Convergence of Weighted Min-Sum Decoding Via Dynamic Programming on Coupled Trees

Yung-Yih Jian and Henry D. Pfister
Department of Electrical and Computer Engineering, Texas A&M University
Email: {yungyih.jian,hpfister}@tamu.edu

***Abstract*—Applying the max-product (and belief-propagation) algorithms to loopy graphs is now quite popular for constraint satisfaction problems. This is largely due to their low computational complexity and impressive performance in practice. Still, there is no general understanding of the conditions required for convergence and/or the optimality of converged solutions. This paper presents an analysis of weighted min-sum (a.k.a. attenuated max-product) decoding for LDPC codes that guarantees convergence to a fixed point when the weight β is sufficiently small. It also shows that, if the fixed point satisfies all the constraints, then it must be both the linear-programming (LP) and maximum-likelihood (ML) solution. For (d_v, d_c)-regular LDPC codes, the weight must satisfy $1/\beta > d_v - 1$ whereas the result of Koetter and Frey requires instead that $1/\beta > (d_v - 1)(d_c - 1)$. A counterexample is also given that shows a fixed point might not be the ML solution if $1/\beta < d_v - 1$. Finally, connections are explored with recent work by Arora et al. on the threshold of LP decoding.**

***Index Terms*—belief propagation, max product, min sum, LDPC codes, linear programming decoding**

I. Introduction

The introduction of turbo codes in 1993 started a revolution in coding and inference that continued with rediscovery of low-density parity-check (LDPC) codes and culminated in optimized LDPC codes that essentially achieve the capacity of practical channels [1], [2], [3], [4]. During this time, Wiberg *et al.* advanced the analysis of iterative decoding by proving a number of results for the min-sum (a.k.a. max-product) decoding algorithm [5], [6], [7]. Richardson and Urbanke also introduced the technique of density evolution (DE) to compute the noise threshold of message-passing decoders for turbo and LDPC codes [8].

For a particular noise realization, the optimality of iterative decoding solution has also been considered by a number of authors. Weiss and Freeman have shown that the max-product (MP) assignment is locally optimal w.r.t. all single-loop and tree perturbations [9]. Unfortunately, this result is typically uninformative for LDPC codes with variables degrees larger than 2. Koetter and Frey have also shown that, with the proper weights and adjustments, the weighted min-sum (WMS) (a.k.a. attenuated max-product) decoder for LDPC codes returns the maximum-likelihood (ML) codeword if it converges to a codeword [7]. For general graphs, Wainwright *et al.* proposed the tree-reweighted max-product message-passing algorithm for computing the MAP assignment on the graph [10]. They have shown that, under some conditions, the fixed point solution gives an optimal configuration for the graph. Their algorithm, though strictly different, has some similarity to the WMS algorithm in [7]. More recently, Arora *et al.* have shown that any codeword satisfying a local optimality condition must also be the globally optimal solution to linear-programming (LP) decoding problem [11].

This work was supported in part by the National Science Foundation under Grant No. 0747470 and the Texas Norman Hackerman Advanced Research Program under Grant No. 000512-0168-2007. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

The results in this paper can be seen as an extension of the work by Koetter and Frey that provides new insight into results of [11]. We view the WMS algorithm as computing the dynamic-programming solution to the optimal discounted-reward problem on a set of coupled trees. This allows us to show that, for any received vector, the one-step update of the WMS algorithm is a contraction on the space of message values when weight parameter is sufficiently small. From this, we deduce that the algorithm converges to unique fixed point. If the resulting fixed point satisfies some consistency conditions, then it must also be the LP optimum solution and, hence, the ML solution.

II. Convergence of Min-Sum Decoding

A. Background

Consider a (d_v, d_c)-regular LDPC code over a bipartite graph $G = (V_L \cup V_R, E)$ whose vertex set consists of variable nodes V_L and check nodes V_R. Let n be the number of variable nodes in G and $N(i)$ be the set of nodes that are neighbors of $i \in G$. Let $\mathcal{T}_i^I$ be the computation tree of depth I rooted at bit $i \in V_L$ (e.g., for $I = 2T$, this tree contains $T+1$ variable node levels) and $N(i, \ell)$ be the set of vertices in the ℓth level of $\mathcal{T}_i^I$ for $\ell \leq I$.

For an LDPC code whose Tanner graph is a tree, the min-sum algorithm is optimal and can be seen as the dynamic programming (DP) solution to the optimal assignment of binary values to edges in computation tree. When the number of iterations is increased beyond half the girth of the Tanner graph, the min-sum algorithm continues to compute the optimal assignment on the coupled set of computation trees (i.e., there is one computation tree for each directed edge). Let $\mu_{i\to j}^{(\ell)}(x)$ be the correlation between the received log-likelihood ratios (LLRs) and the best valid computation tree, of height 2ℓ, whose root node is assigned x. In this case, we find that the min-sum message, $\mu_{i\to j}^{(\ell)} = \mu_{i\to j}^{(\ell)}(0) - \mu_{i\to j}^{(\ell)}(1)$, simply the difference between best 0-root correlation and the best 1-root correlation.

For the ith received bit y_i, let $\gamma_i = \ln \frac{p_{Y|X}(y_i|0)}{p_{Y|X}(y_i|1)}$ be LLR and β_i be the weight factor. Let $\mu_{i\to j}^{(\ell)}$ be the min-sum message, from variable $i \in V_L$ to check $j \in V_R$, in the ℓth iteration. For any WMS message vector $\mu \in \mathbb{R}^{|E|}$, the one-iteration update operator $T : \mathbb{R}^{|E|} \to \mathbb{R}^{|E|}$ is defined by $\nu = T[\mu]$ with

$$\nu_{i\to j} = \gamma_i + \beta_i \sum_{k\in N(i)\setminus j} \left(\prod_{m\in N(k)\setminus i} \operatorname{sgn}(\mu_{m\to k}) \right) \times \min_{m'\in N(k)\setminus i} |\mu_{m'\to k}| .$$

The algorithm is initialized by setting $\mu_{i\to j}^{(0)} = \gamma_i$ and proceeds iteratively by computing $\mu^{(\ell+1)} = T\left[\mu^{(\ell)}\right]$. When needed, the message from check $j \in V_R$ to variable $i \in V_L$, in the ℓth iteration is denoted $\mu_{i\leftarrow j}^{(\ell)}$ and is defined by

$$\mu_{i\leftarrow j}^{(\ell)} = \left(\prod_{m\in N(j)\setminus i} \operatorname{sgn}\left(\mu_{m\to j}^{(\ell)}\right) \right) \min_{m'\in N(j)\setminus i} \left|\mu_{m'\to j}^{(\ell)}\right| .$$

The following generalizes definitions from [5] and [11].

Definition 1: A bit assignment u on $\mathcal{T}_{i_0}^{2T}$ is a *valid deviation* of depth T at $i_0 \in V_L$ or, in short, a *T-local deviation* at i_0, if $u_{i_0} = 1$ and u satisfies all parity checks in $\mathcal{T}_{i_0}^{2T}$. Moreover, u is a *minimal T-local deviation* if, for every check node $j \in \mathcal{T}_{i_0}^{2T}$, at most two neighboring bits are assigned the value 1. Note that a minimal T-local deviation at i_0 can be seen as a subtree of $\mathcal{T}_{i_0}^{2T}$ of height $2T$ rooted at i_0, where every variable node has full degree and every check node has degree 2. Such a tree is referred as a *skinny tree.* If $w = (w_0, \ldots, w_T) \in [0,1]^T$ is a weight vector and u is a minimal T-local deviation at i_0, then $u^{(w)}$ denotes the w-*weighted* deviation

$$u_i^{(w)} = \begin{cases} w_t u_i & \text{if } i \in N(i_0, 2t) \text{ and } 0 \le t \le T, \\ 0 & \text{otherwise.} \end{cases}$$

We say that a bit $i' \in \mathcal{T}_{i_0}^{2T}$ in the computation tree is associated with the bit $i \in V_L$ in the original code (denoted $i' \sim i$) if i' is a copy of i. For any weighted minimal T-local deviation $u^{(w)}$ on $\mathcal{T}_{i_0}^{2T}$, let the projection of $u^{(w)}$ onto the code bit $i \in V_L$ be

$$\pi_i\left(u^{(w)}\right) = \sum_{t=0}^{T} w_t \sum_{m\in N(i_0,2t):\, m\sim i} u_m .$$

Likewise, we let $\pi\left(u^{(w)}\right)$ represent the vector whose elements are $\pi_i\left(u^{(w)}\right)$ for $i \in V_L$. Throughout this paper, the weights are generally chosen to be $w_t = \beta^t$ for some $\beta \in [0,1]$.

In this section, we show that, if the WMS algorithm with $\beta < \frac{1}{d_v-1}$ converges to a set of consistent messages, then the resulting codeword is the LP optimal and, hence, ML codeword.

Definition 2: Consider the WMS algorithm for a (d_v, d_c)-regular LDPC code. In the ℓth iteration, let $\mu_{i\to j}$ be the message passed from the ith bit to the jth check, and $\mu_{i\leftarrow j}$ be the message passed from jth check to ith bit. The messages are called *consistent* if, for each bit $i \in V_L$, we have

1. $\operatorname{sgn}(\mu_{i\to j}) = \operatorname{sgn}(\mu_{i\to j'})$ for $j, j' \in N(i)$
2. $\operatorname{sgn}(\mu_{i\leftarrow j}) = \operatorname{sgn}(\mu_{i\to j})$ for $j \in N(i)$
3. $\operatorname{sgn}\Big(\gamma_i + \beta \sum_{j\in N(i)} \mu_{i\leftarrow j}\Big) = \operatorname{sgn}(\mu_{i\to j'})$ for $j' \in N(i)$.

B. Impossibility of a General ML Certificate

In general, if the WMS algorithm converges to a codeword for $\beta > \frac{1}{d_v-1}$, it may not be the ML codeword. For $(3,4)$ codes, we have found examples supporting this statement for $\beta = 0.6$ via computer simulation of a short LDPC code. For the general, case, the following example gives some intuition.

Example 1: Consider a (d_v, d_c)-regular LDPC code with codeword length n, where d_c is an odd number and $d_v > 3$. Assume an all-zeros codeword is transmitted. Let the channel output LLR be $\gamma = (-1, \ldots, -1)$. Consider the WMS algorithm with $\beta > \frac{2}{d_v-1}$.

At the beginning, all messages from variable nodes to their neighbor check nodes are $\mu_{i\to j}^{(0)} = -1$ for $i = 1, \ldots, n$ and $j \in N(i)$. Consider the message passed from the jth check nodes to its neighbor variable nodes, $\mu_{i\leftarrow j},\ i \in N(j)$. Since the incoming messages are all equal to –1, the update rule of the WMS algorithm at the check node gives

$$\mu_{i\leftarrow j}^{(0)} = \left(\prod_{k\in N(j)\setminus i} \operatorname{sgn}(\mu_{k\to j}) \right) \min_{k'\in N(j)\setminus i} \left|\mu_{k'\to j}^{(0)}\right| = 1,$$

for all $(i,j) \in E$. In the first iteration, the outgoing message from the ith variable node to jth check node is therefore

$$\mu_{i\to j}^{(1)} = \gamma_i + \beta \sum_{k\in N(i)\setminus j} \mu_{i\leftarrow k}^{(0)} > -1 + (d_v - 1)\frac{2}{d_v - 1} = 1.$$

Moreover, one can show that $\mu_{i\to j}^{(\ell)} \to \infty$ as $\ell \to \infty$. Thus, the hard decision output is an all-zero codeword. Unfortunately, given this γ, we know that the ML output must be a nonzero codeword with maximal Hamming weight. Therefore, the WMS decoder cannot provide an ML certificate for $\beta > \frac{2}{d_v-1}$. One might worry that this effect may be related ties between ML codewords, but these can be avoided, without affecting the above result, by adding a very small amount of uniform random noise to the received LLRs.

C. Convergence and Optimality

The main result is based on the following Lemma.

Lemma 1: If we define

$$d_{i,k} \triangleq \left| \left(\prod_{m\in N(k)\setminus i} \operatorname{sgn}(\mu_{m\to k}) \right) \min_{m'\in N(k)\setminus i} |\mu_{m'\to k}| - \left(\prod_{m\in N(k)\setminus i} \operatorname{sgn}(\nu_{m\to k}) \right) \min_{m'\in N(k)\setminus i} |\nu_{m'\to k}| \right|,$$

then we have

$$\max_{m\in N(k)\setminus i} |\mu_{m\to k} - \nu_{m\to k}| \ge d_{i,k}.$$

Proof: The proof is omitted due to space constraints. ■

Theorem 1: For all LLR vectors and message values, the WMS operator T is an $\|\cdot\|_\infty$ contraction if

$$\max_i (\beta_i (|N(i)| - 1)) < 1.$$

Proof: Using Lemma 1, one can upper bound $\|T[\mu] - T[\nu]\|_\infty$ in a straightforward manner to get

$$\begin{aligned}\|T[\mu] - T[\nu]\|_\infty &\le \max_{(i,j)\in E} \beta_i \sum_{k\in N(i)\setminus j} |d_{i,k}| \\ &\le \max_i (\beta_i (|N(i)| - 1)) \max_{(i,j)\in E} |d_{i,k}| \\ &\le \max_i (\beta_i (|N(i)| - 1)) \\ &\quad \times \max_{(i,j)\in E} \max_{m\in N(k)\setminus i} |\mu_{m\to k} - \nu_{m\to k}| \\ &= \max_i (\beta_i (|N(i)| - 1)) \|\mu - \nu\|_\infty .\end{aligned}$$

This implies that T is a $\|\cdot\|_\infty$ contraction. ■

Remark 1: Combining this with the contraction mapping theorem shows that, for an arbitrary (d_v, d_c)-regular LDPC code and any $0 \le \beta < \frac{1}{d_v - 1}$, the WMS algorithm converges to a unique fixed point as the number of iterations goes to infinity. This idea is very similar to the existence proof for optimal stationary policies of discounted Markov decision processes.

Theorem 2: For a (d_v, d_c)-regular LDPC code, if in the ℓth iteration the WMS messages with weight β are consistent, then the hard decisions

$$\hat{x}_i = \frac{1}{2}\left(1 - \operatorname{sgn}\left(\gamma_i + \beta \sum_{j\in N(i)} \mu_{i\leftarrow j}^{(\ell)}\right)\right)$$

for $i = 1, \ldots, n$ give a codeword.

Proof: We prove this result by contrapositive. Assume that $\hat{x}$ is not a codeword. There exists at least one unsatisfied parity check node. Let $j \in V_R$ be the unsatisfied parity check node and $N(j)$ be the neighbors of j. Since $\bigoplus_{i\in N(j)} \hat{x}_i = 1$, we have

$$\prod_{i\in N(j)} \operatorname{sgn}\left(\gamma_i + \beta \sum_{j\in N(i)} \mu_{i\leftarrow j}^{(\ell)}\right) = \prod_{i\in N(j)} \operatorname{sgn}\left(\mu_{i\to j}^{(\ell)}\right) = -1.$$

Consider the message passed from the jth check to the ith bit. From the WMS update rule,

$$\begin{aligned}\mu_{i\leftarrow j}^{(\ell)} &= \beta \left(\prod_{m\in N(j)\setminus i} \operatorname{sgn}\left(\mu_{m\to j}^{(\ell)}\right)\right) \times \min_{m\in N(j)\setminus i} \left|\mu_{m\to j}^{(\ell)}\right| \\ &= -\beta \operatorname{sgn}\left(\mu_{i\to j}^{(\ell)}\right) \times \min_{m\in N(j)\setminus i} \left|\mu_{m\to j}^{(\ell)}\right|.\end{aligned}$$

This contradicts the consistency of the messages. ■

Consider the computation tree $\mathcal{T}_i^{2I}$ rooted at bit $i \in V_L$. For each $i \in V_L$ and root node value $\tilde{x}_i$, the WMS algorithm computes WMS-locally optimal assignments for the bits in $\mathcal{T}_i^{2I}$ assuming the root value is $\tilde{x}_i$. These assignments are not I-locally optimal in the sense of [11], however, because the optimal tree assignment may not be a codeword. To extend the results of [11] to computation trees, we first observe that, if $\beta < \frac{1}{d_v - 1}$, then the binary assignments to the leaf nodes are asymptotically irrelevant for optimality as $I \to \infty$.

Lemma 2: Given the LLR vector $\gamma \in \mathbb{R}^n$, let the assignment $\tilde{x}$, computed by the WMS decoder, for the computation tree $\mathcal{T}_{i_0}^{2I}$ be unique (i.e., there are no ties) after $I \to \infty$ iterations with $\beta < \frac{1}{d_v - 1}$. For any minimal I-local deviation $\tilde{u}$ of depth $I \gg T$ rooted at $i_0 \in V_L$, let the T-level weighted correlation be

$$U_{i_0}^T(\tilde{x}, \tilde{u}) \triangleq \sum_{i=1}^n \sum_{t=0}^T \beta^t \sum_{m\in N(i_0, 2t):\, m\sim i} (-1)^{\tilde{x}_m} \tilde{u}_m \gamma_i,$$

where $N(i_0, \ell)$ is the set of vertices in the ℓth level of $\mathcal{T}_{i_0}^{2I}$. Then, there exists a $T_0 < \infty$ such that $U_{i_0}^T(\tilde{x}, \tilde{u}) > 0$ for all $T \ge T_0$.

Proof: The proof is omitted due to space constraints. ■

The following extends the key result [11, Lemma 4] to our generalized minimal local deviations on the computation tree.

Lemma 3: Let $z \in \mathcal{P}(H)$ be a LP solution of a bit-regular code and consider the set of depth-I computation trees rooted at all non-zero variable nodes. For these trees, there exists a distribution over minimal local deviations such that the expected value, when projected onto the original Tanner graph, is proportional the LP solution z.

Proof: This fact was first observed in [12, Remark 22]. It follows from straightforward modifications to the proof of [11, Lemma 4]. ■

The following theorem shows that if the WMS messages converge to a consistent assignment, then the hard decision bits of the WMS algorithm give a codeword that is both LP optimal and ML. Following [11], we define $u \oplus v$ with

$$[u \oplus v]_i = |u_i - v_i|.$$

Theorem 3: For a given the LLR vector $\gamma \in \mathbb{R}^n$ and a weight $0 \le \beta < \frac{1}{d_v - 1}$, suppose the WMS algorithm converges to a consistent message vector. If the hard decision bits $\hat{x}$ are unique (i.e., there are no ties), then they form a T-locally optimal codeword for some $T < \infty$. Moreover, $\hat{x}$ is the LP optimal and, hence, ML codeword.

Proof: By Theorem 2, we know $\hat{x}$ is a codeword. To prove that $\hat{x}$ is a T-locally optimal codeword, we have to show that for the projection $\pi\left(u^{(w)}\right)$ of any minimal T-local deviation $u^{(w)}$, the inequality

$$\left\langle \hat{x} \oplus \left(\tfrac{1}{d_v - 1} - \beta\right) \pi\left(u^{(w)}\right), \gamma \right\rangle > \langle \hat{x}, \gamma \rangle$$

holds. Without loss of generality, we assume that $u^{(w)}$ is rooted at i_0 and consider the correlation of $\hat{x} \oplus \pi\left(u^{(w)}\right)$ and γ. Let $c = \left(\frac{1}{d_v - 1} - \beta\right)$, then this gives

$$\begin{aligned}\left\langle \hat{x} \oplus c\pi\left(u^{(w)}\right), \gamma \right\rangle &= \sum_{i=1}^n \left|\hat{x}_i - c\pi_i\left(u^{(w)}\right)\right| \gamma_i \\ &= \langle \hat{x}, \gamma \rangle + c \sum_{i=1}^n \sum_{t=0}^T \beta^t \\ &\quad \times \sum_{m\in N(i_0, 2t):\, m\sim i} (-1)^{\hat{x}_m} u_m \gamma_i \\ &= \langle \hat{x}, \gamma \rangle + c U_{i_0}^T(\hat{x}, u).\end{aligned}$$

Note that the constant c is simply required to be small enough so that $c\pi_i\left(u^{(w)}\right) \leq 1$.

To show $U_{i_0}^T(\hat{x}, u) > 0$, consider a tree $\mathcal{T}_{i_0}^I$ with large I. Since the WMS algorithm converges to a consistent message vector, the assignment for the subtree $\mathcal{T}_{i_0}^{2T}$ is the same as $\hat{x}$ for some $T < \infty$. Here, Lemma 2 is required because the leaf assignment may not match a codeword. Also, u can be obtained from the minimal valid deviation $\tilde{u}$ on $\mathcal{T}_{i_0}^{2I}$ by truncating

$$u_m = \begin{cases} \tilde{u}_m & \text{if } m \in N(i_0, 2t) \text{ for some } 0 \leq t \leq 2T \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma 2, we can conclude that $U_{i_0}^T(\hat{x}, u) > 0$. Therefore, $\hat{x}$ is a T-locally optimal codeword.

According to Theorem 4 in [11] or Theorem 6 in [13], the T-local optimality of $\hat{x}$ implies that $\hat{x}$ is the unique optimal LP solution given the LLR γ. Since $\hat{x} \in \{0,1\}^n$ is an integral codeword, $\hat{x}$ is also an ML codeword. ■

Remark 2: In [7], the condition $\beta < \frac{1}{(d_v-1)(d_c-1)}$ is defined by the threshold where the WMS optimal assignments and the resulting correlation is determined almost exactly by a finite number of levels at the "top" of the infinite computation tree. This paper allows relaxes the condition to $\beta < \frac{1}{d_v-1}$ by analyzing the difference between best 0-root assignment and the best 1-root assignment. For this larger value, the WMS optimal assignments are determined by the "top" of the computation tree but the correlations may diverge. The difference messages (i.e., LLRs) are guaranteed to converge though. If the fixed point messages are consistent, then Theorem 3 shows that the ML optimal assignment is also determined by the "top" of the computation tree.

D. Connections with LP Thresholds

The LP decoder for LDPC codes, proposed by Feldman *et al.*, solves a relaxed version of the ML decoding problem [14]. Since its introduction, a number of authors have looked for connections to the MP iterative decoding algorithm [15]. One interesting open question is, "What is the noise threshold of LP decoding?". Several lower bounds for the noise threshold have been proposed [16], [17], [18].

Arora *et al.* showed recently that, for a $(3,6)$-regular LDPC code, LP decoding can tolerate a crossover probability $p = 0.05$ on the BSC(p) [11]. They investigated the primal solution of the LP problem and proposed a local optimality condition for codewords. Since their local optimality conditions are amenable to analysis on tree-like neighborhoods, they perform a DE analysis of the WMS algorithm to obtain BSC noise thresholds for LP decoding. Using DE for memoryless binary-input output symmetric (MBIOS) channels, Halabi *et al.* showed that LP decoding can achieve a noise threshold $\sigma = 0.735$ on the BIAWGN channel[13].

In this subsection, we connect the LP threshold estimation with both converged WMS algorithm and the DE type analysis in [11], [13] based on some of our observations. We have shown that when the WMS algorithm with $\beta < \frac{1}{d_v-1}$ converges to a set of consistent messages, the WMS algorithm returns a codeword which is the same as the LP solution. If the following three conjectures are true, we can conclude that the threshold of the WMS algorithm with $\beta = \frac{1}{d_v-1}$ gives a lower bound for the LP decoding.

Conjecture 1: Consider a (d_v, d_c)-regular LDPC code and particular LLR vector $\gamma \in \mathbb{R}^n$. If the WMS algorithm diverges (i.e., the messages tend to $\pm\infty$) to consistent messages for $\beta = \frac{1}{d_v-1}$, then there is an $\epsilon > 0$ such that it also converges to consistent messages whose hard decisions give the same codeword for $\beta = \frac{1}{d_v-1} - \epsilon$. In this case, the codeword is the LP optimal and, hence, ML codeword.

Conjecture 2: Consider the WMS decoding of (d_v, d_c)-regular LDPC codes with girth $\Omega(\log n)$ over the BSC and let p^* be the bit-error rate threshold for WMS decoding with $\beta = \frac{1}{d_v-1}$. Then, the WMS decoder diverges to consistent messages with high probability for all $p < p^*$.

Remark 3: Conjecture 1 essentially states that WMS decoding, when viewed as a dynamic program, always has a Blackwell optimal policy. Though we are quite confident about this, we have not completed the full proof. Conjecture 2 has been tested via simulation and we are currently pursuing a rigorous proof.

Example 2: Consider a $(3,6)$-regular LDPC code over a BSC(p). From a DE analysis of the WMS algorithm (i.e., not the DE for local optimality proposed in [11]) with $\beta = 1/2$, we found that the WMS algorithm will decode correctly when $p \leq 0.055$.

Remark 4: In the Example 2, the LP threshold lower bound of 0.055 matches the best possible bound using techniques from [11]. The main improvement over [11] is that our analysis (under the conjectures) holds pointwise for any received sequence. It also exposes the fact that the best thresholds from [11] are associated with computation trees where the leaves are asymptotically irrelevant.

III. Numerical Results

The block error rate (BLER) for the WMS algorithms and the probability of not converging to a set of consistent messages are shown in Figure 1. The solid lines are the BLER of the WMS algorithm, and the dashed lines are the probability of non-consistency. The code for this simulation is $(3,6)$-regular LDPC code with $n = 10^4$. Two weight factors, $\beta = 0.49$ and $\beta = 0.5$, are considered, and 500 iterations are performed in decoding one block. Both BSC(p) and BIAWGNC(σ) are tested. As shown in Figure 1, when $\beta = 0.49$, the WMS algorithm may converge to a set of not consistent messages even though the codeword is successfully decoded. However, when $\beta = 0.5$, the gap between those two probabilities becomes smaller as predicted by Conjecture.

To get the lower bound of the LP decoding threshold, a DE-type analysis is employed in [11], [13]. The lower bound provided by the DE-type analysis depends on β though and is plotted in Figure 2. It is worth noting that the best lower bounds are obtained, in all cases, when $\beta = \frac{1}{d_v-1}$ and that there is no threshold effect when $\beta < \frac{1}{d_v-1}$. The threshold effect does not occur because the density of the correlation

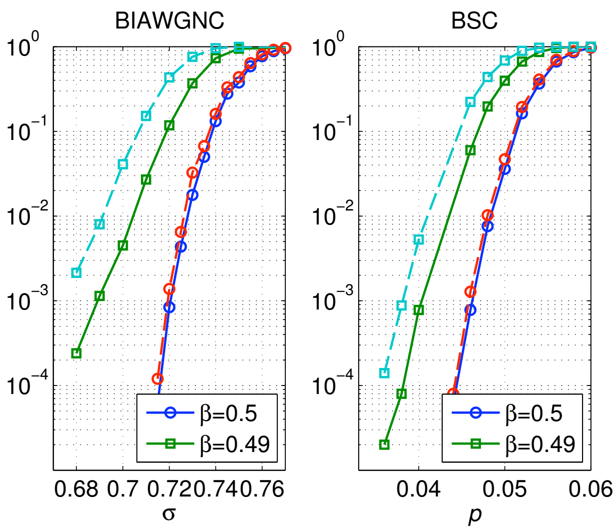

Figure 1. The BLER of WMS algorithm for $(3, 6)$-regular LDPC code and the probability of converging to inconsistent messages.

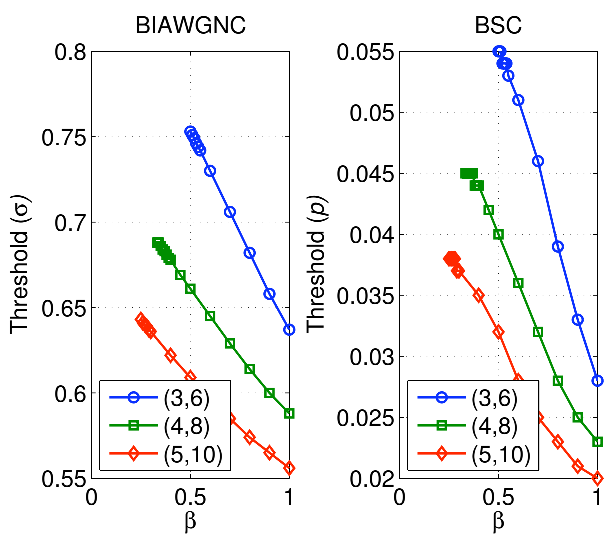

Figure 2. The lower bound of the LP decoding threshold for $(3, 6)$, $(4, 8)$ and $(5, 10)$-regular LDPC codes over BIAWGNC(σ) and BSC(p).

between the best skinny trees and the channel output in [11], [13] converges to a fixed point instead of diverging to $\pm\infty$.

IV. Conclusions and Future Work

For (d_v, d_c)-regular LDPC codes, we show that when the weight factor $\beta < \frac{1}{d_v - 1}$, the WMS algorithm converges to a unique fixed point. We also discuss a sufficient condition for the hard decision output of the weighted min-sum (WMS) algorithm to be a valid codeword. By generalizing the definition of T-local optimality in [11], we also show that if $\beta < \frac{1}{d_v - 1}$ and the WMS algorithm converges to a consistent codeword, then it is the ML solution as well as the LP optimum. This result can be seen as the natural completion of the work initiated by Koetter and Frey in [7]. For weight factors $\beta > \frac{1}{d_v - 1}$, we also provide a counterexample which shows that it is not always possible to provide ML certificates for WMS decoding. Our results have interesting connections with the results of [11] because their best LP thresholds also occur when $\beta = \frac{1}{d_v - 1}$.

In regards to future work, the most interesting open question is whether connections between LP and WMS decoding can be extended beyond $\beta = \frac{1}{d_v - 1}$. In [19], Chen *et al.* studied the optimal attenuation factor for the WMS algorithm. For example, the best β for the $(3, 6)$-regular LDPC code on the BSC is $\beta = 0.8$, and the corresponding threshold is $p = 0.083$. DE also shows that any extension beyond $\beta = \frac{1}{d_v - 1}$ will provide an improved lower bound on the LP threshold.

References

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. Commun.*, vol. 2. Geneva, Switzerland: IEEE, May 1993, pp. 1064–1070.

[2] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA, USA: The M.I.T. Press, 1963.

[3] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, March 1999.

[4] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[5] N. Wiberg, "Codes and decoding on general graphs," Ph.D. dissertation, Linköping University, S-581 83 Linköping, Sweden, 1996.

[6] N. Wiberg, H.-A. Loeliger, and R. Kötter, "Codes and iterative decoding on general graphs," *Eur. Trans. Telecom.*, vol. 6, no. 5, pp. 513–525, Sept. – Oct. 1995.

[7] B. J. Frey and R. Koetter, "Exact inference using the attenuated max-product algorithm," in *Advanced Mean Field Methods: Theory and Practice*, M. Opper and D. Saad, Eds. Cambridge, MA: MIT Press, 2000.

[8] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[9] Y. Weiss and W. T. Freeman, "On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 763–744, Feb. 2001.

[10] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky, "MAP estimation via agreement on trees: Message-passing and linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3697–3717, Nov. 2005.

[11] S. Arora, C. Daskalakis, and D. Steurer, "Message-passing algorithms and improved LP decoding," in *Proceedings of the 41st annual ACM symposium on Symposium on theory of computing*, Bethesda, MD, USA, 2009, pp. 3–12.

[12] P. O. Vontobel, "A factor-graph-based random walk, and its relevance for LP decoding analysis and bethe entropy characterization," in *Proc. Annual Workshop on Inform. Theory and its Appl.*, San Diego, CA, Feb. 2010.

[13] N. Halabi and G. Even, "LP decoding of regular ldpc codes in memoryless channels," *Arxiv preprint cs.IT/1002.3117*, 2010.

[14] J. Feldman, M. J. Wainwright, and D. R. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 954–972, March 2005.

[15] R. Koetter and P. O. Vontbel, "Graph covers and iterative decoding of finite-length codes," in *International Symposium on Turbo Codes and Related Topics*, Brest, France, Sept. 2003, pp. 75–82.

[16] J. Feldman, T. Malkin, R. A. Servedio, C. Stein, and M. J. Wainwright, "LP decoding corrects a constant fraction of errors," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 82–89, Jan. 2007.

[17] C. Daskalakis, A. G. Dimakis, R. M. Karp, and M. J. Wainwright, "Probabilistic analysis of linear programming decoding," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, 2008.

[18] R. Koetter and P. O. Vontbel, "On the block error probability of LP decoding of LDPC codes," in *Proc. 1st Annual Workshop on Inform. Theory and its Appl.*, San Diego, CA, Feb. 2006.

[19] J. Chen and M. P. C. Fossorier, "Density evolution for two improved BP-based decoding algorithms of LDPC codes," *IEEE Commun. Letters*, vol. 6, no. 5, pp. 208–210, 2002.