social Netw. Analys. Mining*, 3(2):151–166, 2013. 10

[7] G. Cornuéjols. General factors of graphs. *J. Combin. Theory Ser. B*, 45(2): 185–198, 1988. 4

[8] R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity*. Springer, 2013. 3

[9] D. Eppstein and E. S. Spiro. The h-index of a graph and its application to dynamic subgraph statistics. *J. Graph Algorithms Appl.*, 16(2):543–567, 2012. 11

[10] M. R. Fellows, J. Guo, H. Moser, and R. Niedermeier. A generalization of Nemhauser and Trotter's local optimization theorem. *J. Comput. System Sci.*, 77(6):1141–1158, 2011. 2

[11] J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer, 2006. 3

[12] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, 1979. 4

[13] P. A. Golovach. Editing to a connected graph of given degrees. *CoRR*, abs/1308.1802, 2013. 2

[14] J. Guo and R. Niedermeier. Invitation to data reduction and problem kernelization. *SIGACT News*, 38(1):31–45, 2007. 3

[15] S. Hartung, A. Nichterlein, R. Niedermeier, and O. Suchý. A refined complexity analysis of degree anonymization on graphs. In *Proc. 40th ICALP*, volume 7966 of *LNCS*, pages 594–606. Springer, 2013. Journal version to appear in Information and Computation. 2, 6, 10, 11

[16] S. Hartung, C. Hoffman, and A. Nichterlein. Improved upper and lower bound heuristics for degree anonymization in social networks. In *Proc. of the 13th SEA*, 2014. Accepted for publication. 2, 11

[17] P. Katerinis and N. Tsikopoulos. Minimum degree and f-factors in graphs. *New Zealand J. Math*, 29(1):33–40, 2000. 6, 7

[18] C. Komusiewicz and R. Niedermeier. New races in parameterized algorithmics. In *Proc. 37th MFCS*, volume 7464 of *LNCS*, pages 19–30. Springer, 2012. 10

[19] K. Liu and E. Terzi. Towards identity anonymization on graphs. In *ACM SIGMOD Conference*, SIGMOD '08, pages 93–106. ACM, 2008. 2, 6, 9, 11

[20] D. Lokshtanov, N. Misra, and S. Saurabh. Kernelization - preprocessing with a guarantee. In *The Multivariate Algorithmic Revolution and Beyond*, pages 129–161, 2012. 3

[21] L. Lovász and M. D. Plummer. *Matching Theory*, volume 29 of *Annals of Discrete Mathematics*. North-Holland, 1986. 7

[22] L. Mathieson and S. Szeider. Editing graphs to satisfy degree constraints: A parameterized approach. *J. Comput. System Sci.*, 78(1):179–191, 2012. 1, 2, 4

[23] H. Moser and D. M. Thilikos. Parameterized complexity of finding regular induced subgraphs. *J. Discrete Algorithms*, 7(2):181–190, 2009. 2

[24] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford University Press, 2006. 3