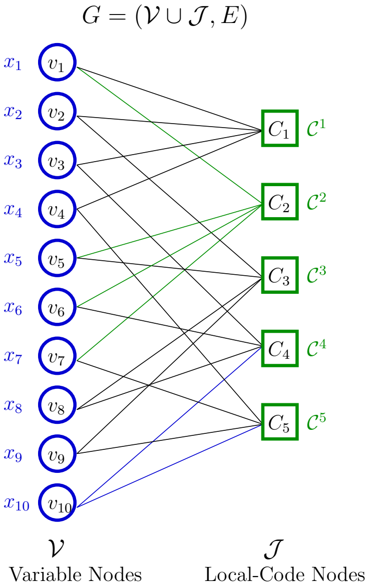
# Hierarchies of Local-Optimality Characterizations in Decoding Tanner Codes


Nissim Halabi    Guy Even

School of Electrical Engineering, Tel-Aviv University


July 6, 2012

# 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 $d^* \triangleq \min_j d_j$

# Decoding of Tanner Codes over MBIOS Channels


- Memoryless binary-input output-symmetric channel (MBIOS) channels characterized by a log-likelihood ratio (LLR) observations $\lambda$:

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

- ML-decoding:

$$\text{ML}(\lambda) \triangleq \arg\min_{x \in \text{conv}(\mathcal{C})} \langle \lambda, x \rangle$$

- LP-decoding [following Feldman-Wainwright-Karger'05]:

$$\text{LP}(\lambda) \triangleq \arg\min_{x \in \mathcal{P}(G,\mathcal{C}^{\mathcal{J}})} \langle \lambda, x \rangle$$

where $\mathcal{P}(G,\mathcal{C}^{\mathcal{J}}) \triangleq$ generalized fundamental polytope of a Tanner code $C(G,\mathcal{C}^{\mathcal{J}})$

# Local-Optimality: Sufficient Condition for Successful Decoding of Finite-Length Codes

- Set of deviations $\mathcal{B}_d^{(w)} \subset \mathbb{R}^N$ : finite set of vectors corresponding to projections of $w$-weighted $d$-trees in computation trees with height $2h$ of the Tanner graph

**Definition ([Even-H'11])**

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

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

**Theorem ([Even-H'11])**

*Let $2 \leqslant d \leqslant d^*$. If $x \in \mathcal{C}$ is $(h, w, d)$-locally optimal w.r.t. $\lambda$, then:*

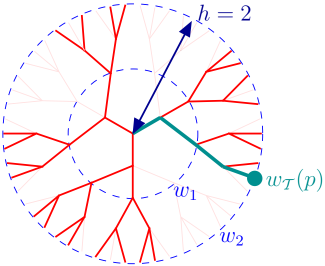
1. *$x$ is the unique ML codeword w.r.t. $\lambda$*
2. *$x$ is the unique optimal solution of the LP-decoder w.r.t. $\lambda$*

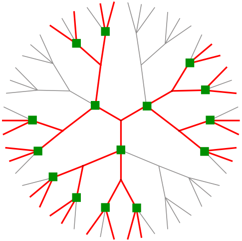
# Parameters $(h, w, d)$ for Deviations of Local-Optimality

- $\mathcal{B}_d^{(w)} \subset \mathbb{R}^N$ : finite set of projections of $w$-weighted $d$-trees with height $2h$ in computation trees of the Tanner graph

$h \in \mathbb{N}$ - tree height $/2$
$w \in \mathbb{R}_+^h$ - level weights

$d \in \mathbb{N}$ - degree of local-codes nodes in the sub-tree, $2 \leqslant d \leqslant d^*$



3-tree ($d = 3$)

- An hierarchy statement: LO with parameter $h \Rightarrow$ LO with parameter $h' > h$

# Height Hierarchy of Local-Optimality: Motivation

- An iterative decoding algorithm (NWMS) is guaranteed to decode LO-certified codeword in $h$ iterations [Even-H'11]
- Questions: what is the effect of increasing the number of iterations? even when number of iterations exceeds the girth?

**Theorem (Hierarchy of local-optimality based on height)**

*An $(h, w, d)$-strongly locally optimal codeword $x$ w.r.t. $\lambda$ is also $(k \cdot h, \tilde{w}, d)$-strongly locally optimal w.r.t. $\lambda$ for any $k \in \mathbb{N}$*

$x$ is SLO with height parameter $h \Rightarrow$ Iterative message-passing decoding by NWMS is guaranteed to decode the ML-certified $x$ after $k \cdot h$ iterations $\forall k \in \mathbb{N}_+$

- Insight on convergence: If a codeword $x$ is SLO-certified after $h$ iterations, then $x$ is the outcome of NWMS infinitely many times

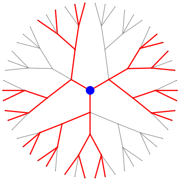
# Strong Local-Optimality

**Definition ($(h, w, d)$-Strong Local Optimality)**

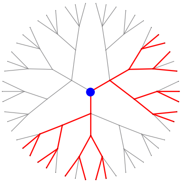
A codeword $x \in \mathcal{C}$ is $(h, w, d)$-strongly locally optimal w.r.t. $\lambda \in \mathbb{R}^N$ if for all vectors $\beta \in \overline{\mathcal{B}}_d^{(w)}$,

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

- $\overline{\mathcal{B}}_d^{(w)} \subset \mathbb{R}^N$ : finite set of projections of $w$-weighted reduced $d$-trees in computation trees with height $h$ of the Tanner graph
- Reduced: $\deg_{\mathcal{T}}(\text{root}) = \deg_G(\text{root}) - 1$ (as if the root itself hangs from an edge)



3−tree
defines deviations for local−optimality



Reduced 3−tree
defines deviations for strong local−optimality

# Strong Local-Optimality vs. Local-Optimality

- Set of pairs $(x, \lambda)$ s.t. $x$ is locally-optimal w.r.t. $\lambda$:

$$\textsc{lo}_{\mathcal{C}}(h, w, d) \triangleq \{(x, \lambda) \in \mathcal{C} \times \mathbb{R} \mid x \text{ is } (h, w, d)\text{–}\textsc{lo} \text{ w.r.t. } \lambda\}$$

- Set of pairs $(x, \lambda)$ s.t. $x$ is strongly locally-optimal w.r.t. $\lambda$:

$$\textsc{slo}_{\mathcal{C}}(h, w, d) \triangleq \{(x, \lambda) \in \mathcal{C} \times \mathbb{R} \mid x \text{ is } (h, w, d)\text{–}\textsc{slo} \text{ w.r.t. } \lambda\}$$

**Lemma**

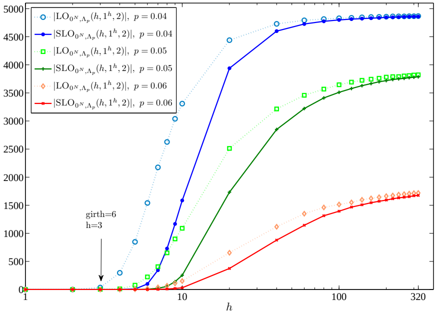
$\textsc{slo}_{\mathcal{C}}(h, w, d) \subseteq \textsc{lo}_{\mathcal{C}}(h, w, d)$

**Corollary**

$\textsc{slo} \Rightarrow$ *unique* $\textsc{ml}$

$\textsc{slo} \Rightarrow$ *unique* $\textsc{lp}$ *opt.*

- Empirically: $\textsc{slo}$ approaches $\textsc{lo}$ as $h$ increases ($h >> \text{girth}(G)$)

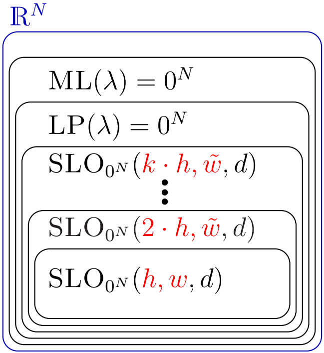


# Height Hierarchy of Local-Optimality

**Theorem** ($h$-hierarchy of local-optimality)

*For every $k \in \mathbb{N}$ and geometric level-weights $w$,*

$$\textsc{slo}_{\mathcal{C}}(h, w, d) \subseteq \textsc{slo}_{\mathcal{C}}(k \cdot h, w, d)$$




- Remark: Assume $x$ is transmitted over an MBIOS channel with some bounded noise level, then $x$ is $\textsc{slo}$ w.r.t. the received LLR with high probability
- Interesting because an iterative message-passing decoding algorithm is guaranteed to find $\textsc{slo}$ codewords

# Proof Method for Height Hierarchy of Local-Optimality

Proof via contrapositive statement:

**Lemma (Symmetry of SLO)**

*$x$ is SLO w.r.t. $\lambda$ iff $0^N$ is SLO w.r.t. $\lambda^0 \triangleq (-1)^x * \lambda$*

$x$ not $(k \cdot h, w, d)$-SLO w.r.t. $\lambda$

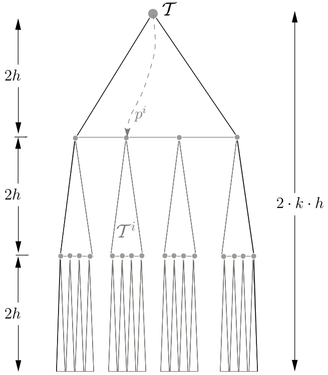
$\Longleftrightarrow$ $0^N$ is not $(k \cdot h, w, d)$-SLO w.r.t. $\lambda^0$ [symmetry]

$\Longleftrightarrow$ $\exists \beta$ based on reduced $d$-tree of height $2 \cdot k \cdot h$ s.t. $\langle \lambda^0, \beta \rangle < 0$ [SLO]

$\Rightarrow$ $\exists \beta'$ based on reduced $d$-tree of height $2 \cdot h$ s.t. $\langle \lambda^0, \beta' \rangle < 0$ ["averaging"]

$\Longleftrightarrow$ $0^N$ is not $(h, w, d)$-SLO w.r.t. $\lambda^0$ [SLO]

$\Longleftrightarrow$ $x$ not $(h, w, d)$-SLO w.r.t. $\lambda$ [symmetry]



# Implications of Height Hierarchy to Message-Passing Decoding

Setting:

- Irregular Tanner codes
- Local-codes = single parity-check code
- Local-optimality: $d = 2$, arbitrary height $h$ (not limited by girth)

Normalized Weighted Min-Sum (NWMS) Algorithm [Even-H'11]

- Normalize: take care of irregular degrees
- Weighted: allow level weights
- Min-Sum: based on Max-Product/Min-Sum algorithm

$\textsc{nwms}(\lambda, h, w)$:

- Input: $\lambda \in \mathbb{R}^N$ - LLRs from the channel
  $h \in \mathbb{N}$ - number of iterations
  $w \in \mathbb{R}_+^h$ - level weights
- Output: $\hat{x} \in \{0,1\}^N$

# BP-Based Decoding - $\textsc{nwms}(\lambda, h, w)$: Init

Initialize:

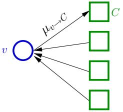
- Init check-to-variable messages with 0:

$$\forall C \in \mathcal{J},\ \forall v \in \mathcal{N}(C): \quad \mu^{(-1)}_{C \to v} \leftarrow 0$$

# BP-Based Decoding - $\mathrm{NWMS}(\lambda, h, w)$: Iterations

Iterate: for $\ell = 0$ to $h-1$:

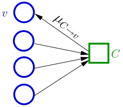
- message: variable node $\to$ check node




$$\mu^{(\ell)}_{v\to C} \leftarrow \frac{w_{h-\ell}}{\deg_G(v)}\lambda_v + \frac{1}{\deg_G(v)-1}\sum_{C'\in\mathcal{N}(v)\setminus\{C\}} \mu^{(\ell-1)}_{C'\to v}$$

[degree normalization, level weights]

- message: check node $\to$ variable node




$$\mu^{(\ell)}_{C\to v} \leftarrow \left(\prod_{u\in\mathcal{N}(C)\setminus\{v\}} \mathrm{sign}\left(\mu^{(\ell)}_{u\to C}\right)\right)\cdot \min_{u\in\mathcal{N}(C)\setminus\{v\}} \left\{\left|\mu^{(\ell)}_{u\to C}\right|\right\}$$

# BP-Based Decoding - $\textsc{nwms}(\lambda, h, w)$: Decision

Decision:

**for all** $v \in \mathcal{V}$ **do**

$$\mu_v \leftarrow \sum_{C \in \mathcal{N}(v)} \mu_{C \to v}^{(h-1)}$$

$$\hat{x}_v \leftarrow \begin{cases} 0 & \text{if } \mu_v > 0, \\ 1 & \text{otherwise.} \end{cases}$$

**end for**

# NWMS: Decoding Guarantee via Local-Optimality

**Theorem (Even-H'11)**

*If $x$ is $(h, w, 2)$-locally optimal w.r.t. $\lambda$ then* $\textsc{nwms}(\lambda, h, w)$ *returns $x$*

**Height Hierarchy of Local-Optimality Implies:**

If NWMS finds an SLO certified codeword $x$ after $h$ iterations, then

1. NWMS outputs $x$ every $k \cdot h$ iterations (infinitely many times)
2. NWMS never outputs a codeword $y \neq x$ for any number of iterations

- Holds for all $h$, decoding guarantee not limited by the girth!
- No issue of convergence/divergence (as in density evolution)
- ML-certificate by local-optimality

# Degree Hierarchy of Local-Optimality Characterization

- What is the effect of increasing the minimum distance of the local codes in Tanner codes?

**Theorem (Hierarchy of local-optimality based on degrees)**

*An $(h, w, d)$-locally optimal codeword $x$ w.r.t. $\lambda$ is also $(h, w, d')$-locally optimal w.r.t. $\lambda$ for any degree parameter $d' > d$*

- Insight on the improvement of suboptimal decodings of expander codes as the minimum distance of local codes increases.

# Summary

## Conclusions

Hierarchies of Local Optimality:

1. Degree hierarchy $\Rightarrow$ take $d$ as large as possible
2. Strong local-optimality $\Rightarrow$ local-optimality ($\Rightarrow$ LP-opt. $\Rightarrow$ ML-opt.)
3. Height hierarchy of local-optimality
   - A strongly locally optimal codeword is infinitely often strongly locally optimal (w.r.t. height parameter)
   - A BP-Based algorithm NWMS decodes $x$ with SLO certificate after $h$ iterations $\Rightarrow$ NWMS decodes $x$ with SLO certificate every $k \cdot h$ iterations, for every $k \in \mathbb{N}$

## Open Questions

- Height hierarchies for belief-propagation (sum-product) algorithm and other BP-based algorithms [no probabilistic assumptions as in monotonicity of DE]