Logical Methods in Computer Science
Volume 22, Issue 2, 2026, pp. 14:1–14:8
https://lmcs.episciences.org/

Submitted Jul. 03, 2025
Published May 06, 2026

ON NP ∩ coNP PROOF COMPLEXITY GENERATORS

JAN KRAJÍČEK

Faculty of Mathematics and Physics, Charles University, Prague
e-mail address: jan.krajicek@protonmail.com

Abstract. Motivated by the theory of proof complexity generators we consider the following Σ_2^p search problem DD_P determined by a propositional proof system P: given a P-proof π of a disjunction $\bigvee_i \alpha_i$, no two α_i having an atom in common, find i such that $\alpha_i \in \mathrm{TAUT}$.

We formulate a hypothesis (ST) that for some strong proof system P the problem DD_P is not solvable in the student-teacher model with a p-time student and a constant number of rounds. The hypothesis follows from the existence of hard one-way permutations.

We prove, using a model-theoretic assumption, that (ST) implies NP ≠ *co*NP. The assumption concerns the existence of extensions of models of a bounded arithmetic theory and it is open at present if it holds.

Introduction

Proof complexity generators are some maps $g : \{0,1\}^* \to \{0,1\}^*$ with $|g(x)|$ determined by and larger than $|x|$. Their purpose is to "generate" hard tautologies: propositional formulas $\tau(g)_b$ expressing that $b \notin Rng(g)$. To be able to express this propositionally we need that g is p-time or more generally NP ∩ coNP (possibly non-uniform).

For NP ∩ coNP maps[1] g sending n bits to $m > n$ bits and $b \in \{0,1\}^m$, the statement $b \notin Rng(g)$ has the form

$$\forall x(|x| = n)\exists i < m \forall y_i \ A_b(x, i, y_i) \tag{0.1}$$

with A_b an open formula in the language L_{PV} having a function symbol for every p-time clocked algorithm, and expressing that y_i does not witness that the i-th bit of $g(x)$ is b_i, the i-th bit of b.

The provability of such a statement in the true universal L_{PV}-theory T_{PV} can be analyzed using the interactive student-teacher model (abbr. S-T) of [KPS90]. In this model S (the student) upon receiving $x := a$ proposes its first candidate solution i_1. If it is correct T (the teacher) will acknowledge this and the computation stops with i_1 as its output. Otherwise she will send to S a counter-example, some w_1 such that $\neg A_b(a, i_1, w_1)$. In that case the computation enters the second round with S proposing a second candidate solution i_2 and T either accepting it or returning a counter-example w_2. The computation goes for a number of rounds until a solution is found.

Key words and phrases: proof complexity, bounded arithmetic, student-teacher computations, search problems.

[1]These are maps for which the binary predicate *the i-th bit of* $g(x)$ is in NP ∩ coNP.

We are interested in the case where the number of rounds is bounded by a constant. The S-T model makes sense for general total search problems where the condition for the acceptance of a solution, here $\forall y \neg A_b(x, i, y)$, starts with a universal quantifier. Here all such search problems will be Σ_2^p and following [Kra25] we denote by

$$\mathrm{ST}[\mathrm{F}, t(n)]$$

the class of total Σ_2^p search problems such that there is a student whose moves are computed by an algorithm from F that solves the problems in $t(n)$ rounds interacting with any teacher.

There are two candidate $\mathrm{NP} \cap \mathrm{coNP}$ proof complexity generators: an NW generator and a gadget generator based on a non-deterministic circuit as the gadget, cf. [Kra25]. An analysis using the S-T model for (0.1) for the NW-generator was performed in [Kra11] and it was shown that (0.1) cannot be witnessed in $\mathrm{ST}[\mathrm{P}/poly, O(1)]$ if a one-way permutation hard for p-size circuits exists. The gadget generator was analyzed in [Kra25] and a similar result was shown assuming a certain plausibly looking computational hardness hypothesis (K), cf. [Kra25, Sec.8.6].

These first-order constructions gave independence results for the theory T_{PV}. There is a close relation between first-order theories and (tacitly propositional) proof systems, based on the notion of propositional translations (for theories in L_{PV} this goes back to [Coo75, KP89]). It is also a classic fact that lengths-of-proofs lower bounds for formulas γ are essentially equivalent to constructing models of a suitable theory where γ can be falsified.

When trying to get from the independence results for the two $\mathrm{NP} \cap \mathrm{coNP}$ generators lengths-of-proofs lower bounds for the corresponding τ-formulas one runs into the issue that the propositional translation is not entirely faithful in the following sense (this is discussed in detail as a "missing reflection" in [Kra11]).

The $\tau(g)$ formulas for an $\mathrm{NP} \cap \mathrm{coNP}$-map g have the form

$$\gamma(p, q) \ := \ \bigvee_i \beta_i(p, q^i) \tag{0.2}$$

with tuples of atoms $p, q = (q^i)_i$ and tuples q^i disjoint. A falsifying assignment to γ exists in a model iff the model satisfies

$$\exists x, y \forall i < m \ \neg A_b(x, i, (y)_i) \tag{0.3}$$

where y is a list of strings $(y)_i$, $i < m$. The constructions in [Kra11, Kra25] yield models satisfying the negation of (0.1)

$$\exists x \forall i < m \exists y_i \ \neg A_b(x, i, y_i) \ . \tag{0.4}$$

To get (0.3) from (0.4) one needs the sharply bounded collection scheme BB of [Bus86] for open L_{PV}-formulas allowing to switch the two inner quantifiers. Unfortunately by [CT06] this principle is not provable in T_{PV} (unless factoring is not hard) while adding it to T_{PV} makes the proof-theoretic analysis of (0.1) from [Kra11, Kra25] invalid (the KPT theorem of [KPT91] does not apply in any usable form to this non-universal theory).

In this note we propose a model-theoretic method how to bypass the BB scheme and get lengths-of-proofs lower bounds from a hypothesis that a certain Σ_2^p search problem is hard for S-T computations. The model-theoretic assumption is of the form saying that certain extensions of models of T_{PV} exist. The caveat is that such extensions are known to exist for models satisfying the BB scheme for open formulas but it is not known in general. However, model theory seems to offer much more room to vary arguments than proof theory does and we think it is worthwhile to try to assert the model theoretic statement needed.

The paper is organized as follows. The relevant search problem and the hypothesis about its S-T intractability are discussed in Section 1, the model theoretic assumption is explained in Section 2 and the theorem reducing lengths-of-proofs lower bounds to Hypothesis (ST) is stated and proved in Section 3.

We assume that the reader is familiar with basic proof complexity theory. We do use quite a lot of notions and facts and it is unfeasible to review it here (the review would be substantially longer than the technical part). However, the reader can find all what we use in [Kra19]. We formulate the result and the arguments without referring to proof complexity generators to make it available also to readers not familiar with that theory, although that is the primary motivation for this research. The reader can find everything that we use in [Kra25].

A convention: letters $p, q, \ldots$, possibly with superscripts $p^i, q^i, \ldots$ denote *tuples* of propositional atoms.

1. The search problem and its S-T computability

A propositional disjunction $\bigvee_{i<r} \alpha_i$ is a **disjoint disjunction**, denoted

$$\dot{\bigvee}_{i<m} \alpha_i \ ,$$

if no two α_i have an atom in common. We often leave m out, and write just $\dot{\bigvee}_i \alpha_i$, as m is bounded by the size of the formula. Recall from [CR79] that a propositional proof system (abbr. pps) is a p-time decidable binary relation $P(x, y)$ such that the condition $\exists x P(x, y)$ defines TAUT. Instead of writing $P(\sigma, \beta)$ we use the suggestive notation $\sigma : P \vdash \beta$ for the provability relation: σ is a P-proof of β. The symbol $P \vdash \beta$ means that β has a P-proof.

Given a pps P, the Σ^p_2 **search problem** $\mathbf{DD}_P$ is:

- *input*: a pair $\pi, \dot{\bigvee}_i \alpha_i$ s.t. $\pi : P \vdash \dot{\bigvee}_i \alpha_i$,
- *solution*: any i s.t. $\alpha_i \in$ TAUT.

The problem can be solved in the S-T model described in the Introduction. S, upon receiving an input, proposes his first candidate solution i_1. If it is correct T acknowledges it. Otherwise she sends to S a counter-example: an assignment falsifying α_{i_1}. S then proposes his second candidate solution, using also the counter-example he learned in the first round, and the computation proceeds analogously for some number of rounds.

Following [Kra25, Def.2.4.3] we define a **strong pps** to be an extension of EF (Extended Frege, cf. [CR79]) by a p-time set of tautologies as extra axioms (see [Kra25, Thm.2.4.4] for their nice properties). Every pps can be p-simulated by a strong one.

Hypothesis (ST): *There exists a strong pps P such that $DD_P \notin ST[P, O(1)]$, P standing here also for the class of p-time algorithms computing functions.*

In fact, we do think that the hypothesis is true for all strong proof systems (which is equivalent to saying that it holds for EF, the Extended Frege proof system) but for the statement this apparently weaker formulation suffices.

The DD_P problem is a special case of a problem with a larger class of inputs applying to formulas of the form (0.2); we shall call the problem here D_P:

- *input*: a triple $\pi, \bigvee_i \beta_i(p, q^i), a$ s.t. $\pi : P \vdash \bigvee_i \beta_i$, only atoms p can occur in more than one β_i, and a is a truth assignment to atoms p,

• *solution*: any i s.t. $\beta_i(a, q^i) \in \mathrm{TAUT}$.

Lemma 1.1.

$$DD_P \notin ST[P, O(1)] \Leftrightarrow D_P \notin ST[P, O(1)]$$

The possibility that $\mathrm{D}_P \in \mathrm{ST}[\mathrm{P}, O(1)]$ holds for strong proof systems was considered in [PS21]. However, it was shown in [Kra20] that this is not true (assuming the existence of strong one-way permutations), essentially for reasons similar to why the feasible interpolation fails for them (cf. [Kra19, Chpt.18]). The problem to witness (0.1) studied in [Kra11] does not have a P-proof as a part of the input and this led to the need to consider non-uniform p-time students there. Another plausibly looking hypothesis about the feasible intractability of a certain search problem related to non-deterministic circuits and implying (ST) was given in [Kra25] (cf. hypothesis (K) there).

We state for the record the plausibility of (ST).

Lemma 1.2 [Kra20]**.** *(ST) is true if a strong one-way permutation exists (no polynomial size circuit can invert it with a non-negligible advantage).*

2. Bounded arithmetic models

The classic theorem we shall recall now uses the correspondence between first-order theories and proof systems mentioned in the Introduction. A prototype of the statement was proved in [KP90, Thm.1] working with theory S^1_2 of [Bus86] and its relation to Extended Frege proof system EF (a consequence of [Coo75]). The same argument applies to any pair of corresponding theories and proof systems using the generalization of [Coo75] developed in [KP89]. The underlying argument is discussed in detail also in [Kra19, Sec.20.1].

As we are aiming at all proof systems at once we can work with the theory T_{PV} and avoid discussing the correspondence between theories and proof systems. The first item of the theorem speaks about propositional proof systems and by those we mean proof systems determined by standard p-time binary relations and not proof systems in the sense of the non-standard model. In particular, a pps P may be incomplete in a model of T_{PV} but it will always be sound as that is expressible by a universal L_{PV}-formula.

Theorem 2.1 (after [KP90, Thm.1])**.** *Assume* $\mathbf{M}$ *is a non-standard model of* T_{PV} *and let* $\varphi \in \mathbf{M}$ *be a propositional formula (in the sense of the model). The following two statements are equivalent:*

(1) φ *has no proof in* $\mathbf{M}$ *in any proof system.*
(2) *There exists an extension* $\mathbf{M}' \supseteq \mathbf{M}$ *such that*

$$\mathbf{M}' \ \models\ T_{PV}\ +\ (\neg\varphi) \in SAT\ .$$

The fact that the construction in [KP90] works with models of S^1_2 implies that $\mathbf{M}'$ preserves (i.e. will satisfy) all Σ^b_1 properties of elements of $\mathbf{M}$ true there (this uses that the models satisfy the sharply bounded collection principle BB for open Σ^b_1-formulas, a consequence of S^1_2 (cf. [Bus86, Kra95]).

Further it is remarked at the end of [KP90] that one can arrange that $\mathbf{M}'$ introduces no new lengths into $\mathbf{M}$:

$$Log(\mathbf{M}') = Log(\mathbf{M}) \tag{2.1}$$

where $Log(\mathbf{M}) = \{|m| \mid m \in \mathbf{M}\}$. A construction which yields this (with an additional requirement on $\mathbf{M}$) is the forcing construction underlying [Kra95, Thm.9.4.2].

The problem we are interested in is whether these two properties, that is item 2 of Theorem 2.1 and (2.1), can be arranged too if we only know that $\mathbf{M} \models T_{\rm PV}$. We shall formulate the problem in a less direct way but closer to the construction in the proof of [Kra95, Thm.9.4.2]. This formulation also relates better to the possibility to use Boolean-valued models discussed at the end of [Kra16].

Problem 2.2. *Assume $\mathbf{M} \models T_{PV}$ and that it satisfies Condition 1 in Theorem 2.1. Does it follow that there are L_{PV}-structures $\mathbf{M}^*, \mathbf{M}'$:*

$$\mathbf{M} \subseteq \mathbf{M}^* \subseteq \mathbf{M}'$$

such that

(1) $\mathbf{M}^*$ *preserves all $\Sigma^b_1(PV)$-properties of elements of* $\mathbf{M}$*,*
(2) $\mathbf{M}' \models T_{PV} + (\neg\varphi) \in SAT$*,*
(3) $Log(\mathbf{M}^*) = Log(\mathbf{M}')$*?*

The assumption we shall use in Theorem 3.1 is that the problem has the affirmative solution. Note that the negative answer implies P $\neq$ NP as otherwise[2] $T_{\rm PV}$ proves $S^1_2({\rm PV})$.

The problem asks about all models of $T_{\rm PV}$ but it would suffice to get the affirmative answer for a sufficiently universal class of models (e.g. countable or recursively saturated). Also, it suffices if $\mathbf{M}^*$ preserves from $\mathbf{M}$ properties of the form $\forall i \exists y B(x, i, y)$ with i sharply bounded, y bounded and B open.

Let us remark that constructing extensions of bounded arithmetic models introducing no new lengths is very close to constructing expansions of non-standard finite structures. Such expansions are at the heart of applications of model theory to lower bounds for circuits or proofs as it is spectacularly demonstrated in [Ajt83, Ajt88]. Developing such constructions is in my view a crucial task in proof complexity; the survey [Kra16] discusses this in detail (cf. also [Kra19, Chpt.20])).

3. The theorem

Universal $L_{\rm PV}$ sentences $\forall x A(x)$ can be translated into a sequence of p-size propositional formulas $\|A\|^n$ such that

$$\forall x(|x| = n) A(x) \leftrightarrow \|A\|^n \in {\rm TAUT}$$

(cf. [Kra95] or [Kra19, Sec.12.6]).

For a pps P let Ref_P be the universal valid $L_{\rm PV}$ formula

$$x : P \vdash y \ \rightarrow (\neg y) \notin {\rm SAT}$$

where we simplify the notation and write $\neg y$ for the p-time function sending a code of a formula to the code of its negation.

Theorem 3.1. *Assume that Problem 2.2 has the affirmative solution. Then Hypothesis (ST) implies NP $\neq$ coNP.*

[2]If f is a p-time algorithm that finds a satisfying assignment for all satisfiable formulas then this fact, being a universal statement, would be in $T_{\rm PV}$ and hence $T_{\rm PV}$ would prove all true bounded formulas.

Proof.

(1)

Assume for the sake of a contradiction that the hypothesis (ST) and NP = coNP both hold. We shall bring it to a contradiction using the model-theoretic assumption. Take a strong pps P that satisfies the condition in (ST) and is also p-bounded.

(2)

Consider theory S in the language extending L_{PV} by two constant symbols: π and α, and having the following axioms:

- T_{PV},
- α is a disjoint disjunction of the form $\dot{\bigvee}_{i<m}\alpha_i$ (m is determined by α),
- $\pi : P \vdash \alpha$,
- $\forall i < m\ (\neg\alpha_i) \in \mathrm{SAT}$.

If S were inconsistent then applying the KPT theorem (cf. [KPT91, Kra95, Kra25]) to the universal theory with axioms listed in the first three items would give p-time functions computing moves of a student that solves DD_P in a constant number of rounds, thus violating (ST).

Let $\mathbf{M}$ be a model (necessarily non-standard) of S.

(3)

Assume $c \geq 1$ is a constant such that any $\beta \in \mathrm{TAUT}$ has a P-proof of size $\leq |\beta|^c$. We shall abbreviate by

$$\sigma : P \vdash_* \beta$$

the formula

$$(\sigma : P \vdash \beta)\ \wedge\ (|\sigma| \leq |\beta|^c)\ .$$

In the following formula the symbols $(y)_i$ and $(z)_i$ denote the i-th string in the list coded by y, z, respectively, and $len(y) \leq |y|$ denotes the number of these strings.

Claim: *For any pps Q the universal formula A_Q:*

$$x : P \vdash \dot{\bigvee}_i (y)_i\ \rightarrow\ Q \not\vdash \|\forall i < len(y)\ (z)_i : P \not\vdash_* (y)_i\| \tag{3.1}$$

(we leave the obvious length-bounds in the translation out) is true and hence in T_{PV}.

The claim holds because some $(y)_i$ has to be a tautology and hence - by the p-boundedness of P - would have a short P-proof.

(4)

All A_Q hold in $\mathbf{M}$ and we can substitute $x := \pi$ and $y := \alpha$. This yields

$$\mathbf{M}\ \models\ Q \not\vdash \varphi(r)$$

where $\varphi(r)$ is obtained from

$$\|\forall i < len(y)\ (z)_i : P \not\vdash_* (y)_i\|(r, q)$$

by substituting the bits of α for atoms q (hence the sharply bounded universal quantifier becomes $\bigwedge_{i<m}$) and the remaining atoms r correspond to (bits of) z.

(5)

Now we conclude the proof by using the assumption that Problem 2.2 has the affirmative solution. This gives us $\mathbf{M} \subseteq \mathbf{M}^* \subseteq \mathbf{M}'$ such that

$$Log(\mathbf{M}^*) = Log(\mathbf{M}')\ , \tag{3.2}$$

and $\mathbf{M}' \models T_{\rm PV} \wedge \neg\varphi \in {\rm SAT}$. Note that $\mathbf{M}^*$ satisfies the last axiom of S:

$$\mathbf{M}^* \ \models\ \forall i < m\ (\neg\alpha_i) \in {\rm SAT} \tag{3.3}$$

as it preserves the $\Sigma^b_1(PV)$-properties from $\mathbf{M}$.

Let $\sigma \in \mathbf{M}'$ be an assignment to the atoms r that falsifies the formula φ in $\mathbf{M}'$. Hence $\sigma = (\sigma_i)_{i<m}$ and we have

$$\mathbf{M}' \ \models\ \sigma_{i_0} : P \vdash \alpha_{i_0}$$

for some $i_0 < m$. But by (3.2) i_o is also already in $\mathbf{M}^*$ and hence by (3.3) we get:

$$\mathbf{M}' \ \models\ (\neg\alpha_{i_0}) \in {\rm SAT}\ .$$

That contradicts Ref_P which holds in $\mathbf{M}'$ as it models $T_{\rm PV}$. □

Note that the proof yields a somewhat stronger conclusion that in no strong P can a proof of some disjunct in a disjoint disjunction be polynomially bounded in size in terms of a proof of the disjunction (i.e. the strong feasible disjunction property fails for strong proof systems, cf.[Kra25])

Acknowledgment

I thank N.Arteche (Lund), O.Ježil (Prague) and M.Müller (Passau) for comments on a draft of the paper.

References

[Ajt83] M. Ajtai. Σ^1_1 - formulas on finite structures. *Annals of Pure and Applied Logic*, 24: 1–48, 1983.

[Ajt88] M. Ajtai. The complexity of the pigeonhole principle. *Proc. IEEE 29th Annual Symp. on Foundation of Computer Science*, 346–355, 1988.

[Bus86] S. R. Buss. *Bounded Arithmetic.* Naples, Bibliopolis, 1986.

[Coo75] S. A. Cook. Feasibly constructive proofs and the propositional calculus, *Proc. 7th Annual ACM Symp. on Theory of Computing*, 83–97, 1975.

[CR79] S. A. Cook and R. A. Reckhow. The relative efficiency of propositional proof systems. *J. Symbolic Logic*, 44(1): 36–50, 1979.

[CT06] S. A. Cook and N. Thapen. The strength of replacement in weak arithmetic. *ACM Transactions on Computational Logic*, 7:4, 2006.

[KP89] J. Krajíček and P. Pudlák. Propositional proof systems, the consistency of first-order theories and the complexity of computations. *J. Symbolic Logic*, 54(3):1063–1079, 1989.

[KP90] J. Krajíček and P. Pudlák. Propositional provability in models of weak arithmetic. *Computer Science Logic (Kaiserlautern, Oct. '89)*, eds. E. Boerger, H. Kleine-Bunning and M.M. Richter, Lecture Notes in Computer Science 440, 193–210, 1990.

[KPS90] J. Krajíček, P. Pudlák, and J. Sgall. Interactive Computations of Optimal Solutions. B. Rovan (ed.): *Mathematical Foundations of Computer Science* (B. Bystrica, August '90), Lecture Notes in Computer Science 452, 48–60, 1990.

[KPT91] J. Krajíček, P. Pudlák and G. Takeuti. Bounded arithmetic and the polynomial hierarchy. *Annals of Pure and Applied Logic*, 52:143–153, 1991.

[Kra95] J. Krajíček. *Bounded arithmetic, propositional logic, and complexity theory.* Encyclopedia of Mathematics and Its Applications, 60, Cambridge University Press, 1995.

[Kra11] J. Krajíček. On the proof complexity of the Nisan-Wigderson generator based on a hard $NP \cap coNP$ function, *J. of Mathematical Logic*, 11(1):11-27, 2011.

[Kra16] J. Krajíček. Expansions of pseudofinite structures and circuit and proof complexity. *Liber Amicorum Alberti*, eds.Jan van Eijck, Rosalie Iemhoff and Joost J. Joosten, Tributes Ser. 30, College Publications, London, 195–203, 2016.

[Kra19] J. Krajíček. *Proof complexity.* Encyclopedia of Mathematics and Its Applications, 170, Cambridge University Press, 2019.

[Kra20] J. Krajíček. A limitation on the KPT interpolation. *Logical Methods in Computer Science*, 16:3, 2020.

[Kra25] J. Krajíček. *Proof complexity generators.* London Mathematical Society Lecture Note Series, No.497, Cambridge University Press, 2025.

[PS21] J. Pich and R. Santhanam. Strong Co-Nondeterministic Lower Bounds for NP Cannot be Proved Feasibly. *Proc. of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, 223–233, 2021.