Near-Optimal Pseudorandom Generators for Constant-Depth Read-Once Formulas

Dean Doron¹ UT Austin \rightarrow Stanford

Pooya Hatami² UT Austin \rightarrow Ohio State William M. Hoza³ UT Austin

July 19 CCC 2019

^{1&}lt;br>Supported by NSF Grant CCF-1705028

 2 Supported by a Simons Investigator Award (#409864, David Zuckerman)

³ Supported by the NSF GRFP under Grant DGE-1610403 and by a Harrington Fellowship from UT Austin

Randomness as a scarce resource

- \blacktriangleright Randomization is a popular algorithmic technique
- \blacktriangleright But randomness is costly

Randomness as a scarce resource

- \blacktriangleright Randomization is a popular algorithmic technique
- \blacktriangleright But randomness is costly
- \triangleright An algorithm that uses fewer random bits is better

▶ Gen "fools" $f: \{0, 1\}^n \to \{0, 1\}$ if

 $\mathbb{E}[f(\mathsf{Gen}(U))] = \mathbb{E}[f(U)] \pm \varepsilon$

▶ Gen "fools" $f: \{0, 1\}^n \to \{0, 1\}$ if

$$
\mathbb{E}[f(\mathsf{Gen}(U))] = \mathbb{E}[f(U)] \pm \varepsilon
$$

 \triangleright Goal: Design PRG that fools an interesting class of functions f

▶ Gen "fools" $f: \{0, 1\}^n \to \{0, 1\}$ if

$$
\mathbb{E}[f(\mathsf{Gen}(U))] = \mathbb{E}[f(U)] \pm \varepsilon
$$

 \triangleright Goal: Design PRG that fools an interesting class of functions f

IMinimize seed length $s = s(n, \varepsilon)$

Read-once formulas

Read-once formulas

In This work: Fool depth-d read-once formulas for $d = O(1)$

Read-once formulas

In This work: Fool depth-d read-once formulas for $d = O(1)$ Read-once version of AC^0

 \triangleright Main result: PRG for read-once AC^0 with seed length

 $\log(n/\varepsilon) \cdot O(d \log \log(n/\varepsilon))^{2d+2}$.

 \triangleright Big open problem: Prove $\mathsf{L} = \mathsf{BPL}$

 \blacktriangleright Big open problem: Prove \blacktriangleright \blacktriangleright BPL

In "Randomness is not necessary for space-efficient computation"

 \blacktriangleright Big open problem: Prove \blacktriangleright \blacktriangleright BPL

 \blacktriangleright "Randomness is not necessary for space-efficient computation"

▶ Main approach: Design optimal PRG for "ROBPs"

 \blacktriangleright Big open problem: Prove \blacktriangleright \blacktriangleright BPL

 \blacktriangleright "Randomness is not necessary for space-efficient computation"

▶ Main approach: Design optimal PRG for "ROBPs"

Bad news: Seed length $O(\log^2 n)$ has not been improved for decades [Nisan '92]

 \blacktriangleright Big open problem: Prove \blacktriangleright \blacktriangleright BPL

 \blacktriangleright "Randomness is not necessary for space-efficient computation"

▶ Main approach: Design optimal PRG for "ROBPs"

- Bad news: Seed length $O(\log^2 n)$ has not been improved for decades [Nisan '92]
- Good news: Can achieve seed length $\widetilde{O}(\log n)$ for increasingly powerful restricted models

 \blacktriangleright Big open problem: Prove \blacktriangleright \blacktriangleright BPL

 \blacktriangleright "Randomness is not necessary for space-efficient computation"

▶ Main approach: Design optimal PRG for "ROBPs"

- Bad news: Seed length $O(\log^2 n)$ has not been improved for decades [Nisan '92]
- Good news: Can achieve seed length $\tilde{O}(\log n)$ for increasingly powerful restricted models

Read-once AC^0 is one of the frontiers of this progress

Seed length $\widetilde{O}(\log n)$

Starting point: Forbes-Kelley PRG

Starting point: Forbes-Kelley PRG

PRGs via pseudorandom restrictions [AW89]

PRGs via pseudorandom restrictions [AW89]

Start by sampling a pseudorandom restriction $X \in \{0, 1, \star\}^n$

PRGs via pseudorandom restrictions [AW89]

Start by sampling a pseudorandom restriction $X \in \{0, 1, \star\}^n$

Restriction notation

Define Res: $\{0,1\}^n \times \{0,1\}^n \to \{0,1,\star\}^n$ by

$$
Res(y, z)_i = \begin{cases} \star & \text{if } y_i = 1\\ z_i & \text{if } y_i = 0 \end{cases}
$$

Restriction notation

Define Res: $\{0,1\}^n \times \{0,1\}^n \to \{0,1,\star\}^n$ by

$$
Res(y, z)_i = \begin{cases} \star & \text{if } y_i = 1\\ z_i & \text{if } y_i = 0 \end{cases}
$$

$$
y = 0 1 1 0 0 1 0 0
$$

$$
z = 0 0 1 1 1 1 0 1
$$

$$
Res(y, z) = 0 \times \times 1 1 \times 0 1
$$

A distribution D over $\{0,1\}^n$ is ε -biased if it fools parities:

$$
S \neq \varnothing \implies \left| \mathbb{E} \left[\bigoplus_{i \in S} D_i \right] - \frac{1}{2} \right| \leq \varepsilon
$$

A distribution D over $\{0,1\}^n$ is ε -biased if it fools parities:

$$
S \neq \varnothing \implies \left| \mathbb{E} \left[\bigoplus_{i \in S} D_i \right] - \frac{1}{2} \right| \leq \varepsilon
$$

Eet D, D' be independent small-bias strings

A distribution D over $\{0,1\}^n$ is ε -biased if it fools parities:

$$
S \neq \varnothing \implies \left| \mathbb{E} \left[\bigoplus_{i \in S} D_i \right] - \frac{1}{2} \right| \leq \varepsilon
$$

Eet D, D' be independent small-bias strings

In Let $X = \text{Res}(D, D')$ (seed length $\widetilde{O}(\log n)$)

A distribution D over $\{0,1\}^n$ is ε -biased if it fools parities:

$$
S \neq \varnothing \implies \left| \mathbb{E} \left[\bigoplus_{i \in S} D_i \right] - \frac{1}{2} \right| \leq \varepsilon
$$

Eet D, D' be independent small-bias strings

- In Let $X = \text{Res}(D, D')$ (seed length $\widetilde{O}(\log n)$)
- **Theorem** [Forbes, Kelley '18]: For any $O(1)$ -width ROBP f,

E $\mathop{\mathbb{E}}_{X,U}[f|_X(U)] \approx \mathop{\mathbb{E}}_{U}[f(U)]$

A distribution D over $\{0,1\}^n$ is ε -biased if it fools parities:

$$
S \neq \varnothing \implies \left| \mathbb{E} \left[\bigoplus_{i \in S} D_i \right] - \frac{1}{2} \right| \leq \varepsilon
$$

Eet D, D' be independent small-bias strings

- In Let $X = \text{Res}(D, D')$ (seed length $\widetilde{O}(\log n)$)
- **Theorem** [Forbes, Kelley '18]: For any $O(1)$ -width ROBP f,

$$
\mathop{\mathbb{E}}_{X,U}[f|_X(U)] \approx \mathop{\mathbb{E}}_{U}[f(U)]
$$

In words, X preserves expectation of f
Forbes-Kelley pseudorandom restriction

A distribution D over $\{0,1\}^n$ is ε -biased if it fools parities:

$$
S \neq \varnothing \implies \left| \mathbb{E} \left[\bigoplus_{i \in S} D_i \right] - \frac{1}{2} \right| \leq \varepsilon
$$

Eet D, D' be independent small-bias strings

- In Let $X = \text{Res}(D, D')$ (seed length $\widetilde{O}(\log n)$)
- **Theorem** [Forbes, Kelley '18]: For any $O(1)$ -width ROBP f,

$$
\mathop{\mathbb{E}}_{X,U}[f|_X(U)] \approx \mathop{\mathbb{E}}_{U}[f(U)]
$$

In words, X preserves expectation of f

I (Proof involves clever Fourier analysis, building on [RSV13, HLV18, CHRT18])

So [FK18] can assign values to half the inputs using $\widetilde{O}(\log n)$ truly random bits

- So [FK18] can assign values to half the inputs using $\widetilde{O}(\log n)$ truly random bits
- After restricting, $f|_X$ is another ROBP

- \triangleright So [FK18] can assign values to half the inputs using $\widetilde{O}(\log n)$ truly random bits
- After restricting, $f|_X$ is another ROBP
- \triangleright So we can apply another pseudorandom restriction

- So [FK18] can assign values to half the inputs using $\widetilde{O}(\log n)$ truly random bits
- After restricting, $f|_X$ is another ROBP
- \triangleright So we can apply another pseudorandom restriction
- Elet $X^{\circ t}$ denote composition of t independent copies of X

- So [FK18] can assign values to half the inputs using $\widetilde{O}(\log n)$ truly random bits
- After restricting, $f|_X$ is another ROBP
- \triangleright So we can apply another pseudorandom restriction
- Elet $X^{\circ t}$ denote composition of t independent copies of X

Let
$$
t = O(\log n)
$$

- \triangleright So [FK18] can assign values to half the inputs using $O(\log n)$ truly random bits
- After restricting, $f|_X$ is another ROBP
- \triangleright So we can apply another pseudorandom restriction
- Elet $X^{\circ t}$ denote composition of t independent copies of X

Let
$$
t = O(\log n)
$$

▶ With high probability, $X^{\circ t} \in \{0, 1\}^n$ (no *)

- So [FK18] can assign values to half the inputs using $\widetilde{O}(\log n)$ truly random bits
- After restricting, $f|_X$ is another ROBP
- \triangleright So we can apply another pseudorandom restriction
- Elet $X^{\circ t}$ denote composition of t independent copies of X

Let
$$
t = O(\log n)
$$

▶ With high probability, $X^{\circ t} \in \{0, 1\}^n$ (no *)

 \blacktriangleright Expectation preserved at every step, so total error is low:

E $\mathop{\mathbb{E}}_{X^{\circ t}}[f(X^{\circ t})] \approx \mathop{\mathbb{E}}_{U}[f(U)]$

- So [FK18] can assign values to half the inputs using $\widetilde{O}(\log n)$ truly random bits
- After restricting, $f|_X$ is another ROBP
- \triangleright So we can apply another pseudorandom restriction
- Elet $X^{\circ t}$ denote composition of t independent copies of X

Let
$$
t = O(\log n)
$$

- ▶ With high probability, $X^{\circ t} \in \{0, 1\}^n$ (no *)
- \triangleright Expectation preserved at every step, so total error is low:

$$
\mathop{\mathbb{E}}_{X^{\circ t}}[f(X^{\circ t})] \approx \mathop{\mathbb{E}}_{U}[f(U)]
$$

 \blacktriangleright Total cost: $\widetilde{O}(\log^2 n)$ truly random bits

Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$

Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$

Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$

 \blacktriangleright Design X to preserve expectation

- Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$
- \blacktriangleright Design X to preserve expectation
- ▶ Design X so that $X^{\circ t}$ also simplifies formula, for $t \ll \log n$

- Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$
- \blacktriangleright Design X to preserve expectation
- ▶ Design X so that $X^{\circ t}$ also simplifies formula, for $t \ll \log n$

- Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$
- \blacktriangleright Design X to preserve expectation
- ▶ Design X so that $X^{\circ t}$ also simplifies formula, for $t \ll \log n$

- Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$
- \blacktriangleright Design X to preserve expectation
- ▶ Design X so that $X^{\circ t}$ also simplifies formula, for $t \ll \log n$

- Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$
- \blacktriangleright Design X to preserve expectation
- ▶ Design X so that $X^{\circ t}$ also simplifies formula, for $t \ll \log n$

- Step 1: Apply pseudorandom restriction $X \in \{0, 1, \star\}^n$
- \blacktriangleright Design X to preserve expectation
- ▶ Design X so that $X^{\circ t}$ also simplifies formula, for $t \ll \log n$

Step 2: Fool restricted formula, taking advantage of simplicity

 \triangleright Assume by recursion: PRG for depth d with seed length $\widetilde{O}(\log n)$

- \triangleright Assume by recursion: PRG for depth d with seed length $O(log n)$
- ► Let's sample $X \in \{0, 1, \star\}^n$ for depth $d+1$

- Assume by recursion: PRG for depth d with seed length $O(log n)$
- ► Let's sample $X \in \{0, 1, \star\}^n$ for depth $d+1$

1. Recursively sample $G_d, G'_d \in \{0,1\}^n$

- **Example 3** Assume by recursion: PRG for depth d with seed length $O(\log n)$
- ► Let's sample $X \in \{0, 1, \star\}^n$ for depth $d+1$

- 1. Recursively sample $G_d, G'_d \in \{0,1\}^n$
- 2. Sample $D, D' \in \{0, 1\}^n$ with small bias

- Assume by recursion: PRG for depth d with seed length $O(\log n)$
- ► Let's sample $X \in \{0, 1, \star\}^n$ for depth $d+1$

- 1. Recursively sample $G_d, G'_d \in \{0,1\}^n$
- 2. Sample $D, D' \in \{0, 1\}^n$ with small bias
- 3. $X = \text{Res}(G_d \oplus D, G'_d \oplus D')$

Preserving expectation

Claim: For any depth- $(d + 1)$ read-once AC^0 formula f,

E $\mathop{\mathbb{E}}_{X,U}[f|_X(U)] \approx \mathop{\mathbb{E}}_{U}[f(U)]$

Preserving expectation

Claim: For any depth- $(d + 1)$ read-once AC^0 formula f,

$$
\mathop{\mathbb{E}}_{X,U}[f|_X(U)] \approx \mathop{\mathbb{E}}_{U}[f(U)]
$$

Proof: Read-once AC^0 can be simulated by constant-width ROBPs [CSV15]

Preserving expectation

Claim: For any depth- $(d + 1)$ read-once AC^0 formula f,

$$
\mathop{\mathbb{E}}_{X,U}[f|_X(U)] \approx \mathop{\mathbb{E}}_{U}[f(U)]
$$

- **Proof**: Read-once AC^0 can be simulated by constant-width ROBPs [CSV15]
- \triangleright So we can simply apply Forbes-Kelley result:

$$
X = \text{Res}(G_d \oplus D, G'_d \oplus D')
$$

Simplification

\triangleright $\Delta(f) \stackrel{\text{def}}{=}$ maximum fan-in of any gate other than root

Simplification

 \triangleright $\Delta(f) \stackrel{\text{def}}{=}$ maximum fan-in of any gate other than root

• Main Lemma: With high probability over $X^{\circ t}$,

 $\Delta(f |_{X^{\circ t}}) \leq$ polylog n,

where $t = O((\log \log n)^2)$

Simplification

 \triangleright $\Delta(f) \stackrel{\text{def}}{=}$ maximum fan-in of any gate other than root

• Main Lemma: With high probability over $X^{\circ t}$,

 $\Delta(f|_{X^{\circ t}}) \leq$ polylog n,

where $t = O((\log \log n)^2)$

Actually we only prove this statement "up to sandwiching"

Let f be a read-once AC^0 formula

Let
$$
f
$$
 be a read-once AC^0 formula

lacktriangleright Let $R = \text{Res}(U, U')$ (truly random restriction)

Let
$$
f
$$
 be a read-once AC^0 formula

lacktriangleright Let $R = \text{Res}(U, U')$ (truly random restriction)

▶ Chen, Steinke, Vadhan '15 \implies W.h.p. over $R^{\circ t}$,

 $\Delta(f|_{R^{\circ t}}) \leq$ polylog n

- Let f be a read-once AC^0 formula
- lacktriangleright Let $R = \text{Res}(U, U')$ (truly random restriction)
- ▶ Chen, Steinke, Vadhan '15 \implies W.h.p. over $R^{\circ t}$,

 $\Delta(f|_{R^{\circ t}}) \leq$ polylog n

In fact the simplification they show is more severe)

Let
$$
f
$$
 be a read-once AC^0 formula

- lacktriangleright Let $R = \text{Res}(U, U')$ (truly random restriction)
- ▶ Chen, Steinke, Vadhan '15 \implies W.h.p. over $R^{\circ t}$,

$$
\Delta(f|_{R^{\circ t}}) \leq \text{polylog } n
$$

In fact the simplification they show is more severe)

Again, these statements are true "up to sandwiching." Proof uses Fourier analysis

Derandomizing simplification

► Let f be a depth- $(d-1)$ read-once AC^0 formula
Derandomizing simplification

► Let f be a depth- $(d-1)$ read-once AC^0 formula let $b \in \{0, 1\}$

Derandomizing simplification

- ► Let f be a depth- $(d-1)$ read-once AC^0 formula
- let $b \in \{0,1\}$
- ▶ Computational problem: Given $y, z \in \{0, 1\}^n$, decide whether

$$
f|_{\mathsf{Res}(y,z)} \equiv b
$$

Derandomizing simplification

- ► Let f be a depth- $(d-1)$ read-once AC^0 formula
- let $b \in \{0,1\}$
- ▶ Computational problem: Given $y, z \in \{0, 1\}^n$, decide whether

$$
f|_{\mathsf{Res}(y,z)} \equiv b
$$

Lemma: Can be decided in depth-d read-once AC^0

Deciding whether $f|_{\text{Res}(y,z)} \equiv b$

Deciding whether $f|_{\text{Res}(y,z)} \equiv b$

Deciding whether $f|_{\text{Res}(y,z)} \equiv b$

Deciding whether $f|_{\text{Res}(y,z)} \equiv b$ (continued)

 \blacktriangleright At bottom, we get one additional layer:

$$
(\mathsf{Res}(y, z)_i \equiv b) \iff (y_i = 0 \land z_i = b)
$$

$$
(\neg \mathsf{Res}(y, z)_i \equiv b) \iff (y_i = 0 \land z_i = 1 - b)
$$

► Let f be a depth- $(d-1)$ read-once \mathbf{AC}^0 formula

let $b \in \{0, 1\}$

► Let f be a depth- $(d-1)$ read-once \mathbf{AC}^0 formula

$$
\blacktriangleright \ \text{Let} \ b \in \{0,1\}
$$

 \blacktriangleright $X = \text{Res}(G_d \oplus D, G'_d \oplus D')$

► Let f be a depth- $(d-1)$ read-once \mathbf{AC}^0 formula

$$
\blacktriangleright \ \text{Let} \ b \in \{0,1\}
$$

- \blacktriangleright $X = \text{Res}(G_d \oplus D, G'_d \oplus D')$
- G_d, G'_{d} fool depth d, so

$$
\Pr_X[f|_X \equiv b] \approx \Pr_R[f|_R \equiv b]
$$

► Let f be a depth- $(d-1)$ read-once \mathbf{AC}^0 formula

$$
\blacktriangleright \ \mathsf{Let} \ b \in \{0,1\}
$$

- \blacktriangleright $X = \text{Res}(G_d \oplus D, G'_d \oplus D')$
- G_d, G'_{d} fool depth d, so

$$
\Pr_X[f|_X \equiv b] \approx \Pr_R[f|_R \equiv b]
$$

 \blacktriangleright Hybrid argument:

$$
\Pr_{X^{\circ t}}[f|_{X^{\circ t}} \equiv b] \approx \Pr_{R^{\circ t}}[f|_{R^{\circ t}} \equiv b]
$$

Bridging the gap from $d-1$ to $d+1$

► So far: Depth- $(d-1)$ formulas collapse with about the right probability

Bridging the gap from $d-1$ to $d+1$

- ► So far: Depth- $(d-1)$ formulas collapse with about the right probability
- \triangleright We were supposed to show that depth- $(d+1)$ formulas simplify w.r.t. Δ w.h.p.

Idea of proof that $\Delta \mapsto$ polylog n

Idea of proof that $\Delta \mapsto$ polylog n

Idea of proof that $\Delta \mapsto$ polylog n

Δ = polylog *n*

► To recap, after $t = O((\log \log n)^2)$ restrictions, $\Delta =$ polylog n

Δ = polylog *n*

Δ = polylog *n*

► To recap, after $t = O((\log \log n)^2)$ restrictions, $\Delta =$ polylog n \triangleright Total cost so far: $\widetilde{O}(\log n)$ truly random bits

▶ Theorem (Meka, Reingold, Tal '19): There is an explicit PRG with seed length $\tilde{O}(\log(n/\varepsilon))$ that fools functions of the form

$$
f=\bigoplus_{i=1}^m f_i,
$$

 \triangleright Theorem (Meka, Reingold, Tal '19): There is an explicit PRG with seed length $O(log(n/\varepsilon))$ that fools functions of the form

$$
f=\bigoplus_{i=1}^m f_i,
$$

where f_1, \ldots, f_m are on disjoint variables and f_i can be computed by an ROBP with width $O(1)$, length polylog n

 \triangleright Theorem (Meka, Reingold, Tal '19): There is an explicit PRG with seed length $O(log(n/\varepsilon))$ that fools functions of the form

$$
f=\bigoplus_{i=1}^m f_i,
$$

where f_1, \ldots, f_m are on disjoint variables and f_i can be computed by an ROBP with width $O(1)$, length polylog n

I (Proof uses GMRTV approach, building on [GY14, CHRT18, Vio09])

 \triangleright Theorem (Meka, Reingold, Tal '19): There is an explicit PRG with seed length $O(log(n/\varepsilon))$ that fools functions of the form

$$
f=\bigoplus_{i=1}^m f_i,
$$

where f_1, \ldots, f_m are on disjoint variables and f_i can be computed by an ROBP with width $O(1)$, length polylog n

I (Proof uses GMRTV approach, building on [GY14, CHRT18, Vio09])

 \blacktriangleright In our case.

$$
f=\bigwedge_{i=1}^m f_i
$$

 \triangleright Theorem (Meka, Reingold, Tal '19): There is an explicit PRG with seed length $O(log(n/\varepsilon))$ that fools functions of the form

$$
f=\bigoplus_{i=1}^m f_i,
$$

where f_1, \ldots, f_m are on disjoint variables and f_i can be computed by an ROBP with width $O(1)$, length polylog n

(Proof uses GMRTV approach, building on [GY14, CHRT18, Vio09])

 \blacktriangleright In our case.

$$
f = \bigwedge_{i=1}^{m} f_i = \sum_{S \subseteq [m]} \frac{(-1)^{|S|}}{2^m} \prod_{i \in S} (-1)^{f_i}
$$

Directions for further research

Directions for further research

Directions for further research

Read-once $\mathbf{AC}^0[\oplus]$

In Natural next step toward derandomizing BPL

In Natural next step toward derandomizing BPL

Best prior PRG: seed length $\widetilde{O}(\log^2 n)$ [FK '18]

In Natural next step toward derandomizing BPL

- Best prior PRG: seed length $\widetilde{O}(\log^2 n)$ [FK '18]
- ▶ Theorem: Our PRG fools read-once $AC^0[\oplus]$ with seed length

 $\widetilde{O}(t + \log(n/\varepsilon))$

where $t = #$ parity gates

▶ Natural next step toward derandomizing **BPL**

- Best prior PRG: seed length $\widetilde{O}(\log^2 n)$ [FK '18]
- ▶ Theorem: Our PRG fools read-once $AC^0[\oplus]$ with seed length $\widetilde{O}(t + \log(n/\varepsilon))$

where $t = #$ parity gates

▶ Fool read-once $AC^0[\oplus]$ with seed length $\widetilde{O}(\log(n/\varepsilon))$?

▶ Natural next step toward derandomizing **BPL**

- Best prior PRG: seed length $\widetilde{O}(\log^2 n)$ [FK '18]
- ▶ Theorem: Our PRG fools read-once $AC^0[\oplus]$ with seed length $\widetilde{O}(t + \log(n/\varepsilon))$

where $t = #$ parity gates

▶ Fool read-once $AC^0[\oplus]$ with seed length $\widetilde{O}(\log(n/\varepsilon))$?

Fianks! Questions?