CS59200-PRS: Pseudorandomness Sept 24, 2025

Lecture 9: INW Generator and Exapander Graphs

Lecturer: Wei Zhan *Scribe: Xiuyu Ye*

1 INW Generator

In Nisan's PRG, we showed that after fixing the distinguisher $A : \{0,1\}^{2\ell} \to [0,1]$, with high probability over pairwise independent hash functions h we have $\mathbb{E}[A(s,s')] \approx \mathbb{E}[A(s,h(s))]$. This is in contrast to the total variation distance, and indeed it is impossible to have a small $d_{\mathrm{TV}}\left((s,s'),(s,h(s))\right)$ for any fixed hash function h. However, we still have

$$d_{\mathrm{TV}}\left((s,s'),(s,h(s))\right) = 0$$

when we consider h as random, and it turns out that this suffices for the fooling argument. In fact, we only need a weaker guarantee on the total variant distance, as the branching program will only store $\log w$ bits of information about s in the first half.

Definition 1 (ε-recycling)**.** *Let $d \in \mathbb{N}^+$. A function $H\colon \{0,1\}^\ell \times [d] \to \{0,1\}^\ell$ is ε-recycling if for every $w \in \mathbb{N}^+$ and every function $F\colon \{0,1\}^\ell \to [w]$,*

$$d_{\mathrm{TV}}\left(\left(F(s),s'\right),(F(s),H(s,r))\right) \le w\varepsilon$$

where s, s' are chosen randomly from $\{0,1\}^\ell$ and r is chosen randomly from $[d]$.

Intuitively, $F(s)$ represents the node at the middle layer of the ROBP, reached by following the PRG output with seed s in the first half. Meanwhile H represents the family of hash functions, and r is the randomness for sampling such a hash function from this family. H effectively recycles the seed s to generate a new seed $H(s,r)$ for the second half of the ROBP. So in the case of Nisan's PRG, H is a pairwise independent family with $d = 2^{O(\ell)}$, and it is 0-recycling.

1.1 The construction

The Impagliazzo-Nisan-Wigderson (INW) generator is build directly upon ε-recycling functions, using the same idea of recursive construction that double the PRG output length as in Nisan's generator. Consider a family of ε-recycling functions with varying inputs lengths

$$\mathcal{H} = \left\{H_k\colon \{0,1\}^{\ell+(k-1)\log d} \times [d] \to \{0,1\}^{\ell+(k-1)\log d}\right\}_{k\in\mathbb{N}}.$$

Initially, let $G_0 \colon \{0,1\}^\ell \to \{0,1\}^\ell$ be $G_0(s) = s$ such that when s is chosen randomly from $\{0,1\}^\ell$, G outputs a random ℓ-bit string. At the k-th step, define

$$G_k(s_k) \;:=\; (G_{k-1}(s_{k-1}), G_{k-1}(H_k(s_{k-1}, r_k)))$$

where s_k stands for the combined seed $s_k = (s_{k-1}, r_k) = (s, r_1, \ldots, r_k)$. Each of the extra randomness $r_1, r_2, \ldots, r_k$ is chosen randomly from $[d]$. For instance,

- $G_1(s, r_1) = (G_0(s), G_0(H_1(s, r_1)))$.
- $G_2(s, r_1, r_2) = (G_1(s, r_1), G_1(H_2(s, r_1, r_2)))$.
- $G_k(s, r_1, \ldots, r_k) = (G_{k-1}(s, r_1, \ldots, r_{k-1}), G_{k-1}(H_k((s, r_1, \ldots, r_{k-1}), r_k)))$.

Remark. Unlike Nisan's PRG, here the seed length is increasing in each step for including the extra randomness (and hence they are part of the PRG parameters instead of subscripts). Why can't we just use the same H_1 to recycle only s? For instance, if we instead let

$$G_2(s, r_1, r_2) = (G_1(s, r_1), G_1(H_1(s, r_2), r_1)),$$

The reused r_1 might introduce some unwanted correlations between the first and the second half. In Nisan's PRG we can fix r_1 since with high probability r_1 is a good choice, but we don't have that privilege here.

Theorem 1. *The INW generator G_k with ε-recycling functions $(2^k - 1)w\varepsilon$-fools all width-w, length-$2^k\ell$ ROBPs.*

Proof. We use induction on k, and assume this is true for $k-1$. Consider an arbitrary width-w, length-$2^k\ell$ ROBP B. With the same hybrid argument as in Nisan's PRG, by triangle inequality,

$$\begin{aligned}
&\left| \mathop{\mathbb{E}}_{s_k \sim \{0,1\}^\ell \times [d]^k} [B(G_k(s_k))] - \mathop{\mathbb{E}}_{x \sim \{0,1\}^{2^k \ell}} [B(x)] \right| \\
\le\; &\left| \mathop{\mathbb{E}}_{s_k \sim \{0,1\}^\ell \times [d]^k} [B(G_k(s_k))] - \mathop{\mathbb{E}}_{s_{k-1}, s'_{k-1} \sim \{0,1\}^\ell \times [d]^{k-1}} \left[B(G_{k-1}(s_{k-1}), G_{k-1}(s'_{k-1})) \right] \right| \\
&+ \left| \mathop{\mathbb{E}}_{s_{k-1}, s'_{k-1}} \left[B(G_{k-1}(s_{k-1}), G_{k-1}(s'_{k-1})) \right] - \mathop{\mathbb{E}}_{\substack{s_{k-1} \\ x_2 \sim \{0,1\}^{2^{k-1}\ell}}} [B(G_{k-1}(s_{k-1}), x_2] \right| \\
&+ \left| \mathop{\mathbb{E}}_{s_{k-1}, x_2} [B(G_{k-1}(s, r_1, \ldots, r_{k-1}), x_2] - \mathop{\mathbb{E}}_{x_1, x_2 \sim \{0,1\}^{2^{k-1}\ell}} [B(x_1, x_2)] \right|.
\end{aligned}$$

The first term in the right hand side is at most $w\varepsilon$ because of the ε-recycling property, and the rest two terms are both bounded by $(2^{k-1} - 1)w\varepsilon$ by induction hypothesis. Therefore the total error is bounded by $w\varepsilon + 2(2^{k-1} + 1)w\varepsilon = (2^k - 1)w\varepsilon$. □

To fooling length-n ROBPs, we can take $\ell = O(1)$, $k = O(\log n)$ and the seed length of the INW generator is $O(\ell + k \log d) = O(\log n \cdot \log d)$, while we need $\varepsilon = O(1/(nw))$. How small could d be? We have the following simple argument.

Theorem 2. *If $H\colon \{0,1\}^\ell \times [d] \to \{0,1\}^\ell$ is ε-recycling, then $d \geq \Omega(\min(\varepsilon^{-1}, 2^\ell))$.*

Proof. Let $F\colon \{0,1\}^\ell \to [w]$ be a function with output as uniform as possible, so that for each $v \in [w]$, $|F^{-1}(v)| \leq \lceil 2^\ell / w \rceil$. Notice that

$$d_{\mathrm{TV}}\left(\left(F(s), s'\right), \left(F(s), H(s,r)\right)\right) = \mathop{\mathbb{E}}_{v}\left[d_{\mathrm{TV}}\left(s', H(s,r) | F(s) = v\right)\right]$$

where v follows the same distribution as $F(s)$, and $H(s,r)|F(s) = v$ is the distribution of $H(s,r)$ conditioned on $F(s) = v$, which has a support of size of $|F^{-1}(v)| \cdot d$.

On the other hand, when we take $w = 1/2\varepsilon$, the total variation distance above is at most $1/2$, and that means for at least one v, the support size of $H(s,r)|F(s) = v$ is at least $2^\ell/2$. Therefore we have

$$2^\ell/2 \leq \lceil 2^\ell / w \rceil \cdot d \leq (2\varepsilon \cdot 2^\ell + 1) \cdot d$$

which implies that $d \geq \Omega(\min(\varepsilon^{-1}, 2^\ell))$. □

Since in the INW construction we will take ε-recycling functions with input length up to $\log n \cdot \log d$, in the above statement 2^ℓ will dominate d and thus we get $d = \Omega(\varepsilon^{-1})$. That means the seed length is $O(\log n \cdot \log(1/\varepsilon)) = O(\log n \cdot \log(nw))$, which is as good as Nisan's PRG.

1.2 Find ε-recycling functions

One way to construction an ε-recycling function, as we mentioned, is by pairwise independent hash functions. In fact, even random constant functions are 0-cycling since we only used the fact that the marginal distribution of $h(s)$ for each s is uniform. However, we cannot use them in the INW generator, because they require $d \geq 2^\ell$ and thus in every step of the recursion we need to take a larger d, which grows double-exponentially. Instead, we want a family of ε-recycling function where d is a constant.

To understand more about the requirement of ε-recycling, let us examine an arbitrary distinguisher function $A\colon [w] \times \{0,1\}^\ell \to 0, 1$. We can write

$$A(F(s), s') = \sum_{v \in [w]} \mathbb{1}_{F(s)=v}(s) \cdot \mathbb{1}_{A(v,s')=1}(s').$$

Here each term in the sum is a *combinatorial rectangle*, i.e. a function on s and s' of form $\mathbb{1}_S(s) \cdot \mathbb{1}_{S'}(s')$, which is a product of two indicator functions on s and s' respectively. If H fools all combinatorial rectangles, then H surely fools all distinguishers A that are sums of w combinatorial rectangles and is thus ε-recycling.

Definition 2 (ε-mixing)**.** *A function $H\colon \{0,1\}^\ell \times [d] \to \{0,1\}^\ell$ is ε-mixing if $(s, H(s,r))$ (with uniformly random $s \sim \{0,1\}^\ell, r \sim [d]$) ε-fools every combinatorial rectangle.*

Lemma 3. *ε-mixing implies ε-recycling.*

Hence our next goal is to find ε-mixing functions such that d is small.

2 Expander Graphs

The property of ε-mixing has a nice graph-theoretic interpretation. Recall that a d-regular graph is a graph where each vertex has d neighbors. We think of H as a d-regular graph over vertex set $\{0,1\}^\ell$, and $H(s,r)$ denotes the r-th neighbor of s.

Equivalent as Definition 2, a d-regular graph H is ϵ-mixing if it is indistinguishable from a complete graph (including self loops) when considering any subset of vertices.

Definition 3 (ε-mixing)**.** *A d-regular graph $H=(V,E)$ is* ε-mixing *if for all subsets $S, S' \subseteq V$,*

$$\left| \frac{e(S,S')}{nd} - \frac{|S|\left|S'\right|}{n^2} \right| \leq \varepsilon$$

where $|V| = n$ and $e(S,S')$ counts the number of directed edges between set S and S', i.e. $e(S,S') = \left|\{(i,j) \in E \mid i \in S, j \in S'\}\right|$.

For a proof of the equivalence, notice that the expectations of $\mathbb{1}_S(s) \cdot \mathbb{1}_{S'}(s')$ where (s,s') is uniformly drawn from the edges of H or the complete graph are $\frac{e(S,S')}{nd}$ and $\frac{|S|\left|S'\right|}{n^2}$ respectively. In other words, we are comparing the fraction of edges from S to S' in both graphs.

As a special case of Definition 3, take S' to be $\overline{S} = V \setminus S$. The ε-mixing property means that the number of edges going from S to $\overline{S}$ cannot be too small; it has to be at least proportional to the size of S. In fact, we have a specific notion as follows.

Definition 4 (α-edge expanding)**.** *A d-regular graph $H=(V,E)$ is* α-edge expanding *if for all $S \subseteq V$, $|S| \leq n/2$,*

$$e(S,\overline{S}) \geq \alpha \cdot d \cdot |S|\,.$$

Remark. ε-mixing is close, but not exactly enough to imply edge expanding, because we could only get

$$e(S,\overline{S}) \geq \frac{d}{n}|S||\overline{S}| - \varepsilon nd \geq \frac{1}{2}d|S| - \varepsilon nd.$$

Another closely related notion is vertex expansion, regarding the number of neighbors for a subset of vertices.

Definition 5 (α-vertex expanding)**.** *A d-regular graph $H=(V,E)$ is* α-vertex expanding *if for all $S \subseteq V, |S| \leq n/2$,*

$$|N(S) \setminus S| \geq \alpha \cdot |S|$$

where $N(S)$ denotes the neighboring vertices of S.

The *expander graphs*, which are graphs with either of the three properties above, are given the name because of the fact that no subset of vertices is badly connected with the rest of graph, and thus the set of reachable vertices is always expanding while taking more and more steps.

Remark. α-edge expanding implies α-vertex expanding, because each vertex in $N(S) \setminus S$ provides at most d edges for $e(S, \overline{S})$. However, the best edge expansion is $1/2$ while the best vertex expansion is 1 (both can be seen by taking a random subset of vertices S with $|S| = n$).

How do we construct expander graphs? We will learn the actual explicit construction in later lectures, but for now let us explore some random construction.

Theorem 4. *Let $G(n, d/n)$ be an Erdös-Rényi graph with n vertices and each edge is sampled independently with probability d/n. Then, $G(n, d/n)$ is ε-mixing with probability $1 - 2^{2n} \cdot 2e^{-\frac{1}{6}\varepsilon^2 dn}$.*

Note that the Erdös-Rényi graph will almost surely not be d-regular, and only the expected degree for each vertex is d.

Lemma 5 (Multiplicative Chernoff Bound)**.** *Let $X = \sum_{i=1}^{n} X_i$, where $X_i \in [0, 1]$ are independent. Let $\mu = \mathbb{E}[X]$. Then, for $\varepsilon > 0$,*

$$\Pr\left[|X - \mu| \geq \epsilon\right] \leq 2e^{-\frac{1}{3}\varepsilon^2 \mu^{-1}}.$$

Proof of Theorem 4. Notice that $e(S, S')$ in $G(n, d/n)$ is a sum of at most $|S||S'|$ independent random variables (the edges between vertices in $S \cap S'$ will be counted twice). Thus

$$\begin{aligned}
&\Pr\left[\left|e(S, S') - \frac{d}{n} \cdot |S| \left|S'\right|\right| \geq \epsilon dn\right] \\
&\leq 2\exp\left(-\frac{1}{6} \cdot \frac{(\epsilon dn)^2}{\frac{d}{n}|S|\left|S'\right|}\right) && \text{(Each edge contribute at most 2 to } e(S, S'), \text{ Lemma 5)} \\
&\leq 2\exp\left(-\frac{1}{6} \cdot \epsilon^2 dn\right) && (|S|, \left|S'\right| < n)
\end{aligned}$$

Then,

$$\begin{aligned}
\Pr\left[G(n, d/n) \text{ is not } \epsilon\text{-mixing}\right] &= \Pr\left[\forall S, S' \subseteq V,\ \left|e(S, S') - \mathbb{E}[e(S, S')]\right| \geq \epsilon dn\right] \\
&\leq \sum_{i=0}^{n} \binom{n}{i} \cdot \sum_{j=0}^{n} \binom{n}{j} \cdot 2\exp\left(-\frac{1}{6} \cdot \epsilon^2 dn\right) && \text{(Union bound)} \\
&= 2^{2n} \cdot 2\exp\left(-\frac{1}{6} \cdot \epsilon^2 dn\right) && \square
\end{aligned}$$

This means that $G(n, d/n)$ is ε-mixing with high probability when $d = O(1/\varepsilon^2)$. However, to show edge and vertex expansion, we need an actual model for random regular graph. A uniformly random regular simple graph is hard to sample, and there are two common ways to sample d-regular graphs which are not guaranteed to be simple.

Example (Permutation Model). Let d be even. Draw $d/2$ permutations $\sigma_1, \sigma_2, \ldots, \sigma_{d/2}$ from permutation group S_n. Let the edge set $E = \{(i, \sigma_j(i)), (\sigma_j(i), i) \mid i \in V, j = 1, 2, \ldots, d/2\}$.

Example (Matching Model). Let n be even. Draw d perfect matchings $M_1, M_2, \ldots, M_d$ on V, and let the edge set E be the union of these matchings.

It can be shown that either random graph will have close to optimal edge and vertex expansion with high probability. We are not going to do that here, and instead in the next lecture we will introduce yet another notion of graph expansion which subsumes all the three above, and construct expander graphs using the new notion.