Lecture 13 - Double Samplers

Today we will define & explore double samplers, which are related to ADX but possibly more flexible.

To start, let us talk about samplers.

**Def:** (Algorithmic definition)

A **sampler** is an algorithm that is given an accuracy parameter \( E \), and an error parameter \( \delta \), and "oracle access" to a function \( f: [N] \rightarrow \{0,1\} \). The sampler probes some \( n \) points and outputs \( \frac{1}{n} \sum_{i} f(x_i) \).

The sampler outputs, with probability at least \( 1 - \delta \), a value that is at most \( E \) away from \( E \) for \( f(x) \).

\[
\Pr \left[ \left| \frac{1}{n} \sum_{i} f(x_i) - E f(x) \right| < \varepsilon \right] > 1 - \delta
\]

**Graph definition:** \( V \)

\[
\begin{array}{c}
\vdots \\
\xrightarrow{2} x_2 \\
\xrightarrow{3} x_3 \\
\vdots \\
\end{array}
\]

A **s**amper is a bipartite graph \((V_s, V_t)\) s.t. each vertex in \( V_s \) has degree \( d \), \( V_s = [N] \) and such that \( \forall f: [N] \rightarrow \{0,1\} \),

\[
\Pr \left[ \left| \frac{1}{n} \sum_{i} f(x_i) - \frac{1}{d} \sum_{i} f(x_i) \right| < \varepsilon \right] > 1 - \delta
\]

Recall \( \chi(b) = \sup_{f: \|f\|_2 \leq 1} \left| \frac{1}{n} \sum_{i} f(x_i) - \frac{1}{d} \sum_{i} f(x_i) \right| \)

**Lemma:** Let \( G \) be a bipartite graph \((V_s, V_t, E)\), right \( \alpha \)-regular.

\( G \) is \( (\varepsilon, \delta) \)-**sampler**

Then \( G \) contains a \( \frac{\varepsilon}{\delta} \)-**expander**

**Proof:** (only items 1)

Let \( f: A \rightarrow \{0,1\} \) \( \ell = \frac{\varepsilon}{\delta} \),

let \( T = \{ x \in B \mid \| f(x) - \frac{1}{d} \delta \|_2 > \varepsilon \} \)

\( M_{\frac{\varepsilon}{\delta}} = M' \), \( \| M' \|_2 \leq \varepsilon \), \( \| f \|_2 \leq \frac{\varepsilon}{\delta} f \). Use \( \| f \|_2 \leq \varepsilon \), \( \| f \|_2 \leq \frac{\varepsilon}{\delta} f \), \( \| f \|_2 \leq \frac{\varepsilon}{\delta} f \).

By Markov's inequality, \( \Pr \left[ \| M' \|_2 > \varepsilon \right] < \frac{\varepsilon}{\delta} f < \frac{\varepsilon}{\delta} f \)
Lemma: Samplers are good for distance amplification.

\[ \text{ABNNR construction} \]

(Spectral) - Double Samplers

\[ V_0 \]
\[ V_1 \]
\[ V_2 \]

**Def:** A 3-partite graph \( V_0, V_1, V_2 \) is a \( \delta \) double sampler if

1. \( \chi(G(V_0, V_1)) \leq \gamma \)
2. \( \chi(G(V_1, V_2)) \leq \gamma \)
3. \( \forall T \in V_2 \) let \( V_T^1, V_T^0 \) be the neighbors of \( T \) and let \( G(V_T^1, V_T^0) \) be the graph induced on these sets. Then \( \forall T \chi(G(V_T^1, V_T^0)) \leq \gamma \).

**Example:** Taking \( V_0 = [n], V_1 = \binom{n}{k}, V_2 = \binom{n}{k} \)
we get a double sampler with \( \lambda^2 \leq \max\left( \frac{1}{t_n^2}, \frac{1}{t_2^2}, \frac{1}{t_3^2}, \frac{1}{t_2t_3} \right) + o(1) \)

**Example 2:** Take \( V_0 = \text{lines in } \mathbb{R}^i \) in a vector space \( V_i = \mathbb{R}^i \) \( V_2 = \mathbb{R}^i \)

However, \( |V_4| \gg |V_0|, |V_6| \gg |V_0| \).

Recall, \( \|L^+ - UD\| < Y \) for all dimensions.

**Theorem:** Let \( X \) be a \( d \)-dimensional \( Y \) \( HDX \) then satisfy

\[
V_0 = x(e) \\
V_i = x(k) \\
V_2 = x(d)
\]

and connect two faces \( s, t \) if set

we get a double sampler with \( \lambda^2 \leq \max\left( \frac{1}{t_n^2}, \frac{1}{t_2^2}, \frac{1}{t_3^2}, \frac{1}{t_2t_3} \right) + o(1) \)

(clearly, for any \( \lambda \) one can choose \( t_3 \) large enough and \( Y \) small enough so that \( \text{RHS} < \lambda \))

and: \( |V_4| = o(|V|) \) \( |V_6| = o(|V|) \).

Recall: \( X(e) \xrightarrow{X(i)} \cdots \xrightarrow{X(d)} \)

up, down operators

**Key lemma:** \( \lambda^2 \left( \mathbb{E}(x(i), x(i+1)) \right) = \lambda \left( D_{i+1} U_i \right) \leq \left( \frac{i+1}{i+2} \right)^2 + o(1) \)
**Corollary:** \( \lambda (G(\chi(k), \chi(d))) \leq \frac{k+1}{d+1} + O\left(\frac{1}{d}\right) \)

Recall, in a bipartite graph \((A, B, E)\)

\[ \lambda(T^*) = \lambda(T^*) \leq \left(\lambda(T)\right)^2 \]

where \( \lambda(T) = \max \{ \|Tf\| \}_{f \perp T} \)

\[ f \perp T \iff \|Tf\| = 1 \]

\[ (so \ A \ f \perp T \|Tf\| \leq \lambda(T) \|f\|) \]

**Proof of corr:** \( \lambda(U_{d-1} \cdots U_k) \leq \lambda(U_{d-1}) \cdots \lambda(U_k) \)

(assume first that \( \chi = 0 \))

\[ \frac{1}{d+1} \lambda^{-1} \leq \frac{k+1}{d+1} + O\left(\frac{1}{d}\right) \]

We used the following fact (inductively on \(d-k\) layers)

If \( G \) is a 3 layer graph on vertices \( A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_t \)

Then \( \lambda(U_{d-1} \cdots U_2 U_1) \leq \prod_{i=1}^{t} \lambda(U_i) \)

**Proof:** If \( f \in \ell_2(A_t) \) \( f \perp T \) then \( Uf \perp T \)

and inductively for all \( i > 1 \) \( U_i \circ U_{i-1} \cdots U_1 \)

so \( \|U_{d-1} \cdots U_2(U_t)\| \leq \chi \cdot \lambda_2 \cdot \lambda_{d-1} \cdot \|U_f\| \leq \chi \cdots \lambda_1 \cdot \|U_f\| \)

inductive hyp
Proof of key lemma:

We prove by induction on $i$ that $\lambda(U_i) = \chi(D_{i+1}U_i) \leq \frac{i+1}{i+2} + \delta(U_{i+1})$.

For $i = 0$: $U_0 : \ell_b(X_0) \to \ell_b(X_1)$ ($D_0 = U_0^*$).

$D_0U_0$ is the upper random walk on the vertices.

This walk is $\frac{1}{2}$-lazy, and $D_0U_0 = \frac{1}{2}I + \frac{1}{2}L_0$.

Additionally, by definition $\|L_0 - U_0 D_0\| \leq \gamma$.

So for any $f \in \ell_b(X_0)$ at $f \perp 1$:

$$\|L_0 f\| = \| (L_0 - U_0 D_0) f\| \leq \gamma \|f\|$$

Therefore:

$$\|D_0U_0 f\| = \| \frac{1}{2}f + \frac{1}{2}L_0 f\| \leq \frac{\gamma}{2} \|f\| + \frac{\gamma}{2} \|f\|$$

$$\Rightarrow \lambda(D_0U_0) \leq \frac{1}{2} + \gamma$$
Assume for \( i \), prove for \( i + 1 \):

\[
U_i : e_i(X(i)) \rightarrow e_i(X(i+1)) \quad D_{i+1} = U_i^*
\]

\( D_{i+1} U_i \) is a lazy upper walk, with \( \frac{1}{i+2} \) probability to stay in place.

\[
D_{i+1} U_i = \frac{1}{i+2} \cdot \text{Id} + \frac{i+1}{i+2} \cdot L_i^+
\]

By assumption \( \| L_i^+ - U_i D_i \| \leq \gamma \).

So for \( f : X(i) \rightarrow \mathbb{R}, f \leq 1 \),

\[
\| L_i^+ f - U_i D_i f \| \leq \gamma \| f \| \quad \text{D inequality}
\]

\[
\Rightarrow \| L_i^+ f \| - \| U_i D_i f \| \leq \gamma \| f \|
\]

\[
\| L_i^+ f \| \leq \left( \gamma + \frac{i+1}{i+2} \cdot \gamma \right) \| f \|
\]

\[
\| D_{i+1} U_i f \| \leq \frac{1}{i+2} + \frac{i+1}{i+2} \cdot \left( \gamma + \frac{i}{i+1} \right)
\]

\[
= \frac{i}{i+2} + \frac{i}{i+2} + \gamma \cdot (i+1) = \frac{i+1}{i+2} + \gamma (i+1)
\]
Double samplers exist, with $O(n)$ vertices and $\lambda \to 0$.

Questions:
- Lower bounds on $C_n$ as $\lambda \to 0$
- Constructions with full regularity.

---

**Summary**

- HDX, local link def
  - RW def \( \| L^+ - UD \| \leq \gamma \)
  - Spectrum of \( X(i) - X(j) \) operators & Double Samplers.
- Expander graphs — ZigZag construction
- Cayley graphs & groups — Characters
- Error correcting codes, $\varepsilon$-biased sets

Constructions of HDX:
- Flags Complex
- Salil's construction
- $K_0$ construction
Testing: BLR lin. testing
Coboundary Expansion & Top overlap
Agreement testing

Harmonic Analysis: on Boolean Cube & Hix