NP-Completeness, Proof Systems, and Disjoint NP-Pairs

Titus Dose
Julius-Maximilians-Universität Würzburg, Germany

Christian Glaßer
Julius-Maximilians-Universität Würzburg, Germany

Abstract

The article investigates the relation between three well-known hypotheses.

$\mathrm{H}_{\mathrm{union}}$: the union of disjoint $\leq_{\mathrm{m}}^{\mathrm{p}}$-complete sets for NP is $\leq_{\mathrm{m}}^{\mathrm{p}}$-complete

$\mathrm{H}_{\mathrm{opps}}$: there exist optimal propositional proof systems

$\mathrm{H}_{\mathrm{cpair}}$: there exist $\leq_{\mathrm{m}}^{\mathrm{pp}}$-complete disjoint NP-pairs

The following results are obtained:

- The hypotheses are pairwise independent under relativizable proofs, except for the known implication $\mathrm{H}_{\mathrm{opps}} \Rightarrow \mathrm{H}_{\mathrm{cpair}}$.
- An answer to Pudlák's question for an oracle relative to which $\neg\mathrm{H}_{\mathrm{cpair}}$, $\neg\mathrm{H}_{\mathrm{opps}}$, and UP has $\leq_{\mathrm{m}}^{\mathrm{p}}$-complete sets.
- The converse of Köbler, Messner, and Torán's implication $\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{coNEE} \Rightarrow \mathrm{H}_{\mathrm{opps}}$ fails relative to an oracle, where $\mathrm{NEE} \stackrel{df}{=} \mathrm{NTIME}(2^{O(2^n)})$.
- New characterizations of $\mathrm{H}_{\mathrm{union}}$ and two variants in terms of coNP-completeness and p-producibility of the set of hard formulas of propositional proof systems.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness; Theory of computation → Proof complexity; Theory of computation → Oracles and decision trees

Keywords and phrases NP-complete, propositional proof system, disjoint NP-pair, oracle

Digital Object Identifier 10.4230/LIPIcs.STACS.2020.9

Related Version Full version available as [10]: `https://eccc.weizmann.ac.il/report/2019/050/`.

Funding *Titus Dose*: supported by the German Academic Scholarship Foundation.

1 Introduction

The three hypotheses studied in this paper came up in the context of fascinating questions. The first one states a simple closure property for the class of NP-complete sets. The second one addresses the existence of optimal propositional proof systems. It is equivalent to the existence of a finitely axiomatized theory that proves the finite consistency of each finitely axiomatized theory by a proof of polynomial length [25]. The third hypothesis is motivated and also implied by the second one.

Below we explain the context in which these hypotheses came up and discuss further connections to complete sets for promise classes like UP, to the security of public-key cryptosystems, and to complete functions for NPSV, the class of single-valued functions computable by NP-machines. At the end of this section we summarize our results.

Hypothesis H_{union}: unions of disjoint $\leq_m^p$-complete sets for NP are $\leq_m^p$-complete

The beauty of hypothesis $\mathrm{H}_{\text{union}}$ lies in its simplicity. It states that the class of NP-complete sets is closed under unions of disjoint sets. The question of whether $\mathrm{H}_{\text{union}}$ holds was raised by Selman [37] in connection with the study of self-reducible sets in NP.[1]

An interesting example for a union of disjoint NP-complete sets is the Clique-Coloring pair, which is due to Pudlák [31]:

$$C_0 = \{(G,k) \mid G \text{ is a graph that has a clique of size } k\}$$
$$C_1 = \{(G,k) \mid G \text{ is a graph that can be colored with } k-1 \text{ colors}\}$$

The sets are NP-complete and disjoint, since a clique of size k cannot be colored with $k-1$ colors. C_0 and C_1 are P-separable [31], which means that there exists an $S \in \mathrm{P}$, the separator, such that $C_0 \subseteq S$ and $C_1 \subseteq \overline{S}$. The P-separability of C_0 and C_1 is a result based on deep combinatorial arguments by Lovász [26] and Tardos [38]. It implies that $C_0 \cup C_1$ is NP-complete.

Glaßer et al. [14, 17] give several equivalent formulations of $\mathrm{H}_{\text{union}}$ and show that the union of disjoint sets that are $\leq_m^p$-complete for NP is complete with respect to strongly nondeterministic, polynomial-time Turing reducibility. Moreover, the union is also nonuniformly polynomial-time many-one complete for NP under the assumption that NP is not infinitely-often in coNP. Moreover, Glaßer et al. [13] provide sufficient and necessary conditions for $\mathrm{H}_{\text{union}}$ in terms of refuters that distinguish languages $L \in \mathrm{NP}$ with $\mathrm{SAT} \cap L = \emptyset$ from $\overline{\mathrm{SAT}}$.

Hypothesis H_{opps}: there exist optimal propositional proof systems

Cook and Reckhow [6] define a propositional proof system (pps) as a polynomial-time computable function f whose range is TAUT, the set of tautologies. If $f(x) = y$, then x is a proof for y. A pps f is simulated by a pps g, if proofs in g are at most polynomially longer than proofs in f. We say that f is P-simulated by g, if additionally for a given proof in f we can compute in polynomial time a corresponding proof in g. A pps g is optimal (resp., P-optimal) if it simulates (resp., P-simulates) each pps.

The question of whether H_{opps} holds was raised by Krajíček and Pudlák [25] in an exciting context:[2] Let $\mathrm{Con}_T(n)$ denote the finite consistency of a theory T, which is the statement that T does not have proofs of contradiction of length $\leq n$. Krajíček and Pudlák [25] showed that H_{opps} is equivalent to the statement that there is a finitely axiomatized theory S that proves the finite consistency $\mathrm{Con}_T(n)$ for each finitely axiomatized theory T by a proof of polynomial length in n. In other words, H_{opps} expresses that a weak version of Hilbert's program (to prove the consistency of all mathematical theories) is possible [30].

Krajíček and Pudlák [25] also show that NE = coNE implies H_{opps} and that E = NE implies the existence of P-optimal pps. The converses of these implications do not hold relative to an oracle constructed by Verbitskii [40]. Köbler, Messner, and Torán [24] prove similar implications with weaker assumptions and reveal a connection to promise classes. For $\mathrm{EE} \stackrel{df}{=} \mathrm{DTIME}(2^{O(2^n)})$ and $\mathrm{NEE} \stackrel{df}{=} \mathrm{NTIME}(2^{O(2^n)})$ they show that $\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{coNEE}$ implies H_{opps}, which in turn implies that $\mathrm{NP} \cap \mathrm{SPARSE}$ has $\leq_m^p$-complete sets. Moreover, $\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{EE}$ implies the existence of P-optimal pps, which in turn implies that UP has $\leq_m^p$-complete sets.

[1] The analog of $\mathrm{H}_{\text{union}}$ in computability theory holds [39], since the many-one complete c.e. sets are creative [27].

[2] The analog of H_{opps} in computability theory holds trivially, since there the notion of simulation does not have any bounds for the length of proofs and hence each proof system is optimal.

Sadowski [36] proves that $\mathrm{H_{opps}}$ is equivalent to the statement that the class of all easy subsets of TAUT is uniformly enumerable. Beyersdorff [2, 3, 4, 5] investigates connections between disjoint NP-pairs and pps, and in particular studies the hypotheses $\mathrm{H_{cpair}}$ and $\mathrm{H_{opps}}$. Pudlák [30, 32] provides comprehensive surveys on the finite consistency problem, its connection to propositional proof systems, and related open questions. In a recent paper, Khaniki [23] shows new relations between the conjectures discussed in [32] and constructs two oracles that separate several of these conjectures. In a couple of further papers [9, 8, 7], one of the authors also builds oracles separating several of the conjectures in [32].

Hypothesis $\mathrm{H_{cpair}}$: there exist $\leq_{\mathrm{m}}^{\mathrm{pp}}$-complete disjoint NP-pairs

Even, Selman, and Yacobi [12, 11] show that the security of public-key cryptosystems depends on the computational complexity of certain promise problems. The latter can be written as disjoint NP-pairs, i.e., pairs (A, B) of disjoint sets $A, B \in \mathrm{NP}$. The Clique-Coloring pair mentioned above is an interesting example for a P-separable disjoint NP-pair. Even, Selman, and Yacobi [12, 11] conjecture that every disjoint NP-pair has a separator that is not $\leq_{\mathrm{T}}^{\mathrm{p}}$-hard for NP. If the conjecture holds, then there are no public-key cryptosystems that are NP-hard to crack. Grollmann and Selman [20] observe that secure public-key cryptosystems exist only if P-inseparable disjoint NP-pairs exist.

The question of whether $\mathrm{H_{cpair}}$ holds was raised by Razborov [34] in the context of pps.[3] To explain this connection we need the notions of reducibility and completeness for disjoint NP-pairs. (A, B) polynomial-time many-one reduces to (C, D), written as $(A, B) \leq_{\mathrm{m}}^{\mathrm{pp}} (C, D)$, if there is a polynomial-time computable h such that $h(A) \subseteq C$ and $h(B) \subseteq D$. A disjoint NP-pair (A, B) is $\leq_{\mathrm{m}}^{\mathrm{pp}}$-complete, if each disjoint NP-pair $\leq_{\mathrm{m}}^{\mathrm{pp}}$-reduces to (A, B). Razborov [34] defines for each pps f a corresponding disjoint NP-pair, the canonical pair of f. He shows that the canonical pair of an optimal pps is an $\leq_{\mathrm{m}}^{\mathrm{pp}}$-complete disjoint NP-pair, i.e.,

$$\mathrm{H_{opps}} \;\Rightarrow\; \mathrm{H_{cpair}}. \tag{1}$$

This means that the open question of whether optimal pps exist can be settled by proving that $\leq_{\mathrm{m}}^{\mathrm{p}}$-complete disjoint NP-pairs do not exist. As we will see, (1) is the only nontrivial implication between the three hypotheses and their negations that holds relative to all oracles. For the relationship between $\mathrm{H_{cpair}}$ and $\mathrm{H_{opps}}$ this is shown by Glaßer et al. [16] who construct two oracles such that $\mathrm{H_{cpair}}$ holds relative to both oracles, but $\mathrm{H_{opps}}$ holds relative to the first one and $\neg\mathrm{H_{opps}}$ relative to the second one.

Pudlák [31] further investigates the connection between pps and disjoint NP-pairs and shows that the canonical pair of the resolution proof system is symmetric. Glaßer, Selman, and Sengupta [15] characterize $\mathrm{H_{cpair}}$ in several ways, e.g., by the uniform enumerability of disjoint NP-pairs and by the existence of $\leq_{\mathrm{m}}^{\mathrm{p}}$-complete functions in NPSV. Glaßer, Selman, and Zhang [18] prove that disjoint NP-pairs and pps have identical degree structures. Moreover, they show the following statement, which connects disjoint NP-pairs, pps, and $\mathrm{H_{union}}$ [19]: If $\mathrm{NP} \neq \mathrm{coNP}$ and each disjoint NP-pair (SAT, B) is strongly polynomial-time many-one equivalent to the canonical pair of a pps, then $\mathrm{H_{union}}$ holds.

Our Contribution

The results of this paper improve our understanding on the three hypotheses and their relationships in the following way.

[3] The analog of $\mathrm{H_{cpair}}$ in computability theory holds [35, Ch. 7., Thm XII(c)].

1. *Relativized independence of the hypotheses.* We show that $\mathrm{H}_{\mathrm{union}}$, $\mathrm{H}_{\mathrm{opps}}$, and $\mathrm{H}_{\mathrm{cpair}}$ are pairwise independent under relativizable proofs (except for the known implication $\mathrm{H}_{\mathrm{opps}} \Rightarrow \mathrm{H}_{\mathrm{cpair}}$). For each two of these hypotheses and every combination of their truth values there exists an appropriate oracle, except for $\mathrm{H}_{\mathrm{opps}} \wedge \neg\mathrm{H}_{\mathrm{cpair}}$ which is impossible. The relativized relationships between $\mathrm{H}_{\mathrm{opps}}$ and $\mathrm{H}_{\mathrm{cpair}}$ were settled by Glaßer et al. [16]. The remaining ones are obtained from an oracle by Ogiwara and Hemachandra [28], an oracle by Homer and Selman [22], and three oracles constructed in the present paper.
2. *Answer to a question by Pudlák.* The oracle in Theorem 11 answers a question by Pudlák [32] who asks for an oracle relative to which $\neg\mathrm{H}_{\mathrm{cpair}}$ and UP has $\leq_{\mathrm{m}}^{\mathrm{p}}$-complete sets, i.e., $\mathsf{DisjNP} \not\Rightarrow \mathsf{UP}$ in the notation of [32] (see subsection 4.1 for definitions). In particular, relative to this oracle there are no P-optimal pps, but UP has $\leq_{\mathrm{m}}^{\mathrm{p}}$-complete sets, i.e., $\mathsf{CON} \not\Rightarrow \mathsf{UP}$. This is interesting, since $\mathsf{CON} \Leftarrow \mathsf{UP}$ is a theorem [24].
3. *Possibility of* $\mathrm{H}_{\mathrm{opps}}$ *without* $\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{coNEE}$*.* The oracle constructed in Theorem 12 shows that the converses of the following implications by Krajíček and Pudlák [25] and Köbler, Messner, and Torán [24] fail relative to an oracle. For the implications (a) and (b) this was known by Verbitskii [40], for the other implications this is a new result. It tells us that $\mathrm{H}_{\mathrm{opps}}$ might be possible under assumptions weaker than $\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{coNEE}$.
   - **(a)** [25] $\mathrm{NE} = \mathrm{coNE} \;\Rightarrow\; \mathrm{H}_{\mathrm{opps}}$
   - **(b)** [25] $\mathrm{E} = \mathrm{NE} \;\Rightarrow$ there exist P-optimal pps
   - **(c)** [24] $\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{coNEE} \;\Rightarrow\; \mathrm{H}_{\mathrm{opps}}$, where $\mathrm{NEE} \overset{df}{=} \mathrm{NTIME}(2^{O(2^n)})$
   - **(d)** [24] $\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{EE} \;\Rightarrow$ there exist P-optimal pps, where $\mathrm{EE} \overset{df}{=} \mathrm{DTIME}(2^{O(2^n)})$
4. *Characterization of* $\mathrm{H}_{\mathrm{union}}$*.* We characterize $\mathrm{H}_{\mathrm{union}}$ and two variants (one is weaker, the other one stronger) in several ways. For instance, $\mathrm{H}_{\mathrm{union}}$ (resp., its stronger version) is equivalent to the statement that for each pps, the set of hard formulas is coNP-complete (resp., p-producible). The latter notion was introduced by Hemaspaandra, Hemaspaandra, and Hempel [21] for the study of inverses of NP-problems.

2 Preliminaries

Throughout this paper let Σ be the alphabet $\{0,1\}$. We denote the length of a word $w \in \Sigma^*$ by $|w|$. The empty word is denoted by ε and the i-th letter of a word w for $0 \leq i < |w|$ is denoted by $w(i)$, i.e., $w = w(0)w(1)\cdots w(|w|-1)$. For $k \leq |w|$ let $\mathrm{pr}_k(w) = w(0)\cdots w(k-1)$ be the length k prefix of w. If v is a prefix (resp., proper prefix) of w, then we write $v \sqsubseteq w$ (resp., $v \sqsubsetneq w$). A function $f : \Sigma^* \to \Sigma^*$ is length-increasing, if $|f(x)| > |x|$ for all $x \in \Sigma^*$. $\mathbb{N}$ (resp., $\mathbb{N}^+$) denotes the set of natural numbers (resp., positive natural numbers). The set of primes is denoted by $\mathbb{P} = \{2,3,5,\ldots\}$, the set of primes $\geq k$ by $\mathbb{P}^{\geq k} = \{n \in \mathbb{P} \mid n \geq k\}$. We identify Σ^* with $\mathbb{N}$ via the polynomial-time-computable, polynomial-time-invertible bijection $w \mapsto \sum_{i<|w|}(1+w(i))2^i$, which is a variant of the dyadic encoding. Hence notations, relations, and operations for Σ^* are transferred to $\mathbb{N}$ and vice versa. In particular, $|n|$ denotes the length of $n \in \mathbb{N}$. We eliminate the ambiguity of the expressions 0^i and 1^i by always interpreting them over Σ^*.

Let $\langle\cdot\rangle : \bigcup_{i\geq 0} \mathbb{N}^i \to \mathbb{N}$ be an injective, polynomial-time-computable, polynomial-time-invertible pairing function such that $|\langle u_1,\ldots,u_n\rangle| = 2(|u_1| + \cdots + |u_n| + n)$.

Given two sets A and B, $A - B = \{a \in A \mid a \notin B\}$. The complement of A relative to the universe U is denoted by $\overline{A} = U - A$. The universe will always be apparent from the context.

FP, P, and NP denote standard complexity classes [29]. Define $\mathrm{co}\mathcal{C} = \{A \subseteq \Sigma^* \mid \overline{A} \in \mathcal{C}\}$ for a class $\mathcal{C}$. Let UP denote the set of problems that can be accepted by a non-deterministic polynomial-time Turing machine that on every input x has at most one accepting path and that accepts if and only if there exists an accepting path. TALLY denotes the class $\{A \mid A \subseteq \{0\}^*\}$. We adopt the following notions from Köbler, Messner, and Torán [24] with the remark that in the literature there exist inequivalent definitions for the double exponential time classes EE and NEE. To avoid confusion, we will recall these definitions where appropriate.

$$\begin{array}{rclcrcl} \mathrm{E} & \stackrel{df}{=} & \mathrm{DTIME}(2^{O(n)}) & \qquad & \mathrm{EE} & \stackrel{df}{=} & \mathrm{DTIME}(2^{O(2^n)}) \\ \mathrm{NE} & \stackrel{df}{=} & \mathrm{NTIME}(2^{O(n)}) & \qquad & \mathrm{NEE} & \stackrel{df}{=} & \mathrm{NTIME}(2^{O(2^n)}) \end{array}$$

We also consider all these complexity classes in the presence of an oracle O and denote the corresponding classes by FP^O, P^O, NP^O, and so on. We use the usual oracle model where the length of queries is *not* bounded, e.g., exponential-time machines can ask queries of exponential length.

Let M be an oracle Turing machine. $M^D(x)$ denotes the computation of M on input x with D as an oracle. For an arbitrary oracle D we let $L(M^D) = \{x \mid M^D(x) \text{ accepts}\}$, where as usual if M is nondeterministic, the computation $M^D(x)$ accepts if and only if it has at least one accepting path. For a deterministic polynomial-time oracle Turing transducer F (i.e., a Turing machine computing a function), depending on the context, $F^D(x)$ either denotes the computation of F on input x with D as an oracle or the output of this computation.

If $A, B \in \mathrm{NP}$ and $A \cap B = \emptyset$, then we call (A, B) a disjoint NP-pair. The set of all disjoint NP-pairs is denoted by DisjNP.

We use the following reducibilities for sets $A, B \subseteq \Sigma^*$. $A \le_{\mathrm{m}}^{\mathrm{p}} B$ if there exists an $f \in \mathrm{FP}$ such that $x \in A \Leftrightarrow f(x) \in B$. $A \le_{\mathrm{m,li}}^{\mathrm{p}} B$ if $A \le_{\mathrm{m}}^{\mathrm{p}} B$ via some length-increasing $f \in \mathrm{FP}$. For disjoint NP-pairs (A, B) and (C, D) we define specific reducibilities. $(A, B) \le_{\mathrm{m}}^{\mathrm{pp}} (C, D)$ (resp., $(A, B) \le_{\mathrm{m,li}}^{\mathrm{pp}} (C, D)$) if there exists an $f \in \mathrm{FP}$ (resp., a length-increasing $f \in \mathrm{FP}$) with $f(A) \subseteq C$ and $f(B) \subseteq D$. We use $A \le_{\mathrm{m}}^{\mathrm{pp}} (C, D)$ as an abbreviation for $(A, \overline{A}) \le_{\mathrm{m}}^{\mathrm{pp}} (C, D)$ and analogous notations for other reducibilities.

When we consider reducibilities in the presence of an oracle O, we write $\le_{\mathrm{m}}^{\mathrm{p},O}$, $\le_{\mathrm{m,li}}^{\mathrm{p},O}$, $\le_{\mathrm{m}}^{\mathrm{pp},O}$, and $\le_{\mathrm{m,li}}^{\mathrm{pp},O}$ to indicate that the reduction function has access to O.

For a complexity class $\mathcal{C}$ and some problem A, we say that A is $\le$-hard for $\mathcal{C}$ if for all $B \in \mathcal{C}$ it holds $B \le A$, where $\le$ is some reducibility. A is called $\le$-complete for $\mathcal{C}$ if A is $\le$-hard for $\mathcal{C}$ and $A \in \mathcal{C}$. Let $\mathrm{NPC}_{\mathrm{m}}^{\mathrm{p}}$ (resp., $\mathrm{NPC}_{\mathrm{m,li}}^{\mathrm{p}}$, $\mathrm{NPC}_{\mathrm{m}}^{\text{io-p/poly}}$) be the set of problems that are $\le_{\mathrm{m}}^{\mathrm{p}}$-complete (resp., $\le_{\mathrm{m,li}}^{\mathrm{p}}$-complete, $\le_{\mathrm{m}}^{\text{io-p/poly}}$-complete) for NP, where the reducibility $\le_{\mathrm{m}}^{\text{io-p/poly}}$ is given in Definition 6 below. If for all $A \in \mathrm{NP}$ it holds $A \le_{\mathrm{m}}^{\mathrm{pp}} (C, D)$, then we say that (C, D) is $\le_{\mathrm{m}}^{\mathrm{pp}}$-hard for NP.

Let SAT denote the set of satisfiable formulas and TAUT the set of tautologies. Without loss of generality, we assume that each word over Σ^* encodes a propositional formula.

► **Definition 1** ([6]). *A function $f \in \mathrm{FP}$ is called a* proof system *for the set* $\mathrm{ran}(f)$*. For $f, g \in \mathrm{FP}$ we say that f* is simulated by *g (resp., f* is P-simulated by *g) denoted by $f \le g$ (resp., $f \le^{\mathrm{p}} g$), if there exists a function π (resp., a function $\pi \in \mathrm{FP}$) and a polynomial p such that $|\pi(x)| \le p(|x|)$ and $g(\pi(x)) = f(x)$ for all x. A function $g \in \mathrm{FP}$ is* optimal *(resp.,* P-optimal*), if $f \le g$ (resp., $f \le^{\mathrm{p}} g$) for all $f \in \mathrm{FP}$ with $\mathrm{ran(f)} = \mathrm{ran(g)}$. Corresponding relativized notions are obtained by using P^O, FP^O, and $\le^{\mathrm{p},O}$ in the definitions above. A* propositional proof system *(pps) is a proof system for* TAUT*.*

► Remark 2. The notion of a *propositional* proof system does not have a canonical relativization. However, in view of Corollary 4 below, it is reasonable to use the following convention. We say that there exist P^O-optimal (resp., optimal) pps relative to an oracle O, if there exists a $\le_{\mathrm{m}}^{\mathrm{p},O}$-complete $A \in \mathrm{coNP}^O$ that has a P^O-optimal (resp., optimal) proof system.

The following proposition states the relativized version of a result by Köbler, Messner, and Torán [24], which they show with a relativizable proof.

▶ **Proposition 3** ([24]). *For every oracle O, if A has a P^O-optimal (resp., optimal) proof system and $\emptyset \neq B \leq_{\mathrm{m}}^{\mathrm{p},O} A$, then B has a P^O-optimal (resp., optimal) proof system.*

▶ **Corollary 4.** *For every oracle O, if there exists a $\leq_{\mathrm{m}}^{\mathrm{p},O}$-complete $A \in \mathrm{coNP}^O$ that has a P^O-optimal (resp., optimal) proof system, then all non-empty sets in coNP^O have P^O-optimal (resp., optimal) proof systems.*

▶ **Definition 5.** *For $f \in \mathrm{FP}$ and a polynomial q, a word $y \in \mathrm{ran}(f)$ is q-hard w.r.t. the proof system f if there does not exist $x \in \Sigma^{\leq q(|y|)}$ such that $f(x) = y$. The set of elements that are q-hard w.r.t. the proof system f is denoted by f_q, i.e., $f_q = \{y \in \mathrm{ran}(f) \mid y \text{ is } q\text{-hard w.r.t. } f\}$.*

We introduce $\leq_{\mathrm{m}}^{\text{io-p/poly}}$-reducibility, which we use to study a weakened variant of $\mathrm{H}_{\text{union}}$: the union of disjoint $\leq_{\mathrm{m}}^{\mathrm{p}}$-complete sets for NP is $\leq_{\mathrm{m}}^{\text{io-p/poly}}$-complete.

P/poly is the class of sets $A \subseteq \Sigma^*$ for which there exist a $B \in \mathrm{P}$ and a function $h : \mathbb{N} \to \Sigma^*$ such that $|h(n)|$ is polynomially bounded in n and for all x it holds that $x \in A \Leftrightarrow (x, h(|x|)) \in B$. FP/poly is the class of total functions $f : \Sigma^* \to \Sigma^*$ for which there exist a $g \in \mathrm{FP}$ and a function $h : \mathbb{N} \to \Sigma^*$ such that $|h(n)|$ is polynomially bounded in n and for all x it holds that $f(x) = g(x, h(|x|))$. Two total functions $f, g : \Sigma^* \to \Sigma^*$ *agree infinitely often*, written as $f \overset{\text{io}}{=} g$, if for infinitely many n it holds that $\forall x \in \Sigma^n, f(x) = g(x)$. Two sets $A, B \subseteq \Sigma^*$ *agree infinitely often*, written as $A \overset{\text{io}}{=} B$, if their characteristic functions agree infinitely often. For a class $\mathcal{C}$ of functions or sets let $\text{io-}\mathcal{C} = \{A \mid \exists B \in \mathcal{C}, A \overset{\text{io}}{=} B\}$.

▶ **Definition 6.** *A set $A \subseteq \Sigma^*$ is* infinitely often P/poly reducible *to a set $B \subseteq \Sigma^*$, written as $A \leq_{\mathrm{m}}^{\text{io-p/poly}} B$, if there exists an $f \in \text{io-FP/poly}$ such that for all x it holds that $x \in A \Leftrightarrow f(x) \in B$.*

It should be mentioned that $\leq_{\mathrm{m}}^{\text{io-p/poly}}$ is an artificial reducibility notion (e.g., it is not transitive), which emerged from the attempt to express the right-hand side of the known implication $\mathrm{H}_{\text{union}} \Rightarrow \mathrm{NP} \neq \mathrm{coNP}$ as a variant of $\mathrm{H}_{\text{union}}$. In Theorem 10 we show that this is possible with $\leq_{\mathrm{m}}^{\text{io-p/poly}}$ reducibility.

In our oracle constructions we use the following notations: If a partial function t is not defined at point x, then $t \cup \{x \mapsto y\}$ denotes the extension t' of t that at x has value y and satisfies $\mathrm{dom}(t') = \mathrm{dom}(t) \cup \{x\}$.

If A is a set, then $A(x)$ denotes the characteristic function at point x, i.e., $A(x)$ is 1 if $x \in A$, and 0 otherwise. An oracle $D \subseteq \mathbb{N}$ is identified with its characteristic sequence $D(0)D(1)\cdots$, which is an ω-word. In this way, $D(i)$ denotes both, the characteristic function at point i and the i-th letter of the characteristic sequence, which are the same. A finite word w describes an oracle that is partially defined, i.e., only defined for natural numbers $x < |w|$. Occasionally, we use w instead of the set $\{i \mid w(i) = 1\}$ and write for example $A = w \cup B$, where A and B are sets. In particular, for an oracle Turing machine M, the notation $M^w(x)$ refers to $M^{\{i|w(i)=1\}}(x)$ (hence, oracle queries that w is not defined for are answered by "no"). Using w instead of $\{i \mid w(i) = 1\}$ additionally allows us to define the following notion: for a nondeterministic oracle Turing machine M, the computation $M^w(x)$ *definitely accepts* if it contains a path that accepts and all queries on this path are $< |w|$. The computation $M^w(x)$ *definitely rejects* if all paths reject and all queries are $< |w|$. We say that the computation $M^w(x)$ *is definite* if it definitely accepts or definitely rejects. Similarly, for a deterministic oracle Turing transducer F, the computation $F^w(x)$ is *definite* if all its queries are $< |w|$.

3 Are Unions of Disjoint NP-Complete Sets NP-Complete?

It is difficult to find out whether $\mathrm{H_{union}}$ is true or not, since each outcome solves a long standing open problem:

$$\begin{array}{lcl} \mathrm{H_{union}} \text{ is true} & \Rightarrow & \mathrm{NP} \neq \mathrm{coNP} \\ \mathrm{H_{union}} \text{ is false} & \Rightarrow & \text{P-inseparable disjoint NP-pairs exist if and only if } \mathrm{P} \neq \mathrm{NP} \end{array}$$

Therefore, researchers approach the hypothesis $\mathrm{H_{union}}$ by proving equivalent, necessary, and sufficient conditions. This section continues this program as follows. In subsection 3.1 we investigate a stronger variant of $\mathrm{H_{union}}$, in 3.2 the original hypothesis, and in 3.3 a weaker variant. We characterize $\mathrm{H_{union}}$ and its variants in several ways, e.g., in terms of p-producibility or coNP-completeness of the set of hard formulas of pps. Within each subsection all hypotheses are equivalent and hence the following implications hold.

$$\begin{array}{ccccc} \text{hypotheses in subsect. 3.1} & \Rightarrow & \text{hypotheses in subsect. 3.2} & \Rightarrow & \text{hypotheses in subsect. 3.3} \\ & & \Updownarrow & & \Updownarrow \\ & & \mathrm{H_{union}} & & \mathrm{NP} \neq \mathrm{coNP} \end{array}$$

Note that under the assumption that all sets in $\mathrm{NPC^p_m}$ are complete w.r.t. length-increasing reductions (which holds for example under the Berman-Hartmanis conjecture), all hypotheses in the subsections 3.1 and 3.2 are equivalent.

3.1 Length-Increasing Polynomial-Time Reducibility

Consider the hypothesis that the union of SAT with a disjoint $B \in \mathrm{NP}$ is $\leq^{\mathrm{p}}_{\mathrm{m,li}}$-complete for NP. We show that this hypothesis can be characterized in terms of the p-producibility of the set of hard formulas of pps. The notion of p-producibility was introduced by Hemaspaandra, Hemaspaandra, and Hempel [21].

▶ **Definition 7** ([21]). *A set A is p-producible if and only if there is some $f \in \mathrm{FP}$ with $|f(x)| \geq |x|$ and $f(x) \in A$ for all x.*

▶ **Theorem 8.** *The following statements are equivalent:*
1. *For all $B \in \mathrm{NP}$ with $\mathrm{SAT} \cap B = \emptyset$ it holds $\mathrm{SAT} \cup B \in \mathrm{NPC^p_{m,li}}$.*
2. *For all $A, B \in \mathrm{NPC^p_{m,li}}$ with $A \cap B = \emptyset$ it holds $A \cup B \in \mathrm{NPC^p_{m,li}}$.*
3. *f_q is p-producible for all pps f and all polynomials q.*

Proof. 1 $\Rightarrow$ 2: Let $A, B \in \mathrm{NPC^p_{m,li}}$ be disjoint and $\mathrm{SAT} \leq^{\mathrm{p}}_{\mathrm{m,li}} A$ via a length-increasing $f \in \mathrm{FP}$. $B' = f^{-1}(B)$ is in NP and disjoint to SAT and hence $\mathrm{SAT} \cup B' \in \mathrm{NPC^p_{m,li}}$. $\mathrm{SAT} \cup B' \leq^{\mathrm{p}}_{\mathrm{m,li}} A \cup B$ via f and thus $A \cup B \in \mathrm{NPC^p_{m,li}}$.

2 $\Rightarrow$ 3: By assumption, $\mathrm{NP} \neq \mathrm{coNP}$. Let f be a pps, q a polynomial, and define

$$B = \{\varphi \mid f(y) = \neg\varphi \text{ for some } y \text{ with } |y| \leq q(|\neg\varphi|)\}.$$

$B \cap \mathrm{SAT} = \emptyset$ and $\mathrm{SAT} \cup B \subsetneq \Sigma^*$. For $A' = 0\mathrm{SAT} \cup 1B$ and $B' = 1\mathrm{SAT} \cup 0B$ it holds $A' \cap B' = \emptyset$ and $A', B' \in \mathrm{NPC^p_{m,li}}$. By 2, $A' \cup B' = \{0,1\}(\mathrm{SAT} \cup B) \in \mathrm{NPC^p_{m,li}}$. In particular, $\mathrm{SAT} \leq^{\mathrm{p}}_{\mathrm{m,li}} \{0,1\}(\mathrm{SAT} \cup B)$. Hence $\mathrm{SAT} \leq^{\mathrm{p}}_{\mathrm{m}} \mathrm{SAT} \cup B$ via $h_1 \in \mathrm{FP}$ with $|x| \leq |h_1(x)|$. Let $h_2 \in \mathrm{FP}$ be length-increasing such that $\mathrm{SAT} \leq^{\mathrm{p}}_{\mathrm{m,li}} \mathrm{SAT}$ via h_2. Thus $\mathrm{SAT} \leq^{\mathrm{p}}_{\mathrm{m,li}} \mathrm{SAT} \cup B$ via $h(x) = h_1(h_2(x))$. We claim that f_q is p-producible via the length-increasing $g(x) = \neg h(x \wedge \neg x)$: As $h(x \wedge \neg x) \notin \mathrm{SAT} \cup B$, $g(x)$ is a tautology. If $g(x) \notin f_q$, then there exists y with $|y| \leq q(|g(x)|)$ and $f(y) = g(x) = \neg h(x \wedge \neg x)$. Hence $h(x \wedge \neg x) \in B$, a contradiction.

$3 \Rightarrow 1$: Choose B according to 1. Consider $B' = \{x \mid x \in B \text{ or } \exists z\, |z| \leq |x| \text{ and } x \vee z \in B\}$ and observe $B' \in \mathrm{NP}$, $B \subseteq B'$, and $B' \cap \mathrm{SAT} = \emptyset$. Let M be an NP-machine with $L(M) = B'$ running in polynomial time q. The following f is a pps.

$$\langle x, z\rangle \mapsto \begin{cases} x & M \text{ accepts } \neg x \text{ on path } z \text{ or } (|z| \geq 2^{|x|} \text{ and } x \text{ is a tautology}) \\ \mathrm{True} & \text{otherwise.} \end{cases}$$

Let q' be a polynomial such that $|\neg x| \leq q'(|x|)$. Choose $r(n) = 2 \cdot (q(q'(n)) + n + 1)$. By 3, f_r is p-producible via some $g \in \mathrm{FP}$ with $|g(x)| \geq |x|$. Consider the length-increasing $h \in \mathrm{FP}$ with $h(x) = \neg g(x) \vee x$. We show $\mathrm{SAT} \leq^{\mathrm{pp}}_{\mathrm{m,li}} (\mathrm{SAT}, \overline{\mathrm{SAT} \cup B})$ via h, which implies $\mathrm{SAT} \leq^{\mathrm{p}}_{\mathrm{m,li}} \mathrm{SAT} \cup B$ via h. As $g(x)$ is a tautology, $x \in \mathrm{SAT} \Leftrightarrow h(x) \in \mathrm{SAT}$. It remains to show $x \notin \mathrm{SAT} \Rightarrow h(x) \notin B$. Let $x \notin \mathrm{SAT}$. If $h(x) = \neg g(x) \vee x \in B$, then due to $|x| \leq |\neg g(x)|$ it holds $\neg g(x) \in B'$. Hence there is some path z such that M accepts $\neg g(x)$ on path z. Thus $|z| \leq q(q'(|g(x)|))$. Consequently, $f(\langle g(x), z\rangle) = g(x)$ and $|\langle g(x), z\rangle| \leq r(|g(x)|)$, in contradiction to $g(x) \in f_r$. ◀

3.2 Polynomial-Time Reducibility

We consider the hypothesis that the union of SAT with a disjoint $B \in \mathrm{NP}$ is $\leq^{\mathrm{p}}_{\mathrm{m}}$-complete for NP. This is equivalent to $\mathrm{H}_{\mathrm{union}}$. We prove one more characterization stating that for each pps f the set of formulas hard for f is coNP-complete. In the following theorem, the equivalence $1 \Leftrightarrow 2$ was shown in [14].

▶ **Theorem 9.** *The following statements are equivalent:*

1. *For all* $B \in \mathrm{NP}$ *with* $\mathrm{SAT} \cap B = \emptyset$ *it holds* $\mathrm{SAT} \cup B \in \mathrm{NPC}^{\mathrm{p}}_{\mathrm{m}}$.
2. *For all* $A, B \in \mathrm{NPC}^{\mathrm{p}}_{\mathrm{m}}$ *with* $A \cap B = \emptyset$ *it holds* $A \cup B \in \mathrm{NPC}^{\mathrm{p}}_{\mathrm{m}}$.
3. f_q *is* $\leq^{\mathrm{p}}_{\mathrm{m}}$*-complete for* coNP *for all pps* f *and all polynomials* q.

Proof. We argue for "$1 \Rightarrow 3$". By definition, $f_q = \{x \in \mathrm{TAUT} \mid \neg\exists z \in \Sigma^{\leq q(|x|)} f(z) = x\}$ and hence $f_q \in \mathrm{coNP}$. Let $B = \{x \in \Sigma^* \mid \exists z \in \Sigma^{\leq q(|\neg x|)} f(z) = \neg x\}$.

Observe that $B \in \mathrm{NP}$ and $\mathrm{SAT} \cap B = \emptyset$. By assumption, $\mathrm{SAT} \cup B \in \mathrm{NPC}^{\mathrm{p}}_{\mathrm{m}}$ and hence $\overline{\mathrm{SAT} \cup B}$ is $\leq^{\mathrm{p}}_{\mathrm{m}}$-complete for coNP. Note $\overline{\mathrm{SAT} \cup B} = \{x \in \Sigma^* \mid \neg x \in \mathrm{TAUT} \wedge \neg\exists z \in \Sigma^{\leq q(|\neg x|)} f(z) = \neg x\}$. Thus $x \in \overline{\mathrm{SAT} \cup B} \Leftrightarrow \neg x \in f_q$ and hence f_q is $\leq^{\mathrm{p}}_{\mathrm{m}}$-complete for coNP.

"$3 \Rightarrow 1$": Let $B \in \mathrm{NP}$ such that $\mathrm{SAT} \cap B = \emptyset$ and let M be a nondeterministic polynomial-time machine that accepts B. Choose a polynomial q such that for all $x \in \Sigma^*$ and all accepting paths y of $M(\neg x)$ it holds that $|\langle x, y\rangle| \leq q(|x|)$. Let

$$f(z) = \begin{cases} x, & \text{if } z = \langle x, y\rangle,\ |y| < 2^{|x|}, \text{ and } y \text{ is an accepting path of } M(\neg x) \\ x, & \text{if } z = \langle x, y\rangle,\ |y| = 2^{|x|}, \text{ and } x \in \mathrm{TAUT} \\ \mathrm{True}, & \text{otherwise.} \end{cases}$$

Observe that f is a pps. By assumption, the set $f_q = \{x \in \mathrm{TAUT} \mid \neg\exists z \in \Sigma^{\leq q(|x|)} f(z) = x\}$ is $\leq^{\mathrm{p}}_{\mathrm{m}}$-complete for coNP. Observe $f_q \cap \Sigma^{\geq n} = \{x \in \mathrm{TAUT} \mid \neg x \notin B\} \cap \Sigma^{\geq n}$ for sufficiently large $n \in \mathbb{N}$. Hence for all $x \in \Sigma^{\geq n}$ it holds that $x \in f_q \Leftrightarrow \neg x \in \overline{\mathrm{SAT} \cup B}$. In the case $\overline{\mathrm{SAT} \cup B} \neq \emptyset$ this shows $f_q \leq^{\mathrm{p}}_{\mathrm{m}} \overline{\mathrm{SAT} \cup B}$ and hence $\mathrm{SAT} \cup B$ is $\leq^{\mathrm{p}}_{\mathrm{m}}$-complete for NP.

It remains to argue that the case $\overline{\mathrm{SAT} \cup B} = \emptyset$ is not possible. If $\overline{\mathrm{SAT} \cup B} = \emptyset$, then $\mathrm{NP} = \mathrm{coNP}$ and hence there exists a polynomially bounded pps f'. Thus for some polynomial q' it holds $f'_{q'} = \emptyset$, which is not $\leq^{\mathrm{p}}_{\mathrm{m}}$-complete for coNP, in contradiction to our assumption. ◀

3.3 Infinitely Often P/poly Reducibility

Consider the hypothesis that the union of SAT with a disjoint $B \in \mathrm{NP}$ is $\leq_{\mathrm{m}}^{\text{io-p/poly}}$-complete for NP. We show that this hypothesis is equivalent to $\mathrm{NP} \neq \mathrm{coNP}$.

▶ **Theorem 10.** *The following statements are equivalent:*

1. *For all* $B \in \mathrm{NP}$ *with* $\mathrm{SAT} \cap B = \emptyset$ *it holds* $\mathrm{SAT} \cup B \in \mathrm{NPC}_{\mathrm{m}}^{\text{io-p/poly}}$.
2. *For all* $A, B \in \mathrm{NPC}_{\mathrm{m}}^{\mathrm{p}}$ *with* $A \cap B = \emptyset$ *it holds* $A \cup B \in \mathrm{NPC}_{\mathrm{m}}^{\text{io-p/poly}}$.
3. $\mathrm{NP} \neq \mathrm{coNP}$ *(i.e., polynomially bounded pps do not exist).*

4 Oracle Constructions

4.1 An Oracle for P = UP and $\neg\mathsf{H}_{\mathsf{cpair}}$

We construct an oracle O relative to which $\mathrm{P} = \mathrm{UP}$ and $\neg\mathrm{H}_{\mathrm{cpair}}$. This answers open questions by Pudlák [32], who lists several conjectures and asks for equivalence proofs and oracles relative to which conjectures are different. Among these are:

$$\begin{aligned}
\mathsf{DisjNP} &\overset{df}{\equiv} \text{“there are no } \leq_{\mathrm{m}}^{\mathrm{pp}}\text{-complete disjoint NP-pairs (i.e., } \neg\mathrm{H}_{\mathrm{cpair}})\text{”} \\
\mathsf{CON} &\overset{df}{\equiv} \text{“there are no P-optimal propositional proof systems”} \\
\mathsf{SAT} &\overset{df}{\equiv} \text{“NP-complete sets do not have P-optimal proof systems”} \\
\mathsf{UP} &\overset{df}{\equiv} \text{“UP does not have } \leq_{\mathrm{m}}^{\mathrm{p}}\text{-complete sets”} \\
\mathsf{NP} \cap \mathsf{coNP} &\overset{df}{\equiv} \text{“NP} \cap \text{coNP does not have } \leq_{\mathrm{m}}^{\mathrm{p}}\text{-complete sets”}
\end{aligned}$$

Relative to O, DisjNP and NP ∩ coNP hold, but UP does not. Hence DisjNP and NP ∩ coNP do not imply UP. Moreover, relative to O, also the following conjectures mentioned by Pudlák [32] do not imply UP (as they are implied by DisjNP relative to all oracles): CON, CON ∨ SAT, and $\mathrm{P} \neq \mathrm{NP}$. The fact that relative to O, CON does not imply UP is of particular interest as the converse implication holds relative to all oracles.

▶ **Theorem 11.** *There exists an oracle O with the following properties.*

1. DisjNP^O *does not have* $\leq_{\mathrm{m}}^{\mathrm{pp},O}$*-complete pairs.*
2. $\mathrm{NP}^O \cap \mathrm{coNP}^O$ *does not have* $\leq_{\mathrm{m}}^{\mathrm{p},O}$*-complete sets.*
3. $\mathrm{P}^O = \mathrm{UP}^O$.

Sketch of the construction: For simplicity, we argue only for 1 and 3. Let $M_0, M_1, \ldots$ be a standard enumeration of nondeterministic, polynomial-time oracle Turing machines and let $F_0, F_1, \ldots$ be a standard enumeration of deterministic, polynomial-time oracle Turing transducers. We assume that for all i the running times of M_i and F_i are bounded by the polynomial $n^i + i$. Adopting an idea by Baker, Gill and Solovay [1], we start with a PSPACE-complete oracle that consists of words of odd length. During the construction we add words of lengths $e(n)$ to the oracle, where $e(0) = 2$ and $e(n+1) = 2^{2^{e(n)}}$. Since $e(n)$ is even, the PSPACE-complete set that we started with will not be damaged.

On the one hand, the construction tries to prevent that $L(M_i)$ and $L(M_j)$ are disjoint. If this is not possible, then M_i and M_j inherently accept disjoint sets. In this case, we make sure that there exists a disjoint NP-pair (A_{ij}, B_{ij}) that does not $\leq_{\mathrm{m}}^{\mathrm{pp}}$-reduce to $(L(M_i), L(M_j))$. This prevents the existence of complete disjoint NP-pairs. On the other hand, we try to prevent that M_i has the uniqueness property “for all x, the computation $M_i(x)$ has at most one accepting path”. If this is not possible, then M_i inherently has the uniqueness property, which allows us to show $L(M_i) \in \mathrm{P}$.

On the technical side, we maintain a growing collection t of properties that we demand in the further construction. If an oracle satisfies the properties defined by t, then we call it t-valid. The collection t contains properties of the following style:

V1: The oracle constructed so far guarantees that $L(M_i) \cap L(M_j) \neq \emptyset$ for all extensions of the oracle.
V2: It is impossible to reach $L(M_i) \cap L(M_j) \neq \emptyset$ and for the oracle constructed so far we have $A_{ij} \cap B_{ij} = \emptyset$. (In the future we restrict to extensions that maintain this property.)
V3: The oracle constructed so far guarantees that for all extensions of the oracle, M_i does not have the uniqueness property.
V4: It is impossible to destroy the uniqueness property of M_i.

The construction successively settles the following tasks:

- task (i, j): If possible, then realize V1 for the pair $(L(M_i), L(M_j))$, otherwise, V2 holds.
- task i: If possible, then realize V3 for M_i, otherwise, V4 holds.
- task (i, j, r): Make sure that F_r does not realize a reduction $(A_{ij}, B_{ij}) \leq_{\mathrm{m}}^{\mathrm{pp}} (L(M_i), L(M_j)))$.

The tasks (i, j) and (i, j, r) make sure that relative to the final oracle, $L(M_i) \cap L(M_j) \neq \emptyset$ or $(L(M_i), L(M_j))$ is not $\leq_{\mathrm{m}}^{\mathrm{pp}}$-complete. The task i ensures that machines having the uniqueness property are very special. An adaption of an argument by Rackoff [33] yields that these machines accept sets in P, hence P = UP.

4.2 An Oracle for $\mathsf{H}_{\mathsf{union}}$ and $\mathsf{H}_{\mathsf{opps}}$

This section constructs an oracle O relative to which the implication $\mathrm{H}_{\mathrm{opps}} \Rightarrow \neg\mathrm{H}_{\mathrm{union}}$ is false. Theorem 22 provides the analogous for the converse implication.

In addition, relative to O there exists a tally set in NEE − coNEE, where $\mathrm{NEE} \overset{\mathit{df}}{=} \mathrm{NTIME}(2^{O(2^n)})$. It shows that two conditions which are sufficient for the existence of an optimal (resp., a P-optimal) pps [24] are not necessary relative to O.

▶ **Theorem 12.** *There exists an oracle O with the following properties.*

1. *There exists a P^O-optimal propositional proof system f.*
2. *If A is $\leq_{\mathrm{m}}^{\mathrm{p},O}$-complete for NP^O and disjoint from $B \in \mathrm{NP}^O$, then $A \cup B$ is $\leq_{\mathrm{m}}^{\mathrm{p},O}$-complete for NP^O.*
3. $\mathrm{NEE}^O \cap \mathrm{TALLY} \not\subseteq \mathrm{coNEE}^O$*, where* $\mathrm{NEE}^O \overset{\mathit{df}}{=} \mathrm{NTIME}^O(2^{O(2^n)})$.

Proof. We only prove statements 1 and 2. Statement 3 follows (in a nontrivial way) from the construction below. Let $M_1, M_3, M_5, \dots$ be a standard enumeration of nondeterministic, polynomial-time oracle Turing machines. Let $F_2, F_4, F_6, \dots$ be a standard enumeration of deterministic, polynomial-time oracle Turing transducers. We assume that the running time of M_i for i odd (resp., F_j for $j > 0$ even) is bounded by the polynomial $n^i + i$ (resp., $n^j + j$).

For a (possibly partial) oracle D we define sets K^D and $K_\vee^D$.

$$K^D = \{\langle 0^i, 0^j, x\rangle \mid i \text{ is odd and } M_i^D(x) \text{ accepts within } j \text{ steps}\}$$
$$K_\vee^D = \{\langle z_1, \dots, z_n\rangle \mid z_1 \in K^D \vee \cdots \vee z_n \in K^D\}$$

▷ Claim 13. For partial oracles v and w and all $y \leq \min(|v|, |w|)$, if $\mathrm{pr}_y(v) = \mathrm{pr}_y(w)$, then $K^w(y) = K^v(y)$ and $K_\vee^w(y) = K_\vee^v(y)$.

Proof. It suffices to show $K^w(y) = K^v(y)$. We may assume $y = \langle 0^i, 0^j, x\rangle$ for suitable i, j, x, since otherwise, $K^w(y) = K^v(y) = 0$. For each q that is queried within the first j steps of $M_i^w(x)$ or $M_i^v(x)$ it holds that $|q| \leq j < |y|$ and thus $q < y$. Hence these queries are answered the same way relative to w and v, showing that $M_i^w(x)$ accepts within j steps if and only if $M_i^v(x)$ accepts within j steps. ◁

K^D and $K_\vee^D$ are $\leq_{\mathrm{m}}^{\mathrm{p},D}$-complete for NP^D and their complements are $\leq_{\mathrm{m}}^{\mathrm{p},D}$-complete for coNP^D. We construct the oracle such that $\overline{K_\vee^D}$ has a P^O-optimal proof system $f \in \mathrm{FP}^O$. As $\overline{K_\vee^O}$ is $\leq_{\mathrm{m}}^{\mathrm{p},O}$-complete for coNP^O, this implies the first statement of the theorem.

For a (possibly partial) oracle D let

$$E^D = \{0^n \mid \exists x \in D \text{ such that } |x| = n\}$$

and observe that $E^D \in \mathrm{NP}^D$. Choose $e \geq 2$ such that $L(M_e^D) = E^D$ for all (possibly partial) oracles D and let $v_n = \langle 0^e, 0^{n^e+e}, 0^n\rangle$. Hence $v_n \in K^D$ if and only if $M_e^D(0^n)$ accepts, i.e., $v_n \in K^D \Leftrightarrow 0^n \in E^D$.

For $i \in 2\mathbb{N}^+$ and $x, y \in \mathbb{N}$ let $c(i,x,y) = \langle 0^i, 0^{(|x|^i+i)^{2ie}}, x, y\rangle$. These words are used to encode proofs into the oracle: if the oracle contains the codeword $c(i,x,y)$, then this means $F_i(x) = y$ and $y \notin K_\vee$, i.e., $c(i,x,y)$ is a proof for $y \notin K_\vee$.

▷ Claim 14. The following holds for all partial oracles w, all $i \in 2\mathbb{N}^+$ and $x, y \in \mathbb{N}$.

1. If $c(i,x,y) \leq |w|$, then $F_i^w(x)$ is definite and $F_i^v(x) = F_i^w(x) < |w|$ for all $v \sqsupseteq w$.
2. If $c(i,x,y) \leq |w|$, then $F_i^w(x)$ is definite and $F_i^w(x) \in K_\vee^w \Leftrightarrow F_i^v(x) \in K_\vee^v$ for all $v \sqsupseteq w$.

Proof. 1: $F_i^w(x)$ is definite, since for each q queried by $F_i^w(x)$ it holds that $|q| \leq |x|^i + i < |c(i,x,y)|$ and hence $q < c(i,x,y) \leq |w|$. The same argument shows $F_i^v(x) = F_i^w(x) < |w|$. 2: Follows from Claims 14.1 and 13. ◁

Preview of construction: On the one hand, the construction tries to prevent that F_i is a proof system for $\overline{K_\vee}$. If this is not possible, then F_i inherently is a proof system for $\overline{K_\vee}$. In this case, the codewords $c(i,x,y)$ are used to encode F_i-proofs into the oracle. These encodings finally yield a P-optimal proof system for $\overline{K_\vee}$. On the other hand, the construction also tries to prevent that M_i accepts a set disjoint from $K_\vee$. If this is not possible, then M_i inherently accepts a set disjoint from $K_\vee$. In this case, there will be a prime p such that the words v_{p^k} for $k \geq 1$ are neither in K nor in $L(M_i)$. It even holds $\langle v_{p^k}, u_1, \ldots, u_n\rangle \notin L(M_i)$ for all $u = \langle u_1, \ldots, u_n\rangle$ of length $\leq |v_{p^k}|$. This means that the v_{p^k} are difficult instances for M_i, since there is no linear-size proof u that allows M_i to recognize that $v_{p^k} \notin K$. Hence adding a sufficiently large v_{p^k} to an instance u does not change the membership to $K_\vee$, but guarantees that the result is not in $L(M_i)$. This yields a reduction $K_\vee \leq_{\mathrm{m}}^{\mathrm{p}} K_\vee \cup L(M_i)$ and implies that $K_\vee \cup L(M_i)$ is NP-complete.

During the construction we maintain a growing list of properties. This list belongs to the set $\mathcal{T} = \{(m_1, \ldots, m_n) \mid n \geq 0,\ m_1, \ldots, m_n \in \mathbb{N}, \text{ and } m_i < m_j \text{ for all } i < j \text{ with } m_j \neq 0\}$. If a partial oracle satisfies the properties defined by a list t, then we call it t-valid. For a list $t = (m_1, \ldots, m_n)$ and $a \in \mathbb{N}$ let $t(i) = m_i$, $|t| = n$, and $t + a = (m_1, \ldots, m_n, a)$. If the list t is a prefix of the list t', then we write $t \sqsubseteq t'$. We start with the empty list $t_0 = ()$, which defines no property. By successively appending an element we obtain lists t_1, t_2, and so on.

A partial oracle $w \in \Sigma^*$ is t-valid, where $t \in \mathcal{T}$, if the following holds:

V1: $w \subseteq \{c(i,x,y) \mid i \in 2\mathbb{N}^+ \text{ and } x, y \in \mathbb{N}\} \cup \{v \mid |v| = p^k \text{ for } p \in \mathbb{P}^{\geq 41} \text{ and } k \geq 1\}$
(meaning: the oracle contains only codewords $c(i,x,y)$ and words of length p^k)

V2: For all $c(i,x,y) \in w$ with $i \in 2\mathbb{N}^+$ and $x, y \in \mathbb{N}$ it holds that $F_i^w(x) = y \notin K_\vee^w$.
(meaning: if the oracle contains the codeword $c(i,x,y)$, then $F_i^w(x)$ outputs $y \notin K_\vee^w$; hence $c(i,x,y) \in w$ is a proof for $y \notin K_\vee^w$)

V3: For all positive even $i \leq |t|$ it holds that $t(i) \in 2\mathbb{N}$ and:

a. If $t(i) = m > 0$, then $c(i,x,y) \in w$ for all $x, y \in \mathbb{N}$ with $F_i^w(x) = y$ and $m \leq c(i,x,y) < |w|$.
(meaning: the oracle maintains codewords for F_i, i.e., if x is large enough and $F_i^w(x)$ outputs y, then w contains a proof for this, namely the codeword $c(i,x,y)$)

b. If $t(i) = 0$, then there exists x such that $F_i^w(x)$ is definite and outputs $y < |w|$ with $y \in K_\vee^w$.
(meaning: F_i is not a proof system for $\overline{K_\vee}$ relative to all extensions of w)

V4: For all odd $i \leq |t|$ it holds that $t(i) \in \{0\} \cup \mathbb{P}^{\geq 41}$ and:

a. If $t(i) = p > 0$, then $\{x \in w \mid |x| = p^k \text{ for } k \geq 1\} = \emptyset$ and for all positive even $j < i$ with $t(j) = 0$ it holds that $\{c(j, x, y) \in w \mid x, y \in \mathbb{N} \text{ and } |c(j, x, y)| \geq p\} = \emptyset$.
(meaning: the first part says $0^{p^k} \notin E^w$ and hence $v_{p^k} \notin K^w$ for all $k \geq 1$; the second part says that if F_j is not a proof system for $\overline{K_\vee}$ and has a smaller index than M_i, then the oracle contains no codewords $c(j, \cdot, \cdot)$ of length $\geq p$)

b. If $t(i) = 0$, then there exists $x < |w|$ such that $x \in K_\vee^w$ and $M_i^w(x)$ definitely accepts.
(meaning: M_i is not disjoint from $K_\vee$ relative to all extensions of w)

▷ Claim 15. The following holds in reference to the definition of t-valid.

1. In V1, the two sets are disjoint.
2. In V2, $F_i^w(x)$ is definite and $F_i^v(x) = y \notin K_\vee^v$ for all $v \sqsupseteq w$.
3. In V3a, $F_i^w(x)$ is definite.
4. In V3b, $y \in K_\vee^v$ for all $v \sqsupseteq w$.
5. In V4b, $x \in K_\vee^v$ for all $v \sqsupseteq w$.

Proof. V1: The union is disjoint, since $|c(i, x, y)|$ is even. V2+V3a: Follows from Claim 14. V3b+V4b: Follows from Claim 13. ◁

▷ Claim 16. Let u and w be t-valid. If $u \sqsubseteq v \sqsubseteq w$, then v is t-valid.

Proof. We show that v satisfies V1–V4. When we consider w and v as sets, then $v \subseteq w$. Therefore, v satisfies V1 and V4a. Moreover, $v \sqsubseteq w$ and Claim 14 imply that v satisfies V2 and V3a. Since u is t-valid, it satisfies V3b and V4b. From $u \sqsubseteq v$, Claim 15.4, and Claim 15.5 it follows that v satisfies V3b and V4b. ◁

Oracle construction: Let $t_0 = ()$ be the empty list and $w_0 = \varepsilon$, which is t_0-valid. We construct a sequence $t_0 \sqsubsetneq t_1 \sqsubsetneq \cdots$ of lists from $\mathcal{T}$ and a sequence $w_0 \sqsubsetneq w_1 \sqsubsetneq \cdots$ of partially defined oracles such that $|t_s| = s$ and w_s is t_s-valid. The final oracle is $O = \lim_{s\to\infty} w_s$. We describe step $s > 0$, which starts with a list t_{s-1} of length $s - 1$ and a t_{s-1}-valid w_{s-1} and which defines a list $t_s \sqsupsetneq t_{s-1}$ of length s and a t_s-valid $w_s \sqsupsetneq w_{s-1}$.

- s even: If there is a t_{s-1}-valid $v \sqsupsetneq w_{s-1}$ such that for some x, $F_s^v(x)$ is definite and has an output $y < |v|$ with $y \in K_\vee^v$, then let $w_s = v$ and $t_s = t_{s-1} + 0$. Otherwise, choose $b \in \{0, 1\}$ such that $w_{s-1}b$ is t_{s-1}-valid, let $w_s = w_{s-1}b$ and $t_s = t_{s-1} + m$ for an even $m > |w_s|$ that is greater than all elements in t_{s-1}.
(meaning: if possible, force that F_s is not a proof system for $\overline{K_\vee}$ relative to all extensions of v; otherwise, we start to maintain codewords for F_s, i.e., if x is large enough and $F_s(x)$ outputs y, then the oracle contains a proof for this, namely the codeword $c(s, x, y)$)
- s odd: If there is a t_{s-1}-valid $v \sqsupsetneq w_{s-1}$ such that for some $x < |v|$, $x \in K_\vee^v$ and $M_s^v(x)$ definitely accepts, then let $w_s = v$ and $t_s = t_{s-1} + 0$. Otherwise, let $w_s = w_{s-1}b$ for $b \in \{0, 1\}$ such that $w_{s-1}b$ is t_{s-1}-valid and $t_s = t_{s-1} + p$ for $p \in \mathbb{P}^{\geq 41}$ large enough such that $(16|v_{p^k}|)^s < 2^{p^k}$ for all $k \in \mathbb{N}^+$, $p > |w_s|$, and p is greater than all elements in t_{s-1}.
(meaning: force $L(M_s) \cap K_\vee \neq \emptyset$ if possible; otherwise, choose a suitable prime p and make sure that the oracle contains no elements of length p^k and hence $v_{p^k} \notin K$ for all $k \geq 1$; the step corresponds to V4)

The subsequent claims refer to the construction above. We start by showing that the construction is possible and how one can extend a t_s-valid $w \sqsupseteq w_s$ by one bit. The proof can be found in [10].

▷ Claim 17. Let $s \in \mathbb{N}$. The choices of w_s and t_s are possible and w_s is t_s-valid. Moreover, for each t_s-valid $w \sqsupseteq w_s$ and $z = |w|$ the following holds.

1. If $z = c(i, x, y)$ for $i \in 2\mathbb{N}^+$, $x, y \in \mathbb{N}$ such that $i \le s$, $t_s(i) > 0$, and $z \ge t_s(i)$, then:
   - **a.** if $F_i^w(x) = y$, then $w1$ is t_s-valid and $w0$ is not.
   - **b.** if $F_i^w(x) \neq y$, then $w0$ is t_s-valid and $w1$ is not.
2. If $z = c(i, x, y)$ for $i \in 2\mathbb{N}^+$, $x, y \in \mathbb{N}$ such that $i \le s$ and $t_s(i) = 0$, then:
   - **a.** $w0$ is t_s-valid.
   - **b.** if $F_i^w(x) = y \notin K_\vee^w$ and there is no odd i' such that $i < i' \le s$, $t_s(i') = p \in \mathbb{P}^{\ge 41}$, and $|z| \ge p$, then $w1$ is t_s-valid.
3. If $z = c(i, x, y)$ for $i \in 2\mathbb{N}^+$, $x, y \in \mathbb{N}$ such that $i > s$, then:
   - **a.** $w0$ is t_s-valid.
   - **b.** if $F_i^w(x) = y \notin K_\vee^w$, then $w1$ is t_s-valid.
4. If $|z| = p^k$ for $p \in \mathbb{P}^{\ge 41}$, $p \notin t_s$, and $k \ge 1$, then $w0$ and $w1$ are t_s-valid.
5. In all other cases $w0$ is t_s-valid.

▷ Claim 18. $M_s^O(\langle v_{p^k}, u_1, \ldots, u_n\rangle)$ rejects for all odd s with $t_s(s) = p \in \mathbb{P}^{\ge 41}$, all $k \in \mathbb{N}^+$, and all $u = \langle u_1, \ldots, u_n\rangle$ with $|u| \le |v_{p^k}|$.

Proof. We assume that $M_s^O(u')$ accepts for $u' = \langle v_{p^k}, u_1, \ldots, u_n\rangle$ and show a contradiction. Choose $j > s$ large enough such that $M_s^{w_j}(u')$ definitely accepts, $|w_j| > u'$, and $|w_j| > q$ for all q with $|q| = p^k$. By construction, w_j is t_j-valid and hence t_{s-1}-valid. Let r be a definitely accepting path of $M_s^{w_j}(u')$. For r we inductively define the set of queries and their dependencies.

$$Q_0 = \{q \mid q \text{ is queried on } r\} \tag{2}$$

$$Q_{n+1} = \bigcup_{\substack{z \in Q_n \text{ with } z = c(i,x,y),\\ i < s,\ x,y \in \mathbb{N},\ t_{s-1}(i) > 0}} \{q \mid q \text{ is queried by } F_i^{w_j}(x)\} \tag{3}$$

Let $Q = \bigcup_{n \ge 0} Q_n$. It holds that $|Q| < 2^{p^k}$, which is seen as follows: For $m_n = \sum_{q \in Q_n} |q|$ we have $m_{n+1} \le m_n/2$, since the sum of lengths of queries induced by $z = c(i, x, y)$ is at most $|x|^i + i \le (|x|^i + i)^{2ie} \le |z|/2$ by the definition of c and $\langle\cdot\rangle$. Thus the m_n form a geometric series. From $|u'| = |u| + 2|v_{p^k}| + 2 \le 4|v_{p^k}|$ it follows $|Q| \le 2m_0 \le 2(|u'|^s + s) \le 4|u'|^s \le (16|v_{p^k}|)^s < 2^{p^k}$, where the latter inequality holds by the choice of p in step s.

Let $\bar{q}$ be the smallest word of length p^k that is not in Q. The word exists, since $|Q| < 2^{p^k}$. By the assumption that $|w_j| > q$ for all q with $|q| = p^k$, it holds in particular $|w_j| > \bar{q}$. By the choice of p in step s we have $p > |w_s|$ and hence $|w_{s-1}| < \bar{q} < |w_j|$. Thus for $v = \mathrm{pr}_{\bar{q}}(w_j)$ it holds that $w_{s-1} \sqsubsetneq v \sqsubsetneq w_j$, where w_{s-1} and w_j are t_{s-1}-valid. By Claim 16, v is t_{s-1}-valid. Moreover, $|v| = \bar{q}$, $|\bar{q}| = p^k$, and $p \notin t_{s-1}$, since step s chooses p greater than all elements in t_{s-1}. From Claim 17.4 it follows that $v1$ is t_{s-1}-valid.

We show that there is a t_{s-1}-valid $w' \sqsupseteq v1$ relative to which r is still a definitely accepting path. More precisely, $|w'| = |w_j|$ and for all $q \in Q$ it holds that $q \in w' \Leftrightarrow q \in w_j$. Below we describe how $v1$ is extended bit by bit to w', i.e., how the word $w \sqsupseteq v1 \sqsupseteq w_{s-1}$ constructed so far is extended by one bit b, where z denotes the length of w. We define b and argue that

$$wb \text{ is } t_{s-1}\text{-valid and if } z \in Q \text{ then } b = w_j(z), \tag{4}$$

where we follow the cases in Claim 17.

1. $z = c(i,x,y)$ for $i \in 2\mathbb{N}^+$, $x,y \in \mathbb{N}$, $i \le s-1$, $t_{s-1}(i) > 0$: If $F_i^w(x) = y$, then $b = 1$ else $b = 0$. Note that $z > \bar{q} > p > t_{s-1}(i)$. By Claim 17.1, wb is t_{s-1}-valid. If $z \in Q$, then by (3), $q \in Q$ for all q queried by $F_i^w(x)$. For these q it holds that $q < z = |w|$ and hence $w(q) = w_j(q)$ by (4). Thus $F_i^w(x) = F_i^{w_j}(x)$. We know that w_j is t_{s-1}-valid and $z > t_{s-1}(i) > 0$. From V2 and V3(a) it follows that $z \in w_j \Leftrightarrow F_i^{w_j}(x) = y \Leftrightarrow F_i^w(x) = y \Leftrightarrow b = 1$. Hence $b = w_j(z)$, which proves (4).
2. $z = c(i,x,y)$ for $i \in 2\mathbb{N}^+$, $x,y \in \mathbb{N}$, $i \le s-1$, $t_{s-1}(i) = 0$: Let $b = 0$. By Claim 17.2, wb is t_{s-1}-valid. Assume $b \ne w_j(z)$, i.e., $z \in w_j$. We are in the situation that w_j is t_j-valid, $s < j$ is odd, $t_j(s) = p$, $i \in 2\mathbb{N}^+$ with $i < s$, and $t_j(i) = 0$. By V4a, the set $\{c(i,x,y) \in w_j \mid x,y \in \mathbb{N} \text{ and } |c(i,x,y)| \ge p\}$ is empty. However, z belongs to this set, as $z = |w| > |v| = \bar{q}$ and hence $|z| \ge p^k \ge p$. This is a contradiction, which shows (4).
3. $z = c(i,x,y)$ for $i \in 2\mathbb{N}^+$, $x,y \in \mathbb{N}$, $i > s-1$: If $z \notin Q \cap w_j$, then $b = 0$ else $b = 1$. If $b = 0$, then wb is t_{s-1}-valid by Claim 17.3. Otherwise, $b = 1$ and $z \in Q \cap w_j$.
   We show $|x|^i + i < p^k$: Assume $|x|^i + i \ge p^k$. From $p \ge 41$, $e \ge 2$, $k \ge 1$, and $i \ge s \ge 1$ it follows that $(41 \cdot p^{ke})^s < p^{2ike}$. Moreover, $|v_{p^k}| = 2(e + p^{ke} + e + p^k + 3) \le 10 \cdot p^{ke}$. Hence we obtain
   $$|c(i,x,y)| > (|x|^i + i)^{2ie} \ge p^{2ike} > (41 \cdot (p^{ke})^s \ge (40 \cdot p^{ke})^s + s \ge (4|v_{p^k}|)^s + s \ge |u'|^s + s.$$
   Thus $|z| > |u'|^s + s \ge m_0 \ge m_1 \ge \cdots$ and hence $z \notin Q$, a contradiction. This proves $|x|^i + i < p^k$.
   We know that w_j is t_j-valid. By V2, $F_i^{w_j}(x) = y \notin K_\vee^{w_j}$. By $|x|^i + i < p^k$, the computation $F_i^{w_j}(x)$ stops within $|x|^i + i < p^k$ steps. Hence it can only ask queries of length $< p^k$ and $|y| < p^k$. Thus $F_i^w(x) = y \notin K_\vee^w$, since w and w_j coincide with respect to all words of length $< p^k$. By Claim 17.3, wb is t_{s-1}-valid.
   To show the second part of (4) assume $z \in Q$. If $b = 1$, then $z \in Q \cap w_j$ and hence $b = w_j(z)$. If $b = 0$, then $z \notin w_j$ and hence $b = w_j(z)$. This proves (4).
4. $|z| = p'^k$ for $p' \in \mathbb{P}^{\ge 41}$, $p' \notin t_s$, $k \ge 1$: Let $b = w_j(z)$. By Claim 17.4, wb is t_{s-1}-valid, which implies (4).
5. Otherwise: Let $b = 0$. By Claim 17.5, wb is t_{s-1}-valid. Assume $b \ne w_j(z)$, i.e., $z \in w_j$. We know that w_j is t_j-valid. From V1 it follows that z must be a word of length p'^k for $p' \in \mathbb{P}^{\ge 41}$ and $p' \in t_{s-1}$ (note that the case $p' \notin t_{s-1}$ has already been considered in 4). Choose s' such that $t_{s-1}(s') = p'$ and note that s' is odd. From V4a it follows that $z \notin w_j$, a contradiction which implies (4).

This shows that there exists a t_{s-1}-valid $w' \sqsupseteq v1 \sqsupsetneq w_{s-1}$ such that $|w'| = |w_j| > u'$ and for all $q \in Q$ it holds that $q \in w' \Leftrightarrow q \in w_j$. Hence $M_s^{w'}(u')$ definitely accepts. Moreover, $|v| = \bar{q}$ and hence $\bar{q} \in w'$. From $|\bar{q}| = p^k$ it follows $v_{p^k} \in K^{w'}$ and $u' \in K_\vee^{w'}$. Therefore, step s of the construction defines $t_s = t_{s-1} + 0$ (and chooses for instance $w_s = w'$), which contradicts the assumption $t_s(s) = p \in \mathbb{P}^{\ge 41}$. ◁

▷ Claim 19. $K_\vee^O \cup B$ is $\le_m^{p,O}$-complete for NP^O for all $B \in \mathrm{NP}^O$ that are disjoint to $K_\vee^O$.

Proof. Choose s odd such that $B = L(M_s^O)$. We claim that $t_s(s) = p \in \mathbb{P}^{\ge 41}$. Otherwise, there exists $x \in K_\vee^{w_s}$ such that $M_s^{w_s}(x)$ definitely accepts. Hence $x \in K_\vee^O$ and $M_s^O(x)$ accepts, which contradicts the assumption $K_\vee^O \cap L(M_s^O) = \emptyset$.

Let $f(\langle u_1, \ldots, u_n \rangle) = \langle u_0, u_1, \ldots, u_n \rangle$, where $u_0 = v_{p^k}$ for the minimal $k \ge 1$ such that $|\langle u_1, \ldots, u_n \rangle| \le |v_{p^k}|$.

It holds that $f \in \mathrm{FP} \subseteq \mathrm{FP}^O$. We argue that f reduces $K_\vee^O$ to $K_\vee^O \cup B$. If $\langle u_1, \ldots, u_n \rangle \in K_\vee^O$, then $f(\langle u_1, \ldots, u_n \rangle) \in K_\vee^O$.

Assume now $\langle u_1, \dots, u_n \rangle \notin K_\vee^O$. From $t_s(s) = p$ it follows that for all $k \geq 1$, O does not contain elements of length p^k and hence $v_{p^k} \notin K^O$. Therefore, $f(\langle u_1, \dots, u_n \rangle) \notin K_\vee^O$. Moreover, by Claim 18, $f(\langle u_1, \dots, u_n \rangle) \notin L(M_s^O) = B$. ◁

▷ Claim 20. If A is $\leq_\mathrm{m}^{\mathrm{p},O}$-complete for NP^O and disjoint to $B \in \mathrm{NP}^O$, then $A \cup B$ is $\leq_\mathrm{m}^{\mathrm{p},O}$-complete for NP^O.

Proof. Otherwise, there are counterexamples A and B. Choose $f \in \mathrm{FP}^O$ such that $K_\vee^O \leq_\mathrm{m}^{\mathrm{p},O} A$ via f and let $B' = f^{-1}(B)$. Observe $B' \in \mathrm{NP}^O$, $K_\vee^O \cap B' = \emptyset$, and $K_\vee^O \cup B' \leq_\mathrm{m}^{\mathrm{p},O} A \cup B$ via f. Hence $K_\vee^O \cup B'$ is not $\leq_\mathrm{m}^{\mathrm{p},O}$-complete for NP^O, which contradicts Claim 19. ◁

▷ Claim 21. $\overline{K_\vee^O}$ has P^O-optimal proof systems.

The straightforward proof of this claim is left due to space restrictions. As $\overline{K_\vee^O}$ is $\leq_\mathrm{m}^{\mathrm{p},O}$-complete for coNP^O, the first statement of the theorem holds. This finishes the proof of Theorem 12. ◀

Köbler, Messner, and Torán [24] prove the following implications (5) and (6).

$$\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{coNEE} \quad \Rightarrow \quad \mathrm{H_{opps}} \tag{5}$$

$$\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{EE} \quad \Rightarrow \quad \exists\ \text{P-optimal pps} \tag{6}$$

Relative to the oracle O constructed above, the converses of (5) and (6) fail, i.e., the premises are stronger than the conclusions. This supports the hope that one can weaken the premises in (5) and (6).

4.3 Further Oracles

We briefly discuss two further oracles.

▶ **Theorem 22.** *There exists an oracle O with the following properties.*

1. DisjNP^O *does not have $\leq_\mathrm{m}^{\mathrm{pp},O}$-complete pairs (and hence $\neg\mathrm{H_{opps}}$ relative to O).*
2. *There are disjoint sets A and B that are $\leq_\mathrm{m}^{\mathrm{p},O}$-complete for NP^O such that $A \cup B$ is not $\leq_\mathrm{m}^{\mathrm{p},O}$-complete for NP^O.*

The construction of this oracle is simpler than the other constructions. In order to achieve statement 1, we proceed similarly as for the oracle in Theorem 11. $\neg\mathrm{H_{union}}$ can be achieved by a straightforward diagonalization.

The following theorem shows that the implication $\mathrm{H_{union}} \Rightarrow \mathrm{H_{cpair}}$ cannot be proven in a relativizable way. Ogiwara and Hemachandra [28] construct an oracle that proves that the converse implication $\mathrm{H_{cpair}} \Rightarrow \mathrm{H_{union}}$ cannot be proven relativizably as well.

▶ **Theorem 23.** *There exists an oracle O with the following properties.*

1. DisjNP^O *does not have $\leq_\mathrm{m}^{\mathrm{pp},O}$-complete pairs (and hence $\neg\mathrm{H_{opps}}$ relative to O).*
2. *If A is $\leq_\mathrm{m}^{\mathrm{p},O}$-complete for NP^O and disjoint to $B \in \mathrm{NP}^O$, then $A \cup B$ is $\leq_\mathrm{m}^{\mathrm{p},O}$-complete for NP^O.*

The construction of this oracle has similarities to the constructions in the Theorems 11 and 12. However, there are less dependencies and thus, the construction is less complicated. Roughly speaking, we achieve $\neg\mathrm{H_{cpair}}$ in the same way as in Theorem 11 and $\mathrm{H_{union}}$ can be obtained similarly as in Theorem 12.

Table 1 Summary of oracles and their properties. Each column corresponds to the oracle mentioned in the topmost cell. We say that there exist P-optimal (resp., optimal) pps relative to an oracle, if relative to this oracle, some $\leq_m^p$-complete $A \in \mathrm{coNP}$ has a P-optimal (resp., optimal) proof system (cf. Remark 2). A disjoint NP-pair (A, B) is $\leq_T^{pp}$-complete, if for every disjoint NP-pair (C, D) and every separator S of (A, B) there exists a separator T of (C, D) such that $T \leq_T^p S$. A disjoint NP-pair (A, B) is $\leq_T^{pp}$-hard for NP, if for every $C \in \mathrm{NP}$ and every separator S of (A, B) it holds that $C \leq_T^p S$. The double exponential time classes are defined as $\mathrm{EE} = \mathrm{DTIME}(2^{O(2^n)})$ and $\mathrm{NEE} = \mathrm{NTIME}(2^{O(2^n)})$.

	[16, T3.8]	[16, T6.1]	[16, T6.7]	[28, L4.7]	[22, T1]	Thm 11	Thm 12	Thm 22	Thm 23
$\exists$ P-optimal pps	false		false			false	true	false	false
$\exists$ optimal pps / $\mathrm{H_{opps}}$	false	true	false	true		false	true	false	false
NPC_m^p closed under disj. union / $\mathrm{H_{union}}$				false	true		true	false	true
$\exists$ $\leq_m^{pp}$-complete disjoint NP-pairs / $\mathrm{H_{cpair}}$	false	true	true	true	true	false	true	false	false
$\exists$ $\leq_T^{pp}$-complete disjoint NP-pairs	false	true	true	true	true			true	
$\exists$ disj. NP-pairs that are $\leq_T^{pp}$-hard for NP	false	false	false	true	false				
$\exists$ P-inseparable disjoint NP-pairs	true	true	true	true	false	true		true	true
$\mathrm{P} \neq \mathrm{UP}$					false	false			
$\mathrm{P} \neq \mathrm{NP}$	true	true	true	true	true	true	true	true	true
$\mathrm{UP} \neq \mathrm{NP}$	true	true	true	true	true	true			
$\mathrm{NP} \neq \mathrm{coNP}$	true	true	true	false	true	true	true	true	true
$\mathrm{NP} \cap \mathrm{SPARSE}$ has $\leq_m^p$-complete sets		true	false	true			true		
$\mathrm{E} \neq \mathrm{NE}$	true	true	true			true	true	true	true
$\mathrm{NE} \neq \mathrm{coNE}$	true	false	true	false		true	true	true	true
$\mathrm{NEE} \cap \mathrm{TALLY} \not\subseteq \mathrm{EE}$	true		true			true	true	true	true
$\mathrm{NEE} \cap \mathrm{TALLY} \not\subseteq \mathrm{coNEE}$	true	false	true	false		true	true	true	true

5 Conclusion and Open Questions

The main goal of this paper is to investigate the hypotheses $\mathrm{H_{union}}$, $\mathrm{H_{opps}}$, and $\mathrm{H_{cpair}}$. We have shown that – except for the known implication $\mathrm{H_{opps}} \Rightarrow \mathrm{H_{cpair}}$ – each two of these hypotheses are independent under relativizable proofs. But what are the connections between the hypotheses if we consider all three at once? At first glance there are 8 possible situations. As $\mathrm{H_{opps}}$ implies $\mathrm{H_{cpair}}$ relative to all oracles, there remain 6 possible situations. Table 1 illustrates that oracles for 4 of the 6 possible situations are known. This leads to the open question: do there also exist oracles for the remaining two situations. More precisely, we ask:

- Does there exist an oracle O_1 with the following properties?
  Relative to O_1, $\neg\mathrm{H_{opps}} \wedge \mathrm{H_{union}} \wedge \mathrm{H_{cpair}}$, i.e., there are no optimal pps, unions of disjoint, $\leq_m^p$-complete NP-sets remain complete, and there are $\leq_m^{pp}$-complete disjoint NP-pairs.
- Does there exist an oracle O_2 with the following properties?
  Relative to O_2, $\neg\mathrm{H_{opps}} \wedge \neg\mathrm{H_{union}} \wedge \mathrm{H_{cpair}}$, i.e., there is no optimal pps, unions of disjoint $\leq_m^p$-complete NP-sets are not always $\leq_m^p$-complete, and DisjNP has $\leq_m^{pp}$-complete elements.

Furthermore we receive new insights on problems related to the main topic. On the one hand, we answer an open question by Pudlák [32] who asks for an oracle relative to which neither $\neg\mathrm{H_{cpair}}$ nor $\neg\mathrm{H_{opps}}$ implies that UP does not have $\leq_m^p$-complete elements (cf. Theorem 11). On the other hand, we show that the converses of Köbler, Messner, and Torán's [24] implications ($\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{coNEE} \Rightarrow \mathrm{H_{opps}}$) and ($\mathrm{NEE} \cap \mathrm{TALLY} \subseteq \mathrm{EE} \Rightarrow$ there exist P-optimal pps) fail relative to an oracle.

References

1 T. Baker, J. Gill, and R. Solovay. Relativizations of the P=NP problem. *SIAM Journal on Computing*, 4:431–442, 1975.
2 O. Beyersdorff. Representable disjoint NP-pairs. In *Proceedings 24th International Conference on Foundations of Software Technology and Theoretical Computer Science*, volume 3328 of *Lecture Notes in Computer Science*, pages 122–134. Springer, 2004.
3 O. Beyersdorff. Disjoint NP-pairs from propositional proof systems. In *Proceedings of Third International Conference on Theory and Applications of Models of Computation*, volume 3959 of *Lecture Notes in Computer Science*, pages 236–247. Springer, 2006.
4 O. Beyersdorff. Classes of representable disjoint NP-pairs. *Theoretical Computer Science*, 377(1-3):93–109, 2007.
5 O. Beyersdorff. The deduction theorem for strong propositional proof systems. *Theory of Computing Systems*, 47(1):162–178, 2010.
6 S. Cook and R. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44:36–50, 1979.
7 T. Dose. P-optimal proof systems for each set in coNP and no complete problems in NP∩coNP relative to an oracle. *CoRR*, abs/1910.08571, 2019. `arXiv:1910.08571`.
8 T. Dose. $\mathrm{P} \neq \mathrm{NP}$ and all sets in NP ∪ coNP have P-optimal proof systems relative to an oracle. *CoRR*, abs/1909.02839, 2019. `arXiv:1909.02839`.
9 T. Dose. An oracle separating conjectures about incompleteness in the finite domain. *Theoret. Comput. Sci.*, 2020. `doi:10.1016/j.tcs.2020.01.003`.
10 T. Dose and C. Glaßer. NP-completeness, proof systems, and disjoint NP-pairs. Technical Report 19-050, Electronic Colloquium on Computational Complexity (ECCC), 2019.
11 S. Even, A. L. Selman, and J. Yacobi. The complexity of promise problems with applications to public-key cryptography. *Information and Control*, 61:159–173, 1984.
12 S. Even and Y. Yacobi. Cryptocomplexity and NP-completeness. In *Proceedings 7th International Colloquium on Automata, Languages and Programming*, volume 85 of *Lecture Notes in Computer Science*, pages 195–207. Springer, 1980.
13 C. Glaßer, J. M. Hitchcock, A. Pavan, and S. Travers. Unions of disjoint NP-complete sets. *ACM Trans. Comput. Theory*, 6(1):3:1–3:10, 2014.
14 C. Glaßer, A. Pavan, A. L. Selman, and S. Sengupta. Properties of NP-complete sets. *SIAM Journal on Computing*, 36(2):516–542, 2006.
15 C. Glaßer, A. L. Selman, and S. Sengupta. Reductions between disjoint NP-pairs. *Information and Computation*, 200:247–267, 2005.
16 C. Glaßer, A. L. Selman, S. Sengupta, and L. Zhang. Disjoint NP-pairs. *SIAM Journal on Computing*, 33(6):1369–1416, 2004.
17 C. Glaßer, A. L. Selman, S. Travers, and K. W. Wagner. The complexity of unions of disjoint sets. *Journal of Computer and System Sciences*, 74(7):1173–1187, 2008.
18 C. Glaßer, A. L. Selman, and L. Zhang. Canonical disjoint NP-pairs of propositional proof systems. *Theoretical Computer Science*, 370:60–73, 2007.
19 C. Glaßer, A. L. Selman, and L. Zhang. The informational content of canonical disjoint NP-pairs. *International Journal of Foundations of Computer Science*, 20(3):501–522, 2009.
20 J. Grollmann and A. L. Selman. Complexity measures for public-key cryptosystems. *SIAM Journal on Computing*, 17(2):309–335, 1988.
21 E. Hemaspaandra, L. A. Hemaspaandra, and H. Hempel. All superlinear inverse schemes are conp-hard. *Theoretical Computer Science*, 345(2-3):345–358, 2005.
22 S. Homer and A. L. Selman. Oracles for structural properties: The isomorphism problem and public-key cryptography. *Journal of Computer and System Sciences*, 44(2):287–301, 1992.
23 Erfan Khaniki. New relations and separations of conjectures about incompleteness in the finite domain. *CoRR*, abs/1904.01362, 2019. `arXiv:1904.01362`.
24 J. Köbler, J. Messner, and J. Torán. Optimal proof systems imply complete sets for promise classes. *Information and Computation*, 184(1):71–92, 2003.

25 J. Krajíček and P. Pudlák. Propositional proof systems, the consistency of first order theories and the complexity of computations. *Journal of Symbolic Logic*, 54:1063–1079, 1989.

26 L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, 1979.

27 J. Myhill. Creative sets. *Mathematical Logic Quarterly*, 1(2):97–108, 1955.

28 M. Ogiwara and L. Hemachandra. A complexity theory of feasible closure properties. *Journal of Computer and System Sciences*, 46:295–325, 1993.

29 C. M. Papadimitriou. *Computational complexity*. Addison-Wesley, Reading, Massachusetts, 1994.

30 P. Pudlák. On the lengths of proofs of consistency. In *Collegium Logicum*, pages 65–86. Springer Vienna, 1996.

31 P. Pudlák. On reducibility and symmetry of disjoint NP pairs. *Theoretical Computer Science*, 295:323–339, 2003.

32 P. Pudlák. Incompleteness in the finite domain. *The Bulletin of Symbolic Logic*, 23(4):405–441, 2017.

33 C. Rackoff. Relativized questions involving probabilistic algorithms. *Journal of the ACM*, 29:261–268, 1982.

34 A. A. Razborov. On provably disjoint NP-pairs. *Electronic Colloquium on Computational Complexity (ECCC)*, 1(6), 1994.

35 H. Rogers Jr. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, New York, 1967.

36 Z. Sadowski. On an optimal propositional proof system and the structure of easy subsets of TAUT. *Theoretical Computer Science*, 288(1):181–193, 2002.

37 A. L. Selman. Natural self-reducible sets. *SIAM Journal on Computing*, 17(5):989–996, 1988.

38 E. Tardos. The gap between monotone and non-monotone circuit complexity is exponential. *Combinatorica*, 8(1):141–142, 1988.

39 S. Travers. *Structural Properties of NP-Hard Sets and Uniform Characterisations of Complexity Classes*. PhD thesis, Julius-Maximilians-Universität Würzburg, 2007.

40 O. V. Verbitskii. Optimal algorithms for coNP-sets and the EXP =? NEXP problem. *Mathematical notes of the Academy of Sciences of the USSR*, 50(2):796–801, August 1991.