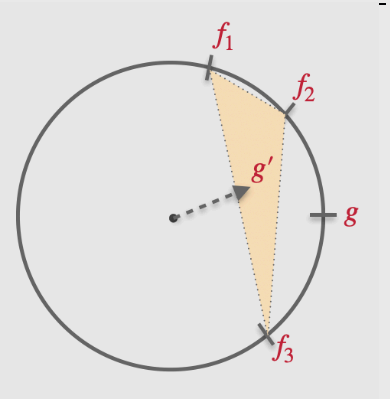
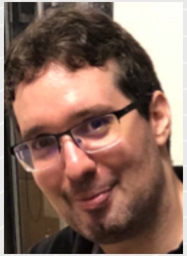
List Decoding of Tanner and Expander Amplified Codes from Distance Certificates



Shashank Srivastava
TTIC



Fernando Granha Jeronimo
(UC Berkeley)

Madhur Tulsiani
(TTIC)

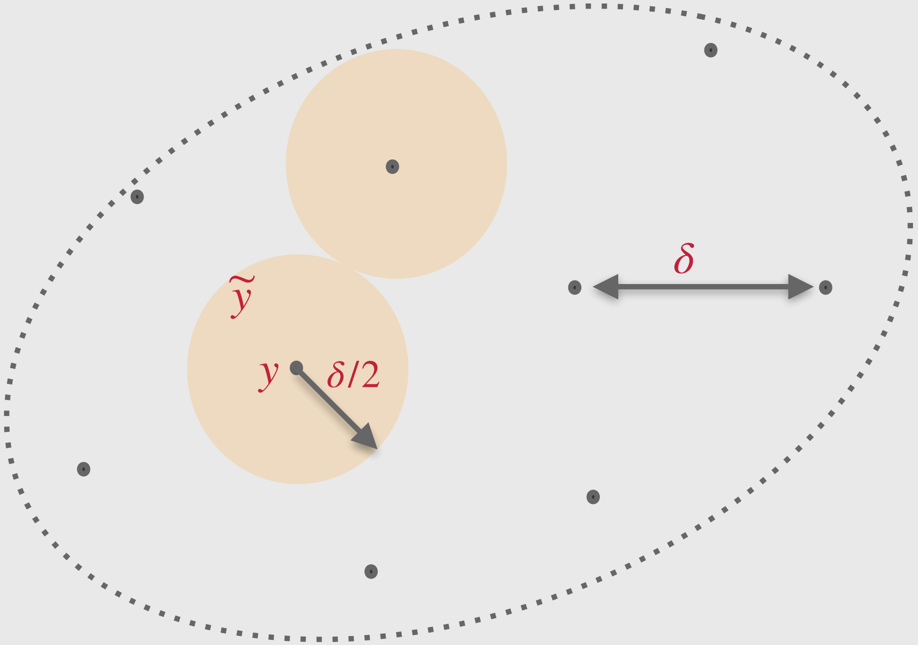
Linear Codes

Code $C \subseteq \{0,1\}^n$

$x \in \{0,1\}^k$ —Encoding→ $y \in C$ → **Noisy Channel** → $\tilde{y} \in \{0,1\}^n$ —Decoding→ y

y: codeword
C: code

- C is linear if it is a subspace of $\mathbb{F}_2^n$.
- $\delta(C) = \min_{y_1 \neq y_2 \in C} \Delta(y_1, y_2)$.
- Rate of C is $\frac{k}{n}$.



List Decoding

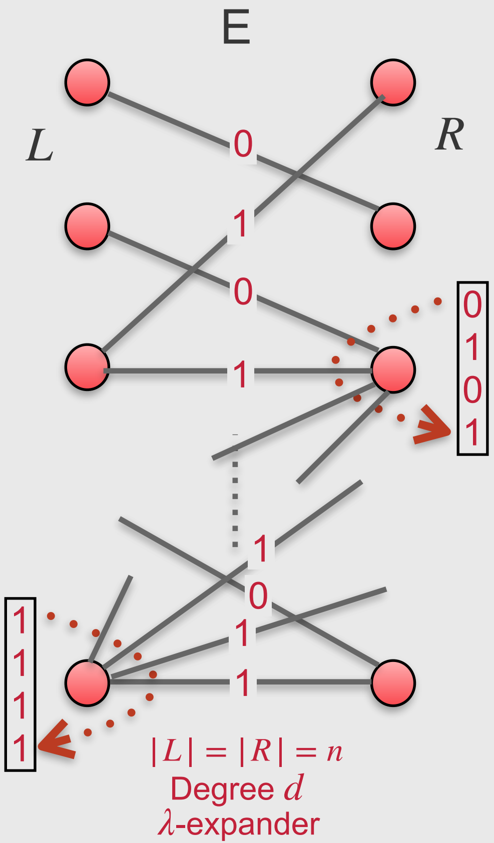
- What happens when number of errors exceeds $\delta/2$?
- Hope: Number of codewords is polynomial, if not 1.
- Johnson bound: Upto $J(\delta)$, list size is bounded.

$$\delta/2 < J(\delta) < \delta$$

- Algorithmic task: find the list.

Tanner Codes

[Tanner'81, Sipser-Spielman'96, Zémor'01]



- Codewords: $\{0,1\}$ assignment to edges.
- Every local view belongs to an inner code C_0.

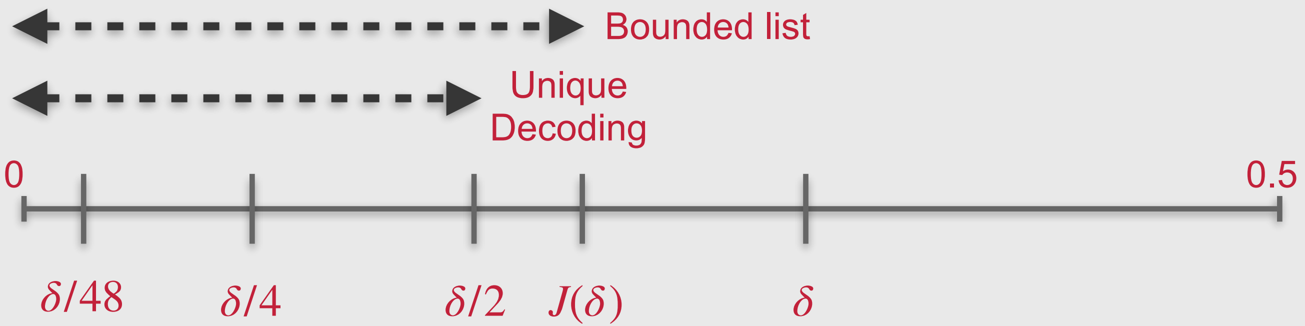
Low-Density Parity Check (LDPC)

Linear-time decoders

Thm (Sipser-Spielman'96): Distance of Tanner code is at least $\delta = \delta_0 (\delta_0 - \lambda)$.

Decoding Tanner Codes

- Sipser-Spielman'96: $\approx \delta/48$
- Zémor'01: $\approx \delta/4$
- Skachek-Roth'03: $\approx \delta/2$



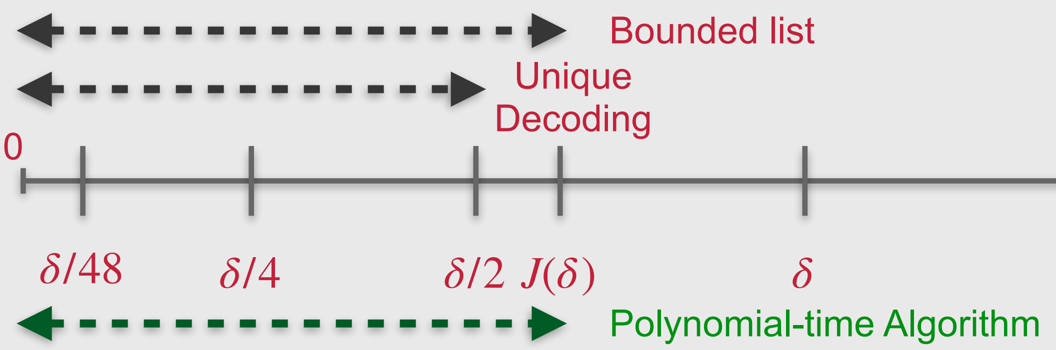
Our Result - list-decoding up to $J(\delta)$.

Main Theorem

Inner Code: C_0 – distance δ_0

Graph: $G - \lambda$-expander

Theorem (Jeronimo-S-Tulsiani):
For any $\epsilon > 0$, the Tanner code C with distance at least $\delta = \delta_0(\delta_0 - \lambda)$ can be list-decoded from radius $J(\delta) - \epsilon$ in time $n^{O_d(1/\epsilon^4)}$.



Why care about list-decoding Tanner codes?

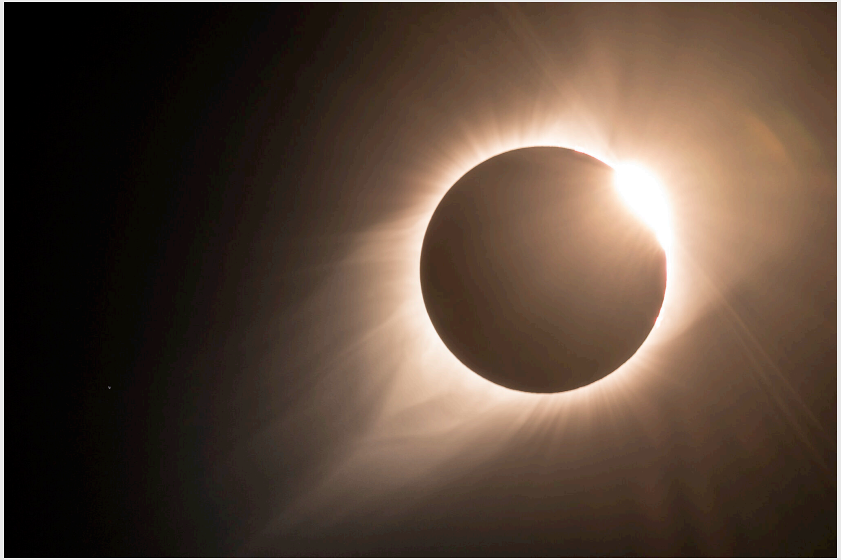
- Unique-decoding to list-decoding requires new ideas.
- Most list-decoding algorithms work for algebraic codes.
- Tanner codes: Source of *linear* time decoders.

Techniques

- Covering Lemma: Algorithm-friendly proof of Johnson bound.
- Proofs-to-Algorithms paradigm for codes.
  - Distance Proof = Local Properties + Spectral Expansion (For local-to-global) ⟹ List Decoding Algorithm upto Johnson bound
  - Used for decoding Ta-Shma code [Richelson-Roy’23]
- Rounding algorithms for convex optimization based decoders.

Covering Lemma

In about an hour, the moon will cover the sun.



Source - Getty Images

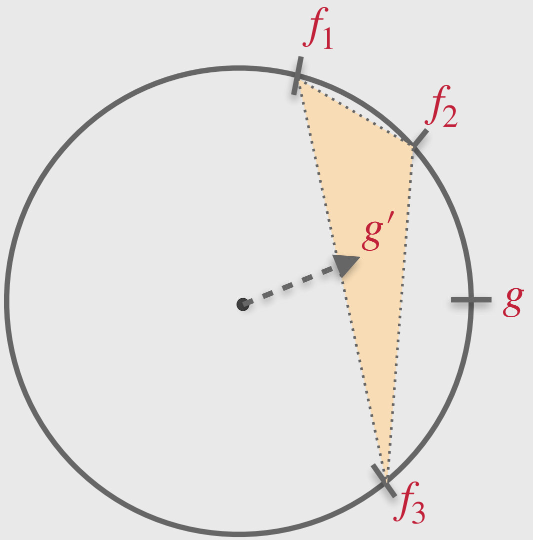
Covering Lemma

Lemma. Given a family $\mathscr{F}$ of unit vectors in $\mathbb{R}^n$, and a unit vector $g \in \mathbb{R}^n$, such that

$$\forall f \in \mathscr{F}, \quad \langle g, f \rangle > \alpha. \qquad \alpha \in (0,1)$$

There exists $g' \in conv(\mathscr{F})$ such that,

$$\forall f \in \mathscr{F}, \quad \langle g', f \rangle > \alpha^2.$$



Proof. g' is the smallest ℓ_2-norm vector in $conv(\mathscr{F})$.

From codes to geometry

Embed $f \in \mathbb{F}_2^n$ into $\mathbb{R}^n$ as $\chi(f)_i = (-1)^{f_i}$.

$$\Delta(f_1, f_2) = \frac{1 - \langle \chi(f_1), \chi(f_2) \rangle}{2}$$

$$\Delta(f_1, f_2) = \frac{1-\beta}{2} \iff \langle \chi(f_1), \chi(f_2) \rangle = \beta$$

- Hamming Distance $\leftrightarrow$ Inner product.
- Hamming Ball $\leftrightarrow$ Half-space.

Johnson Bound:

For $\delta = \dfrac{1-\beta}{2}$, list sizes are polynomial until $J(\delta) = \dfrac{1-\sqrt{\beta}}{2} \in \left(\dfrac{\delta}{2}, \delta\right)$.

Algorithm-friendly proof of Johnson bound

$$0 \to 1$$
$$1 \to -1$$

$$\delta = \frac{1-\beta}{2}$$
$$J(\delta) = \frac{1-\sqrt{\beta}}{2}$$

- For any $h \in \mathscr{L}\left(r, J(\delta)\right)$, it holds that $\langle \chi(r), \chi(h) \rangle > \sqrt{\beta}$.
- Covering Lemma $\Longrightarrow$ There is an $r' \in conv(\mathscr{L})$ such that for any $h \in \mathscr{L}\left(r, J(\delta)\right)$,

$$\langle r', \chi(h) \rangle > \beta.$$

- r' as a convex combination $\to$ distribution $\mathscr{D}$ over C.

$$\mathbb{E}_{f\sim\mathscr{D}}\left[\Delta(f,h)\right] < \delta$$

- Support of $\mathscr{D}$ contains $\mathscr{L}\left(r, J(\delta)\right)$.
- Pick $\mathscr{D}$ with support size $\leq n+1$.

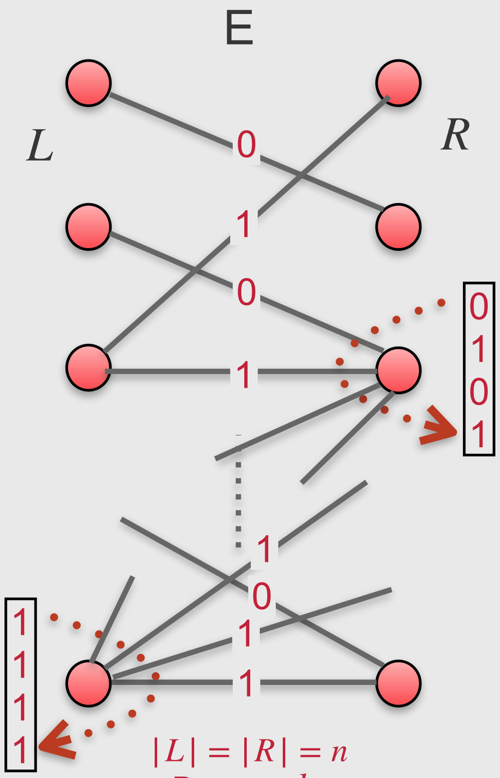
Can take exponential time!

Carathéodory's Theorem

Exponential Time Algorithm

1. Use covering lemma to find distribution $\mathscr{D}$ over C such that for every $h \in \mathscr{L}\left(r, J(\delta)\right)$,
$$\mathbb{E}_{f \sim \mathscr{D}}[\Delta(f, h)] < \delta.$$
2. Sample h' from $\mathscr{D}$.
3. Use distance of C to conclude
$$h' = h$$
with some probability.

Distance Proof of Tanner Code



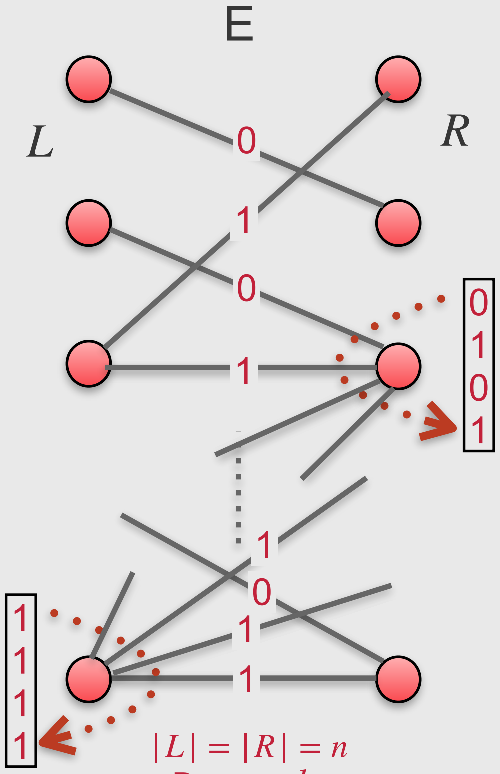
$|L| = |R| = n$
Degree d
λ-expander

Let $F \subseteq E$, $S \subseteq L$, $T \subseteq R$ be positions where $f, g \in \mathbb{F}_2^E$ differ.

Four distances:

1. $\Delta_E(f,g) = \dfrac{|F|}{nd}$
2. $\Delta_L(f,g) = \dfrac{|S|}{n}$
3. $\Delta_R(f,g) = \dfrac{|T|}{n}$
4. $\Delta_{LR}(f,g) = \sqrt{\Delta_L(f,g) \cdot \Delta_R(f,g)}$

Distance Proof of Tanner Code



$|L| = |R| = n$
Degree d
λ-expander

Four distances:

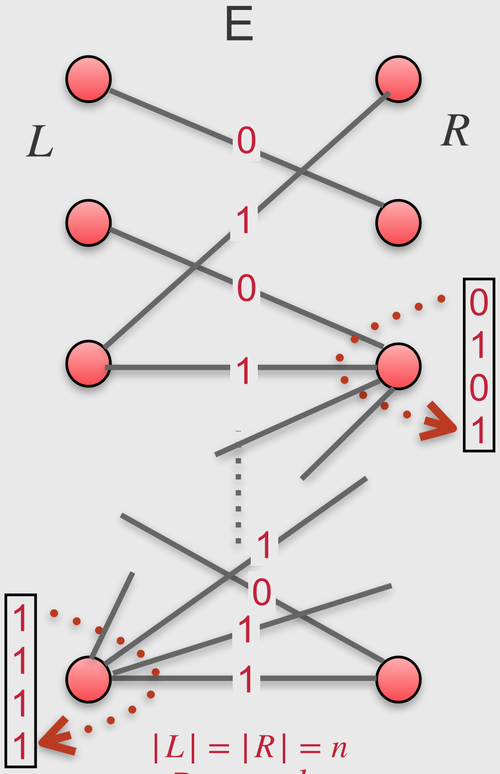
1. $\Delta_E(f, g) = \dfrac{|F|}{nd}$
2. $\Delta_L(f, g) = \dfrac{|S|}{n}$
3. $\Delta_R(f, g) = \dfrac{|T|}{n}$
4. $\Delta_{LR}(f, g) = \sqrt{\Delta_L(f, g) \cdot \Delta_R(f, g)}$

$$|F| \geq |S| \cdot \delta_0 d$$

$$\Delta_E(f, g) \geq \delta_0 \cdot \Delta_L(f, g)$$

$$\Delta_E(f, g) \geq \delta_0 \cdot \Delta_{LR}(f, g)$$

Distance Proof of Tanner Code



$|L| = |R| = n$
Degree d
λ-expander

Four distances:

1. $\Delta_E(f,g) = \dfrac{|F|}{nd}$
2. $\Delta_L(f,g) = \dfrac{|S|}{n}$
3. $\Delta_R(f,g) = \dfrac{|T|}{n}$
4. $\Delta_{LR}(f,g) = \sqrt{\Delta_L(f,g) \cdot \Delta_R(f,g)}$

$$\Delta_E(f,g) \geq \delta_0 \cdot \Delta_{LR}(f,g)$$

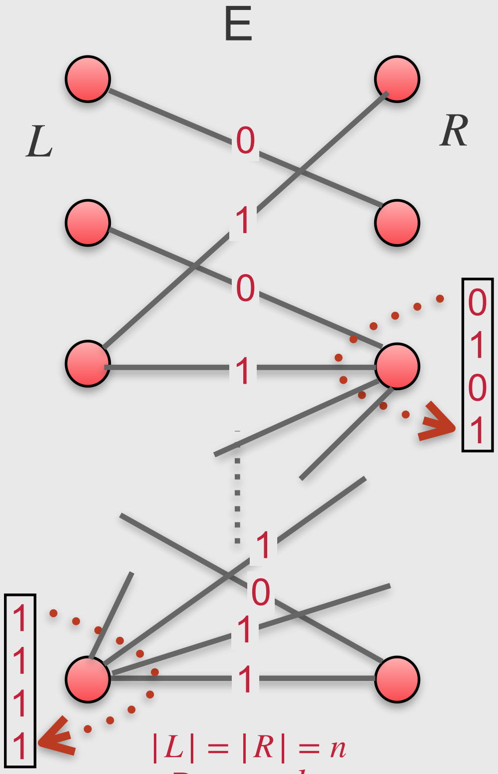
Expander Mixing Lemma

$$F \subseteq E(S,T)$$

$$|F| \leq |E(S,T)| \leq \frac{d}{n}|S| \cdot |T| + \lambda d \sqrt{|S| \cdot |T|}$$

$$\Delta_E(f,g) \leq \Delta_{LR}(f,g)^2 + \lambda \cdot \Delta_{LR}(f,g)$$

Distance Proof of Tanner Code



$|L| = |R| = n$
Degree d
λ-expander

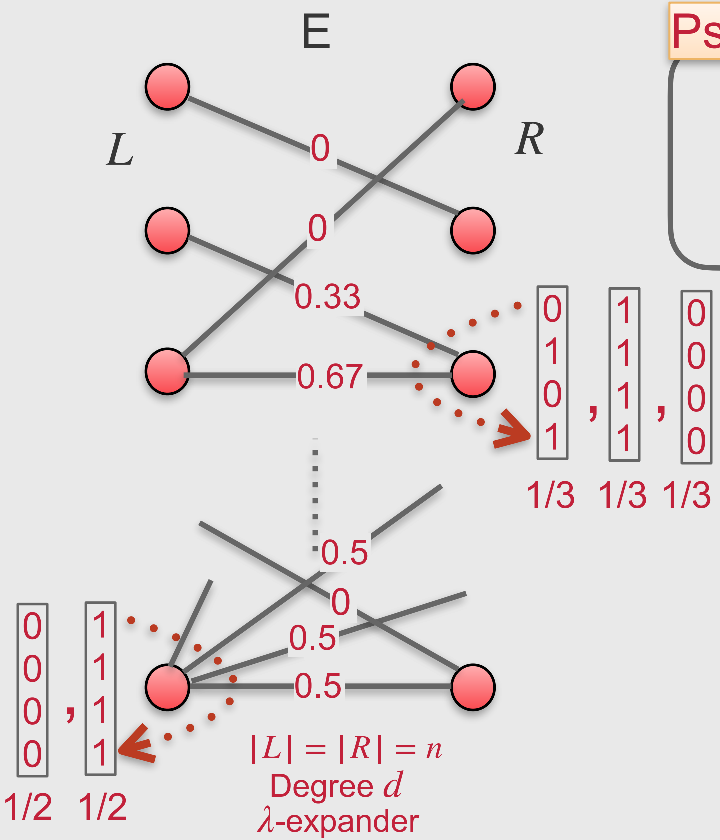
$$\delta_0 \cdot \Delta_{LR}(f,g) \le \Delta_E(f,g) \le \Delta_{LR}(f,g)^2 + \lambda \cdot \Delta_{LR}(f,g)$$

$$\Delta_{LR}(f,g)^2 - (\delta_0 - \lambda) \cdot \Delta_{LR}(f,g) \ge 0$$

$$\Delta_{LR}(f,g) = 0 \text{ or } \Delta_{LR}(f,g) \ge \delta_0 - \lambda$$

$$\implies \Delta_E(f,g) \ge \delta_0(\delta_0 - \lambda)$$

Continuous Relaxation for Tanner Code



Pseudocodeword

Ensemble of distributions

$$\widetilde{\mathscr{D}} = \{\mathscr{D}_\ell\}_{\ell \in L}, \{\mathscr{D}_r\}_{r \in R}$$

Consistency along edges

Used for LP Decoding

Strengthening based on Sum-of-Squares (SoS) “Pseudo-distributions”

Distance Proof for Relaxation of Tanner Code?

$$\mathbb{E}_e[\widetilde{\mathbb{E}}[\mathbf{1}_{f_e \neq 0}]] \geq \mathbb{E}_l[\widetilde{\mathbb{E}}[\delta_0 \cdot \mathbf{1}_{f_l \neq 0}]]$$

$$\Delta_E(\widetilde{\mathscr{D}}, 0) \geq \delta_0 \cdot \Delta_L(\widetilde{\mathscr{D}}, 0)$$

$$\Delta_E(\widetilde{\mathscr{D}}, 0) \geq \delta_0 \cdot \Delta_{LR}(\widetilde{\mathscr{D}}, 0)$$

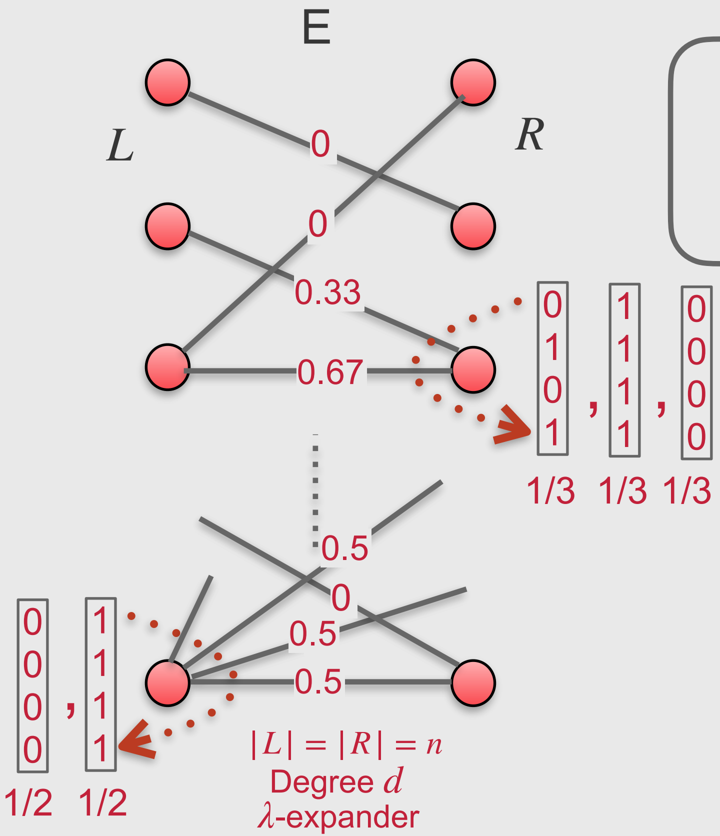
Distance Proof for Relaxation of Tanner Code?

$$\mathbb{E}_e[\widetilde{\mathbb{E}}[\mathbf{1}_{f_e \neq 0}]] \leq \mathbb{E}_{l\sim r}[\widetilde{\mathbb{E}}[\mathbf{1}_{f_l \neq 0} \cdot \mathbf{1}_{f_r \neq 0}]]$$

$$? \quad \leq \mathbb{E}_{l,r}[\widetilde{\mathbb{E}}[\mathbf{1}_{f_l \neq 0} \cdot \mathbf{1}_{f_r \neq 0}]]$$

$$? \quad \leq \mathbb{E}_{l,r}[\widetilde{\mathbb{E}}[\mathbf{1}_{f_l \neq 0}] \cdot \widetilde{\mathbb{E}}[\mathbf{1}_{f_r \neq 0}]]$$

Continuous Relaxation for Tanner Code



Ensemble of distributions

$$\widetilde{\mathscr{D}} = \{\mathscr{D}_\ell\}_{\ell \in L}, \{\mathscr{D}_r\}_{r \in R}$$

Consistency along edges

← Used for LP Decoding

Modifications:

- Enforce positive semidefinite-ness of (global) covariance matrix.
- $\{\mathscr{D}_\ell\}_{\ell \in L}, \{\mathscr{D}_r\}_{r \in R}$ induced by another ensemble of distributions over t-sized sets, for $t \gg d$.

Key steps in the proof

$$\mathbb{E}_{l\sim r}\left[\widetilde{\mathbb{E}}\left[X(f_l)\cdot Y(f_r)\right]\right]$$

$\Downarrow$ λ

Uses PSD-ness/non-negativity of sum-of-squares of polynomials

$$\approx \mathbb{E}_{l,r}\left[\widetilde{\mathbb{E}}\left[X(f_l)\cdot Y(f_r)\right]\right]$$

$\Downarrow$ η

Uses low average correlation obtained by random conditioning.

$$\approx \mathbb{E}_{l,r}\left[\widetilde{\mathbb{E}}\left[X(f_l)\right]\cdot \widetilde{\mathbb{E}}\left[Y(f_r)\right]\right]$$

Exponential Time Algorithm

1. Use covering lemma to find distribution $\mathscr{D}$ over C such that for every $h \in \mathscr{L}\left(r, J(\delta)\right)$,

$$\mathbb{E}_{f \sim \mathscr{D}}[\Delta(f, h)] < \delta$$

2. Sample h' from $\mathscr{D}$.
3. Use distance of C to conclude

$$h' = h.$$

Exponential Time Algorithm

1. Use covering lemma to find distribution $\mathscr{D}$ over C such that for every $h \in \mathscr{L}\left(r, J(\delta)\right)$,

$$\mathbb{E}_{f\sim\mathscr{D}}[\Delta(f,h)] < \delta$$

2. ~~Sample h' from $\mathscr{D}$~~.
   Condition $\mathscr{D}$ on all n coordinates to get h'.

3. Use ~~distance of C~~ $\Delta(h',h)\left(\Delta(h',h) - \delta\right) \geq 0$ to conclude

~~$h' = h$~~

$$\Delta(h',h) = 0.$$

Exponential Time Algorithm

Time 2^n

1. Use covering lemma to find distribution $\mathscr{D}$ over C such that for every $h \in \mathscr{L}\left(r, J(\delta)\right)$,
$$\mathbb{E}_{f\sim\mathscr{D}}[\Delta(f,h)] < \delta$$
2. ~~Sample h' from $\mathscr{D}$.~~
   Condition $\mathscr{D}$ on all n coordinates to get h'.
3. Use ~~distance of C~~ $\Delta(h',h)\big(\Delta(h',h)-\delta\big) \geq 0$ to conclude
$$\sout{h' = h}$$
$$\Delta(h',h) = 0.$$

Polynomial Time Algorithm

Time n^{1/η^2}

1. Use covering lemma to find pseudo-distribution $\widetilde{\mathscr{D}}$ over C such that for every $h \in \mathscr{L}\left(r, J(\delta)\right)$,
$$\widetilde{\mathbb{E}}_{f\sim\widetilde{\mathscr{D}}}[\Delta(f,h)] < \delta$$
2. Condition $\widetilde{\mathscr{D}}$ on $O(1/\eta^2)$ coordinates to get h'.
3. Use $\Delta(h',h)\big(\Delta(h',h)-\delta\big) + \eta \geq 0$ to conclude
$$\Delta(h',h) \leq O(\eta).$$
4. Unique-decode from h'.

Extensions

- Distance Amplification Scheme of Alon-Edmonds-Luby'95

  C_{base}: high-rate positive distance code

  C_{base} —AEL→ Rate R, Distance $1 - R - \gamma$, Alphabet $2^{poly(1/\gamma)}$ + (Our Work) List-decodable upto $J(1 - R - \gamma)$

- Non-binary Tanner codes
- (Weighted) List Recovery
- Concatenated Code upto Johnson bound

Alon-Edmonds-Luby (AEL) Amplification

- Only impose local code constraint on left side
- Local view on the right to be seen as a single alphabet symbol

$$\delta_0 \cdot \Delta_L(f, g) \leq \Delta_E(f, g) \leq \Delta_L(f, g) \cdot \Delta_R(f, g) + \lambda$$

$$\Delta_R(f, g) \geq \delta_0 - \frac{\lambda}{\Delta_L(f, g)}$$

- Choose an (high-rate) outer code C_1 with distance δ_1, and $\lambda = \epsilon \cdot \delta_1$.
- Final code has rate $R(C_1) \cdot R(C_0)$ and distance $\delta_0 - \frac{\lambda}{\delta_1}$.

List Decoding for AEL Amplification

- Typically, inner code is Reed-Solomon, with rate R_0 and distance $1 - R_0$.
- Choose outer code C_1 to be a high-rate code, decodable upto some constant radius.
- Final code has distance $1 - R_0 - \epsilon$.
- Can be list decoded to radius $1 - \sqrt{R_0} - \epsilon_2$.
- Works via reduction to (unique-)decoding of C_1.

Future Directions

- Faster Algorithms
  - Spectral
  - Regularity Lemmas
- Beyond Johnson bound
  - Interesting combinatorially also
- Quantum LDPC Codes
  - [Upcoming work] Can list-decode quantum AEL codes.

Thank you!

Deterministic Algorithm

- All of these algorithms can be made deterministic.
- Try out all conditionings.
  - For degree-t SoS, only n^t many conditionings.
- Use threshold rounding to derandomize the rest.

Correlation Rounding via Conditioning

[Barak, Raghavendra, Steurer '11]

- Suppose $\mathbb{E}_{l,r}[\widetilde{\mathbb{E}}[\mathbf{1}_{f_l \neq 0} \cdot \mathbf{1}_{f_r \neq 0}]]$ and $\mathbb{E}_{l,r}[\widetilde{\mathbb{E}}[\mathbf{1}_{f_l \neq 0}] \cdot \widetilde{\mathbb{E}}[\mathbf{1}_{f_r \neq 0}]]$ are more than η-different.
- Then $\{\mathscr{D}_\ell\}_{\ell \in L}$ and $\{\mathscr{D}_r\}_{r \in R}$ are correlated on average.
- Conditioning $\widetilde{\mathscr{D}}$ on a random $r \in R$ reduces the average variance of $\{\mathscr{D}_\ell\}_{\ell \in L}$ by $\Omega_d(\eta^2)$.
- After $O(1/\eta^2)$ conditionings, must have low correlation on average.
- Can afford to condition this many times if the ensemble was induced by larger degree moments.