Disjoint NP-Pairs from Propositional Proof Systems

(Extended Abstract)

Olaf Beyersdorff

Institut für Informatik, Humboldt-Universität zu Berlin, 10099 Berlin, Germany
beyersdo@informatik.hu-berlin.de

Abstract. For a proof system P we introduce the complexity class $\mathsf{DNPP}(P)$ of all disjoint NP-pairs for which the disjointness of the pair is efficiently provable in the proof system P. We exhibit structural properties of proof systems which make the previously defined canonical NP-pairs of these proof systems hard or complete for $\mathsf{DNPP}(P)$. Moreover we demonstrate that non-equivalent proof systems can have equivalent canonical pairs and that depending on the properties of the proof systems different scenarios for $\mathsf{DNPP}(P)$ and the reductions between the canonical pairs exist.

1 Introduction

Disjoint NP-pairs (DNPP) have been introduced as a complexity theoretic tool to model security aspects of public-key crypto systems [11, 12]. Further, the theory of disjoint NP-pairs is intimately connected to propositional proof complexity with applications to automated theorem proving and lower bounds to the length of proofs [21, 22, 16]. These applications attracted more complexity theoretic research on the structure of the class of disjoint NP-pairs (cf. [8–10, 13, 2]).

Various disjoint NP-pairs have been defined from propositional proof systems which characterize properties of these proof systems. Razborov [22] was the first to associate a canonical pair with a proof system. This pair corresponds to the reflection property of the proof system. Pudlák [21] showed that also the automatizability of the proof system and the feasible interpolation property are expressible by disjoint NP-pairs. In this way disjoint NP-pairs have substantially contributed to the understanding of propositional proof systems.

Conversely, this paper aims to transfer proof-theoretic knowledge to the theory of NP-pairs to gain a more detailed understanding of the structure of the class of disjoint NP-pairs and in particular of the NP-pairs defined from propositional proof systems. For this we define the notions of propositional representations for NP-sets and pairs. The complexity class $\mathsf{DNPP}(P)$ contains all disjoint NP-pairs for which there exist short P-proofs of its disjointness with respect to some representation of the pair. In [22] and [2] similar classes of disjoint NP-pairs corresponding to first order arithmetic theories were considered with the main goal to obtain information on the open problem of the existence of complete pairs for

the class of all DNPP. As theories of bounded arithmetic correspond to strong proof systems the results of [22] and [2] can be transformed into statements about the complexity class $\mathsf{DNPP}(P)$ for strong systems P. However, these results do not apply for weaker systems like resolution or cutting planes which are nevertheless of great interest.

In this paper we demonstrate that also weak proof systems P satisfying certain regularity conditions define reasonable complexity classes $\mathsf{DNPP}(P)$ for which the canonical pairs are complete or hard under the respective reductions. The mentioned regularity conditions are of logical nature: it should be feasible to carry out basic operations like modus ponens or substitutions by constants in the proof system. We also show that proof systems P not satisfying these conditions do not define natural classes $\mathsf{DNPP}(P)$. A recent result of Glaßer et al. [10] states that every DNPP is equivalent to the canonical pair of some proof system. However, the proof systems constructed for this purpose do not satisfy our regularity conditions. The observations of this paper indicate that the Cook-Reckhow framework of propositional proof systems might be too broad for the study of naturally defined classes of disjoint NP-pairs (and in fact for other topics in proof complexity as well). It therefore seems to be natural to make additional assumptions on the properties of proof systems. Consequently, in our opinion, the canonical pairs of these natural proof systems deserve special attention.

The paper is organized as follows. Sections 2 and 3 contain some new results but its main intention is to recall relevant material about propositional proof systems and disjoint NP-pairs. We define and investigate natural properties of proof systems which we use throughout the paper. In Sect. 3 we introduce propositional representations for NP-pairs and the complexity class $\mathsf{DNPP}(P)$.

In Sect. 4 we analyse a weak notion of simulation for proof systems introduced in [17] but not much studied elsewhere. This simulation is provably weaker than the ordinary reduction between proof systems but is equivalent with respect to the existence of optimal proof systems.

In Sect. 5 we provide different ways to construct non-equivalent proof systems with equivalent canonical pairs. A first example for this situation is due to Pudlák [21]. Here we prove that all proof systems that are equivalent with respect to the weak simulation from Sect. 4 possess equivalent canonical pairs.

Section 6 is devoted to the complexity class $\mathsf{DNPP}(P)$. We demonstrate that proof systems P with different properties give rise to different scenarios for $\mathsf{DNPP}(P)$ and the reductions between the NP-pairs associated with P.

Due to space limitations we only sketch proofs or omit them in this extended abstract. The complete paper is available as a technical report [3].

2 Proof Systems with Natural Properties

Propositional proof systems were defined in a very general way by Cook and Reckhow in [7] as polynomial time functions P which have as its range the set of all tautologies. A string π with $P(\pi) = \varphi$ is called a P-proof of the tautology φ. By $P \vdash_{\leq m} \varphi$ we indicate that there is a P-proof of φ of size $\leq m$. If Φ is a set

of propositional formulas we write $P \vdash_* \Phi$ if there is a polynomial p such that $P \vdash_{\leq p(|\varphi|)} \varphi$ for all $\varphi \in \Phi$. If $\Phi = \{\varphi_n \mid n \geq 0\}$ is a sequence of formulas we also write $P \vdash_* \varphi_n$ instead of $P \vdash_* \Phi$.

Proof systems are compared according to their strength by simulations introduced in [7] and [17]. A proof system S *simulates* a proof system P (denoted by $P \leq S$) if there exists a polynomial p such that for all tautologies φ and P-proofs π of φ there is a S-proof π' of φ with $|\pi'| \leq p(|\pi|)$. If such a proof π' can even be computed from π in polynomial time we say that S *p-simulates* P and denote this by $P \leq_p S$. A proof system is called *(p-)optimal* if it (p-)simulates all proof systems. A system P is *polynomially bounded* if $P \vdash_*$ TAUT. By a theorem of Cook and Reckhow [7] polynomially bounded proof systems exist iff $\mathsf{NP} = \mathsf{coNP}$.

In the following we will often consider proof systems satisfying some additional properties. We say that a proof system P is *closed under modus ponens* if there exists a constant c such that $P \vdash_{\leq m} \varphi$ and $P \vdash_{\leq n} \varphi \to \psi$ imply $P \vdash_{\leq m+n+c} \psi$ for all formulas φ and ψ. P is *closed under substitutions* if there exists a polynomial q such that $P \vdash_{\leq m} \varphi$ implies $P \vdash_{\leq q(m+|\sigma(\varphi)|)} \sigma(\varphi)$ for all formulas φ and all substitutions σ. Likewise we say that P is *closed under substitutions by constants* if there exists a polynomial q such that $P \vdash_{\leq m} \varphi(\bar{x}, \bar{y})$ implies $P \vdash_{\leq q(m)} \varphi(\bar{a}, \bar{y})$ for all formulas $\varphi(\bar{x}, \bar{y})$ and constants $\bar{a} \in \{0,1\}^{|\bar{x}|}$. A system P is *closed under disjunctions* if there is a polynomial q such that $P \vdash_{\leq m} \varphi$ implies $P \vdash_{\leq q(m+|\psi|)} \varphi \vee \psi$ for arbitrary formulas ψ. The following property is shared by all systems that simulate the truth-table system: a proof system *evaluates formulas without variables* if these formulas have polynomially long proofs.

We call a proof system *line based* if proofs in the system consist of sequences of formulas, and formulas in such a sequence are derived from earlier formulas in the sequence by the rules available in the proof system. Most of the studied proof systems like resolution, cutting planes and Frege systems are line based in this sense. The most interesting proof system for us will be the *extended Frege proof system EF* that is a usual textbook proof system based on axioms and rules and augmented by the possibility to abbreviate complex formulas by propositional variables to reduce the proof size (see e.g. [14]).

In the following we will often enhance line based proof systems by additional axioms. We will do this in two different ways. Let Φ be a set of tautologies which can be decided in polynomial time. By $P + \Phi$ we denote the proof system P augmented by the possibility to use all formulas from Φ as axiom schemes. This means that formulas from Φ as well as substitution instances of these formulas can be freely introduced as new lines in $P + \Phi$ -proofs. In contrast to this we use the notation $P \cup \Phi$ for the proof system that extends P by formulas from Φ as new axioms. The difference to $P + \Phi$ is that in $P \cup \Phi$ we are only allowed to use formulas from Φ but not their substitution instances in proofs.

We say that a line based proof system P allows *efficient deduction* if there exists a polynomial p such that for all finite sets Φ of tautologies $P \cup \Phi \vdash_{\leq m} \psi$ implies $P \vdash_{\leq p(m+n)} (\bigwedge_{\varphi \in \Phi} \varphi) \to \psi$ where $n = |\bigwedge_{\varphi \in \Phi} \varphi|$. Along the lines of the proof of the deduction theorem for Frege systems (see e.g. [14]) we can prove:

Theorem 1 (Deduction theorem for EF). *EF allows efficient deduction.*

A class of particularly well behaved proof systems is formed by proof systems which correspond to arithmetic theories. To explain this correspondence we have to translate first order arithmetic formulas into propositional formulas. Π^b_1-formulas have only bounded universal quantifiers and describe coNP-predicates. A Π^b_1-formula $\varphi(x)$ is translated into a sequence $\|\varphi(x)\|^n$ of propositional formulas containing one formula per input length for the number x (cf. [14]). We use $\|\varphi(x)\|$ to denote the set $\{\|\varphi(x)\|^n \mid n \geq 1\}$.

The *reflection principle* for a propositional proof system P states a strong form of the consistency of the proof system P. It is formalized by the $\forall\Pi^b_1$-formula

$$\mathrm{RFN}(P) = (\forall\pi)(\forall\varphi)\mathrm{Prf}_P(\pi,\varphi) \rightarrow \mathrm{Taut}(\varphi)$$

where Prf_P and Taut are suitable arithmetic formulas describing P-proofs and tautologies, respectively. A proof system P has the *reflection property* if $P \vdash_* \|\mathrm{RFN}(P)\|^n$ holds.

In [18] a general correspondence between arithmetic theories T and propositional proof systems P is introduced. Pairs (T, P) from this correspondence possess in particular the following two properties:

1. For all $\varphi(x) \in \Pi^b_1$ with $T \vdash (\forall x)\varphi(x)$ we have $P \vdash_* \|\varphi(x)\|^n$.
2. P is the strongest system for which T proves the correctness, i.e. $T \vdash \mathrm{RFN}(P)$ and if $T \vdash \mathrm{RFN}(S)$ for a proof system S, then $S \leq P$.

In the following we call a proof system P *regular* if there exists an arithmetic theory T such that the properties 1 and 2 are fulfilled for (T, P). The most prominent example for this correspondence is the pair (S^1_2, EF). Using this result from [6] we can show that a combination of our extra assumptions on proof systems guarantees the regularity of the system, namely:

Theorem 2. *Let P be a proof system such that $EF \leq P$ and P has reflection and is closed under modus ponens and substitutions. Then $EF + \|\mathrm{RFN}(P)\| \equiv P$. Hence P is regular and corresponds to the theory $S^1_2 + \mathrm{RFN}(P)$.*

3 NP-pairs Defined from Proof Systems

A pair (A, B) is called a disjoint NP-pair (DNPP) if $A, B \in \mathsf{NP}$ and $A \cap B = \emptyset$. The pair (A, B) is *p-separable* if there exists a polynomial time computable set C such that $A \subseteq C$ and $B \cap C = \emptyset$. Grollmann and Selman [11] defined the following reduction between disjoint NP-pairs: $(A, B) \leq_p (C, D)$ if there exists a polynomial time computable function f such that $f(A) \subseteq C$ and $f(B) \subseteq D$. Because elements from $\overline{A \cup B}$ can be mapped to $C \cup D$ a reduction $(A, B) \leq_p (C, D)$ does not imply that A and B are many-one reducible to C and D, respectively. This is, however, the case for the following stronger reduction defined in [13]: $(A, B) \leq_s (C, D)$ if there exists a function $f \in \mathsf{FP}$ with $f^{-1}(C) = A$ and $f^{-1}(D) = B$. As usual we define the equivalence relation $\equiv_p$ as $(A, B) \equiv_p (C, D)$ if $(A, B) \leq_p (C, D)$ and $(C, D) \leq_p (A, B)$, and similarly for $\equiv_s$.

In order to speak about disjoint NP-pairs in proof systems we need to define a propositional encoding of NP-sets.

Definition 3. *Let A be an* NP*-set over the alphabet $\{0,1\}$. A* propositional representation *for A is a sequence of propositional formulas $\varphi_n(\bar{x},\bar{y})$ such that:*

1. *$\varphi_n(\bar{x},\bar{y})$ has propositional variables $\bar{x}$ and $\bar{y}$, and $\bar{x}$ is a vector of n variables.*
2. *There exists a polynomial time algorithm that on input 1^n outputs $\varphi_n(\bar{x},\bar{y})$.*
3. *Let $\bar{a} \in \{0,1\}^n$. Then $\bar{a} \in A$ if and only if $\varphi_n(\bar{a},\bar{y})$ is satisfiable.*

Once we have a propositional description of NP-sets we can also represent disjoint NP-sets in proof systems. This notion is captured by the next definition.

Definition 4. *A disjoint* NP*-pair (A,B) is* representable in a proof system P *if there are representations $\varphi_n(\bar{x},\bar{y})$ of A and $\psi_n(\bar{x},\bar{z})$ of B such that $\bar{x}$ are the common variables of $\varphi_n(\bar{x},\bar{y})$ and $\psi_n(\bar{x},\bar{z})$ and $P \vdash_* \neg\varphi_n(\bar{x},\bar{y}) \vee \neg\psi_n(\bar{x},\bar{z})$.*

By DNPP(P) *we denote the class of all pairs which are representable in P.*

Coding hard tautologies into representations of NP-pairs we can show that the provability of the disjointness of a pair (A,B) in some proof system depends crucially on the choice of the representations for A and B, namely:

Proposition 5. *If optimal proof systems do not exist, then for all proof systems P and all disjoint* NP*-pairs $(A,B) \in$ DNPP(P) there exist representations φ_n of A and ψ_n of B such that $P \nvdash_* \neg\varphi_n \vee \neg\psi_n$.*

Razborov [22] associated a *canonical* disjoint NP-pair $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ with a proof system P where the first component $\mathrm{Ref}(P) = \{(\varphi, 1^m) \mid P \vdash_{\leq m} \varphi\}$ contains information about proof lengths in P and $\mathrm{SAT}^* = \{(\varphi, 1^m) \mid \neg\varphi \in \mathrm{SAT}\}$ is a padded version of SAT. The canonical pair corresponds to the reflection principle of the proof system. Using the above terminology we can express this more precisely as: if P has reflection, then $(\mathrm{Ref}(P), \mathrm{SAT}^*) \in$ DNPP(P). Canonical pairs of strong systems provide candidates for complete NP-pairs. Namely, Razborov showed that if P is an optimal proof system, then the canonical pair of P is $\leq_p$-complete for the class of all DNPP.

The canonical pair is also linked to the automatizability of the proof system, a concept that is of great relevance for automated theorem proving. In [5] a proof system P is called *automatizable* if there exists a deterministic procedure that takes as input a formula φ and outputs a P-proof of φ in time polynomial in the length of the shortest P-proof of φ. This is equivalent to the existence of a deterministic polynomial time algorithm that takes as input $(\varphi, 1^m)$ and produces a P-proof of φ if $(\varphi, 1^m) \in \mathrm{Ref}(P)$. From this reformulation of automatizability it is clear that automatizable proof systems have p-separable canonical pairs. The converse is probably not true as the following proposition shows.

Proposition 6. *There exists a proof system P that has a p-separable canonical pair. But P is not automatizable unless* P $=$ NP*.*

However, Pudlák showed in [21] that the canonical pair of a proof system P is p-separable if and only if there exists an automatizable proof system which simulates P. Therefore proof systems with p-separable canonical pair are called *weakly automatizable*.

Pudlák [21] introduced a second NP-pair for a proof system:

$$I_1(P) = \{(\varphi, \psi, \pi) \mid \mathrm{Var}(\varphi) \cap \mathrm{Var}(\psi) = \emptyset,\ \neg\varphi \in \mathrm{SAT} \text{ and } P(\pi) = \varphi \vee \psi\}$$
$$I_2(P) = \{(\varphi, \psi, \pi) \mid \mathrm{Var}(\varphi) \cap \mathrm{Var}(\psi) = \emptyset,\ \neg\psi \in \mathrm{SAT} \text{ and } P(\pi) = \varphi \vee \psi\}$$

where $\mathrm{Var}(\varphi)$ denotes the set of variables occurring in φ. This pair is p-separable if and only if the proof system P has the efficient interpolation property. Efficient interpolation has been successfully used to show lower bounds to the proof size of a number of proof systems like resolution and cutting planes [4, 15, 20].

In [2] we have defined another kind of canonical pair which is quite similar to the previous pair and which corresponds to the stronger reduction $\leq_s$:

$$U_1(P) = \{(\varphi, \psi, 1^m) \mid \mathrm{Var}(\varphi) \cap \mathrm{Var}(\psi) = \emptyset,\ \neg\varphi \in \mathrm{SAT} \text{ and } P \vdash_{\leq m} \varphi \vee \psi\}$$
$$U_2 = \{(\varphi, \psi, 1^m) \mid \mathrm{Var}(\varphi) \cap \mathrm{Var}(\psi) = \emptyset \text{ and } \neg\psi \in \mathrm{SAT}\}\ .$$

In [2] we investigated classes of disjoint NP-pairs which are representable in theories of bounded arithmetic. As these classes correspond to $\mathsf{DNPP}(P)$ for regular P our results from [2] imply the following:

Theorem 7. *Let P be a regular proof system. Then $(I_1(P), I_2(P))$ and $(U_1(P), U_2)$ are $\leq_s$-complete for $\mathsf{DNPP}(P)$. In particular $(I_1(P), I_2(P)) \equiv_s (U_1(P), U_2)$.*

In Sect. 6 we will analyse this situation for non-regular proof systems.

4 A Weak Reduction Between Proof Systems

Besides $\leq$ and $\leq_p$ we can also study weaker reductions for propositional proof systems. In [17] a weak reduction $\leq'$ is defined between proof systems P and Q as follows: $P \leq' Q$ holds if for all polynomials p there exists a polynomial q such that $P \vdash_{\leq p(|\varphi|)} \varphi$ implies $Q \vdash_{\leq q(|\varphi|)} \varphi$ for all tautologies φ. Using the notation $\vdash_*$ which hides the actual polynomials we can also express the reduction $\leq'$ more compactly as: $P \leq' Q$ iff for all sets Φ of tautologies $P \vdash_* \Phi$ implies $Q \vdash_* \Phi$.

Let us try to motivate the above definition. If we express combinatorial principles in propositional logic we arrive at collections Φ of tautologies that typically contain one tautology per input length. We say that a proof system P proves a combinatorial principle if there exist polynomially long P-proofs of the corresponding collection of tautologies. If $P \leq Q$, then every principle that is provable in P is also provable in Q. The Q-proofs are allowed to be longer than the P-proofs but only up to fixed polynomial amount independent of the principle proven. The reduction $\leq'$ is more flexible as it allows a different polynomial increase for each principle.

It is clear from the above explanation that $\leq$ is a refinement of $\leq'$. We observe that it is indeed a proper refinement, i.e. we can separate $\leq$ and $\leq'$. It is, however, not possible to achieve this separation with regular proof systems.

Proposition 8. *1. Let P be a proof system that is not polynomially bounded. Then there exists a proof system Q such that $P \leq' Q$ but $P \not\leq Q$.*
2. Let P and Q be regular proof systems. Then $P \leq' Q$ implies $P \leq Q$.

However, Krajíček and Pudlák [17] proved that the reductions $\leq$ and $\leq'$ are equivalent with respect to the existence of optimal proof systems.

5 Proof Systems with Equivalent Canonical Pairs

The simulation order of proof systems is reflected in reductions between canonical pairs as the following well known proposition shows (see e.g. [21]):

Proposition 9. *If P and Q are proof systems with $P \leq Q$, then the canonical pair of P is $\leq_p$-reducible to the canonical pair of Q.*

Proof. The reduction is given by $(\varphi, 1^m) \mapsto (\varphi, 1^{p(m)})$ where p is the polynomial from $P \leq Q$. □

If $P \not\leq Q$, then we cannot hope to reduce $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ to $(\mathrm{Ref}(Q), \mathrm{SAT}^*)$ by a reduction of the form $(\varphi, 1^m) \mapsto (\varphi, 1^n)$ that changes only the proof length and not the formula. But unlike in the case of simulations between proof systems the reductions between canonical pairs have the flexibility to change the formula.

The aim of this section is to provide different techniques for the construction of non-equivalent proof systems with equivalent pairs. We first show an analogue of Proposition 9 for $\leq'$.

Proposition 10. *Let P be a proof system that is closed under disjunctions and let Q be a proof system such that $P \leq' Q$. Then $(\mathrm{Ref}(P), \mathrm{SAT}^*) \leq_p (\mathrm{Ref}(Q), \mathrm{SAT}^*)$.*

Proof. The idea of the reduction is to use padding for propositional formulas. For a suitable polynomial q the mapping $(\varphi, 1^m) \mapsto (\varphi \vee \bot^m, 1^{q(m)})$ performs the desired $\leq_p$-reduction where $\bot^m$ stands for $\bot \vee \ldots \vee \bot$ (m disjuncts). □

Combining Propositions 8 and 10 we get the afore mentioned counterexamples to the converse of Proposition 9.

Corollary 11. *Let P be a proof system that is closed under disjunctions and is not polynomially bounded. Then there exists a proof system Q such that $P \not\equiv Q$ and $(\mathrm{Ref}(P), \mathrm{SAT}^*) \equiv_p (\mathrm{Ref}(Q), \mathrm{SAT}^*)$.*

The proof systems P and Q from the last corollary have equivalent canonical pairs and are also $\leq'$-equivalent. Moreover, Proposition 10 implies that the canonical pair is already determined by the $\leq'$-degree of the system:

Proposition 12. *Let P and Q be $\leq'$-equivalent proof systems that are closed under disjunctions. Then $(\mathrm{Ref}(P), \mathrm{SAT}^*) \equiv_p (\mathrm{Ref}(Q), \mathrm{SAT}^*)$.*

Nevertheless we can also construct proof systems that have equivalent canonical pairs but are not $\leq'$-equivalent, namely we can show:

Proposition 13. *Let P be a proof system that is not optimal. Then there exists a proof system Q such that $P \not\equiv' Q$ and $(\mathrm{Ref}(P), \mathrm{SAT}^*) \equiv_p (\mathrm{Ref}(Q), \mathrm{SAT}^*)$.*

Proof. (Idea) For non-optimal proof systems P we can find a polynomial time constructible sequence φ_n with $P \nvdash_* \varphi_n$. Incorporating φ_n as new axioms into P we define a system Q with $Q \vdash_* \varphi_n$ that has the same canonical pair as P. □

The proof systems Q constructed in Proposition 13 have the drawback that they do not satisfy the normality conditions from Sect. 2. In the next theorem we will construct proof systems with somewhat better properties.

Theorem 14. *Let P be a line based proof system that allows efficient deduction and let Φ be a sparse set of tautologies which can be generated in polynomial time. Then $(\mathrm{Ref}(P), \mathrm{SAT}^*) \equiv_p (\mathrm{Ref}(P \cup \Phi), \mathrm{SAT}^*)$.*

Proof. (Idea) The interesting part is to reduce the canonical pair of $P \cup \Phi$ to the canonical pair of P. This is done via $(\psi, 1^m) \mapsto ((\bigwedge_{\varphi \in \Phi_m} \varphi) \to \psi, 1^{p(m)})$ where $\Phi_m = \Phi \cap \Sigma^{\leq m}$, and p is the polynomial from the deduction property of P. □

If we start with a well defined line based system P, then also $P \cup \Phi$ will have good properties (it will lose closure under substitutions). Hence, in contrast to Proposition 13, both P and $P \cup \Phi$ can be chosen to satisfy a reasonable amount of the normality conditions of Sect. 2. As for any non-optimal proof system there exists a sequence of hard tautologies we obtain:

Corollary 15. *For any non-optimal line based proof system P with efficient deduction there exists a sparse set Φ of tautologies which can be generated in polynomial time such that $P \cup \Phi \not\leq' P$ and $(\mathrm{Ref}(P), \mathrm{SAT}^*) \equiv_p (\mathrm{Ref}(P \cup \Phi), \mathrm{SAT}^*)$.*

Because EF admits efficient deduction (Theorem 1) we can formulate the following corollary:

Corollary 16. *Let Φ be a sparse polynomial time set of tautologies. Then we have $(\mathrm{Ref}(EF), \mathrm{SAT}^*) \equiv_p (\mathrm{Ref}(EF \cup \Phi), \mathrm{SAT}^*)$.*

Every proof system P is simulated by $EF + \|\mathrm{RFN}(P)\|$. Clearly $\|\mathrm{RFN}(P)\|$ is a sparse polynomial time set of tautologies. From this information together with Corollary 16 it might be tempting to deduce that the canonical pair of EF is $\leq_p$-complete for the class of all disjoint NP-pairs. The problem, however, is that Corollary 16 only holds for the system $EF \cup \|\mathrm{RFN}(P)\|$ whereas to show the $\leq_p$-completeness of $(\mathrm{Ref}(EF), \mathrm{SAT}^*)$ we would need it for $EF + \|\mathrm{RFN}(P)\|$. We can formulate this observation somewhat differently as:

Theorem 17. *At least one of the following is true:*

1. *The canonical pair of EF is complete for the class of all disjoint NP-pairs.*
2. *There exists a proof system P such that $EF \leq_p EF \cup \|\mathrm{RFN}(P)\| \leq_p EF + \|\mathrm{RFN}(P)\|$ is a chain of pairwise non-equivalent proof systems.*

Both assertions of Theorem 17 contain important information. The first alternative would solve the open problem, posed by Razborov [22], on the existence of complete pairs. But also part 2 is interesting as there is only very limited knowledge about strong proof systems $P \geq EF$.

6 The Complexity Class DNPP(P)

In this section we investigate DNPP(P) for non-regular proof systems. Translating the reductions to the propositional level we have to work with uniform circuit families computing the reduction functions. Since it is possible in resolution to prove the uniqueness of circuit computations we can show the following:

Proposition 18. *Let P be a proof system which simulates resolution and is closed under disjunctions. Then* DNPP(P) *is closed under $\leq_p$.*

Next we show the hardness of the canonical pair for the class DNPP(P):

Theorem 19. *Let P be a proof system that is closed under substitutions by constants and modus ponens and can evaluate formulas without variables. Then* $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ *is $\leq_p$-hard for* DNPP(P).

Proof. (Sketch) Assume that the pair (A, B) is representable in P via the representations $\varphi_n(\bar{x}, \bar{y})$ and $\psi_n(\bar{x}, \bar{z})$, i.e. $P \vdash_* \neg\varphi_n \vee \neg\psi_n$. Then we reduce (A, B) to $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ by $a \mapsto (\neg\psi_{|a|}(\bar{a}, \bar{z}), 1^{p(|a|)})$ with some polynomial p. □

Building on the results of the previous section we construct counterexamples to Theorem 19 under a suitable assumption:

Theorem 20. *There exists a sparse polynomial time constructible set Φ of tautologies such that the canonical pair of $EF \cup \Phi$ is not $\leq_p$-hard for* DNPP($EF \cup \Phi$) *if and only if* $(\mathrm{Ref}(EF), \mathrm{SAT}^*)$ *is not $\leq_p$-complete for all pairs.*

Proof. (Sketch) Assume that $(A, B) \not\leq_p (\mathrm{Ref}(EF), \mathrm{SAT}^*)$. We choose propositional representations φ_n for A and ψ_n for B, and define the set Φ as $\{\neg\varphi_n \vee \neg\psi_n \mid n \geq 0\}$. Then (A, B) is representable in $EF \cup \Phi$ but not reducible to its canonical pair which equals the canonical pair of EF. □

We can interpret Propositions 19 and 20 in such a way that the canonical pairs of sufficiently well defined proof systems like regular proof systems are meaningful as complete pairs for some class of DNPP but that this property is lost for canonical pairs defined from arbitrary proof systems. Therefore the canonical pairs of regular proof systems seem to deserve special attention.

Analogously to Theorem 19 we can prove a propositional variant of Theorem 7, stating the $\leq_s$-hardness of $(U_1(P), U_2)$ for DNPP(P) for proof systems P that are closed under substitutions by constants. In combination with the reflection property we even get completeness results:

Theorem 21. *Let P be a proof system that has the reflection property. Assume further that P is closed under substitutions by constants, modus ponens and disjunctions and can evaluate formulas without variables. Then* $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ *is $\leq_p$-complete for* DNPP(P) *while* $(U_1(P), U_2)$ *is $\leq_s$-complete for* DNPP(P).

What is actually needed for Theorem 21 is not the reflection property of P but the representability of $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ in the proof system P, which is implied by the reflection property of P. However, the next proposition shows that the provability of the reflection principle of a system and the representability of its canonical pair are different concepts.

Proposition 22. *Let P be a regular proof system that is closed under disjunctions. Let further Q be a proof system such that $Q \not\leq P$ but $(\mathrm{Ref}(Q), \mathrm{SAT}^*) \leq_p (\mathrm{Ref}(P), \mathrm{SAT}^*)$. Then $(\mathrm{Ref}(Q), \mathrm{SAT}^*)$ is representable in P but $P \not\vdash_* \|\mathrm{RFN}(Q)\|^n$.*

The following table gives a detailed picture of the properties of the class $\mathsf{DNPP}(P)$ and its associated NP-pairs for three different types of proof systems. Reductions between these NP-pairs and its hardness properties are determined by the properties of the proof system.

weak systems P	resolution, cutting planes
$(\mathrm{Ref}(P), \mathrm{SAT}^*)$	$\leq_p$-hard for $\mathsf{DNPP}(P)$
$(U_1(P), U_2)$	$\leq_s$-hard for $\mathsf{DNPP}(P)$
$(I_1(P), I_2(P))$	p-separable [21]
reductions	$(I_1(P), I_2(P)) \leq_p (U_1(P), U_2) \equiv_p (\mathrm{Ref}(P), \mathrm{SAT}^*)$
	$(U_1(P), U_2) \not\leq_p (I_1(P), I_2(P))$ unless P is weakly automatizable
properties	closed under modus ponens and substitutions by constants
	efficient interpolation [15], no reflection [1]
strong systems P	extensions $EF + \|\Phi\|$ of EF
	by polynomial time computable sets of true Π_1^b-formulas Φ
$(\mathrm{Ref}(P), \mathrm{SAT}^*)$	$\leq_p$-complete for $\mathsf{DNPP}(P)$
$(U_1(P), U_2)$	$\leq_s$-complete for $\mathsf{DNPP}(P)$
$(I_1(P), I_2(P))$	$\leq_s$-complete for $\mathsf{DNPP}(P)$
reductions	$(I_1(P), I_2(P)) \equiv_s (U_1(P), U_2) \equiv_p (\mathrm{Ref}(P), \mathrm{SAT}^*)$
properties	closed under modus ponens and substitutions
	no efficient interpolation under cryptographic assumptions [19]
	reflection property [18], regular
other systems P	extensions $EF \cup \Phi$ of EF by suitable choices
	of polynomial time constructible sets $\Phi \subseteq \mathrm{TAUT}$
$(\mathrm{Ref}(P), \mathrm{SAT}^*)$	not $\leq_p$-hard for $\mathsf{DNPP}(P)$
	unless $(\mathrm{Ref}(EF), \mathrm{SAT}^*)$ is $\leq_p$-hard for all DNPP
reductions	$(I_1(P), I_2(P)) \leq_p (U_1(P), U_2)$, $(\mathrm{Ref}(P), \mathrm{SAT}^*) \leq_p (U_1(P), U_2)$
properties	closed under modus ponens, not closed under substitutions by
	constants unless $(\mathrm{Ref}(EF), \mathrm{SAT}^*)$ is $\leq_p$-hard for all DNPP

Some interesting questions are still unanswered by the last table. For instance, how do $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ and $(U_1(P), U_2)$ compare with respect to the strong reduction $\leq_s$? At least for regular systems we know that $(\mathrm{Ref}(P), \mathrm{SAT}^*) \leq_s (U_1(P), U_2)$. Since $U_1(P)$ is NP-complete the NP-completeness of $\mathrm{Ref}(P)$ is a necessary condition for the opposite reduction to exist. To determine the complexity of $\mathrm{Ref}(P)$ for natural proof systems seems to be an interesting open problem. Approaching this question we note the following:

Proposition 23. *For every proof system P that is closed under disjunctions there is a proof system P' with $P' \equiv_p P$ and $\mathrm{Ref}(P')$ is NP-complete.*

On the other hand there are proof systems P and P' such that $P \equiv_p P'$ and $\mathrm{Ref}(P)$ is decidable in polynomial time while $\mathrm{Ref}(P')$ is NP-complete.

Acknowledgements. For helpful conversations and suggestions on this work I am very grateful to Johannes Köbler, Jan Krajíček, and Pavel Pudlák.

References

1. A. Atserias and M. L. Bonet. On the automatizability of resolution and related propositional proof systems. In *Computer Science Logic, 16th International Workshop*, pages 569–583, 2002.
2. O. Beyersdorff. Representable disjoint NP-pairs. In *Proc. 24th Conference on Foundations of Software Technology and Theoretical Computer Science*, pages 122–134, 2004.
3. O. Beyersdorff. Disjoint NP-pairs from propositional proof systems. Technical Report TR05-083, Electronic Colloquium on Computational Complexity, 2005.
4. M. L. Bonet, T. Pitassi, and R. Raz. Lower bounds for cutting planes proofs with small coefficients. *The Journal of Symbolic Logic*, 62(3):708–728, 1997.
5. M. L. Bonet, T. Pitassi, and R. Raz. On interpolation and automatization for Frege systems. *SIAM Journal on Computing*, 29(6):1939–1967, 2000.
6. S. R. Buss. *Bounded Arithmetic*. Bibliopolis, Napoli, 1986.
7. S. A. Cook and R. A. Reckhow. The relative efficiency of propositional proof systems. *The Journal of Symbolic Logic*, 44:36–50, 1979.
8. C. Glaßer, A. L. Selman, and S. Sengupta. Reductions between disjoint NP-pairs. In *Proc. 19th Annual IEEE Conference on Computational Complexity*, pages 42–53, 2004.
9. C. Glaßer, A. L. Selman, S. Sengupta, and L. Zhang. Disjoint NP-pairs. *SIAM Journal on Computing*, 33(6):1369–1416, 2004.
10. C. Glaßer, A. L. Selman, and L. Zhang. Canonical disjoint NP-pairs of propositional proof systems. In *Proc. 30th International Symposium on the Mathematical Foundations of Computer Science*, pages 399–409, 2005.
11. J. Grollmann and A. L. Selman. Complexity measures for public-key cryptosystems. *SIAM Journal on Computing*, 17(2):309–335, 1988.
12. 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.
13. J. Köbler, J. Messner, and J. Torán. Optimal proof systems imply complete sets for promise classes. *Information and Computation*, 184:71–92, 2003.
14. J. Krajíček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory*, volume 60 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, Cambridge, 1995.
15. J. Krajíček. Interpolation theorems, lower bounds for proof systems and independence results for bounded arithmetic. *The Journal of Symbolic Logic*, 62(2):457–486, 1997.
16. J. Krajíček. Dual weak pigeonhole principle, pseudo-surjective functions, and provability of circuit lower bounds. *The Journal of Symbolic Logic*, 69(1):265–286, 2004.
17. J. Krajíček and P. Pudlák. Propositional proof systems, the consistency of first order theories and the complexity of computations. *The Journal of Symbolic Logic*, 54:1963–1079, 1989.
18. J. Krajíček and P. Pudlák. Quantified propositional calculi and fragments of bounded arithmetic. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 36:29–46, 1990.

19. J. Krajíček and P. Pudlák. Some consequences of cryptographical conjectures for S_2^1 and EF. *Information and Computation*, 140(1):82–94, 1998.
20. P. Pudlák. Lower bounds for resolution and cutting planes proofs and monotone computations. *The Journal of Symbolic Logic*, 62:981–998, 1997.
21. P. Pudlák. On reducibility and symmetry of disjoint NP-pairs. *Theoretical Computer Science*, 295:323–339, 2003.
22. A. A. Razborov. On provably disjoint NP-pairs. Technical Report TR94-006, Electronic Colloquium on Computational Complexity, 1994.