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) \hspace{1cm} y \in C \hspace{1cm}$\tilde{y} \in \{0,1\}^n$ (Decoding)

$\tilde{y}$: codeword \hspace{1cm} 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 $\frac{k}{n}$.

\[
\delta = \frac{\delta}{2}
\]
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.

$$\frac{\delta}{2} < J(\delta) < \delta$$

- 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 \(C_0\).

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

- Sipser-Spielman’96: $\approx \frac{\delta}{48}$
- Zémor’01: $\approx \frac{\delta}{4}$
- Skachek-Roth’03: $\approx \frac{\delta}{2}$

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

Inner Code: $C_0$ – distance $\delta_0$

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 $n^{O_d(1/\epsilon^4)}$. 

Bounded list

Unique Decoding

Polynomial-time Algorithm
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

- Proofs-to-Algorithms paradigm for codes.
  - Distance Proof = Local Properties + Spectral Expansion (For local-to-global)
  - Used for decoding Ta-Shma code [Richelson-Roy’23]
- Rounding algorithms for convex optimization based decoders.

List Decoding Algorithm upto Johnson bound
Covering Lemma

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

Lemma. Given a family $\mathcal{F}$ of unit vectors in $\mathbb{R}^n$, and a unit vector $g \in \mathbb{R}^n$, such that
\[ \forall f \in \mathcal{F}, \; \langle g, f \rangle > \alpha. \quad \alpha \in (0,1) \]

There exists $g' \in \text{conv}(\mathcal{F})$ such that,
\[ \forall f \in \mathcal{F}, \; \langle g', f \rangle > \alpha^2. \]

Proof. $g'$ is the smallest $\ell_2$-norm vector in $\text{conv}(\mathcal{F})$. 
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 $\delta = \frac{1 - \beta}{2}$, list sizes are polynomial until $J(\delta) = \frac{1 - \sqrt{\beta}}{2} \in \left( \frac{\delta}{2}, \delta \right)$. 
Algorithm-friendly proof of Johnson bound

For any \( h \in \mathcal{L}(r, J(\delta)) \), it holds that \( \langle \chi(r), \chi(h) \rangle > \sqrt{\beta} \).

Covering Lemma \( \implies \) There is an \( r' \in \text{conv}(\mathcal{L}) \) such that for any \( h \in \mathcal{L}(r, J(\delta)) \),
\[
\langle r', \chi(h) \rangle > \beta.
\]

\( r' \) as a convex combination \( \rightarrow \) distribution \( \mathcal{D} \) over \( C \).
\[
\mathbb{E}_{f \sim \mathcal{D}} \left[ \Delta(f, h) \right] < \delta
\]

Support of \( \mathcal{D} \) contains \( \mathcal{L}(r, J(\delta)) \).

Pick \( \mathcal{D} \) with support size \( \leq n + 1 \).

\[
\delta = \frac{1 - \beta}{2}, \quad J(\delta) = \frac{1 - \sqrt{\beta}}{2}
\]

Can take exponential time! Carathéodory's Theorem
Exponential Time Algorithm

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

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

Four distances:

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

1. $\Delta_E(f, g) = \frac{|F|}{nd}$
2. $\Delta_L(f, g) = \frac{|S|}{n}$
3. $\Delta_R(f, g) = \frac{|T|}{n}$
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

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

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

\[ \delta_0 \cdot \Delta_{LR}(f, g) \leq \Delta_E(f, g) \leq \Delta_{LR}(f, g)^2 + \lambda \cdot \Delta_{LR}(f, g) \]

\[ \Delta_{LR}(f, g)^2 - (\delta_0 - \lambda) \cdot \Delta_{LR}(f, g) \geq 0 \]

\[ \Delta_{LR}(f, g) = 0 \text{ or } \Delta_{LR}(f, g) \geq \delta_0 - \lambda \]

\[ \implies \Delta_E(f, g) \geq \delta_0(\delta_0 - \lambda) \]

\(|L| = |R| = n\]

Degree \(d\)
\(\lambda\)-expander
Continuous Relaxation for Tanner Code

\[ \mathcal{D} = \{ \mathcal{D}_\ell \}_{\ell \in L}, \{ \mathcal{D}_r \}_{r \in R} \]

Consistency along edges

Used for LP Decoding

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

Ensemble of distributions

| Pseudocodeword |

Degree \( d \)

| L | | R | = n 
1/2 | 1/2

\( \lambda \)-expander
Distance Proof for Relaxation of Tanner Code?

\[ 
\mathbb{E}_e[\mathbb{E}[1_{f_i \neq 0}]] \geq \mathbb{E}_l[\mathbb{E}[\delta_0 \cdot 1_{f_i \neq 0}]] \\
\Delta_E(\mathcal{D}, 0) \geq \delta_0 \cdot \Delta_L(\mathcal{D}, 0) \\
\Delta_E(\mathcal{D}, 0) \geq \delta_0 \cdot \Delta_{LR}(\mathcal{D}, 0) 
\]
Distance Proof for Relaxation of Tanner Code?

\[ E_e[\mathbb{E}[1_{f_c \neq 0}]] \leq E_{l \sim r}[\mathbb{E}[1_{f_l \neq 0} \cdot 1_{f_r \neq 0}]] \]

? \leq E_{l, r}[\mathbb{E}[1_{f_l \neq 0} \cdot 1_{f_r \neq 0}]]

? \leq E_{l, r}[\mathbb{E}[1_{f_l \neq 0}] \cdot \mathbb{E}[1_{f_r \neq 0}]]
Continuous Relaxation for Tanner Code

\[ \tilde{\mathcal{D}} = \{ \mathcal{D}_\ell \}_{\ell \in L}, \{ \mathcal{D}_r \}_{r \in R} \]

Consistency along edges

Modifications:

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

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

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

Uses low average correlation obtained by random conditioning.
Exponential Time Algorithm

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

1. Use covering lemma to find distribution $\mathcal{D}$ over $C$ such that for every $h \in \mathcal{L}(r, J(\delta))$,
   \[ \mathbb{E}_{f \sim \mathcal{D}}[\Delta(f, h)] < \delta \]

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

3. Use distance of $C$ $\Delta(h', h)(\Delta(h', h) - \delta) \geq 0$
   to conclude
   \[ h' = h \]
   \[ \Delta(h', h) = 0. \]
Exponential Time Algorithm

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

Time $2^n$

Polynomial Time Algorithm

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

Time $n^{1/\eta^2}$
Extensions

- Distance Amplification Scheme of Alon-Edmonds-Luby’95
  \[ C_{\text{base}} : \text{high-rate positive distance code} \]

- Non-binary Tanner codes

- (Weighted) List Recovery

- Concatenated Code upto Johnson bound

\[
\begin{align*}
C_{\text{base}} & \quad \text{AEL} \quad \text{Rate } R \\
\text{Distance } 1 - R - \gamma \\
\text{Alphabet } 2^{poly(1/\gamma)} \\
\end{align*}
\]

Our Work
List-decodable upto \( J(1 - R - \gamma) \)
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 \( \delta_1 \), 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 $1 - R_0 - \varepsilon$.
- Can be list decoded to radius $1 - \sqrt{R_0} - \varepsilon_2$.
- Works via reduction to (unique-)decoding of $C_1$.  

\[ R_0 - C_1 \]
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 \( \mathbb{E}_{l,r}[\mathbb{E}[\mathbf{1}_{f_i \neq 0} \cdot \mathbf{1}_{f_r \neq 0}]] \) and \( \mathbb{E}_{l,r}[\mathbb{E}[\mathbf{1}_{f_i \neq 0}] \cdot \mathbb{E}[\mathbf{1}_{f_r \neq 0}]] \) are more than \( \eta \)-different.
- Then \( \{ \mathcal{D}_\ell \}_{\ell \in L} \) and \( \{ \mathcal{D}_r \}_{r \in R} \) are correlated on average.
- Conditioning \( \mathcal{D} \) on a random \( r \in R \) reduces the average variance of \( \{ \mathcal{D}_\ell \}_{\ell \in L} \) 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.