RESEARCH STATEMENT

JÁN PICH

November 2024

Can we automate mathematics? Can we design efficient all-purpose learning algorithms?

I am interested in mathematical logic and complexity theory. In particular, in proof complexity of circuit lower bounds.

Background. Proving lower bounds on the size of circuits computing explicit Boolean functions is one of the most fundamental problems in the theory of computation, underlying questions such as the P versus NP problem. *Circuit lower bounds* have been, however, notoriously difficult to prove. This led to the development of a theory of formal barrier results. A prominent attempt to understand the complexity of circuit lower bounds was made with the discovery of the natural proofs barrier of Razborov and Rudich [37], which ruled out many potential proof methods. The possibility that circuit lower bounds could present even the limits of (feasible) logical reasoning was investigated by Razborov [35], who used natural proofs to derive the unprovability of circuit lower bounds in certain theories of bounded arithmetic. Such unprovability results can be typically interpreted as lower bounds on lengths of proofs in propositional proof systems, which in turn represent the Cook-Reckhow program towards the separation of NP and coNP - proving that there is no propositional proof system with p-size proofs of all tautologies is equivalent to $\mathsf{NP} \neq \mathsf{coNP}$. Unfortunately, Razborov's unprovability result as well as other existing *proof complexity lower bounds* work only for very weak theories and proof systems.

I. Feasible complexity theory

Razborov's conjecture. Razborov's investigation of proof complexity of circuit lower bounds formed a part of the motivation for the development of a theory of proof complexity generators [1, 15, 36], which was introduced in a hope to obtain lower bounds against strong proof systems like Frege - standard textbook systems for propositional logic. Razborov [36] conjectured that any suitable Nisan-Wigderson (NW) generator forms a good proof complexity generator in the sense that tautologies stating the existence of elements outside its range require superpolynomial-size proofs in Frege systems assuming $\mathsf{P/poly}$ is hard on average for NC^1. If true, the conjecture would imply a conditional unprovability of circuit lower bounds in strong mathematical theories. Some specific NW-generators were, in fact, proven to be hard for weak proof systems like Resolution [1]. In [23] I showed that Razborov's conjecture holds for all proof systems with a suitable 'feasible interpolation' property. Unfortunately, the feasible interpolation property is unlikely to hold in strong proof systems like Frege [17, 5].

Feasible provability of P $\neq$ NP. In order to approach stronger systems, I later focused on a first-order version of the problem independently motivated by the question whether $\mathsf{P} \neq \mathsf{NP}$ holds in a strict feasibly constructive sense. In [25] I showed that, under standard hardness assumptions, theories weaker than PV_1 cannot prove (an 'almost everywhere' formulation of) $\mathsf{SAT} \notin \mathsf{P/poly}$.[1] Here, PV_1 is Cook's theory formalizing p-time reasoning [10]. The result was obtained by showing a conditional impossibility of witnessing $\mathsf{P/poly}$ lower bounds by weaker computational models. Unfortunately, this strategy cannot show the unprovability of $\mathsf{SAT} \notin \mathsf{P/poly}$ in PV_1 because standard hardness assumptions imply that $\mathsf{SAT} \notin \mathsf{P/poly}$ can be witnessed by a p-time algorithm. On the positive side, such witnessing algorithms allow us to encode s-size

[1] This result is based on a theorem of Krajíček [16] giving a model-theoretic evidence for Razborov's conjecture and remains the strongest unprovability result concerning $\mathsf{P} \neq \mathsf{NP}$ in first-order theories. However, it does not yield propositional lower bounds. In propositional setting, the strongest lower bound on the hardness of $\mathsf{SAT} \notin \mathsf{P/poly}$ is due to Razborov [36].

circuit lower bounds by 'feasible/succinct' propositional formulas of size $poly(s)$. This realization eventually led to the developments in hardness magnification described below.

Feasible provability of NP $\neq$ coNP. With Santhanam [31] we observed that the proof methods from [25] imply that subexponential average-case co-nondeterministic circuit lower bounds for problems in NP are (unconditionally) unprovable in PV_1. In [30] we gave a propositional analogue of this result: We proved that there is a nonuniform proof system R such that there is no proof system P which for many functions f proves efficiently that R does not prove efficiently $f \notin \mathsf{P/poly}$. Chen, Rothblum and Tell [9] worked out an observation in which I pointed out that their derandomization result implies that under their derandomization assumption there is an explicit nondeterministic algorithm P such that Buss's theory S^1_2 strengthening PV_1 cannot rule out that P is a propositional proof system with p-size proofs of all tautologies. The result is written up for Jeřábek's theory APC_1, but the proof works for S^1_2 as well.

Formalization of complexity theory. Complementing lower bounds, in [26] I aimed at supporting the thesis that a lot of complexity theory is derivable in feasible fragments of arithmetic. I formalized the PCP theorem in the theory PV_1. Further, with Müller [19] we proved $\mathsf{AC}^0, \mathsf{AC}^0[p]$ and monotone circuit lower bounds in a slight extension of PV_1, the theory APC_1 formalizing probabilistic p-time reasoning [13]. In [19] we thus formally strengthened the constructivity of the existing circuit lower bounds. While already Razborov [34] showed the provability of $\mathsf{AC}^0, \mathsf{AC}^0[p]$ and monotone circuit lower bounds in PV_1, we showed the provability of their succinct formulation - in which one is not given the whole truth-table of a hard function but only its polynomially big part or its defining formula. In fact, we gave a succinct version of natural proofs against $\mathsf{AC}^0[p]$ with proofs in a propositional proof system known as WF.

QBF Frege system. An intuitionistic bounded arithmetic S^1_2, developed by Buss, Cook and Urquhart [6, 11], captures the notion of feasibly constructive mathematics more closely than Cook's PV_1. In [4] we showed that a QBF extension of Frege system, introduced by Beyersdorff, Bonacina and Chew [3], presents a QBF equivalent of intuitionistic S^1_2. In fact, it posseses even more constructive properties and can be considered as a formalization of *ultrafinitism*.

II. Hardness magnification

Hardness magnification frontiers. In [24] I suggested to investigate succinct propositional formulas expressing $\mathsf{SAT} \notin \mathsf{P/poly}$ which are exponentially harder than the 'truth-table' formulas expressing $\mathsf{SAT} \notin \mathsf{P/poly}$. In [19] I expanded this approach by observing that if the truth-table formulas encoding a polynomial circuit lower bound require superlinear size proofs in AC^0-Frege systems, then succinct formulas encoding the same polynomial circuit lower bound require Frege proofs of superpolynomial size [19, implicit in Proposition 4.14]. Since AC^0-Frege lower bounds are known, this suggests a way for attacking Frege lower bounds.
Proposition 4.14 [19] inspired Oliveira and Santhanam [21] to develop an analogous strategy in circuit complexity, referred to as *hardness magnification*. They showed that if an average-case version of the minimum circuit size problem MCSP is hard for superlinear-size circuits, then $\mathsf{P} \neq \mathsf{NP}$. Strikingly, their strategy seems to overcome the natural proofs barrier of Razborov and Rudich. In [22] I proved that even if a worst-case version of the minimum circuit size problem MCSP is hard for circuits of superlinear size, then $\mathsf{P} \neq \mathsf{NP}$.[2] Further, by a lower bound of Hirahara-Santhanam [12], the same version of MCSP is hard for formulas of subquadratic size. Since the gap between magnification theorems and known lower bounds is (from several perspectives) seemingly negligible, this raised the hope that closing it was within our reach.

Beyond natural proofs & the locality barrier. In [8] we formally supported the intuition that hardness magnification overcomes the natural proofs barrier: We proved that hardness

[2]Independently, McKay, Murray and Williams [18] proved the implication with a completely different proof.

magnification is in certain cases inherently nonnaturalizable as its successful implementation would imply the non-existence of natural proofs. On the other hand, in [8] we identified a new *locality barrier* which explains why direct adaptations of the existing lower bounds do not yield strong complexity separations via hardness magnification. The phenomenon in the core of the locality barrier says that the existing circuit lower bounds relevant for hardness magnification 'localize' in the sense that they remain valid even if we allow circuits to use arbitrarily powerful oracles with small fan-in. In [29] I showed that a nontrivial localizability is an inherent property of all circuit lower bounds based on the approximation method of Razborov.

III. Learning algorithms from circuit lower bounds

Learning algorithms from witnessing. The formalization results from [19] were partially motivated by an observation [27, 19] pointing out that the ability to prove succinct circuit lower bounds implies the ability to learn outputs of Boolean functions. In [28] I developed this connection further and showed that efficient algorithms witnessing errors of p-size circuits (analogous to the witnessing of circuit lower bounds in bounded arithmetic) are equivalent to the existence of efficient learning algorithms for $\mathsf{P/poly}$. This provided a new characterization of natural proofs and learning algorithms. A similar theorem was previously proved by Carmosino, Impagliazzo, Kabanets and Kolokolova [7], who established an equivalence between learning algorithms and lower bounds defined in terms of natural proofs. Moreover, in [28] I gave an alternative proof of a theorem of Oliveira and Santhanam [20] showing that learning algorithms can be sped up in a generic way. The learning speedup can be also interpreted as a hardness magnification theorem avoiding the locality barrier.

Learning algorithms versus automatability. One of the central notions in proof complexity is the concept of proof-search algorithms formalized by the notion of automatability: a proof system P is automatable if there is an algorithm finding P-proofs of each tautology ϕ in p-time w.r.t. the shortest P-proof of ϕ. Extending the work of Razborov [35], Krajíček [14] and Carmosino et al. [7], in [32] we showed that for every sufficiently strong, well-behaved proof system P such as WF or set theory ZFC (interpreted as a proof system for proving tautologies), if P proves efficiently some sufficiently strong circuit lower bound, then efficient P-provability of efficient learnability of $\mathsf{P/poly}$ is equivalent to efficient P-provability of (a form of) automatability of P on tautologies encoding circuit lower bounds. This shows that in the context of metamathematics it is possible to establish a conditional equivalence between central concepts of complexity theory which we do not know to establish otherwise.

Learning algorithms from breaking cryptography. Do efficient learning algorithms for $\mathsf{P/poly}$ follow from breaking cryptographic pseudorandom generators? Such an implication would significantly strengthen the natural proofs barrier and establish a 'win-win' dichotomy: Either safe cryptography is possible or there are efficient learning algorithms for $\mathsf{P/poly}$. With Santhanam [33] we connected this problem and related questions to a fundamental problem in proof complexity: While the Cook-Reckhow program is an approach to the P vs NP problem, a significant problem with the approach is that we do not know if we ever reach the point of proving a superpolynomial lower bound for all proof systems, if we focus only on concrete ones. In [33] we showed that $\mathsf{P} \neq \mathsf{NP}$ indeed follows from superpolynomial lower bounds for Extended Frege system (EF), if S^1_2 proves (a) a subexponential circuit lower bound for E and that (b) a p-time function transforms circuits breaking one-way functions to p-size circuits computing SAT. If the EF lower bound holds for tautologies expressing suitable circuit lower bounds, (b) can be replaced by the S^1_2-provability of a construction of efficient learning algorithms for $\mathsf{P/poly}$ from circuits breaking one-way functions. This can be interpreted as a conditional proof complexity collapse: If $\mathsf{P} = \mathsf{NP}$ and the above-mentioned assumptions hold, then EF has p-size proofs of

the respective set of tautologies.[3]

References

[1] Alekhnovich M., Ben-Sasson E., Razborov A.A., Wigderson A.; *Pseudorandom generators in propositional proof complexity*; SIAM Journal on Computing, 34(1):67-88, 2004.

[2] Arteche N., Khaniki E., Pich J., Santhanam R.; *From proof complexity to circuit complexity via interactive protocols*; ICALP 2024.

[3] Beyersdorff O., Bonacina I., Chew. L.; *Lower bounds: From circuits to QBF proof systems*; ITCS 2016.

[4] Beyersdorff O., Pich J.; *Understanding Gentzen and Frege systems for QBF*; LICS 2016.

[5] Bonet M.L., Domingo C., Gavalda R., Maciel A., Pitassi T.; *Non-automatizability of bounded-depth Frege proofs*; Computational Complexity, 13:47-68, 2004.

[6] Buss S.R.; *The polynomial hierarchy and intuitionistic bounded arithmetic*; SCT 1986.

[7] Carmosino M., Impagliazzo R., Kabanets V., Kolokolova A.; *Learning algorithms from natural proofs*; CCC 2016.

[8] Chen L., Hirahara S., Oliveira I.C., Pich J., Rajgopal N., Santhanam R.; *Beyond natural proofs: hardness magnification and locality*; ITCS 2020.

[9] Chen L., Rothblum R., Tell R.; *Fiat-Shamir in the plain model from derandomization (Or: Do efficient algorithms believe that* $\mathsf{PSPACE} = \mathsf{NP}$*?)*; ECCC 2024.

[10] Cook S.A.; *Feasibly constructive proofs and the propositional calculus*; STOC 1975.

[11] Cook S.A., Urquhart A.; *Functional interpretations of feasibly constructive arithmetic*; Annals Pure and Applied Logic, 63(2):103-200, 1993.

[12] Hirahara S., Santhanam R.; *On the average-case complexity of MCSP and its variants*; CCC 2017.

[13] Jeřábek E., *Approximate counting in bounded arithmetic*; Journal of Symbolic Logic, 72(3), 2007.

[14] Krajíček J.; *Interpolation theorems, lower bounds for proof systems and independence results for bounded arithmetic*; Journal of Symbolic Logic, 66(2):457-486, 1997.

[15] Krajíček J.; *On the weak pigeonhole principle*; Fundamenta Mathematicae, 170(1-3):123-140, 2001.

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

[17] Krajíček J., Pudlák P.; *Some consequences of cryptographical conjectures for* S_2^1 *and EF*; Information and Computation, 140(1):82-94, 1998.

[18] McKay D., Murray C., Williams R.; *Weak lower bounds on resource-bounded compression imply strong separations of complexity classes*; STOC 2019.

[3] In [2] we avoided the assumption (b) at the expense of a weaker conclusion: We showed that if the implicit Extended Frege system (iEF) proves efficiently a subexponential circuit lower bound for some function, then superpolynomial lower bounds for iEF imply $\mathsf{PSPACE} \not\subseteq \mathsf{P/poly}$.

[19] Müller M., Pich J.; *Feasibly constructive proofs of succinct weak circuit lower bounds*; Annals of Pure and Applied Logic, 2019.

[20] Oliveira I.C., Santhanam R.; *Conspiracies between learning algorithms, circuit lower bounds, and pseudorandomness*; CCC 2017.

[21] Oliveira I.C., Santhanam R.; *Hardness magnification for natural problems*; FOCS 2018.

[22] Oliveira I.C., Pich. J., Santhanam R.; *Hardness magnification near state-of-the-art lower bounds*; CCC 2019.

[23] Pich J.; *Nisan-Wigderson generators in proof systems with forms of interpolation*; Mathematical Logic Quarterly, 57(4), 2011.

[24] Pich J.; *Complexity Theory in Feasible Mathematics*; PhD thesis, Charles University in Prague, 2014.

[25] Pich J.; *Circuit lower bounds in bounded arithmetics*; Annals of Pure and Applied Logic, 166(1), 2015.

[26] Pich J.; *Logical strength of complexity theory and a formalization of the PCP theorem in bounded arithmetic*; Logical Methods in Computer Science, 11(2), 2015.

[27] Pich J.; *Mathesis Universalis*; Literis 2016.

[28] Pich J.; *Learning algorithms from circuit lower bounds*; arXiv 2020.

[29] Pich J.; *Localizability of the approximation method*; Computational Complexity, 33, 12, 2024.

[30] Pich J., Santhanam R.; *Why are proof complexity lower bounds hard?* FOCS 2019.

[31] Pich J., Santhanam R.; *Strong co-nondeterministic lower bounds for* NP *cannot be proved feasibly*; STOC 2021.

[32] Pich J., Santhanam R.; *Learning algorithms versus automatability of Frege systems*; arXiv 2021.

[33] Pich J., Santhanam R.; *Towards $P \neq NP$ from Extended Frege lower bounds*; arXiv 2023.

[34] Razborov A.A.; *Bounded Arithmetic and Lower Bounds in Boolean Complexity*; Feasible Mathematics II, pp. 344-386, 1995.

[35] Razborov A.A; *Unprovability of Lower Bounds on the Circuit Size in Certain Fragments of Bounded Arithmetic*; Izvestiya of the Russian Academy of Science, 59:201-224, 1995.

[36] Razborov A.A; *Pseudorandom Generators Hard for k-DNF Resolution and Polynomial Calculus*; Annals of Mathematics, 181(2):415-472, 2015.

[37] Razborov A.A, Rudich S.; *Natural Proofs*; Journal of Computer and System Sciences, 55(1):24-35, 1997.