

Backdoors to Abduction

Andreas Pfandler* Stefan Rümmele* Stefan Szeider†

Vienna University of Technology, Austria

{pfandler, ruemmele}@dbai.tuwien.ac.at, stefan@szeider.net

Abstract

Abductive reasoning (or Abduction, for short) is among the most fundamental AI reasoning methods, with a broad range of applications, including fault diagnosis, belief revision, and automated planning. Unfortunately, Abduction is of high computational complexity; even propositional Abduction is Σ_2^P -complete and thus harder than NP and co-NP. This complexity barrier rules out the existence of a polynomial transformation to propositional satisfiability (SAT). In this work we use structural properties of the Abduction instance to break this complexity barrier. We utilize the problem structure in terms of small backdoor sets. We present fixed-parameter tractable transformations from Abduction to SAT, which make the power of today’s SAT solvers available to Abduction.

1 Introduction

Abductive reasoning (or *Abduction*, for short) is among the most fundamental reasoning methods. It is used to explain observations by finding appropriate causes. In contrast to deductive reasoning, it is therefore a method for “reverse inference”. Abduction has a broad range of applications in AI, including system and medical diagnosis, planning, configuration, and database updates.

Unfortunately, Abduction is of high computational complexity. Already propositional Abduction, the focus of this paper, is Σ_2^P -complete and thus located at the second level of the Polynomial Hierarchy [Eiter and Gottlob, 1995]. Consequently, Abduction is harder than NP and co-NP. This complexity barrier rules out the existence of a polynomial-time transformation to the propositional satisfiability problem (SAT). As a consequence, one cannot directly apply powerful SAT solvers to Abduction. However, this would be very desirable in view of the enormous power of state-of-the-art SAT solvers that can handle instances with millions of clauses and variables [Gomes *et al.*, 2008; Katebi *et al.*, 2011; Järvisalo *et al.*, 2012].

*Supported by the Austrian Science Fund (FWF): P25518-N23.

†Supported by the European Research Council (ERC), project 239962.

Main contribution. We present a new approach to utilize problem structure in order to break this complexity barrier for Abduction. More precisely, we present transformations from Abduction to SAT that run in quadratic time, with a constant factor that is exponential in the distance of the given propositional theory from being HORN or KROM. We measure the distance in terms of the size of a smallest *strong backdoor set* into the classes HORN or KROM, respectively [Williams *et al.*, 2003]. Thus the exponential blow-up—which is to be expected when transforming a problem from the second level of the Polynomial Hierarchy into SAT—is confined to the size of the backdoor set, whereas the order of the polynomial running time remains constant independent of the distance. Such transformations are known as *fixed-parameter tractable reductions* and are fundamental to Parameterized Complexity Theory [Downey and Fellows, 1999; Flum and Grohe, 2006].

Our new approach to Abduction has several interesting aspects. It provides flexibility and openness. Any additional constraints that one might want to impose on the solution to the Abduction problem can simply be added as additional clauses to the SAT encoding. Hence such constraints can be handled without the need for modifying the basic transformation. The reduction approach readily supports the enumeration of subset-minimal solutions, as we can delegate the enumeration to the employed SAT solver. Any progress made for SAT solvers directly translates into progress for Abduction.

As a by-product, our approach gives rise to several new fixed-parameter tractability results. For instance, Abduction is fixed-parameter tractable when parameterized by the number of hypotheses and the size of a smallest HORN- or KROM-backdoor set. Parameterized by the size of backdoor sets alone, Abduction is para-NP-complete.

Related Work. Methods from parameterized complexity have turned out to be well-suited to tackle hard problems in Knowledge Representation & Reasoning [Gottlob and Szeider, 2008]. In particular, the concept of backdoor sets provides a natural way of parameterizing such problems [Samer and Szeider, 2009; Fichte and Szeider, 2011; Gaspers and Szeider, 2012a; 2012b; Lackner and Pfandler, 2012a; 2012b; Dvorák *et al.*, 2012]. The parameterized complexity of Abduction was subject of previous work, where different parameters have been considered [Gottlob *et al.*, 2010; Fellows *et al.*, 2012]. The most significant difference to our work is that

we use fixed-parameter tractability not to *solve* the Abduction problem itself, but to *reduce* it from the second level of the Polynomial Hierarchy to the first. This way, our parameters can be less restrictive and are potentially smaller for larger classes of Abduction instances. This novel use of fixed-parameter tractability was recently applied in the domain of answer-set programming [Fichte and Szeider, 2013], and we believe it can be applied to other hard reasoning problems as well. Abduction can be transformed in polynomial time to QBF [Egly *et al.*, 2000], but then one remains on the second level of the Polynomial Hierarchy.

2 Preliminaries

Propositional Logic. A formula in conjunctive normal form is a conjunction of disjunctions of literals; we denote the class of all such formulas by CNF. It is convenient to view a formula in CNF also as a set of clauses and a clause as a set of literals. KROM denotes the class of all CNF formulas having clause size at most 2. HORN formulas are CNF formulas with at most one positive literal per clause. Let $\text{var}(\varphi)$ be the set of variables occurring in $\varphi \in \text{CNF}$.

A (partial) *truth assignment* is a mapping $\tau : X \rightarrow \{0, 1\}$ defined for a set X of variables. We write $\text{var}(\tau)$ to denote the domain X . To extend τ to literals we put $\tau(\neg x) = 1 - \tau(x)$ for $x \in X$. By $\text{ta}(X)$ we denote the set of all truth assignments $\tau : X \rightarrow \{0, 1\}$. Let S be a set of variables. We denote by $\text{ta}(X, S)$ the set $\{\tau \in \text{ta}(X) \mid \forall s \in S \cap X : \tau(s) = 1\}$. We say that τ satisfies literal l if $\tau(l) = 1$. A clause is tautological if it contains a variable x and its negation $\neg x$. A truth assignment τ *satisfies* a CNF formula if in each non-tautological clause, there exist a literal that is satisfied by τ . A CNF formula φ is *satisfiable* (or *consistent*) if there exists some truth assignment τ that satisfies φ . If, additionally, $\text{var}(\tau)$ contains all variables of φ , we call it a *model* of φ . The *truth assignment reduct* of a CNF formula φ under a truth assignment τ is the CNF formula $\varphi[\tau]$ obtained from φ by deleting all clauses that are satisfied by τ and by deleting from the remaining clauses all literals that are set to 0 by τ . Let $\varphi, \psi \in \text{CNF}$, then φ entails ψ (denoted by $\varphi \models \psi$) if every model τ of φ with $\text{var}(\psi) \subseteq \text{var}(\tau)$ is also a model of ψ .

Let $\text{Res}(\varphi)$ denote the set of clauses computed from $\varphi \in \text{CNF}$ by iteratively applying resolution and dropping tautological clauses until a fixed-point is reached. Applying resolution adds the clause $C \cup D$ to φ if $C \cup \{x\} \in \varphi$ and $D \cup \{\neg x\} \in \varphi$. Let C be a non-tautological clause then $C \in \text{Res}(\varphi)$ if and only if $\varphi \models \{C\}$. For more information, cf. [Leitsch, 1997].

Let X be a set of variables, we define $\bar{X} := \{\neg x \mid x \in X\}$. For convenience, we view a set of variables X as the conjunction of its elements, whenever it is used as a formula.

Propositional Abduction. A (propositional) *abduction instance* consists of a tuple $\langle V, H, M, T \rangle$, where V is the set of *variables*, $H \subseteq V$ is the set of *hypotheses*, $M \subseteq V$ is the set of *manifestations*, and T is the *theory*, a formula in CNF over V . It is required that $M \cap H = \emptyset$. We define the size of an abduction instance \mathcal{P} to be the size of a reasonable encoding of \mathcal{P} . For instance taking $|V| + |H| + |M| + \sum_{C \in T} |C|$ would do. A set $S \subseteq H$ is a *solution* to (or explanation of) \mathcal{P} if $T \cup S$ is consistent and $T \cup S \models M$ (entailment). $\text{Sol}(\mathcal{P})$

denotes the set of all solutions to \mathcal{P} . The *solvability problem* for *propositional abduction* ABD is the following problem:

ABD
Instance: An abduction instance \mathcal{P} .
Problem: Decide whether $\text{Sol}(\mathcal{P}) \neq \emptyset$.

We will additionally consider a version of the abduction problem where we search for certain solutions only. A solution S is *subset-minimal* if there is no solution $S' \subsetneq S$.

ABD \subseteq
Instance: An abduction instance \mathcal{P} and a hypothesis h .
Problem: Is there a *subset-minimal* solution $S \subseteq H$ to \mathcal{P} with $h \in S$.

Recall that ABD is Σ_2^P -complete in general while it becomes NP-complete when the theory is a HORN or KROM formula. The same holds for ABD \subseteq [Eiter and Gottlob, 1995; Nordh and Zanuttini, 2008; Selman and Levesque, 1990].

Example. In this example we look for explanations why the nicely planned skiing trip ended up so badly. The abduction instance is given as follows:

$V = \{\text{snows, rains, precipitation, warm, hurt, sad}\}$
 $H = \{\text{precipitation, warm, hurt}\}$
 $M = \{\text{sad}\}$
 $T = \{\text{precipitation} \rightarrow \text{rains} \vee \text{snows},$
 $\text{hurt} \rightarrow \text{sad}, \text{warm} \rightarrow \neg \text{snows}, \text{rains} \rightarrow \text{sad}\}$

One can easily verify that $S_1 = \{\text{hurt}\}$ and $S_2 = \{\text{precipitation, warm}\}$ are solutions to this abduction instance. Notice that S_1 and S_2 are the only subset-minimal solutions while, e.g., $S_3 = \{\text{hurt, warm}\}$ is a solution as well.

Parameterized Complexity. We give some basic background on parameterized complexity. For more detailed information we refer to other sources [Downey and Fellows, 1999; Flum and Grohe, 2006]. A *parameterized problem* L is a subset of $\Sigma^* \times \mathbb{N}$ for some finite alphabet Σ . For an instance $(I, k) \in \Sigma^* \times \mathbb{N}$ we call I the *main part* and k the *parameter*. L is *fixed-parameter tractable* if there exists a computable function f and a constant c such that there exists an algorithm that decides whether $(I, k) \in L$ in time $\mathcal{O}(f(k) \|I\|^c)$ where $\|I\|$ denotes the size of I . Such an algorithm is called an *fpt-algorithm*. We will use the $\mathcal{O}^*(\cdot)$ notation which is defined in the same way as $\mathcal{O}(\cdot)$, but ignores polynomial factors.

Let $L \subseteq \Sigma^* \times \mathbb{N}$ and $L' \subseteq \Sigma'^* \times \mathbb{N}$ be two parameterized problems for some finite alphabets Σ and Σ' . An *fpt-reduction* r from L to L' is a many-to-one reduction from $\Sigma^* \times \mathbb{N}$ to $\Sigma'^* \times \mathbb{N}$ such that for all $I \in \Sigma^*$ we have $(I, k) \in L$ if and only if $r(I, k) = (I', k') \in L'$. Thereby, $k' \leq g(k)$ for a fixed computable function $g : \mathbb{N} \rightarrow \mathbb{N}$, and there is a computable function f and a constant c such that r is computable in time $\mathcal{O}(f(k) \|I\|^c)$ where $\|I\|$ denotes the size of I [Flum and Grohe, 2006]. Thus, an fpt-reduction is, in particular, an fpt-algorithm. We would like to note that the theory of fixed-parameter intractability is based on fpt-reductions.

Backdoors. Williams *et al.* [2003] introduced the notion of *backdoors* to explain favorable running times and the heavy-tailed behavior of SAT and CSP solvers on practical instances.

Backdoors are defined with respect to a fixed class \mathcal{C} of CNF formulas, the *base class* (or more figuratively, *island of tractability*). A *strong \mathcal{C} -backdoor set* of a formula $\varphi \in \text{CNF}$ is a set \mathcal{B} of variables such that $\varphi[\tau] \in \mathcal{C}$ for each $\tau \in \text{ta}(\mathcal{B})$. \mathcal{B} is also called a strong \mathcal{C} -backdoor set of an abduction instance $\mathcal{P} = \langle V, H, M, T \rangle$ if \mathcal{B} is a strong \mathcal{C} -backdoor set of T . Observe that the instance from the example above has a strong HORN-backdoor set and a strong KROM-backdoor set of cardinality one (consider, e.g., $\mathcal{B} = \{\text{snows}\}$).

Backdoor Approach. The backdoor approach consists of two phases. First, a backdoor set is computed (*detection*) and afterwards the backdoor is used to solve the problem (*evaluation*). For example, for SAT this approach works as follows. If we know a strong \mathcal{C} -backdoor set of a CNF formula φ of size k , we can reduce the satisfiability of φ to the satisfiability of 2^k easy formulas that belong to the base class. The challenging problem, however, is to find a strong backdoor set of size at most k , if it exists. This problem is NP-hard for all reasonable base classes, but fortunately, fixed-parameter tractable for the base classes KROM and HORN if parameterized by k [Nishimura *et al.*, 2004]. In particular, efficient fixed-parameter algorithms for the 3-HITTING SET problem and for the VERTEX COVER problem can be used for detecting strong backdoor sets for the base classes KROM and HORN. Fastest known fixed-parameter algorithms for these two problems run in time $\mathcal{O}^*(2.270^k)$ and $\mathcal{O}^*(1.2738^k)$ [Niedermeier and Rossmanith, 2003; Chen *et al.*, 2010], respectively. For further information on the parameterized complexity of backdoor set detection we refer to a recent survey [Gaspers and Szeider, 2012b].

For Abduction the detection phase is the same as for SAT, but the evaluation phase becomes the new challenge. We therefore focus on the evaluation phase, and assume that the backdoor set is provided as part of the input. However, whether or not we provide the backdoor set as part of the input does not affect the parameterized complexity of the overall problem, since, as explained above, the detection of strong HORN/KROM-backdoors is fixed-parameter tractable.

3 Transformations using Horn Backdoors

In this section we present the transformation from ABD to SAT using strong HORN-backdoor sets. In our transformation we will build upon ideas from Dowling and Gallier [1984] for computing the unique minimal model (with respect to set-inclusion) of HORN formulas in linear time. Recall that manifestations in the considered Abduction formalism are assumed to be positive literals (i.e., variables). The following lemma captures the evaluation phase of our backdoor approach. The proof can be found in the full version [Pfandler *et al.*, 2013].

Lemma 1. *Let $\mathcal{P} = \langle V, H, M, T \rangle$ be an abduction instance, let $S \subseteq H$, and let $\mathcal{B} \subseteq V$. Then S is a solution to \mathcal{P} if and only if (i) $\exists \tau \in \text{ta}(\mathcal{B}, S)$ such that $T[\tau] \cup S$ is consistent, and (ii) $\forall \tau \in \text{ta}(\mathcal{B}, S)$, $T[\tau] \cup S \models M[\tau]$.*

Based on this lemma, Algorithm 1 checks whether a given candidate is indeed a solution.

Algorithm 1: Solution-Checker_{HORN-bds}

Input : An abduction instance $\mathcal{P} = \langle V, H, M, T \rangle$, a strong HORN-backdoor set \mathcal{B} of \mathcal{P} and a solution candidate $S \subseteq H$ to be checked.
Output : Decision whether S is a solution to \mathcal{P} .

```

1 consistent ← false
2 entailment ← true
3 foreach  $\tau \in \text{ta}(\mathcal{B}, S)$  do
4   if  $\neg$  consistent then
5     if  $T[\tau] \cup S$  is consistent then
6       consistent ← true
7   if  $T[\tau] \cup S$  is consistent then
8      $U \leftarrow$  the unique minimal model of  $T[\tau] \cup S$ 
9     if  $U \not\models M[\tau]$  then // Thus  $T[\tau] \cup S \not\models M[\tau]$ 
10      entailment ← false
11      break
12 return consistent  $\wedge$  entailment
```

Lemma 2. *Let $\mathcal{P} = \langle V, H, M, T \rangle$ be an abduction instance and \mathcal{B} a strong HORN-backdoor set of \mathcal{P} . A set $S \subseteq H$ is a solution to \mathcal{P} if and only if Algorithm 1 returns yes.*

Proof. (\Rightarrow) Assume there exists a solution $S \in \text{Sol}(\mathcal{P})$. By Lemma 1 there exists a $\tau \in \text{ta}(\mathcal{B}, S)$ such that $T[\tau] \cup S$ is consistent. Therefore, for one of the assignments from Line 3, the flag *consistent* will be set to *true* in Line 6. Furthermore, by Lemma 1, for all $\tau \in \text{ta}(\mathcal{B}, S)$ it holds that $T[\tau] \cup S \models M[\tau]$. Therefore, Line 10 will not be reached and the flag *entailment* will remain *true*. Hence, the algorithm returns *yes*.

(\Leftarrow) Assume that the algorithm returns *yes*. Therefore, Line 10 is never reached and $U \models M[\tau]$ for all $\tau \in \text{ta}(\mathcal{B}, S)$. Since $M[\tau]$ is entailed in the minimal model and contains only positive literals, it is entailed in every model and $T[\tau] \cup S \models M[\tau]$ for all $\tau \in \text{ta}(\mathcal{B}, S)$. Furthermore, there exists a $\tau \in \text{ta}(\mathcal{B}, S)$ such that Line 6 is reached and hence $T[\tau] \cup S$ is consistent. It follows from Lemma 1 that $S \in \text{Sol}(\mathcal{P})$. \square

Corollary 3. *Let $\mathcal{P} = \langle V, H, M, T \rangle$ be an abduction instance and \mathcal{B} a strong HORN-backdoor set of \mathcal{P} . We can check whether \mathcal{P} has a solution in time $\mathcal{O}^*(2^{|\mathcal{B}|+|H|})$. Hence, ABD is fixed-parameter tractable when parameterized by $|\mathcal{B}|+|H|$.*

Proof. One has to check for each of the $2^{|\mathcal{B}|}$ -many solution candidates $S \subseteq H$, whether S is a solution to \mathcal{P} . To this end we apply Algorithm 1, which runs in time $\mathcal{O}^*(2^{|\mathcal{B}|})$. \square

If the number of hypotheses is large, this fpt-algorithm is not efficient. To overcome this limitation, we present next an fpt-reduction to SAT using only the backdoor size as the parameter. This is the main result of this section.

Theorem 4. *Given an abduction instance \mathcal{P} of input size n and a strong HORN-backdoor \mathcal{B} of \mathcal{P} of size k , we can create in time $\mathcal{O}(2^k n^2)$ a CNF formula $F_{\text{HORN-Solv}}$ of size $\mathcal{O}(2^k n^2)$ such that $F_{\text{HORN-Solv}}$ is satisfiable if and only if $\text{Sol}(\mathcal{P}) \neq \emptyset$.*

Proof. Let $\mathcal{P} = \langle V, H, M, T \rangle$. We will first construct a propositional formula $F'_{\text{HORN-Solv}}$ which is not in CNF. The required CNF formula $F_{\text{HORN-Solv}}$ can then be obtained from $F'_{\text{HORN-Solv}}$ by means of the well-known transformation due to Tseitin [1968], in which auxiliary variables are introduced.

This transformation produces for a given propositional formula F' in linear time a CNF formula F such that both formulas are equivalent with respect to their satisfiability, and the length of F is linear in the length of F' .

Note that in this encoding a solution $S \subseteq H$ can be obtained by projecting a model of $F'_{\text{HORN-Solv}}$ to the variables in H , i.e., $h \in S$ if and only if $h \in H$ is true in the model. We define

$$F'_{\text{HORN-Solv}} := T \wedge F^{\text{ent}},$$

where F^{ent} is a formula, defined below, that checks the entailment $T \cup S \models M$. Let $\mathcal{B}_1, \dots, \mathcal{B}_{2^k}$ be an enumeration of all the subsets of \mathcal{B} . Each subset implicitly denotes a truth assignment for \mathcal{B} . For each variable $v \in \mathcal{B}$ and each subset \mathcal{B}_i with $1 \leq i \leq 2^k$, we use a propositional constant $B_i(v)$ that is true if and only if $v \in \mathcal{B}_i$. Now we can define F^{ent} :

$$F^{\text{ent}} := \bigwedge_{1 \leq i \leq 2^k} \left(\left(\bigwedge_{h \in H \cap \mathcal{B}_i} (h \rightarrow B_i(h)) \right) \rightarrow F_i^{\text{ent}} \right),$$

where F_i^{ent} , defined below, checks entailment for the i -th truth assignment for \mathcal{B} . By Lemma 1 we have to ensure entailment only for those truth assignments that match the truth value of all $h \in S$. This is done via the implication $(\bigwedge_{h \in H \cap \mathcal{B}_i} (h \rightarrow B_i(h))) \rightarrow F_i^{\text{ent}}$. We define F_i^{ent} using three auxiliary formulas:

$$F_i^{\text{ent}} := (F_i^{\text{lm}} \wedge F_i^{\text{check}}) \rightarrow F_i^{\text{man}}, \quad (1)$$

where F_i^{lm} creates the least model of the HORN theory, F_i^{check} checks whether this model satisfies all constraints of the theory, and F_i^{man} checks if the model satisfies all manifestations. Next we will define F_i^{lm} . The idea behind the construction is to simulate the linear-time algorithm of Dowling and Gallier [1984], where initially all variables are set to false and then a variable is flipped from false to true if and only if it is in the head of a rule where all the variables in the rule body are true (a fact is a rule with empty body). Once a fixed-point is reached, we have obtained the least model. We encode this idea as follows. Since we are interested in the least model of $T \cup S$ instead of just T , we initialize those variables that are contained in S by setting them to true. Let $p := \min\{|T|, |V|\}$ be the maximum number of steps after which the fixed-point is reached. For each i with $1 \leq i \leq 2^k$, we introduce a new set of variables $U_i := \{u_i^j[v] \mid v \in V, 0 \leq j \leq p\}$. The intended meaning of a variable $u_i^j[v]$ is the truth value of the original variable v after the j -th step of the computation of the least model. The following auxiliary formulas encode this computation:

$$\begin{aligned} F_i^{\text{lm}} &:= \bigwedge_{v \in V, 0 \leq j \leq p} F_i^{(v,j)}; \\ F_i^{(v,0)} &:= \begin{cases} u_i^0[v] \leftrightarrow h & \text{if } v = h \in H, \\ u_i^0[v] \leftrightarrow \text{false} & \text{otherwise;} \end{cases} \\ F_i^{(v,j)} &:= u_i^j[v] \leftrightarrow \left(u_i^{j-1}[v] \vee \bigvee_{\substack{r \in \text{Rules}(T[\mathcal{B}_i]), \\ v = \text{Head}(r)}} \bigwedge_{b \in \text{Body}(r)} u_i^{j-1}[b] \right) \quad (\text{for } 1 \leq j \leq p). \end{aligned}$$

As mentioned above, we initially set the variables to false with the exception of the hypotheses. This is done in the formulas $F_i^{(v,0)}$. The computation steps are represented by the formulas $F_i^{(v,j)}$. Thereby we set a variable $u_i^j[v]$ to true if and only if it was already true in the previous step ($u_i^{j-1}[v]$), or there is a rule $r \in T[\mathcal{B}_i]$ such that $\text{Head}(r) = v$ and all body variables $b \in \text{Body}(r)$ were already true in the previous step ($u_i^{j-1}[b]$). In order to check whether the least model satisfies all the constraints (purely negative clauses), we define:

$$F_i^{\text{check}} := \bigwedge_{C \in \text{Constr}(T[\mathcal{B}_i])} \bigvee_{v \in C} \neg u_i^p[v].$$

Finally, we check whether the model satisfies all manifestations with the following formula:

$$F_i^{\text{man}} := \bigwedge_{m \in M \setminus \mathcal{B}} u_i^p[m] \wedge \bigwedge_{m \in M \cap \mathcal{B}} B_i(m).$$

It follows by Lemma 2 and by the construction of the auxiliary formulas that $F'_{\text{HORN-Solv}}$ is satisfiable if and only if $\text{Sol}(\mathcal{P}) \neq \emptyset$. Hence it remains to observe that for each $1 \leq i \leq 2^k$ the auxiliary formula F_i^{lm} can be constructed in quadratic time, and the auxiliary formulas F_i^{check} and F_i^{man} can be constructed in linear time. Therefore, formula F^{ent} can be constructed in time $\mathcal{O}(2^k n^2)$ and has size $\mathcal{O}(2^k n^2)$. \square

4 Transformations using Krom Backdoors

Recall that (in contrast to HORN formulas) a KROM formula might have several (subset) minimal models. Hence we cannot use the above approach for the base class KROM. However, we can exploit special properties of KROM formulas with respect to resolution. Analogously to Section 3 we start with an algorithm for verifying a solution and show fixed-parameter tractability with respect to the combined parameter backdoor size and number of hypotheses. Subsequently we establish the main result of this section, an fpt-reduction with backdoor size as the single parameter.

Lemma 5. *Let $\mathcal{P} = \langle V, H, M, T \rangle$ be an abduction instance and \mathcal{B} be a strong KROM-backdoor set of \mathcal{P} . A set $S \subseteq H$ is a solution to \mathcal{P} if and only if Algorithm 2 returns yes.*

Corollary 6. *Let $\mathcal{P} = \langle V, H, M, T \rangle$ be an abduction instance and \mathcal{B} be a strong KROM-backdoor set of \mathcal{P} . We can check whether \mathcal{P} has a solution in time $\mathcal{O}^*(2^{|\mathcal{B}|+|H|})$. Hence, ABD is fixed-parameter tractable when parameterized by $|\mathcal{B}| + |H|$.*

The transformation in the next theorem will use resolution in a preprocessing step, where we compute for each assignment $\tau \in \text{ta}(\mathcal{B}, S)$ the set containing all (non-tautological) resolvents restricted to variables in $H \cup M$. Computing these resolvents is possible in polynomial time for KROM formulas, since resolution on a KROM formula always yields a KROM formula. We employ the following definition and lemma.

Definition 7 (Fellows et al. [2012], Definition 17). Given an abduction instance for KROM theories $\langle V, H, M, T \rangle$. We define the function $\text{TrimRes}(T, H, M) := \{C \in \text{Res}(T) \mid C \subseteq X\}$, with $X = H \cup M \cup \{\neg x \mid x \in (H \cup M)\}$. In case $\text{TrimRes}(T, H, M)$ contains the empty clause \square , we set $\text{TrimRes}(T, H, M) := \{\square\}$.

Algorithm 2: Solution-Checker_{KROM-bds}

Input : An abduction instance $\mathcal{P} = \langle V, H, M, T \rangle$, a strong KROM-backdoor set \mathcal{B} of \mathcal{P} and a solution candidate $S \subseteq H$ to be checked.

Output : Decision whether S is a solution to \mathcal{P} .

```

1 consistent ← false
2 entailment ← true
3 foreach  $\tau \in \text{ta}(\mathcal{B}, S)$  do
4   if  $\neg$  consistent then
5     if  $T[\tau] \cup S$  is consistent then
6       consistent ← true
7   if  $M[\tau]$  is consistent then
8     foreach  $m \in M \setminus \text{var}(\tau)$  do
9       if  $(T[\tau] \not\models m) \wedge (\forall h \in S: T[\tau] \wedge h \not\models m)$  then
10        entailment ← false
11        break
12  else
13    if  $T[\tau] \cup S$  is consistent then
14      entailment ← false
15      break
16 return consistent  $\wedge$  entailment

```

Lemma 8 (Fellows *et al.* [2012], Lemma 20). *Let $\langle V, H, M, T \rangle$ be an abduction instance for KROM theories, $S \subseteq H$, $m \in M$, and $T \wedge S$ be satisfiable. Then $T \wedge S \models m$ if and only if either $\{m\} \in \text{TrimRes}(T, H, M)$ or there exists some $h \in S$ with $\{\neg h, m\} \in \text{TrimRes}(T, H, M)$.*

We extend this notion by using $\text{TrimRes}(T, H, M, \tau) := \text{TrimRes}(T[\tau], H \setminus \text{var}(\tau), M \setminus \text{var}(\tau))$.

Theorem 9. *Given an abduction instance \mathcal{P} of input size n and a strong KROM-backdoor \mathcal{B} of \mathcal{P} of size k , we can create in time $\mathcal{O}(2^k n^2)$ a CNF formula $F_{\text{KROM-Solv}}$ of size $\mathcal{O}(2^k n^2)$ such that $F_{\text{KROM-Solv}}$ is satisfiable if and only if $\text{Sol}(\mathcal{P}) \neq \emptyset$.*

Proof. By the same argument as in Theorem 4, it suffices to create first a formula $F'_{\text{KROM-Solv}}$ which is not in CNF. Formula $F'_{\text{KROM-Solv}}$ is identical to formula $F'_{\text{HORN-Solv}}$ except for the subformula F_i^{ent} , for which a completely new approach is needed.

Let $\mathcal{P} = \langle V, H, M, T \rangle$ and let $\mathcal{B}_1, \dots, \mathcal{B}_{2^k}$ be an enumeration of all the subsets of \mathcal{B} . Each subset \mathcal{B}_i implicitly defines a truth assignment τ_i of \mathcal{B} . In a preprocessing step, for each assignment τ_i the function $\text{TrimRes}(T, H, M, \tau_i)$ is computed. In order, to connect the result of $\text{TrimRes}(T, H, M, \tau_i)$ with the encoding we use logical constants which are defined as follows: Let C be a clause over $H \cup M$ or the empty clause \square . Then we define TR_i^C to be true if and only if $C \in \text{TrimRes}(T, H, M, \tau_i)$.

We obtain the formula $F'_{\text{KROM-Solv}}$ from $F'_{\text{HORN-Solv}}$ by replacing F_i^{ent} , see Equation (1), by:

$$F_i^{\text{ent}} := \left(\bigwedge_{m \in M \cap \mathcal{B}} B_i(m) \rightarrow \varphi_i^{\text{ent}} \right) \wedge \left(\bigvee_{m \in M \cap \mathcal{B}} \neg B_i(m) \rightarrow \psi_i^{\text{ent}} \right)$$

$$\varphi_i^{\text{ent}} := \bigwedge_{m \in M \setminus \mathcal{B}} \left(TR_i^{\{m\}} \vee \bigvee_{h \in H} (h \wedge TR_i^{\{h \rightarrow m\}}) \right)$$

$$\psi_i^{\text{ent}} := TR_i^{\square} \vee \bigvee_{h \in H} (h \wedge TR_i^{\{\neg h\}}) \vee \bigvee_{h_1, h_2 \in H} (h_1 \wedge h_2 \wedge TR_i^{\{\neg h_1, \neg h_2\}})$$

For each assignment τ_i (represented by $B_i(\cdot)$) we have to check whether the entailment $T[\tau_i] \cup S \models M[\tau_i]$ holds. This is done in F_i^{ent} . The proof of the correctness of this construction can be found in the full version [Pfandler *et al.*, 2013]. Note that F_i^{ent} can be constructed in quadratic time. Thus $F_{\text{KROM-Solv}}$ can be constructed in time and space of $\mathcal{O}(2^k n^2)$. \square

5 Subset Minimality

In abductive reasoning one is often more interested in subset-minimal solutions than in “ordinary” solutions [Eiter and Gottlob, 1995]. Consider the example from the preliminaries. Clearly, S_3 is a solution but actually the hypothesis *warm* is dispensable. In larger settings unnecessary hypotheses in the solution might blur the explanation. We demonstrate now how the previously presented transformations can be modified to produce only subset-minimal solutions and thus solve ABD_{\subseteq} .

Theorem 10. *Given an instance $\langle \mathcal{P}, h^* \rangle$ for ABD_{\subseteq} of size n and a strong HORN- or KROM-backdoor set \mathcal{B} of \mathcal{P} of size k , we can create in time $\mathcal{O}(2^k n^2)$ a CNF formula of size $\mathcal{O}(2^k n^2)$ that is satisfiable if and only if $\langle \mathcal{P}, h^* \rangle$ is a yes-instance for ABD_{\subseteq} .*

Proof. In order to construct a formula $F_{\text{HORN-}\subseteq\text{-Solv}}$ in case of a strong HORN-backdoor set, we modify $F'_{\text{HORN-Solv}}$ from Theorem 4. The first change is that a solution $S \subseteq H$ is no longer represented by the restriction of a model to the variables in H . Instead we introduce a new propositional variable s_h for each $h \in H$ with the intended meaning of s_h being true if and only if $h \in S$. Let V_S denote the set of these variables.

$$F'_{\text{HORN-Solv}} := \bigwedge_{h \in H} (s_h \rightarrow h) \wedge T \wedge F^{\text{ent}}.$$

The new conjunction $\bigwedge_{h \in H} (s_h \rightarrow h)$ ensures that the choice over the variables V_S is propagated to the variables representing the hypotheses. Furthermore, we change F^{ent} and $F_i^{(v,0)}$ so that these formulas use the new variables V_S :

$$F^{\text{ent}} := \bigwedge_{1 \leq i \leq 2^k} \left(\left(\bigwedge_{h \in H \cap \mathcal{B}} (s_h \rightarrow B_i(h)) \right) \rightarrow F_i^{\text{ent}} \right);$$

$$F_i^{(v,0)} := \begin{cases} u_i^0[v] \leftrightarrow s_h & \text{if } v = h \in H, \\ u_i^0[v] \leftrightarrow \text{false} & \text{otherwise.} \end{cases}$$

The other subformulas from the proof of Theorem 4 remain unchanged. We can now write $F'_{\text{HORN-}\subseteq\text{-Solv}}$ as follows.

$$F'_{\text{HORN-}\subseteq\text{-Solv}} := s_{h^*} \wedge F'_{\text{HORN-Solv}} \wedge \bigwedge_{h \in H} (s_h \rightarrow F_{\text{non-ent}}^h);$$

$$F_{\text{non-ent}}^h := \neg h^h \wedge \bigwedge_{v \in H \setminus \{h\}} (s_v \rightarrow v^h) \wedge T^h \wedge \overline{M}^h.$$

where for each $h \in H$ the formula $F_{\text{non-ent}}^h$ enforces that for $S \setminus \{h\}$ the manifestations M are no longer entailed. Thereby

we introduce for each $h \in H$ a new copy v^h of each variable $v \in V$. Let T^h and \overline{M}^h denote T and \overline{M} , respectively, where the variables are replaced by the new copies. It remains to observe that the formula $F'_{\text{HORN-}\underline{C}\text{-Solv}}$ can be constructed in time $\mathcal{O}(2^k n^2)$.

The construction of the formula $F_{\text{KROM-}\underline{C}\text{-Solv}}$ in case of a strong KROM-backdoor set is analogous. We modify the formula $F'_{\text{KROM-Solv}}$ from Theorem 9 similarly to the HORN case above. Again we use new variables s_h to decouple the solution from the hypotheses. The important change is to replace the subformula $F'_{\text{HORN-Solv}}$ by a decoupled version of $F'_{\text{KROM-Solv}}$. This decoupling is achieved by adding the clauses $\bigwedge_{h \in H} (s_h \rightarrow h)$ and replacing each occurrence of h by s_h , and h_1, h_2 by s_{h_1}, s_{h_2} accordingly. \square

6 Completeness for para-NP

A parameterized problem L is contained in the parameterized complexity class para-NP if L can be decided by a *nondeterministic fpt*-algorithm [Flum and Grohe, 2003].

For a non-parameterized problem that is NP-complete, it is considered a bad result if adding a parameter makes it para-NP-complete, since this indicates that the considered parameter does not help. But in the case of abduction, which is Σ_2^P -complete, showing that it becomes para-NP-complete is indeed a positive result. In fact we get the following result as a corollary to Theorems 4, 9, and 10.

Corollary 11. *For $\mathcal{C} \in \{\text{HORN}, \text{KROM}\}$, the problems ABD and $\text{ABD}_{\underline{C}}$ are para-NP-complete when parameterized by the size of a smallest strong \mathcal{C} -backdoor set of the given abduction instance.*

7 Enumeration and Further Extensions

In this section we sketch how the transformations presented above can be used to *enumerate* all (subset-minimal) solutions.

Obtaining a solution to the abduction instance from the models returned by the SAT solver is straightforward. For the formulas $F_{\text{HORN-}\underline{C}\text{-Solv}}$ and $F_{\text{KROM-}\underline{C}\text{-Solv}}$ (Theorem 10) it suffices to restrict the models to the variables s_h , for all $h \in H$. In order to enumerate all possible solutions, one can exclude already found solutions by adding appropriate clauses that eliminate exactly these models.

The formulas $F_{\text{HORN-Solv}}$ and $F_{\text{KROM-Solv}}$ (Theorems 4 and 9) need to be slightly modified, since a solution $S \subseteq H$ is not expressed explicitly. This is for example a problem, if an hypothesis h_1 implies another hypothesis h_2 , because no solution candidate containing only h_1 but not h_2 will be considered. This problem does not occur in the formulas $F'_{\text{HORN-}\underline{C}\text{-Solv}}$ and $F'_{\text{KROM-}\underline{C}\text{-Solv}}$, since there the encoding of a solution is decoupled from the hypotheses via $\bigwedge_{h \in H} (s_h \rightarrow h)$. The same technique can be used in the encoding of $F'_{\text{HORN-Solv}}$ and $F'_{\text{KROM-Solv}}$. Note, that several occurrences of hypotheses h have to be changed to s_h .

The transformations from Sections 3 and 4 can be extended easily to the corresponding *relevance problem*, i.e., asking whether there exists some $S \in \text{Sol}(\mathcal{P})$ containing a certain $h^* \in H$. It suffices to check only those $S \subseteq H$ where $h^* \in S$.

Corollary 12. *Let $\mathcal{C} \in \{\text{HORN}, \text{KROM}\}$. Given an abduction instance $\mathcal{P} = \langle V, H, M, T \rangle$, an atom $h^* \in H$, and a strong \mathcal{C} -backdoor set of \mathcal{P} , we can decide whether h^* belongs to some solution to \mathcal{P} in time $\mathcal{O}^*(2^{|\mathcal{B}|+|H|})$. Thus, the relevance problem is fixed-parameter tractable when parameterized by $|\mathcal{B}| + |H|$. This also holds for $\text{ABD}_{\underline{C}}$.*

Furthermore, the SAT encoding allows us to easily restrict solutions for ABD or $\text{ABD}_{\underline{C}}$ in terms of any constraints that are expressible in CNF. For example, using the encoding of a counter [Sinz, 2005], we can restrict the cardinalities of solutions and therefore solve the variants of ABD as proposed by Fellows *et al.* [2012]. In contrast, adding these constraints directly to the theory of the Abduction instance can increase the size of the backdoor set.

8 Conclusion

We have presented fixed-parameter tractable transformations from various kinds of abduction-based reasoning problems to SAT that utilize small HORN/KROM-backdoor sets in the input. These transformations are *complexity barrier breaking reductions* as they reduce problems from the second level of the Polynomial Hierarchy to the first level. A key feature of our transformations is that the exponential blowup of the target SAT instance can be confined in terms of the size of a smallest backdoor set of the input theory, a number that measures the distance to the “nice” classes of HORN and KROM formulas. There are various possibilities for further reducing the size of the target instance, which would be important for a practical implementation. For instance, one could use more sophisticated computations of the least model combined with target languages that are more compact than propositional CNF [Janhunen, 2004; Janhunen *et al.*, 2009; Thiffault *et al.*, 2004]. An extension of our approach to Abduction with other notions of solution-minimality, as surveyed by Eiter and Gottlob [1995], is left for future work. Adding empty clause detection can lead to smaller backdoors [Dilkina *et al.*, 2007] and thus making our approach applicable to a larger class of instances. While finding such backdoors is not fixed-parameter tractable [Szeider, 2009], one could use heuristics to compute them [Dilkina *et al.*, 2007].

References

- [Chen *et al.*, 2010] Jianer Chen, Iyad A. Kanj, and Ge Xia. Improved upper bounds for vertex cover. *Theoretical Computer Science*, 411(40–42):3736–3756, 2010.
- [Dilkina *et al.*, 2007] Bistra N. Dilkina, Carla P. Gomes, and Ashish Sabharwal. Tradeoffs in the complexity of backdoor detection. In *Proc. CP 2007, LNCS 4741*, pages 256–270. Springer, 2007.
- [Dowling and Gallier, 1984] William F. Dowling and Jean H. Gallier. Linear-time algorithms for testing the satisfiability of propositional Horn formulae. *J. Logic Programming*, 1(3):267–284, 1984.
- [Downey and Fellows, 1999] Rodney G. Downey and Michael R. Fellows. *Parameterized Complexity*. Springer, 1999.

- [Dvorák *et al.*, 2012] Wolfgang Dvorák, Sebastian Ordyniak, and Stefan Szeider. Augmenting tractable fragments of abstract argumentation. *Artificial Intelligence*, 186:157–173, 2012.
- [Egly *et al.*, 2000] Uwe Egly, Thomas Eiter, Hans Tompits, and Stefan Woltran. Solving advanced reasoning tasks using quantified Boolean formulas. In *Proc. AAAI 2000*, pages 417–422, 2000.
- [Eiter and Gottlob, 1995] Thomas Eiter and Georg Gottlob. The complexity of logic-based abduction. *J. of the ACM*, 42(1):3–42, 1995.
- [Fellows *et al.*, 2012] Michael R. Fellows, Andreas Pfandler, Frances A. Rosamond, and Stefan Rümmele. The parameterized complexity of abduction. In *Proc. AAAI 2012*, pages 743–749, 2012.
- [Fichte and Szeider, 2011] Johannes K. Fichte and Stefan Szeider. Backdoors to tractable answer-set programming. In *Proc. IJCAI 2011*, pages 863–868, 2011.
- [Fichte and Szeider, 2013] Johannes K. Fichte and Stefan Szeider. Backdoors to normality for disjunctive logic programs. In *Proc. AAAI 2013*, 2013. To appear. A preliminary version was presented at ASPOCP 2012.
- [Flum and Grohe, 2003] Jörg Flum and Martin Grohe. Describing parameterized complexity classes. *Inf. Comput.*, 187(2):291–319, 2003.
- [Flum and Grohe, 2006] Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Springer, 2006.
- [Gaspers and Szeider, 2012a] Serge Gaspers and Stefan Szeider. Backdoors to acyclic SAT. In *Proc. ICALP 2012, LNCS 7391*, pages 363–374. Springer, 2012.
- [Gaspers and Szeider, 2012b] Serge Gaspers and Stefan Szeider. Backdoors to satisfaction. In Hans L. Bodlaender, et al., editors, *The Multivariate Algorithmic Revolution and Beyond, LNCS 7370*, pages 287–317. Springer, 2012.
- [Gomes *et al.*, 2008] Carla P. Gomes, Henry Kautz, Ashish Sabharwal, and Bart Selman. Satisfiability solvers. In Frank van Harmelen, et al., editors, *Handbook of Knowledge Representation*, pages 89–134. Elsevier, 2008.
- [Gottlob and Szeider, 2008] Georg Gottlob and Stefan Szeider. Fixed-parameter algorithms for artificial intelligence, constraint satisfaction and database problems. *Comput. J.*, 51(3):303–325, 2008.
- [Gottlob *et al.*, 2010] Georg Gottlob, Reinhard Pichler, and Fang Wei. Bounded treewidth as a key to tractability of knowledge representation and reasoning. *Artificial Intelligence*, 174(1):105–132, 2010.
- [Janhunen *et al.*, 2009] Tomi Janhunen, Ilkka Niemelä, and Mark Sevalnev. Computing stable models via reductions to difference logic. In *Proc. LPNMR 2009, LNCS 5753*, pages 142–154. Springer, 2009.
- [Janhunen, 2004] Tomi Janhunen. Representing normal programs with clauses. In *Proc. ECAI 2004*, volume 16, pages 358–362. IOS Press, 2004.
- [Järvisalo *et al.*, 2012] Matti Järvisalo, Daniel Le Berre, Olivier Roussel, and Laurent Simon. The international SAT solver competitions. *AI Magazine*, 33(1):89–92, 2012.
- [Katebi *et al.*, 2011] Hadi Katebi, Karem A. Sakallah, and João P. Marques Silva. Empirical study of the anatomy of modern SAT solvers. In *Proc. SAT 2011, LNCS 6695*, pages 343–356. Springer, 2011.
- [Lackner and Pfandler, 2012a] Martin Lackner and Andreas Pfandler. Fixed-parameter algorithms for closed world reasoning. In *Proc. ECAI 2012*, pages 492–497, 2012.
- [Lackner and Pfandler, 2012b] Martin Lackner and Andreas Pfandler. Fixed-parameter algorithms for finding minimal models. In *Proc. KR 2012*, pages 85–95, 2012.
- [Leitsch, 1997] Alexander Leitsch. *The resolution calculus*. Texts in theoretical computer science. Springer, 1997.
- [Niedermeier and Rossmanith, 2003] Rolf Niedermeier and Peter Rossmanith. An efficient fixed-parameter algorithm for 3-hitting set. *J. Discrete Alg.*, 1(1):89–102, 2003.
- [Nishimura *et al.*, 2004] Naomi Nishimura, Prabhakar Ragde, and Stefan Szeider. Detecting backdoor sets with respect to Horn and binary clauses. In *Proc. SAT 2004*, pages 96–103, 2004.
- [Nordh and Zanuttini, 2008] Gustav Nordh and Bruno Zanuttini. What makes propositional abduction tractable. *Artificial Intelligence*, 172(10):1245–1284, 2008.
- [Pfandler *et al.*, 2013] Andreas Pfandler, Stefan Rümmele, and Stefan Szeider. Backdoors to abduction. *CoRR*, abs/1304.5961, 2013.
- [Samer and Szeider, 2009] Marko Samer and Stefan Szeider. Backdoor sets of quantified Boolean formulas. *J. Autom. Reasoning*, 42(1):77–97, 2009.
- [Selman and Levesque, 1990] Bart Selman and Hector J. Levesque. Abductive and default reasoning: A computational core. In *Proc. AAAI 1990*, pages 343–348, 1990.
- [Sinz, 2005] Carsten Sinz. Towards an optimal CNF encoding of Boolean cardinality constraints. In *Proc. CP 2005, LNCS 3709*, pages 827–831. Springer, 2005.
- [Szeider, 2009] Stefan Szeider. Matched formulas and backdoor sets. *J. on Satisfiability, Boolean Modeling and Computation*, 6:1–12, 2009.
- [Thiffault *et al.*, 2004] Christian Thiffault, Fahiem Bacchus, and Toby Walsh. Solving non-clausal formulas with DPLL search. In *Proc. CP 2004, LNCS 3258*, pages 663–678. Springer, 2004.
- [Tseitin, 1968] Grigori S. Tseitin. On the complexity of derivations in the propositional calculus. *Studies in Constructive Mathematics and Mathematical Logic*, Part II:115–125, 1968.
- [Williams *et al.*, 2003] Ryan Williams, Carla Gomes, and Bart Selman. Backdoors to typical case complexity. In *Proc. IJCAI 2003*, pages 1173–1178, 2003.