



The strength of the dominance rule

Leszek Aleksander Kołodziejczyk ✉ 

Institute of Mathematics, University of Warsaw, Poland

Neil Thapen ✉ 

Institute of Mathematics, Czech Academy of Sciences, Czechia

Abstract

It has become standard that, when a SAT solver decides that a CNF Γ is unsatisfiable, it produces a certificate of unsatisfiability in the form of a refutation of Γ in some proof system. The system typically used is DRAT, which is equivalent to extended resolution (ER) – for example, until this year DRAT refutations were required in the annual SAT competition. Recently [Bogaerts et al. 2023] introduced a new proof system, associated with the tool VeriPB, which is at least as strong as DRAT and is further able to handle certain symmetry-breaking techniques. We show that this system simulates the proof system G_1 , which allows limited reasoning with QBFs and forms the first level above ER in a natural hierarchy of proof systems. This hierarchy is not known to be strict, but nevertheless this is evidence that the system of [Bogaerts et al. 2023] is plausibly strictly stronger than ER and DRAT. In the other direction, we show that symmetry-breaking for a single symmetry can be handled inside ER.

2012 ACM Subject Classification Theory of computation \rightarrow Proof complexity

Keywords and phrases proof complexity, DRAT, symmetry breaking, dominance

Funding *Neil Thapen*: Supported by the Institute of Mathematics, Czech Academy of Sciences (RVO 67985840) and GAČR grant 23-04825S.

1 Introduction

We write Lit for the set of propositional literals and $0, 1$ for the propositional constants. A *clause* is a disjunction which may contain $0, 1$ or a literal together with its negation; in the last two cases we call it *tautologous*. A *conjunctive normal form* formula, or CNF, is a set of clauses, understood as a conjunction. We write \perp for the empty clause and \top for the empty CNF. For a clause $C = x_1 \vee \dots \vee x_k$ we write $\neg C$ for the CNF $\neg x_1 \wedge \dots \wedge \neg x_k$.

A *substitution* is a map $\omega : \text{Lit} \cup \{0, 1\} \rightarrow \text{Lit} \cup \{0, 1\}$ which respects negations and is the identity on $\{0, 1\}$. We view a partial assignment as a kind of substitution. For literals p_1, \dots, p_k we write $\vec{p}_{\upharpoonright\omega}$ for the tuple $\omega(p_1), \dots, \omega(p_k)$. For a clause C , we write $C_{\upharpoonright\omega}$ for the clause $\{\omega(p) : p \in C\}$. For a CNF Γ , we write $\Gamma_{\upharpoonright\omega}$ for $\{C_{\upharpoonright\omega} : C \in \Gamma\}$. We write $\omega \models \Gamma$ to mean that $\Gamma_{\upharpoonright\omega}$ is tautologous, that is, every clause is tautologous. If ω is a partial assignment this is the same as the usual meaning of \models . The composition of substitutions τ, ω is defined by $\tau \circ \omega(p) = \tau(\omega(p))$. Note that $\Gamma_{\upharpoonright\tau \circ \omega} = (\Gamma_{\upharpoonright\omega})_{\upharpoonright\tau}$ and thus $\tau \models \Gamma_{\upharpoonright\omega}$ if and only if $\tau \circ \omega \models \Gamma$.

► **Definition 1.** A *symmetry* of a CNF Γ is a substitution ω such that $\Gamma_{\upharpoonright\omega} = \Gamma$.

Hard combinatorial formulas often have many symmetries. A successful heuristic to make such formulas easier for SAT solvers is *symmetry-breaking*, usually in the form of introducing a *lex-leader* constraint [13]. We illustrate this in the following proposition. Suppose Γ is a CNF in variables z_1, \dots, z_n and suppose we have a formula $[x_1, \dots, x_n \leq_{\text{lex}} y_1, \dots, y_n]$ expressing the lexicographic order on assignments to \vec{x} and \vec{y} . For now we suppress technical issues of how exactly we express this, and tacitly treat $[\vec{x} \leq_{\text{lex}} \vec{y}]$ as though it is a CNF.

► **Proposition 2.** If ω is a symmetry of Γ , then Γ and $\Gamma \cup [\vec{z} \leq_{\text{lex}} \vec{z}_{\upharpoonright\omega}]$ are equisatisfiable.

Proof. Suppose Γ is satisfiable. Let α be a lexicographically minimal assignment such that $\alpha \models \Gamma$. Then by symmetry $\alpha \models \Gamma_{\uparrow\omega}$ and thus $\alpha \circ \omega \models \Gamma$. By minimality we have $\alpha \leq_{\text{lex}} \alpha \circ \omega$ which, working through the definitions, gives us that $\alpha \models [\vec{z} \leq_{\text{lex}} \vec{z}_{\uparrow\omega}]$ as required. \blacktriangleleft

The new formula $\Gamma \cup [\vec{z} \leq_{\text{lex}} \vec{z}_{\uparrow\omega}]$ is potentially much easier to solve than the original Γ , because the extra constraint can substantially shrink the space of partial assignments that the solver has to search through. Note that the proof still works if we simultaneously add lex-leader constraints for several different symmetries.

It is now common for SAT solvers to have a *proof-logging* component which, when it decides that a formula Γ is unsatisfiable, will generate a certificate of unsatisfiability – in other words, a refutation of Γ . For this to be useful, it must be in a well-known proof system for which trusted software exists to verify that the refutation is correct. The standard system used for this is DRAT [30]. However, it is open to what extent DRAT can (feasibly) handle reasoning that uses symmetry-breaking, especially for more than one symmetry [17].

This issue is addressed in [4] which introduces a new proof-logging system with tools to handle quite general symmetry-breaking, extending DRAT. It builds on the system VeriPB [15] and as such uses reasoning using linear inequalities, rather than clauses; furthermore it is equipped to solve optimization problems, rather than just satisfiability.

We study the proof complexity of this system (as a refutation system for CNFs). We show that it is equivalent to G_1 , a system based on limited reasoning with QBFs which is one level above extended resolution (ER) in a natural hierarchy of proof systems [24]. This is in contrast to the redundancy-based systems which have been studied recently, such as DRAT, propagation redundancy and substitution redundancy, which are in their full generality all equivalent to ER [22, 18, 5]. This may represent another step in the strength of reasoning used in SAT algorithms, like from DPLL and treelike resolution, to CDCL and resolution, to the just-mentioned systems and ER [14, 26, 1]. In particular, it is unlikely that the new system in [4] can be simulated by DRAT.

The main tool used in [4] for symmetry-breaking is called the *dominance-based strengthening rule*. It is based on the following principle, which we express here in the language of CNFs and call “informal” because we are sweeping under the carpet the exact nature of the formulas Δ and $[\vec{x} <_{\text{lex}} \vec{y}]$ (intended to express strict lexicographic order).

► **Proposition 3 (informal).** *Let Γ, Δ be formulas, where Γ is in variables \vec{z} and has the property that any assignment satisfying Γ can be extended to satisfy Δ . Suppose we have a clause C and a substitution ω such that*

$$\Gamma \wedge \Delta \wedge \neg C \models \Gamma_{\uparrow\omega} \wedge [\vec{z}_{\uparrow\omega} <_{\text{lex}} \vec{z}]. \quad (1)$$

Then Γ and $\Gamma \wedge C$ are equisatisfiable.

Proof (sketch). Suppose Γ is satisfiable. Let α be a lexicographically minimal assignment such that $\alpha \models \Gamma$. Extend α to $\alpha \cup \beta$ satisfying $\Gamma \wedge \Delta$. If $\alpha \cup \beta \models C$ then we are done. Otherwise, from the entailment (1) we know $\alpha \cup \beta \models \Gamma_{\uparrow\omega} \wedge [\vec{z}_{\uparrow\omega} <_{\text{lex}} \vec{z}]$. Letting $\alpha' = (\alpha \cup \beta) \circ \omega$, we conclude that $\alpha' \models \Gamma$ and $\alpha' <_{\text{lex}} \alpha$, contradicting the minimality of α . \blacktriangleleft

The rule says roughly: if we have derived Γ , and have available a suitable Δ , ω and a proof of entailment (1), then we can derive $\Gamma \wedge C$. (In fact, the full rule in [4] is more general, since it is not restricted to the lexicographic ordering.)

Machinery exists to study the strength of rules of this kind by studying how easy it is to prove their soundness in the first-order setting of bounded arithmetic. Specifically, to carry out the proof of Proposition 3 it is sufficient to know that a nonempty polynomial-time set

(in this case the set of α such that $\alpha \models \Gamma$) always has a lexicographically least element. This puts the system inside the theory T_2^1 , which is associated with G_1 (see below for definitions).

The harder direction, lower-bounding the strength of the rule, uses similar machinery. Consider the transformation $\alpha \mapsto \alpha'$ in the proof above. We claim that, given an arbitrary polynomial time function f , we can construct an instance of the rule where this transformation is given by f . We do this by defining Δ to compute f and store the resulting values in β , and defining ω to pull these values from β back to the \vec{z} variables, as α' . Observe that in the proof, it is not actually necessary for α to be lexicographically minimal; the “local minimality” property that $\alpha \leq_{\text{lex}} f(\alpha)$ is sufficient. We show that a converse holds: that using the rule, we can find a local minimum of this kind, or to say it more precisely, we can efficiently derive a contradiction from the statement that there is no local minimum.

The problem of finding such a local minimum is known as polynomial local search, or PLS [21]. It is known that if we are only interested in sentences of low quantifier complexity, such as “CNFs in this family are unsatisfiable”, then every logical consequence of “every polynomial-time set has a least element”, that is, of T_2^1 , is already a consequence of the apparently weaker statement “every PLS problem has a solution”. Using this we can show that, roughly speaking, every CNF which can be proved unsatisfiable in T_2^1 has a short refutation in the proof system in [4], where the refutation uses an instance of the rule constructed from a PLS problem as described above. It follows that the system simulates G_1 .

The rest of the paper fills in the details of these arguments. In Section 2 we recall the definitions of some proof systems we will need. These are G_1 itself; cutting planes, which is the foundation for the system in [4]; and ER, where we will need to work extensively with derivations as well as refutations. In Section 3 we define our version of the system in [4], where we have removed the machinery for handling optimization problems. We call this the *dominance proof system*, and in it the dominance-based strengthening rule can be used for rather general orderings, not just lexicographic. We will work more with a restriction of it, the *linear dominance proof system*, in which it is limited to essentially the lexicographic ordering. We also define a simpler auxiliary proof system ER-PLS, which uses clauses rather than inequalities and captures the properties of the system that are important for us (illustrated in Proposition 3). It will follow from our results that, as a system for refuting CNFs, it is equivalent to the linear dominance system. In Section 4 we describe some results we need from bounded arithmetic and give a formal definition of PLS. We then show our main result,

► **Theorem 4.** *The linear dominance system is equivalent to G_1 .*

The proof is in three parts. Section 5 contains the main technical work of the paper, showing, in Theorem 18, that ER-PLS simulates G_1 , as sketched above. Section 6 shows that the linear dominance system in turn simulates ER-PLS. For the remaining direction, that G_1 simulates linear dominance, it is enough to show that linear dominance is sound, provably in T_2^1 . This is in Section 7, where we also briefly discuss the difference between linear dominance and the full dominance system.

In Section 8 we study what we *can* prove in ER about fragments of these systems. We show essentially that, in Proposition 3 above, if the mysterious formula Δ is not present then we do not need to use minimality, or even PLS, in the proof. This is because without the step where we extend α to satisfy Δ , the move from α to α' does not involve any computation, but amounts to shuffling the components of α around using the substitution ω , and for simple graph-theoretical reasons we can compute the i th iteration ω^i in polynomial time without invoking any stronger principles. In Section 8.1 we use this observation to show a technical result, that a natural weakening of the linear dominance system is already simulated by ER.

In Section 8.2 we use a similar construction to show¹:

► **Theorem 5.** *Define the system Q_1 as ER plus the power to add a lex-leader constraint for a single symmetry. Then Q_1 is simulated by ER (and thus by DRAT).*

Appendix A contains some technical material postponed from other sections.

We finish this section by addressing the natural question: what is this hierarchy of proof systems above ER, and why should we expect it to be strict? After all, ER is already a very strong system with many tools for proving combinatorial and algebraic statements, and seems to lie well beyond current methods for proving lower bounds [28].

ER was shown in [11] to correspond to the theory PV, which models reasoning with polynomial-time concepts (see Section 4 for definitions). In [24] the systems G_0, G_1, G_2, \dots of quantified Boolean reasoning were introduced, to correspond in the same way to a hierarchy $T_2^0, T_2^1, T_2^2, \dots$ of bounded arithmetic theories extending PV (which we can identify with T_2^0 [20]), where T_2^1 can reason with P^{NP} concepts, T_2^2 with $P^{\Sigma_2^p}$, and so on. In particular, if we ignore issues of uniformity, the unsatisfiable CNFs with short refutations in G_i capture precisely the universal sentences² provable in T_2^i . It is a classical result that the fragments $\text{IS}_0, \text{IS}_1, \dots$ of Peano arithmetic are separated by universal sentences. Specifically, the consistency statement for IS_k has this form and is provable in IS_{k+1} but not in IS_k (see e.g. [16, Chapter I.4(c)]). It is expected, essentially by analogy, that the analogous theories T_2^0, T_2^1, \dots are also separated at the universal level by some kind of consistency statement, although it is known that classical consistency will not work [31].

In the case we are interested in here, of PV and T_2^1 , there is some evidence of separation at the $\forall\Sigma_1^b$ level (one step above universal) since it is a logical version of the question: is the TFNP class PLS different from FP? Here we at least have a relativized separation between PLS and FP [8], although this implies nothing directly about the unrelativized theories.

2 Traditional proof systems

We require that proof systems are sound and that refutations in a given system are recognizable in polynomial time. When comparing two systems P and Q we are usually interested in their behaviour when refuting CNFs, and we use the following basic definition.

► **Definition 6.** *We say that Q simulates P if there is a polynomial-time function which, given a P -refutation of a CNF Γ , outputs a Q -refutation of Γ . Q and P are equivalent if they simulate each other.*

Often it will make sense to discuss not only refutations of formulas, but also *derivations* of one formula from another. We will use the notation e.g. “a derivation $\Gamma \vdash \Delta$ ” instead of “a derivation of Δ from Γ ” and will write $\pi : \Gamma \vdash \Delta$ to express that π is such a derivation.

2.1 Quantified Boolean formulas and G_1

G_1 is a fragment of G , a proof system used for reasoning with quantified Boolean formulas. We give only a brief description of G – for more details see [23, Chapter 4]. We will only

¹ We believe this is more general than the result about single symmetries in [17], since we handle an arbitrary symmetry, not just an involution; see the discussion in [4].

² That is, sentences consisting of a sequence of unbounded universal quantifiers followed by a polynomial time predicate. We could also write $\forall\Pi_1^0$.

consider fragments of G as systems for refuting CNFs; for comparisons of G with some other systems in the context of proving quantified Boolean formulas, see e.g. [3, 9].

A *quantified Boolean formula*, or QBF, is built from propositional variables and connectives in the usual way, and also allows quantification over Boolean variables. That is, if $F(x)$ is a QBF containing a variable x , then so are $\exists x F(x)$ and $\forall x F(x)$. In this context $\exists x$ and $\forall x$ are Boolean quantifiers and these formulas semantically have the expected meanings $F(0) \vee F(1)$ and $F(0) \wedge F(1)$. We stratify QBFs into classes called $\Sigma_1^q, \Pi_1^q, \Sigma_2^q, \Pi_2^q$ etc. in the usual way, by counting quantifier alternations. In particular, Σ_1^q is the closure of the class of (quantifier-free) Boolean formulas under \vee, \wedge and \exists . The strength of proof systems working with QBFs is that they allow us to represent an “exponential-size concept” such as $\bigvee_{\vec{a} \in \{0,1\}^n} F(\vec{a})$ with a polynomial-size piece of formal notation $\exists x_1 \dots x_n F(\vec{x})$.

The proof system G is an extension of the propositional sequent calculus. In this context a *sequent* is an expression of the form

$$A_1, \dots, A_k \longrightarrow B_1, \dots, B_\ell$$

where A_1, \dots, A_k and B_1, \dots, B_ℓ are QBFs. Such a sequent is understood semantically to mean the same as $\bigwedge_i A_i \rightarrow \bigvee_j B_j$, and we say that an assignment *satisfies* a sequent if it satisfies this formula. A derivation in G is a sequence of sequents, each of which is either an axiom of the form $A \longrightarrow A$, or follows from one or two earlier sequents by one of the rules. These rules are sound and complete and we will not list them, as the details are not so important for us (see Section 4 for our justification of this). For $k \in \mathbb{N}$ the system G_k is the restriction of G which only allows formulas from $\Sigma_k^q \cup \Pi_k^q$ to appear in derivations.

We are interested in proof systems as ways of refuting CNFs. To turn G_1 into such a system we have the following definition, where for definiteness we think of Γ as a single QBF (rather than, say, as the cedent given by its clauses) and where \perp is the empty cedent.

► **Definition 7.** A G_1 refutation of a CNF Γ is a G_1 sequent calculus derivation of the sequent $\Gamma \longrightarrow \perp$.

2.2 Pseudo-Boolean constraints and cutting planes

Following [4], we use the term *pseudo-Boolean constraint*, or *PB constraint*, for a linear inequality with integer coefficients over 0/1-valued variables. We will sometimes call a set of such constraints a *PB formula*. PB constraints generalize clauses, since a clause $C = x_1 \vee \dots \vee x_n \vee \neg y_1 \vee \dots \vee \neg y_m$ can be expressed by a constraint of the form $x_1 + \dots + x_n + (1 - y_1) + \dots + (1 - y_m) \geq 1$. We call this constraint C^* and will also write Γ^* for the PB formula obtained by taking C^* for each clause C in a CNF Γ .

If C is a PB constraint $A\vec{x} \geq b$ we write $\neg C$ for the PB constraint $A\vec{x} \leq b - 1$. Note that, although it is semantically the same, this denotes a different piece of syntax from $\neg C$ when C is a clause. Given a substitution ω , we write $C|_\omega$ for the PB constraint obtained by simply replacing each variable x in C with $\omega(x)$, and we use a similar notation for PB formulas.

We use *cutting planes* [12], or CP, as a derivational system for deriving one PB formula G from another PB formula F . A CP derivation is a sequence of PB constraints, including every constraint from G , such that each constraint is either from F , or is a Boolean axiom of the form $x_i \geq 0$ or $x_i \leq 1$, or follows from earlier constraints by one of the rules. These are *addition* – we can derive a new constraint by summing integral multiples of two old constraints; and *rounding* – from a constraint $dA\vec{x} \geq b$, where $d > 0$ is an integral scalar and A is an integral matrix of coefficients, we can derive $A\vec{x} \geq \lceil b/d \rceil$. We use the notation $F \vdash_{\text{CP}} G$ for CP derivations. In a formal CP derivation, coefficients are written in binary.

2.3 Extended resolution as a derivational system

For CNFs Γ and Δ , a *resolution derivation* $\Gamma \vdash \Delta$ is a sequence of clauses, beginning with the clauses of Γ and containing every clause from Δ , such that each clause in the sequence is either in the initial copy of Γ or is derived from earlier clauses by the *resolution* or *weakening* rule. Here the resolution rule derives $C \vee D$ from $C \vee x$ and $D \vee \neg x$, for any variable x , and the weakening rule derives D from C for any $D \supseteq C$ (although see below for a restriction on weakening in the context of extended resolution derivations). Because we allow propositional literals 0 and 1 to occur in clauses, we need to be able to remove them, so we also allow derivation of C from $C \vee 0$ (this can be thought of as resolution with a notional axiom “1”). A *resolution refutation* of Γ is a derivation of $\Gamma \vdash \perp$.

An *extension axiom* has the form of three clauses $\neg u \vee \neg v \vee y$, $\neg y \vee u$ and $\neg y \vee v$ which together express that y is equivalent to $u \wedge v$. The intended use is that u, v are literals and y is a newly-introduced variable. For good behaviour under restrictions we also allow extension axioms of the form $\neg u \vee y$, $\neg y \vee u$ expressing that y is equivalent to a single existing literal, and of the form y or $\neg y$ expressing that y is equivalent to a constant.

► **Definition 8.** For CNFs Γ and Δ , an extended resolution (ER) derivation $\Gamma \vdash \Delta$ is a sequence of clauses, beginning with Γ and including every clause in Δ . Each clause in the sequence either appears in the initial copy of Γ , or is derived from earlier clauses by resolution or weakening (where weakening is not allowed to introduce a variable that has not appeared earlier in the sequence³) or by the extension rule, which allows us to introduce an extension axiom defining a variable that has not appeared earlier in the sequence from variables that have appeared earlier.

Such a derivation is sound in the following sense: any assignment to all variables in Γ which satisfies Γ can be extended to an assignment to all variables in Δ which satisfies Δ . The extension axioms in the derivation tell us explicitly how to extend the assignment.

We will often need to handle many extension axioms at once:

► **Definition 9.** Let \vec{x}, \vec{y} be disjoint tuples of variables. We say that Δ is a set of extension axioms over $\vec{x}; \vec{y}$ if it can be written as a sequence of extension axioms defining variables y_1, \dots, y_r in order, where each y_i is defined in terms of variables from among $\vec{x}, y_1, \dots, y_{i-1}$.

Equivalently, such a Δ can be thought of as describing a Boolean circuit and asserting that, on input \vec{x} , the values computed at the internal nodes are \vec{y} . We also introduce notation for writing sets of extension axioms in this way:

► **Definition 10.** Suppose \vec{x}, \vec{y} are tuples of variables and C is a circuit with gates of fan-in 2. We write $[\vec{y} = C(\vec{x})]$ for the set of extension axioms over $\vec{x}; \vec{y}$ expressing that the non-input nodes of the circuit have values \vec{y} on inputs \vec{x} (we assume \vec{x}, \vec{y} have suitable arities). If the circuit has a distinguished output node we label the corresponding variable in \vec{y} as y^{out} .

The condition that the extension rule must introduce new variables has some counter-intuitive consequences, and we must take extra care when we use ER as a derivational system. For example, there are ER derivations $\top \vdash x$ and $\top \vdash \neg x$, (where \top is the empty CNF) but there is no derivation $\neg x \vdash x$, even though $\neg x$ extends \top , and no derivation $\top \vdash x \wedge \neg x$. In many ways the new variables behave like existentially quantified bound variables.

³ This restriction on weakening is probably not strictly necessary. We include it because it has the helpful consequence that each new variable comes with an explicit definition in terms of the old variables.

Under some reasonable conditions on how extension variables are used, we can avoid problems related to such issues. Below we formally prove two lemmas of this kind, which we will refer to as needed. We could avoid the issue by working with some other system equivalent to ER, such as extended Frege, circuit Frege [19] or even treelike G_1 , but it would then be harder to show that the resulting system is simulated by linear dominance.

We use a convention that, in the context of ER derivations, when we write an expression of the form $\Gamma(\vec{x}) \vdash \Delta$ with some variables \vec{x} displayed on the left, we mean that every variable in \vec{x} is treated as an “old” variable in this derivation and as such is not used as an extension variable in any instance of the extension rule, even if it does not actually appear in the CNF $\Gamma(\vec{x})$.

► **Lemma 11.** *Let $\Gamma(\vec{x}), A(\vec{x}, \vec{y}), B(\vec{x}, \vec{z}), \Delta(\vec{x}, \vec{w})$ be CNFs, where we assume $\vec{x}, \vec{y}, \vec{z}, \vec{w}$ are disjoint and no other variables appear. Suppose we have ER derivations*

$$\pi_1 : \Gamma(\vec{x}) \wedge A(\vec{x}, \vec{y}) \vdash B(\vec{x}, \vec{z}) \quad \text{and} \quad \pi_2 : \Gamma(\vec{x}) \wedge B(\vec{x}, \vec{z}) \vdash \Delta(\vec{x}, \vec{z}, \vec{w}).$$

Then we can construct an ER derivation $\Gamma(\vec{x}) \wedge A(\vec{x}, \vec{y}) \vdash \Delta(\vec{x}, \vec{z}, \vec{w})$ in polynomial time.

Proof. We first copy π_1 . Then we copy π_2 , except that every extension variable in π_2 which is not in \vec{w} is given a new name, to avoid clashes with variables \vec{y} and other extension variables that appeared in π_1 . ◀

► **Lemma 12.** *Given an ER derivation $\pi_1 : \Gamma(\vec{x}) \vdash \Delta \wedge A$, where Δ is a set of extension axioms over $\vec{x}; \vec{y}$, we can construct in polynomial time an ER derivation $\pi_2 : \Gamma(\vec{x}) \wedge \Delta \vdash \Delta \wedge A$.*

Proof. Let Δ' and A' be the same as Δ and A except that we have replaced every variable y_i with a new variable z_i . From $\Gamma \wedge \Delta$, we can derive $\Delta' \wedge A'$ by a copy of π_1 with Δ added to the initial clauses and with each y_i changed to z_i everywhere outside of Δ . Then we can work through \vec{y} and derive, from the relevant extension axioms in Δ and Δ' , using the normal rules of resolution, that $y_i \leftrightarrow z_i$; formally, this is the two clauses $\neg y_i \vee z_i$ and $\neg z_i \vee y_i$. Finally we resolve these clauses with the clauses of A' to derive A . ◀

3 The dominance rule

We define three refutational proof systems using versions of the dominance-based strengthening rule of [4]. The *dominance* proof system is intended to be the same as the system described in [4] except that we have removed the machinery for talking about optimization, that is, everything related to the objective function f . The *linear dominance* proof system restricts this by only allowing a particular kind of ordering to be used in the dominance rule; the practical work in [4] in fact only needs this weaker system. Lastly we introduce our auxiliary system ER-PLS.

In these systems, often we can only apply a rule on the condition that some other derivation $\Gamma \vdash \Delta$ or $\mathcal{C} \vdash_{\text{CP}} \mathcal{D}$ exists, in ER or CP, possibly involving formulas that do not appear explicitly in the proof we are working on; or that some other polynomial-time-checkable object exists, such as a substitution ω . To ensure that correctness of a proof is checkable in polynomial time we implicitly require that, in a formal proof in a dominance-based system, each application of the rule is labelled with an example of the object in question, with the size of the labels (that is, the CP derivations, substitutions etc.) counted towards the size of the formal proof.

3.1 The dominance proof system

This is a system for refuting PB formulas. As in [4], we will call steps in a refutation *configurations*, rather than lines. A configuration is a quadruple $(\mathcal{C}, \mathcal{D}, \mathcal{O}_{\preceq}, \vec{z})$ where

- \mathcal{C} is a set of PB constraints called *core constraints*
- \mathcal{D} is a set of PB constraints called *derived constraints*
- $\mathcal{O}_{\preceq}(\vec{x}, \vec{y})$ is a PB formula where \vec{x} and \vec{y} both have the same arity as \vec{z}
- \vec{z} is a tuple of variables.

We put no conditions on which variables appear in \mathcal{C} and \mathcal{D} , except that the variables \vec{x}, \vec{y} in $\mathcal{O}_{\preceq}(\vec{x}, \vec{y})$ should be thought of as dummy variables that are not related to the rest of the proof. In a valid proof, in every configuration the formula $\mathcal{O}_{\preceq}(\vec{x}, \vec{y})$ defines a preorder and we use this with \vec{z} to define a preorder \preceq on assignments, writing $\alpha \preceq \beta$ if $\mathcal{O}_{\preceq}(\vec{x}, \vec{y})$ is satisfied under the assignment that takes \vec{x} to $\alpha(\vec{z})$ and \vec{y} to $\beta(\vec{z})$.

Semantically a configuration can be thought of as asserting that \mathcal{C} is satisfiable, and that if we order assignments by \mathcal{O}_{\preceq} on \vec{z} as described above, then for any assignment α satisfying \mathcal{C} , some assignment β with $\beta \preceq \alpha$ satisfies $\mathcal{C} \cup \mathcal{D}$ (see Definition 19 below).

A refutation of a PB formula F is then a sequence of configurations, beginning with $(F, \emptyset, \top, \emptyset)$, where \top is the empty PB formula, and ending with a configuration in which \mathcal{C} or \mathcal{D} contains the contradiction \perp , that is, $0 \geq 1$. Each configuration is derived from the previous configuration $(\mathcal{C}, \mathcal{D}, \mathcal{O}_{\preceq}, \vec{z})$ by one of the following rules:

Implicational derivation rule. Derive $(\mathcal{C}, \mathcal{D} \cup \{C\}, \mathcal{O}_{\preceq}, \vec{z})$, if there is a derivation⁴ $\mathcal{C} \cup \mathcal{D} \vdash_{\text{CP}} C$.

(**Objective bound update rule** – this appears in [4], but we omit it from our systems as it only affects the objective function f , which we do not use.)

Redundance-based strengthening rule. Derive $(\mathcal{C}, \mathcal{D} \cup \{C\}, \mathcal{O}_{\preceq}, \vec{z})$ if there is a substitution ω and a derivation $\mathcal{C} \cup \mathcal{D} \cup \{-C\} \vdash_{\text{CP}} (\mathcal{C} \cup \mathcal{D} \cup \{C\})_{\uparrow\omega} \cup \mathcal{O}_{\preceq}(\vec{z}_{\uparrow\omega}, \vec{z})$.

Deletion rule. Derive $(\mathcal{C}', \mathcal{D}', \mathcal{O}_{\preceq}, \vec{z})$ if

1. $\mathcal{D}' \subseteq \mathcal{D}$ and
2. $\mathcal{C}' = \mathcal{C}$ or $\mathcal{C}' = \mathcal{C} \setminus \{C\}$ for some constraint C derivable by the redundance rule above from $(\mathcal{C}', \emptyset, \mathcal{O}_{\preceq}, \vec{z})$ ⁵.

Transfer rule. Derive $(\mathcal{C}', \mathcal{D}, \mathcal{O}_{\preceq}, \vec{z})$ if $\mathcal{C} \subseteq \mathcal{C}' \subseteq \mathcal{C} \cup \mathcal{D}$. In other words, we can copy constraints from \mathcal{D} to \mathcal{C}

Dominance-based strengthening rule. We first give a slightly informal definition: derive $(\mathcal{C}, \mathcal{D} \cup \{C\}, \mathcal{O}_{\preceq}, \vec{z})$ if there is a substitution ω and, informally, a derivation

$$\mathcal{C} \cup \mathcal{D} \cup \{-C\} \vdash_{\text{CP}} \mathcal{C}_{\uparrow\omega} \cup (\vec{z}_{\uparrow\omega} \prec \vec{z})$$

where $\vec{z}_{\uparrow\omega} \prec \vec{z}$ expresses that $\vec{z}_{\uparrow\omega}$ is strictly smaller than \vec{z} in the ordering \mathcal{O}_{\preceq} . However it may be that any PB formula expressing the strict inequality $(\vec{z}_{\uparrow\omega} \prec \vec{z})$ is very large. So

⁴ In [4], the derivation is allowed to use some additional inferences beyond those of CP. For simplicity we omit these, as in the presence of the redundance-based strengthening rule, even strengthening CP here to a system like extended Frege would not make any difference to the overall dominance system. In particular, our proof of the simulation of linear dominance by G_1 in Section 7 would still go through.

⁵ This restriction of the deletion rule ensures that it preserves semantic validity under the intuitive meaning of configurations mentioned above. See Section 7 for an argument.

formally the rule is: derive $(\mathcal{C}, \mathcal{D} \cup \{C\}, \mathcal{O}_{\preceq}, \vec{z})$ if there is a substitution ω and two derivations

$$\begin{aligned} & \mathcal{C} \cup \mathcal{D} \cup \{-C\} \vdash_{\text{CP}} \mathcal{C}_{\upharpoonright\omega} \cup \mathcal{O}_{\preceq}(\vec{z}_{\upharpoonright\omega}, \vec{z}) \\ & \mathcal{C} \cup \mathcal{D} \cup \{-C\} \cup \mathcal{O}_{\preceq}(\vec{z}, \vec{z}_{\upharpoonright\omega}) \vdash_{\text{CP}} \perp. \end{aligned}$$

Order change rule. From $(\mathcal{C}, \emptyset, \mathcal{O}_{\preceq}, \vec{z})$ derive $(\mathcal{C}, \emptyset, \mathcal{O}'_{\preceq}, \vec{z}')$ if \mathcal{O}'_{\preceq} is CP-provably a preorder. That is, if there are derivations $\emptyset \vdash_{\text{CP}} \mathcal{O}_{\preceq}(\vec{u}, \vec{u})$ and $\mathcal{O}_{\preceq}(\vec{u}, \vec{v}) \cup \mathcal{O}'_{\preceq}(\vec{v}, \vec{w}) \vdash_{\text{CP}} \mathcal{O}'_{\preceq}(\vec{u}, \vec{w})$.

3.2 The linear dominance proof system

This restricts the dominance proof system to only use orderings \mathcal{O}_{\preceq} arising from a multilinear objective function. Formally, we require that $\mathcal{O}_{\preceq}(\vec{x}, \vec{y})$ is always of the form $f(\vec{x}) \leq f(\vec{y})$, where f is a multilinear function $\vec{x} \mapsto \sum_i b_i x_i$ for some constants b_i . These constants can be changed using the order change rule. We are no longer required to explicitly prove that $\mathcal{O}_{\preceq}(\vec{x}, \vec{y})$ is an ordering, as CP can always prove this for this restricted form.

The most important ordering of this form is the lexicographic ordering, which we get by setting $f(\vec{x}) \mapsto \sum_i 2^i x_i$ (for a suitable ordering of the variables in \vec{x}).

3.3 The system ER-PLS

This system uses the clausal version of the dominance rule sketched in Proposition 3 in the introduction. The name is intended to suggest that it has a similar connection to polynomial local search “computations” as ER has to polynomial time.

We fix a polynomial-time constructible family of CNFs defining lexicographic ordering. That is, for each k we have a CNF $[x_1, \dots, x_k \leq_{\text{lex}} y_1, \dots, y_k]$, which may also use some auxiliary variables \vec{z} , such that for all $\alpha, \beta \in \{0, 1\}^k$ there is an assignment to \vec{z} satisfying $[\alpha \leq_{\text{lex}} \beta]$ if and only if $\alpha \leq \beta$ lexicographically. It is not too important which CNF we use for $[\vec{x} \leq_{\text{lex}} \vec{y}]$, but we specify a convenient one in Appendix A.1 below.

An ER-PLS refutation of a CNF Γ is formally a sequence of CNFs, beginning with Γ and ending with a CNF containing the empty clause \perp . At each step we apply one of the two rules below to derive the next CNF in the sequence.

ER rule. From Γ derive $\Gamma \wedge \Delta$ if there is an ER derivation $\Gamma \vdash \Delta$, in the sense of Definition 8.

Dominance rule. Let \vec{x} list all variables of Γ in some order and let C be a clause in these variables. From Γ derive $\Gamma \wedge C$, provided we have

1. a set Δ of extension axioms over $\vec{x}; \vec{y}$
2. a substitution ω mapping variables \vec{x} to variables among $\vec{x} \cup \vec{y}$
3. two ER derivations
 - a. $\Gamma \wedge \Delta \wedge \neg C \vdash \Gamma_{\upharpoonright\omega}$
 - b. $\Gamma \wedge \Delta \wedge \neg C \wedge [\vec{x} \leq_{\text{lex}} \vec{x}_{\upharpoonright\omega}] \vdash \perp$

with the technical condition that the auxiliary variables \vec{z} used in $[\vec{x} \leq_{\text{lex}} \vec{x}_{\upharpoonright\omega}]$ may not appear in Γ , Δ , or C .

Informally, condition 3 can be thought of as asking for a single ER derivation $\Gamma \wedge \Delta \wedge \neg C \vdash \Gamma_{\upharpoonright\omega} \wedge [\vec{x}_{\upharpoonright\omega} <_{\text{lex}} \vec{x}]$, as in Proposition 3. Note that we do not have any deletion rule, that is, we can only grow working set of clauses Γ , and never shrink it. This is because Γ is modelled on the *core* constraints \mathcal{C} in the dominance system, which can only be deleted in very specific situations. For simplicity we do not allow deletion at all, as we will not need it.

► **Lemma 13.** *If Γ' is derived from Γ by a rule of ER-PLS, then Γ' and Γ are equisatisfiable.*

Proof. The only nontrivial case is the forward direction of the dominance rule. This is proved in the same way as Proposition 3 in the introduction, with the cosmetic change that we now have $[\vec{x} \leq_{\text{lex}} \vec{x}_{\uparrow\omega}]$ on the left of the entailment rather than $[\vec{x}_{\uparrow\omega} <_{\text{lex}} \vec{x}]$ on the right. We must also deal now with the auxiliary variables in $[\vec{x} \leq_{\text{lex}} \vec{x}_{\uparrow\omega}]$, but since these are not in the domain of the ordering this presents no problem. \blacktriangleleft

4 Bounded arithmetic

We will carry out some arguments in theories of bounded arithmetic, which we will turn into propositional proofs using variants of well-known translations. Here we give a brief overview – for more see e.g. [23, Chapters 9 and 12]. When we write that a formula with free variables is provable in a first-order theory, we mean that its universal closure is provable.

4.1 Theories

PV is the canonical theory for polynomial-time reasoning [11]. Its language contains a function symbol, called a *PV function*, for every polynomial-time algorithm on \mathbb{N} . Its axioms are defining equations for all PV functions, based on Cobham’s characterization of polynomial-time functions. See e.g. [23, Chapter 12.1] for a precise definition (there the theory is called PV_1). Importantly, PV proves the principle of mathematical induction for any property defined by a *PV formula* – that is, by a quantifier-free formula in the language of PV. Such formulas define precisely the polynomial-time properties.

More powerful theories can be obtained by extending PV with stronger induction axioms. A formula in the language of PV is Σ_1^b if it has the form $\exists x \leq t \varphi$, where t is a term not containing x and φ is a PV formula; unsurprisingly, Σ_1^b formulas define exactly properties in NP. A formula is Σ_2^b if it has the form $\exists x \leq t_1 \forall y \leq t_2 \varphi$ for φ a PV formula. The classes Π_1^b and Π_2^b are defined dually.

The theory T_2^1 (a more accurate name would be $T_2^1(\text{PV})$) extends PV by induction axioms for all Σ_1^b formulas. T_2^1 is the weakest theory that suffices to prove the least number principle for Σ_1^b formulas, that is, that every nonempty NP set has a least element; actually, even the least number principle for PV formulas already implies T_2^1 over PV [6].

The theory S_2^1 , intermediate in strength between PV and T_2^1 , extends PV by the *length induction* axioms for Σ_1^b formulas, that is, universal closures of statements of the form

$$\psi(0) \wedge \forall x (\psi(x) \rightarrow \psi(x+1)) \rightarrow \forall x \psi(|x|),$$

where ψ is Σ_1^b and $|\cdot|$ stands for the *length* function that takes a number x to its length in binary notation. The theory S_2^2 is a strengthening of T_2^1 that additionally contains length induction for Σ_2^b formulas. It should be noted that S_2^1 proves length induction also for Π_1^b formulas, T_2^1 proves induction also for Π_1^b formulas, and so on.

Ordered by strength, we have $\text{PV} \subseteq S_2^1 \subseteq T_2^1 \subseteq S_2^2 \subseteq \dots$. There is also partial conservativity between some adjacent theories. In particular, S_2^1 is $\forall\Sigma_1^b$ -conservative over PV and S_2^2 is $\forall\Sigma_2^b$ -conservative over T_2^1 [7]. This means that if $\psi(x)$ is a Σ_1^b formula and S_2^1 proves $\forall x \psi(x)$, then PV proves it as well; analogously for Σ_2^b formulas, S_2^2 and T_2^1 .

There is a well-known connection between propositional proof systems and arithmetic theories, linking ER to PV (and S_2^1) and G_1 to T_2^1 (and S_2^2). The following theorem, which shows that a theory proves the soundness of the corresponding proof system, can be viewed as an upper bound: it says that, for example, G_1 is in some sense no stronger than T_2^1 . We use this for our main result in Section 5.

► **Theorem 14** ([11, 24]). S_2^1 – and by conservativity, PV – proves the CNF-reflection principle for ER: “any CNF refutable in ER is unsatisfiable”. Similarly, S_2^2 – and by conservativity, T_2^1 – proves the CNF-reflection principle for G_1 : “any CNF refutable in G_1 is unsatisfiable”.

Proof sketch. We first consider G_1 . The statement “sequent s , containing only formulas from $\Sigma_1^q \cup \Pi_1^q$, is satisfied by assignment α to its free variables” can be naturally written as a Π_2^b formula $\sigma(s, \alpha)$. Using this we formalize the natural proof of the soundness of G_1 as a length induction on a Π_2^b formula, roughly as follows: given a G_1 derivation π , we show by induction down π that $\forall \alpha \sigma(s, \alpha)$ holds for every line s of π . If the last line has the form $\Gamma \vdash \perp$ for a CNF Γ , this means that Γ cannot be satisfiable.

The part about ER is proved by a similar argument, but with the length induction hypothesis being “every line in the ER refutation π up to the current one is satisfiable”. This can be stated in a Σ_1^b way, holds at the beginning of π if the CNF Γ being refuted is satisfiable, but no longer holds once π reaches the empty clause \perp . ◀

In the other direction, it is possible to translate proofs in an arithmetic theory into uniform families of propositional proofs. We will use this translation for PV and ER, and we give a slight refinement of it in Section 4.2. Such translations can also be used to give a kind of converse to Theorem 14, with a proof similar to our approach in Section 5 below:

► **Theorem 15** ([11, 24]). If PV (equivalently S_2^1) proves the CNF-reflection principle for a propositional proof system Q , then ER simulates Q . Similarly, if T_2^1 (equivalently S_2^2) proves the CNF-reflection principle for Q , then G_1 simulates Q .

Theorems 14 and 15 together give us a close association between T_2^1 and short G_1 proofs, and in fact we prove our results about G_1 indirectly using these theorems, rather than by reasoning about G_1 itself. This is largely the reason that we did not include a complete description of G_1 in Section 2.1.

Another important property of T_2^1 – and, by conservativity, of S_2^2 – is that its $\forall \Sigma_1^b$ consequences can be witnessed by *polynomial local search*. We formally define a PLS problem as a triple (t_w, θ_w, N_w) , where t, N are respectively a unary and a binary PV function, θ is a binary PV formula, and the distinguished argument w written in the subscript is an instance of the problem. The formula θ defines the domain of the problem on instance w (tacitly, any element of the domain is required to be at most polynomially larger than w); t_w is an initial value that should be in the domain; and N_w is a one-place *neighbourhood* function, which attempts to map any value in the domain to a strictly smaller value in the domain. Since the domain has a least element as long as it is nonempty, N_w will sometimes fail, and a solution to the problem on instance w is either t_w , if $\neg \theta_w(t_w)$, or a value y such that $\theta_w(y)$ but either $\neg \theta_w(N_w(y))$ or $N_w(y) \geq y$.

The PLS witnessing theorem for T_2^1 , originally proved in [8], says that if T_2^1 proves $\forall w \exists y \leq t \varphi$, where φ is a PV formula, then the task of finding y given w can be reduced to a PLS problem. Written in a modern form, which also includes an upper bound on the strength of the theory needed to prove correctness of the reduction, we have:

► **Theorem 16** ([2, Theorem 2.5]). Assume that $T_2^1 \vdash \forall w \exists y \leq t \varphi(w, y)$, where φ is a PV formula. Then there is a PLS problem $Q_w = (t_w, \theta_w, N_w)$ and a PV function f such that the following are provable in PV:

1. $\neg \theta_w(t_w) \rightarrow \varphi(w, f(w))$
2. $\theta_w(z) \wedge \neg \theta_w(N_w(z)) \rightarrow \varphi(w, f(z))$
3. $\theta_w(z) \wedge N_w(z) \geq z \rightarrow \varphi(w, f(z))$.

4.2 Propositional translations

We use a version of the translation from PV proofs to polynomial-time constructible families of ER proofs, due to Cook [11].⁶ We first describe how to translate formulas. Consider a PV formula $\theta(\vec{x})$. Let \vec{k} represent a choice of binary bit-lengths for the variables \vec{x} (we will not be very formal about \vec{k} ; we can think of this notation as assigning a length to every free first-order variable in the universe). Supposing $\vec{x} = x_1, \dots, x_\ell$, we will code each x_i using a k_i -tuple of new propositional variables $x_i^1, \dots, x_i^{k_i}$, which tuple we write just as \vec{x}_i .

By definition, θ is a quantifier-free formula built from PV functions. So we can construct, in some canonical way based on the structure of θ , in time polynomial in the bit-lengths k_1, \dots, k_ℓ , a Boolean circuit $C(\vec{x}_1, \dots, \vec{x}_\ell)$ evaluating θ on binary inputs of these lengths. Following Definition 10, we introduce a tuple of new variables \vec{z} , one for each node in C , and define the propositional translation $\llbracket \theta(\vec{x}) \rrbracket_{\vec{k}}$ to be the CNF $[\vec{z} = C(\vec{x}_1, \dots, \vec{x}_\ell)] \wedge z^{\text{out}}$. Unless stated otherwise, we assume that the translations of any two explicitly listed formulas have disjoint auxiliary variables \vec{z} . For example, in Proposition 17, we assume that the translations of $\varphi_1, \dots, \varphi_r, \theta$ all have disjoint auxiliary variables, even if some formula appears twice in this list.

We now state how we translate proofs. For the proof see Appendix A.2.

► **Proposition 17.** *Suppose PV proves a sentence $\forall \vec{x}, \varphi_1(\vec{x}) \wedge \dots \wedge \varphi_r(\vec{x}) \rightarrow \theta(\vec{x})$, where $\varphi_1, \dots, \varphi_r, \theta$ are quantifier-free. Then for any assignment \vec{k} of bit-lengths to the variables \vec{x} , we can construct in time polynomial in \vec{k} an ER derivation*

$$\llbracket \varphi_1(\vec{x}) \rrbracket_{\vec{k}} \wedge \dots \wedge \llbracket \varphi_r(\vec{x}) \rrbracket_{\vec{k}} \vdash \llbracket \theta(\vec{x}) \rrbracket_{\vec{k}}.$$

5 ER-PLS simulates G_1

This section contains the main technical work of the paper. We begin by constructing some PV proofs. We will use propositional translations of these in our simulation.

Let $\text{Sat}(a, x)$ be a natural PV formula expressing that CNF a is satisfied by assignment x . Let $\text{Ref}(a, b)$ be a natural PV formula expressing that b is a G_1 refutation of a . We may take $\forall a, b, x \neg \text{Sat}(a, x) \vee \neg \text{Ref}(a, b)$ as the CNF-reflection principle for G_1 , stating that any CNF refutable in G_1 is unsatisfiable. By Theorem 14 this is provable in T_2^1 .

CNF-reflection is a universally-quantified PV formula, so in particular it is $\forall \Sigma_1^b$. Thus by Theorem 16, there is a PLS problem $Q_w = (t_w, \theta_w, N_w)$, where to save space we think of a, b, x as combined into a single parameter w which we write as a subscript, such that the existence of a solution to Q_w implies $\neg \text{Sat}(a, x) \vee \neg \text{Ref}(a, b)$, provably in PV. Precisely, PV proves the following three formulas, in free variables a, b, x, y (note that since the CNF-reflection principle does not contain an existential quantifier, we do not need the function f that appears in Theorem 16):

1. $\neg \theta_w(t_w) \rightarrow (\neg \text{Sat}(a, x) \vee \neg \text{Ref}(a, b))$
2. $\theta_w(y) \wedge \neg \theta_w(N_w(y)) \rightarrow (\neg \text{Sat}(a, x) \vee \neg \text{Ref}(a, b))$
3. $\theta_w(y) \wedge N_w(y) \geq y \rightarrow (\neg \text{Sat}(a, x) \vee \neg \text{Ref}(a, b))$.

⁶ We emphasize that we are using the Cook translation, rather than the Paris-Wilkie translation of e.g. [27]. The Paris-Wilkie translation is usually used to translate first-order proofs involving an oracle symbol into families of small proofs in relatively weak propositional systems. For example, it translates (a relativized version of) T_2^1 into polylogarithmic-width resolution.

By standard properties of PLS, we may assume in order to simplify some things below that the bit-length of t_w depends only on the components a, b of w and not on the assignment x , and that N_w is hard-wired to never give output bigger than t_w .

Making some small rearrangements and introducing a new variable u for the neighbour of y , we get that PV proves

- F1. $\text{Sat}(a, x) \wedge \text{Ref}(a, b) \wedge y = t_w \rightarrow \theta_w(y)$
- F2. $\text{Sat}(a, x) \wedge \text{Ref}(a, b) \wedge \theta_w(y) \wedge u = N_w(y) \rightarrow \theta_w(u)$
- F3. $\text{Sat}(a, x) \wedge \text{Ref}(a, b) \wedge \theta_w(y) \wedge u = N_w(y) \wedge y \leq u \rightarrow \perp$.

With these proofs in hand we can describe the simulation.

► **Theorem 18.** ER-PLS *simulates* G_1 .

Proof. We are given a CNF A and a G_1 refutation B of A . We want to construct, in polynomial time, an ER-PLS refutation of A . We will build the refutation using propositional translations of the proofs $F1$ - $F3$ above. We begin by calculating the bit-length of the variables a, b, x, y, u which we will use in the translation.

Let n be the number of variables in A and let m and ℓ be the bit-length of the strings coding A and B respectively. We may assume $n \leq m$. We will use m, ℓ, n as the respective bounds on the bit lengths of a, b, x . By our simplifying assumption on the problem Q_w we can find a bound r , polynomial in m and ℓ , on the bit-length of t_w for parameters w of the lengths we are considering. We may also use r as the bit-length for both y and u , since N_w needs at most r bits to encode its output. Thus we use these bounds m, ℓ, n, r, r as the bit-length parameter \vec{k} in all our propositional translations below. As a result, all the CNFs we obtain from the translation will have size polynomial in $m + \ell$. For simplicity of notation, we will omit actually writing the subscript \vec{k} .

Applying Proposition 17 to $F1$, $F2$ and $F3$, we obtain, in time polynomial in $m + \ell$, the following ER derivations:

- $P1$: $\llbracket \text{Sat}(a, x) \rrbracket \wedge \llbracket \text{Ref}(a, b) \rrbracket \wedge \llbracket y = t_w \rrbracket \vdash \llbracket \theta_w(y) \rrbracket$
- $P2$: $\llbracket \text{Sat}(a, x) \rrbracket \wedge \llbracket \text{Ref}(a, b) \rrbracket \wedge \llbracket \theta_w(y) \rrbracket \wedge \llbracket u = N_w(y) \rrbracket \vdash \llbracket \theta_w(u) \rrbracket$
- $P3$: $\llbracket \text{Sat}(a, x) \rrbracket \wedge \llbracket \text{Ref}(a, b) \rrbracket \wedge \llbracket \theta_w(y) \rrbracket \wedge \llbracket u = N_w(y) \rrbracket \wedge \llbracket y \leq u \rrbracket \vdash \perp$.

These formulas and derivations are in propositional variables $\vec{a}, \vec{b}, \vec{x}, \vec{y}, \vec{u}$ that arise from a, b, x, y, u in the translation, using the bit-lengths described above (plus the requisite auxiliary and extension variables).

Now let τ be the substitution (in fact a partial assignment) that first replaces the propositional variables \vec{a} and \vec{b} with the actual bits of A and B . Then, recalling that the CNFs $\llbracket \text{Sat}(a, x) \rrbracket$ and $\llbracket \text{Ref}(a, b) \rrbracket$ are defined in terms of Boolean circuits taking $\vec{a}, \vec{b}, \vec{x}$ as input, we compute the values of all nodes in these circuits that do not depend on \vec{x} (for $\llbracket \text{Ref}(a, b) \rrbracket$ this means all nodes) and let τ assign those values to the corresponding auxiliary variables in the CNFs.

Applying τ to the derivations above, each of $P1_{\uparrow\tau}, P2_{\uparrow\tau}, P3_{\uparrow\tau}$ is still a valid ER refutation. However we may delete $\llbracket \text{Ref}(a, b) \rrbracket_{\uparrow\tau}$ from the assumptions, since by construction τ satisfies every clause in $\llbracket \text{Ref}(a, b) \rrbracket$, because B is in fact a refutation of A . Furthermore we make the following claims, where all the circuits and derivations asserted to exist are constructible in polynomial time from A and B . For the notation $[\vec{z} = C(\vec{e})]$ see Definition 10.

1. Write $A(\vec{x})$ for the CNF A with the variables renamed to x_1, \dots, x_n . Then there is an ER derivation $A(\vec{x}) \vdash \llbracket \text{Sat}(a, x) \rrbracket_{\uparrow\tau}$.

2. There is a circuit D_θ and auxiliary variables \vec{z}_y and \vec{z}_u such that
 - a. $\llbracket \theta_w(y) \rrbracket_{\uparrow\tau}$ has the form $[\vec{z}_y = D_\theta(\vec{x}, \vec{y})] \wedge z_y^{\text{out}}$
 - b. $\llbracket \theta_w(u) \rrbracket_{\uparrow\tau}$ has the form $[\vec{z}_u = D_\theta(\vec{x}, \vec{u})] \wedge z_u^{\text{out}}$.
 Abusing notation, we may write these as $[\hat{\theta}(\vec{x}, \vec{y})]$ and $[\hat{\theta}(\vec{x}, \vec{u})]$.
3. There is a circuit \hat{t} and auxiliary variables \vec{z}_t such that there is an ER derivation $\llbracket (\vec{y}, \vec{z}_t) = \hat{t}(\vec{x}) \rrbracket \vdash \llbracket y = t_w \rrbracket_{\uparrow\tau}$.
4. There is a circuit \hat{N} and auxiliary variables \vec{z}_N such that there is an ER derivation $\llbracket (\vec{u}, \vec{z}_N) = \hat{N}(\vec{x}, \vec{y}) \rrbracket \vdash \llbracket u = N_w(y) \rrbracket_{\uparrow\tau}$.
5. There is an ER derivation $\llbracket \vec{y} \leq_{\text{lex}} \vec{u} \rrbracket \vdash \llbracket y \leq u \rrbracket_{\uparrow\tau}$.

Here claim 2 is true by construction, and for claim 5 note that the CNF $\llbracket y \leq u \rrbracket$ is not changed after restricting by τ . Otherwise we appeal to the well-known strength and robustness of ER and the fact that we are able to choose how to express Sat, and even $y \leq x$, in PV. We give more details of claim 1 in Appendix A.3.

Combining the derivations provided by the claims with $P1_{\uparrow\tau}, P2_{\uparrow\tau}, P3_{\uparrow\tau}$ and appealing to Lemma 11 we get that the following three ER derivations can be constructed in polynomial time from A and B .

$$\begin{aligned}
 R1 : & \quad A(\vec{x}) \wedge [(\vec{y}, \vec{z}_t) = \hat{t}(\vec{x})] \vdash [\hat{\theta}(\vec{x}, \vec{y})] \\
 R2 : & \quad A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{y})] \wedge [(\vec{u}, \vec{z}_N) = \hat{N}(\vec{x}, \vec{y})] \vdash [\hat{\theta}(\vec{x}, \vec{u})] \\
 R3 : & \quad A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{y})] \wedge [(\vec{u}, \vec{z}_N) = \hat{N}(\vec{x}, \vec{y})] \wedge [\vec{y} \leq_{\text{lex}} \vec{u}] \vdash \perp.
 \end{aligned}$$

For example, for $R2$ we first combine the derivation in claim 1, the first identity in claim 2, and the derivation in claim 4 to derive the LHS of $P2_{\uparrow\tau}$ from the LHS of $R2$ (we may need to rename some variables introduced by the extension rule to avoid clashes when we combine derivations, as in the proof of Lemma 11). By Lemma 11 we can then use $P2_{\uparrow\tau}$ to derive $\llbracket \theta_w(u) \rrbracket_{\uparrow\tau}$, which is precisely the RHS of $R2$ by the second identity in claim 2.

The reader should note that R1–R3 still capture the same idea that we began this section with, but now in a nonuniform version: they constitute a proof in ER that if \vec{y} is a solution of a PLS problem related to Q_w , where w is the instance (A, B, \vec{x}) , then $A(\vec{x})$ is false.

We can now describe an ER-PLS refutation of $A(\vec{x})$, and thus one of A . It will use one application of the ER rule and one of the dominance rule. We begin with $A(\vec{x})$. We then introduce the clauses $[\hat{\theta}(\vec{x}, \vec{y})]$ by the ER rule. This is allowed, because we can obtain them from $A(\vec{x})$ by the following ER derivation: first write down the extension axioms $\llbracket (\vec{y}, \vec{z}_t) = \hat{t}(\vec{x}) \rrbracket$, then use $R1$.

To finish the refutation we use the dominance rule to derive the empty clause. That is, in the rule we take the new clause C to be empty. The other ingredients are as follows.

1. We set $\Gamma := A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{y})]$, so Γ consists of all the clauses we have so far.
2. We let $\vec{v} := \vec{y}, \vec{z}_y, \vec{x}$ list all variables that occur in Γ . Here we deliberately put \vec{y} first so that it is most significant in determining the lexicographic order of assignments to \vec{v} .
3. We set $\Delta := [(\vec{u}, \vec{z}_N) = \hat{N}(\vec{x}, \vec{y})] \wedge [\vec{z}_u = D_\theta(\vec{x}, \vec{u})]$. This is a set of extension axioms over $\vec{v}; \vec{u}, \vec{z}_u, \vec{z}_N$.
4. We set ω to be the substitution which maps each variable in \vec{y}, \vec{z}_y to the corresponding variable in \vec{u}, \vec{z}_u and is the identity everywhere else.

The substitution ω is chosen so that $\Gamma_{\uparrow\omega} = A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{u})]$. Also Δ and ω are chosen so that the range of ω is a subset of the variables appearing in Δ , as required by the rule.

The reader should have in mind the following informal process, as sketched in the introduction. Suppose we have an assignment α to \vec{v} satisfying $\Gamma = A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{y})]$. By $R2$, we can use the circuits described in Δ to extend it to an assignment $\alpha \cup \beta$ to $\vec{v}, \vec{u}, \vec{z}_u, \vec{z}_N$

satisfying $\Gamma \upharpoonright \omega = A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{u})]$, and by R3 the \vec{u} -part of β must be strictly smaller than the \vec{y} -part of α . By the construction of ω , if we let $\alpha' := (\alpha \cup \beta) \circ \omega$ then α' again satisfies $A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{y})]$, with the \vec{y} -part of α' the same as the \vec{u} -part of β . In this way Δ and ω work together to simulate one step in the exponential-time algorithm to solve PLS by producing smaller and smaller “feasible solutions” \vec{y} such that $[\hat{\theta}(\vec{x}, \vec{y})]$. Specifically, Δ computes the next solution and writes it on its new variables \vec{u} , then ω copies the values of \vec{u} back to the old variables, overwriting \vec{y} .

Formally, to complete the proof we need to construct two ER derivations

- (a) $\Gamma \wedge \Delta \vdash \Gamma \upharpoonright \omega$
- (b) $\Gamma \wedge \Delta \wedge [\vec{v} \leq_{\text{lex}} \vec{v} \upharpoonright \omega] \vdash \perp$.

Strictly speaking we should also include $\neg C$ in both sets of assumptions, but since C is the empty clause omitting this makes no difference.

Writing out (a) in more detail, what we need to show is

$$A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{y})] \wedge [(\vec{u}, \vec{z}_N) = \hat{N}(\vec{x}, \vec{y})] \wedge [\vec{z}_u = D_\theta(\vec{x}, \vec{u})] \vdash A(\vec{x}) \wedge [\hat{\theta}(\vec{x}, \vec{u})].$$

If the clauses $[\vec{z}_u = D_\theta(\vec{x}, \vec{u})]$ were not present on the left then R2 would already be a derivation of this. However these clauses are part of $[\hat{\theta}(\vec{x}, \vec{u})]$, so we can use R2 with an appeal to Lemma 12.

For (b), we observe that $[\vec{v} \leq_{\text{lex}} \vec{v} \upharpoonright \omega]$ means precisely $[(\vec{y}, \vec{z}_y, \vec{x}) \leq_{\text{lex}} (\vec{u}, \vec{z}_u, \vec{x})]$, from which formula we can easily derive in ER that $[\vec{y} \leq_{\text{lex}} \vec{u}]$. Hence we can use R3 and Lemma 11. \blacktriangleleft

6 Linear dominance simulates ER-PLS

Recall from Section 2.2 the notation C^* and Γ^* for converting clauses and CNFs into equivalent PB constraints and formulas. We will show that, given a derivation of Δ from Γ in ER-PLS, we can construct in polynomial time a derivation of $(\Delta^*, \emptyset, \top, \emptyset)$ from $(\Gamma^*, \emptyset, \top, \emptyset)$ in the linear dominance system, which implies the simulation. So we must show how to handle the two rules of ER-PLS: the ER rule and the dominance rule.

ER rule. This follows straightforwardly by the well-known simulation of resolution by cutting planes [12] and, for extension steps, using the redundance-based strengthening rule (of the linear dominance system) and standard arguments about how to add extension axioms as redundant clauses, see e.g. [25]. We include a detailed proof in Appendix A.4.

Dominance rule. Suppose we have a CNF Γ and a clause C , both in variables x_1, \dots, x_n , plus a set Δ of extension axioms over $\vec{x}; \vec{y}$, a substitution ω mapping variables \vec{x} to variables $\vec{x} \cup \vec{y}$, and two ER derivations

- (a) $\Gamma \wedge \Delta \wedge \neg C \vdash \Gamma \upharpoonright \omega$
- (b) $\Gamma \wedge \Delta \wedge \neg C \wedge [\vec{x} \leq_{\text{lex}} \vec{x} \upharpoonright \omega] \vdash \perp$.

We will describe a derivation from $(\Gamma^*, \emptyset, \top, \emptyset)$ of $(\Gamma^* \cup \{C^*\}, \emptyset, \top, \emptyset)$.

We can combine the derivations for (a) and (b) above to construct an ER derivation

$$\pi : \Gamma \wedge \Delta \wedge \neg C \vdash \Gamma \upharpoonright \omega \wedge [\vec{x} \upharpoonright \omega <_{\text{lex}} \vec{x}].$$

where $[\vec{x} <_{\text{lex}} \vec{y}]$ represents strict lexicographic ordering – see Proposition 25 in Appendix A.1. Furthermore we may assume that π is actually a *resolution* derivation, that is, that it does not include any applications of the extension rule. This is because we can move all extension axioms introduced by that rule from the body of the derivation to Δ , preserving the order in which they appeared in the derivation. That process turns Δ into a set of extension

axioms over $\vec{x}; \vec{y}, \vec{z}$, where \vec{z} now includes all extension variables that were introduced in the original π , and in particular all auxiliary variables in $[\vec{x}_{\uparrow\omega} <_{\text{lex}} \vec{x}]$.

To build the derivation in the linear dominance system, we first change the ordering from the trivial order \top to the lexicographic order on x_1, \dots, x_n , with the most significant bits first. This can be done using the order change rule; see Section 3.2. So we are now in the configuration $(\Gamma^*, \emptyset, \mathcal{O}_{\prec}, \vec{x})$, where \mathcal{O}_{\prec} is the lexicographic order.

We then derive $(\Gamma^*, \Delta^*, \mathcal{O}_{\prec}, \vec{x})$, where we add each extension axiom in Δ in turn using the redundance-based strengthening rule, in the same way that we handle extension axioms in the ER rule. We must check that we satisfy the order condition for this rule, but this is easy, since the substitutions used do not affect \vec{x} variables, which are the only variables relevant to the ordering. Again we have moved the details to Appendix A.4.

Now we use dominance-based strengthening to derive $(\Gamma^*, \Delta^* \cup \{C^*\}, \mathcal{O}_{\prec}, \vec{x})$. We apply the normal translation from resolution into CP to π to get

$$\Gamma^* \cup \Delta^* \cup \{(\neg C)^*\} \vdash_{\text{CP}} (\Gamma_{\uparrow\omega})^* \cup [\vec{x}_{\uparrow\omega} <_{\text{lex}} \vec{x}]^*.$$

It is easy to construct a short derivation $\neg(C^*) \vdash_{\text{CP}} (\neg C)^*$. We can also construct a derivation $[\vec{x}_{\uparrow\omega} <_{\text{lex}} \vec{x}]^* \vdash_{\text{CP}} \mathcal{O}_{\prec}(\vec{x}_{\uparrow\omega}, \vec{x})$ in polynomial time by Lemma 24 in Appendix A.1, where \mathcal{O}_{\prec} is strict lexicographic ordering written in the natural way using the same multilinear function f as \mathcal{O}_{\prec} ; this is the same as $L_{<}$ from Lemma 24, except that $L_{<}$ is reverse-lexicographic. Moreover, $(\Gamma_{\uparrow\omega})^*$ is the same as $(\Gamma^*)_{\uparrow\omega}$. Thus we have

$$\Gamma^* \cup \Delta^* \cup \{(\neg C^*)\} \vdash_{\text{CP}} (\Gamma^*)_{\uparrow\omega} \cup \mathcal{O}_{\prec}(\vec{x}_{\uparrow\omega}, \vec{x}). \quad (2)$$

Finally, from (2) we can trivially construct a derivation

$$\Gamma^* \cup \Delta^* \cup \{(\neg C^*)\} \cup \mathcal{O}_{\prec}(\vec{x}, \vec{x}_{\uparrow\omega}) \vdash_{\text{CP}} \perp. \quad (3)$$

The derivations (2) and (3) are what we need to apply the dominance-based strengthening rule to derive C^* (after weakening \mathcal{O}_{\prec} in (2) to \mathcal{O}_{\preceq}). This completes the proof.

7 G_1 simulates linear dominance

For this result we will use Theorem 15, which states that for a propositional proof system Q , if S_2^2 proves the CNF-reflection principle for Q , then G_1 simulates Q . We take Q to be the linear dominance system, considered as a system for refuting CNFs. That is, a Q -refutation of a CNF Γ is a linear dominance refutation of Γ^* . Thus for the simulation it is enough to prove in S_2^2 that the existence of such a refutation of Γ^* implies that Γ is unsatisfiable.

We do this by formalizing in S_2^2 as much as we can of the proof of soundness of the dominance system from [4]. We run into a problem when dealing with the dominance rule. To show it is sound, it is enough to show that if a CNF is satisfiable, then it has a least satisfying assignment with respect to the ordering \mathcal{O}_{\prec} . However as far as we know the general statement of this form, that an arbitrary ordering has a least element, is not provable in S_2^2 , and is known to be unprovable if the ordering given is by an oracle [29]. It *is* provable in T_2^2 , and hence in S_2^3 , by a simple inductive argument. By the methods in this section it follows from this that the full dominance system is simulated by G_2 .

To stay inside the strength of S_2^2 we chose to work with the linear dominance system instead since T_2^1 , and hence also S_2^2 , can prove that any nonempty set of strings has a

least element in the lexicographic ordering, which is enough to prove the soundness of the dominance rule restricted to such an ordering.⁷

We use a definition from [4].

► **Definition 19.** A configuration $(\mathcal{C}, \mathcal{D}, \mathcal{O}_{\preceq}, \vec{z})$ is called valid if

1. \mathcal{C} is satisfiable
2. For every total assignment $\alpha \models \mathcal{C}$, there is a total assignment β with $\beta \preceq \alpha$ and $\beta \models \mathcal{C} \cup \mathcal{D}$.

Notice that validity is a Π_2^b condition, and in particular S_2^2 is strong enough to do all the basic reasoning we need about sums and inequalities. Working in S_2^2 , suppose for a contradiction we are given a satisfiable CNF Γ and a linear dominance refutation π of Γ^* . We will use Π_2^b length induction, taking as our inductive hypothesis that the i th configuration in π is valid. The base case is the initial configuration $(\Gamma^*, \emptyset, \top, \emptyset)$, which is valid by assumption (since any assignment satisfying Γ already satisfies Γ^*). On the other hand the final configuration in π is not valid, since in that configuration $\mathcal{C} \cup \mathcal{D}$ is not satisfiable. Therefore to derive a contradiction it is enough to show that every rule preserves validity.

Since the soundness of CP is trivially provable in PV, this is easy for the implicational derivation, transfer and order change rules. For the remaining rules, namely redundancy-based strengthening, deletion and dominance-based strengthening, suppose we are at a valid configuration $(\mathcal{C}, \mathcal{D}, \mathcal{O}_{\preceq}, \vec{z})$.

Redundance-based strengthening rule. We have a substitution ω and we know

$$\mathcal{C} \cup \mathcal{D} \cup \{\neg C\} \models (\mathcal{C} \cup \mathcal{D} \cup \{C\})_{\uparrow\omega} \cup \mathcal{O}_{\preceq}(\vec{z}_{\uparrow\omega}, \vec{z}).$$

Let $\alpha \models \mathcal{C}$. By the inductive hypothesis there is $\beta \preceq \alpha$ such that $\beta \models \mathcal{C} \cup \mathcal{D}$, and we want to find $\beta' \preceq \alpha$ such that $\beta' \models \mathcal{C} \cup \mathcal{D} \cup \{C\}$. If $\beta \models C$ then we set $\beta' = \beta$. Otherwise we set $\beta' = \beta \circ \omega$, and the properties of β' follow from the assumption.

Deletion rule. The interesting case is that we derive $(\mathcal{C}', \mathcal{D}', \mathcal{O}_{\preceq}, \vec{z})$ with $\mathcal{D}' \subseteq \mathcal{D}$ and $\mathcal{C}' = \mathcal{C} \setminus \{C\}$ for some C derivable by the redundancy rule from $(\mathcal{C}, \emptyset, \mathcal{O}_{\preceq}, \vec{z})$. Let $\alpha \models \mathcal{C}'$. If $\alpha \models C$ then there is nothing to show. Otherwise, using the notation of the redundancy rule, we let $\alpha' = \alpha \circ \omega$ and know that $\alpha' \preceq \alpha$ and $\alpha' \models \mathcal{C}$. The inductive hypothesis then gives us $\beta' \preceq \alpha'$ with $\beta' \models \mathcal{C} \cup \mathcal{D}$, so in particular $\beta' \preceq \alpha$ and $\beta' \models \mathcal{C}' \cup \mathcal{D}'$.

For all rules so far, the proof that validity is preserved goes through even in PV. For the last rule we will need to minimize the value of a PV function on a polynomial-time computable set, which can be done in S_2^2 , since it extends T_2^1 (see Section 4.1).

Dominance-based strengthening rule. We derive $(\mathcal{C}, \mathcal{D} \cup \{C\}, \mathcal{O}_{\preceq}, \vec{z})$, and for a given substitution ω we know

$$\begin{aligned} \mathcal{C} \cup \mathcal{D} \cup \{\neg C\} \models \mathcal{C}_{\uparrow\omega} \cup f(\vec{z}_{\uparrow\omega}) \leq f(\vec{z}) \\ \mathcal{C} \cup \mathcal{D} \cup \{\neg C\} \cup \{f(\vec{z}) \leq f(\vec{z}_{\uparrow\omega})\} \models \perp \end{aligned}$$

where f is the linear function defining \mathcal{O}_{\preceq} . Let $\alpha \models \mathcal{C}$. We want to find $\beta' \preceq \alpha$ such that $\beta' \models \mathcal{C} \cup \mathcal{D} \cup \{C\}$. Let S be the set of total assignments \preceq -below α satisfying \mathcal{C} . Let β be a

⁷ Whether the dominance system is strictly stronger than the linear dominance system is unclear. Conceivably to bound the strength of the full dominance rule we could make use of the fact that the ordering is not given by an arbitrary relation, but by a relation which is provably an ordering in CP.

member of S for which $f(\beta)$ is minimal (where $f(\beta)$ stands for f applied to the \vec{z} variables of β). Using the least number principle for Σ_1^b formulas available in S_2^2 , we can find such a β .

By the inductive hypothesis (that is, the validity of $(\mathcal{C}, \mathcal{D}, \mathcal{O}_{\leq}, \vec{z})$) we may assume that $\beta \vDash \mathcal{C} \cup \mathcal{D}$. If $\beta \vDash C$ then we set $\beta' = \beta$. Otherwise let $\beta' = \beta \circ \omega$. By the first entailment in the rule, $\beta' \vDash \mathcal{C}$ and $f(\beta') \leq f(\beta)$, so $\beta' \in S$. Therefore by the minimality of $f(\beta)$ we have $f(\beta) \leq f(\beta')$. But this contradicts the second entailment.

This completes the proof that S_2^2 proves the soundness of the linear dominance system, which is thus simulated by G_1 .

8 Simulations of fragments by ER

8.1 Weak linear dominance

Consider the version of the linear dominance system in which we limit the dominance-based strengthening rule by only allowing it to be applied when the set \mathcal{D} of derived clauses is empty. That is, we replace it with the rule: from $(\mathcal{C}, \emptyset, \mathcal{O}_{\leq}, \vec{z})$ derive $(\mathcal{C} \cup \{C\}, \emptyset, \mathcal{O}_{\leq}, \vec{z})$ if there is a substitution ω and derivations

$$\begin{aligned} \mathcal{C} \cup \{\neg C\} \vdash_{\text{CP}} \mathcal{C}_{\upharpoonright\omega} \cup f(\vec{z}_{\upharpoonright\omega}) \leq f(\vec{z}) \\ \mathcal{C} \cup \{\neg C\} \cup \{f(\vec{z}) \leq f(\vec{z}_{\upharpoonright\omega})\} \vdash_{\text{CP}} \perp. \end{aligned}$$

where f is the linear function defining \mathcal{O}_{\leq} . We shall refer to this system as the *weak linear dominance system*.

► **Proposition 20.** *The weak linear dominance system is simulated by ER.*

To prove the simulation, we use a lemma saying that PV knows that there is a polynomial time function that lets us iterate a substitution m times, when m is given in *binary*.

► **Lemma 21.** *There is a polynomial time function $g(\omega, m)$ which takes as input a substitution ω on variables z_1, \dots, z_n and a number m (coded in binary) and outputs the substitution ω^m . Furthermore this works provably in PV, that is, $\text{PV} \vdash g(\omega, m+1) = g(\omega, m) \circ \omega$.*

Proof. Fix a variable z_i and consider the sequence $z_i, \omega(z_i), \dots, \omega^{2n+2}(z_i)$ as a walk through the space $\text{Lit} \cup \{0, 1\}$. The sequence can be produced by a PV function on input $\langle \omega, i \rangle$. By the pigeonhole principle, which is available in PV here since n is small (polynomial in the length of the input), this sequence must touch some point twice. That is, it consists of a walk of length k_i to some $u \in \text{Lit} \cup \{0, 1\}$, followed by a loop of some size ℓ_i , where $0 \leq k_i \leq 2n+1$ and $1 \leq \ell_i \leq 2n$. Again, the numbers k_i, ℓ_i can be computed by a PV function on input $\langle \omega, i \rangle$.

Thus to compute $\omega^m(z_i)$ for $m > 2n+2$ it is enough to calculate the remainder of $m - k_i$ divided by ℓ_i : namely, $\omega^m(z_i) = \omega^{k_i + ((m - k_i) \bmod \ell_i)}(z_i)$. ◀

Proof of Proposition 20. By Theorem 15, it is enough to show that the soundness of the weak linear dominance system is provable in S_2^1 . So, working in S_2^1 , suppose that a CNF Γ is satisfiable but that Γ^* has a refutation π in the system. We will derive a contradiction. We will use length induction, but with a weaker inductive hypothesis than was used in Section 7 for the soundness of full linear dominance. Namely, we will show that for each configuration $(\mathcal{C}, \mathcal{D}, \mathcal{O}_{\leq}, \vec{z})$ in turn in π , $\mathcal{C} \cup \mathcal{D}$ is satisfiable. Satisfiability is a Σ_1^b property, so this is a form of length induction we can carry out in S_2^1 . It yields a contradiction when we get to the last configuration in π .

The first configuration is satisfiable, by the assumption on Γ . It is easy to see that every rule, other than dominance-based strengthening, preserves satisfiability; in the case of the redundancy-based strengthening rule, this is by the standard argument about composing the current assignment once with ω , if necessary.

So suppose we are dealing with the weak dominance-based strengthening rule. We have an assignment α which satisfies the current configuration $(\mathcal{C}, \emptyset, \mathcal{O}_{\leq}, \vec{z})$, and we want to satisfy $(\mathcal{C} \cup \{C\}, \emptyset, \mathcal{O}_{\leq}, \vec{z})$. We have a substitution ω and derivations

$$\begin{aligned} \mathcal{C} \cup \{-C\} \vdash_{\text{CP}} \mathcal{C}_{\uparrow\omega} \cup f(\vec{z}_{\uparrow\omega}) \leq f(\vec{z}) \\ \mathcal{C} \cup \{-C\} \cup \{f(\vec{z}) \leq f(\vec{z}_{\uparrow\omega})\} \vdash_{\text{CP}} \perp. \end{aligned}$$

for a linear f . Suppose for a contradiction that $\mathcal{C} \cup \{-C\}$ is unsatisfiable. Then, since CP derivations are provably sound (even in PV) we know that for any assignment β , if $\beta \models \mathcal{C}$, then $\beta \circ \omega \models \mathcal{C}$ and $f(\beta \circ \omega) < f(\beta)$.

We may assume without loss of generality that f only takes values between 0 and some upper bound m . Writing α_i for $\alpha \circ \omega^i$, we use induction (rather than length induction) on i to show that for all i we have

$$\alpha_i \models \mathcal{C} \quad \text{and} \quad f(\alpha_i) \leq m - i.$$

By Lemma 21, this is a PV formula, so this induction can be carried out in S_2^1 (if the formula were Σ_1^b , we would only be able to use length induction). The base case $i = 0$ is true by the assumptions about α and f , and the inductive step follows from the discussion in the previous paragraphs. We conclude that $f(\alpha_{m+1}) \leq -1$, which is impossible. \blacktriangleleft

8.2 Symmetry breaking in ER

Let us define a proof system Q , which we could call ER plus *static symmetry breaking*. A refutation of a CNF Γ in Q consists of an initial step, in which we list a sequence of symmetries $\omega_1, \dots, \omega_k$ of Γ and write down the corresponding lex-leader constraints (where for each constraint we use fresh auxiliary variables). This is followed by an ER refutation of Γ augmented by these constraints, that is, of $\Gamma' := \Gamma \wedge \bigwedge_i [\vec{z} \leq_{\text{lex}} \vec{z}_{\uparrow\omega_i}]$.

For $k \in \mathbb{N}$ we define Q_k to be Q limited to only adding axioms for k symmetries.

► **Proposition 22.** *The full system Q is sound, and is simulated by G_1 .*

Proof. We repeat the proof of Proposition 2, except this time we fill in some details. To prove soundness, it is enough to show that, supposing Γ is satisfiable, Γ' is satisfiable as well. Let α be a lexicographically minimal assignment to the \vec{z} -variables satisfying Γ . We claim that an extension of α satisfies Γ' . To see this, let ω_i be any symmetry from our list. Then $\alpha \models \Gamma$ implies $\alpha \models \Gamma_{\uparrow\omega_i}$, and thus $\alpha \circ \omega_i \models \Gamma$. By minimality of α we have $\alpha \leq_{\text{lex}} \alpha \circ \omega_i$, and thus, extending α to β which satisfies the extension axioms in the definition of \leq_{lex} , we have that β satisfies the symmetry-breaking axiom $[\vec{z} \leq_{\text{lex}} \vec{z}_{\uparrow\omega_i}]$. In this way we can simultaneously satisfy such axioms for all i , by the assumption that auxiliary variables are disjoint.

For the simulation by G_1 , it is enough to observe that this argument can be cast as a proof of the CNF-reflection principle for Q and carried out in T_2^1 . Then we can appeal to Theorem 15. \blacktriangleleft

The converse direction is presumably false:

► **Proposition 23.** *G_1 is not simulated by Q , assuming G_1 is not simulated by ER.*

Proof. Let Γ_n be a family of CNFs which have polynomial-sized refutations in G_1 but require superpolynomial size in ER. Then it is easy to construct a polynomial-sized CNF A_n such that $\Gamma_n \cup A_n$ has no symmetries; assuming Γ_n has variables x_1, \dots, x_m , a convenient example consists of clauses $x_i \vee y_1 \vee \dots \vee y_i$ for each i , where y_1, \dots, y_m are new variables. Then G_1 refutations of Γ_n still work for $\Gamma_n \cup A_n$ (we may need to add one more weakening step). On the other hand, if π is any Q refutation of $\Gamma_n \cup A_n$, then it must be just an ER refutation, and we can turn it into an ER refutation of Γ_n by applying the restriction which sets every y_i variable to 1. Thus π must have superpolynomial size. \blacktriangleleft

We can now prove Theorem 5 from the introduction, that Q_1 is simulated by ER.

Proof of Theorem 5. We will show that the soundness of Q_1 is provable in S_2^1 . The result then follows by Theorem 15. Let Γ be a CNF and let ω be a symmetry of Γ . Let $\Gamma' := \Gamma \wedge [\vec{z} \leq_{\text{lex}} \vec{z}_{\uparrow\omega}]$ and suppose we are given an ER refutation of Γ' . We will show, with a proof formalizable in S_2^1 , that if Γ is satisfiable then so is Γ' . We can then derive a contradiction, since S_2^1 proves the soundness of ER (see Theorem 14).

Working in S_2^1 , suppose $\alpha \models \Gamma$. As in the proof of Proposition 20 we write α_i for $\alpha \circ \omega^i$, and use the fact that by Lemma 21 this can be computed in polynomial time. Suppose for a contradiction that $\Gamma \wedge [\vec{z} \leq_{\text{lex}} \vec{z}_{\uparrow\omega}]$ is unsatisfiable. It follows that for any assignment β , if $\beta \models \Gamma$ then $\beta \circ \omega <_{\text{lex}} \beta$. On the other hand, if $\beta \models \Gamma$ then we already know $\beta \circ \omega \models \Gamma$, since $\Gamma_{\uparrow\omega} = \Gamma$. Assuming that there are n many z -variables we have $\alpha \leq_{\text{lex}} 2^n - 1$, where we identify $2^n - 1$ with a string of 1s of length n . Thus we can reach a contradiction by a similar induction as in the proof of Proposition 20, showing inductively that for each i we have $\alpha_i \models \Gamma$ and $\alpha_i \leq_{\text{lex}} 2^n - i$. \blacktriangleleft

This proof breaks down immediately even for Q_2 , since we do not have any equivalent of Lemma 21 for arbitrary compositions of two substitutions.

We briefly discuss how one could directly construct an ER refutation from a Q_1 refutation, without going through bounded arithmetic and Theorem 5. The main task is to construct a circuit C which, when given an assignment α such that $\alpha \models \Gamma$, outputs an assignment β such that $\beta \models \Gamma \wedge [\vec{z} \leq_{\text{lex}} \vec{z}_{\uparrow\omega}]$. Furthermore this property of C must be provable in ER, in the sense that we have an ER derivation $\Gamma(\vec{x}) \wedge [\vec{z} = C(\vec{x})] \vdash \Gamma(\vec{z}) \wedge [\vec{z} \leq_{\text{lex}} \vec{z}_{\uparrow\omega}]$ (where we are suppressing auxiliary variables in C and \leq_{lex}). We will just describe C .

We use a subcircuit which takes input \vec{z}, i and computes $\alpha_i := \vec{z}_{\uparrow\omega^i}$ using the algorithm for g in Lemma 21. The circuit C finds i such that the two conditions $\alpha_i \models \Gamma$ and $\alpha_i \leq_{\text{lex}} 2^n - i$ hold for i , but one of them fails for $i + 1$, and outputs α_i . Such an i can be found by binary search, since both conditions hold for $i = 0$ and the second one must fail for $i = 2^n + 1$. Since $\alpha_i \models \Gamma$ and $\alpha_{i+1} = \alpha_i \circ \omega$, we have that $\alpha_{i+1} \models \Gamma$ as ω is a symmetry. We conclude that the second condition fails and $\alpha_{i+1} >_{\text{lex}} 2^n - i - 1$. Thus $\alpha_{i+1} \geq_{\text{lex}} \alpha_i$, meaning that $\alpha_i \leq_{\text{lex}} \alpha_i \circ \omega$ as required.

Acknowledgements. We are grateful to Jakob Nordström for introducing us to this topic and answering our questions about it, and to Sam Buss and Vijay Ganesh for other helpful discussions.

References

- 1 P. Beame, H. Kautz, and A. Sabharwal. Towards understanding and harnessing the potential of clause learning. *Journal of Artificial Intelligence Research*, 22:319–351, 2004. doi:10.1613/jair.1410.

- 2 A. Beckmann and S. R. Buss. Polynomial local search in the polynomial hierarchy and witnessing in fragments of bounded arithmetic. *Journal of Mathematical Logic*, 9(1):103–138, 2009. doi:10.1142/S0219061309000847.
- 3 O. Beyersdorff and J. Pich. Understanding Gentzen and Frege systems for QBF. In *Proceedings of the Annual Symposium on Logic in Computer Science (LICS '16)*, pages 146–155, 2016. doi:10.1145/2933575.2933597.
- 4 B. Bogaerts, S. Gocht, C. McCreesh, and J. Nordström. Certified dominance and symmetry breaking for combinatorial optimisation. *Journal of Artificial Intelligence Research*, 77:1539–1589, 2023. doi:10.1613/jair.1.14296.
- 5 S. Buss and N. Thapen. DRAT and propagation redundancy proofs without new variables. *Logical Methods in Computer Science*, 17, 2021. URL: <https://lmcs.episciences.org/7400>.
- 6 S. R. Buss. *Bounded Arithmetic*. Bibliopolis, 1986.
- 7 S. R. Buss. Axiomatizations and conservation results for fragments of bounded arithmetic. In *Logic and Computation*, volume 106 of *Contemporary Mathematics*, pages 57–84. ACM, 1990.
- 8 S. R. Buss and J. Krajíček. An application of Boolean complexity to separation problems in bounded arithmetic. *Proceedings of the London Mathematical Society*, 69:1–21, 1994. doi:10.1112/plms/s3-69.1.1.
- 9 L. Chew and M. J. H. Heule. Relating existing powerful proof systems for QBF. In *International Conference on Theory and Applications of Satisfiability Testing (SAT '22)*, pages 10:1–10:22, 2022. doi:10.4230/LIPIcs.SAT.2022.10.
- 10 V. Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Mathematics*, 4:305–337, 1973. doi:10.1016/0012-365X(73)90167-2.
- 11 S. A. Cook. Feasibly constructive proofs and the propositional calculus (preliminary version). In *Annual ACM Symposium on Theory of Computing (STOC '75)*, pages 83–97. 1975. doi:10.1145/800116.803756.
- 12 W. Cook, C. R. Coullard, and G. Turán. On the complexity of cutting-plane proofs. *Discrete Applied Mathematics*, 18(1):25–38, 1987. doi:10.1016/0166-218X(87)90039-4.
- 13 J. Crawford, M. Ginsberg, E. Luks, and A. Roy. Symmetry-breaking predicates for search problems. In *Principles of Knowledge Representation and Reasoning (KR '96)*, pages 148–159, 1996.
- 14 M. Davis, G. Logemann, and D. Loveland. A machine program for theorem proving. *Communications of the ACM*, 5:394–397, 1962. doi:10.1145/368273.368557.
- 15 J. Elffers, S. Gocht, C. McCreesh, and J. Nordström. Justifying all differences using pseudo-Boolean reasoning. In *AAAI Conference on Artificial Intelligence*, volume 34, pages 1486–1494, 2020. doi:10.1609/aaai.v34i02.5507.
- 16 P. Hájek and P. Pudlák. *Metamathematics of First-Order Arithmetic*. Springer-Verlag, 1993. doi:10.1007/978-3-662-22156-3.
- 17 M. J. H. Heule, W. A. Hunt, Jr., and N. Wetzler. Expressing symmetry breaking in DRAT proofs. In *International Conference on Automated Deduction*, pages 591–606. Springer, 2015. doi:10.1007/978-3-319-21401-6_40.
- 18 M. J. H. Heule, B. Kiesl, and A. Biere. Strong extension-free proof systems. *Journal of Automated Reasoning*, 64(3):533–554, 2020. doi:10.1007/s10817-019-09516-0.
- 19 E. Jeřábek. Dual weak pigeonhole principle, Boolean complexity, and derandomization. *Annals of Pure and Applied Logic*, 129:1–37, 2004. doi:10.1016/j.apal.2003.12.003.
- 20 E. Jeřábek. The strength of sharply bounded induction. *Mathematical Logic Quarterly*, 52:613–624, 2006. doi:10.1002/malq.200610019.
- 21 D. S. Johnson, C. H. Papadimitriou, and M. Yannakakis. How easy is local search? *Journal of Computer and System Sciences*, 37:79–100, 1988. doi:10.1016/0022-0000(88)90046-3.
- 22 B. Kiesl, A. Rebola-Pardo, and M. J. H. Heule. Extended resolution simulates DRAT. In *International Joint Conference on Automated Reasoning (IJCAR '18)*, pages 516–531, 2018. doi:10.1007/978-3-319-94205-6_34.

- 23 J. Krajčček. *Proof complexity*, volume 170 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2019.
- 24 J. Krajčček and P. Pudlák. Quantified propositional calculi and fragments of bounded arithmetic. *Z. Math. Logik Grundlag. Math.*, 36(1):29–46, 1990. doi:10.1002/ma1q.19900360106.
- 25 O. Kullmann. On a generalization of extended resolution. *Discrete Applied Mathematics*, 96:149–176, 1999. doi:10.1016/S0166-218X(99)00037-2.
- 26 J. P. Marques-Silva and K. A. Sakallah. GRASP—a search algorithm for propositional satisfiability. *IEEE Transactions on Computers*, 48:506–521, 1999. doi:10.1109/12.769433.
- 27 J. B. Paris and A. J. Wilkie. Counting problems in bounded arithmetic. In *Methods in Mathematical Logic*, volume 1130 of *Lecture Notes in Mathematics*, pages 317–340. Springer-Verlag, 1985. doi:10.1007/BFb0075316.
- 28 A. A. Razborov. Propositional proof complexity. In *European Congress of Mathematics 2021*, pages 439–464. EMS Press, 2023.
- 29 S. Riis. Finitisation in bounded arithmetic. *BRICS Report Series*, (23), 1994.
- 30 N. Wetzler, M. J. H. Heule, and W. A. Hunt, Jr. DRAT-trim: Efficient checking and trimming using expressive clausal proofs. In *International Conference on Theory and Applications of Satisfiability Testing (SAT '14)*, pages 422–429, 2014. doi:10.1007/978-3-319-09284-3_31.
- 31 A. J. Wilkie and J. B. Paris. On the scheme of induction for bounded arithmetic formulas. *Annals of Pure and Applied Logic*, 35:261–302, 1987. doi:10.1016/0168-0072(87)90066-2.

A Postponed technical material

A.1 How to formalize ordering (from Section 3.3)

We define CNFs $[\vec{x} <_{\text{lex}} \vec{y}]$ and $[\vec{x} \leq_{\text{lex}} \vec{y}]$ expressing the lexicographic ordering. We will need these to be compatible with how we reason about orderings in the pseudo-Boolean setting.

Let r be the arity of \vec{x} and \vec{y} . For our proof system we will want to compare \vec{x} and \vec{y} with the most significant bits first, but just in this section we prefer to define a CNF for reverse lexicographic order, since it gives more readable notation when we spell out the details.

We work with variables $x_1, \dots, x_r, y_1, \dots, y_r, z_1, \dots, z_{r+1}$ and c_2, \dots, c_{r+1} , representing numbers $x, y < 2^r$, a number $z < 2^{r+1}$ and a string of “carry” bits c . We first write a CNF Δ expressing “ $x + z = y + 2^r$ ”. This is the conjunction of, for each $i = 1, \dots, r + 1$, a CNF expressing

$$x_i + z_i + c_i = y_i + 2c_{i+1} \tag{4}$$

where at the “boundaries” we replace x_{r+1} , c_1 and c_{r+2} with the constant 0 and replace y_{r+1} with the constant 1; this last step corresponds to adding 2^r to the number y . Each expression of the form (4) can be written as a CNF. Moreover, since the number z is uniquely defined, we can introduce extension variables needed for intermediate steps in the verification of (4) in such a way that the whole formula Δ becomes a set of extension axioms defining \vec{z} , \vec{c} and the intermediate extension variables from \vec{x} and \vec{y} .

We write a CNF $P_{<}$ expressing that $\vec{x} < \vec{y}$ by saying that “ $x + z = y + 2^r$ with $z > 2^r$ ”, that is, $\Delta \wedge z_{r+1} \wedge \bigvee_{i=1}^r z_i$.

Now let $L_{<}$ be the natural way to expressing the strict lexicographic ordering by a single PB constraint. That is, $L_{<}$ is the constraint $\sum_{i=1}^r 2^{i-1}x_i < \sum_{i=1}^r 2^{i-1}y_i$. The definition of $P_{<}$ was chosen to give a simple proof of the following lemma.

► **Lemma 24.** *There is a polynomial-time procedure that on input 1^r produces CP derivations $P_{<}(\vec{x}, \vec{y})^* \vdash_{\text{CP}} L_{<}(\vec{x}, \vec{y})$.*

Proof. We will write X_j, Y_j, Z_j for the sums $\sum_{i=1}^j 2^{i-1} x_i$ etc. We will inductively construct derivations $\Delta^* \vdash_{\text{CP}} X_j + Z_j = Y_j + 2^j c_{j+1}$, for $j = 1, \dots, r$, where Δ is the CNF defined above. Here and below the equality $=$ is shorthand for the pair of inequalities \leq and \geq .

For each such j , the CNF Δ contains clauses together expressing $x_i + z_i + c_i = y_i + 2c_{i+1}$, after the substitution $c_1 \rightarrow 0$. Thus, by the implicational completeness of cutting planes [10], we can produce constant-size derivations of this equality (as a pair of PB constraints) from Δ^* . For $i = 1$ this immediately gives the base case of the induction. Now, in the inductive step, suppose $1 \leq j < r$ and we have $X_j + Z_j = Y_j + 2^j c_{j+1}$. Then we add $x_{j+1} + z_{j+1} = y_{j+1} + 2c_{j+2} - c_{j+1}$ multiplied by 2^j to obtain $X_{j+1} + Z_{j+1} = Y_{j+1} + 2^{j+1} c_{j+2}$.

When the inductive construction is completed, we have derived $X_r + Z_r = Y_r + 2^r c_{r+1}$. The final constraints in Δ^* can be used to produce a constant-size derivation of the equality $z_{r+1} + c_{r+1} = 1$. So we can add $2^r z_{r+1} = 2^r - 2^r c_{r+1}$ to get $X_r + Z_{r+1} = Y_r + 2^r$. From the assumptions $z_{r+1} = 1$ and $\bigvee_{i=1}^r z_r$ in $P_{<}$ we can derive $Z_{r+1} > 2^r$, and together these give us $L_{<}$. \blacktriangleleft

We define $P_{\leq}(\vec{x}, \vec{y})$ as the negation of $P_{<}(\vec{y}, \vec{x})$. Precisely, if $P_{<}(\vec{y}, \vec{x})$ has the form $\Delta \wedge z_{r+1} \wedge \bigvee_{i=1}^r z_i$ we take $P_{\leq}(\vec{x}, \vec{y})$ to be the CNF $\Delta \wedge [w \leftrightarrow (z_{r+1} \wedge \bigvee_{i=1}^r z_i)] \wedge \neg w$, where the expression $[w \leftrightarrow \dots]$ is shorthand for a series of extension axioms defining a new variable w . Finally we set the formula $[\vec{x} \leq_{\text{lex}} \vec{y}]$ to be $P_{\leq}(x_r, \dots, x_1, y_r, \dots, y_1)$ and set $[\vec{x} <_{\text{lex}} \vec{y}]$ to be $P_{<}(x_r, \dots, x_1, y_r, \dots, y_1)$. The next proposition is proved in a similar way to the proof of Proposition 17 below.

► **Proposition 25.** *Given an ER derivation $\Gamma \wedge [\vec{x} \leq_{\text{lex}} \vec{y}] \vdash \perp$ we can construct in polynomial time an ER derivation $\Gamma \vdash [\vec{y} <_{\text{lex}} \vec{x}]$.*

A.2 Propositional translations (from Section 4.2)

We give the proof of Proposition 17. That is, suppose PV proves a sentence

$$\forall \vec{x}, \varphi_1(\vec{x}) \wedge \dots \wedge \varphi_r(\vec{x}) \rightarrow \theta(\vec{x}), \quad (5)$$

where $\varphi_1, \dots, \varphi_r, \theta$ are quantifier-free. Then for any assignment \vec{k} of bit-lengths to the variables \vec{x} , we can construct in time polynomial in \vec{k} an ER derivation

$$\llbracket \varphi_1(\vec{x}) \rrbracket_{\vec{k}} \wedge \dots \wedge \llbracket \varphi_r(\vec{x}) \rrbracket_{\vec{k}} \vdash \llbracket \theta(\vec{x}) \rrbracket_{\vec{k}}.$$

Proof of Proposition 17. By the usual form of the translation from PV into ER, as presented for example in [23, Theorem 12.4.2], it follows from the provability of (5) that in time polynomial in \vec{k} we can build an ER *refutation*

$$\pi : \llbracket \varphi_1(\vec{x}) \rrbracket_{\vec{k}} \wedge \dots \wedge \llbracket \varphi_r(\vec{x}) \rrbracket_{\vec{k}} \wedge \llbracket \neg \theta(\vec{x}) \rrbracket_{\vec{k}} \vdash \perp.$$

The details of how to transform π into an ER derivation $\pi' : \Gamma \vdash \llbracket \theta(\vec{x}) \rrbracket_{\vec{k}}$, where we write Γ for $\llbracket \varphi_1(\vec{x}) \rrbracket_{\vec{k}} \wedge \dots \wedge \llbracket \varphi_r(\vec{x}) \rrbracket_{\vec{k}}$, are routine but messy. The translation $\llbracket \theta(\vec{x}) \rrbracket_{\vec{k}}$ has the form $\Delta \wedge z^{\text{out}}$ where Δ is a set of extension axioms over $\vec{x}_1, \dots, \vec{x}_\ell; \vec{z}$ for some variables \vec{z} (which include z^{out}). Modulo trivial modifications, $\llbracket \neg \theta(\vec{x}) \rrbracket_{\vec{k}}$ can be taken to be $\Delta \wedge \neg z^{\text{out}}$. Note that by our conventions on variables, the variables \vec{z} do not appear in Γ . Thus we can begin π' by deriving all of Δ from Γ using the extension rule, and all that remains is to derive z^{out} .

To do this, we work through the refutation $\pi : \Gamma \wedge \Delta \wedge \neg z^{\text{out}} \vdash \perp$ and, for each clause C in π except for the initial clause $\neg z^{\text{out}}$, we derive $C \vee z^{\text{out}}$. If C is a clause of Γ or Δ , we derive $C \vee z^{\text{out}}$ from C by weakening (notice that z^{out} is at this point an “old” variable,

since it appears in Δ); if C is derived in π by resolution on a variable other than z^{out} or by weakening, we derive $C \vee z^{\text{out}}$ by the same inference; if C is derived in π from $B \vee z^{\text{out}}$ and $A \vee \neg z^{\text{out}}$ by resolution on z^{out} , we instead derive $C \vee z^{\text{out}}$ by weakening from $B \vee z^{\text{out}}$ (again, this does not introduce any new variables). If C is part of an extension rule in π , we introduce the same extension axioms and then derive $C \vee z^{\text{out}}$ by weakening, noting that the same variables are new at this point in π and in π' . Thus we finally derive z^{out} , where in π we derived \perp . \blacktriangleleft

A.3 How to formalize satisfiability (from Section 5)

► **Proposition 26.** *There is a polytime procedure that, given a CNF A in n variables, produces an ER derivation $A(\vec{x}) \vdash \llbracket \text{Sat}(a, x) \rrbracket_{\uparrow \tau}$, where the bit-lengths of \vec{a}, \vec{x} in $\llbracket \text{Sat}(a, x) \rrbracket$ are the size of A and n , respectively, τ substitutes the bits of A for \vec{a} , and $A(\vec{x})$ is A with variables renamed to \vec{x} .*

Proof. We may assume that A is already in the variables $\vec{x} = x_1, \dots, x_n$. The details of the procedure will depend on how exactly $\text{Sat}(a, x)$ is formalized by a PV formula, but any two reasonable formalizations will be provably equivalent in PV, so we may focus on one convenient formalization and then invoke Proposition 17 and Lemma 11.

Suppose, for example, that $\text{Sat}(a, x)$ is given as a straightforward procedure that considers each clause of a in succession and checks for each literal whether it belongs to the clause and whether it is satisfied under x ; afterwards, it checks if at least one literal was satisfied in each clause. Let m be the number of clauses in A . Then we may assume that $\llbracket \text{Sat}(a, x) \rrbracket_{\uparrow \tau}$ contains the following auxiliary variables: for each $j = 1, \dots, m$ and each $i = 1, \dots, 2n$ (this being the number of literals in n variables), a variable z_i^j with the intuitive meaning that one of the first i literals appears in the j -th clause and happens to be satisfied; and for each $j = 1, \dots, m$, a variable w_j with the intuitive meaning that each of the first j clauses contains a satisfied literal. We may also assume that $\llbracket \text{Sat}(a, x) \rrbracket_{\uparrow \tau}$ consists of the following conjuncts:

- for $j = 1, \dots, m$ and $i = 1, \dots, 2n$, a constant-size set of clauses logically equivalent to either $z_i^j \leftrightarrow z_{i-1}^j$, if the i -th literal does not appear in the j -th clause, or $z_i^j \leftrightarrow z_{i-1}^j \vee (\neg)x_k$, if the i -th literal appears in the j -th clause and happens to be $(\neg)x_k$ (here z_0^j is the constant 0),
- for $j = 1, \dots, m$, a constant-size set of clauses logically equivalent to with $w_j \leftrightarrow w_{j-1} \wedge z_{2n}^j$ (here w_0 is the constant 1),
- w_m .

The derivation of this from A is the obvious one: first, for each $j = 1, \dots, m$, introduce each $z_i^j \leftrightarrow z_{i-1}^j$ resp. $z_i^j \leftrightarrow z_{i-1}^j \vee (\neg)x_k$ as extension axioms and perform a series of resolutions with the j -th clause of A , eventually obtaining z_{2n}^j . Then, again for each $j = 1, \dots, m$, introduce $w_j \leftarrow w_{j-1} \wedge z_{2n}^j$ as an extension axiom and resolve $w_j \leftarrow w_{j-1} \wedge z_{2n}^j$ with z_{2n}^j and w_{j-1} to obtain w_j . \blacktriangleleft

A.4 Simulating ER-PLS by linear dominance (from Section 6)

We give some details omitted from Section 6. We first show how to simulate the ER rule of ER-PLS in the linear dominance system, and then we fill a gap in how we handled the dominance rule.

ER rule. Suppose $\Gamma \wedge \Delta$ is derived from Γ by the ER rule. Then there is an ER derivation π which begins with Γ and includes every clause in Δ . It is enough to show that from $(\Gamma^*, \emptyset, \top, \emptyset)$ we can derive $(\Gamma^*, \pi^*, \top, \emptyset)$, where we write π^* for the set of translations of clauses appearing

in π . This is because we can then copy all of Δ^* from the derived constraints π^* into the core constraints Γ^* using the transfer rule of the dominance system, to get $((\Gamma \wedge \Delta)^*, \pi^*, \top, \emptyset)$, and finally reset the derived constraints to \emptyset using the deletion rule, to get $((\Gamma \wedge \Delta)^*, \emptyset, \top, \emptyset)$ as required.

Treating each clause in π in turn, either it is an initial clause from Γ ; or it was derived from earlier clauses by resolution or weakening; or it is a clause introduced by the extension rule. The first two cases are easily dealt with by the implicational derivation rule (of the dominance system), plus the well-known simulation of resolution by CP [12]. The remaining case is the extension rule. We use the standard arguments for showing that we can simulate this rule in a system based on adding certain redundant clauses, see e.g. [25].

An extension axiom in π has the form of three clauses $\neg u \vee \neg v \vee y$, $\neg y \vee u$ and $\neg y \vee v$, which we translate into three PB constraints:

$$\begin{aligned} A: & (1 - u) + (1 - v) + y \geq 1 \\ B: & (1 - y) + u \geq 1 \\ C: & (1 - y) + v \geq 1. \end{aligned}$$

We show that these can be derived from any set \mathcal{E} of PB constraints which do not mention the variable y , by three applications of the redundance-based strengthening rule of the dominance system (we may ignore the part of the rule having to do with the ordering, as we are using the trivial ordering \top). That is, we must find substitutions σ, τ, ω and derivations $\mathcal{E} \cup \{\neg A\} \vdash_{\text{CP}} (\mathcal{E} \cup \{A\}) \upharpoonright_{\sigma}$, $\mathcal{E} \cup \{A, \neg B\} \vdash_{\text{CP}} (\mathcal{E} \cup \{A, B\}) \upharpoonright_{\tau}$ and $\mathcal{E} \cup \{A, B, \neg C\} \vdash_{\text{CP}} (\mathcal{E} \cup \{A, B, C\}) \upharpoonright_{\omega}$.

Let σ map $y \mapsto u$ and do nothing else. Then $\mathcal{E} \upharpoonright_{\sigma} = \mathcal{E}$ and $A \upharpoonright_{\sigma}$ is $2 - v \geq 1$, which is the Boolean axiom $v \leq 1$. Thus we have the first CP derivation. For the second CP derivation we can set $\tau = \sigma$ again, since $B \upharpoonright_{\sigma}$ is just $1 \geq 1$.

Now let ω map $y \mapsto 0$ and do nothing else. Then $\mathcal{E} \upharpoonright_{\omega} = \mathcal{E}$ and we have

$$\begin{aligned} A \upharpoonright_{\omega} \text{ is } (1 - u) + (1 - v) \geq 1 & & C \upharpoonright_{\omega} \text{ is } v \geq 0 \\ B \upharpoonright_{\omega} \text{ is } u \geq 0 & & \neg C \text{ is } 1 - y + v \leq 0 \end{aligned}$$

Thus $B \upharpoonright_{\omega}$ and $C \upharpoonright_{\omega}$ are trivially derivable. For $A \upharpoonright_{\omega}$, we can rearrange it as $u + v \leq 1$, which we can derive by starting with $\neg C$ and adding axioms $y \leq 1$ and $u \leq 1$. Thus we have the third CP derivation.

We also allow extension axioms expressing $y \leftrightarrow u$ or $y \leftrightarrow 0$ or $y \leftrightarrow 1$. We can derive the $*$ translations of these in a similar way to the first two derivations above, by simply setting y to be u or the desired value.

Dominance rule. We fill in a step that was omitted in Section 6. Suppose we are in a configuration $(\Gamma^*, \emptyset, \mathcal{O}_{\preceq}, \vec{x})$, where \mathcal{O}_{\preceq} is the lexicographic order. We must derive $(\Gamma^*, \Delta^*, \mathcal{O}_{\preceq}, \vec{x})$, where Δ is a set of extension axioms over $\vec{x}; \vec{y}$.

We do this by adding each extension axiom in Δ in turn using the redundance-based strengthening rule, in the same way that we handled introducing extension axioms in the case of the ER rule above. However this time we must check that we satisfy the order condition for this rule – this did not matter before, as there we had the trivial order.

For example (using the same notation \mathcal{E}, A, σ as above), in the derivation needed to introduce A , the formal requirement is to show

$$\mathcal{E} \cup \{\neg A\} \vdash_{\text{CP}} (\mathcal{E} \cup \{A\}) \upharpoonright_{\sigma} \cup \mathcal{O}_{\preceq}(\vec{x} \upharpoonright_{\sigma}, \vec{x}).$$

However this is trivial as σ does not affect the variables \vec{x} , which are explicitly the only variables compared in the ordering \mathcal{O}_{\preceq} . The same goes for the other substitutions used above.