Today: Graph-Theoretic Codes - II
(Graphically "Generates" Codes)

[Note Spiegelman'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 - 2002ish

1 + 2: Explicit Codes; Good Dist. Rate
2: Efficient algorithms also....
ABNNR Construction

Goal: Code over large alphabet of 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 edges to get symbol on right.
message ∈ Σ^n, encoding ∈ (Σ_{new})^n

- Rate: \frac{\alpha \cdot \delta}{d}

- Distance = ? … Actually terrible; at best \delta.

Construction Part I: ○ use some code \mathcal{C}_0 to map
message in Σ^k to word in Σ^n.

Proposition:
if \delta(\mathcal{C}_0) = \delta \& code is \\mathcal{G}_{\alpha,\delta}(\alpha,\delta) - expander
then \delta(\mathcal{C}_{\text{final}}) ≥ \alpha \cdot \delta \cdot \delta \cdot \delta

Proof: Obvious!

- So how large can \alpha \cdot \delta \cdot \delta be?
- Answer: \sim 1 - \frac{1}{\delta}

Better to look at expansion from right;
sets of size \frac{1}{\delta} expand to (1-\delta) fraction on
(if \delta \gg 1).
Conclusion: ABNNR achieve distance $1 - \frac{1}{d}$ with rate $\mathcal{O}(\frac{1}{d})$ but constant near singleton over large 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. $S$.

B. bip. graph.

Many application of $C$ on disjoint blocks $A \times \alpha$ move symbols $\alpha \leq \epsilon$ along edges.
Parameters: message space $\sum^k$

$C_0: \sum^k \rightarrow \sum^n$ of rate $1-\varepsilon$ and distance $\Omega(\varepsilon)$

$C_1: \sum^n \rightarrow \sum^d$ code of rate $R = \frac{d}{n}$

$B: (\alpha, \delta) - \text{expander}$ (d, d) - regular won't suffice.

Output code $C_{\text{final}}: \sum^k \rightarrow (\sum^d)^{n/2}$

Rate $= \frac{k \cdot e}{dn} = \frac{e}{d} \cdot \frac{k}{n} = (1-\varepsilon) R$

Distance $\delta = \delta = \Omega(\varepsilon) = \text{distance of } C_0$

Expansion insufficient. Let's assume $B$ random

& see what we need.
- Let $S \subseteq \text{Left}$ be vertices that are non-zero.
- Let $T \subseteq \text{Right}$ be non-zero vertices on right.

1. If $B$ is random then for typical $i \in \text{Left}$
   \[ |\Pi(i) \cap T| \approx \frac{|T|}{|\text{Right}|} \cdot \rho \]

2. For $i \in S$, at least $8 \cdot d$ coordinates non-zero
   \[ \Rightarrow |\Pi(i) \cap T| \geq 8d \]

   so if \[ \frac{|T|}{|\text{Right}|} < 8 \Rightarrow S \text{ is atypical} \]

   \[ \Rightarrow |S| \text{ is small.} \]
(ε, δ)-Sampler

Def 1: \[ \Gamma_\delta(T) = \sum_{i \in \text{left} \mid |\Gamma(i) \cap T| > \delta d^2} \]

Def 2: B is (ε, δ) - sampler

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

\[ |\Gamma_\delta(T)| < \varepsilon |\text{left}|. \]

Theorem: if
1. \( C_0 \) is code of rate \( 1 - O(\varepsilon) \)
   & dist. \( \varepsilon \),
2. \( C_1 \) is of rate \( R \) & dist. \( \delta \)
   - length \( d \).
3. \( B \) is (\( \varepsilon, \varepsilon \)) - sampler; \( d \)-regular

then \( C_f \) is code of rate \( (1 - O(\varepsilon)) \cdot R \)
   & dist. \( \varepsilon - \delta \).

over alphabet \( \Delta \)

constant alphabet \( \Delta \)

near singleton!
Algorithms

Flavor:
Assume $C_0$ is $XYZ$-decodable and $B$ is $(\cdot)$-sampler.

Then $C_f$ is $X'Y'Z'$-decodable.

Example: $C_0$ is linear-time decodable with $\varepsilon_2(\varepsilon)$ errors.

$B$ is $(\frac{\varepsilon}{2}, \varepsilon')$-sampler

$\Rightarrow C_f$ is $(\frac{\varepsilon}{2} - \varepsilon')$-error decodable in lin. time.

Proof: Obvious

Example: $C_0$ is list-recoverable from $R+en$-agreement

$\Leftrightarrow B$ is $(\varepsilon, \varepsilon')$-sampler

$\Rightarrow C_f$ is $(\varepsilon - \varepsilon')$-list decodable in poly. time

[Chung, Ran Raz]


Example: $C_0$ is $\mathcal{L}$ lin. tim. list-decodable from $\mathcal{R}(\epsilon)$ Error

$\mathcal{B}$ is $(\epsilon, \epsilon')$-decoder & $C_1$ is $(1-\epsilon')$-list decodable

$\Rightarrow C_5$ is $(\epsilon_5, \epsilon'')$-list-decodable in lin. tim.