Making IP = PSPACE Practical: Efficient Interactive Protocols for BDD Algorithms

Eszter Couillard[1], Philipp Czerner[1(✉)], Javier Esparza[1], and Rupak Majumdar[2]

[1] Technical University of Munich, Munich, Germany
{coillar,czerner,esparza}@in.tum.de
[2] Max Planck Institute for Software Systems (MPI-SWS), Kaiserslautern, Germany
rupak@mpi-sws.org

Abstract. We show that interactive protocols between a prover and a verifier, a well-known tool of complexity theory, can be used in practice to certify the correctness of automated reasoning tools.

Theoretically, interactive protocols exist for all PSPACE problems. The verifier of a protocol checks the prover's answer to a problem instance in probabilistic polynomial time, with polynomially many bits of communication, and with exponentially small probability of error. (The prover may need exponential time.) Existing interactive protocols are not used in practice because their provers use naive algorithms, inefficient even for small instances, that are incompatible with practical implementations of automated reasoning.

We bridge the gap between theory and practice by means of an interactive protocol whose prover uses BDDs. We consider the problem of counting the number of assignments to a QBF instance (#CP), which has a natural BDD-based algorithm. We give an interactive protocol for #CP whose prover is implemented on top of an extended BDD library. The prover has only a linear overhead in computation time over the natural algorithm.

We have implemented our protocol in blic, a certifying tool for #CP. Experiments on standard QBF benchmarks show that blic is competitive with state-of-the-art QBF-solvers. The run time of the verifier is negligible. While loss of absolute certainty can be concerning, the error probability in our experiments is at most 10^{-10} and reduces to 10^{-10k} by repeating the verification k times.

This work was supported by an ERC Advanced Grant (787367: PaVeS), by the Deutsche Forschungsgemeinschaft project 389792660 TRR 248—CPEC, and by the Research Training Network of the Deutsche Forschungsgemeinschaft (DFG) (378803395: ConVeY).

C. Enea and A. Lal (Eds.): CAV 2023, LNCS 13966, pp. 437–458, 2023.
https://doi.org/10.1007/978-3-031-37709-9_21

1 Introduction

Automated reasoning tools often underlie our assertions about the correctness of critical hardware and software components. In recent years, the scope and scalability of these techniques have grown significantly.

Automated reasoning tools are not immune to bugs. If we are to trust their verdict, it is important that they provide evidence of their correct behaviour. A substantial amount of research has gone into proof-producing automated reasoning tools [4,14,16,22,23]. These works define a notion of "correctness certificate" suitable for the reasoning problem at hand, and adapt the reasoning engine to produce independently checkable certificates. For example, SAT solvers produce either a satisfying assignment or a proof of unsatisfiability in some proof system, e.g. resolution (see [16] for a survey). Extending such certificates beyond boolean SAT is an active area of current research [3,4,18,24,29].

In the worst case, the size of certificates grows exponentially in the size of the input, even for boolean unsatisfiability (unless $\mathsf{NP} = \mathsf{coNP}$). If users have limited computational or communication resources, transferring and checking large certificates becomes a burden. Large certificates are not just a theoretical curiosity. In practice, resolution proofs for complex SAT problems may run to petabytes [15]. Ideally, we would prefer "small" certificates (polynomial in the size of the input) which can be checked independently in polynomial time.

The $\mathsf{IP} = \mathsf{PSPACE}$ theorem proves that certification with polynomial verification time is possible for any problem in PSPACE, provided one trades off absolute certainty for certainty with high probability [27]. The complexity class IP consists of those languages for which there is a polynomial-round, complete and sound *interactive protocol* [1,2,13,20]—a sequence of interactions between a (computationally unbounded) prover and a (computationally bounded) verifier after which the verifier decides whether the prover correctly performed a computation. The protocol is complete if, whenever an input belongs to the language, there is an *honest prover* who can convince a polynomial-time randomised verifier in a polynomial number of rounds. The protocol is sound if, whenever an input does not belong to the language, the Verifier will reject the input with high probability—no matter what certificates are provided to the Verifier. That is, a "Prover" cannot fool the certification process.

Since every language in PSPACE has an interactive protocol, there are interactive protocols for UNSAT, QBF, *counting* QBF, safety verification of concurrent state machines, etc. Observe that the prover of a protocol may perform exponential time computations (which is unavoidable unless $\mathsf{P} = \mathsf{PSPACE}$), but the verifier only requires polynomial time in the original input.

If interactive protocols provide a foundation for small and efficiently verifiable certificates (at least for problems in PSPACE), why are they not in widespread practice? We believe the reason to be the following: for asymptotic complexity purposes, it suffices to use honest provers with best-case exponential complexity that naively enumerate all possibilities. Such provers are incompatible with automated reasoning tools, which use more sophisticated data structures and heuristics to scale to real-world examples. So we need to make *practical algorithms*

for automated reasoning *efficiently certifying*. We call an algorithm *efficiently certifying* if, in addition to computing the output, it can execute the steps of an honest prover in an interactive protocol with only polynomial overhead over its running time.

In this paper, we show that algorithms using reduced ordered binary decision diagrams (henceforth called BDDs) [9] can be made efficiently certifying. We consider #CP, the problem of computing the number of satisfying assignments of a *circuit with partial evaluation (CP)*. Besides boolean nodes, a CP contains *partial evaluation* nodes $\pi_{[x:=\mathsf{false}]}$ (resp., $\pi_{[x:=\mathsf{true}]}$) that take a boolean predicate as input, say φ, and output the result of setting x to false (resp., true) in φ. #CP generalises SAT, QBF, and *counting* SAT (#SAT), and has a natural algorithm using BDDs: Compute BDDs for each node of the circuit in topological order, and count the accepting paths of the final BDD.

The theoretical part of the paper proceeds in two steps. First, we present CPCertify, a complete and sound interactive protocol for #CP. CPCertify is similar to the SumCheck protocol [20]. It involves encoding boolean formulas as polynomials over a finite field. The prover is responsible for producing certain polynomials from the original circuit and evaluating them at points of the field chosen by the verifier. These polynomials are either multilinear (all exponents are at most 1) or quadratic (at most 2).

Second, we show that an honest prover in CPCertify can be implemented on top of a suitably extended BDD library. The run times of the certifying BDD algorithms are only a constant overhead over the computation time without certification—they depend linearly on the total number of nodes of the intermediate BDDs computed by the prover to solve the #CP instance. We use two key insights. The first is an encoding of multilinear polynomials as BDDs; we show that the intermediate BDDs represent all the multilinear polynomials a prover needs during the run of CPCertify. The second shows that the quadratic polynomials correspond to *intermediate steps* during the computation of the intermediate BDDs. We extend BDDs with additional "book-keeping" nodes that allow the prover to also compute the quadratic polynomials while solving the problem. So computing the polynomials required by CPCertify has *zero* additional cost; the only overhead is the cost of evaluating the polynomials at the field points chosen by the verifier.

We have implemented a certifying #CP solver based on our extended BDD library. Our experiments show that the solver is competitive with state-of-the-art non-certifying QBF solvers, and can outperform certifying QBF solvers based on BDDs. The number of bytes exchanged between the prover and the verifier are an order of magnitude smaller, and Verifier's run time several orders of magnitude smaller, than current encodings of QBF proofs, while bounding the error probability to below 10^{-10}. Thus, our results open the way for practically efficient, probabilistic certification of automated reasoning problems using interactive protocols.

Additional Related Work. Proof systems for SAT and QBF remain an active area of research—both in theoretical proof complexity and in practical tool devel-

opment. Jussila, Sinz, and Biere [17,28] showed how to extract extended resolution proofs from BDD operations. This is the basis for proof-producing SAT and QBF solvers based on BDDs [6–8]. As in our work, the proof uses intermediate nodes produced in the construction of the BDD operations. We focus on interactive certification instead of extended resolution proofs, which can be exponentially larger than the input formula.

Recently, Luo et al. [21] consider the problem of providing *zero-knowledge* proofs of unsatisfiability, a motivation similar but not equal to ours. Their techniques require the verifier to work in time polynomial in the proof, which can be exponentially bigger than the input formula. In contrast, the verifier of CPCertify runs in polynomial time in the input. Since any language in PSPACE has a zero knowledge proof [5], our protocol can in principle be made zero knowledge. Whether that system scales in practice is left for future work.

Full Version. Detailed proofs can be found in the full version of the paper [11].

2 Preliminaries

The Class IP. An *interactive protocol* between a *Prover* and a *Verifier* consists of a sequence of interactions in which a Verifier asks questions to a Prover, receives responses to the questions, and must ultimately decide if a common input x belongs to a language. The computational power of the Prover is unbounded but the Verifier is a randomised, polynomial-time algorithm.

Formally, let P, V denote (deterministic) Turing machines.

We say that $(r; m_1, ..., m_{2k})$ is a *k-round interaction*, with $r, m_1, ..., m_{2k} \in \{0,1\}^*$, if $m_{i+1} = V(r, m_1, ..., m_i)$ for even i and $m_{i+1} = P(m_1, ..., m_i)$ for odd i. We think of r as an additional sequence of bits given to Verifier V that is chosen randomly. The *output* $\text{out}(P, V)(x, r, k)$ is defined as m_{2k}, where $(r; m_1, ..., m_{2k})$ is the unique k-round interaction with $m_1 = x$.

A language L belongs to IP if there are V, P_H and polynomials p_1, p_2, p_3, s.t. $V(r, x, m_2, ..., m_i)$ runs in time $p_1(|x|)$ for all $r, x, m_2, ..., m_i$, and, for each x and an $r \in \{0,1\}^{p_2(|x|)}$ chosen uniformly at random:

1. (*Completeness*) $x \in L$ implies $\text{out}(P_H, V)(x, r, p_3(|x|)) = 1$ with probability 1, and
2. (*Soundness*) $x \notin L$ implies that for all P we have $\text{out}(P, V)(x, r, p_3(|x|)) = 1$ with probability at most $2^{-|x|}$.

Intuitively, in an interactive protocol, a computationally unbounded Prover interacts with a randomised polynomial-time Verifier for k rounds. In each round, Verifier sends probabilistic "challenges" to Prover, based on the input and the answers to prior challenges, and receives answers from Prover. At the end of k rounds, Verifier decides to accept or reject the input. The completeness property ensures that if the input belongs to the language L, then there is an "honest" Prover P_H who can always convince Verifier that indeed $x \in L$. If the input does not belong to the language, then the soundness property ensures that Verifier

rejects the input with high probability no matter how a (dishonest) Prover tries to convince them.

It is known that IP = PSPACE [20,27], that is, every language in PSPACE has a polynomial-round interactive protocol. The proof exhibits an interactive protocol for the language QBF of true quantified boolean formulae; in particular, the honest Prover is a polynomial space, exponential time algorithm that uses a truth table representation of the formula to implement the protocol.

Polynomials. Interactive protocols make extensive use of polynomials over some prime finite field $\mathbb{F}$.

Let X be a finite set of variables. We use $x, y, z, \ldots$ for variables and $p, q, \ldots$ for polynomials. When we write a polynomial explicitly, we write it in brackets, e.g. $[3xy - z^2]$. We write $\mathbf{1}$ and $\mathbf{0}$ for the polynomials $[1]$ and $[0]$, respectively. We use the following operations on polynomials:

- *Sum, difference, and product.* Denoted $p+q$, $p-q$, $p \cdot q$, and defined as usual. For example, $[3xy-z^2]+[z^2+yz] = [3xy+yz]$ and $[x+y]\cdot[x-y] = [x^2-y^2]$.
- *Partial evaluation.* Denoted $\pi_{[x:=a]}\, p$, it returns the result of setting variable x to the field element a in the polynomial p, e.g. $\pi_{[x:=5]}[3xy-z^2] = [15y-z^2]$.
- *Degree reduction.* Denoted $\delta_x\, p$. It reduces the degree of x in all monomials of the polynomial to 1. For example, $\delta_x[x^3y + 3x^2 + 7z^2] = [xy + 3x + 7z^2]$.

A *(partial) assignment* is a (partial) mapping $\sigma : X \to \mathbb{F}$. We write $\Pi_\sigma\, p$ for $\pi_{[x_1:=\sigma(x_1)]} \ldots \pi_{[x_k:=\sigma(x_k)]}\, p$, where $x_1, \ldots, x_k$ are the variables for which σ is defined. Additionally, we call σ *binary* if $\sigma(x) \in \{0, 1\}$ for each $x \in X$.

Binary and Multilinear Polynomials. A polynomial is *multilinear in x* if the degree of x in p is 0 or 1. A polynomial is *multilinear* if it is multilinear in all its variables. For example, $[xy - y^2]$ is multilinear in x but not in y, and $[3xy - 2zy]$ is multilinear. A polynomial p is *binary* if $\Pi_\sigma\, p \in \{\mathbf{0}, \mathbf{1}\}$ for every binary assignment σ. Two polynomials p, q are *binary equivalent*, denoted $p \equiv_b q$, if $\Pi_\sigma\, p = \Pi_\sigma\, q$ for every binary assignment σ. (Note that non-binary polynomials can be binary equivalent.)

3 Circuits with Partial Evaluation

We introduce circuits with partial evaluation (CP), a compact representation of quantified boolean formulae, and formulate #CP, the problem of counting the number of satisfying assignments of a CP. #CP generalises QBF, the satisfiability problem for quantified boolean formulas. Figure 1 shows an example of a CP. Informally, it is a directed acyclic graph whose nodes are labelled with variables, boolean operators, or *partial evaluation operators* $\pi_{[x:=b]}$. Intuitively, $\pi_{[x:=b]}\varphi$ sets the variable x to the truth value b in the formula φ. In this way, each node of a circuit stands for a boolean function, and the complete circuit stands for the boolean function of the root. Figure 1 shows the formulae represented by each node.

Definition 1. *Let X denote a finite set of* variables *and $S \subseteq X$. A* circuit with partial evaluation and variables in S *(S-CP) has the form*

– true, false, *or* x, *where* $x \in S$,
– $\neg\varphi$, $\varphi \wedge \psi$, *or* $\varphi \vee \psi$, *where* φ, ψ *are* S*-CPs, or*
– $\pi_{[y:=b]}\,\varphi$, *where* $y \in X \setminus S$, $b \in \{\mathsf{true}, \mathsf{false}\}$, *and* φ *is an* $(S \cup \{y\})$*-CP.*

The set of free variables *of a* S*-CP* φ *is* $\mathrm{free}(\varphi) := S$. *The* children *of a CP are inductively defined as follows:* true, false, *and* x *have no children; the children of* $\varphi \wedge \psi$ *and* $\varphi \vee \psi$ *are* φ *and* ψ*; and the only child of* $\neg\varphi$ *and* $\pi_{[y:=b]}\,\varphi$ *is* φ. *The set of* descendants *of* φ *is the smallest set* M *containing* φ *and all children of every element of* M. *The* size *of* φ *is* $|\varphi| := |M|$.

We represent a CP φ as a directed acyclic graph. The nodes of the graph are the descendants of φ. A CP φ encodes a boolean predicate P_φ, which maps assignments $\sigma\colon \mathrm{free}(\varphi) \to \{\mathsf{false}, \mathsf{true}\}$ to a truth value $P_\varphi(\sigma) \in \{\mathsf{false}, \mathsf{true}\}$. It does so in the obvious manner, e.g., $P_x(\sigma) := \sigma(x)$, $P_{\varphi\wedge\psi}(\sigma) := P_\varphi(\sigma) \wedge P_\psi(\sigma)$, etc. We use $\pi_{[x:=b]}$ as partial evaluation operator, so $P_{\pi_{[x:=b]}\varphi}(\sigma) := P_\varphi(\sigma \cup \{x \mapsto b\})$. Intuitively, $\pi_{[x:=b]}\,\varphi$ replaces each occurrence of x in φ by b. An assignment σ *satisfies* φ if $P_\varphi(\sigma) = \mathsf{true}$. We define the macros

$$\forall_x \varphi := \pi_{[x:=0]}\,\varphi \wedge \pi_{[x:=1]}\,\varphi$$
$$\exists_x \varphi := \pi_{[x:=0]}\,\varphi \vee \pi_{[x:=1]}\,\varphi$$

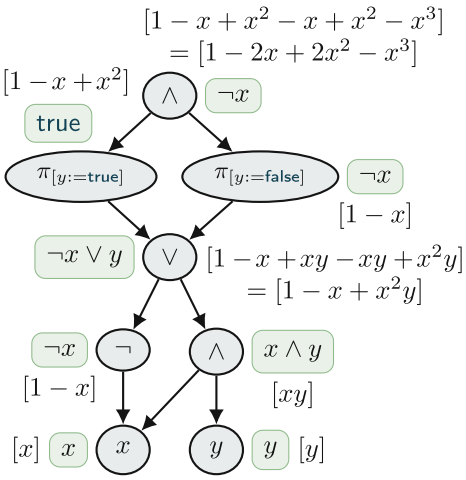

Fig. 1. A CP (Sect. 3), the boolean functions represented by each node (in boxes), and the arithmetisation of the formulae (Sect. 4.1).

Figure 1 shows a CP for the quantified boolean formula $\forall_y(\neg x \vee (x \wedge y))$.

We consider the following problem:

#CP **Input** CP φ.
Output The number of satisfying assignments of φ.

Given a quantified boolean formula, we can use the macros for quantifiers to construct in linear time an equivalent CP, i.e., a CP with the same satisfying assignments. Similarly, #SAT instances can also be reduced to #CP.

Structure of the Rest of the Paper. In Sect. 4, we give an interactive protocol for #CP called CPCertify. In Sect. 5, we implement an honest Prover for CPCertify on top of an extended BDD-based algorithm for #CP. The prover runs in time polynomial in the size of the largest BDD for any of the subcircuits of the initial circuit. Together, these results yield our main result, Theorem 1, showing that any BDD-based algorithm can be modified to run an interactive protocol with small polynomial overhead. Finally, Sect. 6 presents empirical results.

4 An Interactive Protocol for #CP

In this section we describe an interactive protocol for #CP, following the SumCheck protocol of [20]. Section 4.1 introduces arithmetisation, a technique to transform #CP into an equivalent problem about polynomials. Section 4.2 shows how to transform #CP into an equivalent problem about evaluating polynomials *of low degree*. Finally, Sect. 4.3 presents an interactive protocol for this problem.

4.1 Arithmetisation

We define a mapping $[\![\cdot]\!]$ that assigns to each CP φ a polynomial $[\![\varphi]\!]$ over the variables free(φ), called the *arithmetisation* of φ:

- $[\![\mathsf{true}]\!] := \mathbf{1}$; $[\![\mathsf{false}]\!] := \mathbf{0}$; $[\![x]\!] := [x]$ for every $x \in X$; and $[\![\neg\varphi]\!] := \mathbf{1} - [\![\varphi]\!]$;
- $[\![\varphi \wedge \psi]\!] := [\![\varphi]\!] \cdot [\![\psi]\!]$; and $[\![\varphi \vee \psi]\!] := [\![\varphi]\!] + [\![\psi]\!] - [\![\varphi]\!] \cdot [\![\psi]\!]$;
- $[\![\pi_{[x:=b]}\,\varphi]\!] := \pi_{[x:=[\![b]\!]]}[\![\varphi]\!]$, with $x \in \text{free}(\varphi)$, $b \in \{\mathsf{true}, \mathsf{false}\}$.

Figure 1 also shows the polynomials corresponding to the nodes of the CP.

Let $\mathbb{F}$ be a fixed prime finite field. Given an arbitrary truth assignment $\sigma\colon X \to \{\mathsf{true}, \mathsf{false}\}$, let $\overline{\sigma}\colon X \to \mathbb{F}$ be the binary assignment given by $\overline{\sigma}(x) = 1$ if $\sigma(x) = \mathsf{true}$ and $\overline{\sigma}(x) = 0$ if $\sigma(x) = \mathsf{false}$, where 0 and 1 denote the additive and multiplicative identities in $\mathbb{F}$. The mapping $[\![\cdot]\!]$ is defined to satisfy the following property, whose proof is immediate:

Proposition 1. *Let φ be an S-CP encoding some boolean predicate P_φ. Then $P_\varphi(\sigma) = \Pi_{\overline{\sigma}}[\![\varphi]\!]$ for every truth assignment σ to S.*

So, intuitively, the polynomial $[\![\varphi]\!]$ is a conservative extension of the predicate P_φ: It returns the same values for all binary assignments. Accordingly, in the rest of the paper we abuse language and write σ instead of $\overline{\sigma}$ for the binary assignment corresponding to the truth assignment σ.

Observe that #CP can be reformulated as follows: given a CP φ, compute the number of binary assignments σ s.t. $\Pi_\sigma[\![\varphi]\!] = \mathbf{1}$.

4.2 Degree Reduction

Given a CP φ, its associated polynomial can have degree exponential in the height of φ. Since we are ultimately interested in evaluating polynomials over binary assignments, and since $x^2 = x$ for $x \in \{0, 1\}$, we can convert polynomials to low degree without changing their behaviour on binary assignments.

For this, we use a *degree-reduction* operator δ_x for every variable x. The operator $\delta_x p$ reduces the exponent of all powers of x in p to 1. For example, $\delta_x[x^2y + 3xy^2 - 2x^3y^2 + 4] = [xy + 3xy^2 - 2xy^2 + 4]$. Observe that $\delta_x p \equiv_b p$. Instead of working on the input CP directly, we first convert it into a *circuit with partial evaluation and degree reduction* by inserting degree-reduction operators after binary operations. This ensures all intermediate polynomials obtained by arithmetisation have low degree.

Definition 2. *A* circuit with partial evaluation and degree reduction over the set S of variables (S-CPD) *is defined in the same manner as an S-CP, extended as follows:*

- *if φ is an S-CPD and $x \in S$, then $\delta_x\varphi$ is an S-CPD,*
- $[\![\delta_x\varphi]\!] := \delta_x[\![\varphi]\!]$, *and*
- *φ is the only child of $\delta_x\varphi$.*

For an S-CPD φ we define $\mathrm{free}(\varphi)$, $|\varphi|$, *children, descendants, and the graphical representation as for S-CPs.*

We convert a CP φ into a CPD $\mathrm{conv}(\varphi)$ by adding a degree-reduction operator for each free variable before any binary operation.

Definition 3. *Given a CP φ with* $\mathrm{free}(\varphi) = \{x_1, ..., x_k\}$, *its associated CPD* $\mathrm{conv}(\varphi)$ *is inductively defined as follows:*

- $\mathrm{conv}(\mathsf{false}) = \mathsf{false}$, $\mathrm{conv}(\mathsf{true}) := \mathsf{true}$,
- $\mathrm{conv}(\neg\psi) := \neg\,\mathrm{conv}(\psi)$, $\mathrm{conv}(\pi_{[x:=b]}\,\psi) := \pi_{[x:=b]}\,\mathrm{conv}(\psi)$, *and*
- $\mathrm{conv}(\psi_1 \circledast \psi_2) := \delta_{x_1}...\delta_{x_k}(\mathrm{conv}(\psi_1) \circledast \mathrm{conv}(\psi_2))$, *for* $\circledast \in \{\vee, \wedge\}$.

Figure 2 shows the CPD $\mathrm{conv}(\varphi)$ for the CP φ of Fig. 1, together with the polynomials corresponding to each node.

We collect some basic properties of CPDs:

Lemma 1. *Let φ be a CP.*

(a) $[\![\mathrm{conv}(\varphi)]\!]$ *is a binary multilinear polynomial and* $[\![\mathrm{conv}(\varphi)]\!] \equiv_b [\![\varphi]\!]$.
(b) For every descendant ψ of $\mathrm{conv}(\varphi)$, $[\![\psi]\!]$ *has maximum degree* 2.

CPDs have another useful property. Recall that given a CP φ we are interested in its number of satisfying assignments. The next lemma shows that this number can be computed by evaluating the polynomial $[\![\mathrm{conv}(\varphi)]\!]$ on *a single input.*

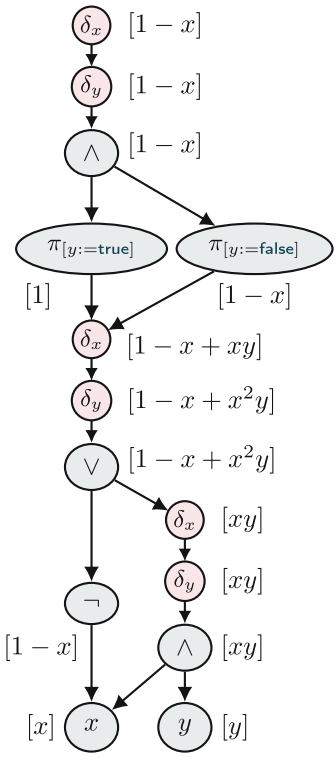

Fig. 2. CPD and polynomials for the CP of Fig. 1.

Lemma 2. *A CP φ with n free variables has $m < |\mathbb{F}|$ satisfying assignments iff $\Pi_\sigma[\![\mathrm{conv}(\varphi)]\!] = m \cdot 2^{-n}$, where σ is the assignment satisfying $\sigma(x) := 2^{-1}$ in the field $\mathbb{F}$ for every variable x.*[1]

[1] Any prime field $\mathbb{F}$ with $|\mathbb{F}| > 2$ has an element c such that $2c = 1$.

4.3 CPCertify: An Interactive Protocol for #CP

We describe an interactive protocol, called CPCertify, for a CP φ with n free variables. Let X denote the variables used in φ. Prover and Verifier fix a finite field with at least $m+1$ elements, where m is an upper bound on the number of assignments (e.g. $m = 2^n$). Prover tries to convince the Verifier that $\Pi_\sigma[\![\mathrm{conv}(\varphi)]\!] = K$ for some $K \in \mathbb{F}$.

In the protocol, Verifier challenges Prover to compute polynomials of the form $\Pi_\sigma([\![\psi]\!])$, where ψ is a node of the CPD $\mathrm{conv}(\varphi)$ and $\sigma\colon \mathrm{free}(\psi) \to \mathbb{F}$ is a (non-binary!) assignment; we call the expression $\Pi_\sigma[\![\mathrm{conv}(\psi)]\!]$ a *challenge*. Observe that all assignments are chosen by Verifier. Prover answers with some $k \in \mathbb{F}$. We call the expression $\Pi_\sigma[\![\mathrm{conv}(\psi)]\!] = k$ a *claim*, or the *answer* to the challenge $\Pi_\sigma[\![\mathrm{conv}(\psi)]\!]$.

CPCertify consists of an initialisation and a number of rounds, one for each descendant of $\mathrm{conv}(\varphi)$. Rounds are executed in topological order, starting at the root, i.e. at $\mathrm{conv}(\varphi)$ itself. The structure of a round for a node ψ of $\mathrm{conv}(\varphi)$ depends on whether ψ is an internal node (including the root), or a leaf.

At each point, Verifier keeps track of a set $\mathcal{C}$ of claims that must be checked.

Initialisation. Verifier sends Prover the challenge $\Pi_\sigma[\![\mathrm{conv}(\varphi)]\!]$, where $\sigma(x) := 2^{-1}$ for every $x \in \mathrm{free}(\varphi)$. Prover returns the claim $\Pi_\sigma[\![\mathrm{conv}(\varphi)]\!] = K$ for some $K \in \mathbb{F}$. (By Lemma 2, this amounts to claiming that φ has $K \cdot 2^n$ satisfying assignments.) Verifier initialises $\mathcal{C} := \{\Pi_\sigma[\![\mathrm{conv}(\varphi)]\!] = K\}$.

Round for an Internal Node. A round for an internal node ψ runs as follows:

(a) Verifier collects all claims $\{\Pi_{\sigma_i}[\![\psi]\!] = k_i\}_{i=1}^m$ in $\mathcal{C}$ relating to ψ, with assignments $\sigma_1, \ldots, \sigma_m\colon \mathrm{free}(\psi) \to \mathbb{F}$ and $k_1, ..., k_m \in \mathbb{F}$. (Initially $\psi = \mathrm{conv}(\varphi)$ and the only claim is $\Pi_\sigma[\![\mathrm{conv}(\varphi)]\!] = K$.)

(b) If $m > 1$, Verifier interacts with Prover to compute a unique claim $\Pi_\sigma[\![\psi]\!] = k$ such that very likely[2] the claim is true only if all claims $\{\Pi_{\sigma_i}[\![\psi]\!] = k_i\}_{i=1}^m$ are true. For this, Verifier sends a number of challenges, and checks that the answers are *consistent* with the prior claims. Based on these answers, Verifier then derives new claims. (See "Description of step (b)" below.)

(c) Verifier interacts with Prover to compute a claim $\Pi_{\sigma'}[\![\psi']\!] = k'$ for each child ψ' of ψ. This is similar to (b): if $\Pi_\sigma[\![\psi]\!] \neq k$, i.e. the unique claim from (b) does not hold, then very likely one of the resulting claims will be wrong. Depending on the type of ψ, the claims are computed based on the answers of Prover to challenges sent by Verifier. (See "Description of step (c)" below.)

(d) In total, Verifier removed the claims $\{\Pi_{\sigma_i}[\![\psi]\!] = k_i\}_{i=1}^m$ from $\mathcal{C}$, and replaced them by one claim $\Pi_{\sigma'}[\![\psi']\!] = k'$ for each child ψ' of ψ.

Observe that, since a node ψ can be a child of several nodes, Verifier may collect multiple claims for ψ, one for each parent node.

Round for a Leaf. If ψ is a leaf, then $\psi = x$ for a variable x, or $\psi \in \{\mathsf{true}, \mathsf{false}\}$. Verifier removes all claims $\{\Pi_{\sigma_i}[\![\psi]\!] = k_i\}_{i=1}^m$ from $\mathcal{C}$, computes the values $c_i := \Pi_{\sigma_i}[\![\psi]\!]$, and rejects if $k_i \neq c_i$ for any i.

[2] The precise bound on the failure probability will be given in Proposition 2.

Observe that if all claims made by Prover about leaves are true, then very likely Prover's initial claim is also true.

Description of Step (b). Let $\{\Pi_{\sigma_i}[\![\psi]\!] = k_i\}_{i=1}^m$ be the claims in $\mathcal{C}$ relating to node ψ. Verifier and Prover conduct step (b) as follows:

(b.1) While there exists $x \in X$ s.t. $\sigma_1(x), \ldots, \sigma_m(x)$ are not pairwise equal:
  (b.1.1) For every $i \in \{1, ..., m\}$, let σ_i' denote the partial assignment which is undefined on x and otherwise matches σ_i. Verifier sends the challenges $\{\Pi_{\sigma_i'}[\![\psi]\!]\}_{i=1}^m$ to Prover. Prover answers with claims $\{\Pi_{\sigma_i'}[\![\psi]\!] = p_i\}_{i=1}^m$. Note that $p_1, \ldots, p_m$ are univariate polynomials with free variable x.
  (b.1.2) Verifier checks whether $k_i = \pi_{[x:=\sigma_i(x)]}\, p_i$ holds for each i. If not, Verifier rejects. Otherwise, Verifier picks $r \in \mathbb{F}$ uniformly at random and updates $\sigma_i(x) := r$ and $k_i := \pi_{[x:=r]} p_i$ for every $i \in \{1, ..., m\}$.
(b.2) If after exiting the loop the values $k_1, ..., k_m$ are not pairwise equal, Verifier rejects. Otherwise (that is, if $k_1 = k_2 = \cdots = k_m$), the set $\mathcal{C}$ now contains a unique claim $\Pi_\sigma[\![\psi]\!] = k$ relating to ψ.

Example 1. Consider the case in which $X = \{x\}$, and Prover has made two claims, $\Pi_{\sigma_1}[\![\psi]\!] = k_1$ and $\Pi_{\sigma_2}[\![\psi]\!] = k_2$ with $\sigma_1(x) = 1$ and $\sigma_2(x) = 2$. In step (b.1.1) we have $\sigma_1' = \sigma_2'$ (both are the empty assignment), and so Verifier sends the challenge $[\![\psi]\!]$ to Prover twice, who answers with claims $[\![\psi]\!] = p_1$ and $[\![\psi]\!] = p_2$. In step (b.1.2) Verifier checks that $p_1(1) = k_1$ and $p_2(2) = k_2$ hold, picks a random number r, and updates $\sigma_1(x) := \sigma_2(x) := r$ and $k_1 := p_1(r), k_2 := p_2(r)$. Now the condition of the while loop fails, so Verifier moves to (b.2) and checks $k_1 = k_2$.

Description of Step (c). Let $\Pi_\sigma[\![\psi]\!] = k$ be the claim computed by Verifier in step (b). Verifier removes this claim from $\mathcal{C}$ and replaces it by claims about the children of ψ, depending on the structure of ψ:

(c.1) If $\psi = \psi_1 \circledast \psi_2$, for a $\circledast \in \{\vee, \wedge\}$, then Verifier sends Prover challenges $\Pi_\sigma[\![\psi_i]\!]$ for $i \in \{1, 2\}$, and Prover sends claims $\Pi_\sigma[\![\psi_i]\!] = k_i$ back. Verifier checks the consistency condition $k = \pi_{[x:=k_1]}\pi_{[y:=k_2]}[\![x \circledast y]\!]$, rejecting if it does not hold. If the condition holds, the claim $\Pi_\sigma[\![\psi_i]\!] = k_i$ is added to $\mathcal{C}$, to be checked in the round for ψ_i.
(c.2) If $\psi = \neg\psi'$, then Verifier adds the claim $\Pi_\sigma[\![\psi']\!] = 1 - k$ to ψ'.
(c.3) If $\psi = \pi_{[x:=b]}\, \psi'$, Verifier sets $\sigma' := \sigma \cup \{x \mapsto b\}$ and adds the claim $\Pi_{\sigma'}[\![\psi']\!] = k$ to $\mathcal{C}$.
(c.4) If $\psi = \delta_x \psi'$, then Verifier sends Prover the challenge $\Pi_{\sigma'}[\![\psi']\!]$, where σ' denotes the partial assignment which is undefined on x and otherwise matches σ. Prover returns the claim $p := \Pi_{\sigma'}[\![\psi']\!]$. Observe that p is a univariate polynomial over x. Verifier checks the consistency condition $\pi_{[x:=\sigma(x)]}\delta_x\, p = k$, rejecting if it does not hold. If it holds, Verifier picks an $r \in \mathbb{F}$ uniformly at random, conducts the updates $\sigma(x) := r$ and $k := \pi_{[x:=r]}\, p$, and adds $\Pi_\sigma[\![\psi']\!] = k$ to the set of claims about ψ'.

This concludes the description of the interactive protocol. We now show CPCertify is complete and sound.

Proposition 2 (CPCertify is complete and sound). *Let φ be a CP with n free variables. Let $\Pi_\sigma[\![\mathrm{conv}(\varphi)]\!] = K$ be the claim initially sent by Prover to Verifier. If the claim is true, then Prover has a strategy to make Verifier accept. If not, for every Prover, Verifier accepts with probability at most $4n|\varphi|/|\mathbb{F}|$.*

If the original claim is correct, Prover can answer every challenge truthfully and all claims pass all of Verifier's checks. So Verifier accepts. If the claim is not correct, we proceed round by round. We bound the probability that the Verifier is tricked in a single step to at most $2/|\mathbb{F}|$ using the Schwartz-Zippel Lemma. We then bound the number of such steps to $2n|\varphi|$ and use a union bound.

5 A BDD-Based Prover

We assume familiarity with *reduced ordered binary decision diagrams* (BDDs) [9]. We use BDDs over $X = \{x_1, \ldots, x_n\}$. We fix the variable order $x_1 < x_2 < \ldots < x_n$, i.e. the root node would decide based on the value of x_n.

Definition 4. *BDDs are defined inductively as follows:*

- $\langle\mathsf{true}\rangle$ *and* $\langle\mathsf{false}\rangle$ *are BDDs of level* 0*;*
- *if $u \neq v$ are BDDs of level ℓ_u, ℓ_v and $i > \ell_u, \ell_v$, then $\langle x_i, u, v\rangle$ is a BDD of level i;*
- *we identify $\langle x_i, u, u\rangle$ and u, for a BDD u of level ℓ_i and $i > \ell_u$.*

The level of a BDD w is denoted $\ell(w)$. The set of descendants *of w is the smallest set S with $w \in S$ and $u, v \in S$ for all $\langle x, u, v\rangle \in S$. The* size *$|w|$ of w is the number of its descendants.*

The arithmetisation *of a BDD w is the polynomial $[\![w]\!]$ defined as follows: $[\![\langle\mathsf{true}\rangle]\!] := \mathbf{1}$, $[\![\langle\mathsf{false}\rangle]\!] := \mathbf{0}$ and $[\![\langle x, u, v\rangle]\!] := [1 - x] \cdot [\![u]\!] + [x] \cdot [\![v]\!]$.*

Figure 3 shows a BDD for the boolean function $\varphi(x, y, z) = (x \wedge y \wedge \neg z) \vee (\neg x \wedge y \wedge z) \vee (x \wedge \neg y \wedge z)$ and the arithmetisation of each node.

BDDSolver: A BDD-based Algorithm for #CP. An instance φ of #CP can be solved using BDDs. Starting at the leaves of φ, we iteratively compute a BDD for each node ψ of the circuit encoding the boolean predicate P_ψ. At the end of this procedure we obtain a BDD for P_φ. The number of satisfying assignments of ψ is the number of accepting paths of the BDD, which can be computed in linear time in the size of the BDD.

For a node $\psi = \psi_1 \circledast \psi_2$, given BDDs representing the predicates P_{φ_1} and P_{φ_2}, we compute a BDD for the predicate $P_\varphi := P_{\varphi_1} \circledast P_{\varphi_2}$, using the $\mathrm{Apply}_\circledast$ operator on BDDs. We name this algorithm for solving #CP "BDDSolver."

From BDDSolver **to** CPCertify**.** Our goal is to modify BDDSolver to play the role of an honest Prover in CPCertify with minimal overhead. In CPCertify, Prover repeatedly performs the same task: evaluate polynomials of the form $\Pi_\sigma[\![\psi]\!]$, where ψ is a descendant of the CPD $\mathrm{conv}(\varphi)$, and σ assigns values to all free variables of ψ except possibly one. Therefore, the polynomials have at most one free variable and, as we have seen, degree at most 2.

Before defining the concepts precisely, we give a brief overview of this section.

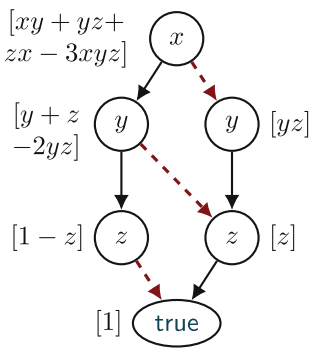

Fig. 3. A BDD and its arithmetisation. For $\langle x, u, v\rangle$, we denote the link from x to v with a solid edge and x to u with a dotted edge. We omit links to $\langle\mathsf{false}\rangle$.

- First (Proposition 3), we show that BDDs correspond to binary multilinear polynomials. In particular, BDDs allow for efficient evaluation of the polynomial. As argued in Lemma 1(a), for every descendant ψ of φ, the CPD $\mathrm{conv}(\psi)$ (which is a descendant of $\mathrm{conv}(\varphi)$) evaluates to a multilinear polynomial. In particular, Prover can use standard BDD algorithms to calculate the corresponding polynomials $\Pi_\sigma[\![\psi]\!]$ for all descendants ψ of $\mathrm{conv}(\varphi)$ that are neither binary operators nor degree reductions.
- Second (the rest of the section), we prove a surprising connection: the intermediate results obtained while executing the BDD algorithms (with slight adaptations) correspond precisely to the remaining descendants of $\mathrm{conv}(\varphi)$.

The following proposition proves that BDDs represent exactly the binary multilinear polynomials.

Proposition 3. *(a) For a BDD w, $[\![w]\!]$ is a binary multilinear polynomial. (b) For a binary multilinear polynomial p there is a unique BDD w s.t. $p = [\![w]\!]$.*

5.1 Extended BDDs

During the execution of CPCertify for a given CPD $\mathrm{conv}(\varphi)$, Prover sends to Verifier claims of the form $\Pi_\sigma[\![\psi]\!]$, where ψ is a descendant of $\mathrm{conv}(\varphi)$, and $\sigma\colon X \to \mathbb{F}$ is a partial assignment. While all polynomials computed by CPCertify are binary, not all are multilinear: some polynomials have degree 2. For these polynomials, we introduce *extended BDDs* (eBDDs) and give eBDD-based algorithms for the following two tasks:

1. Compute an eBDD representing $[\![\psi]\!]$ for every node ψ of $\mathrm{conv}(\varphi)$.
2. Given an eBDD for $[\![\psi]\!]$ and a partial assignment σ, compute $\Pi_\sigma[\![\psi]\!]$.

Computing eBDDs for CPDs: Informal Introduction. Consider a CP φ and its associated CPD $\mathrm{conv}(\varphi)$. Each node of φ induces a chain of nodes in $\mathrm{conv}(\varphi)$, consisting of degree-reduction nodes $\delta_{x_1}, \ldots, \delta_{x_n}$, followed by the node itself (see Fig. 4). Given BDDs u and v for the children of the node in the CP, we can compute a BDD for the node itself using a well-known BDD algorithm $\mathrm{Apply}_{\circledast}(u, v)$ parametric in the boolean operation $\circledast$ labelling the node [9]. Our goal is to transform $\mathrm{Apply}_{\circledast}$ into an algorithm that computes eBDDs *for all nodes in the chain*, i.e. eBDDs for all the polynomials $p_0, p_1, \ldots, p_n$ of Fig. 4.

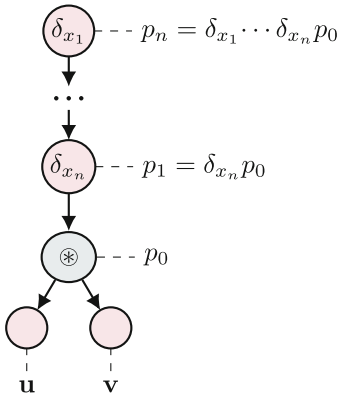

Fig. 4. A node of a CP ($\circledast$) gets a chain of degree reduction nodes in the associated CPD.

Roughly speaking, $\mathrm{Apply}_{\circledast}(u, v)$ recursively computes BDDs $w_0 = \mathrm{Apply}_{\circledast}(u_0, v_0)$ and $w_1 = \mathrm{Apply}_{\circledast}(u_1, v_1)$, where u_b and v_b are the b-children of u and v, and then returns the BDD with w_0 and w_1 as 0- and 1-child, respectively.[3]

Most importantly, we modify $\mathrm{Apply}_{\circledast}$ to run in breadth-first order. Figure 5 shows a graphical representation of a run of $\mathrm{Apply}_{\vee}(u, v)$, where u and v are the two BDD nodes labelled by x. Square nodes represent pending calls to $\mathrm{Apply}_{\circledast}$. Initially there is only one square call $\mathrm{Apply}_{\vee}(u, v)$ (Fig. 5, top left). $\mathrm{Apply}_{\vee}$ calls itself recursively for u_0, v_0 and u_1, v_1 (Fig. 5, top right). Each of the two calls splits again into two; however, the first three are identical (Fig. 5, bottom left), and so reduce to two. These two calls can now be resolved directly; they return nodes true and false, respectively. At this point, the children of $\mathrm{Apply}_{\circledast}(u, v)$ become $\langle y, \mathsf{true}, \mathsf{true}\rangle = \mathsf{true}$, and $\langle y, \mathsf{true}, \mathsf{false}\rangle$, which exists already as well (Fig. 5, bottom right).

We look at the diagrams of Fig. 5 not as a visualisation aid, but as graphs with two kinds of nodes: standard BDD nodes, represented as circles, and *product* nodes, represented as squares. We call them *extended BDDs*. Each node of an extended BDD is assigned a polynomial in the expected way: the polynomial $[\![u]\!]$ of a standard BDD node u with variable x is $x \cdot [\![u_1]\!] + (1 - x) \cdot [\![u_0]\!]$, the polynomial $[\![v]\!]$ of a square $\wedge$-node v is $[\![v_0]\!] \cdot [\![v_1]\!]$, etc. In this way we assign to each eBDD a polynomial. In particular, we obtain the intermediate polynomials p_0, p_1, p_2, p_3 of the figure, one for each level in the recursion. In the rest of the section we show that these are *precisely* the polynomials $p_0, p_1, \ldots, p_n$ of Fig. 4.

Thus, in order to compute eBDDs for all nodes of a CPD $\mathrm{conv}(\varphi)$, it suffices to compute BDDs for all nodes of the CP φ. Since we need to do this anyway to solve #CP, the polynomial certification does not incur any overhead.

[3] In fact, this is only true when u and v are nodes at the same level and $\mathrm{Apply}_{\circledast}(u_0, v_0) \neq \mathrm{Apply}_{\circledast}(u_1, v_1)$, but at this point we only want to convey some intuition.

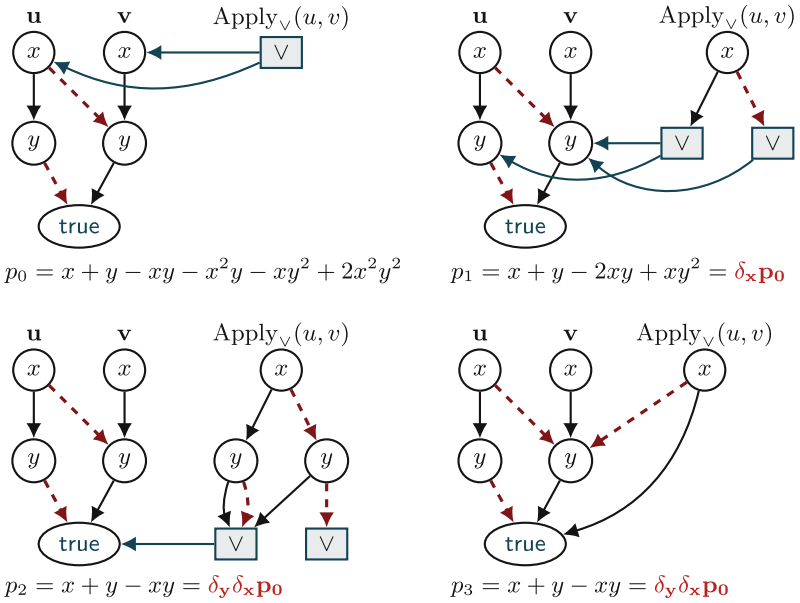

Fig. 5. Run of $\text{Apply}_\vee(u, v)$, but with recursive calls evaluated in breadth-first order. All missing edges go to node false.

Extended BDDs. As for BDDs, we define eBDDs over $X = \{x_1, \ldots, x_n\}$ with the variable order $x_1 < x_2 < \ldots < x_n$.

Definition 5. *Let $\circledast$ be a binary boolean operator. The set of eBDDs (for $\circledast$) is inductively defined as follows:*

- *every BDD is also an eBDD of the same level;*
- *if u, v are BDDs (not eBDDs!), then $\langle u \circledast v \rangle$ is an eBDD of level l where $l := \max\{\ell(u), \ell(v)\}$; we call eBDDs of this form* product nodes*;*
- *if $u \neq v$ are eBDDs and $i > \ell(u), \ell(v)$, then $\langle x_i, u, v \rangle$ is an eBDD of level i;*
- *we identify $\langle x_i, u, u \rangle$ and u for an eBDD u and $i > \ell(u)$.*

The set of descendants *of an eBDD w is the smallest set S with $w \in S$ and $u, v \in S$ for all $\langle u \circledast v \rangle, \langle x, u, v \rangle \in S$ The* size *of w is its number of descendants. For $u, v \in \{\langle \mathsf{true} \rangle, \langle \mathsf{false} \rangle\}$ we identify $\langle u \circledast v \rangle$ with $\langle \mathsf{true} \rangle$ or $\langle \mathsf{false} \rangle$ according to the result of $\circledast$, e.g. $\langle \langle \mathsf{true} \rangle \vee \langle \mathsf{false} \rangle \rangle = \langle \mathsf{true} \rangle$, as $\mathsf{true} \vee \mathsf{false} = \mathsf{true}$. The* arithmetisation *of an eBDD for a boolean operator $\circledast \in \{\wedge, \vee\}$ is defined as for BDDs, with the extensions $[\![\langle u \wedge v \rangle]\!] = [\![u]\!] \cdot [\![v]\!]$ and $[\![\langle u \vee v \rangle]\!] = [\![u]\!] + [\![v]\!] - [\![u]\!] \cdot [\![v]\!]$.*

Example 2. The diagrams in Fig. 5 are eBDDs for $\circledast := \vee$. Nodes of the form $\langle x, u, v \rangle$ and $\langle u \vee v \rangle$ are represented as circles and squares, respectively. Consider the top-left diagram. Abbreviating $x \oplus y := (x \wedge \neg y) \vee (\neg x \wedge y)$ we get $[\![\text{Apply}_\vee(u, v)]\!] = [\![(x \oplus y) \wedge (x \wedge y)]\!] = [\![x \oplus y]\!] \cdot [\![x \wedge y]\!] = (x(1-y) + (1-x) \cdot y - xy(1-x)(1-y)) \cdot xy$, which is the polynomial p_0 shown in the figure.

Table 1. On the left: Algorithm computing eBDDs for the sequence $[\![w]\!]$, $\delta_{x_n}[\![w]\!]$, $\delta_{x_{n-1}}\delta_{x_n}[\![w]\!]$, ..., $\delta_{x_1}\cdots\delta_{x_n}[\![w]\!]$ of polynomials. On the right: Recursive algorithm to evaluate the polynomial represented by an eBDD at a given partial assignment. $P(w)$ is a mapping used to memoize the polynomials returned by recursive calls.

ComputeEBDD(w)

Input: eBDD w
Output: sequence $w_0, ..., w_n$ of eBDDs

$w_0 := w$; output w_0
for $i = 0, \cdots, \ell(w) - 1$ **do**
  $w_{i+1} := w_i$
  for every node $\langle u \circledast v\rangle$ of w_i
    at level $n - i$ **do**
    for $b \in \{0, 1\}$ **do**
      $u_b := \pi_{[x_{n-i}:=b]}\, u$
      $v_b := \pi_{[x_{n-i}:=b]}\, v$
      $t_b := \langle u_b \circledast v_b\rangle$
    $w_{i+1} := w_{i+1}\,[\,\langle u \circledast v\rangle / \langle x_{n-i}, t_0, t_1\rangle\,]$
  output w_{i+1}

EvaluateEBDD(w, σ) =: $E_\sigma(w)$

Input: eBDD w; assignment $\sigma\colon X \to \mathbb{F}$
Output: $\Pi_\sigma[\![w]\!]$

if $P(w)$ is defined return $P(w)$
if $w \in \{\langle\mathsf{true}\rangle, \langle\mathsf{false}\rangle\}$ return $[\![w]\!]$
if $w = \langle u \wedge v\rangle$
  $P(w) := E_\sigma(u) \cdot E_\sigma(v)$
if $w = \langle u \vee v\rangle$
  $P(w) := E_\sigma(u) + E_\sigma(v) - E_\sigma(u)E_\sigma(v)$
if $w = \langle x, u, v\rangle$ and $\sigma(x)$ undefined
  $P(w) := [1 - x] \cdot E_\sigma(u) + [x] \cdot E_\sigma(v)$
if $w = \langle x, u, v\rangle$ and $\sigma(x) = s \in \mathbb{F}$
  $P(w) := [1 - s] \cdot E_\sigma(u) + [s] \cdot E_\sigma(v)$
return $P(w)$

Computing eBDDs for CPDs. Given a node of a CP corresponding to a binary operator $\circledast$, Prover has to compute polynomials p_0, $\delta_{x_1}p_0$, ..., $\delta_{x_n}...\delta_{x_1}p_0$ corresponding to the nodes of the CPD shown on the right. We show that Prover can compute these polynomials by representing them as eBDDs. Table 1 describes an algorithm that gets as input an eBDD w of level n, and outputs a sequence $w_0, w_1, ..., w_{n+1}$ of eBDDs such that $w_0 = w$; $[\![w_{i+1}]\!] = \delta_{x_{n-i}}[\![w_i]\!]$ for every $0 \le i \le \ell(w) - 1$; and w_{n+1} is a BDD. Interpreted as sequence of eBDDs, Fig. 5 shows a run of this algorithm.

Notation. Given an eBDD w and eBDDs u, v such that $\ell(u) \ge \ell(v)$, we let $w[u/v]$ denote the result of replacing u by v in w. For an eBDD $w = \langle x_i, w_0, w_1\rangle$ and $b \in \{0, 1\}$ we define $\pi_{[x_i:=b]}w := w_b$, and for $j > i$ we set $\pi_{[x_j:=b]}w := w$. (Note that $[\![\pi_{[x_j:=b]}w]\!] = \pi_{[x_j:=b]}[\![w]\!]$ holds for any j where it is defined.)

Proposition 4. *Let ψ_1, ψ_2 denote CPs and u_1, u_2 BDDs with $[\![u_i]\!] = [\![\psi_i]\!]$, $i \in \{1, 2\}$. Let $w := \langle u_1 \circledast u_2\rangle$ denote an eBDD. Then* ComputeEBDD(w) *satisfies $[\![w_0]\!] = [\![\psi_1 \circledast \psi_2]\!]$ and $[\![w_{i+1}]\!] = \delta_{x_{n-i}}[\![w_i]\!]$ for every $0 \le i \le n-1$; moreover, w_n is a BDD with $w_n = \mathrm{Apply}_{\circledast}(u_1, u_2)$. Finally, the algorithm runs in time $\mathcal{O}(T)$, where $T \in \mathcal{O}(|u_1| \cdot |u_2|)$ is the time taken by $\mathrm{Apply}_{\circledast}(u_1, u_2)$.*

Evaluating Polynomials Represented as eBDDs. Recall that Prover must evaluate expressions of the form $\Pi_\sigma[\![\psi]\!]$ for some CPD ψ, where σ assigns values to all variables of ψ except for possibly one. We give an algorithm to evaluate arbitrary expressions $\Pi_\sigma[\![w]\!]$, where w is an eBDD, and show that if there is at most one free variable then the algorithm takes linear time in the size of ψ. The algorithm is shown on the right of Table 1. It has the standard structure of BDD procedures: It recurs on the structure of the eBDD, memoizing the result of recursive calls so that the algorithm is called at most once with a given input.

Proposition 5. *Let w denote an eBDD, $\sigma : X \to \mathbb{F}$ a partial assignment, and k the number of variables assigned by σ. Then* EvaluateEBDD *evaluates the polynomial $\Pi_\sigma[\![w]\!]$ in time $\mathcal{O}\big(\mathrm{poly}(2^{n-k}) \cdot |w|\big)$.*

5.2 Efficient Certification

In the CPCertify algorithm, Prover must (a) compute polynomials for all nodes of the CPD, and (b) evaluate them on assignments chosen by Verifier. In the last section we have seen that ComputeEBDD (for binary operations of the CP), combined with standard BDD algorithms (for all other operations), yields eBDDs representing all these polynomials—at no additional overhead, compared to a BDD-based implementation. This covers part (a). Regarding (b), recall that all polynomials computed in (a) have at most one variable. Therefore, using EvaluateEBDD we can evaluate a polynomial in linear time in the size of the eBDD representing it.

The Verifier CPCertify is implemented in a straightforward manner. As the algorithm runs in polynomial size w.r.t. the CP (and not the computed BDDs, which may be exponentially larger), incurring overhead is less of a concern.

Theorem 1 (Main Result). *If* BDDSolver *solves an instance φ of* #CP *with n variables in time T, with $T > n|\varphi|$, then*

(a) Prover computes eBDDs for all nodes of $\mathrm{conv}(\varphi)$ *in time $\mathcal{O}(T)$,*
(b) Prover responds to Verifier's challenges in time $\mathcal{O}(nT)$, and
(c) Verifier executes CPCertify *in time $\mathcal{O}(n^2|\varphi|)$, with failure probability at most $4n|\varphi|/|\mathbb{F}|$.*

As presented above, EvaluateEBDD incurs a factor-of-n overhead, as every node of the CPD must be evaluated. In our implementation, we use a caching strategy to reduce the complexity of Theorem 1(b) to $\mathcal{O}(T)$.

Note that the bounds above assume a uniform cost model. In particular, operations on BDD nodes and finite field arithmetic are assumed to be $\mathcal{O}(1)$. This is a reasonable assumption, as for a constant failure probability $\log|\mathbb{F}| \approx \log n$. Hence the finite field remains small. (It is possible to verify the number of assignments even if it exceeds $|\mathbb{F}|$, see below.)

5.3 Implementation Concerns

We list a number of points that are not described in detail in this paper, but need to be considered for an efficient implementation.

Finite Field Arithmetic. It is not necessary to use large finite fields. In particular, one can avoid the overhead of arbitrarily sized integers. For our implementation we fix the finite field $\mathbb{F} := \mathbb{Z}_p$, with $p = 2^{61} - 1$ (the largest Mersenne prime to fit in 64 bits).

Incremental eBDD Representation. Algorithm ComputeEBDD computes a sequence of eBDDs. These must not be stored explicitly, otherwise one incurs

a space-overhead. Instead, we only store the last eBDD as well as the differences between each subsequent element of the sequence. To evaluate the eBDDs, we then revert to a previous state by applying the differences appropriately.

Evaluation Order. It simplifies the implementation if Prover only needs to evaluate nodes of the CPD in some (fixed) topological order. CPCertify can easily be adapted to guarantee this, by picking the next node appropriately in each iteration, and by evaluating only one child of a binary operator $\psi_1 \circledast \psi_2$. The value of the other child can then be derived by solving a linear equation.

Efficient Evaluation. As stated in Theorem 1, using EvaluateEBDD Prover needs $\Omega(nT)$ time to respond to Verifier's challenges. In our implementation we instead use a caching strategy that reduces this time to $\mathcal{O}(T)$. Essentially, we exploit the special structure of $\mathrm{conv}(\varphi)$: Verifier sends a sequence of challenges $\Pi_{\sigma_0}\delta_{x_1}...\delta_{x_n}w, \Pi_{\sigma_1}\delta_{x_2}...\delta_{x_n}w, ..., \Pi_{\sigma_n}w$, where assignments σ_i and σ_{i+1} differ only in variables x_i and x_{i+1}. The corresponding eBDDs likewise change only at levels i and $i+1$. We cache the linear coefficients of eBDD nodes that contribute to the arithmetisation of the root top-down, and the arithmetised values of nodes bottom up. As a result, only levels $i, i+1$ need to be updated.

Large Numbers of Assignments. If the number of satisfying assignments of a CP exceeds $|\mathbb{F}|$, Verifier would not be able to verify the count accurately. Instead of choosing $|\mathbb{F}| \geq 2^n$, which incurs a significant overhead, Verifier can query the precise number of assignments, and then choose $|\mathbb{F}|$ randomly. This introduces another possibility of failure, but (roughly speaking) it suffices to double $\log|\mathbb{F}|$ for the additional failure probability to match the existing one. Our implementation does not currently support this technique.

6 Evaluation

We have implemented an eBDD library, blic (BDD Library with Interactive Certification)[4], that is a stand-in replacement for BDDs but additionally performs the role of Prover in the CPCertify protocol. We have also implemented a client that executes the protocol as Verifier. The eBDD library is about 900 lines of C++ code and the CPCertify protocol is about 400 lines. We have built a prototype certifying QBF solver in blic, totalling about 2600 lines of code. We aim to answer the following questions in our evaluation:

RQ1. Is a QBF solver with CPCertify-based certification competitive? If so, how high is the overhead of implementing CPCertify on top of the BDD operations?

RQ2. What is the amount of communication for Prover and Verifier in executing the CPCertify protocol, what is the time requirement for Verifier, and how do these numbers compare to proof sizes and proof checking times for certificates based on resolution and other proof systems?

[4] https://gitlab.lrz.de/i7/blic.

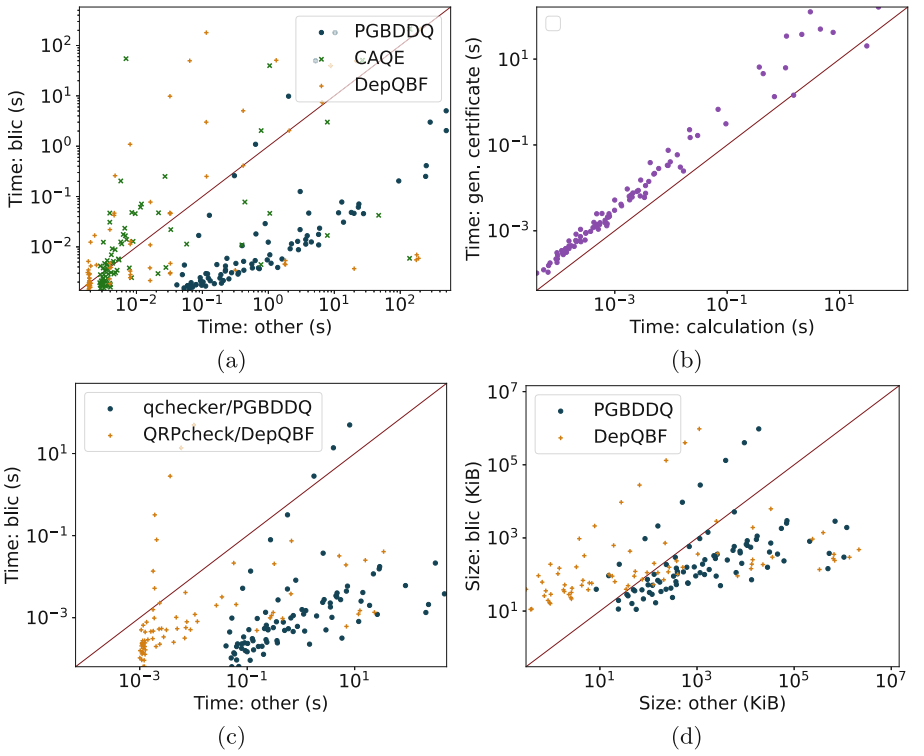

Fig. 6. (a) Time taken on instances (dashed lines are $y = 100x$ and $y = 0.01x$), (b) Cost of generating a certificate over computing the solution, (c) Time to verify the certificate, (d) Size of certificates

RQ1: Performance of blic**.** We compare blic with CAQE, DepQBF, and PGBDDQ, three state-of-the-art QBF solvers. CAQE [10,29] does not provide any certificates in its most recent version. DepQBF [12,19] is a certifying QBF solver. PGBDDQ [7,25] is an independent implementation of a BDD-based QBF solver. Both DepQBF and PGBDDQ provide specialised checkers for their certificates, though PGBDDQ can also proofs in standard QRAT format. Note that PGBDDQ is written in Python and generates proofs in an ASCII-based format, incurring overhead compared to the other tools.

We take 172 QBF instances (all unsatisfiable) from the *Crafted Instances* track of the QBF Evaluation 2022.[5] The *Prenex CNF* track of the QBF competition is not evaluated here. It features instances with a large number of variables. BDD-based solvers perform poorly under these circumstances without additional optimisations. Our overall goal is not to propose a new approach for

[5] CAQE and DepQBF were the winner and runner-up in this category. The configuration we used differs from the competition, as described in the full version of the paper [11].

Table 2. Comparison of certificate generation, bytes exchanged between prover and verifier, and time taken to verify the certificate on a set of QBF benchmarks from [7]. "Solve time" is time taken to solve the instance and to generate a certificate (seconds), "Certificate" is the size of proof encoding for PGBDDQ, and bytes exchanged by CPCertify for blic, and "Verifier time" is time to verify the certificate (Verifier's run time for blic and time taken by qchecker).

Instance		Solve time (s)		Certificate (MiB)		Verifier time (s)	
n	result	blic	PGBDDQ	blic	PGBDDQ	blic	qchecker
10	sat	0.03	3.67	1.20	8.48	0.01	3.80
10	unsat	0.03	3.66	1.20	8.45	0.01	3.83
15	sat	0.13	18.07	4.12	44.25	0.02	18.45
15	unsat	0.13	18.14	4.11	44.20	0.02	18.55
20	sat	0.54	82.92	11.59	198.54	0.07	80.28
20	unsat	0.53	83.02	11.64	198.76	0.06	79.05
25	sat	1.56	261.16	23.94	566.95	0.14	238.99
25	unsat	1.55	261.25	23.86	565.36	0.15	237.94
40	sat	25.22	4863.71	132.43	7464.96	0.95	5141.08
40	unsat	25.25	4827.06	132.67	7467.84	0.99	5463.54

solving QBF, but rather to certify a BDD-based approach, so we wanted to focus on cases where the existing BDD-based approaches are practical.

We ran each benchmark with a 10 min timeout; all tools other than CAQE were run with certificate production. All times were obtained on a machine with an Intel Xeon E7-8857 CPU and 1.58 TiB RAM[6] running Linux. See the full version of the paper [11] for a detailed description. blic solved 96 out of 172 benchmarks, CAQE solved 98, DepQBF solved 87, and PGBDDQ solved 91. Figure 6(a) shows the run times of blic compared to the other tools. The plot indicates that blic is competitive on these instances, with a few cases, mostly from the Lonsing family of benchmarks, where blic is slower than DepQBF by an order of magnitude. Figure 6(b) shows the overhead of certification: for each benchmark (that finishes within a 10min timeout), we plot the ratio of the time to compute the answer to the time it takes to run Prover in CPCertify. The dotted regression line shows CPCertify has a 2.8× overhead over computing BDDs. For this set of examples, the error probability never exceeds $10^{-8.9}$ ($10^{-11.6}$ when Lonsing examples are excluded); running the verifier k times reduces it to $10^{-8.9k}$.

RQ2: Communication Cost of Certification and Verifier Time. We explore **RQ2** by comparing the number of bytes exchanged between Prover and Verifier and the time needed for Verifier to execute CPCertify with the number of bytes in an QBF proof and the time required to verify the proof produced by DepQBF and PGBDDQ, for which we use QRPcheck [24,26] and qchecker [7,25], respectively. Note that the latter is written in Python.

[6] blic uses at most 60 GiB on the shown benchmarks, 5 GiB when excluding timeouts.

We show that the overhead of certification is low. Figure 6(c) shows the run time of Verifier—this is generally negligible for blic, except for the Lonsing and KBKF families, which have a large number of variables, but very small BDDs. Figure 6(d) shows the total number of bytes exchanged between Prover and Verifier in blic against the size of the proofs generated by PGBDDQ and DepQBF. For large instances, the number of bytes exchanged in blic is significantly smaller than the size of the proofs. The exception are again the Lonsing and KBKNF families of instances. For both plots, the dotted line results from a log-linear regression.

In addition to the Crafted Instances, we compare against PGBDDQ on a challenging family of benchmarks used in the PGBDDQ paper (matching the parameters of [7, Table 3]); these are QBF encodings of a linear domino placing game.[7] Our results are summarised in Table 2. The upper bound on Verifier error is $10^{-9.22}$. We show that blic outperforms PGBDDQ both in overall cost of computing the answer and the certificates as well as in the number of bytes communicated and the time used by Verifier.

Our results indicate that giving up absolute certainty through interactive protocols can lead to an order of magnitude smaller communication cost and several orders of magnitude smaller checking costs for the verifier.

7 Conclusion

We have presented a solver that combines BDDs with an interactive protocol. blic can be seen as a self-certifying BDD library able to certify the correctness of arbitrary sequences of BDD operations. In order to trust the result, a user must only trust the verifier (a straightforward program that poses challenges to the prover). We have shown that blic (including certification time) is competitive with other solvers, and Verifier's time and error probabilities are negligible.

Our results show that IP = PSPACE can become an important result not only in theory but also in the practice of automatic verification. From this perspective, our paper is a first step towards practical certification based on interactive protocols. While we have focused on BDDs, we can ask the more general question: which practical automated reasoning algorithms can be made efficiently certifying? For example, whether there is an interactive protocol and an efficient certifying version of modern SAT solving algorithms is an interesting open challenge.

References

1. Arora, S., Barak, B.: Computational Complexity: A Modern Approach. Cambridge University Press, Cambridge (2006). https://theory.cs.princeton.edu/complexity/book.pdf

[7] DepQBF only solved 1 of 10 instances within 120 min, and is thus not compared.

2. Babai, L.: Trading group theory for randomness. In: Sedgewick, R. (ed.) Proceedings of the 17th Annual ACM Symposium on Theory of Computing, 6–8 May 1985, Providence, Rhode Island, USA, pp. 421–429. ACM (1985). https://doi.org/10.1145/22145.22192
3. Balabanov, V., Widl, M., Jiang, J.-H.R.: QBF resolution systems and their proof complexities. In: Sinz, C., Egly, U. (eds.) SAT 2014. LNCS, vol. 8561, pp. 154–169. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-09284-3_12
4. Barbosa, H., et al.: Flexible proof production in an industrial-strength SMT solver. In: Blanchette, J., Kovács, L., Pattinson, D. (eds.) IJCAR 2022. LNCS, vol. 13385, pp. 15–35. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10769-6_3
5. Ben-Or, M., Goldreich, O., Goldwasser, S., Håstad, J., Kilian, J., Micali, S., Rogaway, P.: Everything provable is provable in zero-knowledge. In: Goldwasser, S. (ed.) CRYPTO 1988. LNCS, vol. 403, pp. 37–56. Springer, New York (1990). https://doi.org/10.1007/0-387-34799-2_4
6. Bryant, R.E., Biere, A., Heule, M.J.H.: Clausal proofs for pseudo-boolean reasoning. In: TACAS 2022. LNCS, vol. 13243, pp. 443–461. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-99524-9_25
7. Bryant, R.E., Heule, M.J.H.: Dual proof generation for quantified boolean formulas with a BDD-based solver. In: Platzer, A., Sutcliffe, G. (eds.) CADE 2021. LNCS (LNAI), vol. 12699, pp. 433–449. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-79876-5_25
8. Bryant, R.E., Heule, M.J.H.: Generating extended resolution proofs with a BDD-based SAT solver. In: TACAS 2021. LNCS, vol. 12651, pp. 76–93. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72016-2_5
9. Bryant, R.: Graph-based algorithms for Boolean function manipulation. IEEE Trans. Comput. **C-35**(8), 677–691 (1986)
10. CAQE (2023). https://github.com/ltentrup/caqe. Accessed 03 Feb 2023
11. Couillard, E., Czerner, P., Esparza, J., Majumdar, R.: Making IP=PSPACE practical: efficient interactive protocols for BDD algorithms. CoRR abs/2305.11813 (2023). https://doi.org/10.48550/arXiv.2305.11813
12. DepQBF (2017). https://github.com/lonsing/depqbf. Accessed 03 Feb 2023
13. Goldwasser, S., Micali, S., Rackoff, C.: The knowledge complexity of interactive proof-systems (extended abstract). In: Sedgewick, R. (ed.) Proceedings of the 17th Annual ACM Symposium on Theory of Computing, 6–8 May 1985, Providence, Rhode Island, USA, pp. 291–304. ACM (1985). https://doi.org/10.1145/22145.22178
14. Henzinger, T.A., Necula, G.C., Jhala, R., Sutre, G., Majumdar, R., Weimer, W.: Temporal-safety proofs for systems code. In: Brinksma, E., Larsen, K.G. (eds.) CAV 2002. LNCS, vol. 2404, pp. 526–538. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45657-0_45
15. Heule, M.: Everything's bigger in Texas: "the largest math proof ever". In: Benzmüller, C., Lisetti, C.L., Theobald, M. (eds.) GCAI 2017, 3rd Global Conference on Artificial Intelligence, Miami, FL, USA, 18–22 October 2017. EPiC Series in Computing, vol. 50, pp. 1–5. EasyChair (2017). https://doi.org/10.29007/gdw8
16. Heule, M.J.H.: Proofs of unsatisfiability. In: Biere, A., Heule, M., van Maaren, H., Walsh, T. (eds.) Handbook of Satisfiability - Second Edition, Frontiers in Artificial Intelligence and Applications, vol. 336, pp. 635–668. IOS Press (2021). https://doi.org/10.3233/FAIA200998
17. Jussila, T., Sinz, C., Biere, A.: Extended resolution proofs for symbolic SAT solving with quantification. In: Biere, A., Gomes, C.P. (eds.) SAT 2006. LNCS, vol. 4121, pp. 54–60. Springer, Heidelberg (2006). https://doi.org/10.1007/11814948_8

18. Katz, G., Barrett, C.W., Tinelli, C., Reynolds, A., Hadarean, L.: Lazy proofs for DPLL(T)-based SMT solvers. In: Piskac, R., Talupur, M. (eds.) 2016 Formal Methods in Computer-Aided Design, FMCAD 2016, Mountain View, CA, USA, 3–6 October 2016, pp. 93–100. IEEE (2016). https://doi.org/10.1109/FMCAD.2016.7886666
19. Lonsing, F., Egly, U.: DepQBF 6.0: a search-based QBF solver beyond traditional QCDCL. In: de Moura, L. (ed.) CADE 2017. LNCS (LNAI), vol. 10395, pp. 371–384. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-63046-5_23
20. Lund, C., Fortnow, L., Karloff, H.J., Nisan, N.: Algebraic methods for interactive proof systems. J. ACM **39**(4), 859–868 (1992). https://doi.org/10.1145/146585.146605
21. Luo, N., Antonopoulos, T., Harris, W.R., Piskac, R., Tromer, E., Wang, X.: Proving UNSAT in zero knowledge. In: Yin, H., Stavrou, A., Cremers, C., Shi, E. (eds.) Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, 7–11 November 2022, pp. 2203–2217. ACM (2022). https://doi.org/10.1145/3548606.3559373
22. Namjoshi, K.S.: Certifying model checkers. In: Berry, G., Comon, H., Finkel, A. (eds.) CAV 2001. LNCS, vol. 2102, pp. 2–13. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44585-4_2
23. Necula, G.: Proof-carrying code. In: Principles of Programming Languages, pp. 106–119. ACM Press (1997)
24. Niemetz, A., Preiner, M., Lonsing, F., Seidl, M., Biere, A.: Resolution-based certificate extraction for QBF. In: Cimatti, A., Sebastiani, R. (eds.) SAT 2012. LNCS, vol. 7317, pp. 430–435. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31612-8_33
25. PGBDDQ (2023). https://github.com/rebryant/pgbdd. Accessed 03 Feb 2023
26. QRPcheck (2023). http://fmv.jku.at/qrpcheck/. Accessed 03 Feb 2023
27. Shamir, A.: IP = PSPACE. J. ACM **39**(4), 869–877 (1992). https://doi.org/10.1145/146585.146609
28. Sinz, C., Biere, A.: Extended resolution proofs for conjoining BDDs. In: Grigoriev, D., Harrison, J., Hirsch, E.A. (eds.) CSR 2006. LNCS, vol. 3967, pp. 600–611. Springer, Heidelberg (2006). https://doi.org/10.1007/11753728_60
29. Tentrup, L., Rabe, M.N.: Clausal abstraction for DQBF. In: Janota, M., Lynce, I. (eds.) SAT 2019. LNCS, vol. 11628, pp. 388–405. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24258-9_27