Linear-Programming Decoding of Tanner Codes with Local-Optimality Certificates

Nissim Halabi   Guy Even

School of Electrical Engineering, Tel-Aviv University

July 6, 2012

Communication Over a Noisy Channel

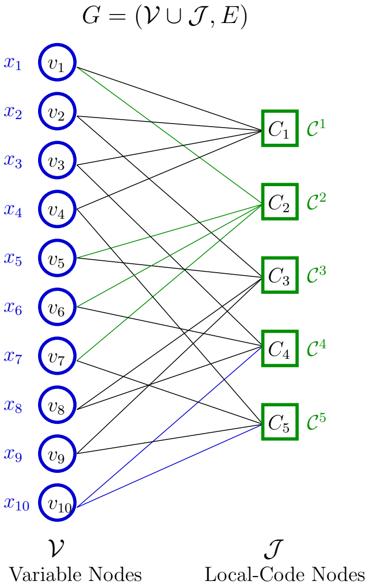
- MBIOS channel: memoryless, binary-input, output-symmetric
- Log-Likelihood-Ratio (LLR):

$$\lambda_i(y_i) \triangleq \ln\left(\frac{\Pr(y_i \mid c_i = 0)}{\Pr(y_i \mid c_i = 1)}\right)$$

- Linear Code: $\mathcal{C} \subseteq \{0,1\}^N$ is subspace of $\mathbb{F}_2^N$ of dimension k.
- Optimal decoding: Maximum Likelihood decoding. Input: y. Output: $\textsc{ml}(y)$.

$$\begin{aligned}\textsc{ml}(y) &\triangleq \arg\max_{x \in \mathcal{C}} \Pr\{y \mid c = x\} \\ &= \arg\min_{x \in \mathcal{C}} \langle \lambda(y), x \rangle\end{aligned}$$

Tanner Graphs and Tanner Codes



- Tanner code $\mathcal{C}(G, \mathcal{C}^{\mathcal{J}})$ represented by bipartite graph
- $x \in \mathcal{C}(G, \mathcal{C}^{\mathcal{J}})$ iff $x \in \mathcal{C}^j$ for every $j \in \{1, \ldots, J\}$
- degrees: can be regular, irregular, bounded, or arbitrary
- can allow arbitrary linear local codes
- minimum local distance $d^* \triangleq \min_j d_j$
- Examples: LDPC codes [Gallager'63], Expander codes [Sipser-Spielman'96]

Linear Programming (LP) Decoding

- $\mathrm{conv}(X) \subseteq \mathbb{R}^N$ - the convex hull a set of points $X \subseteq \mathbb{R}^N$.
- ML-decoding can be rephrased:

$$\mathrm{ML}(y) \triangleq \underset{x \in \mathrm{conv}(\mathcal{C})}{\arg\min} \langle \lambda(y), x \rangle$$

- Generalized fundamental polytope of a Tanner code $C(G, \mathcal{C}^{\mathcal{J}})$ - relaxation of $\mathrm{conv}(\mathcal{C})$ [following Feldman-Wainwright-Karger'05]

$$\mathcal{P}(G, \mathcal{C}^{\mathcal{J}}) \triangleq \bigcap_{\mathcal{C}^j \in \mathcal{C}^{\mathcal{J}}} \mathrm{conv}(\mathcal{C}^j)$$

- LP-decoding:

$$\mathrm{LP}(y) \triangleq \underset{x \in \mathcal{P}(G, \mathcal{C}^{\mathcal{J}})}{\arg\min} \langle \lambda(y), x \rangle$$

LP Decoding with ML Certificate

LP-decode(λ)

solve LP: $\hat{x}^{\mathrm{LP}} \leftarrow \arg\min_{x \in \mathcal{P}(G, \mathcal{C}^{\mathcal{J}})} \langle \lambda, x \rangle$.
if $\hat{x}^{\mathrm{LP}} \in \{0,1\}^N$ **then**
    return $\hat{x}^{\mathrm{LP}}$ is an ML codeword
else
    return fail
end if

- Polynomial time algorithm
- Applies to any MBIOS channel!
- Integral solution $\Rightarrow$ ML-certificate

Goal: Analysis of Finite Length Codes

Problem (Finite Length Analysis)

Design: *Constant rate code* $\mathcal{C}(G, \mathcal{C}^{\mathcal{J}})$ *and an efficient decoding algorithm* DEC.

Analyze: *If* $SNR > t$, *then*

$$Pr(\text{DEC}(\lambda) \neq x | c = x) \leqslant exp(-N^{\alpha})$$

for some $0 < \alpha$.

Goal: *Minimize* t *(lower bound on SNR).*

Remarks:

- Not an asymptotic problem
- Code is not chosen randomly from an ensemble
- Successful decoding $\neq$ ML decoding

Certificate for ML-Optimality / LP-Optimality

Problem (Optimality Certificate)

Input: *Channel observation λ and a codeword $x \in \mathcal{C}$*

Question 1: *Is x ML-optimal with respect to λ? is it unique? (NP-Hard)*

Question 2: *Is x LP-optimal with respect to λ? is it unique?*

Relax: one-sided error test

- A positive answer = certificate for the optimality of x w.r.t. λ
- A negative answer = don't know if optimal or not (allow one sided error)
- Prefer: efficient test via local computations $\Rightarrow$ "Local-Optimality" criterion

Definition of Local-Optimality

- [Feldman'03] For $x \in \{0,1\}^N$ and $f \in [0,1]^N \subseteq \mathbb{R}^N$, define the relative point $x \oplus f$ by $(x \oplus f)_i \triangleq |x_i - f_i|$
- Consider a finite set of "deviations" $\triangleq \mathcal{B} \subset [0,1]^N$

Definition (following [Arora-Daskalakis-Steurer'09])

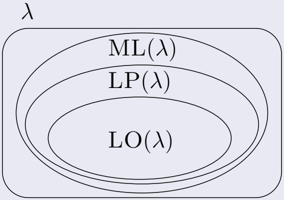
A codeword $x \in \mathcal{C}$ is locally-optimal w.r.t. $\lambda \in \mathbb{R}^N$ if for all vectors $\beta \in \mathcal{B}$,

$$\langle \lambda, x \oplus \beta \rangle > \langle \lambda, x \rangle$$

Goal

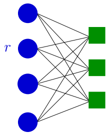
Find a set $\mathcal{B}$ such that:

1. $x \in \text{LO}(\lambda) \Rightarrow x = \text{ML}(\lambda)$ & unique
2. $x \in \text{LO}(\lambda) \Rightarrow x = \text{LP}(\lambda)$ & unique
3. $\Pr_\lambda\{x \in \text{LO}(\lambda) \mid c = x\} = 1 - o(1)$

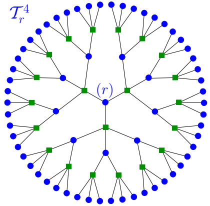


Set $\mathcal{B}$: Projections of Normalized Weighted Subtrees in Computation Trees of the Tanner Graph

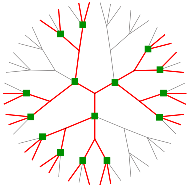
1. Tanner graph G



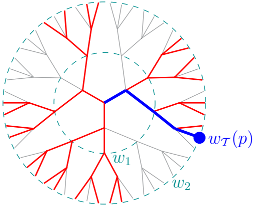
2. Computation tree of G, height$=2h$, root$=var$ node r



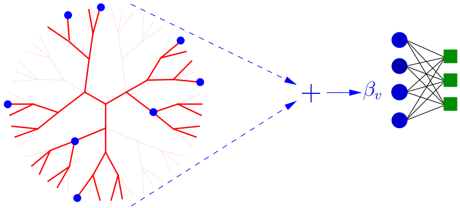
3. d-tree: subtree $\mathcal{T}$ $\deg$(local-code node)$=d$



4. Weight function $w_{\mathcal{T}} : \hat{\mathcal{V}}(\mathcal{T}) \to \mathbb{R}$



5. Projection of weighted d-tree to Tanner graph.
Deviation $\beta \in \mathbb{R}^N=$ projection assignment to var. nodes



Local-Optimality based on Deviations Set $\mathcal{B}_d^{(w)}$

- Set of deviations $\mathcal{B}_d^{(w)}$ = projections of w-weighted d-trees

$$\mathcal{B}_d^{(w)} \triangleq \left\{ \beta \in \mathbb{R}^N \mid \exists w\text{−}weighted\ d\text{−tree}\ \mathcal{T}\ :\ \beta = \pi(\mathcal{T}) \right\}$$

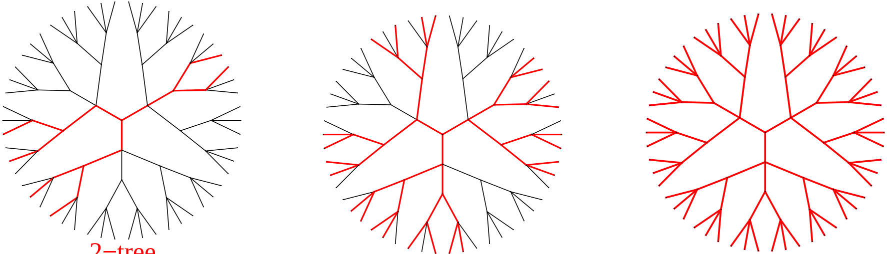
Definition

A codeword $x \in \mathcal{C}$ is (h, w, d)-locally optimal w.r.t. $\lambda \in \mathbb{R}^N$ if for all vectors $\beta \in \mathcal{B}_d^{(w)}$,

$$\langle \lambda, x \oplus \beta \rangle > \langle \lambda, x \rangle$$

- $2h$ - tree height (h levels)
- $w \in \mathbb{R}_+^h$ - tree level weights
- $d \geqslant 2$ - local-code nodes degree

d-Trees Get "Fatter" As d Increases



2−tree (skinny tree [ADS'09]) | 3−tree | 4−tree

- Over an MBIOS channel, the probability of a local-optimality certificate increases as deviations become denser

Local Optimality $\Rightarrow$ unique ML-codeword

Theorem

Let $d \geqslant 2$. If x is (h, w, d)-locally optimal w.r.t. λ, then x is the unique ML-codeword w.r.t. λ.

- hard: is x the unique ML-codeword?
- easy: is x is locally optimal? (dynamic programming)

Proof method:

Lemma (Decomposition Lemma)

Every codeword is a conic combination of projections of weighted d-trees in computation trees of G

$$x = \alpha \cdot \mathbb{E}_{\beta \in_\rho \mathcal{B}_d^{(w)}} [\beta]$$

- Following [ADS'09]: decomposition lemma $\Rightarrow$ unique ML

□

Local Optimality ⇒ unique LP optimality

Theorem

If x is a (h, w, d)-locally optimal codeword w.r.t. λ, then x is also the unique optimal LP solution given λ.

Proof method:

1. Use graph cover decoding [Koetter-Vontobel'05]: In graph covers, realization of LP-opt and ML codeword are the same
2. Lemma: local-optimality is invariant w.r.t. lifting to covering graphs
3. Lift of locally optimal codeword is the unique ML-codeword in the graph cover.

□

	[KV'06][ADS'09][HE'11]	Current
Deviation height	$h < \frac{1}{4}\mathrm{girth}(G)$. Local isomorphism	h is unbounded. Characterization using computation trees
Regularity	Regular Tanner graph	Irregular Tanner graph. Add normalization factors according to node degrees [Von'10]
Constraints	Single parity-check codes	Linear codes. Tighter relaxation for the generalized fundamental polytope (also in [Von'10])
Deviations	"skinny". Locally satisfies parity checks.	"fat". Not necessarily satisfies local codes.
LP solution analysis	Dual/Primal LP. Polyhedral analysis.	Use reduction to ML via characterization of graph cover decoding

- Form of finite length bounds: $\exists c > 1. \exists t. \forall \text{noise} < t.$

$$\Pr\{\text{LP decoder fails}\} \leqslant exp(-c^{\text{girth}})$$

- If girth $= \theta(\log N)$, then

$$\Pr\{\text{LP decoder fails}\} \leqslant exp(-N^{\alpha}), \quad \text{for } 0 < \alpha < 1$$

- $N \to \infty$: t is a lower bound on the threshold of LP-decoding with LO-certificate

	[Skachek-Roth'03]	[Feldman-Stein'05]	Current work
Decoder	Iterative	LP	LP
Channels	Bit-flipping (worst-case)	Bit-flipping (worst-case)	MBIOS (average-case)
Technique	Expansion	Expansion	Density evolution of sum-min-sum random process
Example: BSC(p) threshold $(2, d_R)$-reg Tanner code, Rate$=0.375$	$d_R >> 2$ $d^* >> 2$ $p^{iterat.} > 0.0016$	$d_R >> 2$ $d^* >> 2$ $p^{\text{LP}} > 0.0008$	$d_R = 16$ $d^* = 4$ $p^{\text{LP}} > 0.044$

Summary

Conclusions

Follows line of works based on combinatorial characterizations of local-optimality [KV'06] [ADS'09] [Von'10] [HE'11]:

1. A new combinatorial characterization of local-optimality for irregular Tanner codes
2. Local-opt. $\Rightarrow$ ML-opt.
3. Local-opt. $\Rightarrow$ LP-opt.
4. Efficiently computed certificate (dynamic-programming)
5. Upper bounds on the word error probability of LP-decoding

Open questions

- Prove bounds on noise thresholds for LP-decoding that are better than $p^{\mathrm{LP}} \geqslant 0.05$ for rate-$\frac{1}{2}$ codes [ADS'09]
- Probabilistic analysis for irregular Tanner codes
- Probabilistic analysis beyond the girth