The Journal of Symbolic Logic
Volume 87, Number 3, September 2022

NEW RELATIONS AND SEPARATIONS OF CONJECTURES ABOUT INCOMPLETENESS IN THE FINITE DOMAIN

ERFAN KHANIKI

Abstract. In [20] Krajíček and Pudlák discovered connections between problems in computational complexity and the lengths of first-order proofs of finite consistency statements. Later Pudlák [25] studied more statements that connect provability with computational complexity and conjectured that they are true. All these conjectures are at least as strong as $\mathsf{P} \neq \mathsf{NP}$ [23–25]. One of the problems concerning these conjectures is to find out how tightly they are connected with statements about computational complexity classes. Results of this kind had been proved in [20, 22]. In this paper, we generalize and strengthen these results. Another question that we address concerns the dependence between these conjectures. We construct two oracles that enable us to answer questions about relativized separations asked in [19, 25] (i.e., for the pairs of conjectures mentioned in the questions, we construct oracles such that one conjecture from the pair is true in the relativized world and the other is false and vice versa). We also show several new connections between the studied conjectures. In particular, we show that the relation between the finite reflection principle and proof systems for existentially quantified Boolean formulas is similar to the one for finite consistency statements and proof systems for non-quantified propositional tautologies.

§1. Introduction. In 1989 Krajíček and Pudlák showed in [20] that the following two apparently unrelated statements are equivalent:

(I) There does not exist an optimal propositional proof system. (Optimal means that the lengths of shortest proofs of given tautologies are at most polynomially longer than in any other proof system.)

(II) There is no consistent finitely axiomatized arithmetical theory T such that T proves finite consistency statements by polynomial length proofs for every consistent finitely axiomatized arithmetical theory S. (The finite consistency statement says that there is no first-order proof of contradiction using a standard proof system such as a Hilbert's style proof system from S of length at most n.)

Following [25], we denote statement (I) by $\mathsf{CON}^{\mathsf{N}}$ (some conjectures have two versions and the superscript N indicates the nonuniform one). Krajíček and Pudlák proved that $\mathsf{CON}^{\mathsf{N}}$ implies $\mathsf{NE} \neq \mathsf{coNE}$, which is at least as strong as $\mathsf{P} \neq \mathsf{NP}$. Thus in a way $\mathsf{CON}^{\mathsf{N}}$ is connected with the problem whether nondeterministic computations and conondeterministic computations are polynomially related. They also proved that $\mathsf{CON}^{\mathsf{N}}$ is equivalent to a particular sentence Φ stated purely in computational complexity terms, but Φ does not express a fact about complexity

Received November 29, 2018.
2020 *Mathematics Subject Classification*. 03F20, 03D15, 68Q15, 03F30, 03F40, 68Q17.
Key words and phrases. finite consistency, propositional proof systems, disjoint NE-sets, TFNP class, oracles, relativized worlds.

0022-4812/22/8703-0003
DOI:10.1017/jsl.2021.99

classes, such as $\mathsf{P} \neq \mathsf{NP}$. Although it may not be possible to express $\mathsf{CON}^{\mathsf{N}}$ by a sentence about complexity classes, the question remains how to approximate it as closely as possible by such sentences. We prove several results in this direction, in particular, we show that if we generalize the concept of *optimal proof system* to *nonuniform subExp-optimal proof system* (nonuniform *subExp*-optimal means that the lengths of shortest proofs of given tautologies are at most sub-exponentially longer than in any other proof system), then it is possible to prove a generalization of Krajíček and Pudlák's theorem about the nonexistence of optimal proof system and $\mathsf{NE} \neq \mathsf{coNE}$. In more detail, we prove that if there does not exist a nonuniform *subExp*-optimal proof system, then $\mathsf{NE}_k \neq \mathsf{coNE}_k$ for every $k \geq 1$. (NE_k and coNE_k are the nondeterministic and conondeterministic classes of problems that are decidable in k-times exponential time.)

In order to fully understand these problems, it is useful to put them in a broader context. Therefore more recently Pudlák [25] studied finite reflection principles for Σ_1^b formulas as a generalization of finite consistency statements (Σ_1^b formulas are the logical characterization of NP sets). Given a fixed encoding of arithmetical formulas, these statements say that for every natural number n and k such that k has at most n digits in its binary representation, and every Σ_1^b formula $\phi(x)$ such that the length of the code of $\phi(x)$ is less than n, if there is a proof of $\phi(k)$ with length at most n, then $\phi(k)$ is true. He then considered variants of the following statement:

- There is no consistent arithmetical theory T with polynomial time decidable set of axioms such that for every consistent arithmetical theory S with polynomial time decidable set of axioms, T proves finite Σ_1^b-reflection principles associated with S by polynomial length proofs.

This statement is denoted by $\mathsf{RFN}_1^{\mathsf{N}}$.

The conjectures about finite Σ_1^b-reflection principles are about the unprovability of them in theories of arithmetic by polynomial size first-order proofs. We prove that these statements have an equivalent version in propositional proof complexity terms (Theorem 3.1). In particular, we prove that $\mathsf{RFN}_1^{\mathsf{N}}$ is equivalent to:

- There does not exist an optimal proof system for Σ_1^q-tautologies.

Σ_1^q-formulas are quantified propositional formulas that start with existential quantifiers over some atomic variables and the rest is a usual propositional formula.

After introducing variants of the finite reflection principles, Pudlák proved some relations between statements about finite reflection principles and some other conjectures in proof complexity and computational complexity. In particular, he considered the relationship between the following statements:

(I') There does not exist a p-optimal proof system for satisfiable propositional formulas.

(II') There does not exist a p-optimal propositional proof system (for tautologies).

(III') There is no consistent arithmetical theory T with polynomial time decidable set of axioms such that for every consistent arithmetical theory S with polynomial time decidable set of axioms, there exists a polynomial time computable function f that given n, generates a T-proof of the finite Σ_1^b-reflection principle for length n associated with S.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

Here p-optimality is a similar concept to optimality with the difference that there is a polynomial time computable function that given proofs in the former proof system, generates proofs in the later proof system. Pudlák proved that if at least one of the statements (I') or (II') is true, then statement (III') is also true. We complete his result by proving that if statement (III') is true, then at least one of the statements (I') or (II') is true (Theorem 3.2). As Pudlák mentioned in [25], the relation that he showed is not hard to prove. We see that in the proof of our theorem, the validity of the opposite direction is not trivial. We were able to prove the opposite direction by looking at these statements through the lens of mathematical logic and using the known theorems in this area.

One of the main concerns of [25] is to emphasize that the logical point of view of problems in computational complexity and proof complexity can be very beneficial, so this theorem counts as another example of this fact. It is also important to note that there are several examples of successful application of mathematical logic to problems of computational complexity and propositional proof complexity such as [1, 2, 7, 18, 20, 26]. It is worth noting that except the logical machinery that Pudlák used in [25] for treating the existence of complete problems for promise classes and the existence of optimal proof systems for different sets, there exist other types of machinery for investigating these questions that have more complexity theoretic taste. In particular, Beyersdorff and Sadowski in [4] provided a new complexity theoretic characterizations for treating the existence of complete problems for promise classes and also the existence of optimal proof systems for different sets. Their characterizations work for almost all promise classes C and sets L.

Another source of related conjectures is the research into proof complexity and bounded arithmetic. Much of the research in bounded arithmetic has been concerned with the question of characterizing $\forall\Sigma^b_1$ sentences provable in various bounded arithmetic theories and their fragments. The $\forall\Sigma^b_1$ formulas start with unbounded universal quantifiers, followed by bounded existential quantifiers, and the rest is a Σ^b_0 formula, whose satisfiability is decidable in polynomial time; $\forall\Sigma^b_1$ sentences define TFNP search problems (Total NP search problems), a concept studied extensively in computational complexity. For many theories, the corresponding class of provable TFNP problems has been characterized. These results suggest that with the increasing strength of the theories, the classes of provable TFNP problems grow larger (but we, certainly, cannot prove it, unless we also prove $\mathsf{P} \neq \mathsf{NP}$). This led to the conjecture that the TFNP subclasses increase in general with the increasing strength of the theories in which they are provably total. When stated in a suitable way, the conjecture is equivalent to the simple sentence saying that there is no complete TFNP problem.

Having several plausible conjectures a natural question arises: is there a single natural conjecture that implies all others? Instead of one, two natural conjectures have been proposed which imply most of the studied ones:

(a) There is no complete disjoint pair of NP sets with respect to polynomial reductions.
(b) There is no complete disjoint pair of CoNP sets with respect to polynomial reductions.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

(a) implies $\mathsf{CON}^{\mathsf{N}}$, while (b) implies that there is no complete TFNP problem. In [25] Pudlák asked to either show a dependence between (a) and (b), or to find oracles such that (a) and (b) are independent with respect to these oracles. Also, he asked a similar question about the relations between the other conjectures that are stated in [25]. More specifically, he asked whether there is a dependence between the existence of a p-optimal propositional proof system and the existence of a complete problem for TFNP. Also, Krajíček in [19] stated that it is open whether the existence or nonexistence of a complete problem in TFNP relates in some way to the existence of a p-optimal or optimal propositional proof system (see bibliographical and other remarks in Chapter 19 of [19]). In this paper, we solve these problems. We construct two oracles $\mathcal{V}$ and $\mathcal{W}$ such that relative to $\mathcal{V}$:

- There is no complete disjoint pair of CoNP sets with respect to polynomial reductions.
- $\mathsf{E} = \mathsf{NE}$, which implies that there is a p-optimal propositional proof system.

And relative to $\mathcal{W}$:

- There does not exist an optimal propositional proof system.
- TFNP problems are polynomial time computable and hence TFNP has a complete problem.

These constructions are our main results (Theorems 5.1 and 5.2). Note that the existence of oracle $\mathcal{V}$ shows that (b) does not imply (a) in relativized worlds and hence it solves one direction of the Pudlák's first question. Also, the existence of these oracles implies several independences between the conjectures that are stated in [25] and hence it partially answers Pudlák's second question in general. In the specific case of the relation between the existence of a p-optimal propositional proof system and the existence of a complete problem for TFNP, it shows that these statements are independent in the relativized worlds (Corollary 5.3), hence it answers Pudlák and Krajíček's question in [19, 25].

After this paper was posted as a preprint [17], Dose in [12], constructed an oracle such that relative to it:

- There is no complete disjoint pair of NP sets with respect to polynomial reductions.
- There is a p-optimal proof system for satisfiable propositional formulas.

Together with our result, this answers Pudlák's question about disjoint NP pairs and disjoint CoNP pairs completely. Note, however, that Dose's result and ours are incomparable as we explain in Section 5.

The paper is organized as follows. In Section 2, we explain the basic definitions and notations and review the known results from [25]. In Section 3, we investigate finite reflection principles and prove new results about these principles. In Section 4, we investigate the relationship between the equality of the nondeterministic and deterministic computations and between the conjectures about the existence of p-optimal and optimal proof systems. We explain the construction of the oracles $\mathcal{V}$ and $\mathcal{W}$ in Section 5.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

§2. Preliminaries.

2.1. On first-order theories of arithmetic. Following the notation of [25], we use Buss's first-order theories of arithmetic in a fixed language (see [8]). The language is the standard language of Buss's Bounded Arithmetic, which is

$$\mathcal{L}_{BA} = \{0, S, +, \cdot, |x|, \lfloor x/2 \rfloor, x\#y\}.$$

The intended meaning of the $\lfloor x/2 \rfloor$ is clear. The meaning of the $|x|$ is $\lceil \log_2(x+1) \rceil$. $x\#y$ is interpreted as $2^{|x|\cdot|y|}$.

A sharply bounded quantifier is of the form $Qx < |t|, Q \in \{\forall, \exists\}$. The class of bounded formulas Σ^b_n, Π^b_n, $n \geq 1$ is defined by counting alternations of bounded quantifiers while ignoring sharply bounded quantifiers (see [8]). The class of Δ^b_n formulas is the class of Σ^b_n formulas that have an equivalent Π^b_n definition. The theory S^1_2 consists of basic axioms defining the usual properties of the function symbols and p-induction axioms

$$\phi(0) \wedge \forall x(\phi(\lfloor x/2 \rfloor) \rightarrow \phi(x)) \rightarrow \forall x \phi(x),$$

for every Σ^b_1 formula $\phi(x)$. S^1_2 is the base theory in provability with respect to the Bounded Arithmetic hierarchy like $\mathbf{I}\Sigma_1$ is with respect to Peano arithmetic. One of the main properties of S^1_2 is that Σ^b_1 definable functions of S^1_2 are polynomial time computable (see Chapter 5 of [8]). Additionally, all of the polynomial time computable functions are Δ^b_1 in S^1_2 (a Σ^b_1 formula ϕ is Δ^b_1 in T iff there exists a Π^b_1 formula ψ such that $T \vdash \phi \equiv \psi$). For more information about Bounded Arithmetic, see [8].

Let $\mathcal{T}$ be the set of all consistent theories T in $\mathcal{L}_{BA}$ such that $\mathsf{S}^1_2 \subseteq T$ and the set of axioms of T is polynomial time decidable. The paper [25] focuses on the unprovability and provability of Π_1 sentences with respect to members of $\mathcal{T}$. In [25], some of the well-known conjectures of computational complexity and proof complexity were translated to unprovability statements for members of $\mathcal{T}$.

Next, we explain notations and definitions for proof complexity conjectures and their translations in [25].

2.2. TFNP class. TFNP or Total NP search problem is the class of true $\forall\Sigma^b_1$ sentences. More formally, a total NP search problem is defined by a pair (p, R) such that:

1. $p(x)$ is a polynomial,
2. $R(x, y)$ is a polynomial time computable relation (Δ^b_1 in S^1_2), and
3. $\mathbb{N} \models \forall x \exists y(|y| \leq p(|x|) \wedge R(x, y))$.

For comparing the complexity of TFNP problems, reductions are defined as follows.

Definition 2.1. Suppose P and Q are in TFNP. We say P is polynomially reducible to Q if the search problem P can be solved in polynomial time using an oracle that gives the answers of queries from Q.

Usually, subclasses of TFNP are defined as follows. For a fixed TFNP problem P, there is a natural subclass of TFNP associated with P which is the set of all TFNP problems reducible to P. Another way to compare the complexity of TFNP problems is to measure how much logical strength is needed to prove that the search problem

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

is total. This approach does not mention reductions explicitly, but it produces a similar hierarchy. The next definition formalizes this notion which is defined in [25].

DEFINITION 2.2. Suppose T is in $\mathcal{T}$. We say a TFNP problem (p, R) is provably total in T or $(p, R) \in \mathsf{TFNP}(T)$ iff there exists a pair (q, ϕ) such that:

1. q is a polynomial,
2. $\phi(x, y)$ is Δ_1^b in S_2^1,
3. $\mathbb{N} \models \forall x, y((|y| \leq p(|x|) \wedge R(x, y)) \equiv (|y| \leq q(|x|) \wedge \phi(x, y)))$, and
4. $T \vdash \forall x \exists y(|y| \leq q(|x|) \wedge \phi(x, y))$.

Also, we define $\mathsf{TFNP}^*(T)$ as the class of all TFNP problems that are reducible to a problem in $\mathsf{TFNP}(T)$.

For many bounded arithmetics $T \in \mathcal{T}$ such as Buss's Bounded Arithmetics, $\mathsf{TFNP}(T)$ is characterized. Actually, $\mathsf{TFNP}(T)$ for a bounded arithmetic theory $T \in \mathcal{T}$ can be viewed as a measurement of the strength of the bounded arithmetic T, like the provably total recursive functions for strong theories. The following theorem shows the relationship between the reduction and the provability.

THEOREM 2.1. [25] *The following statements are equivalent*:

1. *There exists a problem* $(p, R) \in \mathsf{TFNP}$ *that is complete with respect to polynomial reductions for the class* TFNP.
2. *There exists a* $T \in \mathcal{T}$ *such that* $\mathsf{TFNP}^*(T) = \mathsf{TFNP}$.

The main conjecture about TFNP is that it does not have a complete problem with respect to polynomial reductions. We denote this conjecture by TFNP-conj.

2.3. Proof systems. Following the definition of Cook–Reckhow (see [11]), a proof system for a set $C \subseteq \mathbb{N}$ is a polynomial time computable function $P : \mathbb{N} \to \mathbb{N}$ (the graph of P is Δ_1^b in S_2^1) such that $\mathsf{Rng}(P) = C$. We assume that different objects such as formulas, proofs, etc. are coded naturally in binary strings, hence every binary code x can be represented by a natural number with binary expansion $1x$, which is denoted by $\llcorner x \lrcorner$. To code a sequence of finite binary strings from x_1 to x_n, which is denoted by $\langle x_1, \ldots, x_n \rangle$, we use the following coding $x_1^* x_2^* \ldots x_{n-1}^* x_n$, for which a binary string z, z^* is obtained from z by doubling its digits and appending the string 01 at the end of it. Note that we can use the same encoding schema to code a finite sequence of natural numbers. By this explanation, we can define proof systems for different sets, such as propositional tautologies (TAUT) or satisfiable propositional formulas (SAT). By the length of an object (formulas, proofs, ...) with the natural number n as its code, we mean $|n|$. For every object A, we use the notation $\ulcorner A \urcorner$ to denote the numerical code of A.

A proof system P for a set C is polynomially bounded iff there exists a polynomial $q(x)$ such that for every $n \in C$, there exists a proof $\pi \in \mathbb{N}$ such that $P(\pi) = n$ and $|\pi| \leq q(|n|)$. One of the most important conjectures in proof complexity is the nonexistence of a polynomially bounded proof system for TAUT. In terms of computational complexity language, this conjecture is equivalent to $\mathsf{NP} \neq \mathsf{CoNP}$. Another concept that is weaker than polynomially boundedness is optimality. The following definition formalizes the components of this concept.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

DEFINITION 2.3. Suppose P and Q are proof systems for a set C. We say that P simulates Q iff there exists a polynomial $h(x)$ such that:

$$\forall \pi \in \mathbb{N}, \forall n \in C(Q(\pi) = n \to \exists \pi' \in \mathbb{N}(|\pi'| \leq h(|\pi|) \wedge P(\pi') = n)).$$

We say that P p-simulates Q iff there exists a polynomial time computable function f such that:

$$\forall \pi \in \mathbb{N}, \forall n \in C(Q(\pi) = n \to P(f(\pi)) = n).$$

We call a proof system P for a set C is optimal (p-optimal) iff for every proof system Q for set C, P simulates (p-simulates) Q. We have the following conjectures about optimality:

- CON: There is no p-optimal proof system for TAUT.
- $\mathsf{CON}^{\mathsf{N}}$: There is no optimal proof system for TAUT.
- SAT-conj: There is no p-optimal proof system for SAT.

To translate these conjectures to provability and unprovability in theories of $\mathcal{T}$, we need to define some machinery. Note that for every $T \in \mathcal{T}$, because the axioms of T are polynomial time decidable, there exists a polynomial time computable relation $Pr_T(x, y)$ which is true iff x is the code of a T-proof in the usual Hilbert style calculi of a formula in $\mathcal{L}_{BA}$ with the code y. One of the important properties of $Pr_T(x, y)$ is stated in the following theorem.

THEOREM 2.2. [8] *For every $T \in \mathcal{T}$, every Σ_1^b formula $\phi(x)$, there exists a polynomial $p(x)$ such that $T \vdash \forall x(\phi(x) \to \exists y(|y| \leq p(|x|) \wedge Pr_T(y, \ulcorner \phi(\dot{x}) \urcorner))$.*

Note that for every nonempty set $C \subseteq \mathbb{N}$, C has a proof system iff C is recursively enumerable. Suppose $C \subseteq \mathbb{N}$ is a nonempty recursively enumerable set. Let $\phi_C(x)$ be a Σ_1 formula in $\mathcal{L}_{BA}$ defining C. To define a proof system for $\phi_C(x)$ from a theory $T \in \mathcal{T}$, we need to express natural numbers in $\mathcal{L}_{BA}$ in an efficient way. The following definition gives us an efficient way of defining the numerals.

DEFINITION 2.4. $$\bar{n} := \begin{cases} 0, & n = 0, \\ SS0 \cdot \bar{k}, & n = 2k, \\ S(SS0 \cdot \bar{k}), & n = 2k + 1. \end{cases}$$

Note that the coded version of $\bar{n}$ needs $O(\log_2 n)$ bits. Additionally, the notation $\ulcorner \phi(\dot{n}) \urcorner$ for formula $\phi(x)$ in $\mathcal{L}_{BA}$ denotes a polynomial time computable function such that it outputs the code of $\phi(\bar{n})$.

Suppose a is in C. Now we define the strong associated proof system of T for C, P_T^C, as follows:

1. Given π, if $\mathbb{N} \models Pr_T(\pi, \ulcorner \phi_C(\dot{n}) \urcorner)$ for some n, then outputs n and
2. otherwise outputs a.

Let $Con_T(n)$ be the formula $\forall x(|x| \leq n \to \neg Pr_T(x, \ulcorner \bot \urcorner))$. Using the above notations and definitions we can express theorems that explain the relationship between optimality of proof systems and provability in members of $\mathcal{T}$.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

THEOREM 2.3. [20] *The following statements are equivalent*:

1. *There exists an optimal proof system for* TAUT.
2. *There exists a* $T \in \mathcal{T}$ *such that for every* $S \in \mathcal{T}$, *the shortest T-proofs of* $Con_S(\bar{n})$ *is bounded by a polynomial in n.*

To work with propositional tautologies and satisfiable formulas, we use the polynomial time computable relation $\mathtt{Sat}(x, y)$, which means the propositional formula with code x is satisfied by assignment with code y. Also, we use Π_1^b notation $\mathtt{Taut}(x) := \forall y(y \leq x \rightarrow \mathtt{Sat}(x, y))$ to define propositional tautologies. In order to work with $\forall\Pi_1^b$ and $\forall\Pi_1^b(\alpha)$ sentences as a family of propositional tautologies, we use the usual translation of $\forall\Pi_1^b$ sentences, and the relativized translation of $\forall\Pi_1^b(\alpha)$ sentences as defined in [3].

THEOREM 2.4. [20] *The following statements are equivalent*:

1. *There exists a p-optimal proof system for* TAUT.
2. *There exists a* $T \in \mathcal{T}$ *such that for every* $S \in \mathcal{T}$, *there exists a polynomial time computable function h that for every n,* $h(n)$ *is a T-proof of* $Con_S(\bar{n})$.
3. *There exists a* $T \in \mathcal{T}$ *such that for every proof system P for* TAUT, *there exists a polynomial time formalization* $P'(x, y)$ *of relation* $P(x) = y$ *such that*

$$T \vdash \forall x, y(P'(x, y) \rightarrow \mathtt{Taut}(y)).$$

The following theorem gives an equivalent statement to the nonexistence of the p-optimal proof system for SAT. This theorem was not mentioned in [25], but its proof is similar to the proof of Theorem 2.4.

THEOREM 2.5. *The following statements are equivalent*:

1. *There exists a p-optimal proof system for. There exists a* $T \in \mathcal{T}$ *such that for every proof system P for* SAT, *there exists a polynomial time formalization* $P'(x, y)$ *of relation* $P(x) = y$ *such that*

$$T \vdash \forall x, y(P'(x, y) \rightarrow \exists z(z < y \wedge \mathtt{Sat}(y, z))).$$

2.4. Disjoint NP pairs, disjoint CoNP pairs. The concept of disjoint NP pairs and disjoint CoNP pairs are studied in [25] with the aim of stating stronger conjectures than TFNP-conj and $\mathsf{CON}^{\mathsf{N}}$. A pair of (Co)NP sets (U, V) is a disjoint (Co)NP pair iff $U \cap V = \varnothing$. We denote this class of pairs by Disj(Co)NP. In order to compare the complexity of disjoint (Co)NP pairs, the reductions are defined as follows:

DEFINITION 2.5. Suppose (U_0, U_1) and (U_0', U_1') are disjoint (Co)NP pairs. We say (U_0, U_1) is polynomially reducible to (U_0', U_1') iff there exists a polynomial time computable function f such that for every $i \in \{0, 1\}$:

$$\forall n \in \mathbb{N}(n \in U_i \rightarrow f(n) \in U_i').$$

Again, another way to compare the complexity of disjoint (Co)NP pairs is to measure how much logical strength is needed to prove that such a pair is disjoint. The next definition formalizes this notion.

DEFINITION 2.6. Suppose that T is in $\mathcal{T}$. We say a (Co)NP pair (U_0, U_1) is provably disjoint in T or $(U_0, U_1) \in \mathsf{Disj(Co)NP}(T)$ iff there exists a (Π_1^b) Σ_1^b pair (ϕ_0, ϕ_1)

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

such that:

1. $\mathbb{N} \models \forall x(x \in U_i \equiv \phi_i(x)), i \in \{0,1\}$.
2. $T \vdash \forall x(\neg\phi_0(x) \vee \neg\phi_1(x))$.

Like Theorem 2.1, the following theorem explains the relationship between the reduction and provability.

THEOREM 2.6. [25] *The following statements are equivalent*:

1. *There exists a pair* $(U, V) \in$ Disj(Co)NP *that is complete with respect to polynomial reductions for class* Disj(Co)NP.
2. *There exists a* $T \in \mathcal{T}$ *such that* $\mathsf{Disj(Co)NP}(T) = \mathsf{Disj(Co)NP}$.

The main conjecture about disjoint (Co)NP pairs is that it does not have a complete problem with respect to polynomial reductions. We denote this conjecture by Disj(Co)NP-conj.

2.5. A finite reflection principle. The conjecture about the finite Σ^b_1-reflection principles was proposed in [25] with the aim of connecting some previous conjectures, in particular those that we have stated in this section. To state the conjecture, we need the following theorem.

THEOREM 2.7. [16] *For every* $i \geq 1$ *there exists a* Σ^b_i *formula* μ_i *such that for every* Σ^b_i *formula* $\phi(x)$ *there exists a natural number* e *and a polynomial* p *such that*:

$$\mathsf{S}^1_2 \vdash \forall x, y\left(|y| \geq p(|x|) \rightarrow (\mu_i(\bar{e}, x, y) \equiv \phi(x))\right).$$

In the above theorem e is a natural encoding of ϕ such that if y is big enough ($|y| \geq p(|x|)$), then $\mu_i(\bar{e}, x, y)$ behaves like $\phi(x)$.

The finite Σ^b_1-reflection principle is defined as follows:

DEFINITION 2.7. For every $T \in \mathcal{T}$, $n \in \mathbb{N}$, the $\Sigma^b_1\mathsf{RFN}_T(\bar{n})$ is defined by

$$\forall e, u, x, z(|e|, |u|, |x|, |z| \leq \bar{n} \wedge Pr_T(u, \ulcorner\mu_1(\dot{e}, \dot{x}, \dot{z})\urcorner) \rightarrow \mu_1(e, x, z)).$$

The following conjectures are stated in [25]:

1. $\mathsf{RFN}^{\mathsf{N}}_1$: For every $T \in \mathcal{T}$, there exists an $S \in \mathcal{T}$ such that the T-proofs of $\Sigma^b_1\mathsf{RFN}_S(\bar{n})$ are not polynomially bounded in n.
2. RFN_1: For every $T \in \mathcal{T}$, there exists an $S \in \mathcal{T}$ such that the T-proofs of $\Sigma^b_1\mathsf{RFN}_S(\bar{n})$ cannot be constructed in polynomial time.

2.6. Relations between the conjectures. Figure 1 shows the relations between the conjectures of this section. For more information about the proofs of these relations see [25].

§3. Proof systems and the finite reflection principles. As we have seen, every conjecture that is discussed in the previous section has two formalizations, one in terms of proof complexity notations, and one in terms of unprovability statements, except RFN_1 and $\mathsf{RFN}^{\mathsf{N}}_1$. Here we want to show that these conjectures have equivalent forms in terms of optimal proof systems for Σ^q_1-TAUT. Σ^q_i (Π^q_i) propositional formulas are quantified propositional formulas. The next theorem is similar to Theorems 2.4 and 2.5 for CON and $\mathsf{CON}^{\mathsf{N}}$.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

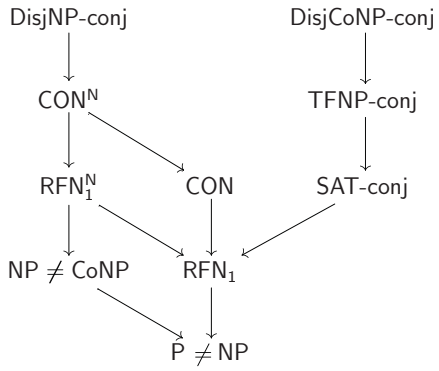

FIGURE 1. Relations between conjectures.

THEOREM 3.1.

1. *The following statements are equivalent*:
   (a) *For every* $T \in \mathcal{T}$, *there exists an* $S \in \mathcal{T}$ *such that the T-proofs of* $\Sigma^b_1\mathsf{RFN}_S(\bar{n})$ *are not polynomially bounded in n.*
   (b) Σ^q_1-TAUT *does not have an optimal proof system.*
2. *The following statements are equivalent*:
   (a) *For every* $T \in \mathcal{T}$, *there exists an* $S \in \mathcal{T}$ *such that the T-proofs of* $\Sigma^b_1\mathsf{RFN}_S(\bar{n})$ *cannot be constructed in polynomial time.*
   (b) Σ^q_1-TAUT *does not have a p-optimal proof system.*
   (c) *For every theory* $T \in \mathcal{T}$, *there exists a proof system P for* Σ^q_1-TAUT *such that T does not prove the soundness of any polynomial time formalization of P.*

PROOF. Here we prove the second part. The proof of the first part is similar.

(a) ⇒ (b). Suppose (b) is false. Let P be a p-optimal proof system for Σ^q_1-TAUT. Let $T := \mathsf{S}^1_2 + \forall\pi\mathtt{Taut}_{\Sigma^q_1}(P(\pi))$ in which $\mathtt{Taut}_{\Sigma^q_1}$ is the Π^b_2 formula that checks whether a Σ^q_1 propositional formula is true or not. Let $S \in \mathcal{T}$. Note that for every $n \in \mathbb{N}$, the translation of $\Sigma^b_1\mathsf{RFN}_S(\bar{n})$ is a Σ^q_1 formula θ_n such that $\mathsf{S}^1_2 \vdash \Sigma^b_1\mathsf{RFN}_S(\bar{n}) \equiv \mathtt{Taut}_{\Sigma^q_1}(\theta_n)$ *and this proof can be constructed in polynomial time* (*see* [9] *for the propositional case.*) (∗).

Let P' be a proof system defined as follows:

$$P'(x) := \begin{cases} \theta_n, & x = \theta_n \text{ for some } n, \\ P(x), & \text{o.w.} \end{cases}$$

Let f be the polynomial time computable function such that $P(f(\pi)) = P'(\pi)$ for every $\pi \in \mathbb{N}$. Note that for every $n \in \mathbb{N}$, the proof of $\mathsf{S}^1_2 \vdash P(f(\theta_n)) = \theta_n$ can be constructed in polynomial time, therefore by soundness of P which is provable in T and (∗), for every

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

$n \in \mathbb{N}$, a T-proof of $\Sigma^b_1\mathsf{RFN}_S(\bar{n})$ can be constructed in polynomial time too.

(b) $\Rightarrow$ (c). Suppose (c) is false. Let $T \in \mathcal{T}$ be a theory that falsifies (c). We want to prove that $P_T^{\Sigma^q_1}$ is p-optimal. Let P' be a proof system and P'' be one of its polynomial time formalizations such that $T \vdash \forall\pi \mathtt{Taut}_{\Sigma^q_1}(P''(\pi))$. Note that there exists a polynomial time computable function f such that

$$T \vdash \forall\pi,\phi(P''(\pi) = \phi \rightarrow Pr_T(f(\pi,\phi), \ulcorner P''(\dot{\pi}) = \dot{\phi}\urcorner)),$$

hence there exists a polynomial time computable function h such that $P''(\pi) = P_T^{\Sigma^q_1}(h(\pi))$, for all $\pi \in \mathbb{N}$.

(b) $\Rightarrow$ (a). Suppose (a) is false. Let $T \in \mathcal{T}$ be a theory that witnesses this fact. We want to show that $P_T^{\Sigma^q_1}$ is p-optimal. Let $\mathtt{Sat}_{\Sigma^q_1}(\phi, v)$ be the Σ^b_1 formula that can check the satisfiability of Σ^q_1 propositional formulas. Let P be a proof system for Σ^q_1-TAUT. Define $T' := \mathsf{S}^1_2 + \forall\pi, v\mathtt{Sat}_{\Sigma^q_1}(P(\pi), v)$. If $P(\pi_\psi) = \psi$, then we can find a proof π' in polynomial time such that $P_{T'}^{\Sigma^q_1}(\pi') = \psi$ (*). Note that there exists a polynomial time computable function f such that

$$\mathbb{N} \models \forall\pi, v, \phi(|v| \leq |\phi| \wedge P_{T'}^{\Sigma^q_1}(\pi) = \phi \rightarrow P_{T'}^{\Sigma^q_1}(f(\pi, v)) = \phi[v/\vec{p}]).$$

Let $T'' := \mathsf{S}^1_2 + \forall\pi, v, \phi(|v| \leq |\phi| \wedge P_{T'}^{\Sigma^q_1}(\pi) = \phi \rightarrow P_{T'}^{\Sigma^q_1}(f(\pi, v)) = \phi[v/\vec{p}])$. Note that T falsifies RFN_1, hence P_T is a p-optimal proof system for TAUT, this means P_T p-simulates $P_{T''}$ (**). The propositional translations of

$$\forall\pi, v, \phi(|v| \leq |\phi| \wedge P_{T'}^{\Sigma^q_1}(\pi) = \phi \rightarrow P_{T'}^{\Sigma^q_1}(f(\pi, v)) = \phi[v/\vec{p}]),$$

have short proofs in $P_{T''}$ and these proofs can be constructed in polynomial time, hence by (*) and (**) we can find a T'-proof π'' of $\forall v(|v| \leq |\psi| \rightarrow P_{T'}^{\Sigma^q_1}(f(\pi', v), \psi[v/\vec{p}]))$ in polynomial time, therefore by constructing a $\Sigma^b_1\mathsf{RFN}_{T'}(n)$ for some suitable n which is polynomial in size of ψ, we can find a proof π^* such that $P_T^{\Sigma^q_1}(\pi^*) = \psi$. So $P_T^{\Sigma^q_1}$ is p-optimal for Σ^q_1-TAUT.

(c) $\Rightarrow$ (b). Suppose (b) is false. Let $T \in \mathcal{T}$ be a theory that witnesses this fact. Thus, the theory $\mathsf{S}^1_2 + \forall\pi \mathtt{Taut}_{\Sigma^q_1} P_T^{\Sigma^q_1}(\pi)$ falsifies (c). $\dashv$

The previous theorem can be generalized for finite reflection principle conjectures for Σ^b_i formulas, as RFN_i.

Figure 1 suggests that the upper conjectures are stronger than those that are below them but it is not known whether an opposite implication can be proved, i.e., a lower conjecture implies an upper one. The next theorem shows a kind of opposite implication. In terms of notations, the next theorem shows that RFN_1 implies CON $\vee$

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

SAT-conj. Note that it is clear from the the Figure 1 that CON ∨ SAT-conj implies RFN_1, hence by the next theorem RFN_1 is equivalent to CON ∨ SAT-conj.

THEOREM 3.2. *At least one of the following statements is true*:

1. *There is no p-optimal proof system for* SAT.
2. *There is no p-optimal proof system for* TAUT.
3. *There exists a* $T \in \mathcal{T}$ *such that for every* $S \in \mathcal{T}$, *the T-proofs of* $\Sigma_1^b\mathsf{RFN}_S(\bar{n})$ *can be constructed in polynomial time.*

PROOF. Suppose (1) and (2) are false. Let $T \in \mathcal{T}$ be the theory that falsifies (1) and (2) simultaneously. Suppose S is in $\mathcal{T}$. We want to show that there exists a polynomial time computable function h such that for every Σ_1^q formula ϕ and every S-proof π of $\forall u(|u| \leq |\phi| \to \mathtt{Sat}_{\Sigma_1^q}(\phi, u))$, $h(\pi)$ is a T-proof of $\forall u(|u| \leq |\phi| \to \mathtt{Sat}_{\Sigma_1^q}(\phi, u))$. Hence $P_T^{\Sigma_1^q}$ is p-optimal and by Theorem 3.1, $\neg\mathsf{RFN}_1$ is true. Note that there exists a polynomial time computable function f such that

$$\mathbb{N} \models \forall \pi, v, \phi(|v| \leq |\phi| \wedge P_S^{\Sigma_1^q}(\pi) = \phi \to P_S^{\mathsf{SAT}}(f(\pi, v)) = \phi\big[v/\vec{p}\big]).$$

Suppose $P_S^{\Sigma_1^q}(\pi_\psi) = \psi$ for a Σ_1^q formula ψ, hence we can find a short T-proof of $P_S^{\Sigma_1^q}(\pi_\psi) = \psi$ in polynomial time $(*)$.

Define

$$S' := \mathsf{S}_2^1 + \forall \pi, v, \phi(|v| \leq |\phi| \wedge P_S^{\Sigma_1^q}(\pi) = \phi \to P(f(\pi, v)) = \phi\big[v/\vec{p}\big]),$$

such that P is a polynomial time formalization of P_S^{SAT} in which soundness of P is provable in T. Because T falsifies (2) and S' has a short proof of translation of

$$\forall v(|v| \leq |\psi| \wedge P_S^{\Sigma_1^q}(\pi_\psi) = \psi \to P(f(\pi_\psi, v)) = \psi\big[v/\vec{p}\big]),$$

we can find a short T-proof of translation of it in polynomial time. Therefore by $(*)$ we get a T-proof of $\forall v(|v| \leq |\psi| \to P(f(\pi_\psi, v)) = \psi\big[v/\vec{p}\big])$ $(**)$. Note that $\psi\big[v/\vec{p}\big]$ does not have free variables, hence there exists a polynomial time computable function g such that T has a short proof of

$$\forall v(|v| \leq |\psi| \to \mathtt{Taut}_{\Sigma_1^q}(\psi\big[v/\vec{p}\big]) \equiv \exists u \mathtt{Sat}(g(\psi\big[v/\vec{p}\big]), u)).$$

Hence by $(**)$ and by the fact that T proves P is a sound proof system for SAT, a T-proof of $\forall v(|v| \leq |\psi| \to \mathtt{Sat}_{\Sigma_1^q}(\psi, v))$ can be constructed in polynomial time. ⊣

§4. Nondeterministic vs deterministic computations and existence of optimal proof systems. In this section, we investigate the relationship between the existence of optimal proof systems and the equality of nondeterministic and deterministic computation. The trivial case is P = NP which implies the existence of polynomial time computable proofs for TAUT. The first step in this direction was done in [20]. They showed that E = NE implies the existence of p-optimal proof systems for TAUT. Later, it was shown in [22] that the condition EE = NEE is sufficient. This phenomenon was investigated further in [3]. In that paper, a general theorem

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

is proved from which it follows that there are oracles such that CON is true, yet EXP = NEXP or EEE = NEEE, and this also holds for higher classes. Apart from the mentioned results, there exists another theorem that explains the situation for the opposite direction. It is proved in [10] that if for every time constructible and increasing function h, $\mathsf{NTIME}(h^{O(1)}) \not\subseteq \mathsf{DTIME}(h^{O(\log h)})$, then TAUT does not have an effective p-optimal proof system (effective p-optimality is a stronger version of p-optimality. For more information see [10]). In the main theorem of this section, we investigate how much optimality we can get by assuming the equality of nondeterministic and (co-non)deterministic computation for complexity classes such as EXP and EEE (Theorem 4.2). Before proving this theorem, we prove a proposition to show the proof method of Theorem 4.2. To prove these statements, we use variants of the proof method of Krajíček and Pudlák in [20].

The next proposition states a sufficient condition for the existence of an optimal and p-optimal proof system for Σ_1^q-TAUT. Note that by Theorem 3.1, the existence of such a proof system is equivalent to $\neg\mathsf{RFN}_1^{\mathsf{N}}$ and $\neg\mathsf{RFN}_1$, respectively. It is shown in [25] that $\mathsf{RFN}_1^{\mathsf{N}}$ implies NP $\neq$ CoNP. The next proposition strengthens this result. To state the next proposition, we need to define the k'th Exponential Time Hierarchy.

DEFINITION 4.1. Define the following functions inductively:

1. $|x|_n := \begin{cases} |x|_0 = x, \\ |x|_{n+1} = ||x|_n|, \end{cases}$
2. $2_n^x := \begin{cases} 2_0^x = x, \\ 2_{n+1}^x = 2^{2_n^x}. \end{cases}$

DEFINITION 4.2. For every k, define the k'th *Exponential Time Hierarchy* (EH_k) as follows:

- For every $L \subseteq \mathbb{N}$, L is in E_k iff there exists a formula $\phi(x)$ that is Δ_1^b in S_2^1 such that $\forall n(n \in L \leftrightarrow \phi(2_k^n))$.
- For every $L \subseteq \mathbb{N}$, L is in $\Sigma_i^{\mathsf{E}_k}$ for some $i > 0$ iff there exists a Σ_i^b formula $\phi(x)$ such that $\forall n(n \in L \leftrightarrow \phi(2_k^n))$.
- For every $L \subseteq \mathbb{N}$, L is in $\Pi_i^{\mathsf{E}_k}$ for some $i > 0$ iff there exists a Π_i^b formula $\phi(x)$ such that $\forall n(n \in L \leftrightarrow \phi(2_k^n))$.

Recall that Σ_i^b (Π_i^b) formulas define exactly the sets that are in Σ_i^p (Π_i^p) of the Polynomial Hierarchy (see [8]). Therefore, by a padding argument E_k, NE_k and CoNE_k are the classes of sets that can be decided by deterministic, nondeterministic and conondeterministic Turing machines in time $2^{O(2_{k-1}^n)}$ respectively.

Note that we do not have a exponentiation function symbol in $\mathcal{L}_{BA}$, therefore by formula $\forall n\phi(2_k^{f(n)})$ for some polynomial time computable function f and some fixed k, we mean $\forall m, n(\psi_{f,k}(m,n) \to \phi(m))$ in which $\psi_{f,k}(m,n)$ is a Δ_1^b formula in S_2^1 that is true iff $m = 2_k^{f(n)}$. Moreover, for a function g and natural numbers n and m such that $g(n) = m$, $\overline{g(n)} := \bar{m}$.

PROPOSITION 4.1. *The following statements are true*:

1. *If for every $T \in \mathcal{T}$, there exists an $S \in \mathcal{T}$ such that the T-proofs of $\Sigma_1^b\mathsf{RFN}_S(\bar{n})$ are not polynomially bounded in n, then* $\mathsf{NE} \neq \Sigma_2^{\mathsf{E}}$.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

2. *If for every* $T \in \mathcal{T}$, *there exists an* $S \in \mathcal{T}$ *such that the* T*-proofs of* $\Sigma_1^b\mathsf{RFN}_S(\bar{n})$ *cannot be constructed in polynomial time, then* $\mathsf{E} \neq \Sigma_2^{\mathsf{E}}$.

Proof. Here we prove the statement (1). Statement (2) has a similar proof. Let $\mathsf{NE} = \Sigma_2^{\mathsf{E}}$. This implies that $\mathsf{NE} = \Pi_2^{\mathsf{E}}$, because $\mathsf{CoNE} \subseteq \Sigma_2^{\mathsf{E}}$. Define the following sets:

1. $L_{\mathsf{NE}} := \{n = \langle e, x, m\rangle \in \mathbb{N} : \mathbb{N} \models \mu_1(e, x, 2^{2^{|m|}})\} \in \mathsf{NE}$.
2. $L_{\Pi_2^{\mathsf{E}}} := \{n = \langle e, x, m\rangle \in \mathbb{N} : \mathbb{N} \models \neg\mu_2(e, x, 2^{2^{|m|}})\} \in \Pi_2^{\mathsf{E}}$.

Note that the above sets are hard for their respective complexity class under linear time reductions. By definition the following predicates exist:

1. There exists a Π_2^b predicate $\mathsf{U}_{\Pi_2^b}$ such that $\mathbb{N} \models \forall n(\mathsf{U}_{\Pi_2^b}(2^n) \leftrightarrow n \in L_{\Pi_2^{\mathsf{E}}})$.
2. There exists a Σ_1^b predicate U_{NP} such that $\mathbb{N} \models \forall n(\mathsf{U}_{\mathsf{NP}}(2^n) \leftrightarrow n \in L_{\mathsf{NE}})$.

Note that $\mathsf{NE} = \Pi_2^{\mathsf{E}}$ implies that there exists a linear time function f such that

$$\mathbb{N} \models \forall n(\mathsf{U}_{\Pi_2^b}(2^n) \leftrightarrow \mathsf{U}_{\mathsf{NP}}(2^{f(n)})),$$

because U_{NP} defines an NE hard set under linear time reductions. Let $T \in \mathcal{T}$ be a theory with the following properties:

1. $\mathbb{N} \models T$,
2. $T \vdash \mathsf{U}_{\mathsf{NP}}(2^n)$ is NE-hard with respect to linear time reductions,
3. $T \vdash \mathsf{U}_{\Pi_2^b}(2^n)$ is Π_2^{E}-hard with respect to linear time reductions, and
4. $T \vdash \forall n(\mathsf{U}_{\Pi_2^b}(2^n) \leftrightarrow \mathsf{U}_{\mathsf{NP}}(2^{f(n)}))$.

Let T' be in $\mathcal{T}$. This implies $\Sigma_1^b\mathsf{RFN}_{T'}(x)$ defines a Π_2^{E} set, so by the mentioned properties of T there exists a linear time function g such that $T \vdash \forall n\big(\Sigma_1^b\mathsf{RFN}_{T'}(n) \leftrightarrow \mathsf{U}_{\mathsf{NP}}(2^{f(g(n))})\big)$. Because $\mathsf{U}_{\mathsf{NP}}(x)$ is Σ_1^b and also $\mathsf{S}_2^1 \subseteq T$, there exists a polynomial $r(x)$ such that

$$T \vdash \forall x\big(\mathsf{U}_{\mathsf{NP}}(x) \to \exists y\big(|y| \leq r(|x|) \wedge Pr_T\big(y, \ulcorner \mathsf{U}_{\mathsf{NP}}(\dot{x})\urcorner\big)\big)\big).$$

This implies that

$$T \vdash \forall x\big(\mathsf{U}_{\mathsf{NP}}(2^{f(g(x))}) \to \exists y\big(|y| \leq r(f(g(x))+1) \wedge Pr_T\big(y, \ulcorner \mathsf{U}_{\mathsf{NP}}(2^{f(g(\dot{x}))})\urcorner\big)\big)\big).$$

Note that $\mathbb{N} \models \forall n\mathsf{U}_{\mathsf{NP}}(2^{f(g(n))})$, so for every $n \in \mathbb{N}$, $T \overset{r(f(g(n))+1)}{\vdash} \mathsf{U}_{\mathsf{NP}}(\overline{2^{f(g(n))}})$, hence there exists a polynomial $p(x)$ such that for every $n \in \mathbb{N}$, $T \overset{p(n)}{\vdash} \Sigma_1^b\mathsf{RFN}_{T'}(\bar{n})$. ⊣

To state Theorem 4.2, we need more definitions. Let $2^{n^{o(1)}}$ and $2^{(\log n)^{O(1)}}$ be sub-exponential (subExp) and quasi-polynomial (Qp) respectively. The concept of simulations and reductions can be stated in terms of other time classes like sub-exponential or quasi-polynomial time instead of polynomial time. Let A be a class of time functions. The concepts of nonuniform A-optimal proof system and A-optimal proof system are the same as the optimal proof system and p-optimal proof system with this difference that we use functions of A instead of polynomials in defining these concepts. It is worth noting that the relations in Figure 1 still remain true even if we use reductions and simulations that have Qp or subExp complexity. Hence it is natural to ask whether these new conjectures are true or not. An oracle is

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

constructed in [15] that DisjNP pairs do not have complete problems with respect to polynomial time reductions. It is not hard to modify that construction to construct an oracle in which DisjNP pairs do not have complete problems with respect to sub-exponential time reductions, hence conjectures that are weaker than it are true with respect to that oracle.

THEOREM 4.2. *The following statements are true*:

1. *If there is no nonuniform subExp-optimal proof system for* TAUT, *then for every* k, $\mathsf{NE}_k \neq \mathsf{CoNE}_k$.
2. *If there is no subExp-optimal proof system for* TAUT, *then for every* k, $\mathsf{E}_k \neq \mathsf{NE}_k$.
3. *If there is no nonuniform Qp-optimal proof system for* TAUT, *then* $\mathsf{NEXP} \neq \mathsf{CoNEXP}$.
4. *If there is no Qp-optimal proof system for* TAUT, *then* $\mathsf{EXP} \neq \mathsf{NEXP}$.

PROOF. Here we only prove the statement (1). The proofs of the other statements are similar. Let $\mathsf{NE}_k = \mathsf{CoNE}_k$ for some $k > 0$. Define the following sets:

1. $L_{\mathsf{NE}_k} := \{n = \langle e, x, m\rangle \in \mathbb{N} : \mathbb{N} \models \mu_1(e, x, 2^{|m|}_{k+1})\} \in \mathsf{NE}_k$.
2. $L_{\mathsf{CoNE}_k} := \{n = \langle e, x, m\rangle \in \mathbb{N} : \mathbb{N} \models \neg\mu_1(e, x, 2^{|m|}_{k+1})\} \in \mathsf{CoNE}_k$.

Note that the above sets are hard for their respective complexity class under linear time reductions. By definition the following predicates exist:

1. There exists a Σ^b_1 predicate U_{NP} such that $\mathbb{N} \models \forall n(\mathsf{U}_{\mathsf{NP}}(2^n_k) \leftrightarrow n \in L_{\mathsf{NE}_k})$.
2. There exists a Π^b_1 predicate $\mathsf{U}_{\mathsf{CoNP}}$ such that $\mathbb{N} \models \forall n(\mathsf{U}_{\mathsf{CoNP}}(2^n_k) \leftrightarrow n \in L_{\mathsf{CoNE}_k})$.

Note that $\mathsf{NE}_k = \mathsf{CoNE}_k$ implies that there exists a linear time function f such that

$$\mathbb{N} \models \forall n(\mathsf{U}_{\mathsf{CoNP}}(2^n_k) \leftrightarrow \mathsf{U}_{\mathsf{NP}}(2^{f(n)}_k)).$$

Let $T \in \mathcal{T}$ be a theory with the following properties:

1. $\mathbb{N} \models T$,
2. $T \vdash \mathsf{U}_{\mathsf{NP}}(2^n_k)$ is NE_k-hard with respect to linear time reductions,
3. $T \vdash \mathsf{U}_{\mathsf{CoNP}}(2^n_k)$ is CoNE_k-hard with respect to linear time functions, and
4. $T \vdash \forall n(\mathsf{U}_{\mathsf{CoNP}}(2^n_k) \leftrightarrow \mathsf{U}_{\mathsf{NP}}(2^{f(n)}_k))$.

Let T' be in $\mathcal{T}$. For every i, define $\mathsf{Con}^i_{T'}(x) := \forall y(|y|_i \leq x \rightarrow \neg Pr_{T'}(y, \ulcorner\bot\urcorner))$, hence $\mathsf{Con}^k_{T'}(x)$ defines a CoNE_k set. So by the mentioned properties of T there exists a linear time function g such that $T \vdash \forall n\big(\mathsf{Con}^k_{T'}(n) \leftrightarrow \mathsf{U}_{\mathsf{NP}}(2^{f(g(n))}_k)\big)$. Because $\mathsf{U}_{\mathsf{NP}}(x)$ is Σ^b_1 and also $\mathsf{S}^1_2 \subseteq T$, there exists a polynomial $r(x)$ such that

$$T \vdash \forall x\big(\mathsf{U}_{\mathsf{NP}}(x) \rightarrow \exists y\big(|y| \leq r(|x|) \wedge Pr_T\big(y, \ulcorner \mathsf{U}_{\mathsf{NP}}(\dot{x})\urcorner\big)\big)\big).$$

This implies

$$T \vdash \forall x\big(\mathsf{U}_{\mathsf{NP}}(2^{f(g(x))}_k) \rightarrow \exists y\big(|y| \leq r(2^{f(g(x))}_{k-1} + 1) \wedge Pr_T\big(y, \ulcorner \mathsf{U}_{\mathsf{NP}}(2^{f(g(\dot{x}))}_k)\urcorner\big)\big)\big).$$

Note that $\mathbb{N} \models \forall n \mathsf{U}_{\mathsf{NP}}(2^{f(g(n))}_k)$, so for every $n \in \mathbb{N}$, $T \vdash \overline{2^{f(g(n))}_k}$, hence there exists a polynomial $p(x)$ such that for every $n \in \mathbb{N}$, $T \vdash$

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

$\mathsf{Con}^k_{T'}(\bar{n})$, hence $T \vdash^{p(2^{f(g(|n|_{k-1}))}_{k-1})} \mathsf{Con}^k_{T'}(\overline{|n|_{k-1}})$, so there exists a polynomial $q(x)$ such that for every $n \in \mathbb{N}$, $T \vdash^{q(2^{f(g(|n|_{k-1}))}_{k-1})} \mathsf{Con}^1_{T'}(\bar{n})$. To complete the proof we need to show that $q(2^{f(g(|n|_{k-1}))}_{k-1}) = 2^{n^{o(1)}}$. For this matter it is sufficient to prove that $2^{(|n|_b)^c}_b = 2^{n^{o(1)}}$ where $b, c > 0$ and this can be proved by induction on b. By the fact that proof of Theorem 2.3 is adaptable in case of quasi-polynomial and sub-exponential, the proof is completed. ⊣

One can easily see that the proof of Theorem 4.2 can be adapted to the cases $\mathsf{RFN}^{\mathsf{N}}_1$ and RFN_1.

§5. Relativized worlds. In this section, we construct two oracles that imply several separations between the conjectures of Figure 1. Relative to the first oracle, DisjCoNP-conj is true, but $\mathsf{E} = \mathsf{NE}$ which implies CON is false (Theorem 5.1). Relative to the second oracle, $\mathsf{CON}^{\mathsf{N}}$ is true, but $\mathsf{TFNP} = \mathsf{FP}$ which implies TFNP-conj is false (Theorem 5.2). The existence of these oracles implies several separations between the conjectures of Figure 1, and in particular answers some open problems from [19, 25] (see bibliographical and other remarks in Chapter 19 of [19]). It is known that some of the conjectures in Figure 1 are true in relativized worlds. For example in [15], an oracle was constructed such that DisjNP-conj is true. In [24], an oracle was constructed such that TFNP-conj is true. Also in [3, 5], it is shown that $\mathsf{CON}^{\mathsf{N}}$ is true in relativized worlds. While in [3], they used a direct construction to satisfy $\mathsf{CON}^{\mathsf{N}}$, in [5], an oracle was constructed such that $\mathsf{NP} \cap \mathsf{SPARSE}$ has no complete sets and then they deduced by known results that relative to the constructed oracle $\mathsf{CON}^{\mathsf{N}}$ is true.

We use the usual definition of forcing in arithmetic to construct the oracles. It is standard to use the forcing relation in constructions of the oracles (see [6, 13]) as it makes the proofs more readable. Here we do not use any results about forcing in set theory.

Definition 5.1. A nonempty set $\mathcal{P}$ of partial functions from natural numbers to $\{0,1\}$ (for every $p \in \mathcal{P}$, $\mathsf{Dom}(p) \subseteq \mathbb{N}$ and $\mathsf{Rng}(p) \subseteq \{0,1\}$) is a forcing notion iff for every $p \in \mathcal{P}$, there exists a $q \in \mathcal{P}$ such that $p \subsetneq q$. We call members of a forcing notion conditions.

Let α be a new unary relation symbol. For every $p \in \mathcal{P}$ and every $\mathcal{L}_{BA}(\alpha)$ sentence ϕ, we define $p \Vdash \phi$ (p forces ϕ) by induction on the complexity of ϕ as follows:

1. $p \nVdash \bot$,
2. $p \Vdash s = t$, iff $\mathbb{N} \models s = t$, for s,t closed terms,
3. $p \Vdash \alpha(t)$ for some closed term t, iff $p(t) = 1$,
4. $p \Vdash \neg\psi$, iff for every $q \in \mathcal{P}$ such that $p \subseteq q$, $q \nVdash \psi$,
5. $p \Vdash \psi \vee \eta$, iff $p \Vdash \psi$ or $p \Vdash \eta$,
6. $p \Vdash \psi \wedge \eta$, iff $p \Vdash \neg(\neg\psi \vee \neg\eta)$,
7. $p \Vdash \exists x \psi(x)$, iff there exists a $n \in \mathbb{N}$ such that $p \Vdash \psi(n)$, and
8. $p \Vdash \forall x \psi(x)$, iff $p \Vdash \neg\exists x \neg\psi(x)$.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

For the next theorem we use the forcing notion

$$\mathcal{P} := \{p : p \text{ is a finite partial function from } \mathbb{N} \text{ to } \{0,1\}\}.$$

To prove the next theorem, we use an idea from [15] and an idea from [21].

In the rest of the paper we use the notation $[n] = \{0, 1, \dots, n\}$. Also, by $t_A(n)$ for some computational machine A (FP functions, Σ^b_i relations, etc.) we mean the time complexity of A on inputs with length of n.

THEOREM 5.1. *There exists an oracle $\mathcal{V}$ such that* $\mathsf{DisjCoNP}^{\mathcal{V}}$ *is true, but* $\mathsf{E}^{\mathcal{V}} = \mathsf{NE}^{\mathcal{V}}$.

PROOF. Let $\{(\phi_i, \psi_i, R_i)\}_{i\in\mathbb{N}}$ be an enumeration of $\Pi^b_1(\alpha) \times \Pi^b_1(\alpha) \times \mathsf{FP}^\alpha$. We want to construct a sequence $p_0 \subseteq p_1 \subseteq p_2 \subseteq \cdots$ of $\mathcal{P}$ such that $\mathcal{V} = \bigcup_i p_i^{-1}(1)$ and $\mathsf{DisjCoNP}^{\mathcal{V}}$ is true, but $\mathsf{E}^{\mathcal{V}} = \mathsf{NE}^{\mathcal{V}}$ if α is interpreted by $\mathcal{V}$.

For every i define the following $\Pi^b_1(\alpha)$ sets:

1. $L^1_i := \{w : \forall |y| = |w|(2\langle i, 1, w, y\rangle \in \alpha)\}$ and
2. $L^2_i := \{w : \forall |y| = |w|(2\langle i, 2, w, y\rangle \in \alpha)\}$.

For every i, let r_i be the first index of occurrence of (ϕ_i, ψ_i) in the enumeration $\{(\phi_i, \psi_i, R_i)\}_{i\in\mathbb{N}}$. We want to construct $\mathcal{V}$ such that for every i, either (ϕ_i, ψ_i) is not disjoint or $(L^1_{r_i}, L^2_{r_i})$ is disjoint and it is not reducible to (ϕ_i, ψ_i) by R_i.

Let L^R_{NE} be the relativized version of the NE-complete problem defined in Proposition 4.1 and $\mathsf{U}^R_{\mathsf{NP}}(x)$ be a $\Sigma^b_1(\alpha)$ predicate such that

$$(\mathbb{N}, A) \models \forall n(n \in L^R_{\mathsf{NE}} \leftrightarrow \mathsf{U}^R_{\mathsf{NP}}(2^n)),$$

for every A. Let $t_{\mathsf{U}^R_{\mathsf{NP}}}(n) \leq n^c + c$ for some $c > 0$. We want to code membership of L^R_{NE} in $\mathcal{V}$ to ensure that $\mathsf{E}^{\mathcal{V}} = \mathsf{NE}^{\mathcal{V}}$. For this matter it is sufficient to have:

$$(\mathbb{N}, \mathcal{V}) \models \forall n(n \in L^R_{\mathsf{NE}} \leftrightarrow 2^{(n+1)^c+c} + 1 \in \alpha).$$

Note that $\mathsf{U}^R_{\mathsf{NP}}(2^n)$ cannot query $2^{(n+1)^c+c} + 1$ and moreover $2^{(n+1)^c+c} + 1$ is computable in polynomial time from the input 2^n.

To satisfy the above requirements, we construct every p_i with the following properties:

- $\mathbf{P}_1$. For every n, if $2^{(n+1)^c+c} + 1 \in \mathsf{Dom}(p_i)$, then $\mathsf{U}^R_{\mathsf{NP}}(2^n)$ is true iff $p_i(2^{(n+1)^c+c} + 1) = 1$ relative to p_i.
- $\mathbf{P}_2$. For every $j \leq i$, (ϕ_j, ψ_j) is not disjoint relative to p_i or (ϕ_j, ψ_j) is not reducible to $(L^1_{r_j}, L^2_{r_j})$ by R_j relative to p_i.
- $\mathbf{P}_3$. For every $j \leq i$, $p_i \not\Vdash \exists x(x \in L^1_{r_j} \wedge x \in L^2_{r_j})$.

Suppose we have constructed $p_{i-1} : \mathsf{Dom}(p_{i-1}) \to \{0,1\}$ such that it satisfies $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$. Let m be big enough (we compute how big m should be) such that

$$\max(t_{\phi_i}(m), t_{\psi_i}(m), t_{R_i}(m)) \leq m^d + d.$$

Define $p_{i-1} \subseteq q$ as follow:

1. $\mathsf{Dom}(q) \subseteq [2^{m^d+d}]$,
2. $\{2\langle r_i, v, x, y\rangle : |x| = |y| = m, v \in \{1,2\}\} \cap \mathsf{Dom}(q) = \varnothing$,

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

3. $(\mathrm{Dom}(q) \setminus \mathrm{Dom}(p_{i-1})) \cap \{2^{(n+1)^c+c}+1 : n \in \mathbb{N}\} = \varnothing$, and
4. $(\{2\langle a, v, x, y\rangle : a, x, y \in \mathbb{N}, v \in \{1,2\}, |x| = |y|, |x| \neq m\} \cap [2^{m^d+d}]) \setminus \mathrm{Dom}(p_{i-1}) \subseteq q^{-1}(0)$.

So we made sure that $[2^{m^d+d}] \setminus \mathrm{Dom}(q)$ are the set of the numbers which we need for encoding and diagonalization. Now we want to extend q to ensure the coding requirement. Let $u_0 = q$. For each $j > 0$ such that $2^{(j+1)^c+c}+1 < 2^{m^d+d}$ we construct u_j by iterating the following rules:

1. If $2^{(j+1)^c+c}+1 \in \mathrm{Dom}(u_{j-1})$, then put $u_j = u_{j-1}$,
2. otherwise,
   (a) if $u_{j-1} \Vdash \neg \mathsf{U}^R_{\mathsf{NP}}(2^j)$, put $u_j = u_{j-1} \cup \{(2^{(j+1)^c+c}+1, 0)\}$ and
   (b) otherwise, extend u_{j-1} to u_j such that:
      i. $u_j \Vdash \mathsf{U}^R_{\mathsf{NP}}(2^j)$,
      ii. $2^{(j+1)^c+c}+1 \in u_j^{-1}(1)$, and
      iii. $|u_j \setminus u_{j-1}| \leq (j+1)^c + c + 1$, we can force this condition because we only need to know the queries of $\mathsf{U}^R_{\mathsf{NP}}(2^j)$ on its accepting path.

Let q' be the unions of u_j for $2^{(j+1)^c+c}+1 < 2^{m^d+d}$. For each x such that $|x| = m$, define $S_x = \{2\langle r_i, v, x, y\rangle : |y| = m, v \in \{1,2\}\}$. Let $k = |\{j \in \mathbb{N} : 2^{(j+1)^c+c}+1 < 2^{m^d+d}\}|$, therefore we have:

$$|q' \setminus q| \leq \sum_{j=0}^{k-1} (j+1)^c + c + 1 \leq k(k^c + c + 1).$$

Because $k \leq (m^d + d - c)^{\frac{1}{c}}$, we have $|q' \setminus q| \leq (m^d + d - c)^{\frac{1}{c}}(m^d + d + 1)$. If m is big enough, then

$$\max\{(m^d + d - c)^{\frac{1}{c}}(m^d + d + 1), 3(m^d + d)\} < 2^m,$$

which means there exists a z with length m such that $S_z \cap \mathrm{Dom}(q') = \varnothing$. By our construction $q' \nVdash \exists x(x \in L^1_{r_i} \wedge x \in L^2_{r_i})$. Now we have enough room to extend q' in such a way that either (ϕ_i, ψ_i) is not disjoint or $(L^1_{r_i}, L^2_{r_i})$ is not reducible to (ϕ_i, ψ_i) by R_i. We compute $R_i(z)$ and answer new oracle questions by the following rule:

1. For every oracle question y, if $y \in S_z$, then accept y and put y in $\mathcal{A}$,
2. if $(y, 1) \in q'$ accept y, and
3. otherwise, reject y.

Let $R_i(z) = z^*$. Let $\mathcal{P}^* \subseteq \mathcal{P}$ such that for every $u \in \mathcal{P}^*$, the following properties are true:

1. $\mathrm{Dom}(u) \subseteq [2^{m^d+d}]$,
2. $u|_{\mathrm{Dom}(q')} = q'$,
3. $\mathcal{A} \subseteq u^{-1}(1)$,
4. $u^{-1}(0) \cap S_z = \varnothing$, and
5. $|\mathrm{Dom}(u) \cap S_z| \leq 2(m^d + d)$.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

Now there are two cases that can occur:

1. If for every $u \in \mathcal{P}^*$, $u \nVdash \neg\phi_i(z^*)$ and also $u \nVdash \neg\psi_i(z^*)$, then define $p' : [2^{m^d+d}] \to \{0,1\}$ by the following definition:

$$p'(c) = \begin{cases} q'(c), & c \in \mathsf{Dom}(q'), \\ 1, & c \in S_z, \\ 0, & \text{o.w.} \end{cases}$$

   Note that $p' \nVdash \neg\phi_i(z^*)$ and also $p' \nVdash \neg\psi_i(z^*)$, because if for example $p' \Vdash \neg\phi_i(z^*)$, then there exists a subset $F \subseteq [2^{m^d+d}]$ such that $p'|_F \in \mathcal{P}^*$ and $p'|_F \Vdash \neg\phi_i(z^*)$ which contradicts our assumption. Hence $p' \nVdash \neg\phi_i(z^*)$ and also $p' \nVdash \neg\psi_i(z^*)$, but this implies $p' \Vdash \phi_i(z^*) \wedge \psi_i(z^*)$, because p' has answers for the oracle questions for all of the numbers with length of less than $m^d + d + 1$. This means that ϕ_i and ψ_i are not disjoint relative to our construction and we define p_i as p'.
2. Otherwise, without loss of generality we can assume that there exists a $u \in \mathcal{P}^*$ such that $u \Vdash \neg\phi_i(z^*)$. Let $S = \{2\langle r_i, 1, z, y\rangle : |y| = m\}$ and define p_i as a condition by the following properties:
   (a) $\mathsf{Dom}(p_i) = [2^{m^d+d}]$,
   (b) $u \subseteq p_i$,
   (c) $S \subseteq p_i^{-1}(1)$, and
   (d) $[2^{m^d+d}] \setminus (\mathsf{Dom}(u) \cup S) \subseteq p_i^{-1}(0)$.
   Therefore, we have the following facts:
   (a) $p_i \Vdash \neg\phi_i(z^*)$ and
   (b) $p_i \Vdash z \in L^1_{r_i}$.
   This implies that $(L^1_{r_i}, L^2_{r_i})$ is not reducible to (ϕ_i, ψ_i) by R_i, relative to our construction.

The above construction guarantees that p_i satisfies $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$ which completes the proof. ⊣

It is worth mentioning that the previous oracle construction still works if want to construct an oracle such that DisjCoNP does not have a complete problem with respect to sub-exponential reductions. Hence by the explanations before Theorem 4.2, the conjectures of Figure 1 are still true with respect to some oracles even if we use sub-exponential reductions and simulations.

In the rest of the paper, we want to construct an oracle $\mathcal{W}$ such that $\mathsf{TFNP}^{\mathcal{W}} = \mathsf{FP}^{\mathcal{W}}$, but there is no optimal proof system for $\mathsf{TAUT}^{\mathcal{W}}$. We use the Kolmogorov generic construction idea that is presented in [6] and also use an idea from [3]. Here we borrow definitions and notations from [6]. As binary strings can code natural numbers and vice versa, we use both natural numbers and strings in the rest of the paper without loss of generality.

Definition 5.2. For every partial computable function $F(x, y)$ and every $x, y \in \{0,1\}^*$, the Kolmogorov complexity of x conditional to y with respect to F, which is denoted as $C_F(x|y)$, is defined as follows:

$$C_F(x|y) := \min\{|e| : e \in \{0,1\}, F(e, y) = x\}.$$

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

We say that $C_F(x|y)$ for some partial computable function $F(x, y)$ is a universal method iff for every partial computable $G(x, y)$, there exists a constant k such that

$$\forall x, y \in \{0,1\}^*(C_F(x|y) \leq C_G(x|y) + k).$$

According to the Solomonoff–Kolmogorov Theorem there exists a universal method. We denote it by $C(x|y)$. Also, we define the unconditional Kolmogorov complexity of x with $C(x) := C(x|\lambda)$ in which λ is the empty string. Here we list some properties of Kolmogorov complexity that are stated in [6].

1. For all x and y, $C(x|y) \leq C(x) + O(1)$.
2. There exists a constant k such that for all x, $C(x) \leq |x| + k$.
3. For all n and m, there is an n bit string x such that $C(x) \geq n - m$. In particular, for every n there is an n bit string x such that $C(x) \geq n$. Such strings are called incompressible.
4. For every computable function $f(x_1, \ldots, x_n)$,

$$C(f(x_1, \ldots, x_n)) \leq 2|x_1| + 2|x_2| + \ldots + 2|x_{n-1}| + |x_n| + O(1).$$

For every $n > 0$ fix a $n2^n$ bit string Z_n such that $C(Z_n) \geq n2^n$. Divide Z_n into 2^n string z_1^n to $z_{2^n}^n$, each of length n and define $Y_n := \{\llcorner\langle i, z_i^n\rangle\lrcorner : i \in \{0,1\}^n\}$. Then $\mathcal{K}$ is $\bigcup_{n\in\mathbb{N}} Y_{2_n^1}$. We define the forcing notion

$$\mathcal{P}_K := \{p : p \text{ is a partial function from } \mathcal{K} \text{ to } \{0,1\}, \mathcal{K} \setminus \mathsf{Dom}(p) \text{ is infinite}\}.$$

THEOREM 5.2. *There exists an oracle $\mathcal{W}$ such that there is no optimal proof system for $\mathsf{TAUT}^{\mathcal{W}}$, but $\mathsf{TFNP}^{\mathcal{W}} = \mathsf{FP}^{\mathcal{W}}$.*

PROOF. Following the argument in [6], we construct an oracle $\mathcal{W}$ such that there is no optimal proof system for $\mathsf{TAUT}^{\mathcal{W}}$, but $\mathsf{TFNP}^{\mathcal{W}} = \mathsf{FP}^{\mathcal{W}}$, assuming $\mathsf{FP} = \mathsf{FPSPACE}$. As we see, the oracle construction still works if we first relativize things with a PSPACE-complete set H and then construct $\mathcal{W}$ with the desired properties. Note that relativizing to H implies $\mathsf{FP}^H = \mathsf{FPSPACE}^H$ and hence we are free from the assumption $\mathsf{FP} = \mathsf{FPSPACE}$. Also, note that relativizing first to H and then relativizing to $\mathcal{W}$ is equivalent to relativizing with $H \oplus \mathcal{W}$ in which $A \oplus B := \{2n : n \in A\} \cup \{2n+1 : n \in B\}$. Let $\{f_i(x)\}_{i\in\mathbb{N}}$ and $\{(r_i, \phi_i(x, y))\}_{i\in\mathbb{N}}$ be enumerations of $\mathsf{FP}(\alpha)$ functions and $\mathbb{N} \times \Delta_1^b(\alpha)$ in which $\phi_i(x, y)$ defines a polynomial time computable relation with access to α. In the rest of the proof we construct a sequence $p_0 \subseteq p_1 \subseteq \cdots$ of $\mathcal{P}_K$ such that $\mathcal{W} = \bigcup_i p_i^{-1}(1)$ and there is no optimal proof system for $\mathsf{TAUT}^{\mathcal{W}}$, but $\mathsf{TFNP}^{\mathcal{W}} = \mathsf{FP}^{\mathcal{W}}$ if α is interpreted by $\mathcal{W}$. For every $i, k \in \mathbb{N}$ define $\theta_{i,k}$ be the relativized translation of the $\Pi_1^b(\alpha)$ sentence $\forall x \left(x < \overline{2^{3n+3}} \rightarrow \left(x < \overline{2^{3n+2}} \vee \neg\alpha(x)\right)\right)$ in which $n = 2^1_{\langle i,k\rangle}$. Note that there is a fixed natural number t such that $|\theta_{i,k}| \leq (2^1_{\langle i,k\rangle})^t + t$ for every $i, k \in \mathbb{N}$. For every $i, j \in \mathbb{N}$ define:

1. $S_j^i := \{\theta_{i,k} : k \geq j\}$.
2. $B_j^i := Y_{2^1_{\langle i,j\rangle}}$.

Note that $|B_j^i| = 2^{2^1_{\langle i,j\rangle}}$.

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

We want to construct every p_i in such a way that each of them satisfies the following properties:

$\mathbf{P}_1$. There exists a natural number b_i such that for every $a \geq \lfloor i/2 \rfloor + 1$

$$\mathsf{Dom}(p_i) \cap \left(\bigcup_{b_i \leq j} B^a_j \right) = \varnothing.$$

$\mathbf{P}_2$. There exists a finite set A_i such that $A_i \subseteq \mathsf{Dom}(p_i)$ and for every $n \in \mathsf{Dom}(p_i) \setminus A_i$, $p_i(n) = 0$ and moreover if $n \in$.

Suppose we have constructed $p_{i-1} : \mathsf{Dom}(p_{i-1}) \to \{0,1\}$ such that p_{i-1} satisfies $\mathbf{P}_1$ and $\mathbf{P}_2$. We extend p_{i-1} to p_i as follows:

1. If $i = 2a$, then we want to ensure that h_a will not be a proof system or that h_a will not have short proofs for members of the set $S^a_{c_a}$ for some c_a relative to $\mathcal{W}$. Let $t_{h_a}(n) \leq n^d + d$. Choose c_a such that $\mathsf{Dom}(p_{i-1}) \cap \left(\bigcup_{c_a \leq j} B^a_j \right) = \varnothing$ and also for every $n \geq c_a$, $(n^t + t)^{d \log_2(n^t+t)} + d < 2^n$. Now, there are two cases that can happen:
   (a) There is a $p_{i-1} \subseteq q \in \mathcal{P}_K$, some $\theta \in S^a_{c_a}$ and $\pi \in \mathbb{N}$ such that
   $$q \Vdash |\pi| \leq |\theta|^{\log_2 |\theta|} \wedge h_a(\pi) = \theta.$$
   This implies that there is a $p_{i-1} \subseteq q' \in \mathcal{P}_K$ such that $|\mathsf{Dom}(q') \setminus \mathsf{Dom}(p_{i-1})| \leq |\theta|^{d \log_2 |\theta|} + d$ and $q' \Vdash |\pi| \leq |\theta|^{\log_2 |\theta|} \wedge h_a(\pi) = \theta$, because h_a only needs at most $|\theta|^{d \log_2 |\theta|} + d$ query answers from $\mathcal{W}$ on the input π. Let θ be $\theta_{a,k}$ for some k. This means
   $$|\theta|^{d \log_2 |\theta|} + d \leq (m^t + t)^{d \log_2 (m^t+t)} + d < 2^m = |B^a_k|,$$
   in which $m = 2^1_{\langle a,k \rangle}$, hence there is a $z \in B^a_k \setminus \mathsf{Dom}(q')$. Define $p_i := q' \cup \{(z,1)\}$. This implies that h_a relative to $\mathcal{W}$ will not be a proof system for $\mathsf{TAUT}^{\mathcal{W}}$, because it proves $\theta_{a,k}$, but $\theta_{a,k}$ is not a tautology relative to $\mathcal{W}$ and
   (b) otherwise, we define $p_i := p_{i-1} \cup \{(x,0) : \exists k \in \mathbb{N}(k \geq c_a \wedge x \in B^a_k)\}$. Note that in this case, for every $\theta \in S^a_{c_a}$, there is no $|\theta|^{\log_2 |\theta|}$ length proof of θ in h_a relative to $\mathcal{W}$.
   So by the construction of p_i we ensured that h_a is not a proof system or h_a is not an optimal proof system for $\mathsf{TAUT}^{\mathcal{W}}$, because $S^a_{c_a}$ is polynomial time decidable.
2. If $i = 2a + 1$, then we want to ensure that $(n^{r_a} + r_a, \phi_a(x,y))$ will not define a TFNP problem relative to $\mathcal{W}$ or it can be computed by some function in $\mathsf{FP}^{\mathcal{W}}$. The construction in this case is very easy. If there is a $p_{i-1} \subseteq q \in \mathcal{P}_K$ such that $q \Vdash \exists x \forall y (|y| \leq |x|^{r_a} + r_a \to \neg \phi_a(x,y))$, then there is some $p_{i-1} \subseteq q' \in \mathcal{P}_K$ such that $|\mathsf{Dom}(q') \setminus \mathsf{Dom}(p_{i-1})|$ is finite and $q' \Vdash \exists x \forall y (|y| \leq |x|^{r_a} + r_a \to \neg \phi_a(x,y))$. In this case we define $p_i := q'$, otherwise if there is no such extension, then we define $p_i := p_{i-1}$.

It is easy to see that in this construction p_i satisfies $\mathbf{P}_1$ and $\mathbf{P}_2$.

To complete the proof we need to show that $\mathsf{TFNP}^{\mathcal{W}} = \mathsf{FP}^{\mathcal{W}}$. Suppose $(n^{r_a} + r_a, \phi_a(x,y))$ defines a TFNP problem relative to $\mathcal{W}$. Now we want to show there is

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

a function $f \in \mathsf{FP}^{\mathcal{W}}$ such that it solves $(n^{r_a} + r_a; \phi_a(x, y))$. Let $t_{\phi_a}(x, y) \leq (|x| + |y|)^b + b$, then on input u with solution v, $\phi_a(u, v)$ asks at most $(|u| + |u|^{r_a} + r_a)^b + b$ questions from $\mathcal{W}$. Choose e such that for all n, $(n + n^{r_a} + r_a)^b + b \leq n^e + e$. The function f works as follows on input x:

Let $m = 2_k^1$ be the biggest tower of two such that $m \leq 4|x|^{2e}$. Note that to compute a solution of this problem we only need to know the oracle answers for members $\bigcup_{i \leq m} Y_i$. First, f asks the value of $\mathcal{W}$ for every member of $A_{2a} \cup \bigcup_{i \leq \log_2 m} Y_i$ and puts the answers in G. Then starting by $Q_1 := \varnothing$ proceeds with the following procedure:

- In the i'th iteration, using the power of PSPACE (we assumed that FP = FPSPACE) find the least v_i such that $|v_i| \leq |x|^{r_a} + r_a$ such that $\phi_a(x, v_i)$ is true relative to $G \cup Q_i$. If $\phi_a(x, v_i)$ is true relative to $\mathcal{W}$, then halt and output v_i, otherwise there is a $u_i \in (\mathcal{W} \cap Y_m) \setminus Q_i$ such that it is the first number in which it is queried in the computation of $\phi_a(x, v_i)$ relative to $\mathcal{W}$ such that $u_i \in \mathcal{W}$, but $u_i \notin Q_i$. Define $Q_{i+1} = Q_i \cup \{u_i\}$ and repeat this procedure.

First, note that in every iteration, this procedure indeed finds a v such that relative to $G \cup Q_i$, $\phi_a(x, v)$ holds, because if this is not the case, then we can find a condition $p_{2a} \subset q \in \mathcal{P}_K$ such that $G \cup Q_i \subseteq q^{-1}(1)$ and hence q forces that $(n^{r_a} + r_a, \phi_a(x, y))$ is not a TFNP problem which contradicts with the construction of p_{2a+1} (if $Y_m \cap \mathcal{W} = \varnothing$, then we should find the solution of the problem relative to $\mathcal{W}$ in the first iteration, hence the construction of the previous conditions which ensures some proof systems are not optimal, will not cause a problem in finding such a q). After some iterations, f will find a solution of this TFNP problem relative to $\mathcal{W}$. If we prove that the number of iterations are polynomial in $|x|$, then we are done. Suppose after l'th iteration we find the solution. This means that $|Q_l| = l - 1$. Let $l' = l - 1$. Note that for every $j < l$, u_j can be described by the code of the polynomial time computable relation $\phi_a(x, y)$, x, $G \cup Q_j$ and an $e \log_2 |x|$ bit string which shows the order number of u_j among the queries of $\phi_a(x, v_j)$, hence Q_l can be described by a string of length $l'(e \log_2 |x|) + O(m \log_2 m) + 2|x| + O(1)$ (note that $G \setminus A_{2a}$ has at most $m + \log_2 m + \log_2 \log_2 m + \cdots$ of strings of length at most $\log_2 m$, hence G can be described by a string of length $O(m \log_2 m)$ bits). Let p be the concatenation of all y's from $\llcorner \langle i, y \rangle \lrcorner \in Y_m \setminus Q_l$ according to the order on i's, hence $|p| = m(2^m - l')$. Note that Z_m can be described using p by inserting the second component of members of Q_l in places that the first component refer to, hence by the fact that

$$C(Q_l) \leq l'(e \log_2 |x|) + O(m \log_2 m) + 2|x| + O(1),$$

we have:

$$m2^m \leq C(Z_m) \leq m(2^m - l') + 2l'(e \log_2 |x|) + O(m \log_2 m) + 4|x| + O(1).$$

This implies

$$l'(m - 2e \log_2 |x|) \leq O(m \log_2 m) + 4|x| + O(1).$$

Note that by definition of m, $4|x|^{2e} < 2^m$, hence $2 + 2e \log_2 |x| < m$. This implies $m - 2e \log_2 |x| > 2$, hence

$$2l' \leq O(m \log_2 m) + 4|x| + O(1),$$

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

which means

$$l \leq O(4|x|2e \log_2(4|x|^{2e})) + 2|x| + O(1),$$

and this completes the proof. ⊣

It is worth mentioning that the forcing notion that was used in [6] is a finite condition forcing, but the forcing notion $\mathcal{P}_K$ permits us to have conditions with an infinite domain. Note that we essentially use this property of $\mathcal{P}_K$ in our construction. We do not know whether optimal proof systems for TAUT exist relative to the original oracle of [6] .

The existence of oracles $\mathcal{V}$ and $\mathcal{W}$ imply several separations between the conjectures of Figure 1. The following corollary shows several independence results (not all of the separations) of the conjectures in Figure 1.

COROLLARY 5.3. *Define the following sets*:

1. $A := \{\mathsf{CON}, \mathsf{CON}^{\mathsf{N}}\}$ *and*
2. $B := \{\mathsf{SAT\text{-}conj}, \mathsf{TFNP\text{-}conj}, \mathsf{DisjCoNP\text{-}conj}\}$.

Then for every conjecture $Q \in A$ and every conjecture $Q' \in B$, Q and Q' do not imply each other in relativized worlds.

PROOF. The corollary follows from Theorems 5.1 and 5.2. ⊣

As we mentioned in the Introduction, the result of Dose in [12] and Theorem 5.2 are incomparable. It is proved in [14] that $\mathsf{TFNP} = \mathsf{FP}$ is equivalent to the statement:

- The standard proof system for SAT is p-optimal.

Note that relative to Dose's oracle, DisjNP-conj is true and hence $\mathsf{CON}^{\mathsf{N}}$ is also true, but relative to $\mathcal{W}$, the standard proof system for SAT is p-optimal which is a stronger statement than ¬SAT-conj, hence Dose's oracle and $\mathcal{W}$ cannot be compared.

§6. Appendix. As we saw in Section 2, the logical conjectures discussed in Section 2 are of the following form:

- For every $T \in \mathcal{T}$ there exists a sentence ϕ that does not have a T-proof with some properties.

The above form works for all of the conjectures that we discussed, except for TFNP-conj. The logical form of TFNP-conj conjecture uses $\mathsf{TFNP}^*(T)$ instead of $\mathsf{TFNP}(T)$. Here we want to investigate what happens if we use $\mathsf{TFNP}(T)$. This new conjecture, which we call TFNP^w-conj, is weaker than TFNP-conj. The next proposition shows that it is at least as strong as SAT-conj.

PROPOSITION 6.1. *If for every $T \in \mathcal{T}$ we have $\mathsf{TFNP}(T) \neq \mathsf{TFNP}$ i.e., if TFNP^w-conj holds true, then there is no p-optimal proof system for* SAT.

PROOF. Suppose P is a p-optimal proof system for SAT. Define $T := \mathsf{S}^1_2 + \forall x \exists y \mathtt{Sat}(P(x), y)$. Let (p, R) be a TFNP problem and (q, ϕ) be one of its formalizations. Suppose F is a proof system for SAT. Let θ_n be the usual propositional translation of polynomial time computable relation $|y| \leq q(|\bar{n}|) \wedge$

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press

$\phi(\bar{n}, y)$. The proof system P_ϕ for SAT is defined as follows:

$$P_\phi(x) := \begin{cases} F(n), & x = 2n, \\ \theta_n, & x = 2n+1. \end{cases}$$

Because P is a p-optimal proof system, there exists a polynomial time computable function h such that $\mathbb{N} \models \forall x (P(h(x)) = P_\phi(x))$. This implies that

$$\mathbb{N} \models \forall x, y \big((|y| \leq q(|x|) \wedge \phi(x, y)) \equiv \mathtt{Sat}(P(h(2x+1)), f(y)) \big),$$

for some polynomial time computable function f, hence $\mathtt{Sat}(P(h(2x+1)), f(y))$ is another formalization of (p, R). Note that by definition of T we have $T \vdash \forall x \exists y \mathtt{Sat}(P(h(2x+1)), f(y))$ which means $(p, R) \in \mathsf{TFNP}(T)$. ⊣

We cannot prove that TFNP^w-conj implies TFNP-conj, but one way to show that the latter conjecture is probably stronger is to find a $T \in \mathcal{T}$ such that $\mathsf{TFNP}(T) \neq \mathsf{TFNP}^*(T)$. It is conjectured that such a T exists, but we observed that the existence of such a T implies $\mathsf{TFNP} \neq \mathsf{FP}$, hence proving that a $T \in \mathcal{T}$ exists such that $\mathsf{TFNP}(T) \neq \mathsf{TFNP}^*(T)$ unconditionally is hard. We need the following lemma to prove the previous implication.

LEMMA 6.2. $\mathsf{TFNP}(\mathsf{S}^1_2) = \mathsf{FP}$.

PROOF. By the fact that Σ^b_1 definable functions of S^1_2 is polynomial time computable we get $\mathsf{TFNP}(\mathsf{S}^1_2) \subseteq \mathsf{FP}$, so it is sufficient to prove $\mathsf{FP} \subseteq \mathsf{TFNP}(\mathsf{S}^1_2)$. Let (p, R) be a TFNP problem which can be solved by the polynomial time computable function f. Let ϕ be the Δ^b_1 formalization of f in S^1_2. Additionally, let (q, ψ) be a formalization of (p, R). Note that $(q, \psi \vee \phi)$ is a formalization of (p, R) and also $\mathsf{S}^1_2 \vdash \forall x \exists y (|y| \leq q(|x|) \wedge (\psi(x, y) \vee \phi(x, y))$, hence $(p, R) \in \mathsf{TFNP}(\mathsf{S}^1_2)$, which implies $\mathsf{FP} \subseteq \mathsf{TFNP}(\mathsf{S}^1_2)$. ⊣

COROLLARY 6.3. *If there exists a* $T \in \mathcal{T}$ *such that* $\mathsf{TFNP}(T) \neq \mathsf{TFNP}^*(T)$, *then* $\mathsf{TFNP} \neq \mathsf{FP}$.

PROOF. Suppose TFNP is equal to FP, hence for every $T \in \mathcal{T}$, $\mathsf{TFNP}(T) \subseteq \mathsf{FP}$, which implies $\mathsf{TFNP}^*(T) \subseteq \mathsf{FP}^{\mathsf{FP}} = \mathsf{FP}$. Also, by definition of T and Lemma 6.2, $\mathsf{FP} = \mathsf{TFNP}(\mathsf{S}^1_2) \subseteq \mathsf{TFNP}(T)$, hence $\mathsf{TFNP}(T) = \mathsf{TFNP}^*(T) = \mathsf{FP}$, which completes the proof. ⊣

Acknowledgments. We are indebted to Pavel Pudlák for many invaluable discussions that we have had about this work. We are also grateful to him for his careful readings of the drafts of this paper, his useful comments and suggestions about it, and also pointing out many small errors that led to improvements in its presentation. Additionally, we are grateful to Moritz Müller for his careful reading of the draft of this paper and his useful suggestions. We thank Lance Fortnow and Michael Rathjen for answering our questions. We thank Grace Kenney and Neil Thapen for pointing out some English errors in the draft of this paper and also we are grateful to Susanna de Rezende for answering our English question. We are also indebted to anonymous referees for pointing out many small errors and suggesting improvements of presentation. This research was partially supported by the ERC Advanced Grant no. 339691 (FEALORA). The current affiliations of the author

are Institute of Mathematics of the Czech Academy of Sciences and Faculty of Mathematics and Physics of Charles University.

REFERENCES

[1] M. Ajtai, Σ^1_1-*formulae on finite structures*. ***Annals of Pure and Applied Logic***, vol. 24 (1983), pp. 1–48.

[2] ———, *The complexity of the pigeonhole principle*. ***Combinatorica***, vol. 14 (1994), no. 4, 417–433.

[3] S. Ben-David and A. Gringauze, *On the existence of propositional proof systems and oracle-relativized propositional logic*, ECCC technical report no. TR98-021, 1998.

[4] O. Beyersdorff and Z. Sadowski, *Characterizing the existence of optimal proof systems and complete sets for promise classes*, ***Computer Science – Theory and Applications: Fourth International Computer Science Symposium in Russia, CSR 2009***, Springer, Berlin, 2009, pp. 47–58.

[5] H. Buhrman, S. Fenner, L. Fortnow, and D. van Melkebeek, *Optimal proof systems and sparse sets*, ***STACS 2000: 17th Annual Symposium on Theoretical Aspects of Computer Science***, Springer, Berlin, 2000, pp. 407–418.

[6] H. Buhrman, L. Fortnow, M. Koucký, J. D. Rogers, and N. Vereshchagin, *Does the polynomial hierarchy collapse if onto functions are invertible*? ***Theory of Computing Systems***, vol. 46 (2010), no. 1, pp. 143–156.

[7] S. Buss, V. Kabanets, A. Kolokolova, and M. Koucký, *Expander construction in* VNC^1. ***Annals of Pure and Applied Logic***, vol. 171 (2020), no. 7, p. 39.

[8] S. R. Buss, ***Bounded Arithmetic***, Bibliopolis, Napoli, 1986.

[9] ———, *Bounded arithmetic and propositional proof complexity*, ***Logic of Computation***, Proceedings of the NATO ASI, Springer, Berlin, 1997, pp. 67–121.

[10] Y. Chen and J. Flum, *On p-optimal proof systems and logics for PTIME*, ***Automata, Languages and Programming: 37th international colloquium, ICALP 2010***, Springer, Berlin, 2010, pp. 321–332.

[11] S. A. Cook and R. A. Reckhow, *The relative efficiency of propositional proof systems*, this Journal, vol. 44 (1979), pp. 36–50.

[12] T. Dose, *An oracle separating conjectures about incompleteness in the finite domain*. ***Theoretical Computer Science***, vol. 809 (2020), pp. 466–481.

[13] S. Fenner, L. Fortnow, S. A. Kurtz, and L. Li, *An oracle builder's toolkit*. ***Information and Computation***, vol. 182 (2003), no. 2, pp. 95–136.

[14] S. A. Fenner, L. Fortnow, A. V. Naik, and J. D. Rogers, *Inverting onto functions*. ***Information and Computation***, vol. 186 (2003), no. 1, pp. 90–103.

[15] C. Glasser, A. L. Selman, S. Sengupta, and L. Zhang, *Disjoint NP-pairs*. ***SIAM Journal on Computing***, vol. 33 (2004), no. 6, pp. 1369–1416.

[16] P. Hájek and P. Pudlák, ***Metamathematics of First-Order Arithmetic***, Springer-Verlag, Berlin, 1993.

[17] E. Khaniki, *New relations and separations of conjectures about incompleteness in the finite domain*, preprint, 2019, arXiv:1904.01362.

[18] J. Krajíček, ***Forcing with Random Variables and Proof Complexity***, London Mathematical Society Lecture Notes Series, vol. 382, Cambridge University Press, Cambridge, 2011.

[19] ———, ***Proof Complexity***, Encyclopedia of Mathematics and its Applications, vol. 170, Cambridge University Press, Cambridge, 2019.

[20] J. Krajíček and P. Pudlák, *Propositional proof systems, the consistency of first order theories and the complexity of computations*. this Journal, vol. 54 (1989), no. 3, pp. 1063–1079.

[21] S. A. Kurtz, *Sparse sets in NP-P: relativizations*. ***SIAM Journal on Computing***, vol. 14 (1985), pp. 113–119.

[22] J. Messner and J. Torán, *Optimal proof systems for propositional logic and complete sets*, ***STACS 98***, Springer, Berlin, 1998, pp. 477–487.

[23] P. Pudlák. ***Logical Foundations of Mathematics and Computational Complexity: A Gentle Introduction***, Springer, Cham, 2013.

[24] ———, *On the complexity of finding falsifying assignments for Herbrand disjunctions*. ***Archive for Mathematical Logic***, vol. 54 (2015), nos. (7–8), pp. 769–783.

[25] ———, *Incompleteness in the finite domain*. ***The Bulletin of Symbolic Logic***, vol. 23 (2017), no. 4, pp. 405–441.

[26] S. Riis, *Count*(q) *versus the pigeon-hole principle*. ***Archive for Mathematical Logic***, vol. 36 (1997), no. 3, pp. 157–188.

DEPARTMENT OF MATHEMATICAL SCIENCES
SHARIF UNIVERSITY OF TECHNOLOGY
TEHRAN, IRAN
E-mail: e.khaniki@gmail.com

https://doi.org/10.1017/jsl.2021.99 Published online by Cambridge University Press