Theoretical Computer Science 412 (2011) 5744–5751

Contents lists available at SciVerse ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Kernels for below-upper-bound parameterizations of the hitting set and directed dominating set problems

G. Gutin *, M. Jones, A. Yeo

Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK

ARTICLE INFO

Article history:
Received 28 October 2010
Received in revised form 28 April 2011
Accepted 14 June 2011
Communicated by T. Erlebach

Keywords:
Hitting set
Parameterized problems
Fixed-parameter tractable
Hypergraph
Kernel

ABSTRACT

In the Hitting Set problem, we are given a collection $\mathcal{F}$ of subsets of a ground set V and an integer p, and asked whether V has a p-element subset that intersects each set in $\mathcal{F}$. We consider two parameterizations of Hitting Set below tight upper bounds, $p = m - k$ and $p = n-k$. In both cases k is the parameter. We prove that the first parameterization is fixed-parameter tractable, but has no polynomial kernel unless coNP $\subseteq$ NP/poly. The second parameterization is W[1]-complete, but the introduction of an additional parameter, the degeneracy of the hypergraph $H = (V, \mathcal{F})$, makes the problem not only fixed-parameter tractable, but also one with a linear kernel. Here the degeneracy of $H = (V, \mathcal{F})$ is the minimum integer d such that for each $X \subset V$ the hypergraph with vertex set $V \setminus X$ and edge set containing all edges of $\mathcal{F}$ without vertices in X, has a vertex of degree at most d.

In Nonblocker (Directed Nonblocker), we are given an undirected graph (a directed graph) G on n vertices and an integer k, and asked whether G has a set X of $n - k$ vertices such that for each vertex $y \notin X$ there is an edge (arc) from a vertex in X to y. Nonblocker can be viewed as a special case of Directed Nonblocker (replace an undirected graph by a symmetric digraph). Dehne et al. (Proc. SOFSEM 2006) proved that Nonblocker has a linear-order kernel. We obtain a linear-order kernel for Directed Nonblocker.

1. Introduction, terminology and notation

In the Hitting Set problem, we are given a collection $\mathcal{F}$ of subsets of a ground set V and an integer p, and asked whether V has a p-element subset that intersects each set in $\mathcal{F}$. It is a well-known problem with various applications, e.g., in software testing [16], in computer networks [18] and in bioinformatics [24]. Hitting Set is equivalent to the set cover problem and several of its special cases are of importance (e.g., the vertex cover and dominating set problems). Hitting Set is NP-complete and its standard parameterization (when p is the parameter) is W[2]-complete. (We provide basic parameterized complexity terminology and notation in Section 1.3.) A few alternative parameterizations of Hitting Set have also been studied and we briefly overview them below. To facilitate our discussion of various parameterizations of Hitting Set, consider the following generic parameterization:

> HitSet(p,κ)
> *Instance:* A set V, a collection $\mathcal{F}$ of subsets of V.
> *Parameter:* κ.
> *Question:* Does $(V, \mathcal{F})$ have a hitting set S of size at most p? (A subset S of V is called a *hitting set* if $S \cap F \neq \emptyset$ for each $F \in \mathcal{F}$.)

In what follows, n stands for the size of V and m for the size of $\mathcal{F}$.

* Corresponding author. Tel.: +44 1784414229.
E-mail addresses: gutin@cs.rhul.ac.uk, greg.gutin@gmail.com (G. Gutin), markj@cs.rhul.ac.uk (M. Jones), anders@cs.rhul.ac.uk (A. Yeo).

0304-3975/$ – see front matter
doi:10.1016/j.tcs.2011.06.018

Aside from HitSet(p, p), the standard parameterization of Hitting Set, the most well-known parameterization is HitSet$(p, p+s)$, where s is the maximum size of a set in $\mathcal{F}$. This parameterization is fixed-parameter tractable and has a kernel of size at most s^p (see Downey and Fellows [11]). Using the Sunflower Lemma, Flum and Grohe [12] obtained a kernel of size $O(sp^s s!)$. Abu-Khzam [1] recently proved that HitSet$(p, p+s)$ has a kernel in which the number of elements in the ground set V is at most $(2s-1)p^{s-1}+p$. Dom et al. [10] proved that HitSet$(p, p+s)$ does not admit a polynomial-size kernel unless coNP⊆NP/poly. Dom et al. [10] also proved that HitSet$(p, p+m)$ and HitSet$(p, p+n)$ have exponential-size kernels but no polynomial-size kernels unless coNP⊆NP/poly.

In this paper, we study two parameterizations: HitSet$(m-k, k)$ and HitSet$(n-k, k)$, as well as HitSet$(n-k, k)$ with an additional parameter. Both parameterizations are of the type "below a tight upper bound". Indeed, both m and n are tight upper bounds as it is easy to see that there is always a hitting set of size at most m (n, respectively) and to construct an infinite family of instances of Hitting Set in which no hitting set is of size less than m (n, respectively). A brief overview is given in Section 1.1 of some well-known results on problems parameterized below tight upper bounds. Section 1.2 is devoted to hypergraph terminology and notation; note that some terminology and notation that we use is new or nonstandard. A very brief introduction to fixed-parameter algorithmics is given in Section 1.3.

In Section 2, we prove that HitSet$(m-k, k)$ has a kernel with at most $k4^k$ elements of V and at most $k4^k$ sets. In our proof, we use a technique called greedy localization, introduced by Chen et al. [5]. This technique is often compared to the well-known iterative compression, see, e.g., Dehne et al. [8]. We also prove that HitSet$(m-k, k)$ has no kernel of polynomial size unless coNP⊆NP/poly. In our proof, we use the result of Dom et al. [10] on HitSet$(p, p+m)$ mentioned above.

In the problem Nonblocker, we are given a graph $G = (V, E)$ and an integer k, and asked whether there is a set $X \subseteq V$ of size at most $|V|-k$ such that each vertex $v \in V \setminus X$ is adjacent to a vertex in X. Here k is the parameter. Note that Nonblocker is a below-tight-upper-bound parameterization of the Dominating Set problem. It is well known that Nonblocker can be reduced to HitSet$(n-k,k)$ (see, e.g., [12], p. 18) and, thus, our no-polynomial-kernel result is in a sharp contrast to a linear-order-kernel result of Dehne et al. [9] for Nonblocker.

In Section 3, we show that HitSet$(n-k, k)$ is W[1]-complete, but the introduction of the second parameter, the corresponding hypergraph degeneracy (defined in Section 1.2), makes the problem not only fixed-parameter tractable, but also one with a linear kernel. Each hypergraph of maximum degree d is d-degenerate, but the family of d-degenerate hypergraphs has its maximum degree unbounded by any function of d. Thus, our result is an extension of the corresponding result when the maximum degree is the additional parameter. The Directed Nonblocker problem is an extension of Nonblocker to directed graphs: we are given a directed graph $G = (V, A)$ and an integer k, and asked whether there is a set $X \subseteq V$ of size at most $|V| - k$ such that for each vertex $v \in V \setminus X$ there is an arc from a vertex of X to v. Using our polynomial-size kernel result for HitSet$(n-k, k)$ with the additional parameter, we show that Directed Nonblocker has a kernel with at most $k^2 + k - 1$ vertices.

In Section 4, we improve the last result by showing that Directed Nonblocker has a kernel with at most $3k - 1$ vertices. To prove this result we use an inequality for the domination number of a digraph with at most one vertex of in-degree zero and no isolated vertices. Further research is discussed in Section 5.

1.1. Problems parameterized below tight upper bounds

Mahajan and Raman [19] were the first to recognize both practical and theoretical importance of parameterizing problems above tight lower bounds or below tight upper bounds. Further arguments for the importance of parameterizing problems above and below tight bounds were given in [20,21]. One example of a problem parameterized below an upper bound is Maximum Clique parameterized below n, the number of vertices in the input graph. Unlike the standard parameterization of Maximum Clique which is W[1]-complete, the parameterization below n is fixed-parameter tractable (and has a linear-order kernel) simply because it is equivalent to the standard parameterization of Vertex Cover. The parameterization below n is important in bioinformatics applications, where the maximum order of a clique is close to n [2].

However, in many cases establishing parameterized complexity of a problem parameterized below a tight upper bound is less straightforward and has been stated as an open question. One such well-known problem is Directed Feedback Vertex Set: given a digraph D and an integer k, decide whether D has an acyclic induced subgraph on at least $n-k$ vertices. The parameterized complexity of Directed Feedback Vertex Set was a long standing open question solved by Chen et al. [6] who established its fixed-parameter tractability. Other well-known examples are Bipartization (decide whether a graph has a bipartite induced subgraph on $n-k$ vertices) which was proved to be fixed-parameter tractable by Reed et al. [23] and Almost 2-Sat (decide whether there is a truth assignment that satisfies at least $m-k$ clauses in a 2-CNF formula with m clauses) which was proved to be fixed-parameter tractable by Razgon and O'Sullivan [22]. Interestingly, no polynomial-size kernel is known for any of the three problems and it is still unknown whether such a kernel exists.

Certainly, not every natural problem parameterized below a tight upper bound is fixed-parameter tractable. A trivial example of such a problem is Almost 3-Sat (decide whether there is a truth assignment that satisfies at least $m-k$ clauses in a 3-CNF formula with m clauses). A less trivial example is the following problem: given a graph G, its maximal matching M and an integer k, decide whether G has a vertex cover with at most $2|M| - k$ vertices. This problem was proved to be W[1]-hard by Gutin et al. [13]. (Here we assume that W[1]≠FPT as widely believed.)

1.2. Hypergraphs

While studying Hitting Set, it will be more convenient for us to use hypergraph terminology and notation, we introduce the relevant terminology and notation in this subsection.

A *hypergraph* $H = (V, \mathcal{F})$ consists of a nonempty set V of *vertices* and a family $\mathcal{F}$ of nonempty subsets of V called *edges* of H. Note that $\mathcal{F}$ may have *parallel* edges, i.e., copies of the same subset of V. For any vertex $v \in V$, and any $\mathcal{E} \subseteq \mathcal{F}$, $\mathcal{E}[v]$ is the set of edges in $\mathcal{E}$ containing v, $N[v]$ is the set of all vertices contained in edges of $\mathcal{F}[v]$, and the *degree* of v is $d(v) = |\mathcal{F}[v]|$. For a subset T of vertices, $\mathcal{F}[T] = \bigcup_{v \in T} \mathcal{F}[v]$.

Deleting an edge e from a *hypergraph* $H = (V, \mathcal{F})$ results in a new hypergraph $H - e$ with vertex set V and edge set $\mathcal{F} \setminus \{e\}$. *Deleting* a vertex v from a *hypergraph* $H = (V, \mathcal{F})$ results in a new hypergraph $H - v$ with vertex set $V \setminus \{v\}$ and edge set $\{e \setminus \{v\} : e \in \mathcal{F}\}$.

A set T of vertices *hits* all edges in $\mathcal{F}[T]$ and an edge e is *hit* by any vertex belonging to it. A set $S \subseteq V$ is called a *hitting set* of a hypergraph $H = (V, \mathcal{F})$ if it hits $\mathcal{F}$. Hitting Set can be formulated as a problem in which we are given a hypergraph H and an integer p and asked whether H contains a hitting set of size at most p.

For a hypergraph $H = (V, \mathcal{F})$ and a set $X \subset V$, the subhypergraph $H \ominus X$ is obtained from H by deleting the set $\mathcal{E}$ of all edges hit by X and all vertices contained only in $\mathcal{E}$. A hypergraph $H = (V, \mathcal{F})$ is *d-degenerate* if, for all $X \subset V$, the subhypergraph $H \ominus X$ contains a vertex of degree at most d. The *degeneracy* $\deg(H)$ of a hypergraph H is the smallest d for which H is d-degenerate.

The degeneracy of a hypergraph can be calculated in linear time using the following algorithm. Pick a vertex v_1 in H of minimum degree d_1, and set $H := H \ominus \{v_1\}$. Pick a vertex v_2 of minimum degree d_2 and set $H := H \ominus \{v_2\}$, and so on. Then $d = \max\{d_i : i \in [n]\}$ is the degeneracy of H. (It is clear that the degeneracy of H must be at least d; the equality follows by observing that for any $X \subset V$ the smallest numbered vertex $u_i \in V \setminus X$ has degree at most d_i in $H \ominus X$.)

1.3. Fixed-parameter tractability and kernels

A *parameterized problem* is a subset $L \subseteq \Sigma^* \times \mathbb{N}$ over a finite alphabet Σ. L is *fixed-parameter tractable* if the membership of an instance (I, k) in $\Sigma^* \times \mathbb{N}$ can be decided in time $f(k)|I|^{O(1)}$ where f is a computable function of the *parameter* k only [11,12,21]. Given a parameterized problem L, a *kernelization of L* is a polynomial-time algorithm that maps an instance (x, k) to an instance (x', k') (the *kernel*) such that (i) $(x, k) \in L$ if and only if $(x', k') \in L$, (ii) $k' \leq h(k)$, and (iii) $|x'| \leq g(k)$ for some functions h and g. The function $g(k)$ is called the *size* of the kernel. It is well known [11,12,21] that a decidable parameterized problem L is fixed-parameter tractable if and only if it has a kernel. Polynomial-size kernels are of main interest, due to applications [11,12,21], but unfortunately not all fixed-parameter problems have such kernels unless coNP⊆NP/poly, see, e.g., [3,4,10].

2. Hitting set parameterized below *m*

In this section we consider HitSet$(m - k, k)$. Let $H = (V, \mathcal{F})$ be a hypergraph.

We begin with some reduction rules. The first two reduction rules have been used by previous Hitting Set algorithms [1,27]. The third is a trivial rule included to simplify later proofs.

Reduction Rule 1. *If there exist distinct $e, e' \in \mathcal{F}$ such that $e \subseteq e'$, set $H := H - e'$ and $k := k - 1$.*

Reduction Rule 2. *If there exist $u, v \in V$ such that $u \neq v$ and $\mathcal{F}[u] \subseteq \mathcal{F}[v]$, set $H := H - u$.*

Reduction Rule 3. *If there exist $v \in V$, $e \in \mathcal{F}$ such that $\mathcal{F}[v] = \{e\}$ and $e = \{v\}$, then delete v and e.*

Lemma 1. *Let (H, k) and (H', k') be instances of* HitSet$(m - k, k)$ *such that (H', k') is derived from (H, k) by repeated applications of Rules 1–3. Then (H', k') is a* Yes*-instance if and only if (H, k) is a* Yes*-instance.*

Proof. Rule 1: Any vertex in V which hits e will also hit e'. Therefore a set $S \subseteq V$ is a hitting set for H if and only if it is also a hitting set for $H - e'$. In $H - e'$ the difference between m and the size of the desired hitting set is one less, so we reduce k by 1.

Rule 2: Any edge which is hit by u is also hit by v. Therefore, if S is a hitting set containing u, we can get another hitting set of equal size or smaller by removing u and adding v. Therefore we may assume u is not in the hitting set and delete u from H.

Rule 3: The proof is trivial. □

Lemma 2. *Let $(H = (V, \mathcal{F}), k)$ be an instance of* HitSet$(m - k, k)$ *which is reduced by Rules 1–3 and $\mathcal{F} \neq \emptyset$. Then for all $v \in V$, $d(v) \geq 2$, and for all $e \in \mathcal{F}$, $|e| \geq 2$.*

Proof. Consider $v \in V$. Suppose $d(v) = 0$. Then trivially, $\mathcal{F}[v] \subseteq \mathcal{F}[u]$ for any $u \in V$, and so Rule 2 applies, a contradiction. Suppose $d(v) = 1$. Then let e be the single edge containing v. Either e contains another vertex u, in which case $\mathcal{F}[v] \subseteq \mathcal{F}[u]$ and Rule 2 applies, or $e = \{v\}$, in which case Rule 3 applies, a contradiction. Thus, $d(v) \geq 2$. A similar argument, using Rule 1 instead of Rule 2, can be used to show that $|e| \geq 2$ for all $e \in \mathcal{F}$. □

We now introduce the concept of a *mini-hitting set*. Lemma 3 shows that the problem of finding a hitting set of size $m-k$ is equivalent to the problem of finding a mini-hitting set.

Definition 1. A *mini-hitting set* is a set $S_{\text{MINI}} \subseteq V$ such that $|S_{\text{MINI}}| \leq k$ and $|\mathcal{F}[S_{\text{MINI}}]| \geq |S_{\text{MINI}}| + k$.

Lemma 3. *A reduced hypergraph $H = (V, \mathcal{F})$ has a hitting set of size at most $m-k$ if and only if it has a mini-hitting set. Moreover,*

1. *Given a mini-hitting set S_{MINI}, we can construct a hitting set S with $|S| \leq m-k$ such that $S_{\text{MINI}} \subseteq S$ in polynomial time.*
2. *Given a hitting set S with $|S| \leq m-k$, we can construct a mini-hitting set S_{MINI} such that $S_{\text{MINI}} \subseteq S$ in polynomial time.*

Proof. 1. For each edge e not hit by S_{MINI}, pick one vertex in e and add it to S_{MINI}. The resulting set S contains at most $m-k$ vertices and hits every edge of $\mathcal{F}$.
2. If $|S| \leq k$ then S itself is a mini-hitting set.
If $|S| > k$, construct S_{MINI} as follows. Let $S_0 = \emptyset$, and for every $0 \leq i \leq m-k-1$, let $S_{i+1} = S_i \cup \{v\}$, where $v \in S \backslash S_i$ is picked to maximize $|\mathcal{F}[v] \backslash \mathcal{F}[S_i]|$. Suppose for a contradiction that $|\mathcal{F}[S_k]| < |S_k| + k$. Then for some $j < k$, $|\mathcal{F}[S_{j+1}]| \leq |\mathcal{F}[S_j]| + 1$. Thus by construction, $|\mathcal{F}[S_{i+1}]| \leq |\mathcal{F}[S_i]| + 1$ for all $i \geq j$. It follows that $|\mathcal{F}[S]| = |\mathcal{F}[S_{m-k}]| < |S_{m-k}| + k = m$, a contradiction. Therefore $|\mathcal{F}[S_k]| \geq |S_k| + k$, and thus S_k is the required S_{MINI}. □

We now describe a greedy algorithm which constructs a set $S^* \subseteq V$. Either S^* is a mini-hitting set, or $\mathcal{F}[S^*]$ has some useful properties which will allow us to bound $|V|$.

Start with $S^* = \emptyset$. While $|\mathcal{F}[S^*]| < |S^*| + k$ and there exists $v \in V$ with $|\mathcal{F}[v] \backslash \mathcal{F}[S^*]| > 1$, do the following: Pick a vertex $v \in V$ such that $|\mathcal{F}[v] \backslash \mathcal{F}[S^*]|$ is as large as possible, and add v to S^*.

If S^* is a mini-hitting set, then by Lemma 3 we are done. We will now assume that S^* is not a mini-hitting set. Let $\mathcal{C} = \mathcal{F}[S^*]$, and let $\mathcal{I} = \mathcal{F} \backslash \mathcal{C}$.

Lemma 4. *Suppose S^* is not a mini-hitting set. Then we have the following:*

1. $|S^*| < k$.
2. $|\mathcal{C}| < 2k$.
3. *For all* $v \in V$, $|\mathcal{C}[v]| \geq 1$ *and* $|\mathcal{I}[v]| \leq 1$.
4. *For all* $v \in V$, $d(v) \leq k$.

Proof. 1. Suppose for a contradiction $|S^*| \geq k$. Then at some point in the construction of S^* we have $|S^*| = k$. Observe that at each stage in the construction of S^*, $|\mathcal{F}[S^*]| \geq 2|S^*|$. It follows that when $|S^*| = k$, $|\mathcal{F}[S^*]| \geq |S^*| + k$, and the algorithm stops. Note that S^* is a mini-hitting set, a contradiction.
2. Suppose for a contradiction that $|\mathcal{C}| \geq 2k$. Then since $|S^*| < k$, $|\mathcal{F}[S^*]| = |\mathcal{C}| \geq |S^*| + k$, and so S^* is a mini-hitting set, a contradiction.
3. Since $|S^*| < k$ but S^* is not a mini-hitting set, the construction of S^* must stop because $|\mathcal{F}[v] \backslash \mathcal{F}[S^*]| \leq 1$ for all $v \in V$, i.e. $|\mathcal{I}[v]| \leq 1$. By Lemma 2, $d(v) \geq 2$ for all $v \in V$. Since $d(v) = |\mathcal{C}[v]| + |\mathcal{I}[v]|$, it follows that $|\mathcal{C}[v]| \geq 1$.
4. Suppose for a contradiction that there exists $v \in V$ with $d(v) > k$. Then in the construction of S^*, we first add a vertex u to S^* with $d(u) > k$. We therefore have a set S^* with $|S^*| = 1$, $|\mathcal{F}[S^*]| \geq k+1 = |S^*| + k$, and so the algorithm terminates and S^* is a mini-hitting set, a contradiction. □

We now have that $\mathcal{F} = \mathcal{C} \uplus \mathcal{I}$, with $|\mathcal{C}| < 2k$, and every vertex in V hits at least one edge in $\mathcal{C}$ and at most one edge in $\mathcal{I}$. Furthermore $2 \leq d(v) \leq k$ for every $v \in V$, and $|e| \geq 2$ for every $e \in \mathcal{F}$. We are no longer interested in S^*.

Using $\mathcal{C}$ and $\mathcal{I}$, we introduce another reduction rule that will bound $|V|$ and consequently $|\mathcal{F}|$.

Reduction Rule 4. *Reduce (H, k) using Rules 1–3, and let $\mathcal{C}$ be as defined above. For any $\mathcal{C}' \subseteq \mathcal{C}$, let $V(\mathcal{C}') = \{v \in V : \mathcal{C}[v] = \mathcal{C}'\}$. If $|V(\mathcal{C}')| > k$, pick a vertex $v \in V(\mathcal{C}')$ and set $H := H - v$.*

Lemma 5. *Let (H, k) and (H', k) be instances of* HitSet$(m-k, k)$ *such that (H, k) is reduced under Rules 1–3 and (H', k) is derived from (H, k) by an application of Rule 4. Then (H', k) is a* Yes*-instance if and only if (H, k) is a* Yes*-instance.*

Proof. Let v be the vertex removed from V during an application of Rule 4. By Lemma 3, (H, k) is a Yes-instance if and only if H has a mini-hitting set. It is therefore enough to show that H has a mini-hitting set if and only if $H - v$ has a mini-hitting set.

Suppose $S_{\text{MINI}} \subseteq V$ is a mini-hitting set for H, and assume that $v \in S_{\text{MINI}}$. By Rule 2, each $u \in V(\mathcal{C}')$ is in a different edge $e_u \in \mathcal{I}$. Furthermore, by Lemma 4 part 3, $e_u \cap e_{u'} = \emptyset$ for any $u \neq u' \in V(\mathcal{C}')$. As $|V(\mathcal{C}')| > k$, it follows that there exists $u \in V(\mathcal{C}')$ such that $e_u \cap S_{\text{MINI}} = \emptyset$, i.e. $e_u \notin \mathcal{F}[S_{\text{MINI}}]$. Therefore $S'_{\text{MINI}} = (S_{\text{MINI}} \backslash \{v\}) \cup \{u\}$ hits e_u, which is not hit by S_{MINI}, and the only edge which is hit by S_{MINI} but not S'_{MINI} is e_v. Therefore $|S'_{\text{MINI}}| = |S_{\text{MINI}}|$ and $|\mathcal{F}[S'_{\text{MINI}}]| \geq |\mathcal{F}[S_{\text{MINI}}]| \geq |S_{\text{MINI}}| + k = |S'_{\text{MINI}}| + k$, so S'_{MINI} is a mini-hitting set that does not contain v. Therefore S'_{MINI} is a mini-hitting set for $H - v$.

The reverse direction is trivial: If S_{MINI} is a mini-hitting set for $H - v$ then it is also a mini-hitting set for H. □

Note that although the number of subsets of $\mathcal{C}$ can be exponential in k, Rule 4 can be run in polynomial time. This is because we do not need to check every subset $\mathcal{C}' \subseteq \mathcal{C}$; it is enough to calculate $\mathcal{C}[v]$ for each $v \in V$ and only calculate $|V(\mathcal{C}')|$ if there exists $v \in V$ for which $\mathcal{C}[v] = \mathcal{C}'$.

Theorem 1. HitSet$(m-k, k)$ *has a kernel with at most* $k4^k$ *vertices and at most* $k4^k$ *edges.*

Proof. Let (H, k) be an instance of HitSet$(m-k, k)$ irreducible by the above four reduction rules and let $H = (V, \mathcal{F})$. The number of possible subsets $\mathcal{C}' \subseteq \mathcal{C}$ is $2^{|\mathcal{C}|} < 2^{2k}$. Therefore by Rule 4 $n = |V| < k2^{2k} = k4^k$.

To bound $m = |\mathcal{F}|$ recall that $d(v) \leq k$ for all $v \in V$, and $|e| \geq 2$ for all $e \in \mathcal{F}$. It follows that $|\mathcal{F}| \leq k|V|/2 < k^2 2^{2k-1}$. We can improve this bound as follows.

Here we make use of Theorem 5, which we prove in the next section. If $m \leq n$ then obviously $m \leq k4^k$. Suppose that $m > n$. Then finding a hitting set of size $m-k$ is equivalent to finding a hitting set of size $n-k'$, where $k' = (n-m)+k$ is less than k. By Theorem 5, this has a kernel with $m \leq d(d+1)k' < d(d+1)k$, where d is the degeneracy of H. By Lemma 4 part 4, we have $d \leq k$ and, thus, $m \leq k^2(k+1)$. Therefore we have a kernel with $m \leq k2^{2k}$ in either case. □

We now show that our exponential kernel for HitSet$(m-k, k)$ cannot be improved to a polynomial size one, given certain complexity assumptions. We make use of a result of Dom et al. [10] who proved the following:

Theorem 2. HitSet$(p, m+p)$ *does not have a polynomial kernel, unless coNP* $\subseteq$ *NP/poly.*

We may now prove the following theorem:

Theorem 3. HitSet$(m-k, k)$ *does not have a polynomial kernel, unless coNP* $\subseteq$ *NP/poly.*

Proof. Assume that HitSet$(m-k, k)$ has a polynomial kernel. We will show that HitSet$(p, m+p)$ has a polynomial kernel, a contradiction unless *coNP* $\subseteq$ *NP/poly*.

Consider an instance of HitSet$(p, m+p)$, in which we are given a hypergraph $H = (V, \mathcal{F})$ with $|V| = n$, $|\mathcal{F}| = m$, together with an integer p, and are asked to find a hitting set in H of size p. Let $k = m - p$. Observe that H has a hitting set of size p if and only if H has a hitting set of size $m-k$. By our assumption, there is a transformation which produces a hypergraph $H' = (V', \mathcal{F}')$ with $|V'| = n'$, $|\mathcal{F}'| = m'$ together with an integer k', such that H has a hitting set of size $m-k$ if and only if H' has a hitting set of size $m'-k'$. Furthermore, $m', n' \leq P(k)$ for some polynomial P, and the transformation takes time polynomial in n and m. We may assume without loss of generality that P is an increasing function.

Let $p' = m' - k'$, and observe that H has a hitting set of size p if and only if H' has a hitting set of size p'. Therefore H' with parameter p' is an equivalent instance of HitSet$(p, m+p)$ which can be constructed in time polynomial in m and n, and $m', n' \leq P(k) = P(m-p) \leq P(m+p)$, i.e. the size of the instance is bounded by a polynomial in the original parameter. It remains to show that the new parameter $m' + p'$ is also bounded by a function of the original parameter, but this follows from the fact that $p' \leq m'$. □

3. Hitting Set parameterized below *n*

Unlike HitSet$(m-k, k)$, HitSet$(n-k, k)$ is not fixed-parameter tractable unless $FPT = W[1]$.

Theorem 4. HitSet$(n-k, k)$ *is* $W[1]$*-complete.*

Proof. To show hardness, we use a well-known reduction from independent set, in which we are given a graph $G = (V, E)$ and are asked whether it contains an independent set $V' \subseteq V$ set of size k, where k is the parameter. In our instance of HitSet$(n-k, k)$, we let H be G viewed as a hypergraph. Then for any $V' \subseteq V$ with $|V'| = k$, V' is an independent set in the graph if and only if every edge contains a member of $V \backslash V'$, i.e. $V \backslash V'$ is a hitting set.

To show membership in $W[1]$, we reduce HitSet$(n-k, k)$ to the problem p-WSat$(\Gamma^-_{2,1})$, described in Flum and Grohe [12]. $\Gamma^-_{2,1}$ is the class of CNF formulas which contain only negative literals. In the parameterized problem p-WSat$(\Gamma^-_{2,1})$, we are given a formula in $\Gamma^-_{2,1}$ and an integer parameter k, and we are asked whether the formula has a satisfying assignment in which exactly k variables are assigned True. It follows from Theorem 7.29 in Flum and Grohe [12] that p-WSat$(\Gamma^-_{2,1})$ is in $W[1]$ (a more general problem is in $W[1]$).

For an instance of HitSet$(n-k, k)$, let $V = \{v_1, \ldots v_n\}$ be the vertices and $\mathcal{F} = \{e_1, \ldots e_m\}$ the edges in H. For each edge $e \in \mathcal{F}$, we let the clause $C_e = \bigvee_{v_i \in e} \bar{x}_i$, and let our formula be $\bigwedge_{j \in [m]} C_{e_j}$. Then there is a hitting set of size $(n-k)$ if and only if the formula has a satisfying assignment in which exactly k variables are assigned True. This is precisely the problem p-WSat$(\Gamma^-_{2,1})$, and so we are done. □

Note that in the hardness proof above, every set in the HitSet$(n-k, k)$ instance was of size 2. This means that HitSet$(n-k, k)$ is $W[1]$-hard even for the subcase where the edge size is bounded by r, for any $r \geq 2$. Therefore if we let the parameter be $k + \max_{e \in \mathcal{F}} |e|$, the problem is still $W[1]$-hard.

Another approach would be to consider the degree of the vertices as an additional parameter. Under this parameterization the problem does turn out to be fixed-parameter tractable; in fact we prove a stronger result by showing that the problem is fixed-parameter tractable with respect to $k+d$, where d is the degeneracy of H. This is the problem HitSet$(n-k, k+d)$.

We begin with the following simple result on the chromatic number of a d-degenerate hypergraph. For a hypergraph $H = (V, \mathcal{F})$, a mapping $c : V \to [t]$ is called a *proper t-coloring* if each edge e of H of cardinality at least 2 is not *monochromatic*, i.e., e has vertices u, v such that $c(u) \neq c(v)$. Here $c(u)$ is the *color* of u. The *chromatic number* $\chi(H)$ of a hypergraph H is the minimum integer t for which H has a proper t-coloring.

Lemma 6. *The chromatic number of a d-degenerate hypergraph is at most* $d + 1$.

Proof. The proof is by induction on the number n of vertices of H. If $n = 1$ then H has no edge of cardinality 1, and so $\chi(H) = 1$. Now assume that $n \geq 2$. Let v be a vertex of minimum degree q in $H = (V, \mathcal{F})$. By the induction hypothesis and definition of a d-degenerate hypergraph, $\chi(H \ominus \{v\}) \leq d + 1$. Consider a $(d + 1)$-coloring of $H \ominus \{v\}$ and edges $e_1, \ldots, e_q$ of cardinality at least 2 containing v. Note that $q \leq d$ and form a set C of colors by picking one color used in each e_i (if any vertex in e_i is colored). If C is empty, add to it color 1. Clearly, $|C| \leq d$ and, thus, there is a color t not in C among colors in $[d + 1]$. Assign v using color t and use one of the colors in C to color all other uncolored vertices. Observe that none of $e_1, \ldots, e_q$ is monochromatic. □

To get rid of edges of cardinality 1, we use the following rule whose correctness is easy to see.

Reduction Rule 5. *If there exist* $v \in V$, $e \in \mathcal{F}$ *such that* $e = \{v\}$, *then replace* $H = (V, \mathcal{F})$ *by* $H \ominus \{v\}$. *Keep k the same.*

For a hypergraph H, a set S of vertices is *independent* if S does not contain any edge of H.

Theorem 5. *The problem* HitSet$(n - k, k + d)$ *admits a kernel with less than* $(d + 1)k$ *vertices and* $d(d + 1)k$ *edges.*

Proof. Let H be a d-degenerate hypergraph. Using Rule 5 as long as possible, we reduce H to a d-degenerate hypergraph with no edge of cardinality 1. By Lemma 6, $\chi(H) \leq d+1$. Consider a proper $\chi(H)$-coloring of H and a largest set S of vertices of H assigned the same color. Clearly, $|S| \geq |V|/(d + 1)$.

Now observe that T is a hitting set of $H = (V, \mathcal{F})$ if and only if $V \setminus T$ is an independent set. Thus, if $|V|/(d + 1) \geq k$, the answer to HitSet$(n - k, k + d)$ is Yes. Otherwise, $|V| < (d + 1)k$.

To prove that $|\mathcal{F}| < d(d + 1)k$, choose a vertex v of minimum degree and observe that $d(v) \leq d$. Now delete v from V and $\mathcal{F}[v]$ from $\mathcal{F}$, and choose a vertex v of minimum degree again, and observe that $d(v) \leq d$. Continuing this procedure we will delete all edges in $\mathcal{F}$ and thus $|\mathcal{F}| \leq d|V| < d(d + 1)k$. □

In a directed graph $G = (V, A)$, a *dominating set* is a set $V' \subseteq V$ such that for every vertex $u \in V \setminus V'$, there is a vertex $v \in V'$ such that there is an arc from v to u. Recall that in Directed Nonblocker, we are given a directed graph G with n vertices and an integer k, and asked whether G has a dominating set with at most $n - k$ vertices.

Corollary 1. Directed Nonblocker *has a kernel with at most* $k^2 + k - 1$ *vertices.*

Proof. Let $(G = (V, A), k)$ be an instance of Directed Nonblocker with $|V| = n$. If G has a vertex v of out-degree at least k, then $V \setminus \{w \in V : vw \in A\}$ is a dominating set of size at most $n - k$. Thus, we may assume that the maximum out-degree of G is at most $k - 1$.

We construct an instance of HitSet$(n-k, k)$ as follows. Let $H = (V, \mathcal{F})$, where $\mathcal{F} = \{N^-[v] : v \in V\}$, $N^-[v] = \{v\} \cup \{u \in V : uv \in A\}$. Observe that $N^-[v]$ is hit by a set $S \subseteq V$ if and only if $v \in S$ or v is dominated by a vertex in S. Therefore, H has a hitting set of size $|\mathcal{F}| - k = |V| - k$ if and only if G has a dominating set of size $|V| - k$.

Since the maximum out-degree of G is at most $k - 1$, the maximum degree of a vertex in H is at most k. Thus, the degeneracy d of H is at most k and the result follows from Theorem 5. □

4. Directed Nonblocker

In this section, we improve the bound of Corollary 1 for $k > 2$.

It is well known that every hypergraph $H = (V, \mathcal{F})$ in which each edge has at least two vertices, has a hitting set of cardinality at most $(|V| + |\mathcal{F}|)/3$, cf. [26]. We start from a minor extension of this result.

Lemma 7. *Let* $H = (V, \mathcal{F})$ *be a hypergraph such that every edge has at least two vertices apart from, possibly, one edge that has just one vertex. If H has a one-vertex edge* $e = \{v\}$, *let there be another edge f of H containing v. Then H has a hitting set of cardinality at most* $(|V| + |\mathcal{F}|)/3$.

Proof. Let $n = |V|$ and $m = |\mathcal{F}|$. The proof is by induction on $n \geq 2$. If $n = 2$, then H has a hitting set of cardinality 1 and $n+m \geq 3$. Now assume that $n \geq 3$ and let $t(H)$ be the minimum cardinality of a hitting set in H. If H has a one-vertex edge $e = \{v\}$, then set $u = v$. Otherwise, let u be a vertex of H of maximum degree. Remove u from H together with all edges containing u and all vertices contained only in the removed edges. Denote the resulting hypergraph by H' and let n' and m' be the number of vertices and edges, respectively, in H'. Then $3t(H) \leq 3 + 3t(H') \leq 3 + n' + m' \leq n + m$. The second inequality in this chain of inequalities is by the induction hypothesis and the third inequality is due to the fact that either we remove at least two edges and one vertex or at least two vertices and one edge. □

For a digraph D, let $\gamma(D)$ denote the minimum size of a dominating set in D. Using Lemma 7, it is easy to prove the following key lemma of this section.

Lemma 8. *Let D a digraph on n vertices, none of which are isolated, and let D have at most one vertex of in-degree zero. Then* $\gamma(D) \leq 2n/3$.

Proof. We construct an instance $H = (V, \mathcal{F})$ of HitSet$(n-k, k)$ as in the proof of Corollary 1. The lemma follows from that facts that H satisfies the conditions of Lemma 7, $|V| = |\mathcal{F}|$, and the minimum cardinalities of a hitting set in H and a dominating set in D coincide. □

Theorem 6. Directed Nonblocker *has a kernel with at most* $3k - 1$ *vertices.*

Proof. Let D be a digraph with n vertices. If D has isolated vertices, then delete them without changing the answer to Directed Nonblocker as all of them must be in any dominating set of D. Thus, we may assume that D has no isolated vertices. Let S be the set of all vertices of D of in-degree zero. Assume that $|S| > 1$. Then contract all vertices of S into one vertex s which dominates all vertices dominated by S. Let D' be the resulting digraph. Since all vertices of S must be in any dominating set of D, the answers to Directed Nonblocker on D and on D' are the same. Thus, we may assume that $|S| \leq 1$. Then, by Lemma 8, $\gamma(D) \leq 2n/3$ and, thus, if $n - k \geq 2n/3$, the answer to Directed Nonblocker is Yes. Otherwise, $n - k < 2n/3$ and $n \leq 3k - 1$. □

To obtain a smaller kernel for Directed Nonblocker, it might be helpful to use further results on hitting sets of hypergraphs with a lower bound on the minimum size of an edge. Chvátal and McDiarmid [7] and Tuza [25] proved independently that a hypergraph $H = (V, \mathcal{F})$ with minimum edge size equal three, has a hitting set of size at most $(|V| + |\mathcal{F}|)/4$. Thomassé and Yeo [26] showed that if the minimum edge in a hypergraph $H = (V, \mathcal{F})$ is four and the minimum size of a hitting set of H is t, then $21t \leq 5|V| + 4|\mathcal{F}|$.

5. Further research

For a hypergraph H, let $\alpha(H)$ be the maximum size of an independent set of H. In the proof of Theorem 5, we observed that HitSet$(n-k, k+d)$ is equivalent to problem of deciding whether $\alpha(H) \geq k$ for a d-degenerate hypergraph H. This observation and the inequality $\alpha(H) \geq n/(d+1)$, where n is the number of vertices in H, have allowed us to obtain a linear kernel for HitSet$(n-k, k+d)$. However, the inequality $\alpha(H) \geq n/(d+1)$ suggests that, in fact, to have the parameter small in relevant cases (as it should be in the spirit of parameterized algorithmics) it makes more sense to consider the following parameterization above tight lower bound: decide whether for a d-degenerate hypergraph H we have $\alpha(H) \geq n/(d+1)+\kappa$, where κ is the new parameter. (Problems parameterized above tight lower bounds were studied in several papers including [13–15,19,20].)

It would be interesting to determine the parameterized complexity of the last problem even in the case of graphs. To the best of our knowledge, the only related result was obtained by Gutin et al. [13] who observed that the last problem has a linear kernel for graphs of maximum degree at most d. Indeed, in the case of graphs with maximum degree at most d, the existence of a linear kernel is an easy consequence of the following Brooks' Theorem [28]: for a graph G with maximum degree at most d we have $\chi(G) \leq d$ unless one of the connectivity components of G is K_{d+1} or, if $d = 2$ and one of the connectivity components of G is an odd cycle. Brook's Theorem was extended to hypergraphs by Kostochka et al. [17] who proved that for a connected hypergraph H with all edges of cardinality at least 2 and of maximum degree at most d we have $\chi(H) \leq d$ unless H has only one edge of cardinality at least 3 or H is a graph (in which case the odd cycle and complete cases apply). Thus, the observation of [13] can be extended to hypergraphs.

References

[1] F.N. Abu-Khzam, A kernelization algorithm for d-hitting set, J. Comput. System Sci. 76 (2010) 524–531.
[2] F.N. Abu-Khzam, R.L. Collins, M.R. Fellows, M.A. Langston, W.H. Sutters, C.T. Symons, Kernelization algorithms for the vertex cover problem: theory and experiments, in: Proc. Workshop Algorithm Engineering and Experiments, ALENEX'04, 2004, pp. 62–69.
[3] H.L. Bodlaender, R.G. Downey, M.R Fellows, D. Hermelin, On problems without polynomial kernels, J. Comput. System Sci. 75 (8) (2009) 423–434.
[4] H.L. Bodlaender, S. Thomasse, A. Yeo, Kernel Bounds for Disjoint Cycles and Disjoint Paths, in: Proc. 17th Europ. Symp. Algorithms, in: Lect. Notes Comput. Sci., vol. 5757, 2009, pp. 635–646.
[5] J. Chen, D.K. Firesen, W. Jia, I.A. Kanj, Using nondeterminism to design efficient deterministic algorithms, Algorithmica 40 (2) (2004) 83–97.
[6] J. Chen, Y. Liu, S. Lu, B. O'Sullivan, I. Razgon, A fixed-parameter algorithm for the directed feedback vertex set problem, J. ACM 55 (5) (2008) 1–19.
[7] V. Chvátal, C. McDiarmid, Small transversals in hypergraphs, Combinatorica 12 (1992) 19–26.
[8] F. Dehne, M.R. Fellows, F.A. Rosamond, P. Shaw, Greedy localizaion, iterative compression, and modeled crown reduction: new FPT techniques, an improved algorithm for set splitting, and a novel $2k$ kernelization for Vertex Cover, in: Proc. 1st International Workshop on Parameterized and Exact Computation, in: Lect. Notes Comput. Sci., vol. 3162, 2004, pp. 271–280.
[9] F. Dehne, M. Fellows, H. Fernau, E. Prieto, F. Rosamond, Nonblocker: parameterized algorithmics for Minimum Dominating Set, in: Proc. 32nd International Conference on Current Trends in Theory and Practice of Computer Science, in: Lect. Notes Comput. Sci., vol. 3831, 2006, pp. 237–245.
[10] M. Dom, D. Lokshtanov, S. Saurabh, Incompressibility though Colors and IDs, in: Proc. 36th International Colloquium on Automata, Languages and Programming, Part I, in: Lect. Notes Comput. Sci., vol. 5555, 2009, pp. 378–389.
[11] R.G. Downey, M.R. Fellows, Parameterized Complexity, Springer Verlag, 1999.
[12] J. Flum, M. Grohe, Parameterized Complexity Theory, Springer Verlag, 2006.
[13] G. Gutin, E.J. Kim, M. Lampis, V. Mitsou, Vertex cover problem parameterized above and below tight bounds, Theory Comput. Syst. 48 (2011) 402–410.
[14] G. Gutin, E. J. Kim, M. Mnich, A. Yeo, Betweenness parameterized above tight lower bound, J. Comput. System Sci. 76 (2010) 872–878.
[15] G. Gutin, E.J. Kim, S. Szeider, A. Yeo, A probabilistic approach to problems parameterized above or below tight bounds, J. Comput. Syst. Sci. 77 (2011) 422–429.
[16] J.A. Jones, M.J. Harrold, Test-suite reduction and prioritization for modified condition/decision coverage, IEEE Trans. Softw. Eng. 29 (3) (2009) 195–209.
[17] A. Kostochka, M. Stiebitz, B. Wirth, The colour theorems of Brooks and Gallai extended, Discrete Math. 162 (1–3) (1996) 299–303.
[18] F. Kuhn, P. von Rickenbach, R. Wattenhofer, E. Welzl, A. Zollinger, Interference in cellular networks: the minimum membership set cover problem, in: Proc. 11th Annual International Conference on Computing and Combinatorics, COCOON'05, in: Lect. Notes Comput. Sci., vol. 3595, 2005, pp. 188–198.

[19] M. Mahajan, V. Raman, Parameterizing above guaranteed values: MaxSat and MaxCut, J. Algorithms 31 (2) (1999) 335–354.
[20] M. Mahajan, V. Raman, S. Sikdar, Parameterizing above or below guaranteed values, J. Comput. Syst. Sci. 75 (2) (2009) 137–153.
[21] R. Niedermeier, Invitation to Fixed-Parameter Algorithms, Oxford University Press, 2006.
[22] I. Razgon, B. O'Sullivan, Almost 2-SAT is fixed-parameter tractable, J. Comput. Syst. Sci. 75 (8) (2009) 435–450.
[23] B. Reed, K. Smith, A. Vetta, Finding odd cycle transversals, Oper. Res. Lett. 32 (4) (2004) 299–301.
[24] D. Ruchkys, S. Song, A parallel approximation hitting set algorithm for gene expression analysis, in: Proc. 14th Symposium on Computer Architecture and High Performance Computing, IEEE Comput. Soc., 2002, pp. 75–81.
[25] Z. Tuza, Covering all cliques of a graph, Discrete Math. 86 (1990) 117–126.
[26] S. Thomassé, A. Yeo, Total domination of graphs and small transversals of hypergraphs, Combinatorica 27 (4) (2007) 473–487.
[27] K. Weihe, Covering trains by stations or the power of data reduction, in: Proc. Workshop Algorithm Engineering and Experiments, ALENEX'98, 1998, pp. 1–8.
[28] D.B. West, Introduction to Graph Theory, Prentice Hall, 2001.