Towards $P \neq NP$ from Extended
Frege lower bounds

Ján Pich
University of Oxford

joint work with Rahul Santhanam
Proof complexity

\[ \neg \exists \text{ p-bounded pps} \iff \text{NP} \neq \text{coNP} \]
Proof complexity

\[ \neg \exists \text{ p-bound pps} \iff \text{NP} \neq \text{coNP} \]
Proof complexity

\[ \neg \exists \text{ p-bounded pps} \iff \text{NP} \neq \text{coNP} \]

Frege
AC\(^0\)-Frege
Resolution

Cook-Reckhow program
Proof complexity

$\neg \exists \text{ p-bounded pps} \iff \text{NP} \neq \text{coNP}$

Cook-Reckhow program
\[ \neg \exists \text{p-bounded pps} \iff \text{NP} \neq \text{coNP} \]

EF lower bounds \[\Rightarrow \text{P} \neq \text{NP}\]
\neg \exists \ p\text{-bounded} \ pps \iff NP \neq coNP

lifting \Rightarrow \text{monotone P/poly lbs}

IPS lb \Rightarrow VP \neq VNP

R lb \Rightarrow P \neq NP \ T_R\text{-consistent}

EF lower bounds \Rightarrow P \neq NP
Cook-Reckhow program

$\neg \exists \text{ p-bounded pps } \iff \text{NP} \neq \text{coNP}$

- Frege
- AC$^0$-Frege
- Resolution

lifting $\Rightarrow$ monotone P/poly lbs
IPS lb $\Rightarrow$ VP $\neq$ VNP
R lb $\Rightarrow$ P $\neq$ NP $T_R$ -consistent

EF lower bounds $\Rightarrow$ P $\neq$ NP
Impagliazzo’s worlds shortly before collision
Proof complexity

Impagliazzo’s worlds shortly before collision
Self-provability of $P=NP$

$P=NP \not\implies ZFC \neg P=NP$
**Self-provability of P=NP**

\[ \text{SAT}_n(x, y) \equiv \text{“formula } x \text{ satisfied by assignment } y\” \]

\[ \text{SAT}_n \notin \text{Circuit}[n^{10k}] \]

\[ \Rightarrow \exists \text{ p-time } f \text{ s.t. } \forall C \in \text{Circuit}[n^k] \]
\[ \text{SAT}_n(f_1(C), f_2(C)) \land \neg\text{SAT}_n(f_1(C), C(f_1(C))) \]
Self-provability of $P=NP$

$\text{SAT}_n(x, y) \equiv \text{"formula } x \text{ satisfied by assignment } y\text{"}$

$\text{Witnessing } P \neq NP$

$\text{SAT}_n \not\in \text{Circuit}[n^{10k}] \Rightarrow \exists \text{ p-time } f \text{ s.t. } \forall C \in \text{Circuit}[n^k]$

$\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))$

random

$h$ is one-way $\Rightarrow \text{"}h(x) = h(a)\text{" is a hard SAT-instance}
Self-provability of P=NP

\[ \text{Witnessing } P \neq NP \]

\[ \text{SAT}_n \not\in \text{Circuit}[n^{10^k}] \]

⇒

\[ \exists \text{ p-time } f \text{ s.t. } \forall C \in \text{Circuit}[n^k] \]

\[ \text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C))) \]

\[ h \text{ is one-way } \Rightarrow \text{“} h(x) = h(a) \text{” is a hard SAT-instance} \]

E hard for subexponential-size circuits
Self-provability of \( P = NP \)

Witnessing \( P \neq NP \)

\[
\text{SAT}_n \not\in \text{Circuit}[n^{10k}] \\
\Rightarrow \\
\exists \text{p-time } f \text{ s.t. } \forall C \in \text{Circuit}[n^k] \\
\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))
\]

\( h \) is one-way \( \Rightarrow \) “\( h(x) = h(a) \)” is a hard SAT-instance

\( E \) hard for subexponential-size circuits

[Gutfreund Shaltiel Ta-Shma]-style constructions in uniform setting
Self-provability of P=NP

\[ \text{SAT}_n(x, y) \equiv \text{“formula } x \text{ satisfied by assignment } y \” \]

Witnessing P ≠ NP

\[ \exists \text{ p-time } f \text{ s.t. } w^k_n(f) \in \text{TAUT?} \]

\[ w^k_n(f) := [\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, C(x))] \lor [\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))] \]

variables: \( x, y, C \)
Self-provability of P=NP

\[ \text{SAT}_n(x, y) \equiv \text{“formula } x \text{ satisfied by assignment } y\text{”} \]

Witnessing P ≠ NP

\[ \exists \text{ p-time } f \text{ s.t. } w_n^k(f) \in \text{TAUT?} \]

\[ w_n^k(f) := [\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, C(x))] \vee [\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))] \]

variables: x, y, C

encodes \( n^k \)-size circuits

\[ \exists \text{ p-time } f \text{ s.t. } \forall C \in \text{Circuit}[n^k] \]

\[ \text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C))) \]
Self-provability of $P=NP$

$\text{SAT}_n(x, y) \equiv \text{“formula } x \text{ satisfied by assignment } y\text{”}$

Witnessing $P \neq NP$

$\exists \text{ p-time } f \text{ s.t. } w_n^k(f) \in \text{TAUT}$?

$w_n^k(f) := [\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, C(x))] \lor [\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))]$
Self-provability of $P=NP$

\[
\text{SAT}_n(x, y) \equiv \text{"formula } x \text{ satisfied by assignment } y\"
\]

\[\exists \text{ p-time } f \text{ s.t. } w_n^k(f) \in \text{TAUT}\]

\[w_n^k(f) \equiv [\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, C(x))] \lor [\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))]\]

\[\exists \text{ p-time } f \text{ s.t. } \forall C \in \text{Circuit}[n^k] \text{, SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))\]

\[w_n^k(f) \in \text{TAUT} \downarrow \text{EF + } w_n^k(f)\]

1. \(\vdash A \rightarrow (B \rightarrow A)\)
2. \(\vdash (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))\)
3. \(\vdash (\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)\)
Self-provability of P=NP

\[ \text{SAT}_n(x, y) \equiv \text{“formula } x \text{ satisfied by assignment } y\text{”} \]

\[ \text{SAT}_n \notin \text{Circuit}[n^{10k}] \]

\[ \exists \text{ p-time } f \text{ s.t. } w_n^k(f) \in \text{TAUT} \]

\[ w_n^k(f) := [\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, C(x))] \lor [\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))] \]

variables: \(x, y, C\)

encodes \(n^k\)-size circuits

\[ \text{SAT}_n \in \text{Circuit}[n^{k/10}] \Rightarrow \text{EF} + w^k(f) \vdash \text{“SAT}_n \in \text{Circuit}[n^k]” \]
Self-provability of P=NP

\[ \text{SAT}_n(x, y) \equiv \text{“formula } x \text{ satisfied by assignment } y\text{”} \]

- **Witnessing P ≠ NP**
  - \( \exists \text{ p-time } f \text{ s.t. } w^k_n(f) \in \text{TAUT?} \)
  - \( w^k_n(f) := [\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, C(x))] \lor [\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))] \)

- Variables: \( x, y, C \)
- Encodes \( n^k \)-size circuits

- \( \text{SAT}_n \notin \text{Circuit}[n^{10k}] \)
  \[ \Rightarrow \exists \text{ p-time } f \text{ s.t. } \forall C \in \text{Circuit}[n^k] \]
  \[ \text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C))) \]

**Conclusion**

- \( \text{EF} + w^k_n(f) \vdash \text{“SAT}_n \in \text{Circuit}[n^k]” \)
- \( \Rightarrow \text{EF} + w^k_n(f) \text{ is p-bounded} \)
Self-provability of $P=NP$

$\text{SAT}_n(x, y) \equiv \text{"formula } x \text{ satisfied by assignment } y\text{"}$

Witnessing $P \neq NP$

$\exists \text{ p-time } f \text{ s.t. } w_n^k(f) \in \text{TAUT}$?

$w_n^k(f) := [\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, C(x))] \lor [\text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C)))]$

variables: $x, y, C$

encodes $n^k$-size circuits

$\text{SAT}_n \in \text{Circuit}[n^{k/10}] \Rightarrow \text{EF} + w^k(f) \vdash \text{"SAT}_n \in \text{Circuit}[n^k"]$

$\Rightarrow \text{EF} + w^k(f) \text{ is p-bounded}$

$(\phi \in \text{TAUT} \Rightarrow \text{EF} \vdash \neg \text{SAT}(\neg \phi, C(\neg \phi)) \Rightarrow \text{EF} + w^k(f) \vdash \neg \text{SAT}(\neg \phi, y) \Rightarrow \text{EF} + w^k(f) \vdash \phi)$
Circuit complexity $\iff$ proof complexity & witnessing of $P \neq NP$
Circuit complexity $\iff$ proof complexity & witnessing of $P \neq NP$

**Theorem 1**

Let $k \geq 1$ be a constant.

1. Suppose that there is a $p$-time function $f$ such that for each big enough $n$, $w_n^k(f)$ is a tautology.

In Items 1 and 2, $\epsilon > 0$ is a universal constant (independent of $k$).
Theorem 1

Let $k \geq 1$ be a constant.

1. Suppose that there is a p-time function $f$ such that for each big enough $n$, $w^k_n(f)$ is a tautology. If $EF + w^k(f)$ is not p-bounded, then $\text{SAT}_n \notin \text{Circuit}[n^{\epsilon k}]$ for infinitely many $n$.

In Items 1 and 2, $\epsilon > 0$ is a universal constant (independent of $k$).
Theorem 1
Let $k \geq 1$ be a constant.

1. Suppose that there is a p-time function $f$ such that for each big enough $n$, $w^{k}_n(f)$ is a tautology. If $\text{EF} + w^{k}_n(f)$ is not p-bounded, then $\text{SAT}_n \not\in \text{Circuit}[n^{ck}]$ for infinitely many $n$.

2. Suppose that there is a p-time function $f$ such that for some $n_0$, $S_2^1 \vdash W^{k}_{n_0}(f)$. If $\text{EF}$ is not p-bounded, then $\text{SAT}_n \not\in \text{Circuit}[n^{ck}]$ for infinitely many $n$.

In Items 1 and 2, $\epsilon > 0$ is a universal constant (independent of $k$).
Circuit complexity $\iff$ proof complexity & witnessing of $P \neq NP$

**Theorem 1**

Let $k \geq 1$ be a constant.

1. Suppose that there is a p-time function $f$ such that for each big enough $n$, $w^k_n(f)$ is a tautology. If $\text{EF} + w^k(f)$ is not p-bounded, then $\text{SAT}_n \notin \text{Circuit}[n^{ck}]$ for infinitely many $n$.

2. Suppose that there is a p-time function $f$ such that for some $n_0$, $\mathcal{S}_2^1 \vdash W^k_{n_0}(f)$. If $\text{EF}$ is not p-bounded, then $\text{SAT}_n \notin \text{Circuit}[n^{ck}]$ for infinitely many $n$.

In Items 1 and 2, $\epsilon > 0$ is a universal constant (independent of $k$).

- Generalizes to stronger systems
Theorem 1
Let \( k \geq 1 \) be a constant.

1. Suppose that there is a \( p \)-time function \( f \) such that for each big enough \( n \), \( w^k_n(f) \) is a tautology. If \( \text{EF} + w^k(f) \) is not \( p \)-bounded, then \( \text{SAT}_n \notin \text{Circuit}[n^{ck}] \) for infinitely many \( n \).

2. Suppose that there is a \( p \)-time function \( f \) such that for some \( n_0 \), \( S^1_2 \vdash W^k_{n_0}(f) \). If \( \text{EF} \) is not \( p \)-bounded, then \( \text{SAT}_n \notin \text{Circuit}[n^{ck}] \) for infinitely many \( n \).

In Items 1 and 2, \( \epsilon > 0 \) is a universal constant (independent of \( k \)).
Circuit complexity $\iff$ proof complexity & witnessing of $\mathsf{P} \neq \mathsf{NP}$

**Theorem 1**

Let $k \geq 1$ be a constant.

1. Suppose that there is a p-time function $f$ such that for each big enough $n$, $w_n^k(f)$ is a tautology. If $\text{EF} + w^k(f)$ is not $p$-bounded, then $\text{SAT}_n \notin \text{Circuit}[n^{\epsilon k}]$ for infinitely many $n$.

2. Suppose that there is a p-time function $f$ such that for some $n_0$, $S_2^1 \vdash W_{n_0}^k(f)$. If $\text{EF}$ is not $p$-bounded, then $\text{SAT}_n \notin \text{Circuit}[n^{\epsilon k}]$ for infinitely many $n$.

In Items 1 and 2, $\epsilon > 0$ is a universal constant (independent of $k$).

**Open problem:** $w_n^k(f) \in \text{TAUT}$?

For each p-time $f$ some circuit looks like it solves $\text{SAT}$?
Open problem: \( w^k_n(f) \in \text{TAUT} \)?

\[ \forall k \exists f, \ w^k_n(f) \in \text{TAUT} \Rightarrow \text{NEXP} \not\subseteq \text{P/poly} \]
Nonuniform witnessing

\[ \alpha^s_n := (\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, B(x))) \lor \bigvee_{z \in A} C(z) \neq \text{SAT}_n(z) \]
Nonuniform witnessing

\[ \alpha_n^s := (\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, B(x))) \lor \left( \bigvee_{z \in A} C(z) \neq \text{SAT}_n(z) \right) \]
Nonuniform witnessing

\[ \alpha_n^s := (\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, B(x))) \lor \left( \bigvee_{z \in A} C(z) \neq \text{SAT}_n(z) \right) \]

\[ \exists \text{poly}(s)-\text{size } A \mid \text{SAT}_n \notin \text{Circuit}[s^3] \implies \forall \text{s-size } C, \bigvee_{x \in A} C(x) \neq \text{SAT}_n(x) \]

anti-checkers
Nonuniform witnessing

\[ \alpha_n^s := (\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, B(x))) \lor \left( \bigvee_{z \in A} C(z) \not= \text{SAT}_n(z) \right) \]

∃ \text{s\-size } B' \quad \text{SAT}_n \in \text{Circuit}[s^3] \iff \forall x \in \{0, 1\}^n, B'(x) = \text{SAT}_n(x)

∃ \text{poly}(s\)-size } A \quad \text{SAT}_n \not\in \text{Circuit}[s^3] \Rightarrow \forall \text{s\-size } C, \bigvee_{x \in A} C(x) \not= \text{SAT}_n(x) \quad \text{anti-checkers}
Nonuniform witnessing

\[ \alpha_n^s := (\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, B(x))) \lor (\bigvee_{z \in A} C(z) \neq \text{SAT}_n(z)) \]

fixed p-size circuit

fixed p-size set

\[ \exists s^3\text{-size } B' \quad \text{SAT}_n \in \text{Circuit}[s^3] \iff \forall x \in \{0, 1\}^n, B'(x) = \text{SAT}_n(x) \]

\[ \exists \text{poly(s)-size } A \quad \text{SAT}_n \notin \text{Circuit}[s^3] \quad \Rightarrow \forall s\text{-size } C, \bigvee_{x \in A} C(x) \neq \text{SAT}_n(x) \quad \text{anti-checkers} \]

Theorem 2 (Circuit complexity from nonuniform proof complexity).

Let \( k \geq 3 \) be a constant. If there are tautologies without p-size EF-derivations from substitutional instances of tautologies \( \alpha_n^{n^k} \), then \( \text{SAT}_n \notin \text{Circuit}[n^k] \) for infinitely many \( n \).
Nonuniform witnessing

fixed p-size circuit

\( \alpha_n^s := (\text{SAT}_n(x, y) \rightarrow \text{SAT}_n(x, B(x))) \lor (\bigvee_{z \in A} C(z) \neq \text{SAT}_n(z)) \)

Open problem: Feasible MinMax?

**Theorem 2** (Circuit complexity from nonuniform proof complexity).

Let \( k \geq 3 \) be a constant. If there are tautologies without p-size EF-derivations from substitutional instances of tautologies \( \alpha_n^{nk} \), then \( \text{SAT}_n \notin \text{Circuit}[n^k] \) for infinitely many \( n \).
Collapsing Impagliazzo’s worlds
OWF ⇐ P ≠ NP
Proof complexity collapse from “OWF \iff P \neq NP” & hardness of E

\[ S_2^1 \vdash \text{E hard on average for subexponential-size circuits} \]

\&

\[ S_2^1 \vdash \text{OWF} \iff P \neq NP \]

\[ \implies \]

EF not p-bounded \implies P \neq NP
Proof complexity collapse from “OWF ⇐ P≠NP” & hardness of E

Theorem

\[ S^1_2 \vdash E \text{ hard on average for subexponential-size circuits} \]

\[ \& \]

\[ S^1_2 \vdash \text{OWF} \leftrightarrow P\neq NP \]

\[ \Rightarrow \]

EF not p-bounded \implies P\neq NP

• No need for the provability of “E is hard” if EF replaced by EF+“E is hard”
Proof complexity collapse from “OWF ⇐ P≠NP” & hardness of E

\[ S^1_2 \vdash E \text{ hard on average for subexponential-size circuits} \]
\[ \& \]
\[ S^1_2 \vdash \text{OWF} ⇐ P≠NP \]

\[ \implies \]

EF not p-bounded \implies P≠NP

• No need for the provability of “E is hard” if EF replaced by EF+“E is hard”
• Generalizes to stronger systems, e.g. ZFC
Proof complexity collapse from “OWF $\iff P \neq NP$” & hardness of E

\[ S_2^1 \vdash E \text{ hard on average for subexponential-size circuits} \]
&
\[ S_2^1 \vdash \text{OWF} \iff P \neq NP \]

$\implies$

EF not p-bounded $\Rightarrow P \neq NP$

- No need for the provability of “E is hard” if EF replaced by EF+“E is hard”
- Generalizes to stronger systems, e.g. ZFC
- Requires \textit{p-time reductions} witnessing that OWF $\iff P \neq NP$
random
\[ \Rightarrow \]
\( h \text{ is one-way} \Rightarrow "h(x) = h(a)" \) is a hard \( \textsf{SAT} \)-instance

\( \mathbf{E} \) hard on average for subexponential-size circuits

\[ \Rightarrow \]
\[ \exists \text{ p-time } f \text{ s.t. } \forall C \in \text{ Circuit}[n^k] \]
\[ \text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C'), C(f_1(C'))) \]
Proof

$h$ is one-way $\Rightarrow$ “$h(x) = h(a)$” is a hard SAT-instance

$E$ hard on average for subexponential-size circuits

$\exists$ p-time $f$ s.t. $\forall C \in \text{Circuit}[n^k]$

$SAT_n(f_1(C), f_2(C)) \land \neg SAT_n(f_1(C), C(f_1(C)))$
Proof

\[ S_2^1 \vdash P \neq NP \]

\[ \downarrow \]

\[ h \text{ is one-way} \Rightarrow "h(x) = h(a)" \text{ is a hard SAT-instance} \]

\[ \downarrow \]

\[ E \text{ hard on average for subexponential-size circuits} \]

\[ \iff \]

\[ \exists \text{ p-time } f \text{ s.t. } \forall C \in \text{ Circuit}[n^k] \]

\[ \text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C))) \]
Proof

\( \exists \) p-time \( f \) s.t. \( \forall C \in \text{Circuit}[n^k] \)
\( \text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C))) \)
Proof

\[ S_2^1 \vdash P \neq NP \]

\[ \downarrow \]

\[ h \text{ is one-way} \Rightarrow \text{“} h(x) = h(a) \text{” is a hard SAT-instance} \]

\[ \downarrow \]

\[ E \text{ hard on average for subexponential-size circuits} \]

\[ S_2^1 \vdash P = NP \text{ or} \]

\[ S_2^1 \vdash \exists \text{ p-time } f \text{ s.t. } \forall C \in \text{Circuit}[n^k] \]

\[ \text{SAT}_n(f_1(C), f_2(C)) \land \neg \text{SAT}_n(f_1(C), C(f_1(C))) \]

\[ \downarrow \]

\[ S_2^1 \vdash w_n^k(f) \in \text{TAUT} \]
EF not p-bounded $\Rightarrow$ $P \neq NP$

**Theorem**

$S_2^1 \vdash E$ hard on average for subexponential-size circuits

&

$S_2^1 \vdash OWF \iff P \neq NP$

$\iff$

EF not p-bounded $\Rightarrow P \neq NP$

- Can replace “$OWF \iff P \neq NP$” by “**Learning or Crypto**”
  
  if EF lower bounds replaced by EF lower bounds for tautologies expressing circuit lower bounds
Learning or Crypto

S\(^1\) \models E \text{ hard on average for subexponential-size circuits}

\&

S\(^1\) \models OWF or Learning P/poly

\implies

EF \notin \text{ circuit lower bound } \implies P \neq NP

• Can replace “OWF \iff P \neq NP” by “Learning or Crypto”
  if EF lower bounds replaced by EF lower bounds for tautologies expressing circuit lower bounds
Automatability or OWF

Theorem

\( S_2^1 \vdash E \text{ hard on average for subexponential-size circuits} \)

&

\( S_2^1 \vdash \text{OWF or EF automatable} \)

\( \implies \)

EF \( \not\exists \) circuit lower bound \( \Rightarrow \) P \( \neq \) NP

• Can replace “OWF \( \iff \) P \( \neq \) NP” by “Automatability or OWF”
  if EF lower bounds replaced by EF lower bounds for tautologies expressing circuit lower bounds
Concluding remarks
Concluding remarks

fundamental connection between
logic
crypto &
learning
Concluding remarks

fundamental connection between logic crypto & learning

Thank You