On Raz and Reingold PRG and A New White Box WPRG, and Error Reduction For Weighted PRGs Against Read Once Branching Programs

Master's Thesis Presentation

Oren Renard
Advisor: Dr. Gil Cohen

Tel-Aviv University

October 12, 2021

Table of Contents

BPL vs. L

Definition (Space Bounded Classes)

1. L is the class of all languages decidable by logarithmic space TM,
2. BPL, RL defined as all languages that are decidable by logarithmic space probabilistic TM with *two* or *one* sided error, respectively.

BPL vs. L

Definition (Space Bounded Classes)

1. L is the class of all languages decidable by logarithmic space TM,
2. BPL, RL defined as all languages that are decidable by logarithmic space probabilistic TM with *two* or *one* sided error, respectively.

The problem

What are the relations between BPL, RL and L?

Read Once Branching Programs

The uniform PTM model is difficult to work with directly.

Lemma

Every PTM with fixed input length n that uses space $s(n)$ can be represented by an $(n, w = 2^{O(s(n))})$ Read Once Branching Program (ROBP, BP).

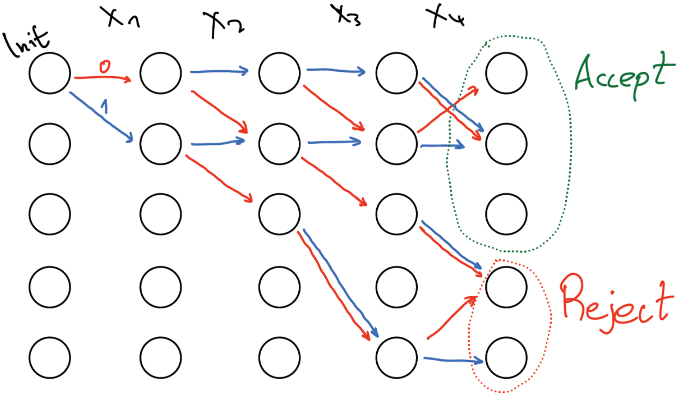


Figure: An $(n = 4, w = 5)$ BP.

Approaches for Derandomization

Goal: approximate acceptance probability of all (n, w) machines.

Approaches for Derandomization

Goal: approximate acceptance probability of all (n, w) machines.

Definition (Der. Algorithm)

Given an (n, w) BP, approximate its acceptance probability up to error ε.

This is a **White Box** technique, as its allowed to *inspect* the input.

Approaches for Derandomization (cont.)

Another approach is in oblivious manner, i.e. **Black Box** techniques.

Definition (PRG)

A function $G : \{0,1\}^s \to \{0,1\}^n$ is called an (n, w, ε) PRG, if for every (n, w) BP M,

$$|\mathbb{E}[M(U_n)] - \mathbb{E}[M(G(U_s))]| \leq \varepsilon.$$

Approaches for Derandomization (cont.)

Another approach is in oblivious manner, i.e. **Black Box** techniques.

Definition (PRG)

A function $G : \{0,1\}^s \to \{0,1\}^n$ is called an (n, w, ε) PRG, if for every (n, w) BP M,

$$|\mathbb{E}[M(U_n)] - \mathbb{E}[M(G(U_s))]| \leq \varepsilon.$$

By the probabilistic method, $\exists\ (n, w, \varepsilon)$ PRG with seed length $s = O(\log n + \log w + \log \varepsilon^{-1})$.

Approaches for Derandomization (cont.)

Another approach is in oblivious manner, i.e. **Black Box** techniques.

Definition (PRG)

A function $G : \{0,1\}^s \to \{0,1\}^n$ is called an (n, w, ε) PRG, if for every (n, w) BP M,

$$|\mathbb{E}[M(U_n)] - \mathbb{E}[M(G(U_s))]| \leq \varepsilon.$$

By the probabilistic method, $\exists$ (n, w, ε) PRG with seed length $s = O(\log n + \log w + \log \varepsilon^{-1})$.

Definition (WPRG)

A pair of functions $(G, \mu) : \{0,1\}^s \to \{0,1\}^n \times \mathbb{R}$ are called (n, w, ε) Weighted-PRG, if for every (n, w) BP M,

$$\left| \mathbb{E}[M(U_n)] - \mathop{\mathbb{E}}_{x \sim U_s} [\mu(x) \cdot M(G(x))] \right| \leq \varepsilon.$$

Approaches for Derandomization (cont.)

Another approach is in oblivious manner, i.e. **Black Box** techniques.

Definition (PRG)

A function $G : \{0,1\}^s \to \{0,1\}^n$ is called an (n, w, ε) PRG, if for every (n, w) BP M,

$$|\mathbb{E}[M(U_n)] - \mathbb{E}[M(G(U_s))]| \leq \varepsilon.$$

By the probabilistic method, $\exists$ (n, w, ε) PRG with seed length $s = O(\log n + \log w + \log \varepsilon^{-1})$.

Definition (WPRG)

A pair of functions $(G, \mu) : \{0,1\}^s \to \{0,1\}^n \times \mathbb{R}$ are called (n, w, ε) Weighted-PRG, if for every (n, w) BP M,

$$\left|\mathbb{E}[M(U_n)] - \mathop{\mathbb{E}}_{x \sim U_s}[\mu(x) \cdot M(G(x))]\right| \leq \varepsilon.$$

For every machine M, the Derandomization via PRG/WPRG (G, μ) follows as:

1. Enumerate seeds $x \in \{0,1\}^s$,
2. Compute $G(x)$,
3. Avg. the result of $M(G(x))$,
   - For WPRG: consider the *weighted* avg.

Approaches for Derandomization (cont.)

Another approach is in oblivious manner, i.e. **Black Box** techniques.

Definition (PRG)

A function $G : \{0,1\}^s \to \{0,1\}^n$ is called an (n, w, ε) PRG, if for every (n, w) BP M,

$$|\mathbb{E}[M(U_n)] - \mathbb{E}[M(G(U_s))]| \leq \varepsilon.$$

By the probabilistic method, $\exists$ (n, w, ε) PRG with seed length $s = O(\log n + \log w + \log \varepsilon^{-1})$.

Definition (WPRG)

A pair of functions $(G, \mu) : \{0,1\}^s \to \{0,1\}^n \times \mathbb{R}$ are called (n, w, ε) Weighted-PRG, if for every (n, w) BP M,

$$\left| \mathbb{E}[M(U_n)] - \mathop{\mathbb{E}}_{x \sim U_s} [\mu(x) \cdot M(G(x))] \right| \leq \varepsilon.$$

For every machine M, the Derandomization via PRG/WPRG (G, μ) follows as:

1. Enumerate seeds $x \in \{0,1\}^s$,
2. Compute $G(x)$,
3. Avg. the result of $M(G(x))$,
   - For WPRG: consider the *weighted* avg.

$\Longrightarrow$ Computing $\mathbb{E}[M(U_n)] \pm \varepsilon$ takes space: $space(G) + seed(G)$

Brief History of Derandomization

Space	Ref
$O(\lg n \cdot \lg \frac{nw}{\varepsilon})$	Sav 70', BCP 83', Nis 92
$O(\sqrt{\lg n} \cdot \lg \frac{nw}{\varepsilon})$	SZ 95'
$O(\sqrt{\lg n} \cdot \lg(nw) + \lg\lg_{nw} \frac{1}{\varepsilon})$	AKMPVS 21
$O(\sqrt{\lg n} \cdot \lg(\frac{nw}{\varepsilon}) \cdot \frac{1}{\sqrt{\lg\lg n}})$	Hoza 21

Der. of (n,w) BP

Seed	Type	Ref
$O(\lg n \cdot \lg \frac{nw}{\varepsilon})$	PRG	Nis 92', INW 94'
$O\left(\frac{\lg n \cdot \lg \frac{nw}{\varepsilon}}{\max\{1, \lg\lg w - \lg\lg \frac{n}{\varepsilon}\}}\right)$	PRG	Arm 98'
$\tilde{O}(\lg n \cdot \lg(nw) + \lg \frac{1}{\varepsilon})$	WPRG	BCG 18', CL20', CDRSTS 21', PV 21'
$O(\lg n \cdot \lg(nw) + \lg \frac{1}{\varepsilon})$	WPRG	Hoza 21'

PRG & WPRG for (n,w) BP

Brief History of Derandomization

Space	Ref
$O\left(\lg n \cdot \lg \frac{nw}{\varepsilon}\right)$	Sav 70', BCP 83', Nis 92
$O\left(\sqrt{\lg n} \cdot \lg \frac{nw}{\varepsilon}\right)$	SZ 95'
$O\left(\sqrt{\lg n} \cdot \lg(nw) + \lg\lg_{nw} \frac{1}{\varepsilon}\right)$	AKMPVS 21
$O\left(\sqrt{\lg n} \cdot \lg\left(\frac{nw}{\varepsilon}\right) \cdot \frac{1}{\sqrt{\lg\lg n}}\right)$	Hoza 21

Der. of (n,w) BP

Seed	Type	Ref
$O\left(\lg n \cdot \lg \frac{nw}{\varepsilon}\right)$	PRG	Nis 92', INW 94'
$O\left(\frac{\lg n \cdot \lg \frac{nw}{\varepsilon}}{\max\{1, \lg\lg w - \lg\lg \frac{n}{\varepsilon}\}}\right)$	PRG	Arm 98'
$\tilde{O}\left(\lg n \cdot \lg(nw) + \lg \frac{1}{\varepsilon}\right)$	WPRG	BCG 18, CL 20', CDRSTS 21', PV 21'
$O\left(\lg n \cdot \lg(nw) + \lg \frac{1}{\varepsilon}\right)$	WPRG	Hoza 21'

PRG & WPRG for (n,w) BP

Table of Contents

Rough recipe for PRGs by [Nis92; INW94]

Simplification: (1) all the layers of M are equal, (2) denote $M \in \mathbb{R}^{w \times w}$ as the transition matrix.

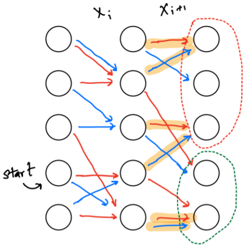
Rough recipe for PRGs by [Nis92; INW94]

Simplification: (1) all the layers of M are equal, (2) denote $M \in \mathbb{R}^{w \times w}$ as the transition matrix.
One step. Prove the Recycle lemma for arbitrary $w \in \mathbb{N}$ and $\varepsilon_{\mathcal{P}} > 0$:

Recycle Lemma

$\exists$ an explicit construction of pseudo random family $\mathcal{P}$ s.t. for every $(2, w)$ BP M,

$$p \sim U_{\mathcal{P}} \implies M_p \approx_{\varepsilon_{\mathcal{P}}} M^2$$



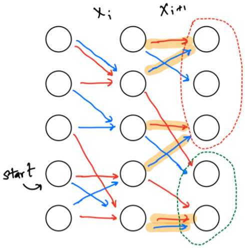
Rough recipe for PRGs by [Nis92; INW94]

Simplification: (1) all the layers of M are equal, (2) denote $M \in \mathbb{R}^{w \times w}$ as the transition matrix.
One step. Prove the Recycle lemma for arbitrary $w \in \mathbb{N}$ and $\varepsilon_{\mathcal{P}} > 0$:

Recycle Lemma

$\exists$ an explicit construction of pseudo random family $\mathcal{P}$ s.t. for every $(2, w)$ BP M,

$$p \sim U_{\mathcal{P}} \quad \Longrightarrow \quad M_p \approx_{\varepsilon_{\mathcal{P}}} M^2$$



Deduce PRG recursively.
Sample *independently* $\overline{p} = (p_1, \ldots, p_h)$ so for every $(2^h, w)$ BP M,

$$M_{p_1,\ldots,p_h} \approx_{\varepsilon(1)} M^2_{p_1,\ldots,p_{h-1}} \approx_{\varepsilon(2)} \ldots \approx_{\varepsilon(h)} M^{2^h}$$

Rough recipe for PRGs by [Nis92; INW94] (cont.)

Analysis. The family size in both constructions is

$$\log|\mathcal{P}| = O(\log w + \log \varepsilon_{\mathcal{P}}^{-1}).$$

Rough recipe for PRGs by [Nis92; INW94] (cont.)

Analysis. The family size in both constructions is

$$\log |\mathcal{P}| = O(\log w + \log \varepsilon_{\mathcal{P}}^{-1}).$$

At level h they claim that $M_{p_1,\dots p_h} \approx_{\varepsilon(h)} M^{2^h}$, where

$$\varepsilon(h) = 2\varepsilon(h-1) + \varepsilon_{\mathcal{P}}.$$

Thus $\varepsilon \stackrel{\text{def}}{=} \varepsilon(\log n) = n \cdot \varepsilon_{\mathcal{P}}$, so the seed length becomes

$$\begin{aligned} s &= \log n \cdot (\log |\mathcal{P}|) \\ &= O(\log n \cdot (\log w + \log \varepsilon_{\mathcal{P}}^{-1})) \\ &= O(\log n \cdot (\log n + \log w + \log \varepsilon^{-1})) \end{aligned}$$

Rough recipe for PRGs by [Nis92; INW94] (cont.)

Analysis. The family size in both constructions is

$$\log |\mathcal{P}| = O(\log w + \log \varepsilon_{\mathcal{P}}^{-1}).$$

At level h they claim that $M_{p_1,\ldots p_h} \approx_{\varepsilon(h)} M^{2^h}$, where

$$\varepsilon(h) = 2\varepsilon(h-1) + \varepsilon_{\mathcal{P}}.$$

Thus $\varepsilon \overset{\text{def}}{=} \varepsilon(\log n) = n \cdot \varepsilon_{\mathcal{P}}$, so the seed length becomes

$$\begin{aligned} s &= \log n \cdot (\log |\mathcal{P}|) \\ &= O(\log n \cdot (\log w + \log \varepsilon_{\mathcal{P}}^{-1})) \\ &= O(\log n \cdot (\log n + \log w + \log \varepsilon^{-1})) \end{aligned}$$

Gil's Road map to space bounded computation

Improving Nisan [Nis92] PRG.
Better analysis of the error may lead to PRG with seed length

$$s_{\text{romantic}} = O(\log n \cdot (\log w) + \log \varepsilon^{-1})$$

Gil's Road map to space bounded computation

Improving Nisan [Nis92] PRG.
Better analysis of the error may lead to PRG with seed length

$$\mathsf{s}_{\mathsf{romantic}} = O(\log n \cdot (\log w) + \log \varepsilon^{-1})$$

Improving Saks and Zhou [SZ99] derandomization.

1. The WPRG of Braverman, Cohen, and Garg [BCG18] already has seed length

$$\mathsf{s}_{\mathsf{BCG}} = \widetilde{O}(\log n \cdot (\log n + \log w) + \log \varepsilon^{-1}),$$

2. while Raz and Reingold [RR99] obtained *conditional* PRG with seed length

$$\mathsf{s}_{\mathsf{RR}} = \widetilde{O}(\log n \cdot (\log n + \log \varepsilon^{-1}) + \log w).$$

Gil's Road map to space bounded computation

Improving Nisan [Nis92] PRG.
Better analysis of the error may lead to PRG with seed length

$$s_{\text{romantic}} = O(\log n \cdot (\log w) + \log \varepsilon^{-1})$$

Improving Saks and Zhou [SZ99] derandomization.

1. The WPRG of Braverman, Cohen, and Garg [BCG18] already has seed length

$$s_{\text{BCG}} = \widetilde{O}(\log n \cdot (\log n + \log w) + \log \varepsilon^{-1}),$$

2. while Raz and Reingold [RR99] obtained *conditional* PRG with seed length

$$s_{\text{RR}} = \widetilde{O}(\log n \cdot (\log n + \log \varepsilon^{-1}) + \log w).$$

So maybe combine somehow [BCG18] and [RR99] to get

$$s_{\text{hopefully}} = \widetilde{O}(\log n \cdot (\log n) + \log w + \log \varepsilon^{-1}).$$

Plugging such a good seeded PRG into [SZ99] framework would yield $\mathsf{BPL} \subseteq \mathsf{L}^{4/3}$.

Contribution summary

Theorem ([CDR+21; PV21])

Let $G_0 : \{0,1\}^{s_0} \to \{0,1\}^n$ *be an* $(n, w, \varepsilon_0 = 1/n^2)$ *Black Box PRG. Then, for every error parameter* $0 < \varepsilon < \varepsilon_0$ *there exists an* (n, w, ε) *Black Box WPRG with seed length*

$$s_0 + O\left((\log w + \log \varepsilon^{-1}) \cdot \log\log_n(1/\varepsilon)\right)$$

computable in space $O\left(space(G_0) + \log\log_n(1/\varepsilon) \cdot (\log\log(w/\varepsilon))^2\right)$.

Contribution summary

Theorem ([CDR+21; PV21])

Let $G_0 : \{0,1\}^{s_0} \to \{0,1\}^n$ *be an* $(n, w, \varepsilon_0 = 1/n^2)$ *Black Box PRG. Then, for every error parameter* $0 < \varepsilon < \varepsilon_0$ *there exists an* (n, w, ε) *Black Box WPRG with seed length*

$$s_0 + O\left((\log w + \log \varepsilon^{-1}) \cdot \log\log_n(1/\varepsilon)\right)$$

computable in space $O\left(space(G_0) + \log\log_n(1/\varepsilon) \cdot (\log\log(w/\varepsilon))^2\right)$.

Re organization of [RR99]

1. An attempt to simplify the proof of [RR99]
2. Reducing their conditional result to **explicit** one

Contribution summary

Theorem ([CDR+21; PV21])

Let $G_0 : \{0,1\}^{s_0} \to \{0,1\}^n$ be an $(n, w, \varepsilon_0 = 1/n^2)$ Black Box PRG. Then, for every error parameter $0 < \varepsilon < \varepsilon_0$ there exists an (n, w, ε) Black Box WPRG with seed length

$$s_0 + O\left((\log w + \log \varepsilon^{-1}) \cdot \log\log_n(1/\varepsilon)\right)$$

computable in space $O\left(space(G_0) + \log\log_n(1/\varepsilon) \cdot (\log\log(w/\varepsilon))^2\right)$.

Re organization of [RR99]

1. An attempt to simplify the proof of [RR99]
2. Reducing their conditional result to **explicit** one

Theorem

There exists an (n, w, ε) White Box WPRG with seed length

$$\mathsf{s} = \widetilde{O}(\log n \cdot (\log n) + \log w + \log \varepsilon^{-1}),$$

that is computable in space

$$\widetilde{O}(\log n \cdot (\log n) + \sqrt{\log n} \cdot (\log w) + \log \varepsilon^{-1}).$$

Contribution summary

Theorem ([CDR+21; PV21])

Let $G_0 : \{0,1\}^{s_0} \to \{0,1\}^n$ *be an* $(n, w, \varepsilon_0 = 1/n^2)$ *Black Box PRG. Then, for every error parameter* $0 < \varepsilon < \varepsilon_0$ *there exists an* (n, w, ε) *Black Box WPRG with seed length*

$$s_0 + O\left((\log w + \log \varepsilon^{-1}) \cdot \log \log_n (1/\varepsilon)\right)$$

computable in space $O\left(space(G_0) + \log\log_n(1/\varepsilon) \cdot (\log\log(w/\varepsilon))^2\right)$.

Re organization of [RR99]

1. An attempt to simplify the proof of [RR99]
2. Reducing their conditional result to **explicit** one

Theorem

There exists an (n, w, ε) *White Box WPRG with seed length*

$$\mathsf{s} = \widetilde{O}(\log n \cdot (\log n) + \log w + \log \varepsilon^{-1}),$$

that is computable in space

$$\widetilde{O}(\log n \cdot (\log n) + \sqrt{\log n} \cdot (\log w) + \log \varepsilon^{-1}).$$

Table of Contents

Table of Contents

Brief exposition of Extractors

Definition

A distribution $X \subseteq \{0,1\}^{n}$ is called (n, k) **source** if

$$\max_{x \in \mathrm{supp}(X)} \Pr[X = x] \leq 2^{-k}.$$

Brief exposition of Extractors

Definition

A distribution $X \subseteq \{0,1\}^n$ is called (n,k) **source** if

$$\max_{x \in \mathrm{supp}(X)} \Pr[X = x] \leq 2^{-k}.$$

Definition

$\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is called (k,ε) **extractor** if for every (n,k) source X,

$$\mathrm{Ext}(X, U_d) \approx_\varepsilon U_m.$$

Brief exposition of Extractors

Definition

A distribution $X \subseteq \{0,1\}^n$ is called (n,k) **source** if

$$\max_{x \in \text{supp}(X)} \Pr[X = x] \le 2^{-k}.$$

Definition

$\text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is called (k,ε) **extractor** if for every (n,k) source X,

$$\text{Ext}(X, U_d) \approx_\varepsilon U_m.$$

Theorem (Lower bound)

A (k,ε) extractor is **optimal** *if*

$$d = O(\log(n/\varepsilon))$$
$$m = k + d - O(\log \varepsilon^{-1}).$$

For simplicity we assume such extractors are fully explicit (i.e. computable in linear space)...

Toy example of [INW94]

We focus on fooling $(n+m, w)$ BPs.

1. Let Ext be a $(k_{\mathrm{INW}} = n - (\log w - \log \varepsilon^{-1} - 1), \varepsilon/2)$ *optimal* extractor, where

$$\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m.$$

2. The PRG $\mathrm{INW} : \{0,1\}^{n+d} \to \{0,1\}^{n+m}$ is defined as

$$\mathrm{INW}(x \circ y) = x \circ \mathrm{Ext}(x,y).$$

Toy example of [INW94]

We focus on fooling $(n+m, w)$ BPs.

1. Let Ext be a $(k_{\mathrm{INW}} = n - (\log w - \log \varepsilon^{-1} - 1), \varepsilon/2)$ *optimal* extractor, where

$$\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m.$$

2. The PRG $\mathrm{INW} : \{0,1\}^{n+d} \to \{0,1\}^{n+m}$ is defined as

$$\mathrm{INW}(x \circ y) = x \circ \mathrm{Ext}(x,y).$$

Claim

INW is an ε-PRG, i.e. for every $(n+m, w)$ BP M,

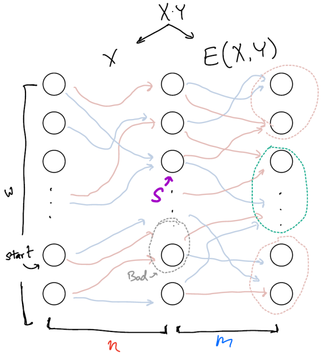
$$|\Pr[M(U_{n+m}) \text{ acc}] - \Pr[M(\mathrm{INW}(U_{n+d})) \text{ acc}]| \leq \varepsilon.$$

Toy example of [INW94] (cont.)

INW PRG

Let Ext be $(k_{\mathrm{INW}} = n - (\log w - \log \varepsilon^{-1} - 1), \varepsilon/2)$ extractor.
Then, $\mathrm{INW}(x \circ y) = x \circ \mathrm{Ext}(x, y)$.

Analysis.
Let $X \circ Y \sim U_{n+d}$, and M be some BP. Let $s \sim M(X)$.



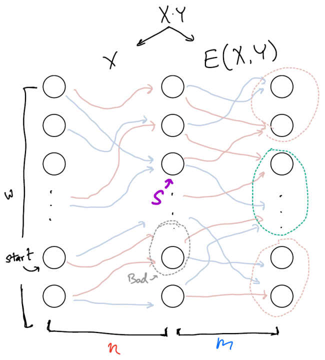
Toy example of [INW94] (cont.)

INW PRG

Let Ext be $(k_{\mathrm{INW}} = n - (\log w - \log \varepsilon^{-1} - 1), \varepsilon/2)$ extractor.
Then, $\mathrm{INW}(x \circ y) = x \circ \mathrm{Ext}(x, y)$.

Analysis.
Let $X \circ Y \sim U_{n+d}$, and M be some BP. Let $s \sim M(X)$.
Define

$$\mathrm{Bad} \stackrel{\text{def}}{=} \{s : \Pr[M(X) = s] < \varepsilon/2w\}.$$



Toy example of [INW94] (cont.)

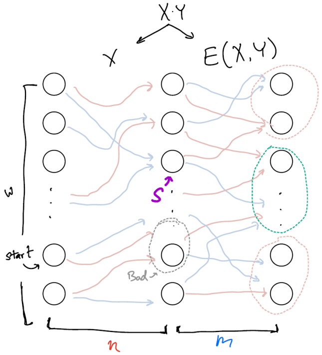
INW PRG

Let Ext be $(k_{\text{INW}} = n - (\log w - \log \varepsilon^{-1} - 1), \varepsilon/2)$ extractor.
Then, $\text{INW}(x \circ y) = x \circ \text{Ext}(x, y)$.

Analysis.
Let $X \circ Y \sim U_{n+d}$, and M be some BP. Let $s \sim M(X)$.
Define

$$\text{Bad} \stackrel{\text{def}}{=} \{s : \Pr[M(X) = s] < \varepsilon/2w\}.$$

Define X_s as the uniform dist. over $M^{-1}(s)$.



Toy example of [INW94] (cont.)

INW PRG

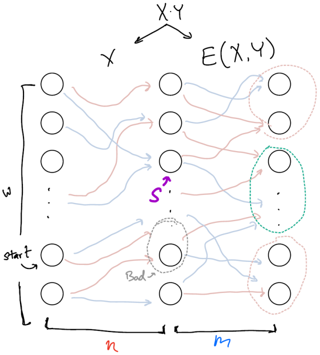
Let Ext be $(k_{\mathrm{INW}} = n - (\log w - \log \varepsilon^{-1} - 1), \varepsilon/2)$ extractor.
Then, $\mathrm{INW}(x \circ y) = x \circ \mathrm{Ext}(x, y)$.

Analysis.
Let $X \circ Y \sim U_{n+d}$, and M be some BP. Let $s \sim M(X)$.
Define

$$\mathrm{Bad} \overset{\text{def}}{=} \{s : \Pr[M(X) = s] < \varepsilon/2w\}.$$

Define X_s as the uniform dist. over $M^{-1}(s)$.
It is not hard to prove:

$$s \notin \mathrm{Bad} \implies X_s \text{ is an } (n, k_{\mathrm{INW}}) \text{ source.}$$



Toy example of [INW94] (cont.)

INW PRG

Let Ext be $(k_{\mathrm{INW}} = n - (\log w - \log \varepsilon^{-1} - 1), \varepsilon/2)$ extractor.
Then, $\mathrm{INW}(x \circ y) = x \circ \mathrm{Ext}(x, y)$.

$$\mathrm{Bad} \stackrel{\text{def}}{=} \{s : \Pr[M(X) = s] < \varepsilon/2w\}.$$

Toy example of [INW94] (cont.)

INW PRG

Let Ext be $(k_{\mathrm{INW}} = n - (\log w - \log \varepsilon^{-1} - 1), \varepsilon/2)$ extractor.
Then, $\mathrm{INW}(x \circ y) = x \circ \mathrm{Ext}(x, y)$.

$$\mathrm{Bad} \stackrel{\text{def}}{=} \{s : \Pr[M(X) = s] < \varepsilon/2w\}.$$

Thus,

$$\begin{aligned}
&|\Pr[M(U_{n+m}) \text{ acc}] - \Pr[M(\mathrm{INW}(U_{n+d})) \text{ acc}]| \\
&\quad = |\Pr[M(X \circ U_m) \text{ acc}] - \Pr[M(\mathrm{INW}(X, Y)) \text{ acc}]| \\
&\quad = \left| \sum_{s \in [w]} \Pr[M(X) = s] \cdot (\Pr[M_s(U_m) \text{ acc}] - \Pr[M_s(\mathrm{Ext}(X_s, Y)) \text{ acc}]) \right| \\
&\quad \le \left| \sum_{s \notin \mathrm{Bad}} \Pr[M(X) = s] \cdot \mathsf{SD}(U_m, \mathrm{Ext}(X_s, Y)) \right| + \left| \sum_{s \in \mathrm{Bad}} \Pr[M(X) = s] \cdot 1 \right| \\
&\quad \le 1 \cdot \frac{\varepsilon}{2} + w \cdot \frac{\varepsilon}{2w} \\
&\quad \le \varepsilon.
\end{aligned}$$

Full construction of [INW94]

The PRG $\mathrm{INW}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ defined as

$$\mathrm{INW}_{h+1}(x \circ y) \stackrel{\text{def}}{=} \mathrm{INW}(x) \circ \mathrm{INW}(\mathrm{Ext}_h(x,y))$$
$$\mathrm{INW}_1(x) \stackrel{\text{def}}{=} x_0$$

where $\mathrm{Ext}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ is a $(k_{h+1}, \varepsilon_{\mathrm{Ext}})$ extractor.

Full construction of [INW94]

The PRG $\mathrm{INW}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ defined as

$$\mathrm{INW}_{h+1}(x \circ y) \stackrel{\text{def}}{=} \mathrm{INW}(x) \circ \mathrm{INW}(\mathrm{Ext}_h(x,y))$$
$$\mathrm{INW}_1(x) \stackrel{\text{def}}{=} x_0$$

where $\mathrm{Ext}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ is a $(k_{h+1}, \varepsilon_{\mathrm{Ext}})$ extractor.

Parameters.

1. The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Full construction of [INW94]

The PRG $\mathrm{INW}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ defined as

$$\mathrm{INW}_{h+1}(x \circ y) \stackrel{\text{def}}{=} \mathrm{INW}(x) \circ \mathrm{INW}(\mathrm{Ext}_h(x,y))$$

$$\mathrm{INW}_1(x) \stackrel{\text{def}}{=} x_0$$

where $\mathrm{Ext}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ is a $(k_{h+1}, \varepsilon_{\mathrm{Ext}})$ extractor.

Parameters.

1. The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.
2. Using optimal extractors, $m_h = k_{h+1} - O(\log \varepsilon_{\mathrm{Ext}}^{-1})$ and $d_h = O(\log \ell_h + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Full construction of [INW94]

The PRG $\mathrm{INW}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ defined as

$$\mathrm{INW}_{h+1}(x \circ y) \stackrel{\text{def}}{=} \mathrm{INW}(x) \circ \mathrm{INW}(\mathrm{Ext}_h(x,y))$$

$$\mathrm{INW}_1(x) \stackrel{\text{def}}{=} x_0$$

where $\mathrm{Ext}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ is a $(k_{h+1}, \varepsilon_{\mathrm{Ext}})$ extractor.

Parameters.

1. The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.
2. Using optimal extractors, $m_h = k_{h+1} - O(\log \varepsilon_{\mathrm{Ext}}^{-1})$ and $d_h = O(\log \ell_h + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Thus, the seed develops as

$$\begin{aligned}\ell_{h+1} &= \ell_h + d_h + O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1}) \\ &= \ell_h + O(\log \ell_h + \log \varepsilon_{\mathrm{Ext}}^{-1}) + O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1}) \\ &= O(h \cdot (\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})).\end{aligned}$$

Full construction of [INW94]

The PRG $\mathrm{INW}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ defined as

$$\mathrm{INW}_{h+1}(x \circ y) \stackrel{\text{def}}{=} \mathrm{INW}(x) \circ \mathrm{INW}(\mathrm{Ext}_h(x,y))$$
$$\mathrm{INW}_1(x) \stackrel{\text{def}}{=} x_0$$

where $\mathrm{Ext}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ is a $(k_{h+1}, \varepsilon_{\mathrm{Ext}})$ extractor.

Parameters.

1. The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.
2. Using optimal extractors, $m_h = k_{h+1} - O(\log \varepsilon_{\mathrm{Ext}}^{-1})$ and $d_h = O(\log \ell_h + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Thus, the seed develops as

$$\begin{aligned}\ell_{h+1} &= \ell_h + d_h + O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1}) \\ &= \ell_h + O(\log \ell_h + \log \varepsilon_{\mathrm{Ext}}^{-1}) + O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1}) \\ &= O(h \cdot (\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})).\end{aligned}$$

And since the error develops as

$$\varepsilon(h) = 2\varepsilon(h-1) + \varepsilon_{\mathrm{Ext}} = 2^h \cdot \varepsilon_{\mathrm{Ext}},$$

so $\varepsilon = \varepsilon(\log n) = n \cdot \varepsilon_{\mathrm{Ext}}$ forces seed length

$$s_{\mathrm{INW}} \stackrel{\text{def}}{=} \ell_{\log n} = O(\log n \cdot (\log n + \log w + \log \varepsilon^{-1}))$$

Full construction of [INW94]

The PRG $\mathrm{INW}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ defined as

$$\mathrm{INW}_{h+1}(x \circ y) \stackrel{\text{def}}{=} \mathrm{INW}(x) \circ \mathrm{INW}(\mathrm{Ext}_h(x,y))$$
$$\mathrm{INW}_1(x) \stackrel{\text{def}}{=} x_0$$

where $\mathrm{Ext}_{h+1} : \{0,1\}^{\ell_{h+1}} \times \{0,1\}^{d_{h+1}} \to \{0,1\}^{m_h}$ is a $(k_{h+1}, \varepsilon_{\mathrm{Ext}})$ extractor.

Parameters.

1. The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.
2. Using optimal extractors, $m_h = k_{h+1} - O(\log \varepsilon_{\mathrm{Ext}}^{-1})$ and $d_h = O(\log \ell_h + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Thus, the seed develops as

$$\begin{aligned}\ell_{h+1} &= \ell_h + d_h + O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1}) \\ &= \ell_h + O(\log \ell_h + \log \varepsilon_{\mathrm{Ext}}^{-1}) + O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1}) \\ &= O(h \cdot (\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})).\end{aligned}$$

And since the error develops as

$$\varepsilon(h) = 2\varepsilon(h-1) + \varepsilon_{\mathrm{Ext}} = 2^h \cdot \varepsilon_{\mathrm{Ext}},$$

so $\varepsilon = \varepsilon(\log n) = n \cdot \varepsilon_{\mathrm{Ext}}$ forces seed length

$$s_{\mathrm{INW}} \stackrel{\text{def}}{=} \ell_{\log n} = O(\log n \cdot (\log n + \log w + \log \varepsilon^{-1}))$$

Table of Contents

Can we do better?

The cause for seed length $\Omega(\log n \cdot \log w)$

The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Can we do better?

The cause for seed length $\Omega(\log n \cdot \log w)$

The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\text{Ext}}^{-1})$.

Question. Do we have to lose $\log w$ entropy at each recursion level to recycle X?

Can we do better?

The cause for seed length $\Omega(\log n \cdot \log w)$

The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Question. Do we have to lose $\log w$ entropy at each recursion level to recycle X?
Relaxed Question. Could we do better given that we know how much the machine learnt?

Can we do better?

The cause for seed length $\Omega(\log n \cdot \log w)$

The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Question. Do we have to lose $\log w$ entropy at each recursion level to recycle X?
Relaxed Question. Could we do better given that we know how much the machine learnt?

Definition (Estimator)

$\mathrm{A} : \{(n, w) \text{ BP}\} \times [nw]^2 \to \mathbb{R}^+$ is called an (n, w, r) **estimator** if for every two states a, b,

$$2^{-r} \cdot p_{a,b} \leq \mathrm{A}(M, a, b) \leq 2^{r} \cdot p_{a,b},$$

where, by denoting the layers distance of them as t,

$$p_{a,b} \stackrel{\text{def}}{=} \Pr[M_a(U_t) = b].$$

Can we do better?

The cause for seed length $\Omega(\log n \cdot \log w)$

The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\mathrm{Ext}}^{-1})$.

Question. Do we have to lose $\log w$ entropy at each recursion level to recycle X?
Relaxed Question. Could we do better given that we know how much the machine learnt?

Definition (Estimator)

$\mathrm{A} : \{(n, w)\ \text{BP}\} \times [nw]^2 \to \mathbb{R}^+$ is called an (n, w, r) **estimator** if for every two states a, b,

$$2^{-r} \cdot p_{a,b} \leq \mathrm{A}(M, a, b) \leq 2^r \cdot p_{a,b},$$

where, by denoting the layers distance of them as t,

$$p_{a,b} \stackrel{\text{def}}{=} \Pr[M_a(U_t) = b].$$

But thinking about estimation raises doubts...

1. There is no known explicit construction of *multiplicative* error objects,

Can we do better?

The cause for seed length $\Omega(\log n \cdot \log w)$

The entropy is set to $k_{h+1} = \ell_{h+1} - O(\log w + \log \varepsilon_{\text{Ext}}^{-1})$.

Question. Do we have to lose $\log w$ entropy at each recursion level to recycle X?
Relaxed Question. Could we do better given that we know how much the machine learnt?

Definition (Estimator)

$\mathrm{A} : \{(n, w) \text{ BP}\} \times [nw]^2 \to \mathbb{R}^+$ is called an (n, w, r) **estimator** if for every two states a, b,

$$2^{-r} \cdot p_{a,b} \leq \mathrm{A}(M, a, b) \leq 2^r \cdot p_{a,b},$$

where, by denoting the layers distance of them as t,

$$p_{a,b} \stackrel{\text{def}}{=} \Pr[M_a(U_t) = b].$$

But thinking about estimation raises doubts...

1. There is no known explicit construction of *multiplicative* error objects,
2. The estimator implies multiplicative derandomization, so PRG construction seems irrelevant.

Settling the doubts

Doubts about estimators

1. There is no known explicit construction of *multiplicative* error objects,
2. The estimator implies multiplicative derandomization, so PRG construction seems irrelevant.

We resolve the issues:

Settling the doubts

Doubts about estimators

1. There is no known explicit construction of *multiplicative* error objects,
2. The estimator implies multiplicative derandomization, so PRG construction seems irrelevant.

We resolve the issues:

1. They do exists: a simple reduction of “der. alg $\Longrightarrow$ estimators”.

Settling the doubts

Doubts about estimators

1. There is no known explicit construction of *multiplicative* error objects,
2. The estimator implies multiplicative derandomization, so PRG construction seems irrelevant.

We resolve the issues:

1. They do exists: a simple reduction of "der. alg $\Longrightarrow$ estimators".
2. 1. PRGs are much more powerful than derandomizations! (as Saks and Zhou [SZ99] showed us)

Settling the doubts

Doubts about estimators

1. There is no known explicit construction of *multiplicative* error objects,
2. The estimator implies multiplicative derandomization, so PRG construction seems irrelevant.

We resolve the issues:

1. They do exists: a simple reduction of "der. alg $\Longrightarrow$ estimators".
2.
   1. PRGs are much more powerful than derandomizations! (as Saks and Zhou [SZ99] showed us)
   2. We combine solution (1) with recent developments to conclude a new white box Weighted PRG

Derandomization implies Estimators

Let A be an $(n, w, \varepsilon_{\mathbf{A}})$ (additive) derandomization, M be an (n, w) BP with two states s, t.
We wish to compute $\widetilde{p_{s,t}}$ s.t.

$$2^{-r} \cdot p_{s,t} \leq \widetilde{p_{s,t}} \leq 2^{r} \cdot p_{s,t}.$$

Derandomization implies Estimators

Let A be an $(n, w, \varepsilon_{\mathbf{A}})$ (additive) derandomization, M be an (n, w) BP with two states s, t. We wish to compute $\widetilde{p_{s,t}}$ s.t.

$$2^{-r} \cdot p_{s,t} \leq \widetilde{p_{s,t}} \leq 2^{r} \cdot p_{s,t}.$$

Set $\widetilde{p_{s,t}} \stackrel{\text{def}}{=} \mathrm{A}(M, s, t)$, and due to A promise,

$$2^{-r} \cdot p_{s,t} \leq p_{s,t} \pm \varepsilon_{\mathbf{A}} \leq 2^{r} \cdot p_{s,t} \quad \Longrightarrow \quad r \geq \log\left(1 \pm \frac{\varepsilon_{\mathbf{A}}}{p_{s,t}}\right).$$

Clearly $\varepsilon_{\mathbf{A}} = O(p_{s,t})$ implies (multiplicative) estimation of $r = O(1)$.

Derandomization implies Estimators

Let A be an $(n, w, \varepsilon_{\mathrm{A}})$ (additive) derandomization, M be an (n, w) BP with two states s, t.
We wish to compute $\widetilde{p_{s,t}}$ s.t.

$$2^{-r} \cdot p_{s,t} \leq \widetilde{p_{s,t}} \leq 2^{r} \cdot p_{s,t}.$$

Set $\widetilde{p_{s,t}} \stackrel{\text{def}}{=} \mathrm{A}(M, s, t)$, and due to A promise,

$$2^{-r} \cdot p_{s,t} \leq p_{s,t} \pm \varepsilon_{\mathrm{A}} \leq 2^{r} \cdot p_{s,t} \quad \Longrightarrow \quad r \geq \log\left(1 \pm \frac{\varepsilon_{\mathrm{A}}}{p_{s,t}}\right).$$

Clearly $\varepsilon_{\mathrm{A}} = O(p_{s,t})$ implies (multiplicative) estimation of $r = O(1)$.
While potentially $p_{s,t} = 2^{-\Theta(n)}$, its easy to discard such low probabilities **in the analysis**.

PRG Analysis

Say we wish to construct an $(n, w, \boldsymbol{\varepsilon_G})$ PRG G.
By setting $\boldsymbol{\varepsilon_G} \leftarrow \boldsymbol{\varepsilon_G} + \gamma \cdot nw$, one may assume that **always**

$$p_{s,t} = \Pr[M_s(U) = t] > \gamma.$$

Derandomization implies Estimators

Let A be an $(n, w, \varepsilon_{\mathrm{A}})$ (additive) derandomization, M be an (n, w) BP with two states s, t.
We wish to compute $\widetilde{p_{s,t}}$ s.t.

$$2^{-r} \cdot p_{s,t} \leq \widetilde{p_{s,t}} \leq 2^{r} \cdot p_{s,t}.$$

Set $\widetilde{p_{s,t}} \stackrel{\text{def}}{=} \mathrm{A}(M, s, t)$, and due to A promise,

$$2^{-r} \cdot p_{s,t} \leq p_{s,t} \pm \varepsilon_{\mathrm{A}} \leq 2^{r} \cdot p_{s,t} \quad \Longrightarrow \quad r \geq \log\left(1 \pm \frac{\varepsilon_{\mathrm{A}}}{p_{s,t}}\right).$$

Clearly $\varepsilon_{\mathrm{A}} = O(p_{s,t})$ implies (multiplicative) estimation of $r = O(1)$.
While potentially $p_{s,t} = 2^{-\Theta(n)}$, its easy to discard such low probabilities **in the analysis**.

PRG Analysis

Say we wish to construct an (n, w, ε_G) PRG G.
By setting $\varepsilon_G \leftarrow \varepsilon_G + \gamma \cdot nw$, one may assume that **always**

$$p_{s,t} = \Pr[M_s(U) = t] > \gamma.$$

Thus,

$$r = O(\varepsilon_{\mathrm{A}}/\gamma).$$

Parameters. Set $\gamma = 1/nw$ and let A be [SZ99] with $\varepsilon_{\mathrm{A}} = 1/nw \quad \Longrightarrow \quad r = O(1)$.

A New PRG

First, we conclude an explicit white box PRG from [RR99] using [SZ99] as an estimator:

Theorem (PRG based [RR99])

There exists an (n, w, ε) *white box PRG with seed length*

$$\mathsf{s}_{\mathsf{RR}} = \widetilde{O}(\log n \cdot (\log n + \log \varepsilon^{-1}) + \log w),$$

that is computable in space

$$\widetilde{O}(\log n \cdot (\log n + \log \varepsilon^{-1}) + \sqrt{\log n} \cdot (\log w)).$$

A New PRG

First, we conclude an explicit white box PRG from [RR99] using [SZ99] as an estimator:

Theorem (PRG based [RR99])

There exists an (n, w, ε) *white box PRG with seed length*

$$\mathsf{s}_{\mathsf{RR}} = \widetilde{O}(\log n \cdot (\log n + \log \varepsilon^{-1}) + \log w),$$

that is computable in space

$$\widetilde{O}(\log n \cdot (\log n + \log \varepsilon^{-1}) + \sqrt{\log n} \cdot (\log w)).$$

Now use the black box error reduction of [CDR+21; PV21] to conclude:**

Theorem

There exists an (n, w, ε) *white box Weighted-PRG with seed length*

$$\mathsf{s} = \widetilde{O}(\log n \cdot (\log n) + \log w + \log \varepsilon^{-1}),$$

that is computable in space

$$\widetilde{O}(\log n \cdot (\log n) + \sqrt{\log n} \cdot (\log w) + \log \varepsilon^{-1}).$$

A New PRG

First, we conclude an explicit white box PRG from [RR99] using [SZ99] as an estimator:

Theorem (PRG based [RR99])

There exists an (n, w, ε) *white box PRG with seed length*

$$\mathsf{s}_{\mathsf{RR}} = \widetilde{O}(\log n \cdot (\log n + \log \varepsilon^{-1}) + \log w),$$

that is computable in space

$$\widetilde{O}(\log n \cdot (\log n + \log \varepsilon^{-1}) + \sqrt{\log n} \cdot (\log w)).$$

Now use the black box error reduction of [CDR+21; PV21] to conclude:**

Theorem

There exists an (n, w, ε) *white box Weighted-PRG with seed length*

$$\mathsf{s} = \widetilde{O}(\log n \cdot (\log n) + \log w + \log \varepsilon^{-1}),$$

that is computable in space

$$\widetilde{O}(\log n \cdot (\log n) + \sqrt{\log n} \cdot (\log w) + \log \varepsilon^{-1}).$$

Table of Contents

Preliminaries

Notations:

1. For any event E, we define $\mathbf{H}(E) = \log \frac{1}{\Pr[E]}$.
2. For every $q \in [n]$ and $s \in [w]$,

$$S^{\mathrm{ideal,q}}(s) \stackrel{\mathsf{def}}{=} \Pr[M_{s_{\mathrm{init}}}(U_q) = s].$$

3. For simplicity, assume A is an $(n, w, r = 0)$ **perfect estimator**.
4. The estimated entropy of a given state is

$$\widetilde{\mathbf{H}}(s) \stackrel{\mathsf{def}}{=} \log \frac{1}{\mathrm{A}(M, s_{\mathrm{init}}, s)},$$

and since A is perfect,

$$\widetilde{\mathbf{H}}(s) = \log \frac{1}{\mathrm{A}(M, s_{\mathrm{init}}, s)} = \log \frac{1}{S^{\mathrm{ideal,q}}(s)} = \mathbf{H}(S^{\mathrm{ideal,q}} = s).$$

Preliminaries

Notations:

1. For any event E, we define $\mathbf{H}(E) = \log \frac{1}{\Pr[E]}$.
2. For every $q \in [n]$ and $s \in [w]$,

$$S^{\text{ideal,q}}(s) \stackrel{\text{def}}{=} \Pr[M_{s_{\text{init}}}(U_q) = s].$$

3. For simplicity, assume A is an $(n, w, r = 0)$ **perfect estimator**.
4. The estimated entropy of a given state is

$$\widetilde{\mathbf{H}}(s) \stackrel{\text{def}}{=} \log \frac{1}{\mathrm{A}(M, s_{\text{init}}, s)},$$

and since A is perfect,

$$\widetilde{\mathbf{H}}(s) = \log \frac{1}{\mathrm{A}(M, s_{\text{init}}, s)} = \log \frac{1}{S^{\text{ideal,q}}(s)} = \mathbf{H}(S^{\text{ideal,q}} = s).$$

The PRG

Begin with the INW PRG, and we modify it gradually:

$$G(x, y) \stackrel{\text{def}}{=} G^{\mathrm{A}}(x, y) \stackrel{\text{def}}{=} x \circ \mathrm{Ext}(x, y).$$

Improving the entropy loss

Analysis.
Let $X \circ Y \sim U_{n+d}$. Let $S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}$ be state distribution after walking via U_n or $G^{\text{A}}(X)$.

Improving the entropy loss

Analysis.
Let $X \circ Y \sim U_{n+d}$. Let $S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}$ be state distribution after walking via U_n or $G^{\mathrm{A}}(X)$.

Entropy analysis. Sample $s_{\text{mid}} \sim S_{\text{mid}}^{\text{gen}}$ and observe,

$$\begin{aligned} k_{s_{\text{mid}}} &= \mathbf{H}_\infty(X \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}) \geq \mathbf{H}_\infty(X) - \log \frac{1}{\Pr[S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}]} \\ &= n - \mathbf{H}(S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}). \end{aligned}$$

Improving the entropy loss

Analysis.
Let $X \circ Y \sim U_{n+d}$. Let $S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}$ be state distribution after walking via U_n or $G^{\text{A}}(X)$.

Entropy analysis. Sample $s_{\text{mid}} \sim S_{\text{mid}}^{\text{gen}}$ and observe,

$$
\begin{aligned}
k_{s_{\text{mid}}} &= \mathbf{H}_\infty(X \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}) \geq \mathbf{H}_\infty(X) - \log \frac{1}{\Pr[S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}]} \\
&= n - \mathbf{H}(S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}).
\end{aligned}
$$

Now note:

1. $\text{SD}(S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}) = 0$ since X is uniform,
2. $\widetilde{\mathbf{H}}(s_{\text{mid}}) = \mathbf{H}(S_{\text{mid}}^{\text{ideal}} = s_{\text{mid}})$ since the estimator is perfect,

Improving the entropy loss

Analysis.
Let $X \circ Y \sim U_{n+d}$. Let $S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}$ be state distribution after walking via U_n or $G^{\text{A}}(X)$.

Entropy analysis. Sample $s_{\text{mid}} \sim S_{\text{mid}}^{\text{gen}}$ and observe,

$$k_{s_{\text{mid}}} = \mathbf{H}_\infty(X \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}) \geq \mathbf{H}_\infty(X) - \log \frac{1}{\Pr[S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}]}$$
$$= n - \mathbf{H}(S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}).$$

Now note:

1. $\text{SD}(S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}) = 0$ since X is uniform,
2. $\widetilde{\mathbf{H}}(s_{\text{mid}}) = \mathbf{H}(S_{\text{mid}}^{\text{ideal}} = s_{\text{mid}})$ since the estimator is perfect,

$$\implies \quad \widetilde{\mathbf{H}}(s_{\text{mid}}) = \mathbf{H}(S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}) \quad \implies \quad k_{s_{\text{mid}}} \text{ is } \textit{computable}$$

Improving the entropy loss (cont.)

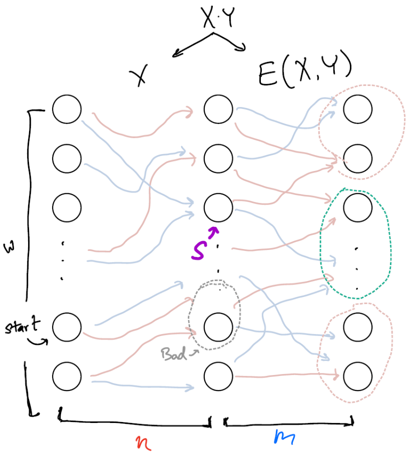


Figure: INW

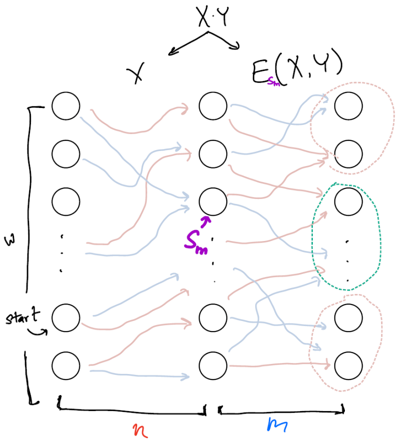


Figure: G^{A}

Improving the entropy loss (cont.)

So we summarize:

1. $k_{s_{\text{mid}}} = n - \widetilde{\mathbf{H}}(s_{\text{mid}})$,
2. $k_{\text{INW}} = n - \log w - \log \varepsilon^{-1} - 1$.

Is it always the case $k_{s_{\text{mid}}} \gg k_{\text{INW}}$?

Improving the entropy loss (cont.)

So we summarize:

1. $k_{s_{\mathrm{mid}}} = n - \widetilde{\mathbf{H}}(s_{\mathrm{mid}})$,
2. $k_{\mathrm{INW}} = n - \log w - \log \varepsilon^{-1} - 1$.

Is it always the case $k_{s_{\mathrm{mid}}} \gg k_{\mathrm{INW}}$?
Depends on $\widetilde{\mathbf{H}}(s_{\mathrm{mid}})$... Actually, it could be $k_{s_{\mathrm{mid}}} \ll k_{\mathrm{INW}}$!

Discarding states with low probability.

Using the same trick as before, we discard all states that satisfies

$$S^{\mathrm{ideal,q}}(s) \leq \varepsilon/(2 \cdot (n+m)w)$$

and so increase G error by

$$< (n+m)w \cdot \varepsilon/(2 \cdot (n+m)w) = \varepsilon/2.$$

Improving the entropy loss (cont.)

So we summarize:

1. $k_{s_{\text{mid}}} = n - \widetilde{\mathbf{H}}(s_{\text{mid}})$,
2. $k_{\text{INW}} = n - \log w - \log \varepsilon^{-1} - 1$.

Is it always the case $k_{s_{\text{mid}}} \gg k_{\text{INW}}$?
Depends on $\widetilde{\mathbf{H}}(s_{\text{mid}})$... Actually, it could be $k_{s_{\text{mid}}} \ll k_{\text{INW}}$!

Discarding states with low probability.

Using the same trick as before, we discard all states that satisfies

$$S^{\text{ideal,q}}(s) \leq \varepsilon/(2 \cdot (n+m)w)$$

and so increase G error by

$$< (n+m)w \cdot \varepsilon/(2 \cdot (n+m)w) = \varepsilon/2.$$

Now we can bound for every $s_{\text{mid}} \in \text{supp}(S^{\text{gen}}_{\text{mid}})$:

$$k_{s_{\text{mid}}} = n - \widetilde{\mathbf{H}}(s_{\text{mid}}) \geq n - \log w - \log \varepsilon^{-1} - 1 - \log n - \log m.$$

so we gained nothing...

Recycling the input states

Observation: when we condition the source on s_{mid}, the input state s_{in} is no longer conditioned.

Recycling the input states

Observation: when we condition the source on s_{mid}, the input state s_{in} is no longer conditioned.
The idea: recycle s_{in} when going right, i.e. $\mathrm{Ext}(X \circ s_{\mathrm{in}}, Y)$.

Recycling the input states

Observation: when we condition the source on s_{mid}, the input state s_{in} is no longer conditioned.
The idea: recycle s_{in} when going right, i.e. $\mathrm{Ext}(X \circ s_{\mathrm{in}}, Y)$.

The input length

Let $s \in [nw]$. Define $\ell_s \stackrel{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.
The input source X is redefined to have length $\ell_{S_{\mathrm{in}}^{\mathrm{ideal}}}$, i.e. *random variable*.

Recycling the input states

Observation: when we condition the source on s_{mid}, the input state s_{in} is no longer conditioned.
The idea: recycle s_{in} when going right, i.e. $\text{Ext}(X \circ s_{\text{in}}, Y)$.

The input length

Let $s \in [nw]$. Define $\ell_s \stackrel{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.
The input source X is redefined to have length $\ell_{S_{\text{in}}^{\text{ideal}}}$, i.e. *random variable*.

The input to the PRG is now $S_{\text{in}}^{\text{ideal}} \circ X$.

Recycling the input states

Observation: when we condition the source on s_{mid}, the input state s_{in} is no longer conditioned.
The idea: recycle s_{in} when going right, i.e. $\text{Ext}(X \circ s_{\text{in}}, Y)$.

The input length

Let $s \in [nw]$. Define $\ell_s \stackrel{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.
The input source X is redefined to have length $\ell_{S_{\text{in}}^{\text{ideal}}}$, i.e. *random variable*.

The input to the PRG is now $S_{\text{in}}^{\text{ideal}} \circ X$.
What is $\mathbf{H}_\infty(S_{\text{in}}^{\text{ideal}} \circ X)$?

Recycling the input states

Observation: when we condition the source on s_{mid}, the input state s_{in} is no longer conditioned.
The idea: recycle s_{in} when going right, i.e. $\text{Ext}(X \circ s_{\text{in}}, Y)$.

The input length

Let $s \in [nw]$. Define $\ell_s \stackrel{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.
The input source X is redefined to have length $\ell_{S_{\text{in}}^{\text{ideal}}}$, i.e. *random variable*.

The input to the PRG is now $S_{\text{in}}^{\text{ideal}} \circ X$.
What is $\mathbf{H}_\infty(S_{\text{in}}^{\text{ideal}} \circ X)$?

$\mathbf{H}_\infty(S_{\text{in}}^{\text{ideal}} \circ X) = n$

We circumvent the min entropy definition and use $\mathbf{H}$ instead. For every $s_{\text{in}} \circ x$,

$$\begin{aligned}\mathbf{H}(S_{\text{in}}^{\text{ideal}} \circ X = s_{\text{in}} \circ x) &= \mathbf{H}(S_{\text{in}}^{\text{ideal}} = s_{\text{in}}) + \ell_{s_{\text{in}}} \\ &= \mathbf{H}(S_{\text{in}}^{\text{ideal}} = s_{\text{in}}) + n - \widetilde{\mathbf{H}}(s_{\text{in}}) \\ &= n.\end{aligned}$$

□

One step of recycling

Reminder

Let $s \in [nw]$. Define $\ell_s \stackrel{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.

Let $X \circ Y \sim U_{\ell_{S_{\text{in}}^{\text{ideal}}}} \times U_d$. Let $S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}$ be state distribution after walking via U_n or $G^{\text{A}}(X)$.

One step of recycling

Reminder

Let $s \in [nw]$. Define $\ell_s \stackrel{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.

Let $X \circ Y \sim U_{\ell_{S_{\text{in}}^{\text{ideal}}}} \times U_d$. Let $S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}$ be state distribution after walking via U_n or $G^{\mathrm{A}}(X)$. Since X is uniform,

$$\mathsf{SD}(S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}) = 0.$$

Let $s_{\text{mid}} \sim S_{\text{mid}}^{\text{gen}}$.

One step of recycling

Reminder

Let $s \in [nw]$. Define $\ell_s \overset{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.

Let $X \circ Y \sim U_{\ell_{S_{\text{in}}^{\text{ideal}}}} \times U_d$. Let $S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}$ be state distribution after walking via U_n or $G^{\mathrm{A}}(X)$.
Since X is uniform,

$$\mathsf{SD}(S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}) = 0.$$

Let $s_{\text{mid}} \sim S_{\text{mid}}^{\text{gen}}$.
Analyzing the entropy. What is $\mathbf{H}_\infty(S_{\text{in}}^{\text{ideal}} \circ X \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}})$?

One step of recycling

Reminder

Let $s \in [nw]$. Define $\ell_s \overset{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.

Let $X \circ Y \sim U_{\ell_{S_{\text{in}}^{\text{ideal}}}} \times U_d$. Let $S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}$ be state distribution after walking via U_n or $G^{\text{A}}(X)$.
Since X is uniform,

$$\text{SD}(S_{\text{mid}}^{\text{ideal}}, S_{\text{mid}}^{\text{gen}}) = 0.$$

Let $s_{\text{mid}} \sim S_{\text{mid}}^{\text{gen}}$.
Analyzing the entropy. What is $\mathbf{H}_\infty(S_{\text{in}}^{\text{ideal}} \circ X \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}})$?

$\mathbf{H}_\infty(S_{\text{in}}^{\text{ideal}} \circ X \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}) = \ell_{s_{\text{mid}}}$

We again circumvent $\mathbf{H}_\infty$ with $\mathbf{H}$:

$$\begin{aligned}
&\mathbf{H}(S_{\text{in}}^{\text{ideal}} \circ X = s_{\text{in}} \circ x \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}) \\
&\quad \geq \mathbf{H}(S_{\text{in}}^{\text{ideal}} \circ X = s_{\text{in}} \circ x) - \mathbf{H}(S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}) \\
&\quad = \mathbf{H}(S_{\text{in}}^{\text{ideal}} = s_{\text{in}}) + \ell_{s_{\text{in}}} - \mathbf{H}(S_{\text{mid}}^{\text{gen}} = s_{\text{mid}}) \\
&\quad = \mathbf{H}(S_{\text{in}}^{\text{ideal}} = s_{\text{in}}) + \ell_{s_{\text{in}}} - \mathbf{H}(S_{\text{mid}}^{\text{ideal}} = s_{\text{mid}}) \\
&\quad = \ell_{s_{\text{mid}}} + (\ell_{s_{\text{in}}} - \ell_{s_{\text{mid}}}) + \mathbf{H}(S_{\text{in}}^{\text{ideal}} = s_{\text{in}}) - \mathbf{H}(S_{\text{mid}}^{\text{ideal}} = s_{\text{mid}}) \\
&\quad = \ell_{s_{\text{mid}}} + (-\widetilde{\mathbf{H}}(s_{\text{in}}) + \widetilde{\mathbf{H}}(s_{\text{mid}})) + \mathbf{H}(S_{\text{in}}^{\text{ideal}} = s_{\text{in}}) - \mathbf{H}(S_{\text{mid}}^{\text{ideal}} = s_{\text{mid}}) \\
&\quad = \ell_{s_{\text{mid}}}
\end{aligned}$$

One step of recycling (cont.)

Reminder

Let $s \in [nw]$. Define $\ell_s \stackrel{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.

We want to recycle $(S_{\text{in}}^{\text{ideal}} \circ X \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}})$. We choose $(k_{s_{\text{mid}}} = \ell_{s_{\text{mid}}}, \varepsilon)$ extractor

$$\text{Ext}_{s_{\text{mid}}} : \{0,1\}^{n'_{S_{\text{in}}^{\text{ideal}}}} \times \{0,1\}^d \to \{0,1\}^m$$

where

$$n'_{S_{\text{in}}^{\text{ideal}}} = \log(nw) + \ell_{S_{\text{in}}^{\text{ideal}}}$$

Using optimal extractors,

$$m = k_{s_{\text{mid}}} - O(\log(1/\varepsilon)) = \ell_{s_{\text{mid}}} - O(\log(1/\varepsilon))$$
$$d = O(\log(n'_{S_{\text{in}}^{\text{ideal}}}/\varepsilon)) = O(\log n + \log(1/\varepsilon) + \log\log w)$$

One step of recycling (cont.)

Reminder

Let $s \in [nw]$. Define $\ell_s \overset{\text{def}}{=} n - \widetilde{\mathbf{H}}(s)$.

We want to recycle $(S_{\text{in}}^{\text{ideal}} \circ X \mid S_{\text{mid}}^{\text{gen}} = s_{\text{mid}})$. We choose $(k_{s_{\text{mid}}} = \ell_{s_{\text{mid}}}, \varepsilon)$ extractor

$$\text{Ext}_{s_{\text{mid}}} : \{0,1\}^{n'_{S_{\text{in}}^{\text{ideal}}}} \times \{0,1\}^d \to \{0,1\}^m$$

where

$$n'_{S_{\text{in}}^{\text{ideal}}} = \log(nw) + \ell_{S_{\text{in}}^{\text{ideal}}}$$

Using optimal extractors,

$$m = k_{s_{\text{mid}}} - O(\log(1/\varepsilon)) = \ell_{s_{\text{mid}}} - O(\log(1/\varepsilon))$$
$$d = O(\log(n'_{S_{\text{in}}^{\text{ideal}}}/\varepsilon)) = O(\log n + \log(1/\varepsilon) + \log\log w)$$

Thus, since the analysis holds for every s_{mid},

$$S_{\text{mid}}^{\text{gen}} \circ \text{Ext}(S_{\text{in}}^{\text{ideal}} \circ X \mid S_{\text{mid}}^{\text{gen}}, Y) \approx_\varepsilon S_{\text{mid}}^{\text{gen}} \circ U_{m_{S_{\text{mid}}^{\text{gen}}}}$$
$$= S_{\text{mid}}^{\text{ideal}} \circ U_{m_{S_{\text{mid}}^{\text{ideal}}}}$$
$$= S_{\text{mid}}^{\text{ideal}} \circ U_{\ell_{S_{\text{mid}}^{\text{ideal}}} - O(\log 1/\varepsilon)}$$

One step of recycling (cont.)

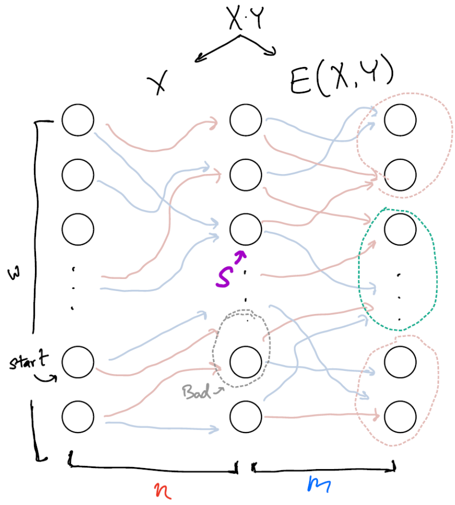


Figure: INW

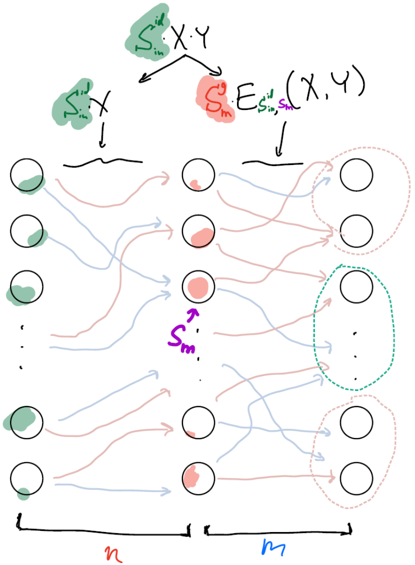


Figure: G^{A}

A full construction

In similar fashion to INW, we can devise (n, w, ε) white box PRG with seed length

$$\mathsf{s}_0 = O(\log n \cdot (\log n + \log \varepsilon^{-1} + \log \log w) + \log w)$$

but its space complexity is

$$O(\log n \cdot (\mathsf{s}_0 + \log w)) + \mathrm{space}_{\mathrm{A}}(n, w, r = 0)$$

A full construction

In similar fashion to INW, we can devise (n, w, ε) white box PRG with seed length

$$\mathsf{s}_0 = O(\log n \cdot (\log n + \log \varepsilon^{-1} + \log \log w) + \log w)$$

but its space complexity is

$$O(\log n \cdot (\mathsf{s}_0 + \log w)) + \mathrm{space}_{\mathrm{A}}(n, w, r = 0)$$

To solve the problems...

1. Save only 2 states instead of up to $\log n$ (which multiplied by $\times \log w$)
2. Use global buffers to maintain linear space
3. Use condensers to collect the extractors unavoidable loss

Table of Contents

Table of Contents

Stochastic Matrix Powering and Derandomization

Let M be (n, w) BP.

1. By increasing the width $w \mapsto \mathrm{poly}(n, w)$, wlog all layers are identical.
2. Abuse notation to denote $\mathbf{M}$ the stochastic *transition matrix* of every layer.
3. Define $\mathbf{M} = \frac{1}{2}(\mathbf{M}^{(0)} + \mathbf{M}^{(1)})$ where $\mathbf{M}^{(0)}, \mathbf{M}^{(1)} \in \{0,1\}^{w \times w}$.
4. The derandomization task is equivalent to approximation of

$$\mathbf{M}^n = \mathop{\mathbb{E}}_{\sigma \sim \{0,1\}^n} \mathbf{M}^{(\sigma)}$$

   where $\mathbf{M}^{(\sigma)} = \mathbf{M}^{(\sigma_1)} \cdots \mathbf{M}^{(\sigma_n)}$.
5. We use $\|\cdot\|$ as the infinity norm, i.e. $\|\mathbf{M}\| \stackrel{\text{def}}{=} \max_{j \in [w]} \sum_{i \in [w]} |\mathbf{M}_{i,j}|$.

Stochastic Matrix Powering and Derandomization

Let M be (n, w) BP.

1. By increasing the width $w \mapsto \mathrm{poly}(n, w)$, wlog all layers are identical.
2. Abuse notation to denote $\mathbf{M}$ the stochastic *transition matrix* of every layer.
3. Define $\mathbf{M} = \frac{1}{2}(\mathbf{M}^{(0)} + \mathbf{M}^{(1)})$ where $\mathbf{M}^{(0)}, \mathbf{M}^{(1)} \in \{0,1\}^{w \times w}$.
4. The derandomization task is equivalent to approximation of

$$\mathbf{M}^n = \mathop{\mathbb{E}}_{\sigma \sim \{0,1\}^n} \mathbf{M}^{(\sigma)}$$

   where $\mathbf{M}^{(\sigma)} = \mathbf{M}^{(\sigma_1)} \cdots \mathbf{M}^{(\sigma_n)}$.

5. We use $\|\cdot\|$ as the infinity norm, i.e. $\|\mathbf{M}\| \stackrel{\text{def}}{=} \max_{j \in [w]} \sum_{i \in [w]} |\mathbf{M}_{i,j}|$.

Redefining PRG and WPRG

An (n, w, ε) PRG $G : \{0,1\}^s \to \{0,1\}^n$ satisfies

$$\left\| \mathop{\mathbb{E}}_{\sigma \sim \{0,1\}^n} [\mathbf{M}^{(\sigma)}] - \mathop{\mathbb{E}}_{x \in \{0,1\}^s} [\mathbf{M}^{(G(x))}] \right\| \leq \varepsilon,$$

An (n, w, ε) WPRG $G = (I, \mu) : \{0,1\}^s \to \mathbb{R} \times \{0,1\}^n$ satisfies

$$\left\| \mathop{\mathbb{E}}_{\sigma \sim \{0,1\}^n} [\mathbf{M}^{(\sigma)}] - \mathop{\mathbb{E}}_{x \in \{0,1\}^s} [\mu(x) \cdot \mathbf{M}^{(I(x))}] \right\| \leq \varepsilon.$$

Encoding powers in Laplacians

Goal: given $\mathbf{M} \in \mathbb{R}^{w \times w}$, $\varepsilon > 0$, output $\widetilde{\mathbf{M}^n}$ s.t. $\left\| \widetilde{\mathbf{M}^n} - \mathbf{M}^n \right\| \leq \varepsilon$.

Encoding powers in Laplacians

Goal: given $\mathbf{M} \in \mathbb{R}^{w \times w}$, $\varepsilon > 0$, output $\widetilde{\mathbf{M}^n}$ s.t. $\left\| \widetilde{\mathbf{M}^n} - \mathbf{M}^n \right\| \leq \varepsilon$.

Inverse of Laplacians

If $\mathbf{I} - \mathbf{A}$ is invertible, then

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \ldots + \mathbf{A}^n + \ldots$$

Encoding powers in Laplacians

Goal: given $\mathbf{M} \in \mathbb{R}^{w \times w}$, $\varepsilon > 0$, output $\widetilde{\mathbf{M}^n}$ s.t. $\left\| \widetilde{\mathbf{M}^n} - \mathbf{M}^n \right\| \leq \varepsilon$.

Inverse of Laplacians

If $\mathbf{I} - \mathbf{A}$ is invertible, then

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \ldots + \mathbf{A}^n + \ldots$$

Traced back to [Coo85], there is a simple reduction of "Laplacian inverse $\Longrightarrow$ Matrix powering":

$$\mathbf{P}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \ \mathbf{P}_4 \otimes \mathbf{M} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{M} & 0 & 0 & 0 \\ 0 & \mathbf{M} & 0 & 0 \\ 0 & 0 & \mathbf{M} & 0 \end{pmatrix}, \ (\mathbf{P}_4 \otimes \mathbf{M})^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \mathbf{M}^2 & 0 & 0 & 0 \\ 0 & \mathbf{M}^2 & 0 & 0 \end{pmatrix}$$

Encoding powers in Laplacians

Goal: given $\mathbf{M} \in \mathbb{R}^{w \times w}$, $\varepsilon > 0$, output $\widetilde{\mathbf{M}^n}$ s.t. $\left\| \widetilde{\mathbf{M}^n} - \mathbf{M}^n \right\| \le \varepsilon$.

Inverse of Laplacians

If $\mathbf{I} - \mathbf{A}$ is invertible, then

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \ldots + \mathbf{A}^n + \ldots$$

Traced back to [Coo85], there is a simple reduction of "Laplacian inverse $\Longrightarrow$ Matrix powering":

$$\mathbf{P}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \; \mathbf{P}_4 \otimes \mathbf{M} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{M} & 0 & 0 & 0 \\ 0 & \mathbf{M} & 0 & 0 \\ 0 & 0 & \mathbf{M} & 0 \end{pmatrix}, \; (\mathbf{P}_4 \otimes \mathbf{M})^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \mathbf{M}^2 & 0 & 0 & 0 \\ 0 & \mathbf{M}^2 & 0 & 0 \end{pmatrix}$$

so since $(\mathbf{A} \otimes \mathbf{B})^k = \mathbf{A}^k \otimes \mathbf{B}^k$ and $\mathbf{P}_{n+1}^{n+1} = 0$,

$$\begin{aligned} (\mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M})^{-1} &= \mathbf{I} + (\mathbf{P}_{n+1} \otimes \mathbf{M}) + (\mathbf{P}_{n+1} \otimes \mathbf{M})^2 + \ldots (\mathbf{P}_{n+1} \otimes \mathbf{M})^n + \ldots \\ &= \mathbf{I} + (\mathbf{P}_{n+1} \otimes \mathbf{M}) + \ldots + (\mathbf{P}_{n+1} \otimes \mathbf{M})^n \\ &= \mathbf{I} + (\mathbf{P}_{n+1} \otimes \mathbf{M}) + \ldots + (\mathbf{P}_{n+1}^n \otimes \mathbf{M}^n) \end{aligned}$$

Encoding powers in Laplacians

Goal: given $\mathbf{M} \in \mathbb{R}^{w \times w}$, $\varepsilon > 0$, output $\widetilde{\mathbf{M}^n}$ s.t. $\left\| \widetilde{\mathbf{M}^n} - \mathbf{M}^n \right\| \leq \varepsilon$.

Inverse of Laplacians

If $\mathbf{I} - \mathbf{A}$ is invertible, then

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \ldots + \mathbf{A}^n + \ldots$$

Traced back to [Coo85], there is a simple reduction of "Laplacian inverse $\Longrightarrow$ Matrix powering":

$$\mathbf{P}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \; \mathbf{P}_4 \otimes \mathbf{M} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{M} & 0 & 0 & 0 \\ 0 & \mathbf{M} & 0 & 0 \\ 0 & 0 & \mathbf{M} & 0 \end{pmatrix}, \; (\mathbf{P}_4 \otimes \mathbf{M})^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \mathbf{M}^2 & 0 & 0 & 0 \\ 0 & \mathbf{M}^2 & 0 & 0 \end{pmatrix}$$

so since $(\mathbf{A} \otimes \mathbf{B})^k = \mathbf{A}^k \otimes \mathbf{B}^k$ and $\mathbf{P}_{n+1}^{n+1} = 0$,

$$\begin{aligned} (\mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M})^{-1} &= \mathbf{I} + (\mathbf{P}_{n+1} \otimes \mathbf{M}) + (\mathbf{P}_{n+1} \otimes \mathbf{M})^2 + \ldots (\mathbf{P}_{n+1} \otimes \mathbf{M})^n + \ldots \\ &= \mathbf{I} + (\mathbf{P}_{n+1} \otimes \mathbf{M}) + \ldots + (\mathbf{P}_{n+1} \otimes \mathbf{M})^n \\ &= \mathbf{I} + (\mathbf{P}_{n+1} \otimes \mathbf{M}) + \ldots + (\mathbf{P}_{n+1}^n \otimes \mathbf{M}^n) \end{aligned}$$

Encoding powers in Laplacians (cont.)

$$(\mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M})^{-1} = \mathbf{I} + (\mathbf{P}_{n+1} \otimes \mathbf{M}) + \ldots + (\mathbf{P}_{n+1}^{n} \otimes \mathbf{M}^{n})$$

As an example,

$$(\mathbf{I} - \mathbf{P}_4 \otimes \mathbf{M})^{-1} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{M} & \mathbf{I} & 0 & 0 \\ \mathbf{M}^2 & \mathbf{M} & \mathbf{I} & 0 \\ \mathbf{M}^3 & \mathbf{M}^2 & \mathbf{M} & \mathbf{I} \end{pmatrix}$$

Encoding powers in Laplacians (cont.)

$$(\mathbf{I}-\mathbf{P}_{n+1}\otimes\mathbf{M})^{-1}=\mathbf{I}+(\mathbf{P}_{n+1}\otimes\mathbf{M})+\ldots+(\mathbf{P}_{n+1}^{n}\otimes\mathbf{M}^{n})$$

As an example,

$$(\mathbf{I}-\mathbf{P}_{4}\otimes\mathbf{M})^{-1}=\begin{pmatrix}\mathbf{I} & 0 & 0 & 0\\ \mathbf{M} & \mathbf{I} & 0 & 0\\ \mathbf{M}^2 & \mathbf{M} & \mathbf{I} & 0\\ \mathbf{M}^3 & \mathbf{M}^2 & \mathbf{M} & \mathbf{I}\end{pmatrix}$$

And using some notations:

1. Denote $\mathbf{L}\stackrel{\text{def}}{=}\mathbf{I}-\mathbf{P}_{n+1}\otimes\mathbf{M}$,
2. So we are interested in approximating $\mathbf{L}^{-1}$ (that contains $\mathbf{M}^n$).

Table of Contents

Lemma (Precondition Richardson)

Let $\mathbf{L} \in \mathbb{R}^{w \times w}$ *invertible matrix. Let* $\widetilde{\mathbf{L}^{-1}}$ *s.t.*

$$\left\|\widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1}\right\| \leq \boldsymbol{\varepsilon_0}.$$

Define $\mathbf{R}_k \stackrel{\text{def}}{=} \sum_{i=0}^{k}(\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \cdot \widetilde{\mathbf{L}^{-1}}$. *Then,*

$$\left\|\mathbf{R}_k - \mathbf{L}^{-1}\right\| \leq \left\|\mathbf{L}^{-1}\right\| \cdot \|\mathbf{L}\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1}.$$

Richardson Iterations

Lemma (Precondition Richardson)

Let $\mathbf{L} \in \mathbb{R}^{w \times w}$ *invertible matrix. Let* $\widetilde{\mathbf{L}^{-1}}$ *s.t.*

$$\left\| \widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1} \right\| \leq \varepsilon_{\mathbf{0}}.$$

Define $\mathbf{R}_k \stackrel{\text{def}}{=} \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \cdot \widetilde{\mathbf{L}^{-1}}$. *Then,*

$$\left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq \left\| \mathbf{L}^{-1} \right\| \cdot \left\| \mathbf{L} \right\|^{k+1} \cdot \varepsilon_{\mathbf{0}}{}^{k+1}.$$

Reminder: Inverse of Laplacians

If $\mathbf{I} - \mathbf{A}$ is invertible, then

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \ldots + \mathbf{A}^n + \ldots$$

The intuition is pretty simple, as for $k = \infty$,

$$\mathbf{R}_\infty = \left(\sum_{i=0}^{\infty} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \right) \cdot \widetilde{\mathbf{L}^{-1}} = \left(\mathbf{I} - \left(\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L} \right) \right)^{-1} \cdot \widetilde{\mathbf{L}^{-1}} = \left(\widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L} \right)^{-1} \cdot \widetilde{\mathbf{L}^{-1}} = \mathbf{L}^{-1}$$

Richardson Iterations

Lemma (Precondition Richardson)

Let $\mathbf{L} \in \mathbb{R}^{w \times w}$ *invertible matrix. Let* $\widetilde{\mathbf{L}^{-1}}$ *s.t.*

$$\left\| \widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1} \right\| \leq \boldsymbol{\varepsilon_0}.$$

Define $\mathbf{R}_k \stackrel{\text{def}}{=} \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \cdot \widetilde{\mathbf{L}^{-1}}$. *Then,*

$$\left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq \left\| \mathbf{L}^{-1} \right\| \cdot \|\mathbf{L}\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1}.$$

Reminder: Inverse of Laplacians

If $\mathbf{I} - \mathbf{A}$ is invertible, then

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \ldots + \mathbf{A}^n + \ldots$$

The intuition is pretty simple, as for $k = \infty$,

$$\mathbf{R}_\infty = \left(\sum_{i=0}^{\infty} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \right) \cdot \widetilde{\mathbf{L}^{-1}} = \left(\mathbf{I} - \left(\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L} \right) \right)^{-1} \cdot \widetilde{\mathbf{L}^{-1}} = \left(\widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L} \right)^{-1} \cdot \widetilde{\mathbf{L}^{-1}} = \mathbf{L}^{-1}$$

Richardson Iterations (cont.)

And for arbitrary k, use geometric sum:

$$\mathbf{R}_k = \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \cdot \widetilde{\mathbf{L}^{-1}} = \frac{\mathbf{I} - (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^{k+1}}{\mathbf{I} - (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})} \cdot \widetilde{\mathbf{L}^{-1}} = (\mathbf{I} - (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^{k+1}) \cdot \mathbf{L}^{-1}$$

Richardson Iterations (cont.)

And for arbitrary k, use geometric sum:

$$\mathbf{R}_k = \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \cdot \widetilde{\mathbf{L}^{-1}} = \frac{\mathbf{I} - (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^{k+1}}{\mathbf{I} - (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})} \cdot \widetilde{\mathbf{L}^{-1}} = (\mathbf{I} - (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^{k+1}) \cdot \mathbf{L}^{-1}$$

Thus,

$$\begin{aligned} \left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| &= \left\| (\mathbf{I} - (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^{k+1}) \cdot \mathbf{L}^{-1} - \mathbf{L}^{-1} \right\| \\ &\leq \left\| (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^{k+1} \right\| \cdot \left\| \mathbf{L}^{-1} \right\| \\ &\leq \left\| \mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L} \right\|^{k+1} \cdot \left\| \mathbf{L}^{-1} \right\| \\ &\leq \left\| (\mathbf{L}^{-1} - \widetilde{\mathbf{L}^{-1}}) \cdot \mathbf{L} \right\|^{k+1} \cdot \left\| \mathbf{L}^{-1} \right\| \\ &\leq (\boldsymbol{\varepsilon_0} \cdot \|\mathbf{L}\|)^{k+1} \cdot \left\| \mathbf{L}^{-1} \right\| \end{aligned}$$

Table of Contents

Error reduction recipe

Richardson in one line

$$\left\|\widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1}\right\| \leq \boldsymbol{\varepsilon_0} \quad \Longrightarrow \quad \left\|\mathbf{R}_k - \mathbf{L}^{-1}\right\| \leq \left\|\mathbf{L}^{-1}\right\| \cdot \|\mathbf{L}\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1} \stackrel{\text{def}}{=} \varepsilon$$

Error reduction recipe

Richardson in one line

$$\left\| \widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1} \right\| \leq \boldsymbol{\varepsilon_0} \quad \Longrightarrow \quad \left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq \left\| \mathbf{L}^{-1} \right\| \cdot \left\| \mathbf{L} \right\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1} \stackrel{\text{def}}{=} \varepsilon$$

Recall that given BP $\mathbf{M} \in \mathbb{R}^{w \times w}$, we wish to approximate $\mathbf{M}^n$. So...

Error reduction recipe

Richardson in one line

$$\left\|\widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1}\right\| \leq \varepsilon_0 \quad \Longrightarrow \quad \left\|\mathbf{R}_k - \mathbf{L}^{-1}\right\| \leq \left\|\mathbf{L}^{-1}\right\| \cdot \|\mathbf{L}\|^{k+1} \cdot \varepsilon_0{}^{k+1} \stackrel{\text{def}}{=} \varepsilon$$

Recall that given BP $\mathbf{M} \in \mathbb{R}^{w \times w}$, we wish to approximate $\mathbf{M}^n$. So...

1. Construct a modest approx. $\left\|\widetilde{\mathbf{M}^i} - \mathbf{M}^i\right\| \leq \varepsilon_0 / n$ for $i \in [n]$,

Error reduction recipe

Richardson in one line

$$\left\|\widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1}\right\| \leq \varepsilon_0 \quad \Longrightarrow \quad \left\|\mathbf{R}_k - \mathbf{L}^{-1}\right\| \leq \left\|\mathbf{L}^{-1}\right\| \cdot \|\mathbf{L}\|^{k+1} \cdot \varepsilon_0{}^{k+1} \stackrel{\text{def}}{=} \varepsilon$$

Recall that given BP $\mathbf{M} \in \mathbb{R}^{w \times w}$, we wish to approximate $\mathbf{M}^n$. So...

1. Construct a modest approx. $\left\|\widetilde{\mathbf{M}^i} - \mathbf{M}^i\right\| \leq \varepsilon_0/n$ for $i \in [n]$,
2. Denote $\mathbf{L} = \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M}$, and encode

$$\widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 \\ \vdots & \vdots & \widetilde{\mathbf{M}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^n} & \widetilde{\mathbf{M}^{n-1}} & \cdots & \mathbf{M} & \mathbf{I} \end{pmatrix}$$

Error reduction recipe

Richardson in one line

$$\left\|\widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1}\right\| \le \boldsymbol{\varepsilon_0} \implies \left\|\mathbf{R}_k - \mathbf{L}^{-1}\right\| \le \left\|\mathbf{L}^{-1}\right\| \cdot \|\mathbf{L}\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1} \stackrel{\text{def}}{=} \boldsymbol{\varepsilon}$$

Recall that given BP $\mathbf{M} \in \mathbb{R}^{w \times w}$, we wish to approximate $\mathbf{M}^n$. So...

1. Construct a modest approx. $\left\|\widetilde{\mathbf{M}^i} - \mathbf{M}^i\right\| \le \boldsymbol{\varepsilon_0}/n$ for $i \in [n]$,
2. Denote $\mathbf{L} = \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M}$, and encode

$$\widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 \\ \vdots & \vdots & \widetilde{\mathbf{M}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^n} & \widetilde{\mathbf{M}^{n-1}} & \cdots & \mathbf{M} & \mathbf{I} \end{pmatrix}$$

3. Compute $\mathbf{R}_k = \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \cdot \widetilde{\mathbf{L}^{-1}}$ for $k = \frac{\log \boldsymbol{\varepsilon}^{-1}}{\log(n/\boldsymbol{\varepsilon_0})}$

Error reduction recipe

Richardson in one line

$$\left\|\widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1}\right\| \leq \boldsymbol{\varepsilon_0} \quad \Longrightarrow \quad \left\|\mathbf{R}_k - \mathbf{L}^{-1}\right\| \leq \left\|\mathbf{L}^{-1}\right\| \cdot \|\mathbf{L}\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1} \stackrel{\text{def}}{=} \boldsymbol{\varepsilon}$$

Recall that given BP $\mathbf{M} \in \mathbb{R}^{w \times w}$, we wish to approximate $\mathbf{M}^n$. So...

1. Construct a modest approx. $\left\|\widetilde{\mathbf{M}^i} - \mathbf{M}^i\right\| \leq \boldsymbol{\varepsilon_0}/n$ for $i \in [n]$,
2. Denote $\mathbf{L} = \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M}$, and encode

$$\widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 \\ \vdots & \vdots & \widetilde{\mathbf{M}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^n} & \widetilde{\mathbf{M}^{n-1}} & \cdots & \mathbf{M} & \mathbf{I} \end{pmatrix}$$

3. Compute $\mathbf{R}_k = \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \cdot \widetilde{\mathbf{L}^{-1}}$ for $k = \frac{\log \boldsymbol{\varepsilon}^{-1}}{\log(n/\boldsymbol{\varepsilon_0})}$
4. Output the bottom left block so $\|(\mathbf{R}_k)[n+1, 1] - \mathbf{M}^n\| \leq \boldsymbol{\varepsilon}$.

Error reduction recipe

Richardson in one line

$$\left\|\widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1}\right\| \leq \varepsilon_0 \quad \Longrightarrow \quad \left\|\mathbf{R}_k - \mathbf{L}^{-1}\right\| \leq \left\|\mathbf{L}^{-1}\right\| \cdot \|\mathbf{L}\|^{k+1} \cdot {\varepsilon_0}^{k+1} \overset{\text{def}}{=} \varepsilon$$

Recall that given BP $\mathbf{M} \in \mathbb{R}^{w \times w}$, we wish to approximate $\mathbf{M}^n$. So...

1. Construct a modest approx. $\left\|\widetilde{\mathbf{M}^i} - \mathbf{M}^i\right\| \leq \varepsilon_0/n$ for $i \in [n]$,
2. Denote $\mathbf{L} = \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M}$, and encode

$$\widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 \\ \vdots & \vdots & \widetilde{\mathbf{M}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^n} & \widetilde{\mathbf{M}^{n-1}} & \cdots & \mathbf{M} & \mathbf{I} \end{pmatrix}$$

3. Compute $\mathbf{R}_k = \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^i \cdot \widetilde{\mathbf{L}^{-1}}$ for $k = \frac{\log \varepsilon^{-1}}{\log(n/\varepsilon_0)}$
4. Output the bottom left block so $\|(\mathbf{R}_k)[n+1, 1] - \mathbf{M}^n\| \leq \varepsilon$.

Examples

Consider $k = 1$ and $n = 3$. Then,

$$\mathbf{L} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ -\mathbf{M} & \mathbf{I} & 0 & 0 \\ 0 & -\mathbf{M} & \mathbf{I} & 0 \\ 0 & 0 & -\mathbf{M} & \mathbf{I} \end{pmatrix}, \widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^3} & \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} \end{pmatrix}$$

and

$$\mathbf{R}_{k=1} = \widetilde{\mathbf{L}^{-1}} + (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^1 \cdot \widetilde{\mathbf{L}^{-1}}.$$

Examples

Consider $k = 1$ and $n = 3$. Then,

$$\mathbf{L} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ -\mathbf{M} & \mathbf{I} & 0 & 0 \\ 0 & -\mathbf{M} & \mathbf{I} & 0 \\ 0 & 0 & -\mathbf{M} & \mathbf{I} \end{pmatrix}, \quad \widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^3} & \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} \end{pmatrix}$$

and

$$\mathbf{R}_{k=1} = \widetilde{\mathbf{L}^{-1}} + (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^1 \cdot \widetilde{\mathbf{L}^{-1}}.$$

We examine $\mathbf{R}_{k=1}$:

$$(\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{M} - \widetilde{\mathbf{M}} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}}\mathbf{M} - \widetilde{\mathbf{M}^2} & \mathbf{M} - \widetilde{\mathbf{M}} & 0 & 0 \\ \widetilde{\mathbf{M}^2}\mathbf{M} - \widetilde{\mathbf{M}^3} & \widetilde{\mathbf{M}}\mathbf{M} - \widetilde{\mathbf{M}^2} & \mathbf{M} - \widetilde{\mathbf{M}} & 0 \end{pmatrix}$$

$$\mathbf{R}_{k=1} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{M} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{M}}\mathbf{M} + \mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}^2} & \mathbf{M} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^2}\mathbf{M} - \widetilde{\mathbf{M}^2}\widetilde{\mathbf{M}} + \widetilde{\mathbf{M}}\mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}}\widetilde{\mathbf{M}^2} + \mathbf{M}\widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}}\mathbf{M} + \mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}^2} & \mathbf{M} & \mathbf{I} \end{pmatrix}$$

Examples (cont.)

Now take $k = 2$ and $n = 3$. Then,

$$\mathbf{L} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ -\mathbf{M} & \mathbf{I} & 0 & 0 \\ 0 & -\mathbf{M} & \mathbf{I} & 0 \\ 0 & 0 & -\mathbf{M} & \mathbf{I} \end{pmatrix}, \ \widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^3} & \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} \end{pmatrix}$$

and

$$\mathbf{R}_{k=2} = \widetilde{\mathbf{L}^{-1}} + (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^1 \cdot \widetilde{\mathbf{L}^{-1}} + (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^2 \cdot \widetilde{\mathbf{L}^{-1}}$$

Examples (cont.)

Now take $k = 2$ and $n = 3$. Then,

$$\mathbf{L} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ -\mathbf{M} & \mathbf{I} & 0 & 0 \\ 0 & -\mathbf{M} & \mathbf{I} & 0 \\ 0 & 0 & -\mathbf{M} & \mathbf{I} \end{pmatrix}, \quad \widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{M}} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^3} & \widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}} & \mathbf{I} \end{pmatrix}$$

and

$$\mathbf{R}_{k=2} = \widetilde{\mathbf{L}^{-1}} + (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^1 \cdot \widetilde{\mathbf{L}^{-1}} + (\mathbf{I} - \widetilde{\mathbf{L}^{-1}} \cdot \mathbf{L})^2 \cdot \widetilde{\mathbf{L}^{-1}}$$

So $\mathbf{R}_{k=2}$ looks like:

$$\mathbf{R}_{k=2} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{M} & \nwarrow & 0 & 0 \\ \mathbf{M}^2 - \mathbf{M}\widetilde{\mathbf{M}} + \widetilde{\mathbf{M}}^2 + \widetilde{\mathbf{M}}\mathbf{M} - \widetilde{\mathbf{M}^2} & \nwarrow & \nwarrow & 0 \\ \star & \nwarrow & \nwarrow & \nwarrow \end{pmatrix}$$

$$\star = \widetilde{\mathbf{M}}\mathbf{M}^2 - 2\widetilde{\mathbf{M}^2}\mathbf{M} + 3\widetilde{\mathbf{M}}^3 + \mathbf{M}\widetilde{\mathbf{M}}\mathbf{M} - \mathbf{M}\widetilde{\mathbf{M}}^2$$
$$+\mathbf{M}^2\widetilde{\mathbf{M}} + \widetilde{\mathbf{M}^2}\mathbf{M} - \widetilde{\mathbf{M}^2}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}}\widetilde{\mathbf{M}^2} + \mathbf{M}\widetilde{\mathbf{M}^2}$$

Deducing the improved accuracy of $\mathbf{M}^n$

Reminder

$\mathbf{M}$ is some BP, $\mathbf{L} \stackrel{\text{def}}{=} \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M}$, and

$$\mathbf{L}^{-1} = \begin{pmatrix} \mathbf{M}^0 & 0 & 0 \\ \vdots & \ddots & 0 \\ \mathbf{M}^n & \cdots & \mathbf{M}^0 \end{pmatrix}, \widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \widetilde{\mathbf{M}^0} & 0 & 0 \\ \vdots & \ddots & 0 \\ \widetilde{\mathbf{M}^n} & \cdots & \widetilde{\mathbf{M}^0} \end{pmatrix}$$

where $\left\| \widetilde{\mathbf{M}^i} - \mathbf{M}^i \right\| \leq \boldsymbol{\varepsilon_0}/n$.

Richardson

$$\left\| \widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1} \right\| \leq \boldsymbol{\varepsilon_0} \quad \Longrightarrow \quad \left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq \left\| \mathbf{L}^{-1} \right\| \cdot \left\| \mathbf{L} \right\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1} \stackrel{\text{def}}{=} \boldsymbol{\varepsilon}.$$

Deducing the improved accuracy of $\mathbf{M}^n$

Reminder

$\mathbf{M}$ is some BP, $\mathbf{L} \stackrel{\text{def}}{=} \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M}$, and

$$\mathbf{L}^{-1} = \begin{pmatrix} \mathbf{M}^0 & 0 & 0 \\ \vdots & \ddots & 0 \\ \mathbf{M}^n & \cdots & \mathbf{M}^0 \end{pmatrix}, \ \widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \widetilde{\mathbf{M}^0} & 0 & 0 \\ \vdots & \ddots & 0 \\ \widetilde{\mathbf{M}^n} & \cdots & \widetilde{\mathbf{M}^0} \end{pmatrix}$$

where $\left\| \widetilde{\mathbf{M}^i} - \mathbf{M}^i \right\| \leq \boldsymbol{\varepsilon_0}/n$.

Richardson

$$\left\| \widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1} \right\| \leq \boldsymbol{\varepsilon_0} \quad \Longrightarrow \quad \left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq \left\| \mathbf{L}^{-1} \right\| \cdot \left\| \mathbf{L} \right\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1} \stackrel{\text{def}}{=} \boldsymbol{\varepsilon}.$$

It's not hard to convince that

$$\|\mathbf{L}\| \leq 2 \quad , \quad \left\| \mathbf{L}^{-1} \right\| \leq n + 1 \leq 2n$$

Deducing the improved accuracy of $\mathbf{M}^n$

Reminder

$\mathbf{M}$ is some BP, $\mathbf{L} \overset{\text{def}}{=} \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M}$, and

$$\mathbf{L}^{-1} = \begin{pmatrix} \mathbf{M}^0 & 0 & 0 \\ \vdots & \ddots & 0 \\ \mathbf{M}^n & \cdots & \mathbf{M}^0 \end{pmatrix}, \quad \widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \widetilde{\mathbf{M}^0} & 0 & 0 \\ \vdots & \ddots & 0 \\ \widetilde{\mathbf{M}^n} & \cdots & \widetilde{\mathbf{M}^0} \end{pmatrix}$$

where $\left\| \widetilde{\mathbf{M}^i} - \mathbf{M}^i \right\| \leq \boldsymbol{\varepsilon_0}/n$.

Richardson

$$\left\| \widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1} \right\| \leq \boldsymbol{\varepsilon_0} \quad \Longrightarrow \quad \left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq \left\| \mathbf{L}^{-1} \right\| \cdot \|\mathbf{L}\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1} \overset{\text{def}}{=} \boldsymbol{\varepsilon}.$$

It's not hard to convince that

$$\|\mathbf{L}\| \leq 2 \quad , \quad \left\| \mathbf{L}^{-1} \right\| \leq n + 1 \leq 2n$$

thus,

$$\left\| (\mathbf{R}_k - \mathbf{L}^{-1})_{n+1,1} \right\| \leq \left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq 2n \cdot (2 \cdot \boldsymbol{\varepsilon_0})^{k+1} \overset{\text{def}}{=} \boldsymbol{\varepsilon},$$

so we need to take $k \overset{\text{def}}{=} \frac{\log \boldsymbol{\varepsilon}^{-1}}{\log(4n^2/\boldsymbol{\varepsilon_0})}$.

Deducing the improved accuracy of $\mathbf{M}^n$

Reminder

$\mathbf{M}$ is some BP, $\mathbf{L} \stackrel{\text{def}}{=} \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{M}$, and

$$\mathbf{L}^{-1} = \begin{pmatrix} \mathbf{M}^0 & 0 & 0 \\ \vdots & \ddots & 0 \\ \mathbf{M}^n & \cdots & \mathbf{M}^0 \end{pmatrix}, \widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \widetilde{\mathbf{M}^0} & 0 & 0 \\ \vdots & \ddots & 0 \\ \widetilde{\mathbf{M}^n} & \cdots & \widetilde{\mathbf{M}^0} \end{pmatrix}$$

where $\left\| \widetilde{\mathbf{M}^i} - \mathbf{M}^i \right\| \leq \boldsymbol{\varepsilon_0} / n$.

Richardson

$$\left\| \widetilde{\mathbf{L}^{-1}} - \mathbf{L}^{-1} \right\| \leq \boldsymbol{\varepsilon_0} \quad \Longrightarrow \quad \left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq \left\| \mathbf{L}^{-1} \right\| \cdot \left\| \mathbf{L} \right\|^{k+1} \cdot \boldsymbol{\varepsilon_0}^{k+1} \stackrel{\text{def}}{=} \boldsymbol{\varepsilon}.$$

It's not hard to convince that

$$\left\| \mathbf{L} \right\| \leq 2 \quad , \quad \left\| \mathbf{L}^{-1} \right\| \leq n + 1 \leq 2n$$

thus,

$$\left\| (\mathbf{R}_k - \mathbf{L}^{-1})_{n+1,1} \right\| \leq \left\| \mathbf{R}_k - \mathbf{L}^{-1} \right\| \leq 2n \cdot (2 \cdot \boldsymbol{\varepsilon_0})^{k+1} \stackrel{\text{def}}{=} \boldsymbol{\varepsilon},$$

so we need to take $k \stackrel{\text{def}}{=} \frac{\log \boldsymbol{\varepsilon}^{-1}}{\log(4n^2/\boldsymbol{\varepsilon_0})}$.

Implied WPRG

Let $G : \{0,1\}^{s_0} \to \{0,1\}^n$ be some $\boldsymbol{\varepsilon_0}$ PRG, and denote $G_i(x) = G(x)_{0,\ldots,i-1}$.

Implied WPRG

Let $G : \{0,1\}^{s_0} \to \{0,1\}^n$ be some $\boldsymbol{\varepsilon_0}$ PRG, and denote $G_i(x) = G(x)_{0,\dots,i-1}$.
Recall,

$$\mathbf{R}_{k=1} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{M} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{M}}\mathbf{M} + \mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}^2} & \mathbf{M} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^2}\mathbf{M} - \widetilde{\mathbf{M}^2}\widetilde{\mathbf{M}} + \widetilde{\mathbf{M}}\mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}}\widetilde{\mathbf{M}^2} + \mathbf{M}\widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}}\mathbf{M} + \mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}^2} & \mathbf{M} & \mathbf{I} \end{pmatrix}$$

Implied WPRG

Let $G : \{0,1\}^{s_0} \to \{0,1\}^n$ be some ε_0 PRG, and denote $G_i(x) = G(x)_{0,\ldots,i-1}$.
Recall,

$$\mathbf{R}_{k=1} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{M} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{M}}\mathbf{M} + \mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}^2} & \mathbf{M} & \mathbf{I} & 0 \\ \widetilde{\mathbf{M}^2}\mathbf{M} - \widetilde{\mathbf{M}^2}\widetilde{\mathbf{M}} + \widetilde{\mathbf{M}}\mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}}\widetilde{\mathbf{M}^2} + \mathbf{M}\widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}}\mathbf{M} + \mathbf{M}\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}^2} & \mathbf{M} & \mathbf{I} \end{pmatrix}$$

and for general k, it can be shown easily that

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j_r=1}^{O(k)} \widetilde{\mathbf{M}^{i_{j_r}}} \right) = \sum_{r=1}^{n^k} \left(\pm \prod_{j_r=1}^{O(k)} \mathop{\mathbb{E}}_{x_{j_r} \in \{0,1\}^{s_0}} [\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}] \right)$$

Implied WPRG

Let $G:\{0,1\}^{s_0}\to\{0,1\}^n$ be some ε_0 PRG, and denote $G_i(x)=G(x)_{0,\dots,i-1}$.
Recall,

$$\mathbf{R}_{k=1}=\begin{pmatrix}\mathbf{I} & 0 & 0 & 0\\ \mathbf{M} & \mathbf{I} & 0 & 0\\ \widetilde{\mathbf{M}}\mathbf{M}+\mathbf{M}\widetilde{\mathbf{M}}-\widetilde{\mathbf{M}^2} & \mathbf{M} & \mathbf{I} & 0\\ \widetilde{\mathbf{M}^2}\mathbf{M}-\widetilde{\mathbf{M}^2}\widetilde{\mathbf{M}}+\widetilde{\mathbf{M}}\mathbf{M}\widetilde{\mathbf{M}}-\widetilde{\mathbf{M}}\widetilde{\mathbf{M}^2}+\mathbf{M}\widetilde{\mathbf{M}^2} & \widetilde{\mathbf{M}}\mathbf{M}+\mathbf{M}\widetilde{\mathbf{M}}-\widetilde{\mathbf{M}^2} & \mathbf{M} & \mathbf{I}\end{pmatrix}$$

and for general k, it can be shown easily that

$$\mathbf{R}_k[n+1,1]=\sum_{r=1}^{n^k}\left(\pm\prod_{j_r=1}^{O(k)}\widetilde{\mathbf{M}^{i_{j_r}}}\right)=\sum_{r=1}^{n^k}\left(\pm\prod_{j_r=1}^{O(k)}\mathop{\mathbb{E}}_{x_{j_r}\in\{0,1\}^{s_0}}[\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}]\right)$$

So how much does it cost to sample k **different** seeds from G?

Implied WPRG (cont.)

Let's say we take G as [Nis92] (actually, it's the best we can...),

Nisan [Nis92] PRG

For every n, w, ε there exists PRG against (n, w) BP with seed

$$O(\log n \cdot (\log n + \log w + \log \varepsilon^{-1}))$$

Recall that we need $\boldsymbol{\varepsilon_0} \stackrel{\text{def}}{=} 1/4n^2$, so

$$\begin{aligned} s_0 &= O(\log n \cdot (\log n + \log w + \log \boldsymbol{\varepsilon_0}^{-1})) \\ &= O(\log n \cdot (\log n + \log w)). \end{aligned}$$

Implied WPRG (cont.)

Let's say we take G as [Nis92] (actually, it's the best we can...),

Nisan [Nis92] PRG

For every n, w, ε there exists PRG against (n, w) BP with seed

$$O(\log n \cdot (\log n + \log w + \log \varepsilon^{-1}))$$

Recall that we need $\boldsymbol{\varepsilon_0} \stackrel{\text{def}}{=} 1/4n^2$, so

$$\begin{aligned} s_0 &= O(\log n \cdot (\log n + \log w + \log \boldsymbol{\varepsilon_0}^{-1})) \\ &= O(\log n \cdot (\log n + \log w)). \end{aligned}$$

Since we need $O(k)$ different seeds, where

$$O(k) = O\left(\frac{\log \boldsymbol{\varepsilon}^{-1}}{\log \boldsymbol{\varepsilon_0}^{-1}}\right) = O\left(\frac{\log \boldsymbol{\varepsilon}^{-1}}{\log n}\right)$$

the new seed length $\mathsf{s}_{\mathsf{new}}$ is

$$\begin{aligned} \mathsf{s}_{\mathsf{new}} &= O(k \cdot s_0) \\ &= O\left(\frac{\log \boldsymbol{\varepsilon}^{-1}}{\log n} \cdot \log n \cdot (\log n + \log w)\right) \\ &= \log \boldsymbol{\varepsilon}^{-1} \cdot (\log n + \log w) \end{aligned}$$

and sadly, we've got longer seed $\mathsf{s}_{\mathsf{new}} \gg s_0$, as $\boldsymbol{\varepsilon} \ll 1/n$...

Derandomizing the seeds selection

Recall that

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j_r=1}^{O(k)} \mathop{\mathbb{E}}_{x_{j_r} \in \{0,1\}^{s_0}} [\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}] \right)$$

Derandomizing the seeds selection

Recall that

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j_r=1}^{O(k)} \mathop{\mathbb{E}}_{x_{j_r} \in \{0,1\}^{s_0}} [\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}] \right)$$

By setting

$$\mathbf{N}^{(j_r)} \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{x_{j_r} \in \{0,1\}^{s_0}} [\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}]$$

Derandomizing the seeds selection

Recall that

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j_r=1}^{O(k)} \mathop{\mathbb{E}}_{x_{j_r} \in \{0,1\}^{s_0}} [\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}] \right)$$

By setting

$$\mathbf{N}^{(j_r)} \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{x_{j_r} \in \{0,1\}^{s_0}} [\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}]$$

we observe that $\mathbf{R}_k[n+1,1]$ can bee seen as BP with alphabet $\Sigma = \{0,1\}^{s_0}$ (rather then $\{0,1\}$):

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j=1}^{O(k)} \mathbf{N}^{(j_r)} \right)$$

Reminder

Let $\mathbf{M}$ be an $(n, w, \Sigma = \{0,1\})$ BP with all identical layers. Our goal is to approximate $\mathbf{M}^n$ where,

$$\mathbf{M}^n = \left(\frac{1}{2} \left(\mathbf{M}^{(0)} + \mathbf{M}^{(1)} \right) \right)^n = \prod_{i=1}^{n} \mathop{\mathbb{E}}_{i \in \{0,1\}} \mathbf{M}^{(i)}$$

Derandomizing the seeds selection

Recall that

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j_r=1}^{O(k)} \mathop{\mathbb{E}}_{x_{j_r} \in \{0,1\}^{s_0}} [\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}] \right)$$

By setting

$$\mathbf{N}^{(j_r)} \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{x_{j_r} \in \{0,1\}^{s_0}} [\mathbf{M}^{(G_{i_{j_r}}(x_{j_r}))}]$$

we observe that $\mathbf{R}_k[n+1,1]$ can bee seen as BP with alphabet $\Sigma = \{0,1\}^{s_0}$ (rather then $\{0,1\}$):

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j=1}^{O(k)} \mathbf{N}^{(j_r)} \right)$$

Reminder

Let $\mathbf{M}$ be an $(n, w, \Sigma = \{0,1\})$ BP with all identical layers. Our goal is to approximate $\mathbf{M}^n$ where,

$$\mathbf{M}^n = \left(\frac{1}{2} \left(\mathbf{M}^{(0)} + \mathbf{M}^{(1)} \right) \right)^n = \prod_{i=1}^{n} \mathop{\mathbb{E}}_{i \in \{0,1\}} \mathbf{M}^{(i)}$$

Derandomizing the seeds selection (cont.)

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j=1}^{O(k)} \mathbf{N}^{(j_r)} \right) \quad , \quad \text{where } \mathbf{N}^{(j_r)} \text{ is an } (n', w', \Sigma = \{0,1\}^{s_0}) \text{ BP}$$

Impagliazzo, Nisan, and Wigderson [INW94] PRG

There exists $(n', w', \Sigma, \varepsilon_{\mathrm{INW}})$ PRG with seed length

$$\mathsf{s}_{\mathsf{INW}} = O(\log n' \cdot (\log n' + \log w' + \log \varepsilon_{\mathrm{INW}}^{-1})) + \log |\Sigma|$$

Derandomizing the seeds selection (cont.)

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j=1}^{O(k)} \mathbf{N}^{(j_r)} \right) \quad , \quad \text{where } \mathbf{N}^{(j_r)} \text{ is an } (n', w', \Sigma = \{0,1\}^{s_0}) \text{ BP}$$

Impagliazzo, Nisan, and Wigderson [INW94] PRG

There exists $(n', w', \Sigma, \varepsilon_{\mathrm{INW}})$ PRG with seed length

$$\mathsf{s}_{\mathsf{INW}} = O(\log n' \cdot (\log n' + \log w' + \log \varepsilon_{\mathrm{INW}}^{-1})) + \log |\Sigma|$$

Our settings are:

$$n' = O(k) = O\left(\frac{\log \boldsymbol{\varepsilon}^{-1}}{\log \boldsymbol{\varepsilon_0}^{-1}}\right), \qquad |\Sigma| = 2^{s_0}$$

$$w' = \boldsymbol{w}, \qquad \varepsilon_{\mathrm{INW}} = \boldsymbol{\varepsilon}/n^k = \boldsymbol{\varepsilon}^2$$

Derandomizing the seeds selection (cont.)

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j=1}^{O(k)} \mathbf{N}^{(j_r)} \right) \quad , \quad \text{where } \mathbf{N}^{(j_r)} \text{ is an } (n', w', \Sigma = \{0,1\}^{s_0}) \text{ BP}$$

Impagliazzo, Nisan, and Wigderson [INW94] PRG

There exists $(n', w', \Sigma, \varepsilon_{\mathrm{INW}})$ PRG with seed length

$$\mathsf{s}_{\mathsf{INW}} = O(\log n' \cdot (\log n' + \log w' + \log \varepsilon_{\mathrm{INW}}^{-1})) + \log |\Sigma|$$

Our settings are:

$$n' = O(k) = O\left(\frac{\log \boldsymbol{\varepsilon}^{-1}}{\log \boldsymbol{\varepsilon_0}^{-1}}\right), \qquad |\Sigma| = 2^{s_0}$$

$$w' = \boldsymbol{w}, \qquad \varepsilon_{\mathrm{INW}} = \boldsymbol{\varepsilon}/n^k = \boldsymbol{\varepsilon}^2$$

Derandomizing the seeds selection (cont.)

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k} \left(\pm \prod_{j=1}^{O(k)} \mathbf{N}^{(j_r)} \right) \quad , \quad \text{where } \mathbf{N}^{(j_r)} \text{ is an } (n', w', \Sigma = \{0,1\}^{s_0}) \text{ BP}$$

Impagliazzo, Nisan, and Wigderson [INW94] PRG

There exists $(n', w', \Sigma, \varepsilon_{\text{INW}})$ PRG with seed length

$$s_{\text{INW}} = O(\log n' \cdot (\log n' + \log w' + \log \varepsilon_{\text{INW}}^{-1})) + \log |\Sigma|$$

Our settings are:

$$n' = O(k) = O\left(\frac{\log \boldsymbol{\varepsilon}^{-1}}{\log \boldsymbol{\varepsilon_0}^{-1}}\right), \qquad |\Sigma| = 2^{s_0}$$

$$w' = \boldsymbol{w}, \qquad \varepsilon_{\text{INW}} = \boldsymbol{\varepsilon}/n^k = \boldsymbol{\varepsilon}^2$$

thus, the INW seed is:

$$\begin{aligned} s_{\text{INW}} &= O(\log k \cdot (\log k + \log w + \log \varepsilon_{\text{INW}}^{-1})) + s_0 \\ &= s_0 + O((\log \boldsymbol{\varepsilon}^{-1} + \log w) \cdot \log\log(1/\boldsymbol{\varepsilon})). \end{aligned}$$

Derandomizing the seeds selection (cont.)

$$\mathbf{R}_k[n+1,1] = \sum_{r=1}^{n^k}\left(\pm\prod_{j=1}^{O(k)} \mathbf{N}^{(j_r)}\right) \quad , \quad \text{where } \mathbf{N}^{(j_r)} \text{ is an } (n', w', \Sigma = \{0,1\}^{s_0}) \text{ BP}$$

Impagliazzo, Nisan, and Wigderson [INW94] PRG

There exists $(n', w', \Sigma, \varepsilon_{\mathrm{INW}})$ PRG with seed length

$$\mathsf{s}_{\mathsf{INW}} = O(\log n' \cdot (\log n' + \log w' + \log \varepsilon_{\mathrm{INW}}^{-1})) + \log|\Sigma|$$

Our settings are:

$$n' = O(k) = O\left(\frac{\log \boldsymbol{\varepsilon}^{-1}}{\log \boldsymbol{\varepsilon_0}^{-1}}\right), \qquad |\Sigma| = 2^{s_0}$$

$$w' = \boldsymbol{w}, \qquad \varepsilon_{\mathrm{INW}} = \boldsymbol{\varepsilon}/n^k = \boldsymbol{\varepsilon}^2$$

thus, the INW seed is:

$$\begin{aligned}\mathsf{s}_{\mathsf{INW}} &= O(\log k \cdot (\log k + \log w + \log \varepsilon_{\mathrm{INW}}^{-1})) + s_0 \\ &= s_0 + O((\log \boldsymbol{\varepsilon}^{-1} + \log w) \cdot \log\log(1/\boldsymbol{\varepsilon})).\end{aligned}$$

So we conclude an $\boldsymbol{\varepsilon}$ Weighted PRG against $(n, \boldsymbol{w})$ with seed length

$$\mathsf{s}_{\mathsf{new}} = \mathsf{s}_{\mathsf{INW}} = s_0 + \widetilde{O}(\log \boldsymbol{\varepsilon}^{-1} + \log w). \qquad ☺$$

Table of Contents

Subsequent work

Further improvement. Hoza [Hoz21] showed how to combine an idea from Armoni [Arm98] to remove the $\log\log$ factors for PRGs with *inherent gap* between their seed length and space complexity.
Since Nisan's PRG indeed "blessed" with that gap, he concluded a WPRG with seed length

$$O(\log n \cdot (\log n + \log w) + \log \varepsilon^{-1})$$

Error reduction from $\varepsilon_0 \gg 1/n$. For the restricted class of *permutation* branching programs, Pyne and Vadhan [PV21] established the analysis under another norm $\|\cdot\|$, and so concluded error reduction when $\varepsilon_0 = 1/\log n$.

Table of Contents

Future Directions

1. **Randomize** [RR99] to get better space in price of seed
2. Eliminate the restriction of the error reduction from $\boldsymbol{\varepsilon_0} \approx 1/n$
3. Compositions of Weighted PRGs would yield **immediate improvement** of Nisan [Nis92] PRG

Future Directions

1. **Randomize** [RR99] to get better space in price of seed
2. Eliminate the restriction of the error reduction from $\boldsymbol{\varepsilon_0} \approx 1/n$
3. Compositions of Weighted PRGs would yield **immediate improvement** of Nisan [Nis92] PRG

Thanks!