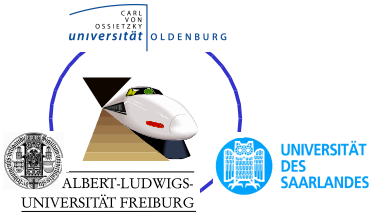


AVACS – Automatic Verification and Analysis of Complex Systems

REPORTS

of SFB/TR 14 AVACS

Editors: Board of SFB/TR 14 AVACS

Preprocessing for DQBF

by

Ralf Wimmer Karina Gitina Jennifer Nist
Christoph Scholl Bernd Becker

AVACS Technical Report No. 110
July 2015
ISSN: 1860-9821

Publisher: Sonderforschungsbereich/Transregio 14 AVACS
(Automatic Verification and Analysis of Complex Systems)
Editors: Bernd Becker, Werner Damm, Bernd Finkbeiner, Martin Fränzle,
Ernst-Rüdiger Olderog, Andreas Podelski
ATRs (AVACS Technical Reports) are freely downloadable from `www.avacs.org`

Author(s) contact: Ralf Wimmer(`wimmer@informatik.uni-freiburg.de`).

Preprocessing for DQBF*

Ralf Wimmer, Karina Gitina, Jennifer Nist, Christoph Scholl, and Bernd Becker

Albert-Ludwigs-Universität Freiburg, Germany
{wimmer | gitina | nistj | scholl | becker}@informatik.uni-freiburg.de

Abstract. For SAT and QBF formulas many techniques are applied in order to reduce/modify the number of variables and clauses of the formula, before the formula is passed to the actual solving algorithm. It is well known that these preprocessing techniques often reduce the computation time of the solver by orders of magnitude. In this paper we generalize different preprocessing techniques for SAT and QBF problems to dependency quantified Boolean formulas (DQBF) and describe how they need to be adapted to work with a DQBF solver core. We demonstrate their effectiveness both for CNF- and non-CNF-based DQBF algorithms.

1 Introduction

Many problems, practically relevant and at the same time hard from a complexity theoretic point of view, can be reduced to solving quantifier-free (SAT) or quantified (QBF) Boolean formulas. Such applications range, among many others, from verification and test of hard- and software [1, 2] to planning [3], product configuration [4], and cryptanalysis [5]. During the last three decades, the development of very efficient algorithms to solve such formulas has paved the way from academic interest to industrial application of solver techniques. SAT-formulas with hundred thousands of variables and millions of clauses can be solved nowadays, with QBF about two orders of magnitude behind.

In this paper, we consider the more general, still practically relevant formalism of *dependency quantified Boolean formulas* (DQBF). "Standard" quantified Boolean formulas (in prenex normal form) have the restriction that each existential variable depends on all universal variables in whose scope it is. This restriction is relaxed for DQBF, which allows arbitrary dependencies at the cost of a higher complexity for the decision problem – for SAT it is NP-complete [6], for QBF PSPACE-complete [7], and for DQBF it is NEXPTIME-complete [8]. However, some applications like the verification of incomplete circuits [9] or the synthesis of safe controllers [10] require the higher expressiveness of DQBF. Therefore, first solvers for DQBF have been presented recently: iDQ [11] reduces the solution of a DQBF to the solution of a series of SAT instantiations. HQS [12] applies quantifier elimination to solve the formula.

Part of the success of SAT and QBF solving is due to efficient preprocessing of the formula under consideration. The goal of preprocessing is to simplify the

* This work was partly supported by the German Research Council (DFG) as part of the Transregional Collaborative Research Center AVACS (SFB/TR 14).

formula by reducing/modifying the number of variables, clauses and quantifier alternations, such that it can be solved more efficiently afterwards. However, there is typically a trade-off between the number of variables and the number of clauses; e. g., eliminating variables by resolution can increase the number of clauses significantly, which in turn increases memory consumption and the cost of subsequent operations on the formula. Removing redundant clauses is also not always beneficial: search-based SAT and QBF solvers add implied clauses to the formula to drive the search away from unsatisfiable parts of the search space [13, 14], which often reduces computation times considerably.

For SAT and QBF, efficient and effective preprocessing tools are available like SatELite [15], Coprocessor [16] for SAT and squeezeBF [17], bloqqer [18] for QBF. Both available DQBF solvers, however, still lack a preprocessing phase before the actual solving process. Due to the success of preprocessing in SAT and QBF, one can expect that preprocessing is beneficial for DQBF, too – even more because the actual solving process is more costly than for QBF. This raises the question which techniques can be generalized from SAT and QBF to DQBF. Which adaptations need to be made to make them correct for the more general formalism? After suitable adaptations have been found, the correctness proofs have to be re-done for DQBF carefully because for QBF they often exploit the fact that dependencies in QBF follow a linear order. But also techniques like the detection of backbone literals [19, 20], which work for DQBF in the same way as for SAT and QBF, have to be re-thought: in SAT only incomplete, but cheap syntactic tests for the special case of unit literals are useful – determining backbone literals completely is as expensive as solving the SAT problem itself. For DQBF the situation is different as the decision problem is much harder. Even solving QBF approximations [9, 21] of the formula at hand as an incomplete decision procedure can be beneficial. Additionally the higher flexibility regarding the dependency sets in DQBF makes some techniques more powerful compared to QBF and enables new techniques.

Taken together, in this paper for the first time preprocessing techniques are made available for DQBF solving. We provide several extensions and adaptations and provide techniques demonstrating that preprocessing for DQBF also conceptually goes beyond standard SAT/QBF techniques. We generalize successful preprocessing techniques for QBF to DQBF like blocked clause elimination (BCE) [22, 18], equivalence reasoning [17], structure extraction [23], and variable elimination by resolution [24]. All correctness proofs are available in the appendix of this paper. We present experimental results which show the effectiveness of these techniques for DQBF. We demonstrate that the applied techniques have to be chosen depending on the solving techniques applied in the solver core. For example, BCE prevents an effective undoing of Tseitin transformation [25], which is used to transform a formula into conjunctive normal form (CNF). Therefore, it is better to disable BCE if the underlying solver core does not rely on a formula in CNF, and to use BCE if undoing Tseitin transformation is not possible because the solver core requires a formula in CNF. The experiments show that

preprocessing both reduces the computation times and significantly increases the number of solved instances of both solvers, iDQ and HQS.

Structure of this paper. The next section introduces the necessary foundations of DQBF. Section 3 reviews incomplete, but cheap decision procedures for DQBF, Section 4 describes the preprocessing techniques for DQBF that we apply in our tool to simplify the DQBF at hand. Section 5 gives an experimental evaluation of the described techniques, and Section 6 concludes the paper.

2 Preliminaries

In this section, we briefly review the necessary foundations regarding dependency quantified Boolean formulas.

Let φ, κ be quantifier-free Boolean formulas over the set V of variables and $v \in V$. We denote by $\varphi[\kappa/v]$ the Boolean formula which results from φ by replacing all occurrences of v (simultaneously) by κ. For a set $V' \subseteq V$ we denote by $\mathcal{A}(V')$ the set of Boolean assignments for V', i. e., $\mathcal{A}(V') = \left\{\nu \,\middle|\, \nu : V' \to \{0,1\}\right\}$.

Definition 1 (DQBF). *Let $V = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ be a set of Boolean variables. A* dependency quantified Boolean formula *(DQBF) ψ over V has the form $\psi := \forall x_1 \forall x_2 \ldots \forall x_n \exists y_1(D^\psi_{y_1}) \exists y_2(D^\psi_{y_2}) \ldots \exists y_m(D^\psi_{y_m}) : \varphi$ where $D^\psi_{y_i} \subseteq \{x_1, \ldots, x_n\}$ for $i = 1, \ldots, m$ is the* dependency set *of y_i, and φ is a Boolean formula over V, the* matrix *of ψ.*

To simplify the notation, we often write $\psi = Q : \varphi$ with the quantifier prefix Q and the matrix φ. Throughout the whole paper we assume, unless explicitly stated differently, that a DQBF $\psi = Q : \varphi$ as in Definition 1 with φ in CNF is given. We denote its set of universal variables by $V^\psi_\forall = \{x_1, \ldots, x_n\}$ and its set of existential variables by $V^\psi_\exists = \{y_1, \ldots, y_m\}$. If we do not need to distinguish between existential and universal variables, we write $v \in V$. $Q \setminus \{v\}$ denotes the prefix that results from removing a variable $v \in V$ from Q together with its quantifier. If v is existential, then its dependency set is removed as well; if v is universal, then all occurrences of v in the dependency sets of existential variables are removed. Similarly we use $Q \cup \left\{\exists y(D^\psi_y)\right\}$ to add existential variables to the prefix. The order in which the variables appear in the prefix is irrelevant. We introduce the dependency function $\mathrm{dep}_\psi : V \to 2^V$ by $\mathrm{dep}_\psi(v) = D^\psi_v$ if $v \in V^\psi_\exists$, and $\mathrm{dep}_\psi(v) = \{v\}$ for $v \in V^\psi_\forall$.

Definition 2 (Semantics of DQBF). *Let ψ be a DQBF with matrix φ as above. ψ is* satisfied *(written $\vDash \psi$) iff there are functions $s_{y_i} : \mathcal{A}(D^\psi_{y_i}) \to \{0,1\}$ for $1 \leq i \leq m$ such that replacing each y_i by (a Boolean expression for) s_{y_i} turns φ into a tautology. Then s_{y_i} is called a* Skolem function *for y_i.*

Two DQBFs ψ_1 and ψ_2 are *equivalent* iff $\vDash \psi_1 \;\Leftrightarrow\; \vDash \psi_2$ holds.

Definition 3 (QBF). *A* quantified Boolean formula *(QBF)*[1] *is a DQBF ψ such that $D^\psi_y \subseteq D^\psi_{y'}$ or $D^\psi_{y'} \subseteq D^\psi_y$ holds for any pair $y, y' \in V^\psi_\exists$ of existential variables.*

[1] We only consider closed QBFs in prenex form here, i. e., QBFs in which all variables are bound by a quantifer and in which the quantifiers precede the matrix.

In the following we assume that the matrix φ is given in *conjunctive normal form* (CNF). A formula is in CNF if it is a conjunction of *clauses*; a clause is a disjunction of *literals*, and a literal is either a variable v or its negation $\neg v$. We identify a formula in CNF with its set of clauses and a clause with its set of literals, e. g., we write $\big\{\{x_1, \neg x_2\}, \{x_2, \neg x_3\}\big\}$ for the formula $(x_1 \vee \neg x_2) \wedge (x_2 \vee \neg x_3)$. A clause C subsumes a clause C' iff $C \subseteq C'$. For a literal ℓ, $\mathrm{var}(\ell)$ denotes the corresponding variable, i. e., $\mathrm{var}(v) = \mathrm{var}(\neg v) = v$ and $\mathrm{dep}_\psi(\ell) = \mathrm{dep}_\psi\big(\mathrm{var}(\ell)\big)$. Moreover, we define the "sign" sgn of a literal as $\mathrm{sgn}(v) = 1$ and $\mathrm{sgn}(\neg v) = 0$.

Each DQBF can be transformed such that the matrix is in CNF. While transforming the matrix directly into CNF can cause an exponential blow-up in size, Tseitin transformation [25] can do this with only a linear increase in size at the cost of additional existential variables. The idea is to introduce auxiliary existential variables that store the truth value of sub-expressions. Since the values of these variables are uniquely determined by the sub-expression, they can simply depend on all universal variables.

We assume that none of the clauses of the CNF φ under consideration is tautological, i. e., there is no variable v such that $\{v, \neg v\} \subseteq C$ for all $C \in \varphi$. The preprocessing operations we present check the modified or added clauses whether they are tautologies and, if this is the case, remove or ignore them.

Resolution is a central operation on formulas in CNF:

Definition 4 (Resolution). *Let φ be a formula in CNF, ℓ a literal, and $C, C' \in \varphi$ clauses such that $\ell \in C$ and $\neg\ell \in C'$. The* resolvent *of C and C' w. r. t. to the* pivot literal *ℓ is given by $C \otimes_\ell C' := \big(C \setminus \{\ell\}\big) \cup \big(C' \setminus \{\neg\ell\}\big)$.*

Resolvents are implied by the formula, i. e., if R is a resolvent of two clauses in φ, then φ and $\varphi \cup \{R\}$ are equivalent [26, Sec. 3.2.1].

Currently, three solvers for DQBF have been proposed: An extension of the DPLL algorithm, typically applied for solving SAT and QBF formulas, has been described in [27]. However, no implementation thereof is available. The second solver is iDQ [11], which relies on a formula in CNF and uses instantiation-based solving, i. e., it reduces deciding a DQBF to deciding a series of SAT problems. Finally, there is the solver HQS [12], which applies quantifier elimination on And-Inverter Graphs (AIGs) to solve the formula. An AIG is essentially a circuit which consists of AND and inverter gates only. Although HQS reads the same CNF-based input format as iDQ, its back-end can handle Boolean formulas of arbitrary structure. We use both iDQ and HQS for the evaluation of the preprocessing techniques presented in the following.

3 Incomplete, but Cheap Decision Procedures

Before we present our preprocessing techniques for DQBF, we review an incomplete, but cheap decision procedure for DQBF. Experiments showed that it is beneficial to use this procedure as filter before the solving process. Our approach is as follows: First we apply preprocessing for DQBF, which is helpful for both

the filter technique and the actual solver core. Then we run the filter technique, and only if it finishes with an inconclusive result, we apply the solver core.

The filter is based on QBF approximations: By using an appropriate quantifier prefix and the same matrix, a DQBF ψ can be over-approximated by a QBF $\Psi^\uparrow$ such that the unsatisfiability of $\Psi^\uparrow$ implies the unsatisfiability of ψ [9]. Similarly one can construct an under-approximation $\Psi^\downarrow$ such that the satisfiability of $\Psi^\downarrow$ implies the satisfiability of ψ. As the under-approximation was inconclusive for all instances in our experiments, we focus on over-approximations which allow to show unsatisfiability of DQBFs. The theory for under-approximations is analogous as for over-approximations.

Definition 5 (QBF over-approximation). *A QBF $\Psi^\uparrow = Q' : \varphi$ is an* over-approximation *of ψ (written $\psi \sqsubseteq \Psi^\uparrow$) if for all existential variables $y \in V_\exists^\psi$ $D_y^{\Psi^\uparrow} \supseteq D_y^\psi$ holds,.*

Lemma 1. *Let ψ be a DQBF and $\psi \sqsubseteq \Psi^\uparrow$. If $\Psi^\uparrow$ is unsatisfiable, so is ψ.*

This lemma directly follows from the fact that Skolem functions for ψ are Skolem functions for $\Psi^\uparrow$, too.

Typically, there are several QBF over-approximations of a DQBF. They can be more or less precise. Let $\Psi_1^\uparrow = Q_1 : \varphi$ and $\Psi_2^\uparrow = Q_2 : \varphi$ be two QBF over-approximations of the same DQBF $\psi = Q : \varphi$. We call $\Psi_1^\uparrow$ *stronger* than $\Psi_2^\uparrow$ (written $\Psi_1^\uparrow \sqsubseteq \Psi_2^\uparrow$) if for all $y \in V_\exists^\psi$ we have $D_y^{\Psi_1^\uparrow} \subseteq D_y^{\Psi_2^\uparrow}$. If $\Psi_1^\uparrow \sqsubseteq \Psi_2^\uparrow$ and $\Psi_2^\uparrow$ is unsatisfiable, so is $\Psi_1^\uparrow$. A QBF over-approximation is a *strongest* QBF over-approximation if there is no different QBF over-approximation that is stronger. Strongest over-approximations are as close to the original DQBF w. r. t. the the dependency sets as possible. Therefore it is desirable to solve a strongest over-approximation as an incomplete decision procedure for DQBF.

Finkbeiner and Tentrup [21] improve this by constructing a series of more and more precise QBF formulas, starting with a strongest QBF over-approximation. To make this possible they modify both the sets of variables and the matrix of the DQBF: The idea is to use $k \geq 1$ copies of the matrix and its variables. It is required that the existential variables are assigned consistently over all copies and that all copies of the matrix are satisfied. Consistent means that if the universal variables in the dependency set of an existential variable are assigned the same values in two copies, then the existential variables have to carry the same value. This is expressed with the following formula:

$$\mathrm{Cons}(Y,k) := \bigwedge_{y \in Y} \bigwedge_{i=1}^{k} \bigwedge_{j=i+1}^{k} \left((y^i \equiv y^j) \vee \bigvee_{x \in D_y^\psi} (x^i \not\equiv x^j) \right).$$

Let $Q = \forall x_1 \ldots \forall x_n \exists y_1(D_{y_1}^{\Psi^\uparrow}) \ldots \exists y_m(D_{y_m}^{\Psi^\uparrow})$ be the prefix of a strongest QBF approximation $\Psi^\uparrow$ of the DQBF ψ. We define QBF $\Psi(k)$ for a parameter $k \geq 1$

by[2]:

$$\Psi(k) := \forall x_1^1 \ldots \forall x_1^k \forall x_2^1 \ldots \forall x_2^k \ldots \forall x_n^1 \ldots \forall x_n^k \exists y_1^1(D_{y_1}^{\Psi^\uparrow}) \ldots \exists y_1^k(D_{y_1}^{\Psi^\uparrow})$$
$$\exists y_2^1(D_{y_2}^{\Psi^\uparrow}) \ldots \exists y_2^k(D_{y_2}^{\Psi^\uparrow}) \ldots \exists y_m^1(D_{y_m}^{\Psi^\uparrow}) \ldots \exists y_m^k(D_{y_m}^{\Psi^\uparrow}) : \mathrm{Cons}\big(V_\exists^\psi, k\big) \wedge \bigwedge_{i=1}^{k} \varphi^k \,.$$

Theorem 1 ([21]). *The DQBF ψ is unsatisfiable iff $\Psi(k)$ is unsatisfiable for some $k \geq 1$.*

Experiments show that this technique can identify many unsatisfiable instances with fairly small values of k (with the majority of unsatisfiable instances identified already by $k = 1$ when $\Psi(k)$ is equal to a strongest QBF over-approximation). Since the sizes of the QBF instances grow considerably with increasing values of k, in most cases only values $k \leq 3$ seem beneficial. For more details we refer the reader to [21].

4 Preprocessing Techniques for DQBF

In this section we describe techniques which can be applied to preprocess a DQBF. The proofs of the main theorems and lemmas are given in the appendix of this paper.

4.1 Backbones, Monotonic and Equivalent Variables

Here we describe techniques which reduce both the number of variables in the formula and the number of clauses.

Unit and pure variables are well-known concepts from SAT and QBF solving. They can be replaced by constant values without influencing the formula's truth value. Typically a variable is defined as unit if the matrix contains a clause consisting only of this variable. A variable is pure if it occurs in the whole matrix either only positive or only negative:

Definition 6 (Unit and pure literals). *A literal ℓ is a* unit literal *if $\{\ell\} \in \varphi$; ℓ is a* pure literal *if $\neg\ell$ does not appear in any clause of φ.*

These are syntactic criteria that can be checked efficiently. This is necessary because in particular the detection of unit literals is one of the main operations of search-based SAT and QBF solvers as a part of Boolean constraint propagation (BCP).

For DQBF preprocessing, it is possible to use more expensive checks to determine variables which may be replaced by constants. Therefore we give a more general semantic definition:

[2] For consistency reasons we have negated the formula compared to [21].

Definition 7 (Backbones and monotonic variables). *A variable $v \in V$ is a positive (negative)* backbone *if $\varphi[0/v]$ ($\varphi[1/v]$, resp.) is unsatisfiable. A literal ℓ is a backbone, if $\ell = v$ and v a positive backbone, or if $\ell = \neg v$ and v a negative backbone.*

A variable $v \in V$ is positive (negative) monotonic *if $\varphi[0/v] \wedge \neg\varphi[1/v]$ ($\varphi[1/v] \wedge \neg\varphi[0/v]$, resp.) is unsatisfiable.*

The following theorem states how we can exploit backbones and monotonic variables to reduce the size of the formula:

Theorem 2. *Let $\psi = Q : \varphi$ be a DQBF and $v \in V$ a backbone or a monotonic variable.*

If v is a positive or negative backbone and universal, ψ is unsatisfiable. Otherwise ψ is equivalent to ψ' where

- *$\psi' = Q \setminus \{v\} : \varphi[1/v]$ if v is existential and either a positive backbone or positive monotonic, or v is universal and negative monotonic;*
- *$\psi' = Q \setminus \{v\} : \varphi[0/v]$ if v is existential and either a negative backbone or negative monotonic, or v is universal and positive monotonic.*

This theorem has been proven formally in [28]. Checks whether a variable is a backbone or monotonic can be done using a SAT solver. As already mentioned, in the SAT and QBF context typically efficient (sound but not complete) syntactic criteria are applied to detect backbones and monotonic variables. It is easy to show that unit literals are backbones and pure literals are monotonic.

Another cheap criterion to identify backbones uses the binary implication graph of a formula (which later also used to identify equivalent literals):

Definition 8. *Let $\varphi^2 = \{C \in \varphi \mid |C| = 2\}$ be the set of binary clauses. The* binary implication graph *of ψ is the directed graph $\mathrm{BIP}(\psi) = (L, E)$ with the set $L = \{v, \neg v \mid v \in V\}$ of literals as its set of nodes and $E = \{(v, \neg w), (\neg v, w) \mid \{v, w\} \in \varphi^2\}$ the set of edges.*

Then the following lemma holds:

Lemma 2. *A literal ℓ is a backbone if there is a path in $\mathrm{BIP}(\psi)$ from $\neg\ell$ to ℓ.*

If there is a path from literal ℓ to literal ℓ', we can derive the clause $\{\neg\ell, \ell'\}$ by resolution. In case of the lemma, the path from $\neg\ell$ to ℓ implies that we can derive the clause $\{\neg\neg\ell, \ell\} = \{\ell\}$. Since this is a resolvent of clauses in φ, it may be added to φ. Then we can apply Definition 6 to obtain the result.

Unit and pure literals, according to Definition 6, and backbones according to Lemma 2, can be determined efficiently by traversing the matrix or, respectively, the binary implication graph. Since solving a DQBF is much harder than solving a SAT (or even QBF) problem and the gain by eliminating one variable is larger, it often pays off to additionally use semantic checks (cf. Definition 7) for backbones and monotonic variables, which are based on solving a sequence of SAT problems. For backbones in the QBF context this observation has been made in [29].

Definition 9 (Equivalent literals). *The literals ℓ and μ are* equivalent *w. r. t. a propositional formula φ iff φ is equivalent to $\varphi \wedge (\ell \equiv \mu)$.*

Theorem 3. *Let ℓ and μ be equivalent literals. We assume, w. l. o. g., that* $\operatorname{sgn}(\ell) = 1$.

If $\operatorname{var}(\ell), \operatorname{var}(\mu) \in V_\forall^\psi$*, then ψ is unsatisfiable. Otherwise, we assume w. l. o. g. that* $\operatorname{var}(\ell) \in V_\exists^\psi$*. If* $\operatorname{var}(\mu) \in V_\forall^\psi$ *and* $\operatorname{var}(\mu) \notin D^\psi_{\operatorname{var}(\ell)}$*, then ψ is unsatisfiable. If* $\operatorname{var}(\mu) \in V_\forall^\psi$ *and* $\operatorname{var}(\mu) \in D^\psi_{\operatorname{var}(\ell)}$*, then ψ is equivalent to* $Q \setminus \{\operatorname{var}(\ell)\} : \varphi[\mu/\ell]$. *If* $\operatorname{var}(\ell), \operatorname{var}(\mu) \in V_\exists^\psi$*, then ψ is equivalent to*

$$\psi' := \big(Q \setminus \{\operatorname{var}(\mu), \operatorname{var}(\ell)\}\big) \cup \big\{\exists \operatorname{var}(\mu)(D^\psi_{\operatorname{var}(\mu)} \cap D^\psi_{\operatorname{var}(\ell)})\big\} : \varphi[\mu/\ell]\,.$$

A proof can be found in the appendix of this paper.

To detect equivalent literals, we exploit the following lemma:

Lemma 3. *Two literals ℓ, μ are equivalent if there is a path in* $\mathrm{BIP}(\psi)$ *from ℓ to μ and vice versa.*

The path from ℓ to μ allows us to derive the clause $\{\neg\ell, \mu\}$ by resolution, the path from μ to ℓ the corresponding clause $\{\neg\mu, \ell\}$. Both clauses may be added to φ as they are resolvents. Because $(\neg\ell \vee \mu) \wedge (\ell \vee \neg\mu)$ is equivalent to $(\ell \equiv \mu)$, this implies according to Definition 9 that ℓ and μ are equivalent.

We decompose $\mathrm{BIP}(\psi)$ into strongly connected components (SCCs) using Tarjan's SCC algorithm [30]. SCCs have the property that there is a path between each pair of nodes in an SCC. Therefore literals within one SCC are equivalent. They are replaced by one representative by applying Theorem 3. This procedure was described e. g., in [31–34, 16] for SAT preprocessing. Further equivalent literals can be found using structure extraction (see Section 4.5). Of course, even SAT checks based on Definition 9 may be beneficial in the DQBF context.

4.2 Reduction of Dependency Sets

In a DQBF, a universal variable $x \in V_\forall^\psi$ may be contained in the dependency set D_y^ψ of an existential variable $y \in V_\exists^\psi$, but actually, due to the structure of the matrix, the Skolem function for y does not need to exploit the information about x's value to satisfy the formula. If such a situation is detected, x can be removed from D_y^ψ. This potentially reduces the number of copies of variables, if universal expansion according to Theorem 6 is used for solving a DQBF.

An example for a situation when dependency sets may be reduced is when a circuit is transformed into CNF by Tseitin transformation. The dependency set D_y^ψ of a Tseitin variable y can be an arbitrary superset of the universal variables in its cone-of-influence. The variables in D_y^ψ that are not in the cone-of-influence of y can be removed from D_y^ψ without affecting the truth value of the formula.

Definition 10. *An existential variable $y \in V_\exists^\psi$ is* independent *of a universal variable $x \in V_\forall^\psi$ if either $x \notin D_y^\psi$ or replacing D_y^ψ by $D_y^\psi \setminus \{x\}$ does not change the truth value of ψ.*

Deciding whether two variables are independent has the same complexity as deciding the DQBF itself [35]. Therefore one resorts to sufficient criteria to show independence. The most simple ones are based on the incidence graph of the matrix:

The *variable-clause incidence graph* $G_{V,\varphi} = (V \cup \varphi, E)$ of the formula is an undirected graph with $E = \big\{\{v, C\} \in V \times \varphi \,\big|\, v \in C \vee \neg v \in C\big\}$.

Theorem 4 (Standard dependency scheme). *An existential variable $y \in V_\exists^\psi$ is independent of a universal variable $x \in V_\forall^\psi$ if there is no path in $G_{V,\varphi}$ from x to y, visiting only variables in $\{z \in V_\exists^\psi \,|\, x \in D_z^\psi\}$ in between.*

For a proof for this theorem, which generalizes a theorem from [35], see the appendix of this paper.

In the QBF context more powerful dependency schemes have been developed which can possibly identify more variables as independent, see, e. g., [35–39]. A generalization of these techniques will have an immediate benefit for DQBF solving by increasing the potential to save variable copies during universal expansion.

4.3 Universal Reduction, Resolution, and Universal Expansion

Universal reduction, *resolution*, and *universal expansion* are well-known techniques used during the solution of QBFs. Universal reduction removes a universal variable from a clause if the clause does not contain any existential variable which depends upon it. This technique has already been generalized to DQBF in [40, 11].

Lemma 4 (Universal reduction, [40, 11]). *Let $Q : \varphi \wedge C$ be a DQBF and $\ell \in C$ a universal literal such that for all $k \in C$ with $k \neq \ell$ we have $\mathrm{var}(\ell) \notin \mathrm{dep}_\psi(k)$. Then $Q : \varphi \wedge C$ and $Q : \varphi \wedge \big(C \setminus \{\ell\}\big)$ are equivalent.*

For QBF resolution and universal reduction together are able to derive the empty clause iff the formula is unsatisfiable. This does not hold for DQBF [40]. Resolution in QBF formulas allows to eliminate an existential variable by replacing the clauses containing this variable with their resolvents. While adding resolvents is sound for DQBF as well, eliminating existential variables by resolution [15] only works under certain conditions. Here we give a set of sufficient conditions which allow variable elimination by resolution for DQBF. In particular when the formula is created by Tseitin transformation [25], variable elimination by resolution is applicable to a large subset of the formula's existential variables.

Theorem 5 (Variable elimination by resolution). *Let $y \in V_\exists^\psi$ be an existential variable of ψ. We partition φ into the sets $\varphi^y = \{C \in \varphi \,|\, y \in C\}$, $\varphi^{\neg y} = \{C \in \varphi \,|\, \neg y \in C\}$, and $\varphi^\emptyset = \varphi \setminus (C^y \cup C^{\neg y})$.*
If one of the following conditions is satisfied:

- *for all $C \in \varphi^y$ and all $k \in C$ we have $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(y)$,*
- *for all $C' \in \varphi^{\neg y}$ and all $k \in C'$ we have $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(y)$, or*
- *y is the defined variable of a functional definition, i. e., there are clauses encoding the relationship $y \equiv f(V')$ for some function f and arguments $V' \subseteq V \setminus \{y\}$, $\mathrm{dep}_\psi(v) \subseteq \mathrm{dep}_\psi(y)$ for all $v \in V'$ (cf. Sec. 4.5),*

then ψ *is equivalent to* $\psi' := Q \setminus \{y\} : \varphi^{\emptyset} \wedge \bigwedge_{C \in \varphi^y} \bigwedge_{C' \in \varphi^{\neg y}} C \otimes_y C'$.

Proof sketch. Resolvents are implied by the matrix, i. e., adding resolvents to the matrix yields an equivalent formula. If ψ is satisfied, then removing the clauses in φ^y and $\varphi^{\neg y}$ cannot make the formula unsatisfied, i. e., ψ' is satisfied.

Assume that ψ' is satisfied by Skolem functions s_z for $z \in V_\exists^\psi \setminus \{y\}$. We define $s_y := \neg\varphi^y\big[0/y\big]\big[s_z/z \text{ for } z \in V_\exists^\psi \setminus \{y\}\big]$ in the first case, $s_y := \neg\varphi^{\neg y}\big[1/y\big]\big[s_z/z \text{ for } z \in V_\exists^\psi \setminus \{y\}\big]$ in the second case, and $s_y := f(V')\big[s_z/z \text{ for } z \in V_\exists^\psi \setminus \{y\}\big]$ in the third case. It is not hard to show that s_y is an admissible Skolem function for y and that $\varphi[s_v/v \text{ for } v \in V_\exists^\psi]$ is indeed a tautology. Details can be found in the appendix. □

Theorem 5 does not provide a decision algorithm for arbitrary DQBFs, since it is possible that the conditions do not hold for any existential variable. Moreover, eliminating all existential variables fulfilling the conditions of Theorem 5 is in general not feasible because the number of clauses can grow considerably during elimination. We first create a list of variables that may be eliminated. For each such variable y we estimate the cost c_y of elimination, i. e., $c_y := \frac{|\varphi^{\emptyset}|+|\varphi^y|\cdot|\varphi^{\neg y}|}{|\varphi|}$. We eliminate one variable y with minimum cost provided that c_y is less than a user-specified factor $\varepsilon > 1$. After resolving variables we check for subsumed clauses, i. e., clauses C such that there is a clause C' with $C' \subseteq C$. Then C can be deleted [26, Sec. 3.2].

Universal expansion [41, 42, 40, 9] is the corresponding method for eliminating universal variables. It is the main operation which the solver HQS [12] uses to transform the DQBF at hand into an equivalent QBF. This QBF can be solved by an arbitrary QBF solver.

Theorem 6 (Universal expansion). *Let* $x_i \in V_\forall^\psi$, *and* $E_{x_i}^\psi = \big\{y_i \in V_\exists^\psi \,\big|\, x_i \in \mathrm{dep}_\psi(y_j)\big\}$. *Then* ψ *is equivalent to*

$$\big(Q\setminus\{x_i\}\big)\cup\big\{\exists y_j'(D_{y_j}^\psi\setminus\{x_i\}) \,\big|\, y_j \in E_{x_i}^\psi\big\} : \varphi[1/x_i]\wedge\varphi[0/x_i][y_j'/y_j \text{ for all } y_j \in E_{x_i}^\psi]\,.$$

A formal proof of this theorem is given, e. g., in [9]. In order to avoid unnecessary variable copies, we check using the standard dependency scheme (cf. Theorem 4) which existential variables actually depend on the expanded universal variable.

4.4 Blocked Clause Elimination

The concept of blocked clauses was introduced by Järvisalo et al. for SAT in [22] and later generalized to QBF by Biere et al. in [18]. Blocked clauses can be removed from a formula without changing its truth value. Before checking for blockedness, clauses can be extended by so-called hidden and covered literals [43, 44, 18]. This does not change the truth value of the formula, but increases the chance that a clause is blocked.

In this section, we first generalize the notion of blocked clauses to DQBF such that blocked clauses satisfy the same properties as in SAT and QBF. Then we investigate how to generalize hidden and covered literals to DQBF.

For a *QBF* $Q : \varphi \wedge C$, a clause C containing an existential literal $\ell \in C$ can be omitted (resulting in an equivalent formula), if 'ℓ is blocking for C', which means that for all $C' \in \varphi$ with $\neg\ell \in C'$ there is a variable k such that $\{k, \neg k\} \subseteq C \otimes_\ell C'$ and k precedes ℓ in the quantifier prefix (which means in DQBF notions: $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$). In the QBF context the intuitive background of blocked clause elimination is simple: Consider a solving approach to QBF which always removes the innermost existential quantifiers (which depend on all universal ones) by resolution[3] and the innermost universal quantifiers (upon which no existential variable depends) by universal reduction until all quantifiers have been removed [24]. If ℓ is blocking for C, all resolvents resulting from C contain $\{k, \neg k\}$, i. e., are tautological, and their addition makes no contribution. The condition 'k precedes ℓ in the quantifier prefix' ensures that $\mathrm{var}(k)$ has not been removed before ℓ in the process sketched above, i. e., the reason $\{k, \neg k\}$ for the resolvents being tautological has not been removed. This implies that we can alternatively remove C from $\varphi \wedge C$ in the very beginning without changing the result of the solving process.

Fortunately, we can show that the notion of blocked clauses has a natural generalization to DQBF. However, the proof idea of blocked clause elimination sketched above does not work anymore, since in DQBF there is no linear order for the quantifiers such that 'removing quantifiers starting with the innermost' does not have a counterpart in DQBF; the correctness proof has to be re-done for DQBF carefully taking into account that arbitrary dependencies may be defined in a DQBF. We first give the generalized definition of blocked clauses:

Definition 11 (Blocked clauses). *Let $Q : \varphi \wedge C$ be a DQBF and C a clause with $\ell \in C$. Literal ℓ is a* blocking literal *for C if ℓ is existential and for all $C' \in \varphi$ with $\neg\ell \in C'$ there is a variable k such that $\{k, \neg k\} \subseteq C \otimes_\ell C'$ and $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$. A clause is* blocked *if it contains a blocking literal.*

Now we can prove results that are analogous to QBF and SAT.

Theorem 7 (Blocked clause elimination, BCE). *Let $Q : \varphi \wedge C$ be a DQBF with a blocked clause C. Then $Q : \varphi \wedge C$ and $Q : \varphi$ are equivalent.*

Proof sketch. The theorem can be shown by induction on the number $|\mathrm{dep}_\psi(\ell)|$ of ℓ's dependencies. The base case $\mathrm{dep}_\psi(\ell) = \emptyset$ works analogously to the QBF case, see [18]. For the induction step, we choose an arbitrary universal variable $x \in \mathrm{dep}_\psi(\ell)$ and eliminate it by universal expansion (see Theorem 6). In the resulting formula, ℓ and its copy ℓ' depend on one variable less. One can show that both copies of C in this formula are either blocked or tautological. Therefore they can be removed by the induction assumption. Un-doing the expansion step yields the result. A more detailed proof can be found in the appendix. □

[3] Adding all possible resolvents with pivot variable v and then removing all clauses containing v or $\neg v$ corresponds to existential quantification of v.

Lemma 5. *BCE for DQBF has a unique fixed point.*

That means that the result of elimination does not depend on the order in which the clauses are considered.

The purpose of the following techniques is to extend clauses by redundant literals. This increases the chance that the clause is blocked and can be deleted. If the extended clause is not blocked, the additional literals are removed again.

Definition 12 (Hidden literals). *Let $Q : \varphi \wedge C$ be a DQBF. A literal $\ell \notin C$ is a* hidden literal *for C if there is a clause $\{\ell_1, \ldots, \ell_n, \neg\ell\} \in \varphi$ such that $\{\ell_1, \ldots, \ell_n\} \subseteq C$.*

Theorem 8 (Hidden literal addition, HLA). *Let $Q : \varphi \wedge C$ be a DQBF and ℓ a hidden literal for C. Then $Q : \varphi \wedge C$ and $Q : \varphi \wedge \big(C \cup \{\ell\}\big)$ are equivalent.*

The idea of hidden literal addition is based on *self-subsuming resolution* [15]. The resolvent $(C \cup \{\ell\}) \otimes_\ell \{\ell_1, \ldots, \ell_n, \neg\ell\}$ is equal to C and subsumes $C \cup \{\ell\}$. Thus after adding the resolvent C, $C \cup \{\ell\}$ can be removed, leading to an equivalent formula. Note that the argument for hidden literal addition is based on a consideration of the matrix only, thus in this case the argumentation is exactly the same as for SAT and QBF.

This is in contrast to the 'covered literal addition' described in the following. For covered literals we need a careful generalization of the QBF definition together with a non-trivial proof of the generalization to DQBF.

Definition 13 (Covered literals). *Let $\psi = Q : \varphi \wedge C$ be a DQBF and let ℓ be an existential literal with $\ell \in C$. The set of* resolution candidates *for C w. r. t. ℓ is the set $R_\psi(C, \ell) = \big\{C' \in \varphi \,\big|\, \neg\ell \in C' \wedge \forall v \in V : (\{v, \neg v\} \subseteq C \otimes_\ell C' \;\Rightarrow \mathrm{dep}_\psi(v) \not\subseteq \mathrm{dep}_\psi(\ell))\big\}$.*

A literal k is a covered literal *for C w. r. t. ℓ if $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$ and $k \in \bigcap R_\psi(C, \ell) \setminus \{\neg\ell\}$.*

Theorem 9 (Covered literal addition, CLA). *Let $Q : \varphi \wedge C$ be DQBF and k a covered literal for C. Then $Q : \varphi \wedge C$ and $Q : \varphi \wedge \big(C \cup \{k\}\big)$ are equivalent.*

Proof sketch. Assume that k is a covered literal for C w. r. t. ℓ. We show the theorem by induction on the number $|\mathrm{dep}_\psi(\ell)|$ of dependencies of ℓ. The induction base where $\mathrm{dep}_\psi(\ell) = \emptyset$ is similar to the QBF case (cf. [18]). For the induction step, we apply universal expansion of an arbitrary variable in $\mathrm{dep}_\psi(\ell)$ (see Theorem 6) to obtain a formula in which ℓ and its copy ℓ' both depend on one variable less. It is rather technical to show that adding k (k') to the copies of C in this formula leads to an equivalent formula, since these copies are either tautological or k (k') is a covered literal. By undoing the expansion step we obtain the desired result. For a detailed proof we refer to the appendix. □

A rough basic intuition for covered literal addition is as follows: "If a literal k is already contained in all non-tautological resolvents of a clause C with pivot literal ℓ, then k may be added to C resulting in an equivalent formula." In

addition to this basic idea we need the condition $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$ and a bigger set of resolution candidates $R_\psi(C,\ell) = \left\{C' \in \varphi \,\middle|\, \neg\ell \in C' \wedge \forall v \in V : (\{v, \neg v\} \subseteq C \otimes_\ell C' \;\Rightarrow\; \mathrm{dep}_\psi(v) \not\subseteq \mathrm{dep}_\psi(\ell))\right\}$ instead of $R_\psi(C,\ell) = \left\{C' \in \varphi \,\middle|\, \neg\ell \in C' \wedge \nexists v \in V : \{v, \neg v\} \subseteq C \otimes_\ell C'\right\}$ in order to be able to lead the (rather involved) proof of Theorem 9, see the appendix.

In order to reduce the size of the formula, we determine for each clause C the set H of hidden and the set K of covered literals. Then we check if $C \cup H \cup K$ is blocked or tautological. If this is the case, C is removed; otherwise C remains unchanged. This is iterated until we reach a fixed point.

Note that if a hidden or covered literal is universal, its addition can be helpful not only because it can make a clause blocked. If a CNF-based solver core uses elimination of universal variables to decide the formula, all clauses which contain an existential variable that depends on the eliminated universal variable have to be doubled [9]. If the clause contains the universal variable to be eliminated, one of these copies is satisfied and can therefore be omitted (cf. [45]).

4.5 Structure Extraction

The DQBF's matrix in CNF is often created from a circuit or a Boolean expression by Tseitin transformation [25], where a new existential variable v_e is created for each sub-expression e (or gate output). Clauses encoding the relationship $v_e \equiv e$ are added and the sub-expression e is replaced by the variable v_e. If a solver (like HQS) does not rely on a matrix in CNF, this transformation step can be undone. This removes all artificially introduced variables. Structure extraction is used in the QBF solver AIGsolve [23].

For example, a k-input AND gate $y \equiv \mathrm{AND}(\ell_1, \ldots, \ell_k)$ has a Tseitin encoding consisting of $(k+1)$ clauses $\{\neg y, \ell_1\}, \ldots, \{\neg y, \ell_k\}, \{y, \neg\ell_1, \ldots, \neg\ell_k\}$. In a functional definition $y \equiv f(\ell_1, \ldots, \ell_k)$, y is called the *defined variable*, f is the *definition* of y, and the clauses corresponding to the relationship $y \equiv f(\ell_1, \ldots, \ell_k)$ are the *defining clauses*.

Theorem 10. *Let $\psi = Q : \varphi$ be a DQBF and $\varphi^f \subseteq \varphi$ the defining clauses for the relationship $y \equiv f(\ell_1, \ldots, \ell_k)$. Then ψ is equivalent to $Q \setminus \{y\} : (\varphi \setminus \varphi^f)[f(\ell_1, \ldots, \ell_k)/y]$ if $y \in V_\exists^\psi$ and for $i = 1, \ldots, k$ we have $\mathrm{dep}_\psi(\ell_i) \subseteq \mathrm{dep}_\psi(y)$.*

Our implementation checks for defining clauses for (multi-input) (N)AND gates and 2-input XOR gates, both with arbitrarily negated inputs. We do not extract definitions that lead to cyclic dependencies.

Gate detection can be used as the last step of the preprocessing routine. If a relationship $y \equiv f(\ell_1, \ldots, \ell_k)$ is detected which does not lead to cyclic dependencies, we remove y from the prefix and the defining clauses from the matrix. We additionally use a data structure which assigns to each defined variable its definition. To create an AIG representation that can be passed to a non-CNF-based solver core like HQS, we convert the remaining clause into an AIG and then substitute the defined variables by their definitions.

The same structure extraction procedure can also be used to identify equivalent variables and unnecessary variable dependencies. For both purposes, the relationships are only detected, but neither are the defining clauses removed nor is the data structure that stores the relationships updated. Therefore this can also be used if the solver back-end requires a matrix in CNF: If there is the relationship $y \equiv f(\ell_1, \ldots, \ell_k)$ and $\bigcup_{i=1}^{k} \mathrm{dep}_\psi(\ell_i) \subsetneq D_y^\psi$, then D_y^ψ can be replaced by $\bigcup_{i=1}^{k} \mathrm{dep}_\psi(\ell_i)$. If two defined variables y, y' with the same definition are detected, i. e., $y \equiv f(\ell_1, \ldots, \ell_k)$ and $y' \equiv f(\ell_1, \ldots, \ell_k)$, then y and y' are equivalent and Theorem 3 can be applied to remove one of them.

5 Experimental Results

We have implemented the described techniques in C++ as a preprocessor for our DQBF solver HQS. To support other back-end solvers, too, it is able to write the resulting formula into a file in DQDIMACS format, which can be read by the currently only competing solver iDQ [11].

As benchmark instances we use 4381 formulas, resulting from the verification of incomplete circuits [9, 21, 11] and controller synthesis [10]. The synthesis benchmarks are those shipped with the tool *Demiurge 1.1.0* [10]. We used the encoding described in [10] to create a DQBF formulation.

All experiments were run on one Intel Xeon E5-2650v2 core at 2.60 GHz with 64 GB of main memory, running Ubuntu Linux 12.04 in 64-bit mode as operating system. We aborted all experiments whose computation time exceeded 900 seconds or which required more than 8 GB of memory. For solving QBFs, we use DepQBF 4.0 [46, 47] with the QBF preprocessor bloqqer [18] (version 35) if the matrix is in CNF, and AIGsolve [23] if the matrix is given as an AIG.

We used two parameter settings for preprocessing, in the following called V_1 and V_2. Both use the detection of backbones (by syntactic and semantic checks), monotonic variables (by syntactic checks), and equivalent variables (both using the binary implication graph and structure extraction). We reduce the dependency sets of the existential variables using the standard dependency scheme and structure extraction. For these operations, the functional definitions are only detected, but neither are the defined variables replaced by their definition nor are the defining clauses removed.

- V_1 additionally enables structure extraction, which replaces the defined variables by their definitions. V_1 does not yield a CNF representation, but rather an And-Inverter Graph (AIG) [48] for the formula. Since iDQ requires a CNF representation of the matrix, V_1 can only be combined with HQS.
- V_2 applies BCE after adding hidden and covered literals and variable elimination by resolution ($\varepsilon = 1.1$), but disables structure extraction. V_2 yields a matrix in CNF; therefore it can be combined with both iDQ and HQS.

Table 1 shows the number of *solved instances* (out of 4381) for different combinations of preprocessing, filtering using QBF over-approximations see Section 3), and the HQS or iDQ solver cores. Preprocessing alone can only solve a small fraction

of all instances (80 for V_1 and 57 for V_2). The filter solves already 935 instances for $k = 1$ (slightly more with higher values of k). The combination of preprocessing V_1 with the filter allows to decide 2459 instances (2240 with V_2). In spite of using bloqqer as preprocessor for simplifying the QBF over-approximations for filtering, doing DQBF preprocessing before reduces the solving times for the QBFs. Without DQBF preprocessing, solving the QBF approximation runs into a timeout frequently.

For HQS as solver back-end, the trend is similar: without preprocessing and filtering, HQS is able to solve 1537 instances, with V_1 preprocessing this number increases to 3629 instances, and if filtering is used thereafter, 3752 instances can be solved. We can also see that BCE largely prevents structure extraction: if all described techniques are enabled, only 2174 instances can be solved successfully. Increasing the value of k to 2 does not seem beneficial at least if a time limit of 15 min is used. For larger time limit, $k = 2$ can slightly increase the number of solved instances. Finally, if we combine V_2 with filtering ($k = 1$) and HQS, we can also observe a positive effect on the number of solved instances (3542); however, it is not as strong as with V_1, which includes structure extraction instead of BCE.

Table 1. Effect of preprocessing

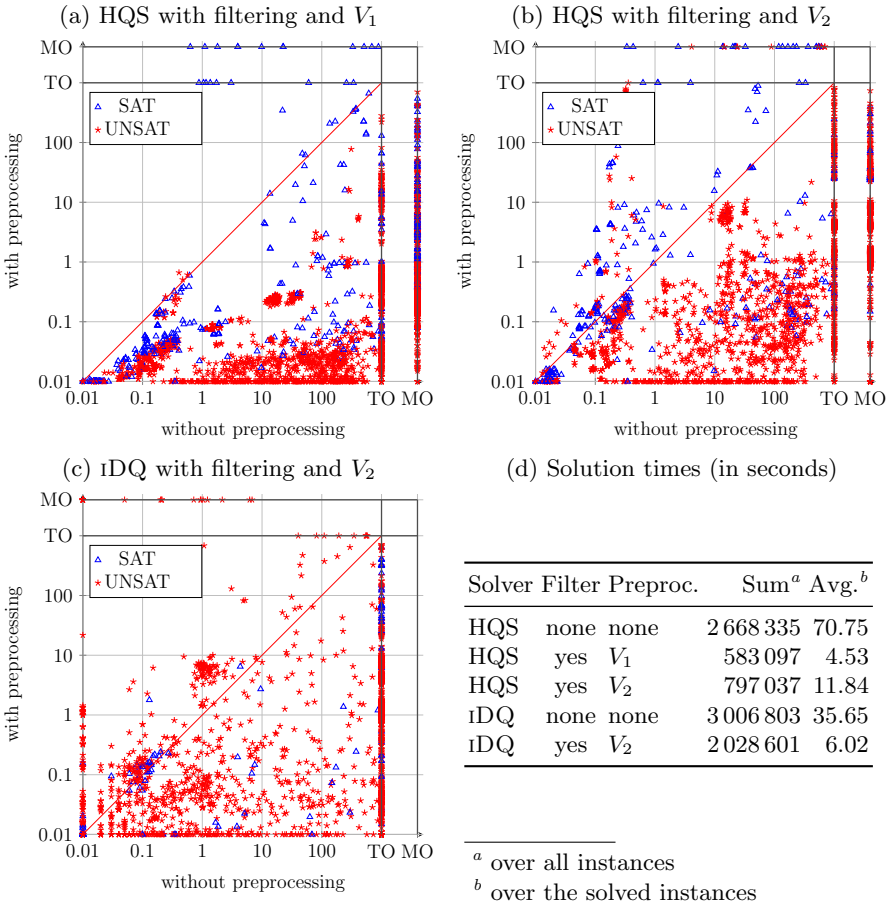
Solver	Filter	Preproc.	Solved
none	$k = 1$	none	935
none	none	V_1	80
none	none	V_2	57
none	$k = 1$	V_1	2459
none	$k = 2$	V_1	2733
none	$k = 1$	V_2	2240
HQS	none	none	1537
HQS	none	V_1	3629
HQS	$k = 1$	V_1	3752
HQS	$k = 1$	V_1 + BCE	2174
HQS	$k = 2$	V_1	3737
HQS	$k = 1$	V_2	3542
iDQ	none	none	1073
iDQ	none	V_2	1378
iDQ	$k = 1$	none	1359
iDQ	$k = 1$	V_2	2714

iDQ without filtering and preprocessing solves 1073 instances. This number is increased to 1378 by preprocessing (V_2) and to 1359 instances by filtering ($k = 1$). The combination with filtering and preprocessing yields 2714 solved instances.

In summary, the combination of filtering and preprocessing significantly increases the number of solved instances by a factor of up to 2.44 (for HQS) and 2.52 (for iDQ). The best results are obtained if the preprocessing techniques are chosen according to the solver core.

Now we focus on the *size of the instances* before and after preprocessing. Preprocessing variant V_2 reduces the number of clauses by 64 % on average, the number of existential variables by 76 % on average, but leaving the number of universal variables essentially unchanged. As preprocessing variant V_1 does not yield a CNF representation, we cannot compare the number of clauses. Instead we compare the size of the AIG representation of the matrix before and after preprocessing. V_1 reduces the number of existential variables by 97 % on average (including all Tseitin variables), the number of AIG nodes by 84 %, leaving the number of universal variables almost unchanged, too.

If the CNF structure of the matrix needs to be preserved (as in V_2) not all Tseitin variables can be removed by identifying functional definitions and by elimination by resolution, since this leads to a significant increase in size of the



Solver	Filter	Preproc.	Sum[a]	Avg.[b]
HQS	none	none	2 668 335	70.75
HQS	yes	V_1	583 097	4.53
HQS	yes	V_2	797 037	11.84
iDQ	none	none	3 006 803	35.65
iDQ	yes	V_2	2 028 601	6.02

[a] over all instances
[b] over the solved instances

Fig. 1. Running times (in seconds) for HQS and iDQ with and without preprocessing.

CNF (at least for some intermediate instances). This effect is lessened by BCE, in particular if HLA and CLA are enabled.

Finally, we take a closer look at the *solving times* of the instances. For the instances which were solved with or without preprocessing and filtering, Fig. 1 compares the computation times when using only the solver core and when using the solver core after preprocessing and filtering. The times include everything from reading the input files to termination. The upper two pictures show HQS with V_1 (Fig. 1(a)) and V_2 (Fig. 1(b)) and filtering using $k = 1$, compared to HQS without preprocessing and filtering. Fig. 1(c) shows iDQ with V_2, compared to iDQ without preprocessing. In Fig. 1(d) we present the accumulated running times over all instances (unsolved instances contributing the time limit of 900 seconds) and the average running time of the solved instances.

In all three cases, preprocessing and filtering reduce the computation times for the vast majority of instances significantly, often by orders of magnitude. The very few exceptions in case of iDQ are instances that are very easy to solve such that the overhead for preprocessing exceeds the solving time. We can also observe that many instances, for which the solver core alone ran into a time out or memory out, can be solved successfully after preprocessing and filtering.

6 Conclusion

We have shown how preprocessing techniques for SAT and QBF can be generalized to DQBF. Experiments have demonstrated that they can reduce the running time of the actual solving process by orders of magnitude, both for CNF-based and non-CNF-based solver cores.

In future we want to investigate more powerful dependency schemes and how the flexibility in the dependency sets can be exploited when choosing sets of universal variables to eliminate in order to obtain a QBF.

Acknowledgements. We thank Sven Reimer for fruitful discussions and Florian Pigorsch for providing us with his AIGsolve [29] implementation.

References

1. Biere, A., Cimatti, A., Clarke, E.M., Strichman, O., Zhu, Y.: Bounded model checking. Advances in Computers **58** (2003) 117–148
2. Czutro, A., Polian, I., Lewis, M.D.T., Engelke, P., Reddy, S.M., Becker, B.: TIGUAN: thread-parallel integrated test pattern generator utilizing satisfiability analysis. In: Int'l Conf. on VLSI Design, New Delhi, India, IEEE Computer Society (2009) 227–232
3. Rintanen, J.: Constructing conditional plans by a theorem-prover. Journal of Artificial Intelligence Research **10** (1999) 323–352
4. Sinz, C., Kaiser, A., Küchlin, W.: Formal methods for the validation of automotive product configuration data. AI EDAM **17**(1) (2003) 75–97
5. Mironov, I., Zhang, L.: Applications of SAT solvers to cryptanalysis of hash functions. In Biere, A., Gomes, C.P., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 4121 of LNCS, Seattle, WA, USA, Springer (2006) 102–115
6. Cook, S.A.: The complexity of theorem-proving procedures. In: Annual ACM Symposium on Theory of Computing (STOC), ACM Press (1971) 151–158
7. Meyer, A.R., Stockmeyer, L.J.: Word problems requiring exponential time: Preliminary report. In: Annual ACM Symposium on Theory of Computing (STOC), ACM Press (1973) 1–9
8. Peterson, G., Reif, J., Azhar, S.: Lower bounds for multiplayer non-cooperative games of incomplete information. Computers & Mathematics with Applications **41**(7–8) (2001) 957–992
9. Gitina, K., Reimer, S., Sauer, M., Wimmer, R., Scholl, C., Becker, B.: Equivalence checking of partial designs using dependency quantified Boolean formulae. In: IEEE Int'l Conf. on Computer Design (ICCD), Asheville, NC, USA, IEEE Computer Society (2013) 396–403

10. Bloem, R., Könighofer, R., Seidl, M.: SAT-based synthesis methods for safety specs. In McMillan, K.L., Rival, X., eds.: Int'l Conf. on Verification, Model Checking, and Abstract Interpretation (VMCAI). Vol. 8318 of LNCS, San Diego, CA, USA, Springer (2014) 1–20
11. Fröhlich, A., Kovásznai, G., Biere, A., Veith, H.: iDQ: Instantiation-based DQBF solving. In Berre, D.L., ed.: Int'l Workshop on Pragmatics of SAT (POS). Vol. 27 of EPiC Series, Vienna, Austria, EasyChair (2014) 103–116
12. Gitina, K., Wimmer, R., Reimer, S., Sauer, M., Scholl, C., Becker, B.: Solving DQBF through quantifier elimination. In: Int'l Conf. on Design, Automation & Test in Europe (DATE), Grenoble, France, IEEE (2015)
13. Jr., R.J.B., Schrag, R.: Using CSP look-back techniques to solve real-world SAT instances. In Kuipers, B., Webber, B.L., eds.: National Conference on Artificial Intelligence / Innovative Applications of Artificial Intelligence Conference (AAAI/IAAI), Providence, Rhode Island, USA, AAAI Press / The MIT Press (1997) 203–208
14. Silva, J.P.M., Sakallah, K.A.: GRASP: A search algorithm for propositional satisfiability. IEEE Transactions on Computers **48**(5) (1999) 506–521
15. Eén, N., Biere, A.: Effective preprocessing in SAT through variable and clause elimination. In Bacchus, F., Walsh, T., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 3569 of LNCS, St. Andrews, UK, Springer (2005) 61–75
16. Manthey, N.: Coprocessor 2.0 – A flexible CNF simplifier – (tool presentation). In Cimatti, A., Sebastiani, R., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 7317 of LNCS, Trento, Italy, Springer (2012) 436–441
17. Giunchiglia, E., Marin, P., Narizzano, M.: sQueezeBF: An effective preprocessor for QBFs based on equivalence reasoning. In Strichman, O., Szeider, S., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 6175 of LNCS, Edinburgh, UK, Springer (2010) 85–98
18. Biere, A., Lonsing, F., Seidl, M.: Blocked clause elimination for QBF. In Bjørner, N., Sofronie-Stokkermans, V., eds.: Int'l Conf. on Automated Deduction (CADE). Vol. 6803 of LNCS, Springer (2011) 101–115
19. Kilby, P., Slaney, J.K., Thiébaux, S., Walsh, T.: Backbones and backdoors in satisfiability. In Veloso, M.M., Kambhampati, S., eds.: National Conference on Artificial Intelligence / Int'l Conf. on Innovative Applications of Artificial Intelligence (IAAI), Pittsburgh, Pennsylvania, USA, AAAI Press / The MIT Press (2005) 1368–1373
20. Janota, M., Lynce, I., Marques-Silva, J.: Algorithms for computing backbones of propositional formulae. AI Communications **28**(2) (2015) 161–177
21. Finkbeiner, B., Tentrup, L.: Fast DQBF refutation. In: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 8561 of LNCS, Springer (2014) 243–251
22. Järvisalo, M., Biere, A., Heule, M.: Blocked clause elimination. In Esparza, J., Majumdar, R., eds.: Int'l Conf. on Tools and Algorithms for the Construction and Analysis of Systems (TACAS). Vol. 6015 of LNCS, Springer (2010) 129–144
23. Pigorsch, F., Scholl, C.: Exploiting structure in an AIG based QBF solver. In: Int'l Conf. on Design, Automation & Test in Europe (DATE), Nice, France, IEEE (2009) 1596–1601
24. Biere, A.: Resolve and expand. In Hoos, H.H., Mitchell, D.G., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 3542 of LNCS, Vancouver, BC, Canada, Springer (2004) 59–70

25. Tseitin, G.S.: On the complexity of derivation in propositional calculus. Studies in Constructive Mathematics and Mathematical Logic **Part 2** (1970) 115–125
26. Biere, A., Heule, M., van Maaren, H., Walsh, T., eds.: Handbook of Satisfiability. Vol. 185 of Frontiers in Artificial Intelligence and Applications. IOS Press (2008)
27. Fröhlich, A., Kovásznai, G., Biere, A.: A DPLL algorithm for solving DQBF. In: Int'l Workshop on Pragmatics of SAT (POS), Trento, Italy (2012)
28. Gitina, K., Wimmer, R., Reimer, S., Sauer, M., Scholl, C., Becker, B.: Solving DQBF through quantifier elimination. Reports of SFB/TR 14 AVACS 107 (2015) Online available at `http://www.avacs.org`.
29. Pigorsch, F., Scholl, C.: An AIG-based QBF-solver using SAT for preprocessing. In Sapatnekar, S.S., ed.: ACM/IEEE Design Automation Conference (DAC), Anaheim, CA, USA, ACM Press (2010) 170–175
30. Tarjan, R.E.: Depth-first search and linear graph algorithms. SIAM Journal on Computing **1**(2) (1972) 146–160
31. Brafman, R.I.: A simplifier for propositional formulas with many binary clauses. IEEE Transactions on Systems, Man, and Cybernetics, Part B **34**(1) (2004) 52–59
32. Gelder, A.V.: Toward leaner binary-clause reasoning in a satisfiability solver. Ann. Math. Artif. Intell. **43**(1) (2005) 239–253
33. Gershman, R., Strichman, O.: Cost-effective hyper-resolution for preprocessing CNF formulas. In Bacchus, F., Walsh, T., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 3569 of LNCS, St. Andrews, UK, Springer (2005) 423–429
34. Heule, M., Järvisalo, M., Biere, A.: Efficient CNF simplification based on binary implication graphs. In Sakallah, K.A., Simon, L., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 6695 of LNCS, Ann Arbor, MI, USA, Springer (2011) 201–215
35. Samer, M., Szeider, S.: Backdoor sets of quantified Boolean formulas. Journal of Automated Reasoning **42**(1) (2009) 77–97
36. Samer, M.: Variable dependencies of quantified CSPs. In: Int'l Conf. on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR). Vol. 5330 of LNCS, Doha, Qatar, Springer (2008) 512–527
37. Lonsing, F., Biere, A.: Efficiently representing existential dependency sets for expansion-based QBF solvers. Electronic Notes in Theoretical Computer Science **251** (2009) 83–95
38. Gelder, A.V.: Variable independence and resolution paths for quantified Boolean formulas. In Lee, J.H., ed.: Int'l Conf. on Principles and Practice of Constraint Programming (CP). Vol. 6876 of LNCS, Perugia, Italy, Springer (2011) 789–803
39. Slivovsky, F., Szeider, S.: Computing resolution-path dependencies in linear time. In Cimatti, A., Sebastiani, R., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 7317 of LNCS, Trento, Italy, Springer (2012) 58–71
40. Balabanov, V., Chiang, H.K., Jiang, J.R.: Henkin quantifiers and Boolean formulae: A certification perspective of DQBF. Theoretical Computer Science **523** (2014) 86–100
41. Bubeck, U., Büning, H.K.: Dependency quantified Horn formulas: Models and complexity. In Biere, A., Gomes, C.P., eds.: Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT). Vol. 4121 of LNCS, Seattle, WA, USA, Springer (2006) 198–211
42. Bubeck, U.: Model-based transformations for quantified Boolean formulas. PhD thesis, University of Paderborn (2010)

43. Heule, M., Järvisalo, M., Biere, A.: Clause elimination procedures for CNF formulas. In Fermüller, C.G., Voronkov, A., eds.: Int'l Conf. on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR). Vol. 6397 of LNCS, Yogyakarta, Indonesia, Springer (2010) 357–371
44. Heule, M., Järvisalo, M., Biere, A.: Covered clause elimination. In Voronkov, A., Sutcliffe, G., Baaz, M., Fermüller, C.G., eds.: Int'l Conf. on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR) (Short papers). Vol. 13 of EPiC Series, Yogyakarta, Indonesia, EasyChair (2010) 41–46
45. Heule, M., Seidl, M., Biere, A.: Blocked literals are universal. In Havelund, K., Holzmann, G., Joshi, R., eds.: NASA Formal Methods Symposium (NFM). Vol. 9058 of LNCS, Pasadena, CA, USA, Springer (2015) 436–442
46. Lonsing, F., Egly, U.: Incremental QBF solving by DepQBF. In Hong, H., Yap, C., eds.: Int'l Congress on Mathematical Software (ICMS). Vol. 8592 of LNCS, Seoul, South Korea, Springer (2014) 307–314
47. Lonsing, F., Biere, A.: DepQBF: A dependency-aware QBF solver. Journal on Satisfiability, Boolean Modelling and Computation **7**(2-3) (2010) 71–76
48. Kuehlmann, A., Paruthi, V., Krohm, F., Ganai, M.K.: Robust Boolean reasoning for equivalence checking and functional property verification. IEEE Transactions on CAD of Integrated Circuits and Systems **21**(12) (2002) 1377–1394

Appendix

A Proofs of the Main Lemmas and Theorems

A.1 Backbones, Monotonic and Equivalent Variables

Lemma A.1 *Unit literals are backbones, and pure literals monotonic.*

Proof. Let ℓ be a literal such that $\{\ell\} \in \varphi$, i. e., $\varphi = \varphi' \cup \big\{\{\ell\}\big\}$. W. l. o. g., we assume $\operatorname{sgn}(\ell) = 1$. Replacing ℓ by 0 yields $\varphi[0/\ell] = \varphi'[0/\ell] \cup \big\{\{0\}\big\}$, which is unsatisfiable as the clause $\{0\}$ is false. Therefore ℓ is a backbone.

Now consider the case that φ does not contain $\neg\ell$. Then we can partition φ into the clauses C^ℓ which contain ℓ and the clauses $C^\emptyset$ not containing ℓ, i. e., $\varphi = \varphi^\ell \wedge \varphi^\emptyset$. We can write φ as $\varphi = (\ell \vee \bar{\varphi}^\ell) \wedge \varphi^\emptyset$ with an appropriate formula $\bar{\varphi}^\ell$ that does not contain ℓ. We have

$$\begin{aligned}\varphi[0/\ell] \wedge \neg\varphi[1/\ell] &= (0 \vee \bar{\varphi}^\ell) \wedge \varphi^\emptyset \wedge \neg\big((1 \vee \bar{\varphi}^\ell) \wedge \varphi^\emptyset\big) \\ &= \bar{\varphi}^\ell \wedge \varphi^\emptyset \wedge \neg\varphi^\emptyset \\ &= 0\end{aligned}$$

Hence, ℓ is monotonic. □

To prove Lemmas 2 and 3 we need the following property of binary implication graphs:

Lemma A.2 *Let $G_\psi = (L, E)$ be the binary implication graph of ψ and $\ell_1, \ell_2 \in L$ two literals. If there is a path from ℓ_1 to ℓ_2 in G_ψ, then φ implies the clause $\{\neg\ell_1, \ell_2\}$.*

Proof. Let $\ell_1 = \kappa_1 \to \kappa_2 \to \cdots \to \kappa_k = \ell_2$ be a path in G_ψ from ℓ_1 to ℓ_2. This means that φ contains the clauses $C_i := \{\neg\kappa_i, \kappa_{i+1}\}$ for $i = 1, \ldots, k-1$. We set $r_1 = \{\neg\kappa_1, \kappa_2\}$ and $r_i = r_{i-1} \otimes_{\kappa_i} C_i = \{\neg\kappa_1, \kappa_{i+1}\}$ for $i = 2, \ldots, k-1$. So we can derive $r_{k-1} = \{\neg\kappa_1, \kappa_k\} = \{\neg\ell_1, \ell_2\}$ from $C_1, \ldots, C_{k-1}$ by resolution. As resolvents are implied by the formula, they can be added without changing the formula's truth value. This shows the claim. □

Lemma 2. *A literal ℓ is a backbone if there is a path in G_ψ from $\neg\ell$ to ℓ.*

Proof. Assume that there is a path in G_ψ from $\neg\ell$ to ℓ. By Lemma A.2 we can derive the clause $\{\neg\neg\ell, \ell\} = \{\ell\}$ by resolution. Therefore $Q : \varphi$ is equivalent to $Q : \varphi \wedge \{\ell\}$. By Definition 6 we have that ℓ is unit in $\varphi \wedge \{\ell\}$. Therefore ℓ is a backbone in $\varphi \wedge \{\ell\}$ and also in φ. □

Theorem 3. *Let ℓ and μ be equivalent literals. We assume w. l. o. g. that* $\operatorname{sgn}(\ell) = 1$.

If $\operatorname{var}(\ell), \operatorname{var}(\mu) \in V_\forall^\psi$*, then ψ is unsatisfiable. Otherwise, we assume w. l. o. g. that* $\operatorname{var}(\ell) \in V_\exists^\psi$*. If* $\operatorname{var}(\mu) \in V_\forall^\psi$ *and* $\operatorname{var}(\mu) \notin D^\psi_{\operatorname{var}(\ell)}$*, then ψ is unsatisfiable. If*

$\mathrm{var}(\mu) \in V_\forall^\psi$ *and* $\mathrm{var}(\mu) \in D^\psi_{\mathrm{var}(\ell)}$*, then* ψ *is equivalent to* $Q \setminus \{\mathrm{var}(\ell)\} : \varphi[\mu/\ell]$*. If* $\mathrm{var}(\ell), \mathrm{var}(\mu) \in V_\exists^\psi$*, then* ψ *is equivalent to*

$$\psi' := \big(Q \setminus \{\mathrm{var}(\mu), \mathrm{var}(\ell)\}\big) \cup \big\{\exists\, \mathrm{var}(\mu)(D^\psi_{\mathrm{var}(\mu)} \cap D^\psi_{\mathrm{var}(\ell)})\big\} : \varphi[\mu/\ell]\,.$$

Proof. Assume that ℓ and μ are equivalent, i. e., φ is equivalent to $\varphi \wedge (\neg\ell \vee \mu) \wedge (\ell \vee \neg\mu)$, which is the same as $\varphi \wedge (\ell \equiv \mu)$.

- We first consider the case where $\mathrm{var}(\ell), \mathrm{var}(\mu) \in V_\forall^\psi$.
  Since $(\ell \equiv \mu)$ does not contain an existential variable and universal quantifiers distribute over $\wedge$, ψ is equivalent to
  $$\big(Q : \varphi\big) \wedge \big(\forall\, \mathrm{var}(\ell) \forall\, \mathrm{var}(\mu) : (\ell \equiv \mu)\big).$$
  Obviously, $(\ell \equiv \mu)$ is not a tautology, and therefore ψ is unsatisfiable.
- Next, let $\mathrm{var}(\ell) \in V_\exists^\psi$, $\mathrm{var}(\mu) \in V_\forall^\psi$, and $\mathrm{var}(\mu) \notin D^\psi_{\mathrm{var}(\ell)}$.
  Assume ψ were satisfiable. Then there were Skolem functions that turn $\psi \wedge (\ell \equiv \mu)$ into a tautology. The only Skolem function for $\mathrm{var}(\ell)$ which is able to turn $(\ell \equiv \mu)$ into a tautology is $s_{\mathrm{var}(\ell)} = \mu$ (note that we assume $\mathrm{sgn}(\ell) = 1$). However, this Skolem function is not admissible as $\mathrm{var}(\mu) \notin D^\psi_{\mathrm{var}(\ell)}$. Therefore ψ is unsatisfiable.
- Now consider the case that $\mathrm{var}(\ell) \in V_\exists^\psi$, $\mathrm{var}(\mu) \in V_\forall^\psi$, and $\mathrm{var}(\mu) \in D^\psi_{\mathrm{var}(\ell)}$.
  With a similar argumentation as in the previous case, we can derive that if ψ is satisfiable, the only Skolem function for $\mathrm{var}(\ell)$ is $s_{\mathrm{var}(\ell)} = \mu$, which is admissible. Therefore, replacing ℓ by μ again yields a satisfiable formula. On the other hand, if ψ is unsatisfiable, replacing the existential variables by any admissible functions does not turn φ into a tautology. Therefore replacing ℓ by μ yields an unsatisfiable formula again.
- Finally, we consider $\mathrm{var}(\ell), \mathrm{var}(\mu) \in V_\exists^\psi$.
  First assume that ψ is unsatisfiable, i. e., there is no set of Skolem functions for the existential variables which turns the matrix into a tautology. In particular, this also holds for all sets of Skolem functions in which $s_{\mathrm{var}(\ell)}$ and $s_{\mathrm{var}(\mu)}$ are identical (modulo negation). Therefore replacing ℓ by μ and restricting the dependency set accordingly yields an unsatisfiable formula again.
  Now assume that ψ is satisfied. Because the Skolem functions for $\mathrm{var}(\ell)$ and $\mathrm{var}(\mu)$ have to turn $(\ell \equiv \mu)$ into a tautology, they have to be equal because of $(\ell \equiv \mu)$. Since the Skolem function must be admissible for $\mathrm{var}(\ell)$, it must not depend on $D^\psi_{\mathrm{var}(\mu)} \setminus D^\psi_{\mathrm{var}(\ell)}$. Similarly, because it has to be admissible for $\mathrm{var}(\mu)$, it must not depend on $D^\psi_{\mathrm{var}(\ell)} \setminus D^\psi_{\mathrm{var}(\mu)}$. Therefore it may only depend on $D^\psi_{\mathrm{var}(\ell)} \cap D^\psi_{\mathrm{var}(\mu)}$. So we may replace ℓ by μ if we restrict $\mathrm{var}(\mu)$'s dependency set to $D^\psi_{\mathrm{var}(\ell)} \cap D^\psi_{\mathrm{var}(\mu)}$.

□

Lemma 3. *Two literals* ℓ, μ *are equivalent if there is a path in* $\mathrm{BIP}(\psi)$ *from* ℓ *to* μ *and vice versa.*

Proof. The path from ℓ to μ allows us, according to Lemma A.2, to derive the clause $\{\neg\ell, \mu\}$, the path from μ to ℓ the corresponding clause $\{\neg\mu, \ell\}$. Both clauses may be added to φ as they are resolvents. This implies according to Definition 9 that ℓ and μ are equivalent. □

A.2 Reduction of Dependency Sets

To prove Theorem 4 we use the following definition:

Definition 14. *Let $G_{V,\varphi} = (V \cup \varphi, E)$ the variable-clause incidence graph of ψ and $A \subseteq V$. Two variables $v, v' \in V$ are* connected *w. r. t. A if there is a path from v to v', visiting only clauses and variables in A in between.*

Using this definition, Theorem 4 reads as follows:

Theorem 4 (Standard dependency scheme). *An existential variable $y \in V_\exists^\psi$ is independent of a universal variable $x \in V_\forall^\psi$ with $x \in D_y^\psi$ if they are not connected w. r. t. $Z := \{z \in V_\exists^\psi \mid x \in D_z^\psi\}$.*

Proof. W. l. o. g. we assume that $x = x_1$ and $y = y_1$.

Let $y_1 \in V_\exists^\psi$ and $x_1 \in V_\forall^\psi$ be variables such that $x_1 \in D_{y_1}^\psi$, but they are not connected w. r. t. Z.

Let $\psi' = Q' : \varphi$ result from $\psi = Q : \varphi$ by removing x_1 from the dependency set $D_{y_1}^\psi$ of y_1. We have to show that ψ' is satisfied iff ψ is satisfied.

Since there is no path from x_1 to y_1 which only visits clauses and variable nodes in Z, we can partition the set $V_\exists^\psi$ of existential variables into the three pairwise disjoint subsets

$$
\begin{aligned}
V_A^\psi &= \{a \in V_\exists^\psi \mid x_1 \notin D_a^\psi\},\\
V_B^\psi &= \{b \in V_\exists^\psi \mid x_1 \in D_b^\psi \wedge b \text{ is connected w. r. t. } Z \text{ to } x_1\},\\
V_C^\psi &= \{c \in V_\exists^\psi \mid x_1 \in D_c^\psi \wedge c \text{ is not connected w. r. t. } Z \text{ to } x_1\}.
\end{aligned}
$$

Using these sets of variables, we define the following two sets of clauses:

$$
\begin{aligned}
\varphi_{x_1} &= \{C \in \varphi \mid \operatorname{var}(C) \cap V_B^\psi \neq \emptyset\},\\
\varphi_{y_1} &= \{C \in \varphi \mid \operatorname{var}(C) \cap V_B^\psi = \emptyset\}.
\end{aligned}
$$

There is no variable $y_C \in V_C^\psi$ which appears in φ_{x_1}. Assume the contrary, i. e., there is a clause $C \in \varphi_{x_1}$ and a variable y_C such that $y_C \in V_C^\psi \cap \operatorname{var}(C)$. Due to the definition of φ_{x_1}, $\operatorname{var}(C)$ contains a variable y_B with $y_B \in \operatorname{var}(C) \cap V_B^\psi$. That means, y_B is connected w. r. t. Z to x_1 and $G_{V,\varphi}$ contains edges between y_B and C as well as C and y_C. As $x_1 \in D_{y_B}^\psi$ and $x_1 \in D_{y_C}^\psi$, x_1 and y_C are connected w. r. t. Z, which contradicts $y_C \in V_C^\psi$.

Using this partitioning, we can write ψ as

$$\begin{aligned}\psi &= \forall x_1 \dots \forall x_n \exists y_1(D^\psi_{y_1}) \dots \exists y_m(D^\psi_{y_m}) : \varphi \\ &= \forall x_1 \dots \forall x_n \exists y_1(D^\psi_{y_1}) \dots \exists y_m(D^\psi_{y_m}) : \varphi_{x_1} \wedge \varphi_{y_1}\end{aligned}$$

φ_{x_1} does not contain any variable in V^ψ_C, and φ_{y_1} none of the variables in V^ψ_B. Therefore we can move the existential quantifier inside [40]:

$$\equiv \forall x_1 \dots \forall x_n \underbrace{\exists a(D^\psi_a)}_{\text{for } a \in V^\psi_A} : \Big(\underbrace{\exists b(D^\psi_b)}_{\text{for } b \in V^\psi_B} : \varphi_{x_1} \Big) \wedge \Big(\underbrace{\exists c(D^\psi_c)}_{\text{for } c \in V^\psi_C} : \varphi_{y_1} \Big)$$

The variables in V^ψ_A do not depend on x_1 and universal quantifiers distribute over $\wedge$:

$$\equiv \forall x_2 \dots \forall x_n \underbrace{\exists a(D^\psi_a)}_{\text{for } a \in V^\psi_A} : \Big(\forall x_1 \underbrace{\exists b(D^\psi_b)}_{\text{for } b \in V^\psi_B} : \varphi_{x_1} \Big) \wedge \Big(\forall x_1 \underbrace{\exists c(D^\psi_c)}_{\text{for } c \in V^\psi_C} : \varphi_{y_1} \Big)$$

As φ_{y_1} does not contain x_1, we can remove x_1 from any dependency set D^ψ_c with $c \in V^\psi_C$, in particular from $D^\psi_{y_1}$:

$$\equiv \forall x_2 \dots \forall x_n \underbrace{\exists a(D^\psi_a)}_{\text{for } a \in V^\psi_A} : \Big(\forall x_1 \underbrace{\exists b(D^\psi_b)}_{\text{for } b \in V^\psi_B} : \varphi_{x_1} \Big) \wedge \Big(\forall x_1 \exists y_1(D^\psi_{y_1} \setminus \{x_1\}) \underbrace{\exists c(D^\psi_c)}_{\text{for } c \in V^\psi_C \setminus \{y_1\}} : \varphi_{y_1} \Big)$$

Bringing the formula back to prenex form yields:

$$\begin{aligned}&\equiv \forall x_1 \dots \forall x_n \exists y_1(D^\psi_{y_1} \setminus \{x_1\}) \exists y_2(D^\psi_{y_2}) \dots \exists y_m(D^\psi_{y_m}) : \varphi_{x_1} \wedge \varphi_{y_1} \\ &= \forall x_1 \dots \forall x_n \exists y_1(D^\psi_{y_1} \setminus \{x_1\}) \exists y_2(D^\psi_{y_2}) \dots \exists y_m(D^\psi_{y_m}) : \varphi\,.\end{aligned}$$

□

A.3 Universal Reduction, Resolution, and Universal Expansion

Theorem 5 (Variable elimination by resolution). *Let $Q : \varphi$ be DQBF and $y \in V^\psi_\exists$ an existential variable. We partition φ into the sets $\varphi^y = \{C \in \varphi \mid y \in C\}$, $\varphi^{\neg y} = \{C \in \varphi \mid \neg y \in C\}$, and $\varphi^\emptyset = \varphi \setminus (C^y \cup C^{\neg y})$.*

If one of the following conditions is satisfied:

1. *for all $C \in \varphi^y$ and all $k \in C$ we have $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(y)$,*
2. *for all $C' \in \varphi^{\neg y}$ and all $k \in C'$ we have $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(y)$, or*
3. *y is the output of a gate, i. e., there are clauses encoding the relationship $y \equiv f(V')$ for some function f and arguments $V' \subseteq V \setminus \{y\}$ (cf. Sec. 4.5),*

then $Q : \varphi$ is equivalent to

$$Q' : \varphi^{\emptyset} \wedge \bigwedge_{C \in \varphi^y} \bigwedge_{C' \in \varphi^{\neg y}} C \otimes_y C'.$$

where Q' results from Q by removing the quantification of the existential variable y from the quantifier prefix Q.

Proof. Since the variable y does not occur in $\varphi^{\emptyset} \wedge \bigwedge_{C \in \varphi^y} \bigwedge_{C' \in \varphi^{\neg y}} C \otimes_y C'$, it is clear that

$$Q' : \varphi^{\emptyset} \wedge \bigwedge_{C \in \varphi^y} \bigwedge_{C' \in \varphi^{\neg y}} C \otimes_y C'$$

and

$$Q : \underbrace{\varphi^{\emptyset} \wedge \bigwedge_{C \in \varphi^y} \bigwedge_{C' \in \varphi^{\neg y}} C \otimes_y C'}_{\varphi'} \tag{1}$$

are equivalent. So we have to prove the equivalence of $Q : \varphi$ and $Q : \varphi^{\emptyset} \wedge \bigwedge_{C \in \varphi^y} \bigwedge_{C' \in \varphi^{\neg y}} C \otimes_y C'$. First let φ be satisfied. Since adding resolvents of φ to φ does not change φ, we have

$$Q : \varphi \quad \equiv \quad Q : \varphi \wedge \bigwedge_{C \in \varphi^y} \bigwedge_{C' \in \varphi^{\neg y}} C \otimes_y C'.$$

Since deleting clauses from a satisfied formula yields a satisfied formula again, we can conclude that (1) is satisfied.

Now let (1) be satisfied. We show that this implies the satisfaction of $Q : \varphi$ if one of the conditions in the theorem is satisfied.

1. First assume that $\mathrm{dep}_{\psi}(k) \subseteq \mathrm{dep}_{\psi}(y)$ holds for all $C \in \varphi^y$ and all $k \in C$. If (1) is satisfied, there are Skolem functions s'_{y_i} for all existential variables y_i ($i = 1, \ldots, m$) such that replacing y_i by s'_{y_i} turns φ' into a tautology. We define the following set of Skolem functions for φ:

$$s_z = \begin{cases} s'_z, & \text{if } z \neq y, \\ \neg\varphi^y[0/y][s'_{y_i}/y_i \text{ for all } i = 1, \ldots, m \text{ with } y_i \neq y], & \text{if } z = y. \end{cases}$$

   That means: for all existential variables except y we use the Skolem functions from (1). The Skolem function for y is chosen such that it assigns 1 to y iff there is a clause in φ^y which is not already satisfied by (the Skolem functions of) the other literals. s_y defined in the given way depends on a subset of variables from $\mathrm{dep}_{\psi}(y)$, since we have $\mathrm{dep}_{\psi}(k) \subseteq \mathrm{dep}_{\psi}(y)$ for all $C \in \varphi^y$ and all $k \in C$.
   Clearly, $s_{y_1}, \ldots, s_{y_m}$ satisfy the clauses in $\varphi^{\emptyset}$. We distinguish two cases for proving that $s_{y_1}, \ldots, s_{y_m}$ satisfy all clauses in $\varphi^{\neg y}$ and $\varphi^{\neg y}$ as well (for all variable assignments to the universal variables). Consider an arbitrary assignment ν of values to the universal variables.

- Case 1:
  Each clause $C \in \varphi^y$ contains a literal $k \neq y$ such that $s_{\mathrm{var}(k)}(\nu_{|\mathrm{dep}_\psi(k)}) = \mathrm{sgn}(k)$, if $k \in V_\exists^\psi$, $\nu(\mathrm{var}(k)) = \mathrm{sgn}(k)$, if $k \in V_\forall^\psi$, i. e., the formula φ^y evaluates to 1, if we replace existential variables by Skolem functions and evaluate w. r. t. ν. This is independent from the value for y. According to the definition of $s_y = \neg\varphi^y[0/y][s'_{y_i}/y_i$ for all $i = 1, \ldots, m$ with $y_i \neq y]$ we then have $s_y(\nu_{|\mathrm{dep}_\psi(y)}) = 0$. Thus, all clauses in $C^{\neg y}$ evaluate to 1, if we replace existential variables by Skolem functions and evaluate w. r. t. ν, since all those clauses contain $\neg y$.
- Case 2:
  There is a clause $C \in \varphi^y$ such that for all literals $k \in C$ with $k \neq y$ it holds $s_{\mathrm{var}(k)}(\nu_{|\mathrm{dep}_\psi(k)}) = \neg\,\mathrm{sgn}(k)$, if $k \in V_\exists^\psi$, $\nu(\mathrm{var}(k)) = \neg\,\mathrm{sgn}(k)$, if $k \in V_\forall^\psi$, i. e., the formula φ^y would evaluate to 0, if we would replace y by 0. According to the definition of $s_y = \neg\varphi^y[0/y][s'_{y_i}/y_i$ for all $i = 1, \ldots, m$ with $y_i \neq y]$ we then have $s_y(\nu_{|\mathrm{dep}_\psi(y)}) = 1$ and thus all clauses in φ^y evaluate to 1 for ν after replacing existential variables by Skolem functions. We have to show that in this case all clauses in $\varphi^{\neg y}$ evaluate to 1 as well. Assume the opposite, i. e., there is a clause $C' \in \varphi^{\neg y}$ that does not evaluate to 1. Then consider the resolvent $C \otimes_y C'$. As for all literals k in $C \setminus \{y\}$ it holds $s_{\mathrm{var}(k)}(\nu_{|\mathrm{dep}_\psi(k)}) = \neg\,\mathrm{sgn}(k)$, if $k \in V_\exists^\psi$, $\nu(\mathrm{var}(k)) = \neg\,\mathrm{sgn}(k)$, if $k \in V_\forall^\psi$, and for all literals ℓ in $C' \setminus \{\neg y\}$ it holds $s_{\mathrm{var}(k)}(\nu_{|\mathrm{dep}_\psi(\ell)}) = \neg\,\mathrm{sgn}(\ell)$, if $\ell \in V_\exists^\psi$, $\nu(\mathrm{var}(\ell)) = \neg\,\mathrm{sgn}(\ell)$, if $\ell \in V_\forall^\psi$, the resolvent $(C \otimes_y C')[s_{y_i}/y_i$ for all $i = 1, \ldots, m]$ as well evaluates to 0 for assignment ν. This contradicts the assumption that (1) is satisfied by the Skolem functions s'_{y_i} ($C \otimes_y C'$ does not contain y and the Skolem functions s_{y_i} are equal to s'_{y_i} for all $y_i \neq y$). Thus $\varphi^{\neg y}[s_{y_i}/y_i$ for all $i = 1, \ldots, m]$ evaluates to 1 for assignment ν.

2. The second condition is dual to the first one. As Skolem functions for φ we choose $s_{y_i} = s'_{y_i}$ for all $i = 1, \ldots, m$ with $y_i \neq y$ and
$$s_y = \neg\varphi^{\neg y}[1/y][s'_{y_i}/y_i \text{ for all } i = 1, \ldots, m \text{ with } y_i \neq y].$$
3. The third condition requires that there are clauses which define the relationship $y \equiv f(V')$ and $\mathrm{dep}_\psi(v) \subseteq \mathrm{dep}_\psi(y)$ for all $v \in V'$. This implies that given Skolem functions for the remaining existential variables in (1), there is only one Skolem function for y, which is given by the gate definition. Therefore we can increase $\mathrm{dep}_\psi(y)$ to $V_\forall^\psi$ without changing the truth value of the DQBF, i. e., we set the dependency set $\mathrm{dep}_\psi(y) := V_\forall^\psi$. Then we can eliminate the variable y as described in [9], e.g.. Eliminating a variable by existential quantification is then equivalent to elimination using resolution.

□

A.4 Blocked Clause Elimination

Theorem 7. *Let $Q : \varphi \wedge C$ be a DQBF with a blocked clause C. Then $Q : \varphi \wedge C$ and $Q : \varphi$ are equivalent.*

Proof. Let $\psi := Q : \varphi \wedge C$ be a DQBF and C a clause that is blocked by literal $\ell \in C$. Furthermore, we set $\psi' := Q : \varphi$ as the formula after deleting the blocked clause C. We have to show that $\psi \equiv \psi'$.

If ψ is satisfied, ψ' is satisfied, too, as the matrix of ψ' is a subset of the matrix of ψ. So we assume that ψ' is satisfied and show that this also holds for ψ. We prove this claim by induction on the number $|\mathrm{dep}_\psi(\ell)|$ of dependencies of the blocking literal ℓ.

Since ψ' is satisfied, there are Skolem functions $s_{y_i} : \mathcal{A}\big(\mathrm{dep}_\psi(y_i)\big) \to \{0,1\}$ for the existential variables $y_i \in V_\exists^\psi$ such that replacing each y_i by s_{y_i} turns φ into a tautology.

For the induction base, we assume that $|\mathrm{dep}_\psi(\ell)| = 0$, i. e., $\mathrm{dep}_\psi(\ell) = \emptyset$, $s_{\mathrm{var}(\ell)}$ is a constant function (0 or 1). Now we construct Skolem functions s'_{y_i} for $\varphi \wedge C$ by only modifying $s_{\mathrm{var}(\ell)}$. For this we have to distinguish two cases:

- Case 1: Replacing each y_i by s_{y_i} turns C into a tautology. Then the replacement turns $\varphi \wedge C$ into a tautology. We just set $s'_{y_i} := s_{y_i}$.
- Case 2: Replacing each y_i by s_{y_i} does *not* turn C into a tautology. Then $s_{\mathrm{var}(\ell)} = \neg\,\mathrm{sgn}(\ell)$.[4] We set $s'_{\mathrm{var}(\ell)} := \mathrm{sgn}(\ell)$ and $s'_{y_i} := s_{y_i}$ for all other existential variables. Thus, C is turned into a tautology by replacing variables by Skolem functions. All clauses $C' \in \varphi$ with $\neg\ell \notin C'$ remain tautologies after replacing variables by Skolem functions. So consider a clause $C' \in \varphi$ with $\neg\ell \in C'$. Since C is blocked by ℓ, there is a literal k with $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell) = \emptyset$, $k \in C'$, and $\neg k \in C$. Since $\mathrm{dep}_\psi(k) = \emptyset$, $s'_{\mathrm{var}(k)} = s_{\mathrm{var}(k)}$ is a constant function. Since $\neg k \in C$ and replacing variables by Skolem functions s_{y_i} does not turn C into a tautology, $s'_{\mathrm{var}(k)} = s_{\mathrm{var}(k)} = \mathrm{sgn}(k)$. Therefore C' with $k \in C'$ remains a tautology after replacing variables by Skolem functions (even though $s_{\mathrm{var}(\ell)}$ is changed).

For the induction step, assume that $|\mathrm{dep}_\psi(\ell)| > 0$. We expand a universal variable $x \in \mathrm{dep}_\psi(\ell)$ (cf. Theorem 6) and show that the two clauses induced by C in the resulting formula $\tilde{\psi}$ are either satisfied or blocked by ℓ or its copy ℓ'. By induction assumption they can be removed as in $\tilde{\psi}$ they depend on one variable less. Undoing the expansion step yields a formula without the blocked clause that is equivalent to ψ. This shows the claim.

Eliminating x from ψ by universal expansion (cf. Theorem 6) yields [41, 42, 40, 9]:

$$\begin{aligned}\tilde{\psi} &= \tilde{Q} : \big(\varphi \wedge C\big)\big[1/x\big] \wedge \big(\varphi \wedge C\big)\big[0/x\big]\big[y'/y \text{ for all } y \text{ with } x \in \mathrm{dep}_\psi(y)\big] \\ &= \tilde{Q} : \underbrace{\varphi\big[1/x\big] \wedge C\big[1/x\big]}_{(*)} \wedge \underbrace{\varphi\big[0/x\big]\big[y'/y\ldots\big] \wedge C\big[0/x\big]\big[y'/y\ldots\big]}_{(**)}\end{aligned}$$

where $\tilde{Q}$ results from Q by removing x from all dependency sets and by adding existential quantifiers for the copied existential variables y' with the same dependency sets as for variables y. Since $x \in \mathrm{dep}_\psi(\ell)$, ℓ is replaced by the copied variable ℓ' in $(**)$.

[4] Remember that $\mathrm{sgn}(v) = 1$ and $\mathrm{sgn}(\neg v) = 0$ for a variable v.

Both $(*)$ and $(**)$ contain a copy of C. We have to show that both can be removed.

We distinguish the following three cases:

- Case 1: $x \in C$:
  Then $C[1/x]$ is satisfied and can be removed from $(*)$.
  We show that in $(**)$, $C[0/x][y'/y\ldots] = C[y'/y\ldots] \setminus \{x\}$ is a blocked clause that is blocked by literal $\ell' \in C[y'/y\ldots] \setminus \{x\}$: Consider an arbitrary clause in $\tilde{\psi}$ which contains $\neg\ell'$. This clause has to be in $(**)$ and has the form $C'[0/x][y'/y\ldots]$. Since C is blocked by ℓ in ψ, C contains a literal k, C' a literal $\neg k$ and $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$. If $k = x$, we do not need to consider $C'[0/x][y'/y\ldots]$, since C' is satisfied by replacing x by 0, i. e., $C'[0/x][y'/y\ldots]$ can be removed from $(**)$. $k \neq \neg x$, since otherwise C would be tautological. Thus we can assume $k \notin \{x, \neg x\}$ in the following. $C[0/x][y'/y\ldots]$ contains $\tilde{k}$, $C'[0/x][y'/y\ldots]$ contains $\neg\tilde{k}$, i. e., $\{\tilde{k}, \neg\tilde{k}\} \subseteq C[0/x][y'/y\ldots] \otimes_\ell C'[0/x][y'/y\ldots]$ and $\mathrm{dep}_\psi(\tilde{k}) = \mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell')$. (To simplify notations we always write $\tilde{k}$ for the copy k' of k in $(**)$, if $(**)$ contains k', and otherwise for k.) $C[0/x][y'/y\ldots]$ is indeed blocked. Because ℓ' does not depend on x anymore in $\tilde{\psi}$ (i. e., it depends on one variable less), we can apply the induction assumption and remove the blocked clause.
- Case 2: Neither x nor $\neg x$ are contained in C:
  Again we show that in $(**)$ $C[0/x][y'/y\ldots]$ is a blocked clause that is blocked by literal $\ell' \in C[0/x][y'/y\ldots]$. As in Case 1 all clauses in $\tilde{\psi}$ containing $\neg\ell'$ are restricted to $(**)$ and have the form $C'[0/x][y'/y\ldots]$. The proof that $C[0/x][y'/y\ldots]$ is blocked by $\ell' \in C[0/x][y'/y\ldots]$ is exactly the same as in Case 1 (with the only difference that now $k \notin \{x, \neg x\}$, since $k \in C$, but $x \notin C$, $\neg x \notin C$ due to the assumption of Case 2).
  Moreover, we show that in $(*)$, $C[1/x]$ is a blocked clause that is blocked by literal $\ell \in C[1/x]$. All clauses in $\tilde{\psi}$ containing $\neg\ell$ are restricted to $(*)$ and have the form $C'[1/x]$. The proof that $C[1/x]$ is blocked by ℓ is analogous to the first part for $C[0/x][y'/y\ldots]$. $C[0/x][y'/y\ldots]$ and $C[1/x]$ are blocked and because ℓ and ℓ' do not depend anymore on x in $\tilde{\psi}$ (i. e., they depend on one variable less), we can apply the induction assumption and remove the blocked clauses.
- Case 3: $\neg x \in C$:
  This case is analogous to the first case.

This shows

$$\psi \quad\equiv\quad \tilde{\psi} \quad\equiv\quad \tilde{Q} : \varphi\big[1/x\big] \wedge \varphi\big[0/x\big]\big[y'/y\ldots\big] \quad\equiv\quad \psi'.$$

□

Lemma 5. *BCE for DQBF has a unique fixed point.*

Proof. The proof is similar as for BCE on SAT problems [22]: If $\psi := Q : \varphi \wedge C$ is a DQBF and C a blocked clause w. r. t. ψ. Then any clause $C' \in \varphi$ which is

blocked w. r. t. ψ is also blocked w. r. t. $Q : \varphi$. Therefore the result of BCE is independent of the order in which the clauses are removed, and hence BCE has a unique fixed point. □

Theorem 8 (Hidden literal addition, HLA). *Let $Q : \varphi \wedge C$ be a DQBF and ℓ a hidden literal for C. Then $Q : \varphi \wedge C$ and $Q : \varphi \wedge \big(C \cup \{\ell\}\big)$ are equivalent.*

Proof. Assume that ℓ is a hidden literal for C according to Definition 12. We set $C' := C \cup \{\ell\}$. ℓ being a hidden literal for C means that there is a clause $D := \{\ell_1, \ldots, \ell_n, \neg\ell\} \in \varphi$ with $\{\ell_1, \ldots, \ell_n\} \subseteq C \subseteq C'$. For the resolvent of C' and D w. r. t. ℓ we have $C' \otimes_\ell D = C$. Adding a resolvent to a CNF yields an equivalent CNF, i. e.,

$$\varphi \wedge C' \quad \equiv \quad \varphi \wedge C' \wedge C.$$

C subsumes $C' = C \cup \{\ell\}$; therefore C' can be removed from the formula. Hence, $\varphi \wedge C$ and $\varphi \wedge C'$ are logically equivalent. Replacing the matrix of a DQBF with an equivalent formula does not change the truth value of a DQBF. □

Theorem 9 (Covered literal addition, CLA). *Let $Q : \varphi \wedge C$ be DQBF and k a covered literal for C. Then $Q : \varphi \wedge C$ and $Q : \varphi \wedge \big(C \cup \{k\}\big)$ are equivalent.*

Proof. Let k be covered for C w. r. t. an existential variable ℓ, i. e., $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$ and $k \in \bigcap R_\psi(C, \ell) \setminus \{\neg\ell\}$.

We have to show

$$\vDash Q : \varphi \wedge C \quad \Leftrightarrow \quad \vDash Q : \varphi \wedge \big(C \cup \{k\}\big).$$

The direction "⇒" is trivial as each satisfying assignment of $\varphi \wedge C$ also satisfies $\varphi \wedge \big(C \cup \{k\}\big)$. So let $Q : \varphi \wedge \big(C \cup \{k\}\big)$ be satisfied. We show by induction on $|\mathrm{dep}_\psi(\ell)|$ that this implies the satisfaction of $Q : \varphi \wedge C$.

Since $Q : \varphi \wedge \big(C \cup \{k\}\big)$ is satisfied, there are Skolem functions $s_{y_i} : \mathcal{A}\big(\mathrm{dep}_\psi(y_i)\big) \to \{0, 1\}$ for the existential variables $y_i \in V_\exists^\psi$ such that replacing each y_i by s_{y_i} turns $\varphi \wedge \big(C \cup \{k\}\big)$ into a tautology.

For the induction base, we assume that $|\mathrm{dep}_\psi(\ell)| = 0$, i. e., $\mathrm{dep}_\psi(\ell) = \emptyset$, $s_{\mathrm{var}(\ell)}$ is a constant function (0 or 1). Now we construct Skolem functions s'_{y_i} for $\varphi \wedge C$ by only modifying $s_{\mathrm{var}(\ell)}$. For this we have to distinguish two cases:

- Case 1: Replacing each y_i by s_{y_i} turns C into a tautology. Then the replacement turns $\varphi \wedge C$ into a tautology. We just set $s'_{y_i} := s_{y_i}$.
- Case 2: Replacing each y_i by s_{y_i} does *not* turn C into a tautology. Then $s_{\mathrm{var}(\ell)} = \neg\,\mathrm{sgn}(\ell)$, because $\ell \in C$.[5] We set $s'_{\mathrm{var}(\ell)} := \mathrm{sgn}(\ell)$ and $s'_{y_i} := s_{y_i}$ for all other existential variables. Thus, C is turned into a tautology by replacing variables by Skolem functions. All clauses $C' \in \varphi$ with $\neg\ell \notin C'$ remain tautologies after replacing variables by Skolem functions. So consider a clause $C' \in \varphi$ with $\neg\ell \in C'$.

  We consider two cases: The first case is that $C' \notin R_\psi(C, \ell)$. Since $\neg\ell \in C'$,

[5] Remember that $\mathrm{sgn}(v) = 1$ and $\mathrm{sgn}(\neg v) = 0$ for a variable v.

there has to be a literal $m \neq \ell$ with $\{m, \neg m\} \subseteq C \otimes_\ell C'$, but $\mathrm{dep}_\psi(m) \subseteq \mathrm{dep}_\psi(\ell)$, i. e., $\mathrm{dep}_\psi(m) = \emptyset$. We assume w. l. o. g. $m \in C'$ and $\neg m \in C$. Due to $\mathrm{dep}_\psi(m) = \emptyset$, the Skolem function $s_{\mathrm{var}(m)}$ has to be a constant function. Since $\neg m \in C$ and replacing variables by Skolem functions s_{y_i} does not turn C into a tautology, $s'_{\mathrm{var}(m)} = s_{\mathrm{var}(m)} = \mathrm{sgn}(m)$. Therefore C' with $m \in C'$ remains a tautology after replacing variables by Skolem functions (even though $s_{\mathrm{var}(\ell)}$ is changed).
The second case is that $C' \in R_\psi(C, \ell)$. Then we can conclude that the covered literal k is contained in C'. Since $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$, i. e., $\mathrm{dep}_\psi(k) = \emptyset$, the Skolem function $s_{\mathrm{var}(k)}$ has to be a constant function. Since replacing the existential variables y_i by Skolem functions s_{y_i} turns $\big(C \cup \{k\}\big)$ into a tautology, but not C, it follows $s'_{\mathrm{var}(k)} = s_{\mathrm{var}(k)} = \mathrm{sgn}(k)$. Because of $k \in C'$, replacement of y_i by Skolem functions s'_{y_i} turns C' into a tautology.

Altogether the constructed Skolem functions s'_{y_i} turn $\varphi \wedge C$ into a tautology, i. e., $Q : \varphi \wedge C$ is satisfied.

For the induction step assume $|\mathrm{dep}_\psi(\ell)| > 0$. We do universal expansion (cf. Theorem 6) of $x \in \mathrm{dep}_\psi(\ell)$ in ψ and obtain the formula [41, 42, 40, 9]:

$$\begin{aligned}\tilde{\psi} &= \tilde{Q} : \big(\varphi \wedge C\big)\big[1/x\big] \wedge \big(\varphi \wedge C\big)\big[0/x\big]\big[y'/y \text{ for all } y \text{ with } x \in \mathrm{dep}_\psi(y)\big] \\ &= \tilde{Q} : \underbrace{\varphi\big[1/x\big] \wedge C\big[1/x\big]}_{(*)} \wedge \underbrace{\varphi\big[0/x\big]\big[y'/y\ldots\big] \wedge C\big[0/x\big]\big[y'/y\ldots\big]}_{(**)}\end{aligned}$$

where $\tilde{Q}$ results from Q by removing x from all dependency sets and by adding existential quantifiers for the copied existential variables y' with the same dependency sets as for variables y.

To simplify notations in the proof, we always write $\tilde{y}$ for the copies y' of variables y in $(**)$, if $(**)$ contains y', and otherwise for y. Of course, dependency sets of literals are defined w. r. t. to the DQBFs containing the literals. If it is necessary to make clear which DQBF we mean, we use subscripts: The dependency set of a literal p in ψ is written as $\mathrm{dep}_\psi(p)$, the corresponding dependency set in $\tilde{\psi}$ as $\mathrm{dep}_{\tilde{\psi}}(p)$.

The basic proof idea is as follows: We show that k is a covered literal w. r. t. ℓ in $C[1/x]$ and $\tilde{k}$ a covered literal w. r. t. $\tilde{\ell}$ in $C[0/x][y'/y\ldots]$.

Since ℓ and ℓ' depend in $\tilde{\psi}$ on one variable less than in ψ, we can apply the induction assumption and add the covered literals k and $\tilde{k}$ to $C[1/x]$ and $C[0/x][y'/y\ldots]$, resp., resulting in the equivalent formula

$$\tilde{\psi}' := \tilde{Q} : \big(\varphi \wedge (C \cup \{k\})\big)\big[1/x\big] \wedge \big(\varphi \wedge (C \cup \{k\})\big)\big[0/x\big]\big[y'/y \text{ for all } y \text{ with } x \in \mathrm{dep}_\psi(y)\big],$$

which is equivalent to $Q : \varphi \wedge (C \cup \{k\})$. This then shows the claim.

It remains to reason why k and $\tilde{k}$ are covered literals. We distinguish the following three cases:

– Case 1: $x \in C$:
Then $C[1/x]$ is satisfied and arbitrary literals (including k) can be added to

$C[1/x]$.
It remains to be proven that k may be added to C in $C[0/x][y'/y\ldots]$. We begin with the special case $k = x$. In that case adding k to C does not change C, since already $x \in C$ according to the case assumption.

Since $x \in \mathrm{dep}_\psi(\ell)$, ℓ is replaced by the copied variable ℓ' in $(**)$.
Then $C[0/x][y'/y\ldots] = C[y'/y\ldots] \setminus \{x\}$ contains ℓ', the copy of ℓ. All resolution candidates for $C[0/x][y'/y\ldots]$ in $R_{\tilde{\psi}}(C[0/x][y'/y\ldots], \ell')$ contain $\neg\ell'$, are included in $(**)$, and have the form $C'[0/x][y'/y\ldots]$. Now we consider an arbitrary resolution candidate $C'[0/x][y'/y\ldots] \in R_{\tilde{\psi}}(C[0/x][y'/y\ldots], \ell')$ and we prove that $C' \in R_\psi(C, \ell)$. We assume the contrary, i. e., $C' \notin R_\psi(C, \ell)$. Then there exists a literal m with $m \in C'$, $\neg m \in C$ and it holds $\mathrm{dep}_\psi(m) \subseteq \mathrm{dep}_\psi(\ell)$ in ψ. $m \neq x$, since otherwise C would be tautological ($\neg x \in C$ and $x \in C$ due to case assumption). $m \neq \neg x$, since otherwise $\neg x \in C'$, $C'[0/x][y'/y\ldots]$ would be satisfied, removed from $(**)$ and thus $C'[0/x][y'/y\ldots] \notin R_{\tilde{\psi}}(C[0/x][y'/y\ldots], \ell')$. Therefore $\tilde{m} \in C'[0/x][y'/y\ldots]$ and $\neg\tilde{m} \in C[0/x][y'/y\ldots]$. If $m \in V_\exists^\psi$, then $\mathrm{dep}_{\tilde{\psi}}(\tilde{m}) = \mathrm{dep}_\psi(m) \setminus \{x\}$, $\mathrm{dep}_{\tilde{\psi}}(\ell') = \mathrm{dep}_\psi(\ell) \setminus \{x\}$, i. e., $\mathrm{dep}_\psi(m) \subseteq \mathrm{dep}_\psi(\ell)$ in ψ implies $\mathrm{dep}_{\tilde{\psi}}(\tilde{m}) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell')$ in $\tilde{\psi}$. If $m \in V_\forall^\psi$, then $\mathrm{dep}_\psi(m) = \mathrm{var}(m)$, $\mathrm{dep}_{\tilde{\psi}}(\tilde{m}) = \mathrm{var}(\tilde{m})$, and as before $\mathrm{dep}_{\tilde{\psi}}(\ell') = \mathrm{dep}_\psi(\ell) \setminus \{x\}$, i. e., $\mathrm{dep}_\psi(m) \subseteq \mathrm{dep}_\psi(\ell)$ in ψ and $\mathrm{var}(m) \neq x$ implies $\mathrm{dep}_{\tilde{\psi}}(\tilde{m}) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell')$ in $\tilde{\psi}$. In both cases we have $\tilde{m} \in C'[0/x][y'/y\ldots]$, $\neg\tilde{m} \in C[0/x][y'/y\ldots]$, and $\mathrm{dep}_{\tilde{\psi}}(\tilde{m}) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell')$ which is a contradiction to $C'[0/x][y'/y\ldots] \in R_{\tilde{\psi}}(C[0/x][y'/y\ldots], \ell')$. Thus the assumption $C' \notin R_\psi(C, \ell)$ was wrong.
Since k is a covered literal for C w. r. t. ℓ, we then have $k \in C'$ and $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$. In the case $k = \neg x$ we then observe that $C'[0/x][y'/y\ldots]$ is satisfied, removed from $(**)$ and thus $C'[0/x][y'/y\ldots] \notin R_{\tilde{\psi}}(C[0/x][y'/y\ldots], \ell')$. This means that $R_{\tilde{\psi}}(C[0/x][y'/y\ldots], \ell') = \emptyset$. $R_{\tilde{\psi}}(C[0/x][y'/y\ldots], \ell') = \emptyset$ implies that $C[0/x][y'/y\ldots]$ is a blocked clause with blocking literal ℓ'. This in turn implies that $(C \cup \{k\})[0/x][y'/y\ldots]$ is blocked as well with blocking literal ℓ'. Thus we can remove $C[0/x][y'/y\ldots]$ from $\tilde{\psi}$ and add $(C \cup \{k\})[0/x][y'/y\ldots]$ resulting in an equivalent formula. In the case $k \notin \{x, \neg x\}$ we have $\mathrm{dep}_{\tilde{\psi}}(\tilde{k}) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell')$ (with the same arguments as for $\mathrm{dep}_{\tilde{\psi}}(\tilde{m})$ and $\mathrm{dep}_{\tilde{\psi}}(\ell')$ given above). Since $\tilde{k}$ is contained in every $C'[0/x][y'/y\ldots] \in R_{\tilde{\psi}}(C[0/x][y'/y\ldots], \ell')$ and $\mathrm{dep}_{\tilde{\psi}}(\tilde{k}) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell')$, $\tilde{k}$ is a covered literal for $C[0/x][y'/y\ldots]$. Since $\mathrm{dep}_{\tilde{\psi}}(\ell')$ contains one variable less than $\mathrm{dep}_\psi(\ell)$, we can apply the induction assumption and conclude that we can replace $C[0/x][y'/y\ldots]$ in $\tilde{\psi}$ by $(C \cup \{k\})[0/x][y'/y\ldots]$ resulting in an equivalent DQBF.

– Case 2: Neither x nor $\neg x$ are contained in C:
First we prove that k may be added to C in $C[1/x]$. We begin with the special case $k = \neg x$. In that case we can safely add k to C, since $C[1/x]$ and $(C \cup \{\neg x\})[1/x]$ are identical.
Since $x \in \mathrm{dep}_\psi(\ell)$, ℓ is replaced by the copied variable ℓ' in $(**)$. All resolution candidates for $C[1/x]$ in $R_{\tilde{\psi}}(C[1/x], \ell)$ contain $\neg\ell$, are included in $(*)$, and

have the form $C'[1/x]$. Now we consider an arbitrary resolution candidate $C'[1/x] \in R_{\tilde{\psi}}(C[1/x], \ell)$ and we prove that $C' \in R_\psi(C, \ell)$. We assume the contrary, i. e., $C' \notin R_\psi(C, \ell)$. Then there exists a literal m with $m \in C'$, $\neg m \in C$ and it holds $\mathrm{dep}_\psi(m) \subseteq \mathrm{dep}_\psi(\ell)$. $m \notin \{x, \neg x\}$ due to case assumption $x \notin C$ and $\neg x \notin C$. Therefore $m \in C'[1/x]$ and $\neg m \in C[1/x]$. Exactly as in Case 1 we can conclude that $\mathrm{dep}_\psi(m) \subseteq \mathrm{dep}_\psi(\ell)$ implies $\mathrm{dep}_{\tilde{\psi}}(m) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell)$. Altogether we have $m \in C'[1/x]$, $\neg m \in C[1/x]$, and $\mathrm{dep}_{\tilde{\psi}}(m) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell)$ which is a contradiction to $C[1/x] \in R_{\tilde{\psi}}(C[1/x], \ell)$. Thus the assumption $C' \notin R_\psi(C, \ell)$ was wrong.
Since k is a covered literal for C w. r. t. ℓ, we then have $k \in C'$ and $\mathrm{dep}_\psi(k) \subseteq \mathrm{dep}_\psi(\ell)$. Therefore $k \in C'[1/x]$. In the case $k = x$ we then observe that $C'[1/x]$ is satisfied, removed from $(*)$ and thus $C'[1/x] \notin R_{\tilde{\psi}}(C[1/x], \ell)$. This means that $R_{\tilde{\psi}}(C[1/x], \ell) = \emptyset$. $R_{\tilde{\psi}}(C[1/x], \ell) = \emptyset$ implies that $C[1/x]$ is a blocked clause with blocking literal ℓ. This in turn implies that $(C \cup \{k\})[1/x]$ is blocked as well with blocking literal ℓ. Thus we can remove $C[1/x]$ from $\tilde{\psi}$ and add $(C \cup \{k\})[1/x]$ resulting in an equivalent formula. In the case $k \notin \{x, \neg x\}$ we have $\mathrm{dep}_{\tilde{\psi}}(k) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell)$ (with the same arguments as for $\mathrm{dep}_{\tilde{\psi}}(m)$ and $\mathrm{dep}_{\tilde{\psi}}(\ell)$ given above). Since k is contained in every $C'[1/x] \in R_{\tilde{\psi}}(C[1/x], \ell)$ and $\mathrm{dep}_{\tilde{\psi}}(k) \subseteq \mathrm{dep}_{\tilde{\psi}}(\ell)$, k is a covered literal for $C[1/x]$. Since $\mathrm{dep}_{\tilde{\psi}}(\ell)$ contains one variable less than $\mathrm{dep}_\psi(\ell)$, we can apply the induction assumption and conclude that we can replace $C[1/x]$ in $\tilde{\psi}$ by $(C \cup \{k\})[1/x]$ resulting in an equivalent DQBF.

The proof that k may be added to C in $C[0/x][y'/y \ldots]$ is analogous. It starts with the special case $k = x$ where k can be safely added to C, since $C[0/x][y'/y \ldots]$ and $(C \cup \{x\})[0/x][y'/y \ldots]$ are identical. For the case $k \neq x$ we observe that $R_{\tilde{\psi}}(C[0/x][y'/y \ldots], \ell') = \emptyset$, both $C[0/x][y'/y \ldots]$ and $(C \cup \{x\})[0/x][y'/y \ldots]$ are blocked with blocking literal ℓ', i. e., we can remove $C[0/x][y'/y \ldots]$ from $\tilde{\psi}$ and add $(C \cup \{x\})[0/x][y'/y \ldots]$ resulting in an equivalent formula. In the case $k \notin \{x, \neg x\}$ we conclude that k is a covered literal for $C[0/x][y'/y \ldots]$ and we apply the induction assumption.

– Case 3: $\neg x \in C$:
This case is analogous to the first one. □

A.5 Structure Extraction

Theorem 10. *Let $\psi = Q : \varphi$ be a DQBF, $\varphi^f \subseteq \varphi$ the defining clauses for the relationship $y \equiv f(\ell_1, \ldots, \ell_k)$. Then ψ is equivalent to*

$$Q \setminus \{y\} : (\varphi \setminus \varphi^f)[f(\ell_1, \ldots, \ell_k)/y]$$

if the following conditions are satisfied:

1. $y \in V_\exists^\psi$,
2. *for $i = 1, \ldots, k$ we have* $\mathrm{dep}_\psi(\ell_i) \subseteq \mathrm{dep}_\psi(y)$.

Proof. We set $\psi' := Q \setminus \{y\} : (\varphi \setminus \varphi^f)[f(\ell_1, \ldots, \ell_k)/y]$ and show that ψ and ψ' are equivalent.

First assume that ψ is unsatisfiable. Then there is no set of Skolem functions which turns φ into a tautology. In particular sets of Skolem functions for which $s_y = f(\ell_1, \ldots, \ell_k)[s_{y'}/y' \text{ for } y' \in \{\text{var}(\ell_1), \ldots, \text{var}(\ell_k)\} \cap V_\exists^\psi]$ holds do not turn φ into a tautology. (Note that s_y defined in that way is an admissible Skolem function due to condition (2) in the theorem.) Hence, $Q \setminus \{y\} : \varphi[f(\ell_1, \ldots, \ell_k)/y]$ is unsatisfiable. Since φ^f is equivalent to $y \equiv f(\ell_1, \ldots, \ell_k)$, $\varphi^f[f(\ell_1, \ldots, \ell_k)/y]$ is a tautology. Therefore $\varphi[f(\ell_1, \ldots, \ell_k)/y]$ and $(\varphi \setminus \varphi^f)[f(\ell_1, \ldots, \ell_k)/y]$ are equivalent. This shows that ψ' is unsatisfiable.

Now assume that ψ is satisfiable. This implies that there are Skolem functions $s_{y_1}, \ldots, s_{y_m}$ for $y_1, \ldots, y_m$. Because of the defining clauses that encode $y \equiv f(\ell_1, \ldots, \ell_k)$, the Skolem function s_y for y satisfies the relationship $s_y = f(\ell_1, \ldots, \ell_k)[s_{y'}/y' \text{ for } y' \in V_\exists^\psi \setminus \{y\}]$. Thus $\{s_{y_1}, \ldots, s_{y_m}\} \setminus \{s_y\}$ is a set of Skolem functions for $Q \setminus \{y\} : \varphi[f(\ell_1, \ldots, \ell_k)/y]$. This means that $Q \setminus \{y\} : \varphi[f(\ell_1, \ldots, \ell_k)/y]$ is satisfiable. Satisfiability of $Q \setminus \{y\} : \varphi[f(\ell_1, \ldots, \ell_k)/y]$ implies satisfiability of $Q \setminus \{y\} : (\varphi \setminus \varphi^f)[f(\ell_1, \ldots, \ell_k)/y]$. □