Error Reduction for Weight PRG against ROBP

Gil Cohen

Joint work with Dean Doron, Oren Renard, Ori Sberlo and Amnon Ta-Shma

August 12, 2021

Outline

The **BPL** vs. **L** Problem

The Problem

Derandomize with minimal overhead in space.

Given a randomized algorithm with space complexity $s = s(n)$, n being the input length, devise a deterministic algorithm with comparable space complexity $s' = s'(s)$.

Whether or not derandomization with constant overhead in space $s' = O(s)$ is possible is known as the **BPL** $=$ **L** problem.

What is known?

- Savitch's Theorem (1970) implies $\mathbf{RL} \subseteq \mathbf{NL} \subseteq \mathbf{L}^2$.
- Borodin-Cook-Pippenger (1983) established $\mathbf{BPL} \subseteq \mathbf{L}^2$.
- Nisan (1992, 94) proved that $\mathbf{BPL} \subseteq \mathbf{SC}$.
- Saks-Zhou (1999) proved $\mathbf{BPL} \subseteq \mathbf{L}^{3/2}$. A slight improvement has been made very recently by Hoza (2021).
- Exciting developments in recent years (see STOC 2020 - Workshop 6: Derandomizing Space-Bounded Computation) but no improvement for the general case.

Outline

Pseudorandom generators

Denote by $\mathcal{F}_n$ the set of all functions $f : \{0,1\}^n \to \{0,1\}$.

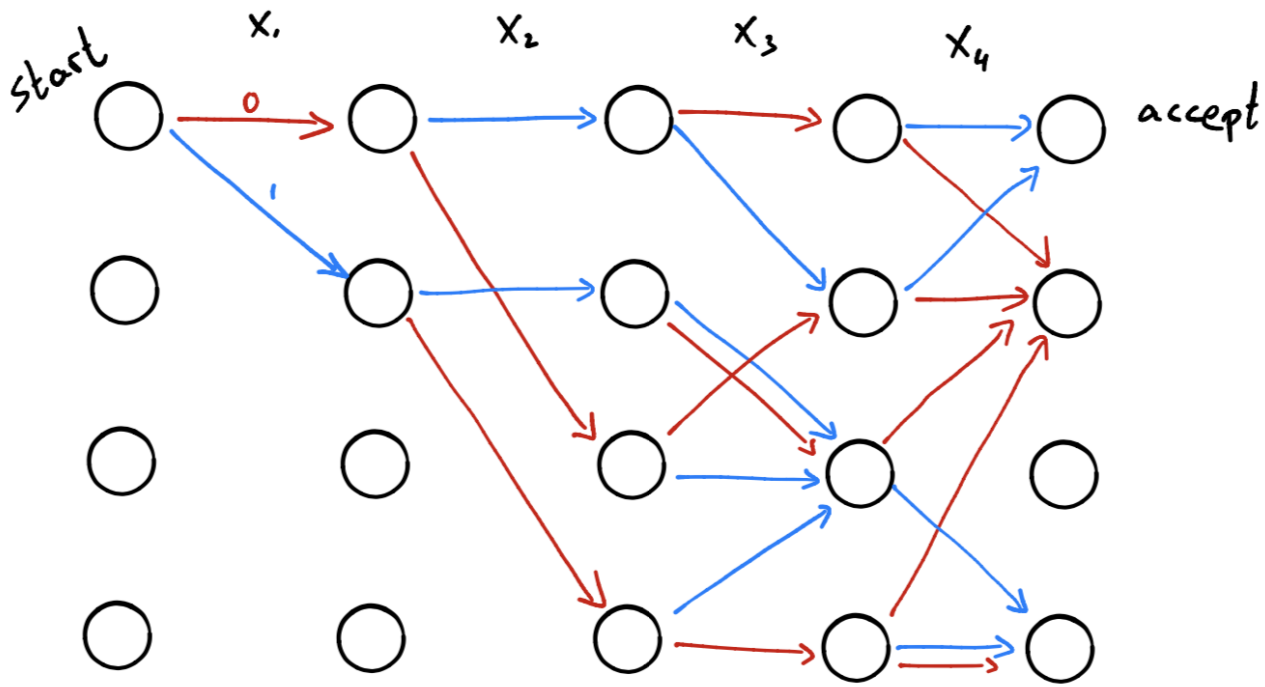
Definition

Let $\mathcal{C} \subset \mathcal{F}_n$ be a class of functions. A Pseudorandom generator (PRG) with error ε against $\mathcal{C}$ is a function $\mathrm{G} : \{0,1\}^s \to \{0,1\}^n$ s.t for every $f \in \mathcal{C}$

$$\left|\mathbb{E}[f(U_n)] - \mathbb{E}[f(\mathrm{G}(U_s))]\right| \leq \varepsilon.$$

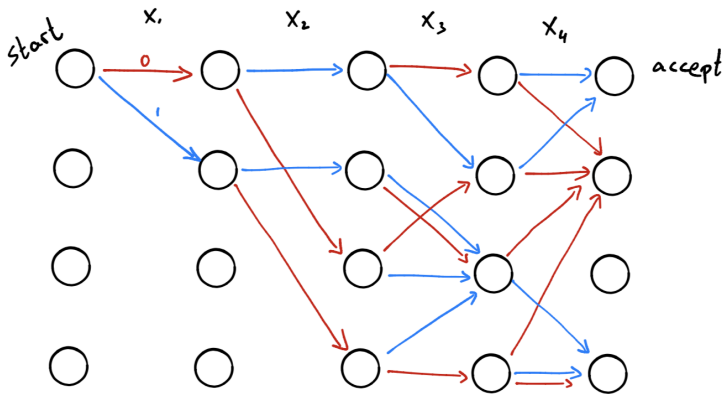
s is called the seed length.

Read-once branching-programs



An example of a width $w = 4$, length $n = 4$ ROBP.

PRG for ROBP



Proposition

For every n, w, ε, there exists a PRG for width-w length-n ROBP with

$$s_{\text{opt}} = O\left(\log n + \log w + \log \varepsilon^{-1}\right).$$

Theorem (Nisan 1992)

For every n, w, ε there exists a space-efficient PRG for (w, n)-ROBP with

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

The three decouplings

Theorem (Nisan 1992)

For every n, w, ε there exists a space-efficient PRG for (w, n)-ROBP with

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

For derandomizing **BPL** via the naïve derandomization, $w = n^{O(1)}$, $\varepsilon = O(1)$ and so $s = O(\log^2 n)$ which gives $\mathbf{BPL} \subseteq \mathbf{L}^2$.

Saks-Zhou applies Nisan's PRG in a nontrivial way in a regime in which $w, \varepsilon^{-1} \gg n$ to get their result $\mathbf{BPL} \subseteq \mathbf{L}^{3/2}$.

The three decouplings

Braverman-Cohen-Garg (2018) had two observations that motivated their work:

Observation 1. A PRG with seed length

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

when used in the Saks-Zhou framework, would yield $\mathbf{BPL} \subseteq \mathbf{L}^{4/3}$.

Observation 2. The reason for the $\log n \cdot \log n$ term is the way that the error evolves in Nisan's recursive construction.

Thus, improving the way the error evolves may just solve both problem, leaving us with seed length

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

The error parameter

The main result of BCG is essentially a PRG with seed length

$$s_{\mathrm{BCG}} = \widetilde{O}\left(\log n \cdot (\log n + \log w) + \log \varepsilon^{-1}\right).$$

More precisely, BCG introduced and constructed weighted PRG.

Definition

Let $\mathcal{C} \subset \mathcal{F}_n$ be a class of functions. A weighted pseudorandom generator (WPRG) with error ε against $\mathcal{C}$ is a function

$$(\mathrm{G}, \omega) : \{0,1\}^s \to \{0,1\}^n \times \mathbb{R}$$

s.t. $\forall f \in \mathcal{C}$

$$\left|\mathbb{E}[f(U_n)] - \mathbb{E}[\omega(U_s) \cdot f(\mathrm{G}(U_s))]\right| \leq \varepsilon.$$

The weights can be both positive and negative and not necessarily bounded. WPRG are as good as PRG for the naïve derandomization and also for the Saks-Zhou framework.

A somewhat simplified construction was obtained by Chattopadhyay and Liao (2020).

The width parameter

Raz-Reingold (1999) suggested a beautiful idea on how to to decouple the width so to obtain seed length

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

BCG's construction is quite involved (construction and analysis is over 40 pages). This makes it extremely difficult to combine with RR so to, hopefully, get

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

and thus (when combined also with Saks-Zhou) to yield $\mathbf{BPL} \subseteq \mathbf{L}^{4/3}$.

New result. The main result of CDRSTS (2021) is a much simpler way to decouple ε. More precisely, a way to reduce a modest error $\approx \frac{1}{nw}$ to a desired one ε.

The result was obtained independently and concurrently by Pyne and Vadhan (2021).

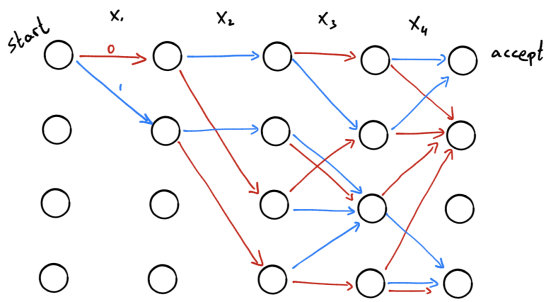
Outline

From ROBP to matrix powering

1. By blowing up $w \to wn$, we may assume all layers are the same.
2. If $\mathbf{A}$ is the transition matrix of the layer, and assuming the accept state is the first in its layer, our goal is to approximate $(\mathbf{A}^n)_{1,1}$.
3. Instead, we can approximate $\mathbf{A}^n$ in $\|\cdot\|_{\max}$.
4. It is more convenient to work with $\|\cdot\|_\infty$ as it is sub-multiplicative.
5. More precisely, if $\mathbf{A} = \frac{\mathbf{A}^{(0)}+\mathbf{A}^{(1)}}{2}$ where $\mathbf{A}^{(0)}, \mathbf{A}^{(1)}$ are zero-one stochastic matrices, then we wish to approximate

$$\mathbf{A}^n = \mathbb{E}_{\sigma\in\{0,1\}^n} \mathbf{A}^{(\sigma)}$$

where $\mathbf{A}^{(\sigma)} = \prod_{i=1}^n \mathbf{A}^{(\sigma_i)}$.



From ROBP to matrix powering

1. More precisely, if $\mathbf{A} = \frac{\mathbf{A}^{(0)}+\mathbf{A}^{(1)}}{2}$ where $\mathbf{A}^{(0)}, \mathbf{A}^{(1)}$ are zero-one stochastic matrices, then we wish to approximate

$$\mathbf{A}^n = \mathbb{E}_{\sigma\in\{0,1\}^n}\mathbf{A}^{(\sigma)}$$

   where $\mathbf{A}^{(\sigma)} = \prod_{i=1}^{n}\mathbf{A}^{(\sigma_i)}$.
2. A PRG $\mathrm{G} : \{0,1\}^s \to \{0,1\}^n$ with seed length s is redefined to be

$$\|\mathbf{A}^n - \mathbb{E}_{\tau\in\{0,1\}^s}[\mathbf{A}^{(\mathrm{G}(\tau))}]\| \leq \varepsilon.$$

3. A WPRG $(\mathrm{G},\omega) : \{0,1\}^s \to \{0,1\}^n \times \mathbb{R}$ is redefined to be

$$\|\mathbf{A}^n - \mathbb{E}_{\tau\in\{0,1\}^s}[\omega(\tau)\cdot\mathbf{A}^{(\mathrm{G}(\tau))}]\| \leq \varepsilon.$$

Outline

Powers and inverse

Say we want to compute $\mathbf{A}^n$. Observe that, symbolically,

$$(\mathbf{I}-\mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^n + \cdots$$

To avoid this "interference" of all powers we can consider the tensor with the directed path graph. E.g.,

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

As $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD})$ and since $\mathbf{P}_{n+1}^{n+1} = 0$,

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

Powers and inverse

$$(\mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{A})^{-1} = \mathbf{I} + (\mathbf{P}_{n+1} \otimes \mathbf{A}) + (\mathbf{P}_{n+1}^2 \otimes \mathbf{A}^2) + \cdots + (\mathbf{P}_{n+1}^n \otimes \mathbf{A}^n).$$

For example,

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

Those with spectral graph theory background will recognize $\mathbf{L} = \mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{A}$ as the Laplacian of the directed graph $\mathbf{P}_{n+1} \otimes \mathbf{A}$.

Richardson iterations

Say $\widetilde{\mathbf{L}^{-1}}$ is a modest approximation to $\mathbf{L}^{-1}$. Specifically,

$$\|\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L}\| \leq \varepsilon_0$$

(rather than $\|\widetilde{\mathbf{L}^{-1}} - \mathbf{L}\| \leq \varepsilon_0$). Define

$$\mathbf{L}_k = \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})^i \widetilde{\mathbf{L}^{-1}}.$$

Note that, symbolically,

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

More generally,

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

Richardson iterations

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

So

$$\mathbf{I} - \mathbf{L}_k\mathbf{L} = (\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})^{k+1},$$

and

$$\|\mathbf{I} - \mathbf{L}_k\mathbf{L}\| = \|(\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})^{k+1}\| \leq \|(\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})\|^{k+1} \leq \varepsilon_0^{k+1}.$$

To summarize,

$$\|\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L}\| \leq \varepsilon_0 \quad \Longrightarrow \quad \|\mathbf{I} - \mathbf{L}_k\mathbf{L}\| \leq \varepsilon_0^{k+1},$$

where

$$\mathbf{L}_k = \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})^i \widetilde{\mathbf{L}^{-1}}.$$

Richardson iterations

$$\|\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L}\| \leq \varepsilon_0 \quad \Longrightarrow \quad \|\mathbf{I} - \mathbf{L}_k\mathbf{L}\| \leq \varepsilon_0^{k+1},$$

Thus, to obtain a good ε approximation of $\mathbf{A}^n$, we

1. Compute a modest ε_0 approximation $\widetilde{\mathbf{A}}^i$ of $\mathbf{A}^i$ for $1 \leq i \leq n$. Namely, $\|\widetilde{\mathbf{A}}^i - \mathbf{A}^i\| \leq \varepsilon_0$.
2. Construct
$$\widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ \widetilde{\mathbf{A}} & \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}} & \mathbf{I} & 0 & 0 \\ \vdots & \vdots & \widetilde{\mathbf{A}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{A}^n} & \widetilde{\mathbf{A}^{n-1}} & \dots & \widetilde{\mathbf{A}} & \mathbf{I} \end{pmatrix}.$$
3. Compute $\mathbf{L}_k = \sum_{i=0}^{k} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})^i \widetilde{\mathbf{L}^{-1}}$ for $k = \frac{\log \varepsilon^{-1}}{\log \varepsilon_0^{-1}}$.
4. Return the bottom-left block of $\mathbf{L}_k$.

Example $k = 1, n = 3$

Consider a single iteration ($k = 1$) with $n = 3$:

$$\mathbf{L}_1 = \sum_{i=0}^{k=1} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})^i \widetilde{\mathbf{L}^{-1}},$$

where recall

$$\widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{A}} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{A}^3} & \widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}} & \mathbf{I} \end{pmatrix} \qquad \mathbf{L} = \mathbf{I} - \mathbf{P}_4 \otimes \mathbf{A} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ -\mathbf{A} & \mathbf{I} & 0 & 0 \\ 0 & -\mathbf{A} & \mathbf{I} & 0 \\ 0 & 0 & -\mathbf{A} & \mathbf{I} \end{pmatrix}$$

Then,

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

Example $k = 1, n = 3$

Let us consider a single iteration ($k = 1$) with $n = 3$:

$$\mathbf{L}_1 = \sum_{i=0}^{k=1} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})^i \widetilde{\mathbf{L}^{-1}},,$$

where recall

$$\widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{A}} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{A}^3} & \widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}} & \mathbf{I} \end{pmatrix} \qquad \mathbf{L} = \mathbf{I} - \mathbf{P}_4 \otimes \mathbf{A} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ -\mathbf{A} & \mathbf{I} & 0 & 0 \\ 0 & -\mathbf{A} & \mathbf{I} & 0 \\ 0 & 0 & -\mathbf{A} & \mathbf{I} \end{pmatrix}$$

Then,

$$\mathbf{L}_1 = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{A} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{A}}\mathbf{A} + \mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^2} & \mathbf{A} & \mathbf{I} & 0 \\ \widetilde{\mathbf{A}^2}\mathbf{A} - \widetilde{\mathbf{A}^2}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}\mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}\widetilde{\mathbf{A}^2} + \mathbf{A}\widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}}\mathbf{A} + \mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^2} & \mathbf{A} & \mathbf{I} \end{pmatrix}.$$

Example $k = 1, n = 4$

$$\widetilde{\mathbf{L}^{-1}} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ \widetilde{\mathbf{A}} & \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{A}^3} & \widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}} & \mathbf{I} & 0 \\ \widetilde{\mathbf{A}^4} & \widetilde{\mathbf{A}^3} & \widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}} & \mathbf{I} \end{pmatrix} \quad \mathbf{L} = \mathbf{I} - \mathbf{P}_5 \otimes \mathbf{A} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ -\mathbf{A} & \mathbf{I} & 0 & 0 & 0 \\ 0 & -\mathbf{A} & \mathbf{I} & 0 & 0 \\ 0 & 0 & -\mathbf{A} & \mathbf{I} & 0 \\ 0 & 0 & 0 & -\mathbf{A} & \mathbf{I} \end{pmatrix}$$

Thus,

$$\mathbf{L}_1 = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ \mathbf{A} & \mathbf{I} & 0 & 0 & 0 \\ \widetilde{\mathbf{A}}\mathbf{A} + \mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^2} & \mathbf{A} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{A}^2}\mathbf{A} - \widetilde{\mathbf{A}^2}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}\mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}\widetilde{\mathbf{A}^2} + \mathbf{A}\widetilde{\mathbf{A}^2} & \nwarrow & \mathbf{A} & \mathbf{I} & 0 \\ \widetilde{\mathbf{A}^3}\mathbf{A} + \widetilde{\mathbf{A}}\widetilde{\mathbf{A}^2}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^3}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}\mathbf{A}\widetilde{\mathbf{A}^2} - \widetilde{\mathbf{A}^2}\widetilde{\mathbf{A}^2} + \mathbf{A}\widetilde{\mathbf{A}^3} - \widetilde{\mathbf{A}}\widetilde{\mathbf{A}^3} & \nwarrow & \nwarrow & \mathbf{A} & \mathbf{I} \end{pmatrix}.$$

Example $k = 2, n = 3$

Recall

$$\mathbf{L}_2 = \sum_{i=0}^{k=2} (\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L})^i \widetilde{\mathbf{L}^{-1}}$$

Then,

$$\mathbf{L}_2 = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{A} & 0 & 0 & 0 \\ \mathbf{A}^2 - \mathbf{A}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}^2 + \widetilde{\mathbf{A}}\mathbf{A} - \widetilde{\mathbf{A}^2} & \mathbf{A} & 0 & 0 \\ \text{too long} & \nwarrow & \mathbf{A} & 0 \end{pmatrix}.$$

$$\text{too long} = \widetilde{\mathbf{A}}\mathbf{A}^2 - 2\widetilde{\mathbf{A}}^2\mathbf{A} + 3\widetilde{\mathbf{A}}^3 + \mathbf{A}\widetilde{\mathbf{A}}\mathbf{A} - 3\mathbf{A}\widetilde{\mathbf{A}}^2 + \mathbf{A}^2\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}^2}\mathbf{A} - \widetilde{\mathbf{A}^2}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}\widetilde{\mathbf{A}^2} + \mathbf{A}\widetilde{\mathbf{A}^2}$$

Analysis.

$$\|\widetilde{\mathbf{A}}^i - \mathbf{A}^i\| \leq \varepsilon_0 \quad \Longrightarrow \quad \|\mathbf{L}^{-1} - \widetilde{\mathbf{L}^{-1}}\| \leq n\varepsilon_0$$

and so

$$\|\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L}\| = \|(\mathbf{L}^{-1} - \widetilde{\mathbf{L}^{-1}})\mathbf{L}\| \leq \|\mathbf{L}^{-1} - \widetilde{\mathbf{L}^{-1}}\| \cdot \|\mathbf{L}\| \leq n\varepsilon_0 \cdot \|\mathbf{L}\|.$$

Recall that

$$\|\mathbf{L}\| = \|\mathbf{I} - \mathbf{P}_{n+1} \otimes \mathbf{A}\| \leq \|\mathbf{I}\| + \|\mathbf{P}_{n+1}\| \cdot \|\mathbf{A}\| \leq 2,$$

and so

$$\|\mathbf{I} - \widetilde{\mathbf{L}^{-1}}\mathbf{L}\| \leq 2n\varepsilon_0.$$

Thus,

$$\|\mathbf{I} - \mathbf{L}_k\mathbf{L}\| \leq (2n\varepsilon_0)^{k+1}.$$

Analysis. So far

$$\|\mathbf{I} - \mathbf{L}_k\mathbf{L}\| \leq (2n\varepsilon_0)^{k+1}.$$

We, however, care about

$$\begin{aligned}\|\mathbf{L}^{-1} - \mathbf{L}_k\| &= \|(\mathbf{I} - \mathbf{L}_k\mathbf{L})\mathbf{L}^{-1}\| \\ &\leq \|\mathbf{I} - \mathbf{L}_k\mathbf{L}\| \cdot \|\mathbf{L}^{-1}\| \\ &\leq (n+1)(2n\varepsilon_0)^{k+1}.\end{aligned}$$

Thus by taking, e.g., $\varepsilon_0 = \frac{1}{4n^2}$ we get final error ε by taking $k \approx \log_n \varepsilon^{-1}$.

Outline

Error reduction via spectral methods

Let us redo this in the "black box" setting. Say that

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

is given such that each

$$\widetilde{\mathbf{A}}^i = \mathbb{E}_{\tau \sim \{0,1\}^s}[\mathbf{A}^{(\mathrm{G}_i(\tau))}]$$

for some $\mathrm{G}_i : \{0,1\}^s \to \{0,1\}^i$.

Error reduction via spectral methods

Recall

$$\mathbf{L}_1 = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{A} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{A}}\mathbf{A} + \mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^2} & \mathbf{A} & \mathbf{I} & 0 \\ \widetilde{\mathbf{A}^2}\mathbf{A} - \widetilde{\mathbf{A}^2}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}\mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}\widetilde{\mathbf{A}^2} + \mathbf{A}\widetilde{\mathbf{A}^2} & \widetilde{\mathbf{A}}\mathbf{A} + \mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^2} & \mathbf{A} & \mathbf{I} \end{pmatrix}.$$

A general entry of $\mathbf{L}_k$ is a sum of terms of the form

$$\pm \prod_{j=1}^{O(k)} \widetilde{\mathbf{A}^{i_j}} = \pm \prod_{j=1}^{O(k)} \mathbb{E}_{\tau_j \sim \{0,1\}^s} [\mathbf{A}^{(\mathrm{G}_{i_j}(\tau_j))}]$$

that consists of at most n^k summands. As we saw,

$$\|\widetilde{\mathbf{A}}^i - \mathbf{A}^i\| \le \varepsilon_0 \quad \Longrightarrow \quad \|\mathbf{L}^{-1} - \mathbf{L}_k\| \le (n+1)(2n\varepsilon_0)^{k+1}.$$

By choosing $\varepsilon_0 \approx n^{-2}$ and $k = \log_n \varepsilon^{-1}$ we get error ε.

Error reduction via spectral methods

Recall $\varepsilon_0 \approx n^{-2}$ and $k = \log_n \varepsilon^{-1}$.

The seed length is then $O(ks)$ where

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

and so

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

which is even worse than what we started with.

Error reduction via spectral methods

Recall that a general entry of $\mathbf{L}_k$ is a sum of terms of the form

$$\pm \prod_{j=1}^{O(k)} \widetilde{\mathbf{A}^{i_j}} = \pm \prod_{j=1}^{O(k)} \mathbb{E}_{\tau_j \sim \{0,1\}^s} [\mathbf{A}^{(\mathrm{G}_{i_j}(\tau_j))}].$$

The key observation is that we can avoid using ks random bits as the above product "comes from" a length $O(k)$ width w ROBP with arity 2^s.

Indeed, recall that an arity 2 ROBP can be expressed as

$$\prod_{j=1}^{n} \mathbb{E}_{i \sim \{0,1\}} [\mathbf{A}_j^{(i)}].$$

Luckily, there is a PRG for general arity m with seed length

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

due to Impagliazzo, Nisan and Wigderson (1994). Furthermore, $k \ll n$.

Hence, we define a matrix $\widetilde{\mathbf{L}_k}$ that is obtained by derandomizing the products inside $\mathbf{L}_k$. E.g., if

$$\mathbf{L}_1 = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{A} & \mathbf{I} & 0 & 0 \\ \mathbf{A}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}\mathbf{A} - \widetilde{\mathbf{A}^2} & \mathbf{A} & \mathbf{I} & 0 \\ \mathbf{A}^2\mathbf{A} - \mathbf{A}^2\widetilde{\mathbf{A}} + \mathbf{A}\widetilde{\mathbf{A}}\mathbf{A} - \mathbf{A}\widetilde{\mathbf{A}^2} + \widetilde{\mathbf{A}}\mathbf{A}^2 & \widetilde{\mathbf{A}}\mathbf{A} + \mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^2} & \mathbf{A} & \mathbf{I} \end{pmatrix}.$$

then $\widetilde{\mathbf{L}_1}$ will look like

$$\widetilde{\mathbf{L}_1} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{A} & \mathbf{I} & 0 & 0 \\ \widetilde{\mathbf{A}}\mathbf{A} + \mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^2} & \mathbf{A} & \mathbf{I} & 0 \\ \widetilde{\mathbf{A}^2\mathbf{A} - \mathbf{A}^2\widetilde{\mathbf{A}} + \mathbf{A}\widetilde{\mathbf{A}}\mathbf{A} - \mathbf{A}\widetilde{\mathbf{A}^2} + \widetilde{\mathbf{A}}\mathbf{A}^2} & \widetilde{\widetilde{\mathbf{A}}\mathbf{A} + \mathbf{A}\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}^2}} & \mathbf{A} & \mathbf{I} \end{pmatrix}.$$

Error reduction via spectral methods

If the "outer" PRG has error ε_1 (and $\varepsilon_0 = n^{-2}$, $k = \log_n \varepsilon^{-1}$) then we get an additional error of $\varepsilon_1 n^k$ and so by setting $\varepsilon_1 = n^{-k}\varepsilon$ we error $O(\varepsilon)$ with seed length

$$\begin{aligned} s_{\text{final}} &= s + O(\log k \cdot (\log k + \log w + \log \varepsilon_1^{-1})) \\ &= s + (\log w + \log \varepsilon^{-1}) \log\log \varepsilon^{-1}. \end{aligned}$$

Since $s = O(\log n \cdot (\log n + \log w))$, we have that

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

Space complexity

The space complexity of our error reduction is of the order of

$$m + \left(\log\log\frac{w}{\varepsilon}\right)^3,$$

where m is the space complexity of the given PRG.

When instantiated with Nisan's PRG, we obtain a WPRG with space complexity of the order of

$$\log nw + \left(\log\log\varepsilon^{-1}\right)^3.$$

Outline

Summary

Related work. Hoza (2021) reduced the seed length further to

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

Moreover, concurrent to our work, based on heavier spectral machinery (developed in the context of fast Laplacian solvers), Pyne and Vadhan (2021) obtained better WPRG for permutation ROBP–better than what any PRG can yield.

Future research. Do better.

Thank you!