CS229r - LECTURE 14 3/9/2017

TODAY: GRAPH-THEORETIC CODES - II
(GRAPHICALLY "GENERATED" CODES)

[Note Spielman's Lin. Time Encodable codes already fit this notion... but we'll do different things today]

Three main references:

1. ALON-BRUCK-NAOR-NAOR-ROTH '90ish
2. ALON-LUBY - 95ish
3. GURUSWAMI - INDYK - 2002 ish

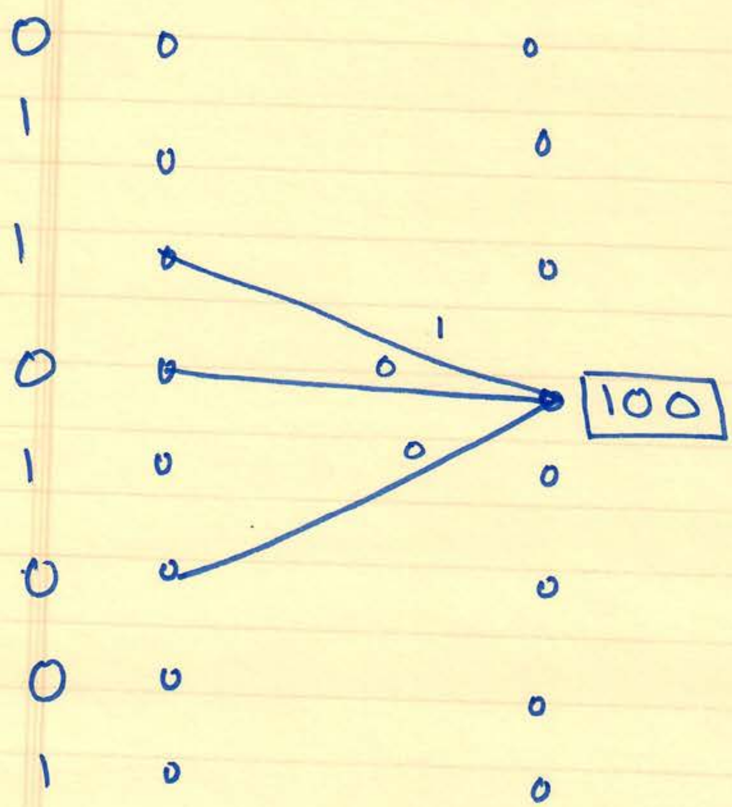
1 + 2 : EXPLICIT CODES; GOOD DIST. vs. RATE.

3 : Efficient algorithms also....

ABNNR Construction

Goal: Code over large alphabet of ~~high~~ positive rate & high distance.

Construction (Part II): ① Start with Bip. regular expander. (d,d) on n-vertices each.



② Message = Assignment to left vertices

③ Encoding: Move bits to right vertex & concatenate all bits, ~~on~~ on edges to get symbol on right.

③

message $\in \Sigma^n$ encoding $\in (\underbrace{\Sigma^d}_{})^n$

new - alphabet.

- Rate : ~~a~~ ~~b~~ $= \frac{1}{d}$.

- Distance = ? ... Actually terrible; at best d.

Construction Part I: ⓪ use some decent code C_0 to ~~map~~ map message in Σ^k to word in Σ^n.

<u>Proposition:</u>

if $\delta(C_0) = \delta$ & ~~code~~ Graph is ~~S~~ (α, δ) - expander

then $\delta(C_{final}) \geq \alpha \cdot d \cdot \delta$

<u>Proof:</u> Obvious!

- So how large can $\alpha \cdot d \cdot \delta$ be?

- Answer $\sim 1 - \frac{1}{d}$

Better to look at expansion from right;

sets of size $\frac{1}{d}$ expand to $(1-\delta)$ fraction on ~~right~~ left.

(if $\alpha \to 1$).

Conclusion: ABNNR achieve distance $1 - \frac{1}{d}$

with rate $\Omega\left(\frac{1}{d}\right)$

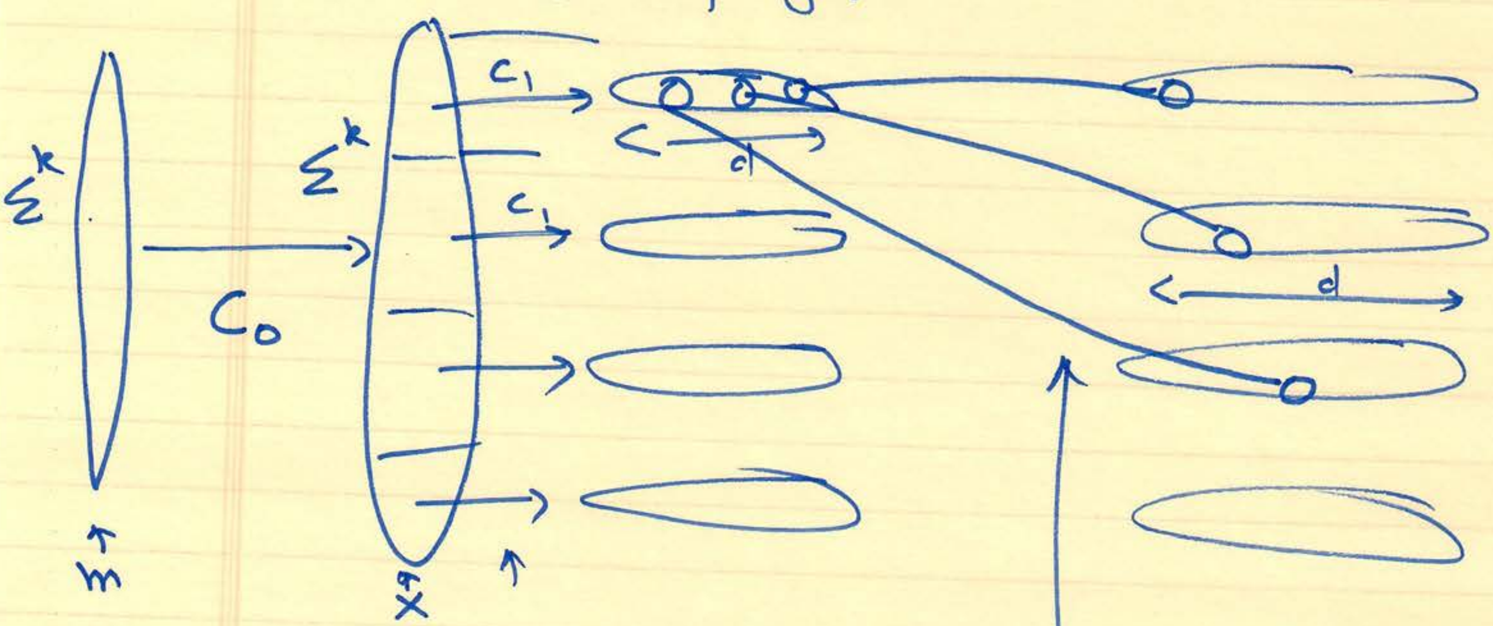
" Near Singleton over large " but constant alphabet

Can we get rate $> \frac{1}{2}$? ALon-Luby !!

---

3 ingredients: C_0 - big code ; C_1 - small code of Rate R_1 dist. δ.

B. bip. graph.



many applications of C_1 on disjoint blocks of x

move symbols of Σ along edges.

<u>Parameters</u> - message space Σ^k.

$C_0 : \cancel{\text{Rate}}\ \Sigma^k \to \Sigma^n$ of rate $1-\epsilon$ & distance $\Omega(\epsilon)$

$C_1 : \Sigma^\ell \to \Sigma^d \to$ code of rate $R = \frac{\ell}{d}$ & dist δ.

B : (α, δ_1) - expander (d,d)- regular. on n/ℓ vertices

won't suffice.

Output code $C_{final} : \Sigma^k \to (\Sigma^d)^{n/\ell}$

$$\text{Rate} = \frac{k\ell}{dn} = \frac{\ell}{d} \cdot \frac{k}{n} = (1-\epsilon) R.$$

Distance $\stackrel{?}{=}$ $\cancel{\alpha}$ $\delta_1 = \Omega(\epsilon)$ = distance of C_0

$\cancel{\Rightarrow \delta(C_{final}) = \alpha \cdot d \cdot \delta_1}$ Expansion Insufficient.

Lets assume B random

& see what we need.

- Let $S \subseteq$ Left be vertices that are non-zero
- Let $T \subseteq$ Right be non-zero vertices on right.

1. If B is random then for typical $i \in$ Left

$$|\Gamma(i) \cap T| \approx \frac{|T|}{|\text{Right}|} \cdot d$$

2. For $i \in S$, at least $\delta \cdot d$ coordinates non-zero

$$\Rightarrow \ |\Gamma(i) \cap T| \geq \delta d$$

so if $\dfrac{|T|}{|\text{Right}|} < \delta \ \Rightarrow$ S is atypical

$\Rightarrow$ $|S|$ is small.

(δ,ϵ)-Sampler

Def 1: $\Gamma_\delta(T) \triangleq \{ i \in \text{Left} \mid |\Gamma(i) \cap T| \geq \delta \cdot d \}$

Def 2: B is (δ,ϵ)-sampler

if $\forall\, T \subseteq \text{Right}$, $|T| \leq (\delta-\epsilon)|\text{Right}|$

$$|\Gamma_\delta(T)| < \epsilon \cdot |\text{Left}|.$$

Theorem: if ① C_0 is codes of Rate $1-O(\epsilon)$ & Dist. ϵ,

② C_1 is of Rate R & dist. δ - length d.

③ B is (δ,ϵ)-sampler ; d-regular

then C_f is code of Rate $(1-O(\epsilon)) \cdot R$ & Distance $\delta-\epsilon$.

over alphabet Σ^d

long!

constant alphabet

Near Singleton!

Algorithms

Flavor:

Assume C_0 is XYZ-Decodable & B is ( )-sampler.

then C_f is X'Y'Z'-Decodable.

Example: C_0 is linear-time decodable with $\Omega(\epsilon)$ errors.

& B is $\left(\frac{\delta}{2}, \epsilon'\right)$-sampler

$\Rightarrow C_f$ is $\left(\frac{\delta}{2} - \epsilon'\right)$-error decodable in lin. time.

Proof: Obvious

Example: C_0 is list-recoverable from $k + \epsilon n$-agreement

& B is (δ, ϵ')-sampler

$\Rightarrow C_f$ is $(\delta - \epsilon')$-list decodable in poly time

[Guruswami - Rudra].

<u>Example</u>: C_0 is lin. time list-decodable from $\Omega(\epsilon)$ error

& B is (δ, ϵ')-sampler & C_1 is $(1-\epsilon')$-list decodable

$\Rightarrow$ C_δ is $\cancel{(\delta, \epsilon')}$ $(1-\epsilon'')$-list-decodable in lin. time.