

Local-Optimality Guarantees for Optimal Decoding Based on Paths

Guy Even ^{*} Nissim Halabi [†]

Abstract

This paper presents a unified analysis framework that captures recent advances in the study of local-optimality characterizations for codes on graphs. These local-optimality characterizations are based on combinatorial structures embedded in the Tanner graph of the code. Local-optimality implies both unique maximum-likelihood (ML) optimality and unique linear-programming (LP) decoding optimality. Also, an iterative message-passing decoding algorithm is guaranteed to find the unique locally-optimal codeword, if one exists.

We demonstrate this proof technique by considering a definition of local-optimality that is based on the simplest combinatorial structures in Tanner graphs, namely, paths of length h . We apply the technique of local-optimality to a family of Tanner codes. Inverse polynomial bounds in the code length are proved on the word error probability of LP-decoding for this family of Tanner codes.

1 Introduction

The method of decoding error correcting codes based on a linear programming (LP) relaxation of maximum-likelihood (ML) decoding was introduced by Feldman, Wainwright and Karger [FK02, FWK05]. Successful decoding by LP-decoding was analyzed for different codes and channels (see e.g., [FMS⁺07, FS05, DDKW08, Ska11]). Significant advances in the analysis of successful LP-decoding have been achieved recently. Following Koetter and Vontobel [KV06], Arora *et al.* [ADS09] obtained improved bounds for $(3, 6)$ -regular low-density parity-check (LDPC) codes over the binary symmetric channel (BSC). These techniques were extended to memoryless binary-input output-symmetric (MBIOS) channels [HE11] and to Tanner codes [EH11]. The proofs in these papers are based on complicated graphical structures and on a sophisticated analysis of a random min-sum process. Our goal in this paper is to present a simple analysis based on this proof technique that proves nontrivial bounds for a broad family of Tanner codes.

The proof technique in [ADS09] is based on the following steps:

^{*}School of Electrical Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel.
E-mail: guy@eng.tau.ac.il.

[†]School of Electrical Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel.
E-mail: nissimh@eng.tau.ac.il.

1. Define a set of deviations. A deviation is induced by combinatorial structures in the Tanner graph or the computation tree [Wib96, KV06, ADS09, Von10, EH11].
2. Define local-optimality. Local-optimality is a combinatorial characterization of a codeword $x \in \{0, 1\}^N$ with respect to a channel output $\lambda \in \mathbb{R}^N$. This definition is based on: (1) a definition of a relative point for a codeword x and each deviation β [Fel03], and (2) a definition of the cost of each relative point with respect to the log-likelihood ratios (LLR) vector λ . Loosely speaking, a codeword x is locally-optimal if its cost is smaller than the cost of every relative point.
3. Prove that if x is locally-optimal codeword w.r.t. an LLR vector λ , then x is the unique maximum-likelihood (ML) codeword. The proof is based on a decomposition lemma that states that every codeword is a conical sum of deviations.
4. Prove that if x is locally-optimal codeword w.r.t. an LLR vector λ , then x is the unique linear-programming (LP) codeword. This proof is based on a lifting lemma [HE11] that states that local-optimality is invariant under liftings of codewords to covering graphs.
5. Analyze the probability that there does not exist a locally-optimal codeword.

We demonstrate this proof technique by considering the simplest combinatorial structures in Tanner graphs, namely, paths of length h . We attach to each path p a deviation $\beta \in \mathbb{R}^N$ that equals the multiplicity of each variable node along p divided by its degree times $h + 1$. Surprisingly, all the ingredients of the proof technique can be demonstrated with respect to this primitive set of deviations.

The family of codes we consider is the set of Tanner codes that satisfy two properties: (1) the local codes contain only codewords of even weight, and (2) all the variable node degrees are even. We refer to this family of Tanner codes as *even Tanner codes*¹. Among the codes in this family are: (1) LDPC codes with even left degrees, (2) irregular repeat accumulate codes where the repetition factors are even, and (3) expander codes with even variable node degrees and even weighted local codes.

Apart from offering a simple application of a powerful proof technique, we shed light on two issues. (1) The deviations should *not* be limited by the girth of the Tanner graph. Indeed, we consider paths of arbitrary length h which need not be simple for steps 1-4 of the proof technique. One should note that step 5 that bounds the probability of the existence of a locally-optimal codeword requires independence of the LLRs in the deviation. Hence, the bound here holds only for simple paths, thus limiting h by the girth. (2) The Tanner graph need not be regular. Irregular degrees are handled by dividing the deviations by the node degrees (see [Von10]).

Local-optimality is also related to iterative decoding. An iterative message-passing decoding algorithm is presented in [EH11] with the guarantee that, if a locally-optimal codeword exists, then the decoding algorithm finds it. Moreover, since local-optimality can be verified efficiently, an ML-certificate is obtained whenever there exists a locally-optimal codeword.

We prove inverse polynomial bounds in the code length on the word error probability for LP-decoding of even Tanner codes whose Tanner graphs have logarithmic girth. For certain

¹The parity of the node degrees is used to prove that the subgraph induced by codewords is Eulerian, a key component in the decomposition lemma of step 3.

sub-classes of even Tanner codes this technique provides new bounds on the word error probability of LP-decoding. In the case of repeat accumulate codes, we provide a simple proof of previous results [FK02, GB11].

Organization. The remainder of the paper is organized as follows. Section 2 provides background on ML-decoding and LP-decoding of Tanner codes over MBIOS channels. Section 3 contains steps 1-4 in the proof technique of local-optimality applied to even Tanner codes. In Section 4 we demonstrate step 5 by a simple probabilistic analysis of the event where no locally-optimal codeword exists. We conclude with a discussion in Section 5.

2 Preliminaries

Graph Terminology. Let $G = (V, E)$ denote an undirected graph. Let $\mathcal{N}_G(v)$ denote the set of neighbors of node $v \in V$. Let $\deg_G(v) \triangleq |\mathcal{N}_G(v)|$ denote the edge degree of node v in graph G . A path $p = (v, \dots, u)$ in G is a sequence of vertices such that there exists an edge between every two consecutive nodes in the sequence p . A *simple* path is a path with no repeated vertex. A *simple cycle* is a closed path where the only repeated vertex is the first and last vertex. A path p is *backtrackless* if every three consecutive vertices along p are distinct (i.e., a subpath (u, v, u) is not allowed). Let $|p|$ denote the length of a path p , i.e., the number of edges in p . Let $\text{girth}(G)$ denote the length of the shortest cycle in G . The *subgraph of G induced by $S \subseteq V$* consists of S and all edges in E , both endpoints of which are contained in S . Let G_S denote the subgraph of G induced by S .

Tanner-codes and Tanner graph representation. Let $G = (\mathcal{V} \cup \mathcal{J}, E)$ denote an edge-labeled bipartite-graph, where $\mathcal{V} = \{v_1, \dots, v_N\}$ is a set of N vertices called *variable nodes*, and $\mathcal{J} = \{C_1, \dots, C_J\}$ is a set of J vertices called *local-code nodes*. We associate with each local-code node C_j a linear code $\bar{\mathcal{C}}^j$ of length $\deg_G(C_j)$. Let $\bar{\mathcal{C}}^{\mathcal{J}} \triangleq \{\bar{\mathcal{C}}^j : 1 \leq j \leq J\}$ denote the set of *local codes*, one for each local-code node. We say that v_i *participates* in $\bar{\mathcal{C}}^j$ if (v_i, C_j) is an edge in E .

A word $x = (x_1, \dots, x_N) \in \{0, 1\}^N$ is an assignment to variable nodes in \mathcal{V} where x_i is assigned to v_i . Let \mathcal{V}_j denote the set $\mathcal{N}_G(C_j)$ ordered according to labels of edges incident to C_j . Denote by $x_{\mathcal{V}_j} \in \{0, 1\}^{\deg_G(C_j)}$ the projection of the word $x = (x_1, \dots, x_N)$ onto entries associated with \mathcal{V}_j .

The *Tanner code* $\mathcal{C}(G, \bar{\mathcal{C}}^{\mathcal{J}})$ based on the labeled *Tanner graph* G is the set of vectors $x \in \{0, 1\}^N$ such that $x_{\mathcal{V}_j}$ is a codeword in $\bar{\mathcal{C}}^j$ for every $j \in \{1, \dots, J\}$. Let \mathcal{C}^j denote the *extension* of the local code $\bar{\mathcal{C}}^j$ from length $\deg(C_j)$ to length N defined by $\mathcal{C}^j \triangleq \{x \in \{0, 1\}^N \mid x_{\mathcal{V}_j} \in \bar{\mathcal{C}}^j\}$. The Tanner code is simply the intersection of the extensions of the local codes, i.e., $\mathcal{C}(G, \bar{\mathcal{C}}^{\mathcal{J}}) = \bigcap_{j \in \{1, \dots, J\}} \mathcal{C}^j$.

We consider a family of Tanner codes defined as follows.

Definition 1 (even Tanner codes). A *Tanner code* $\mathcal{C}(G, \bar{\mathcal{C}}^{\mathcal{J}})$ based on a *Tanner graph* $G = (\mathcal{V} \cup \mathcal{J}, E)$ is called an *even Tanner code* if: (1) $\deg_G(v)$ is even for every $v \in \mathcal{V}$, and (2) every codeword in each local code $\bar{\mathcal{C}}^j \in \bar{\mathcal{C}}^{\mathcal{J}}$ has even weight.

LP decoding of Tanner codes over memoryless channels. Let $c_i \in \{0, 1\}$ denote the i th transmitted binary symbol (channel input), and let $y_i \in \mathbb{R}$ denote the i th received symbol (channel output). A *memoryless binary-input output-symmetric* (MBIOS) channel is defined by a conditional probability density function $f(y_i|c_i = a)$ for $a \in \{0, 1\}$, that satisfies $f(y_i|0) = f(-y_i|1)$. The binary erasure channel (BEC), binary symmetric channel (BSC) and binary-input additive white Gaussian noise (BI-AWGN) channel are examples for MBIOS channels. In MBIOS channels, the *log-likelihood ratio* (LLR) vector $\lambda \in \mathbb{R}^N$ is defined by $\lambda_i(y_i) \triangleq \ln \left(\frac{f(y_i|c_i=0)}{f(y_i|c_i=1)} \right)$ for every input bit i . For a code \mathcal{C} , *Maximum-Likelihood (ML) decoding* is equivalent to

$$\hat{x}^{\text{ML}}(y) = \arg \min_{x \in \text{conv}(\mathcal{C})} \langle \lambda(y), x \rangle, \quad (1)$$

where $\text{conv}(\mathcal{C}) \subset [0, 1]^N$ denotes the convex hull of the codewords in \mathcal{C} .

In general, solving the optimization problem in (1) for linear codes is intractable [BMvT78]. Feldman *et al.* [Fel03, FWK05] introduced a linear programming relaxation for the problem of ML decoding of Tanner codes with single parity-check codes acting as local codes. We consider an extension of this definition to the case in which the local codes are arbitrary as follows. The *generalized fundamental polytope* $\mathcal{P} \triangleq \mathcal{P}(G, \bar{\mathcal{C}}^{\mathcal{J}})$ of a Tanner code $\mathcal{C} = \mathcal{C}(G, \bar{\mathcal{C}}^{\mathcal{J}})$ is defined by

$$\mathcal{P} \triangleq \bigcap_{\mathcal{C}^j \in \mathcal{C}^{\mathcal{J}}} \text{conv}(\mathcal{C}^j). \quad (2)$$

Given an LLR vector λ for a received word y , LP-decoding is defined by the following linear program:

$$\hat{x}^{\text{LP}}(y) \triangleq \arg \min_{x \in \mathcal{P}(G, \bar{\mathcal{C}}^{\mathcal{J}})} \langle \lambda(y), x \rangle. \quad (3)$$

The difference between ML-decoding and LP-decoding is that the fundamental polytope $\mathcal{P}(G, \bar{\mathcal{C}}^{\mathcal{J}})$ may strictly contain the convex hull of \mathcal{C} . Vertices of $\mathcal{P}(G, \bar{\mathcal{C}}^{\mathcal{J}})$ that are not codewords of \mathcal{C} must have fractional components and are called *pseudocodewords*.

3 A Local Combinatorial Certificate for an Optimal Codeword

In this section we define a simple type of local-optimality characterization that is based on backtrackless paths of arbitrary length h in the Tanner graph. We prove that for codewords of even Tanner codes, this characterization suffices both for ML-optimality and LP-optimality.

Definition 2 (normalized characteristic vector). *Consider a Tanner graph $G = (\mathcal{V} \cup \mathcal{J}, E)$. The normalized characteristic vector $\chi_G(p) \in \mathbb{R}^{|\mathcal{V}|}$ of a path p is defined as follows. For every $v \in \mathcal{V}$ the component $[\chi_G(p)]_v$ equals to the multiplicity of the variable node v in the path p divided by its degree in G . Formally,*

$$[\chi_G(p)]_v = \frac{1}{\deg_G(v)} \cdot |\{v \mid v \in p\}|. \quad (4)$$

If p is closed (i.e., a cycle), then we count the multiplicity of the endpoints only once.

The normalization by $\deg_G(v)$ is needed if G has irregular variable node degrees.

For any fixed h , let $\mathcal{B}^{(h)}$ denote the set of normalized characteristic vectors of backtrackless paths of length h in G scaled by a factor $\frac{1}{h+1}$. That is,

$$\mathcal{B}^{(h)} \triangleq \left\{ \frac{\chi_G(p)}{h+1} \mid p \text{ is a backtrackless path of length } h \right\}. \quad (5)$$

Vectors in $\mathcal{B}^{(h)}$ are called *deviations*. Note that $\mathcal{B}^{(h)} \in [0, 1]^{|V|}$ because every variable node appears less than $h+1$ times in a path of length h .

For two vectors $x \in \{0, 1\}^N$ and $f \in [0, 1]^N$, let $x \oplus f \in [0, 1]^N$ denote the *relative point* defined by $(x \oplus f)_i \triangleq |x_i - f_i|$ [Fel03]. The following definition characterizes local-optimality based on backtrackless paths for even Tanner codes over MBIOS channels.

Definition 3 (path-based local-optimality). *Let $\mathcal{C}(G) \subset \{0, 1\}^N$ denote an even Tanner code and let $h \in \mathbb{N}_+$. A codeword $x \in \mathcal{C}(G)$ is h -locally optimal with respect to $\lambda \in \mathbb{R}^N$ if for all vectors $\beta \in \mathcal{B}^{(h)}$,*

$$\langle \lambda, x \oplus \beta \rangle > \langle \lambda, x \rangle. \quad (6)$$

For two vectors $y, z \in \mathbb{R}^N$, let “ \ast ” denote coordinate-wise multiplication, i.e., $(y \ast z)_i \triangleq y_i \cdot z_i$. For a word $x \in \{0, 1\}^N$, let $(-1)^x \in \{\pm 1\}^N$ denote the vector whose i th component equals $(-1)^{x_i}$. The following proposition and corollary state that the mapping $(x, \lambda) \mapsto (0^N, (-1)^x \ast \lambda)$ preserves local-optimality.

Proposition 4. [EH11] *For every $\lambda \in \mathbb{R}^N$ and every $\beta \in [0, 1]^N$,*

$$\langle (-1)^x \ast \lambda, \beta \rangle = \langle \lambda, x \oplus \beta \rangle - \langle \lambda, x \rangle. \quad (7)$$

Corollary 5 (symmetry of local-optimality). *For every $x \in \mathcal{C}$, x is h -locally optimal w.r.t. λ if and only if 0^N is h -locally optimal w.r.t. $(-1)^x \ast \lambda$.*

Proof. By Proposition 4, $\langle \lambda, x \oplus \beta \rangle - \langle \lambda, x \rangle = \langle (-1)^x \ast \lambda, \beta \rangle$. □

Corollary 5 suggests that a codeword x can be verified to be h -locally optimal w.r.t. a given LLR λ by verifying that each backtrackless path of length h has positive normalized cost w.r.t. $(-1)^x \ast \lambda$. That is, if the minimum normalized cost of every path with length h w.r.t. $(-1)^x \ast \lambda$ is positive, then x is h -locally optimal w.r.t. λ . A min-cost path of length h in a graph can be computed by a simple dynamic programming algorithm (Floyd’s algorithm) in time $O(h \cdot |E|)$. Hence, a codeword can be efficiently verified to be locally-optimal w.r.t. λ .

3.1 Local-Optimality Implies ML-Optimality

In the following section we show that local-optimality is sufficient for ML-optimality (Theorem 8). The proof of Theorem 8 is based on the representation of every codeword as a conical combination of deviations in $\mathcal{B}^{(h)}$ (Corollary 7). We first prove that every codeword in an even Tanner code is a conical combination of normalized characteristic vectors of simple cycles in the Tanner graph (Lemma 6). Then we show that every cycle of length ℓ is a conical combination of ℓ deviations in $\mathcal{B}^{(h)}$ for any arbitrary h , which implies Corollary 7.

Lemma 6 (simple cycles decomposition). *Let $\mathcal{C}(G)$ denote an even Tanner code, and let Γ denote the set of simple cycles in G . For every codeword $x \neq 0^N$, there exists a distribution ρ over the set Γ and an $\alpha > 1$, such that*

$$x = \alpha \cdot \mathbb{E}_{\gamma \in \rho \Gamma} [\chi_G(\gamma)]. \quad (8)$$

Proof. Let $\mathcal{V}_x \triangleq \{v \mid x_v = 1\}$, and let G_x denote the subgraph of the Tanner graph G induced by $\mathcal{V}_x \cup \mathcal{N}_G(\mathcal{V}_x)$. Because x is a codeword in an even Tanner code, the degree of every node (both variable nodes and local-code nodes) in G_x is even. Therefore, each connected component in G_x is Eulerian. Denote by $\{G_x^{(j)}\}$ the set of connected components of G_x , and let $\psi^{(j)}$ denote an Eulerian cycle in $G_x^{(j)}$.

Consider a variable node v in the connected component $G_x^{(j)}$, then the multiplicity of v in $\psi^{(j)}$ equals $\frac{\deg_{G_x}(v)}{2}$. Therefore, $2 \cdot \sum_j [\chi_G(\psi^{(j)})]_v = 1$ (and, by definition, $x_v = 1$).

Every Eulerian cycle $\psi^{(j)}$ can be decomposed into a set of edge disjoint simple cycles. Let $\Gamma^{(j)}$ denote the decomposition of $\psi^{(j)}$ into simple cycles. Then, $\chi_G(\psi^{(j)}) = \sum_{\gamma \in \Gamma^{(j)}} \chi_G(\gamma)$. Thus,

$$x_v = 2 \cdot \sum_j \sum_{\gamma \in \Gamma^{(j)}} \chi_G(\gamma).$$

Let ρ denote the uniform distribution over $\cup_j \Gamma^{(j)}$ and let $s \triangleq |\cup_j \Gamma^{(j)}|$. Then,

$$x = 2s \cdot \mathbb{E}_{\gamma \in \rho \Gamma} [\chi_G(\gamma)].$$

□

Corollary 7. *Let $\mathcal{C}(G)$ denote an even Tanner code, and let $h \in \mathbb{N}_+$. For every codeword $x \neq 0^N$, there exists a distribution ρ over the set $\mathcal{B}^{(h)}$ and an $\alpha' > 1$ such that*

$$x = \alpha' \cdot \mathbb{E}_{\beta \in \rho \mathcal{B}^{(h)}} [\beta].$$

Proof. Following Lemma 6, it suffices to show that, for every simple cycle γ , the set $\{\psi_i\}$ of paths in γ of length h satisfy

$$\exists \delta_\gamma \geq 1 : \chi_G(\gamma) = \delta_\gamma \cdot \sum_{i=0}^{|\gamma|-1} \frac{1}{h+1} \cdot \chi_G(\psi_i). \quad (9)$$

Indeed, let $\beta_i(\gamma) \triangleq \frac{1}{h+1} \cdot \chi_G(\psi_i)$, then

$$\begin{aligned} x &= \alpha \cdot \sum_{\gamma} \rho(\gamma) \cdot \chi_G(\gamma) \\ &= \alpha \cdot \sum_{\gamma} \left(\rho(\gamma) \cdot \delta_\gamma \cdot \sum_i \beta_i(\gamma) \right) \\ &= \alpha \cdot \sum_{\beta} (\rho(\gamma_\beta) \cdot \delta_\gamma \cdot \beta). \end{aligned}$$

Note that in the last line, we use the fact that the cycles in G_x are decomposed into edge disjoint cycles, and hence each deviation β appears in exactly one cycle, denoted by γ_β . The corollary follows because the coefficients $\rho(\gamma_\beta) \cdot \delta_\gamma$ are nonnegative and their sum is at least one.

We now prove Equation (9). Let $\gamma = (v_0, v_1, \dots, v_{\ell-1}, v_\ell = v_0)$ be a simple cycle in G of length ℓ . For every $0 \leq i \leq \ell - 1$, let $\psi_i = (v_i, v_{i+1 \bmod \ell}, \dots, v_{i+h \bmod \ell})$ denote a segment of γ that starts at node v_i and contains h edges. (Note that if $h \geq \text{girth}(G)$, then a single segment may traverse a node in the cycle more than once.) For every $0 \leq j \leq h$, a node v appears exactly once as the j th node in one of the paths $\{\psi_i\}_{i=0}^{\ell-1}$.

If h is not a multiple of ℓ , then the multiplicity of every vertex $v \in \gamma$ in $\cup_{i=0}^{\ell-1} \psi_i$ equals $h + 1$. Therefore,

$$\chi_G(\gamma) = \sum_{i=0}^{\ell-1} \frac{1}{h+1} \cdot \chi_G(\psi_i).$$

Otherwise, ℓ divides h , and every ψ_i is a cycle whose both endpoints v_i are counted as one occurrence in $\chi_G(\psi_i)$. Hence, the multiplicity of every vertex $v \in \gamma$ in $\cup_{i=0}^{\ell-1} \psi_i$ equals h , and

$$\chi_G(\gamma) = \frac{h+1}{h} \sum_{i=0}^{\ell-1} \frac{1}{h+1} \cdot \chi_G(\psi_i).$$

□

Theorem 8 (local-optimality is sufficient for ML). *Let $\mathcal{C}(G)$ denote an even Tanner code. Let $\lambda \in \mathbb{R}^N$ denote the LLR vector received from the channel and let $h \in \mathbb{N}_+$. If x is h -locally optimal codeword w.r.t. λ , then x is also the unique maximum-likelihood codeword w.r.t. λ .*

Proof. The proof follows [ADS09, proof of Theorem 2] and [HE11, proof of Theorem 6]. We use the decomposition proved in Corollary 7 to show that for every codeword $x' \neq x$, $\langle \lambda, x' \rangle > \langle \lambda, x \rangle$. Let $z \triangleq x \oplus x'$. By linearity, $z \in \mathcal{C}(G)$. Moreover, $z \neq 0^N$ because $x \neq x'$. By Corollary 7 there exists a distribution over the set $\mathcal{B}^{(h)}$, such that $\mathbb{E}_{\beta \in \mathcal{B}^{(h)}} \beta = \delta \cdot z$, where $\delta \triangleq \frac{1}{\alpha'} < 1$. Let $f : [0, 1]^N \rightarrow \mathbb{R}$ be the affine linear function defined by $f(\beta) \triangleq \langle \lambda, x \oplus \beta \rangle = \langle \lambda, x \rangle + \sum_{i=1}^N (-1)^{x_i} \lambda_i \beta_i$. Then,

$$\begin{aligned} \langle \lambda, x \rangle &< \mathbb{E}_{\beta \in \mathcal{B}^{(h)}} \langle \lambda, x \oplus \beta \rangle \quad (\text{by local-optimality of } x) \\ &= \langle \lambda, x \oplus \mathbb{E}_{\beta \in \mathcal{B}^{(h)}} \beta \rangle \quad (\text{by linearity of } f \text{ and linearity of expectation}) \\ &= \langle \lambda, x \oplus \delta z \rangle \quad (\text{by Lemma 6}) \\ &= \langle \lambda, (1 - \delta)x + \delta(x \oplus z) \rangle \\ &= \langle \lambda, (1 - \delta)x + \delta x' \rangle \\ &= (1 - \delta) \langle \lambda, x \rangle + \delta \langle \lambda, x' \rangle. \end{aligned}$$

which implies that $\langle \lambda, x' \rangle > \langle \lambda, x \rangle$ as desired. □

Remark: Lemma 6 allows one to define local-optimality with respect to deviations induced by simple cycles. Local-optimality based on backtrackless paths of arbitrary length h decouples the definition of local-optimality from the girth of the Tanner graph. The implication of this decoupling on iterative decoding is discussed in Section 5.

3.2 Local-Optimality Implies LP-Optimality

In the following section we show that local-optimality is sufficient for LP-optimality (Theorem 11). We consider graph cover decoding introduced by Vontobel and Koetter [VK05] and

its extension to Tanner codes [Hal11, Chapter 2.6]. The proof of Theorem 11 is based on Lemma 10 that states that local-optimality is preserved under lifting to any M -cover graph.

The presentation in this section uses the terms and notation of Vontobel and Koetter [VK05] (see also [Hal11, Chapter 2.6]). Let \tilde{G} denote an M -cover of G . Let $\tilde{x} = x^{\uparrow M} \in \mathcal{C}(\tilde{G})$ and $\tilde{\lambda} = \lambda^{\uparrow M} \in \mathbb{R}^{N \cdot M}$ denote the M -lifts of x and λ , respectively.

Proposition 9 (local-optimality of the all-zero codeword is preserved by M -lifts). *0^N is h -locally optimal codeword w.r.t. $\lambda \in \mathbb{R}^N$ if and only if $0^{N \cdot M}$ is h -locally optimal codeword w.r.t. $\tilde{\lambda}$.*

Proof. Consider the surjection φ of paths with length h in \tilde{G} to paths in G . This surjection is based on the covering map between \tilde{G} and G . Given a path $\tilde{\psi}$ in \tilde{G} , let $\psi \triangleq \varphi(\tilde{\psi})$. Let $\tilde{\beta} \triangleq \frac{1}{h+1} \chi_{\tilde{G}}(\tilde{\psi})$ and $\beta \triangleq \frac{1}{h+1} \chi_G(\psi)$. The proposition follows because $\langle \lambda, \beta \rangle = \langle \tilde{\lambda}, \tilde{\beta} \rangle$. \square

The following lemma states that local-optimality is preserved by lifting to an M -cover.

Lemma 10. *x is h -locally optimal w.r.t. λ if and only if \tilde{x} is h -locally optimal w.r.t. $\tilde{\lambda}$.*

Proof. Assume that \tilde{x} is h -locally optimal codeword w.r.t. $\tilde{\lambda}$. By Corollary 5, $0^{N \cdot M}$ is h -locally optimal w.r.t. $(-1)^{\tilde{x}} * \tilde{\lambda}$. By Proposition 9, 0^N is h -locally optimal w.r.t. $(-1)^x * \lambda$. By Corollary 5, x is h -locally optimal w.r.t. λ . Each of these implications is necessary and sufficient, and the lemma follows. \square

The following theorem is obtained as a corollary of Theorem 8 and Lemma 10. The proof is based on arguments utilizing properties of graph cover decoding. Those arguments are used for a reduction from ML-optimality to LP-optimality [HE11, Theorem 8].

Theorem 11 (local-optimality is sufficient for LP optimality). *If x is a h -locally optimal codeword w.r.t. λ , then x is also the unique optimal LP solution given λ .*

Proof. Suppose that x is h -locally optimal codeword w.r.t. $\lambda \in \mathbb{R}^N$. By [VK05, Proposition 10], for every basic feasible solution $z \in [0, 1]^N$ of the LP, there exists an M -cover \tilde{G} of G and an assignment $\tilde{z} \in \{0, 1\}^{N \cdot M}$ such that $\tilde{z} \in \mathcal{C}(\tilde{G})$ and $z = \zeta(\tilde{z})$, where $\zeta(\tilde{z})$ is the image of the scaled projection of \tilde{z} in G (i.e., the pseudo-codeword associated with \tilde{z}). Moreover, since the number of basic feasible solutions is finite, we conclude that there exists a finite M -cover \tilde{G} such that every basic feasible solution of the LP admits a valid assignment in \tilde{G} .

Let z^* denote an optimal LP solution given λ . Without loss of generality z^* is a basic feasible solution. Let $\tilde{z}^* \in \{0, 1\}^{N \cdot M}$ denote the 0 – 1 assignment in the M -cover \tilde{G} that corresponds to $z^* \in [0, 1]^N$. By [VK05, Proposition 10] and the optimality of z^* it follows that \tilde{z}^* is a codeword in $\mathcal{C}(\tilde{G})$ that minimizes $\langle \tilde{\lambda}, \tilde{z} \rangle$ for $\tilde{z} \in \mathcal{C}(\tilde{G})$, namely \tilde{z}^* is the ML codeword in $\mathcal{C}(\tilde{G})$ w.r.t. $\tilde{\lambda}^{\uparrow M}$.

Let $\tilde{x} = x^{\uparrow M}$ denote the M -lift of an h -locally optimal codeword x . Note that because x is a codeword, i.e., $x \in \{0, 1\}^N$, there is a unique pre-image of x in \tilde{G} , which is the M -lift of x . Lemma 10 implies that \tilde{x} is h -locally optimal codeword w.r.t. $\tilde{\lambda}$. By Theorem 8, we also get that \tilde{x} is the ML codeword in $\mathcal{C}(\tilde{G})$ w.r.t. $\tilde{\lambda}^{\uparrow M}$. Moreover, Theorem 8 guarantees the uniqueness of an ML optimal solution. Thus, $\tilde{x} = \tilde{z}^*$. Because $\tilde{x} = \tilde{z}^*$, when projected to G , we get that $x = z^*$ and uniqueness follows, as required. \square

4 Probabilistic Analysis of Path-Based Local-Optimality

In the previous section, we showed that LP-decoding succeeds if a locally-optimal codeword exists w.r.t. the received LLR. In this section we analyze the probability that a locally-optimal codeword exists for even Tanner codes in MBIOS channels. The following equation justifies the all-zero codeword assumption for analyses based on local-optimality characterizations.

$$\begin{aligned}
\Pr\{\text{LP decoding fails}\} &\triangleq \Pr\{x \neq \hat{x}^{\text{LP}}(\lambda) \mid c = x\} \\
&\stackrel{(1)}{\leq} \Pr\{x \text{ is not } h\text{-locally optimal w.r.t. } \lambda \mid c = x\} \\
&\stackrel{(2)}{=} \Pr\{0^N \text{ is not } h\text{-locally optimal w.r.t. } (-1)^x * \lambda \mid c = x\} \\
&\stackrel{(3)}{=} \Pr\{0^N \text{ is not } h\text{-locally optimal w.r.t. } \lambda \mid c = 0^N\} \\
&\stackrel{(4)}{=} \Pr\{\exists \beta \in \mathcal{B}^{(h)} \text{ such that } \langle \lambda, \beta \rangle \leq 0 \mid c = 0^N\}. \tag{10}
\end{aligned}$$

Inequality (1) is the contrapositive statement of Theorem 11. Equality (2) follows Corollary 5. For MBIOS channels, $\Pr(\lambda_i \mid c_i = 0) = \Pr(-\lambda_i \mid c_i = 1)$. Therefore, the mapping $(x, \lambda) \mapsto (0^N, b * \lambda)$ where $b_i \triangleq (-1)^{x_i}$ is a measure preserving mapping. Equality (3) follows by applying this mapping to $(x, b * \lambda) \mapsto (0^N, b * b * \lambda)$. Equality (4) follows by the definition of path-based local-optimality.

Following Equation (10), our goal is to prove an upper bound on the probability that there exists a path of length h in G whose normalized characteristic vector has non-positive cost w.r.t. λ .

We use the following notation. For a path ψ , the *normalized cost of ψ w.r.t. λ* is defined by $\text{cost}(\psi) \triangleq \langle \lambda, \chi_G(\psi) \rangle$. Let $d_L^{\min} \triangleq \min\{\deg_G(v) \mid v \in \mathcal{V}\}$, $d_L^{\max} \triangleq \max\{\deg_G(v) \mid v \in \mathcal{V}\}$, and $d_R^{\max} \triangleq \max\{\deg_G(C) \mid C \in \mathcal{J}\}$. Let $D \triangleq d_L^{\max} \cdot d_R^{\max}$.

The intuition of the analysis is that long simple paths are unlikely to have non-positive cost. We restrict the path length by $h < \text{girth}(G)$ only for the probabilistic analysis. Tanner graphs with logarithmic girth can be constructed explicitly (see e.g. [Gal63]). In particular, we assume that $\text{girth}(G) > \log_D(N)$.

The following theorem presents an analytical bound on the word error probability of the LP decoder over the BSC.

Theorem 12. *Let $\mathcal{C}(G)$ denote an even Tanner code of length N such that $g \triangleq \log_D(N) < \text{girth}(G)$. Consider a BSC with crossover probability p . For any $\epsilon > 0$, if $p < D^{-2 \cdot (1 + \frac{d_L^{\min}}{d_L^{\max}}) \cdot (\epsilon + \frac{3}{2} + \frac{1}{2} \log_D(2))}$, then the LP decoder fails to decode the transmitted codeword with a probability of at most $N^{-\epsilon}$.*

For example, for left-regular codes (i.e., $d_L^{\max} = d_L^{\min}$), set $\epsilon = \frac{1}{2}$. If $p \leq \frac{D^{-8}}{4}$, then the word error probability of LP-decoding is at most $\frac{1}{\sqrt{N}}$.

Proof. By Equation (10) we may assume that the all-zero codeword is transmitted, i.e., $c = 0^N$. Hence $\lambda_v = 1$ w.p. $(1 - p)$ and $\lambda_v = -1$ w.p. p . We bound the word error probability P_w using a union bound over all the events where simple paths of length g have non-positive normalized cost in G w.r.t. λ .

Let ψ be a particular path of length g ; ψ contains $\frac{g}{2}$ variable nodes. Each variable node in the path is assigned $+1$ with probability $(1 - p)$ and -1 with probability p . Because of the

degree normalization, at least $\frac{g}{2} \cdot \frac{d_L^{\min}}{d_L^{\min} + d_L^{\max}}$ of the variable nodes in ψ must be assigned -1 to obtain $\text{cost}(\psi) \leq 0$. Let $\delta \triangleq \frac{d_L^{\min}}{d_L^{\min} + d_L^{\max}}$. Therefore,

$$\Pr\{\text{cost}(\psi) \leq 0\} \leq \left(\frac{g}{2}\right) p^{\frac{g}{2} \cdot \delta} \leq 2^{\frac{g}{2}} \cdot p^{\frac{g}{2} \cdot \delta}. \quad (11)$$

There are at most $|\mathcal{V}| \cdot D^{\frac{g}{2}}$ different simple paths of length g in G . By the union bound,

$$\begin{aligned} P_w &\leq |\mathcal{V}| \cdot D^{\frac{g}{2}} \cdot 2^{\frac{g}{2}} \cdot p^{\frac{g}{2} \cdot \delta} \\ &\leq N \cdot D^{\frac{1}{2} \log_D(N)} \cdot 2^{\frac{1}{2} \log_D(N)} \cdot p^{\frac{1}{2} \log_D(N) \cdot \delta} \\ &\leq N \cdot N^{\frac{1}{2}} \cdot N^{\frac{1}{2} \log_D(2)} \cdot N^{\frac{1}{2} \log_D(p) \cdot \delta} \\ &= N^{\frac{3}{2} + \frac{1}{2} \log_D(2) + \frac{1}{2} \log_D(p) \cdot \delta} \leq N^{-\epsilon}. \end{aligned} \quad (12)$$

□

For the case of BI-AWGN channel, we derive a bound on the word error probability for left-regular even Tanner codes, i.e., $d_L = d_L^{\min} = d_L^{\max}$. The extension to the case of irregular even Tanner codes requires exhaustive notation and computations.

Theorem 13. *Let $\mathcal{C}(G)$ denote a left-regular even Tanner code of length N such that $g = \log_D(N)$ where $D \triangleq d_L \cdot d_R^{\max}$. Consider a BI-AWGN channel with variance σ^2 . For any $\epsilon > 0$, if $\sigma^2 < \frac{\log_D(e)}{6+4\epsilon}$, then the LP decoder fails to decode the transmitted codeword with a probability of at most $\frac{\sigma}{\sqrt{\pi \log_D(N)}} \cdot N^{-\epsilon}$.*

Proof. By Equation (10) we assume that the all-zero codeword is transmitted, i.e., $c = 0^N$. Hence $\lambda_v = 1 + \phi_i$ where $\phi_i \sim \mathcal{N}(0, \sigma^2)$ is a zero-mean Gaussian random variable with variance σ^2 . We bound the word error probability P_w using a union bound over all the events where simple paths of length g have non-positive cost in G .

Let ψ be a particular path of length g . If the sum of the costs of the variable nodes in ψ is non-positive, then $\text{cost}(\psi) \leq 0$. Hence,

$$\begin{aligned} \Pr\{\text{cost}(\psi) \leq 0\} &= \Pr\left\{\sum_{i=1}^{\frac{g}{2}} (1 + \phi_i) \leq 0\right\} \\ &= \Pr\left\{\sum_{i=1}^{\frac{g}{2}} \phi_i \leq -\frac{g}{2}\right\}. \end{aligned} \quad (13)$$

The sum of independent Gaussian random variables (RVs) with zero mean is a zero-mean Gaussian RV whose variance equals to the sum of the variances of the accumulated variables. Let $\Phi = \sum_{i=1}^{\frac{g}{2}} \phi_i$, then $\Phi \sim \mathcal{N}(0, \sigma^2 \cdot \frac{g}{2})$ is a zero-mean Gaussian RV with variance $\sigma^2 \cdot \frac{g}{2}$. Moreover, the Gaussian distribution function is symmetric around 0. Therefore,

$$\begin{aligned} \Pr\{\text{cost}(\psi) \leq 0\} &= \Pr\left\{\Phi \leq -\frac{g}{2}\right\} \\ &= \Pr\left\{\Phi \geq \frac{g}{2}\right\}. \end{aligned} \quad (14)$$

For a Gaussian RV $\phi \sim \mathcal{N}(0, \sigma^2)$ with zero mean and variance σ^2 , the inequality

$$\Pr\{\phi \geq x\} \leq \frac{\sigma}{x\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad (15)$$

holds for every $x > 0$ [Fel68]. We conclude that

$$\Pr\{\text{cost}(\psi) \leq 0\} \leq \frac{\sigma}{\sqrt{\pi g}} e^{-\frac{g}{4\sigma^2}}. \quad (16)$$

There are at most $|\mathcal{V}| \cdot D^{\frac{g}{2}}$ different simple paths of length g in G . By the union bound,

$$\begin{aligned} P_w &\leq |\mathcal{V}| \cdot D^{\frac{g}{2}} \cdot \Pr\{\text{cost}(\psi) \leq 0\} \\ &\leq N \cdot D^{\frac{g}{2}} \frac{\sigma}{\sqrt{\pi g}} e^{-\frac{g}{4\sigma^2}} \\ &\leq N^{\frac{3}{2}} \cdot \frac{\sigma}{\sqrt{\pi \log_D(N)}} e^{-\frac{\log_D(N)}{4\sigma^2}} \\ &= \frac{\sigma}{\sqrt{\pi \log_D(N)}} N^{\frac{3}{2} - \frac{1}{4\sigma^2} \log_D(e)} \\ &\leq \frac{\sigma}{\sqrt{\pi \log_D(N)}} \cdot N^{-\epsilon}. \end{aligned} \quad (17)$$

□

5 Discussion

5.1 Finding Locally-Optimal Codewords with Message-Passing Algorithm

A message-passing decoding algorithm, called normalized weighted min-sum (NWMS), was presented in [EH11] for Tanner codes with single parity-check (SPC) local codes. The NWMS decoder is guaranteed to compute the ML-codeword in h iterations provided that a locally-optimal codeword with height h exists. The local-optimality characterization in [EH11] is stronger and is based on subtrees of computation trees of the Tanner graph. We note that for even Tanner codes with SPC local codes, the NWMS decoder with uniform weights (i.e., $w = 1^h$) is guaranteed to compute the ML-codeword in h iterations provided that a path-based h -locally optimal codeword exists. The number of iterations h may exceed the girth of the Tanner graph.

Consider an even Tanner code and an LLR vector λ . Assume that there exists a codeword x that is h -locally optimal w.r.t. λ . Given this assumption, maximum-likelihood decoding is equivalent to solving the following problem: Find an assignment $x \in \{0, 1\}^N$ to variable nodes such that every path of length h in G w.r.t. vertex weights $(-1)^{x_v} \cdot \lambda_v / \deg_G(v)$ has positive weight. By Theorem 11, LP-decoding computes such a valid assignment x . The iterative NWMS decoding algorithm also computes such a valid assignment x provided that the local codes are restricted to SPC codes.

5.2 Punctured Tanner Codes

Local-optimality characterization and its analysis remain valid under puncturing of a codeword. Puncturing of a code is specified by a subset of variable nodes that are not transmitted; this subset is called the *puncturing pattern*. A punctured code can be analyzed simply by zeroing the LLR values of the punctured variable nodes.

How does puncturing affect the probability that a locally-optimal codeword exists? Consider a probabilistic analysis for local-optimality that is based on a simple union bound over the set of deviations of a Tanner code. We say that a variable node v *participates* in deviation β if $\beta_v \neq 0$. Suppose that the puncturing pattern is chosen so that for every deviation β , at most a constant fraction α of participating variable nodes are punctured. Then, the same union bound analysis can be applied directly to the punctured Tanner code. Such an analysis implies bounds that are similar to the bounds obtained for the (unpunctured) Tanner code; the only difference is that some parameters need to be scaled by a constant factor that is proportional to α .

5.3 Analysis of Repeat-Accumulate Codes via Local-Optimality

Feldman and Karger [FK02, FK04] introduced the concept of linear-programming based decoding for repeat-accumulate RA(q) turbo-like codes. In an RA(q) codes: (1) an information word is repeated q times, (2) the repeated information word is permuted by an interleaver, and (3) codeword bit i equals to the parity of the sum of the first i bits (prefix) of the permuted repeated information word. For RA(2) codes over the BSC, they proved that the word error probability of LP-decoding is bounded by an inverse polynomial in the code length, given a certain constant threshold on the noise. A similar claim was also proved for the BI-AWGN channel. These bounds were further improved based on exact combinatorial characterization of an error event and a refined algorithmic analysis [HE05].

Recently, Goldenberg and Burshtein [GB11] generalized the analysis of Feldman and Karger [FK04] to RA(q) codes with even repetition $q \geq 4$. For this family of codes, they proved inverse polynomial bounds in the code length on the word error probability of LP-decoding. These bounds are based on analyzing complicated graph structures, called hyper-promenades in hypergraphs, that were defined in [FK04].

In fact, the same bounds can be obtained by local-optimality with deviations induced by short paths. The idea is to consider (non-systematic) irregular repeat-accumulate codes with even repetition factors as punctured Tanner codes (as illustrated in Figure 1). Namely, set the LLR of each systematic variable node to zero. Notice that, for every backtrackless path in Tanner graphs of repeat accumulate codes, at most half of the variable nodes are systematic. Therefore, at most half of the variable nodes in each deviation of path-based local-optimality are punctured. We can now analyze the word error probability by replacing $\frac{q}{2}$ by $\frac{q}{4}$ in the proofs of Theorems 12 and 13. Hence the unified technique of local-optimality with deviations induced by short paths yields the same results proved by Feldman and Karger [FK04] and Goldenberg and Burshtein [GB11]. Improving these bounds on successful decoding of RA(q) codes (for $q \geq 3$) remains an intriguing open question.

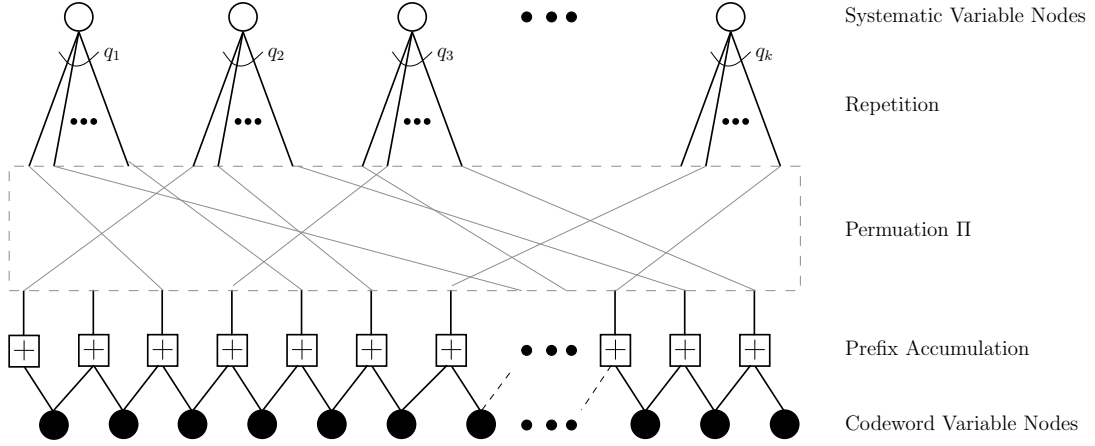


Figure 1: Tanner graph representation of an irregular repeat-accumulate code.

The variable nodes in the top row of nodes (illustrated by white circles) correspond to the k systematic bits. The repetition of every systematic bit corresponds to the degree q_i of each systematic variable node (we assume that q_i is even for every i). The interleaving process corresponds to the permutation of edges. The prefix accumulation corresponds to the alternating chain of single-parity check code nodes (illustrated by plus-squares) and codeword variable nodes (denoted by black circles) in the bottom.

6 Conclusions

We present a simple application of the proof technique in [ADS09] for bounds on word error probability with LP-decoding. The set of deviations used for defining local-optimality is induced by paths in the Tanner graph. We apply the proof technique to a family of codes, called even Tanner codes, that contains repeat accumulate codes with a even repetition factors, LDPC codes with even left degrees, and expander codes with even variable node degrees and even weighted local codes. Inverse polynomial error bounds are proved for these codes for the BSC and AWGN channel.

Stronger error bounds have been obtained for LDPC codes [KV06, ADS09, HE11] and Tanner codes [EH11] (without the restriction to even degrees and even weighted local codes) by considering more complicated graphical structures and a more sophisticated analysis. In these cases, inverse exponential error bounds and improved bounds on noise thresholds were presented for regular codes whose Tanner graphs have logarithmic girth. For example, the local-optimality characterization for Tanner codes presented in [EH11] is based on projections of weighted subtrees in computation trees of the Tanner graph. The error bounds in this case are based on an analysis of a sum-min-sum random process on trees. While this local-optimality characterization applies to any regular and irregular Tanner code, the probabilistic analysis and the error bounds were restricted to regular Tanner codes. The simplicity of local-optimality based on paths enables us to obtain (weak) bounds even for irregular codes. Two interesting open questions related to proving stronger bounds on the error probability are (i) extend the analysis of inverse exponential bounds to irregular Tanner graphs, and (ii) obtain bounds with respect to local-optimality even “beyond the girth”.

References

- [ADS09] S. Arora, C. Daskalakis, and D. Steurer, “Message passing algorithms and improved LP decoding,” in *Proc. of the 41st annual ACM Symp. Theory of Computing (STOC’09)*, Bethesda, MD, USA, May 31 - June 02, 2009, pp. 3–12.
- [BMvT78] E. Berlekamp, R. McEliece, and H. van Tilborg, “On the inherent intractability of certain coding problems,” *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 384–386, May 1978.
- [DDKW08] C. Daskalakis, A. G. Dimakis, R. M. Karp, and M. J. Wainwright, “Probabilistic analysis of linear programming decoding,” *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3565–3578, Aug. 2008.
- [EH11] G. Even and N. Halabi, “On decoding irregular Tanner codes with local-optimality guarantees,” *CoRR*, <http://arxiv.org/abs/1107.2677>, Jul. 2011.
- [Fel68] W. Feller, *An introduction to probability theory and its applications*, 3rd ed., New York: Wiley, 1968, vol. I.
- [Fel03] J. Feldman, “Decoding error-correcting codes via linear programming,” Ph.D. dissertation, MIT, Cambridge, MA, 2003.
- [FK02] J. Feldman and D. R. Karger, “Decoding turbo-like codes via linear programming” in *Proc. 43rd Annu. IEEE Symp. Foundations of Computer Science (FOCS)*, Nov. 2002.
- [FK04] J. Feldman and D. R. Karger, “Decoding turbo-like codes via linear programming,” *J. Comput. Syst. Sci.*, vol. 68, no. 4, pp. 733–752, June 2004.
- [FMS⁺07] J. Feldman, T. Malkin, R. A. Servedio, C. Stein, and M. J. Wainwright, “LP decoding corrects a constant fraction of errors,” *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 82–89, Jan. 2007.
- [FS05] J. Feldman and C. Stein, “LP decoding achieves capacity,” in *Proc. Symp. Discrete Algorithms (SODA’05)*, Vancouver, Canada, Jan. 2005, pp. 460–469.
- [FWK05] J. Feldman, M. J. Wainwright, and D. R. Karger, “Using linear programming to decode binary linear codes,” *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 954–972, Mar. 2005.
- [Gal63] R. G. Gallager, *Low-Density Parity-Check Codes*. MIT Press, Cambridge, MA, 1963.
- [GB11] I. Goldenberg and D. Burshtein, “Error bounds for repeat-accumulate codes decoded via linear programming,” *Adv. Math. Commun.*, vol. 5, no. 4, pp. 555–570, Nov. 2011.
- [Hal11] N. Halabi, “On decoding Tanner codes with local-optimality guarantees,” Ph.D. dissertation, submitted, Tel-Aviv University.

- [HE05] N. Halabi and G. Even, “Improved bounds on the word error probability of RA(2) codes with linear-programming-based decoding,” *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 265–280, Jan. 2005.
- [HE11] N. Halabi and G. Even, “LP decoding of regular LDPC codes in memoryless channels,” *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 887–897, Feb. 2011.
- [KV06] R. Koetter and P. O. Vontobel, “On the block error probability of LP decoding of LDPC codes,” in *Proc. Inaugural Workshop of the Center for Information Theory and its Applications*, La Jolla, CA, USA, Feb. 2006.
- [Ska11] V. Skachek, “Correcting a Fraction of Errors in Nonbinary Expander Codes With Linear Programming,” *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3698–3706, Jun. 2011.
- [VK05] P. O. Vontobel and R. Koetter, Graph-cover decoding and finite-length analysis of message-passing iterative decoding of LDPC codes, *CoRR*, <http://www.arxiv.org/abs/cs.IT/0512078>, Dec. 2005.
- [Von10] P. Vontobel, “A factor-graph-based random walk, and its relevance for LP decoding analysis and Bethe entropy characterization,” in *Proc. Information Theory and Applications Workshop*, UC San Diego, LA Jolla, CA, USA, Jan. 31-Feb. 5, 2010.
- [Wib96] N. Wiberg, “Codes and decoding on general graphs”, Ph.D. dissertation, Linköping University, Linköping, Sweden, 1996.