On the Relevance of Graph Covers and Zeta Functions for the Analysis of SPA Decoding of Cycle Codes

Henry D. Pfister
Department of ECE
Texas A&M University
College Station, TX 77843, USA
Email: hpfister@tamu.edu

Pascal O. Vontobel
Hewlett–Packard Laboratories
1501 Page Mill Road
Palo Alto, CA 94304, USA
Email: pascal.vontobel@ieee.org

***Abstract*—For an arbitrary binary cycle code, we show that sum-product algorithm (SPA) decoding after infinitely many iterations equals symbolwise graph-cover decoding. We do this by characterizing the Bethe free energy function of the underlying normal factor graph (NFG) and by stating a global convergence proof of the SPA. We also show that the set of log-likelihood ratio vectors for which the SPA converges to the all-zero codeword is given by the region of convergence of the edge zeta function associated with the underlying NFG. The results in this paper justify the use of graph-cover pseudo-codewords and edge zeta functions to characterize the behavior of SPA decoding of cycle codes. These results have also implications for the analysis of attenuated sum-product and max-product algorithm decoding of low-density parity-check (LDPC) codes beyond cycle codes.**

I. Introduction

Graph-based codes and message-passing iterative (MPI) decoders have become increasingly popular in the last fifteen years. It is fair to say that these codes and decoding algorithms (and ideas related to them) have thoroughly changed much of modern communications (see, *e.g.*, [1], [2]). Before this backdrop, a good understanding of these types of communication techniques is highly desirable, especially the understanding of MPI decoding of finite-length codes.

In this paper, we focus on binary cycle codes, *i.e.*, binary low-density parity-check (LDPC) codes with a parity-check matrix such that all columns have Hamming weight two (*cf.* [3] and references therein). Although the relevance of this class of codes is somewhat limited because its members have relatively modest minimum Hamming distance, studying this class of codes is justified for at least two reasons: as shown in this paper, the behavior of the MPI decoding, in particular of sum-product algorithm (SPA) decoding, can be characterized analytically, and, as not shown in this paper due to space limitations, some of these results have implications for the analysis of attenuated sum-product and max-product algorithm decoding of general LDPC codes.

In [4], [5], the authors suggested that a better understanding of locally operating algorithms like the SPA can be obtained by studying finite graph covers of the base factor graph. This analysis approach is based on the observation that locally operating algorithms "cannot distinguish" if they are operating on the base factor graph or any of its finite covers. Indeed, for factor graphs with cycles, this "non-distinguishability" observation implies fundamental limitations on the conclusions that can be reached by locally operating algorithms because finite graph covers of such factor graphs are in general "non-trivial" in the sense that they contain valid configurations that "cannot be explained" by valid configurations in the base factor graph. A possibility for formalizing this "non-distinguishability" observation is the symbolwise graph-cover decoder (SGCD) [6]. A priori this is a heuristic, but in this paper we show that SGCD gives indeed the correct predictions for SPA decoding of cycle codes after infinitely many iterations.

Ideally, one would want to characterize the set of pseudo-codewords derived from valid configurations of computation trees, because, as shown by Wiberg [7], SPA decoding of LDPC codes is equivalent to running the SPA on computation trees. However, characterizing this set turns out to be rather difficult, part of the reason being that the neighborhood of a leaf of a computation tree is different than the neighborhood of the corresponding vertex in the base factor graph. In contrast to this, the set of pseudo-codewords derived from valid configurations of finite graph covers is much more amenable to analysis, one of the main reasons for this being that the neighborhood of any vertex in a finite graph cover equals the neighborhood of the corresponding vertex in the base factor graph.

In this paper, we also extend earlier results that show the relevance of graph zeta functions for the analysis of SPA decoding of cycles codes. Namely, consider a binary cycle code of length n. This code can be represented by a normal factor graph (NFG) N [8]–[10] whose edge set $\mathcal{I}$ has size n. Let $\zeta_{\mathsf{N}}(\mathbf{U}) \triangleq \sum_{\boldsymbol{k}} \zeta_{\mathsf{N},\boldsymbol{k}} \mathbf{U}^{\boldsymbol{k}}$ be the Taylor series expansion of the edge zeta function [11] associated with N, where $\mathbf{U} = (\mathsf{U}_i)_{i \in \mathcal{I}}$ are indeterminates. In [12] it was shown that N has an unscaled pseudo-codeword $\boldsymbol{k}$ if and only if $\zeta_{\mathsf{N},\boldsymbol{k}} \neq 0$. Moreover, for an arbitrary unscaled pseudo-codeword $\boldsymbol{\omega}$ of N, it was demonstrated in [13] that the asymptotic behavior of $\zeta_{\mathsf{N},s\cdot\boldsymbol{\omega}}$, $s \to \infty$, is characterized to first order by the induced Bethe entropy function that is associated with N. In this paper, we show that the SPA converges to the all-zero codeword for the log-likelihood ratio (LLR) vector $\boldsymbol{\gamma} = (\gamma_i)_{i \in \mathcal{I}}$ if and only if $\mathbf{U} = \big(\exp(-\gamma_i)\big)_{i \in \mathcal{I}}$ is in the region of convergence of $\zeta_{\mathsf{N}}(\mathbf{U})$. Note that this statement links two different types of convergence statements: the convergence of an algorithm and the convergence of a power series. In the context of this paragraph, it is worthwhile to mention that also the evaluation of the edge zeta function can contain valuable information about a factor graph. Namely, for the class of factor graphs studied in [14], [15], one can express the determinant of the Hessian of the Bethe free energy function (or, equivalently, the determinant of the negative Hessian of the Bethe entropy function) in terms of a suitably evaluated edge zeta function.

Overall, the results in this paper justify the use of graph-cover pseudo-codewords and edge zeta functions to characterize the behavior of SPA decoding of cycle codes. These results have also implications (which are omitted due to space reasons) for the analysis of attenuated sum-product and max-product algorithm decoding of general LDPC codes, in particular there are important connections to the results in [16], [17].

A. Overview of the Paper

Section II gives some basic facts about cycle codes, SPA decoding, the Bethe free energy function, and the edge zeta function. Afterwards, Section III discusses the shape of the Bethe free energy and related functions, Section IV presents SPA convergence results, and Section V lists implications of these results for the analysis of SPA decoding of cycle codes via graph covers and graph zeta functions.

B. Notation

We let $\mathbb{F}_2$ be the finite field of size two and $\mathbb{R}$ be the real field. We use the ∇ symbol as follows: for a function $g : \mathbb{R} \to \mathbb{R}$, ∇g denotes its derivative; for a function $g : \mathbb{R}^N \to \mathbb{R}$, ∇g denotes its gradient; for a function $g : \mathbb{R}^N \to \mathbb{R}^N$, ∇g denotes its Jacobian. All logarithms are natural logarithms.

II. Basics

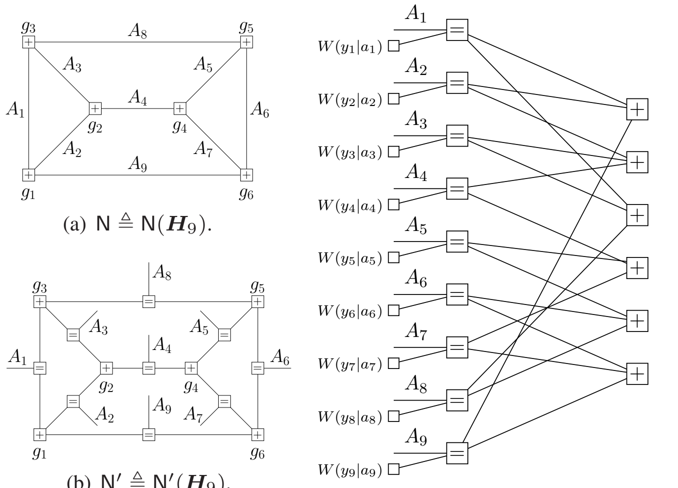
A. Cycle Codes

All definitions, lemmas, and theorems in this paper will be given in terms of a fixed binary linear code $\mathcal{C}_{\mathrm{ch}}$ defined by some parity-check matrix $\boldsymbol{H}$ of size $m \times n$. For such a parity-check matrix $\boldsymbol{H}$, we let $\mathcal{I} \triangleq \mathcal{I}(\boldsymbol{H}) \triangleq \{1, \ldots, n\}$ be the set of codeword indices, we let $\mathcal{J} \triangleq \mathcal{J}(\boldsymbol{H}) \triangleq \{1, \ldots, m\}$ be the set of check indices, we let $\mathcal{J}_i \triangleq \mathcal{J}_i(\boldsymbol{H}) \triangleq \big\{ j \in \mathcal{J} \bigm| [\boldsymbol{H}]_{j,i} = 1 \big\}$ be the set of check indices that involve the ith codeword index, and we let $\mathcal{I}_j \triangleq \mathcal{I}_j(\boldsymbol{H}) \triangleq \big\{ i \in \mathcal{I} \bigm| [\boldsymbol{H}]_{j,i} = 1 \big\}$ be the set of codeword indices that are involved in the jth check. In the following, we assume that $|\mathcal{J}_i| = 2$ for all $i \in \mathcal{I}$, *i.e.*, we assume that $\mathcal{C}_{\mathrm{ch}}$ is a binary cycle code.

Example 1 *Consider the binary linear* $[9, 4, 3]$ *code* $\mathcal{C}_9$ *that is defined by the parity-check matrix (see also [7, Ch. 6])*

$$\boldsymbol{H}_9 \triangleq \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

(Note that $\boldsymbol{H}_9$ *does not have full rank over* $\mathbb{F}_2$*; in fact, one row is redundant.) For the parity-check matrix* $\boldsymbol{H}_9$ *we have the sets* $\mathcal{I} = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $\mathcal{J} = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{I}_1 = \{1, 2, 9\}$, $\mathcal{I}_2 = \{2, 3, 4\}$, etc., $\mathcal{J}_1 = \{1, 3\}$, $\mathcal{J}_2 = \{1, 2\}$, etc. *Figure 1(a) shows an NFG called* $\mathsf{N} \triangleq \mathsf{N}(\boldsymbol{H}_9)$ *whose global function equals the indicator function of* $\mathcal{C}_9$. *Here, for all* $i \in \mathcal{I}$, *the alphabet of* A_i *equals* $\{0, 1\}$, *and for all* $j \in \mathcal{J}$, *the local function* g_j *equals the indicator function of a single parity-check constraint of degree* $|\mathcal{I}_j|$. *Moreover, Figure 1(b) shows an NFG called* $\mathsf{N}' \triangleq \mathsf{N}'(\boldsymbol{H}_9)$ *whose global function is such that its marginal function w.r.t. the half-edge variables equals the indicator function of* $\mathcal{C}_9$. □



(a) $\mathsf{N} \triangleq \mathsf{N}(\boldsymbol{H}_9)$.

(b) $\mathsf{N}' \triangleq \mathsf{N}'(\boldsymbol{H}_9)$.

(c) $\mathsf{N}'' \triangleq \mathsf{N}''(\boldsymbol{H}_9)$.

Fig. 1. NFGs for Examples 1 and 2.

Pseudo-codewords of binary cycle codes were studied in, *e.g.*, [18].

B. SPA Decoding

In the following, we assume that $\mathsf{N} \triangleq \mathsf{N}(\boldsymbol{H})$ is connected and that the code $\mathcal{C}_{\mathrm{ch}}$ is used for data transmission over a memoryless binary-input channel with channel law $W(y|x)$. Let $\boldsymbol{\gamma} = (\gamma_i)_{i \in \mathcal{I}} \triangleq \log\big(W(y_i|0)/W(y_i|1)\big)$ be the LLR vector that is computed at the receiver based on the observed channel output sequence $\boldsymbol{y} = (y_i)_{i \in \mathcal{I}}$. If not stated otherwise, we assume throughout this paper that the vectors $\boldsymbol{y}$ and $\boldsymbol{\gamma}$ are fixed and that all components of $\boldsymbol{\gamma}$ are finite.

Example 2 *We continue Example 1. Figure 1(c) shows an NFG called* $\mathsf{N}'' \triangleq \mathsf{N}''(\boldsymbol{H}_9)$ *whose global function is such that its marginal function w.r.t. half-edge variables is proportional to* $\prod_{i \in \mathcal{I}} W(y_i|a_i)$ *for* $\boldsymbol{a} = (a_i)_{i \in \mathcal{I}} \in \mathcal{C}_9$, *and zero otherwise. (Here we have followed the approach of [6, Section VIII] to draw such an NFG.) Note that for every* $i \in \mathcal{I}$ *there is an equality function node of degree* $|\mathcal{J}_i| + 2$ *and for every* $j \in \mathcal{J}$ *there is a single parity-check function node of degree* $|\mathcal{I}_j|$. *Note also that this NFG is such that the omission of the function nodes* $W(y_i|a_i)$, $i \in \mathcal{I}$, *along with their adjacent edges, yields the NFG in Figure 1(b).* □

For $i \in \mathcal{I}$, $j \in \mathcal{J}_i$, and integer $t \geqslant 0$, let $\overrightarrow{\lambda}_{i,j}^{(t)}$ and $\overleftarrow{\lambda}_{i,j}^{(t)}$ be the, respectively, left-to-right and right-to-left LLR-based SPA messages along edge (i, j) at iteration t [8]. Moreover, let $\overrightarrow{\boldsymbol{\lambda}}^{(t)} = \big(\overrightarrow{\lambda}_{i,j}^{(t)}\big)_{i \in \mathcal{I}, j \in \mathcal{J}_i}$ and $\overleftarrow{\boldsymbol{\lambda}}^{(t)} = \big(\overleftarrow{\lambda}_{i,j}^{(t)}\big)_{i \in \mathcal{I}, j \in \mathcal{J}_i}$ collect, respectively, all the left-to-right and all the right-to-left LLR-based SPA messages at iteration t. In the following, we assume that $\overrightarrow{\boldsymbol{\lambda}}^{(0)} = \overleftarrow{\boldsymbol{\lambda}}^{(0)} = \boldsymbol{0}$.

Definition 3 *Let the operator $S_{\boldsymbol{\gamma}}$ map the vector $\overrightarrow{\boldsymbol{\lambda}} = (\overrightarrow{\lambda}_{i,j})_{i\in\mathcal{I},j\in\mathcal{J}_i}$ to the vector $\overrightarrow{\boldsymbol{\lambda}}' = (\overrightarrow{\lambda}'_{i,j})_{i\in\mathcal{I},j\in\mathcal{J}_i}$ such that*

$$\overrightarrow{\lambda}'_{i,j} = \gamma_i + \sum_{j'\in\mathcal{J}_i\setminus j} 2\,\mathrm{artanh}\left(\prod_{i'\in\mathcal{I}_{j'}\setminus i} \tanh\left(\overrightarrow{\lambda}^{(t)}_{i',j'}/2\right)\right)$$

for all $i\in\mathcal{I}$, $j\in\mathcal{J}_i$. □

The SPA message update rules in [8] imply that $\overrightarrow{\boldsymbol{\lambda}}^{(t)} = S_{\boldsymbol{\gamma}}\big(\overrightarrow{\boldsymbol{\lambda}}^{(t-1)}\big)$. Note that for cycle codes the set $\mathcal{J}_i\setminus j$ contains exactly one element for every $i\in\mathcal{I}$, $j\in\mathcal{J}_i$.

C. The Bethe Free Energy and Related Functions

For the NFG $\mathsf{N}'' = \mathsf{N}''(\boldsymbol{H})$ we define the set of half-edges $\mathcal{E}_{\mathrm{half}}$, the local marginal polytope $\mathcal{B}$, the fundamental polytope $\mathcal{P}$, the Bethe average energy function U_{B}, the Bethe entropy function H_{B}, and the Bethe free energy function F_{B} as in [6, Sections V and VIII].

Towards defining the induced Bethe entropy function, we define the following mappings [13] (see also [6, Section VIII]). First, we define the surjective mapping $\boldsymbol{\psi}$: $\mathcal{B}\to\mathcal{P}$, $\boldsymbol{\beta}\mapsto(\beta_{e,1})_{e\in\mathcal{E}_{\mathrm{half}}}$, that maps pseudo-marginal vectors to scaled pseudo-codewords. Second, we define the mapping $\boldsymbol{\Psi}_{\mathrm{BME}}$: $\mathcal{P}\to\mathcal{B}$, $\boldsymbol{\omega}\mapsto\arg\max_{\boldsymbol{\beta}\in\mathcal{B}:\,\boldsymbol{\psi}(\boldsymbol{\beta})=\boldsymbol{\omega}} H_{\mathrm{B}}(\boldsymbol{\beta})$, which gives for each $\boldsymbol{\omega}\in\mathcal{P}$ the $\boldsymbol{\beta}$ among all the $\boldsymbol{\psi}$-pre-images of $\boldsymbol{\omega}$ that has the maximal Bethe entropy function value (here "BME" stands for "Bethe Max-Entropy"). Finally, the induced Bethe entropy function is defined to be H_{B}: $\mathcal{P}\to\mathbb{R}$, $\boldsymbol{\omega}\mapsto H_{\mathrm{B}}\big(\boldsymbol{\Psi}_{\mathrm{BME}}(\boldsymbol{\omega})\big)$. Note that the argument of H_{B} determines if H_{B} denotes the Bethe entropy function or the induced Bethe entropy function.

For the NFG $\mathsf{N}'' = \mathsf{N}''(\boldsymbol{H})$ we define the induced Bethe average energy function to be U_{B}: $\mathcal{P}\to\mathbb{R}$, $\boldsymbol{\omega}\mapsto\mathrm{const}(\boldsymbol{y}) - \sum_{i\in\mathcal{I}}\omega_i\gamma_i$, where $\mathrm{const}(\boldsymbol{y})$ is an expression that depends only on the received vector $\boldsymbol{y}$. Moreover, we define the induced Bethe free energy function (at temperature 1) to be F_{B}: $\mathcal{P}\to\mathbb{R}$, $\boldsymbol{\omega}\mapsto U_{\mathrm{B}}(\boldsymbol{\omega}) - H_{\mathrm{B}}(\boldsymbol{\omega})$.

D. The Edge Zeta Function

In this subsection we introduce the edge zeta function of an undirected graph $\mathsf{N}\triangleq\mathsf{N}(\boldsymbol{H})$ like the NFG in Figure 1(a). More details can be found in [11], [12].

Definition 4 ([11], [19]) *Let Γ be a path in the undirected graph N with edge set $\mathcal{I}$; write $\Gamma = (i_1,\ldots,i_k)$ to indicate that Γ begins with the edge i_1 and ends with the edge i_k.*

- *The* monomial of Γ *is given by $\gamma(\Gamma)\triangleq\mathsf{U}_{i_1}\cdots\mathsf{U}_{i_k}$, where the U_i's are indeterminates.*
- *The* edge zeta function *of N is defined to be the power series $\zeta_{\mathsf{N}}(\mathbf{U})\triangleq\prod_{[\Gamma]\in\mathcal{P}(\mathsf{N})}\frac{1}{1-\gamma(\Gamma)}$, with coefficients in $\mathbb{Z}$ and with indeterminates $\mathbf{U}=\{\mathsf{U}_i\}_{i\in\mathcal{I}}$. Here, $\mathcal{P}(\mathsf{N})$ is the collection of equivalence classes of backtrackless, tailless, primitive cycles in N.* □

Although the product in the definition of the edge zeta function is, in general, infinite, it turns out that the edge zeta function is a rational function.

Theorem 5 ([11]) *The edge zeta function $\zeta_{\mathsf{N}}(\mathbf{U})$ is a rational function. More precisely, for any directed edge matrix $\boldsymbol{M}$ of N (see [11], [12] for a definition of this matrix), we have $\zeta_{\mathsf{N}}(\mathbf{U}) = 1\big/\det(\boldsymbol{I}-\boldsymbol{M}[\![\mathbf{U}]\!])$, where $\boldsymbol{I}$ is the identity matrix of size $2|\mathcal{I}|\times2|\mathcal{I}|$ and $[\![\mathbf{U}]\!]\triangleq\mathrm{diag}(\mathsf{U}_1,\ldots,\mathsf{U}_{|\mathcal{I}|},\mathsf{U}_1,\ldots,\mathsf{U}_{|\mathcal{I}|})$ is a diagonal matrix of indeterminates.* □

Based on $\boldsymbol{M}$ and $[\![\mathbf{U}]\!]$ from Theorem 5, we define the matrix

$$\boldsymbol{Z}_{\boldsymbol{\gamma}}\triangleq\boldsymbol{M}\cdot[\![\mathbf{U}]\!]\Big|_{\mathsf{U}_i\triangleq\exp(-\gamma_i)}, \qquad (1)$$

and denote its spectral radius by $\rho_{\boldsymbol{\gamma}}$. In the following, for ease of exposition, we assume that $\boldsymbol{\gamma}$ is such that $\rho_{\boldsymbol{\gamma}}\neq1$.

III. The Shape of the Bethe Free Energy and Related Functions

In this section we study the convexity / concavity and other properties of the Bethe free energy and related functions.

Theorem 6 *The induced Bethe entropy function is concave.*

Proof: Although the proof strategy is rather straightforward, namely studying the Hessian of the induced Bethe entropy function, the details are quite involved and therefore omitted due to space reasons. ■

Theorem 7 *The induced Bethe free energy function is convex.*

Proof: Because $U_{\mathrm{B}}(\boldsymbol{\omega})$ is a linear function of $\boldsymbol{\omega}$, this result follows immediately from Theorem 6. ■

The following theorem characterizes the induced Bethe free energy function around the all-zero vertex of $\mathcal{P}$. Analogous statements can be made for any other codeword vertex of $\mathcal{P}$.

Theorem 8 *Let $\boldsymbol{\omega}(\tau)\triangleq\tau\cdot\hat{\boldsymbol{\omega}}$ for some vector $\hat{\boldsymbol{\omega}}\in\mathcal{P}$ with $\|\hat{\boldsymbol{\omega}}\|_1 = 1$. Then, for $0<\tau<1$, we have*

$$F_{\mathrm{B}}\big(\boldsymbol{\omega}(\tau)\big)\geqslant\mathrm{const}(\boldsymbol{y})-\tau\cdot\log(\rho_{\boldsymbol{\gamma}})+O(\tau^2), \qquad (2)$$

where $\rho_{\boldsymbol{\gamma}}$ is the spectral radius of $\boldsymbol{Z}_{\boldsymbol{\gamma}}$ (cf. (1)). Note that equality holds in (2) for a certain vector $\hat{\boldsymbol{\omega}}$ that is derived from the left and right eigenvectors of $\boldsymbol{Z}_{\boldsymbol{\gamma}}$ with eigenvalue $\rho_{\boldsymbol{\gamma}}$.

Proof: The proof is along similar lines as the proof of Theorem 26 in [20]. The details are omitted. ■

Corollary 9 *Let $\rho_{\boldsymbol{\gamma}}$ be the spectral radius of $\boldsymbol{Z}_{\boldsymbol{\gamma}}$ (cf. (1)).*

- *If $\rho_{\boldsymbol{\gamma}}<1$ then $F_{\mathrm{B}}(\boldsymbol{\omega})$ has its unique minimum at $\boldsymbol{\omega}=\mathbf{0}$.*
- *If $\rho_{\boldsymbol{\gamma}}>1$ then $F_{\mathrm{B}}(\boldsymbol{\omega})$ is not minimal at $\boldsymbol{\omega}=\mathbf{0}$.*

Proof: From Theorem 8 we know that $F_{\mathrm{B}}\big(\boldsymbol{\omega}(\tau)\big)\geqslant\mathrm{const}(\boldsymbol{y})-\tau\cdot\log(\rho_{\boldsymbol{\gamma}})+O(\tau^2)$, with equality for the vector $\hat{\boldsymbol{\omega}}$ that was specified there. Moreover, from Theorem 7 we know that $F_{\mathrm{B}}(\boldsymbol{\omega})$ is convex. Therefore, if $\log(\rho_{\boldsymbol{\gamma}})<0$ (*i.e.*, $\rho_{\boldsymbol{\gamma}}<1$) then $F_{\mathrm{B}}(\boldsymbol{\omega})$ has a unique minimum at $\boldsymbol{\omega}=\mathbf{0}$. On the other hand, if $\log(\rho_{\boldsymbol{\gamma}})>0$ (*i.e.*, $\rho_{\boldsymbol{\gamma}}>1$) then $F_{\mathrm{B}}(\boldsymbol{\omega})$ cannot be minimal at $\boldsymbol{\omega}=\mathbf{0}$. Note that for $\log(\rho_{\boldsymbol{\gamma}})=0$ (*i.e.*, $\rho_{\boldsymbol{\gamma}}=1$), the minimality / non-minimality of $F_{\mathrm{B}}(\boldsymbol{\omega})$ at $\boldsymbol{\omega}=\mathbf{0}$ is determined by the $O(\tau^2)$ term. ■

IV. Convergence of the SPA

In the following, we assume that the parity-check matrix $\boldsymbol{H}$ of $\mathcal{C}_{\mathrm{ch}}$ is such that all checks nodes have degree at least three, *i.e.*, $|\mathcal{I}_j| \geqslant 3$ for all $j \in \mathcal{J}$. We start by stating a contraction result for the SPA.[1]

Theorem 10 *Assume that the message vectors $\overrightarrow{\boldsymbol{\lambda}}$ and $\overrightarrow{\boldsymbol{\nu}}$ have finite components. Then the mapping $S_{\boldsymbol{\gamma}}$ is a strict contraction under the max-norm,* i.e.,

$$\left\| S_{\boldsymbol{\gamma}}(\overrightarrow{\boldsymbol{\lambda}}) - S_{\boldsymbol{\gamma}}(\overrightarrow{\boldsymbol{\nu}}) \right\|_{\infty} < \left\| \overrightarrow{\boldsymbol{\lambda}} - \overrightarrow{\boldsymbol{\nu}} \right\|_{\infty}.$$

Proof: Let $\overrightarrow{\boldsymbol{\lambda}}' \triangleq S_{\boldsymbol{\gamma}}(\overrightarrow{\boldsymbol{\lambda}})$. The main part of the proof consists in showing that $\|\nabla \overrightarrow{\lambda}'_{i,j}(\overrightarrow{\boldsymbol{\lambda}})\|_1 < 1$ for all $i \in \mathcal{I}, j \in \mathcal{J}_i$ and any $\overrightarrow{\boldsymbol{\lambda}}$ with finite components. This implies[2] $\|\nabla S_{\boldsymbol{\gamma}}\|_{\infty} = \max_{i,j} \|\nabla \overrightarrow{\lambda}'_{i,j}(\overrightarrow{\boldsymbol{\lambda}})\|_1 < 1$ for any $\overrightarrow{\boldsymbol{\lambda}}$ with finite components. Finally, combining this inequality with Lemma 1 from [21], yields the desired result.

Let us show $\|\nabla \overrightarrow{\lambda}'_{i,j}(\overrightarrow{\boldsymbol{\lambda}})\|_1 < 1$ for some $i \in \mathcal{I}$ and $j \in \mathcal{J}_i$. Note that the partial derivative of $\overrightarrow{\lambda}'_{i,j}(\overrightarrow{\boldsymbol{\lambda}})$ w.r.t. $\overrightarrow{\lambda}_{i',j'}$ is only non-zero if $j' \in \mathcal{J}_i \setminus j$ and $i' \in \mathcal{I}_{j'} \setminus i$. In the following, let j' be the unique element in $\mathcal{J}_i \setminus j$. Moreover, assume for the moment that $\overrightarrow{\lambda}_{i',j'} \neq 0$ for all $i' \in \mathcal{I}_{j'} \setminus i$. With this, we obtain

$$\begin{aligned}
\left\|\nabla \overrightarrow{\lambda}'_{i,j}(\overrightarrow{\boldsymbol{\lambda}})\right\|_1 &= \sum_{i' \in \mathcal{I}_{j'} \setminus i} \left| \frac{\partial}{\partial \overrightarrow{\lambda}_{i',j'}} \overrightarrow{\lambda}'_{i,j}(\overrightarrow{\boldsymbol{\lambda}}) \right| \\
&\overset{\text{(a)}}{=} \sum_{i' \in \mathcal{I}_{j'} \setminus i} \left| \frac{\partial}{\partial \overrightarrow{\lambda}_{i',j'}} 2 \operatorname{artanh}\left(\prod_{i'' \in \mathcal{I}_{j'} \setminus i} \tanh\left(\overrightarrow{\lambda}_{i'',j'}/2\right) \right) \right| \\
&\overset{\text{(b)}}{=} \sum_{i' \in \mathcal{I}_{j'} \setminus i} \left| \frac{\frac{1}{\cosh^2\left(\overrightarrow{\lambda}_{i',j'}/2\right)} \prod_{i'' \in \mathcal{I}_{j'} \setminus \{i,i'\}} \tanh\left(\overrightarrow{\lambda}_{i'',j'}/2\right)}{1 - \prod_{i'' \in \mathcal{I}_{j'} \setminus i} \tanh^2\left(\overrightarrow{\lambda}_{i'',j'}/2\right)} \right| \\
&\overset{\text{(c)}}{=} \sum_{i' \in \mathcal{I}_{j'} \setminus i} \left| \frac{\frac{2}{\sinh\left(\overrightarrow{\lambda}_{i',j'}\right)} \prod_{i'' \in \mathcal{I}_{j'} \setminus i} \tanh\left(\overrightarrow{\lambda}_{i'',j'}/2\right)}{1 - \prod_{i'' \in \mathcal{I}_{j'} \setminus i} \tanh^2\left(\overrightarrow{\lambda}_{i'',j'}/2\right)} \right| \\
&\overset{\text{(d)}}{=} \sum_{i' \in \mathcal{I}_{j'} \setminus i} \left| \frac{\frac{2}{\sinh\left(\overrightarrow{\lambda}_{i',j'}\right)} \tanh\left(\overleftarrow{\lambda}_{i,j'}/2\right)}{1 - \tanh^2\left(\overleftarrow{\lambda}_{i,j'}/2\right)} \right| \\
&\overset{\text{(e)}}{=} \sum_{i' \in \mathcal{I}_{j'} \setminus i} \left| \frac{\sinh\left(\overleftarrow{\lambda}_{i,j'}\right)}{\sinh\left(\overrightarrow{\lambda}_{i',j'}\right)} \right| = \sum_{i' \in \mathcal{I}_{j'} \setminus i} \frac{\sinh\left(|\overleftarrow{\lambda}_{i,j'}|\right)}{\sinh\left(|\overrightarrow{\lambda}_{i',j'}|\right)} \\
&\overset{\text{(f)}}{=} \frac{\sum_{i' \in \mathcal{I}_{j'} \setminus i} \sinh\left(z(|\overrightarrow{\lambda}_{i',j'}|)\right)}{\sinh\left(z(|\overleftarrow{\lambda}_{i,j'}|)\right)} \overset{\text{(g)}}{=} \frac{\sum_{i' \in \mathcal{I}_{j'} \setminus i} \sinh\left(z(|\overrightarrow{\lambda}_{i',j'}|)\right)}{\sinh\left(\sum_{i' \in \mathcal{I}_{j'} \setminus i} z(|\overrightarrow{\lambda}_{i',j'}|)\right)} \\
&\overset{\text{(h)}}{<} \frac{\sum_{i' \in \mathcal{I}_{j'} \setminus i} \sinh\left(z(|\overrightarrow{\lambda}_{i',j'}|)\right)}{\sum_{i' \in \mathcal{I}_{j'} \setminus i} \sinh\left(z(|\overrightarrow{\lambda}_{i',j'}|)\right)} = 1,
\end{aligned}$$

where step (a) uses Definition 3; where step (b) uses $\nabla \operatorname{artanh}(\xi) = 1/(1-\xi^2)$; where step (c) uses the equality $\sinh(2\xi) = 2\sinh(\xi)\cosh(\xi)$; where at step (d) we have defined $\overleftarrow{\lambda}_{i,j'} \triangleq 2 \operatorname{artanh}\big(\prod_{i' \in \mathcal{I}_{j'} \setminus i} \tanh(\overrightarrow{\lambda}^{(t)}_{i',j}/2)\big)$, which implies $\tanh(\overleftarrow{\lambda}_{i,j'}/2) = \prod_{i' \in \mathcal{I}_{j'} \setminus i} \tanh(\overrightarrow{\lambda}^{(t)}_{i',j}/2)$; where step (e) uses $\tanh(\xi) = \sinh(\xi)/\cosh(\xi)$, $\sinh(2\xi) = 2\sinh(\xi)\cosh(\xi)$, and $\cosh^2(\xi) - \sinh^2(\xi) = 1$; where step (f) uses $\sinh(\xi) = 1/\sinh(z(\xi))$, which holds for the function $z(\xi) \triangleq -\log\big(\tanh(\xi/2)\big)$; where step (g) uses $z\big(|\overleftarrow{\lambda}_{i,j'}|\big) = \sum_{i' \in \mathcal{I}_{j'} \setminus i} z\big(|\overrightarrow{\lambda}_{i',j'}|\big)$, which follows from taking the negative logarithm on both sides of the expression that was used at step (d); and where step (h) uses the fact that $|\mathcal{I}_{j'} \setminus i| \geqslant 3 - 1 = 2$ and the fact that $\sinh(\xi_1 + \xi_2) > \sinh(\xi_1) + \sinh(\xi_2)$, which holds for positive ξ_1 and ξ_2 due to the strict convexity of $\sinh(\,\cdot\,)$ and due to $\sinh(0) = 0$.

We consider now the case where there is at least one $i' \in \mathcal{I}_{j'} \setminus i$ such that $\overrightarrow{\lambda}_{i',j'} = 0$. Let i^* be such an index. Continuing the above derivation after step (b), we obtain

$$\left\|\nabla \overrightarrow{\lambda}'_{i,j}(\overrightarrow{\boldsymbol{\lambda}})\right\|_1 \overset{\text{(i)}}{=} \prod_{i'' \in \mathcal{I}_{j'} \setminus \{i,i^*\}} \tanh\left(|\overrightarrow{\lambda}_{i'',j'}|/2\right) \overset{\text{(j)}}{<} 1,$$

where step (i) uses $\cosh(0) = 0$ and the fact that in the sum $\sum_{i' \in \mathcal{I}_{j'} \setminus i}$ only the term for $i' = i^*$ can be non-zero because $\tanh(0) = 0$, and where step (j) uses the inequalities $0 \leqslant \tanh\big(|\overrightarrow{\lambda}_{i'',j'}|/2\big) < 1$. ■

With the LLR-based SPA message $\overrightarrow{\lambda}^{(t)}_{i,j}$, we associate the inverse likelihood ratio SPA message $\overrightarrow{\mathrm{V}}^{(t)}_{i,j} \triangleq \exp(-\overrightarrow{\lambda}^{(t)}_{i,j})$.

Lemma 11 *For $i \in \mathcal{I}$, $j \in \mathcal{J}_i$, and $t \geqslant 1$, it holds that $\overrightarrow{\mathrm{V}}^{(t)}_{i,j} \leqslant \exp(-\gamma_i) \cdot \sum_{i' \in \mathcal{I}_{j'} \setminus i} \overrightarrow{\mathrm{V}}^{(t-1)}_{i',j'}$ where j' is the unique element of $\mathcal{J}_i \setminus j$.*

Proof: For $j \in \mathcal{J}$, let $\mathcal{A}_j$ bet the set that contains all vectors of $\mathbb{F}_2^{|\mathcal{I}_j|}$ with even Hamming weight. From the SPA message update rules [8] it follows that

$$\overrightarrow{\mathrm{V}}^{(t)}_{i,j} = \exp(-\gamma_i) \cdot \frac{\sum_{\boldsymbol{a}_{j'} \in \mathcal{A}_{j'} :\, a_{j',i} = 1} \prod_{i' \in \mathcal{I}_{j'} \setminus i :\, a_{j',i'} = 1} \overrightarrow{\mathrm{V}}^{(t-1)}_{i',j'}}{\sum_{\boldsymbol{a}_{j'} \in \mathcal{A}_{j'} :\, a_{j',i} = 0} \prod_{i' \in \mathcal{I}_{j'} \setminus i :\, a_{j',i'} = 1} \overrightarrow{\mathrm{V}}^{(t-1)}_{i',j'}}.$$

With this, showing the inequality in the lemma statement is equivalent to showing the inequality

$$\sum_{\substack{\boldsymbol{a}_{j'} \in \mathcal{A}_{j'} \\ a_{j',i} = 1}} \prod_{\substack{i' \in \mathcal{I}_{j'} \setminus i \\ a_{j',i'} = 1}} \overrightarrow{\mathrm{V}}^{(t-1)}_{i',j'} \leqslant \left(\sum_{i' \in \mathcal{I}_{j'} \setminus i} \overrightarrow{\mathrm{V}}^{(t-1)}_{i',j'} \right) \cdot \left(\sum_{\substack{\boldsymbol{a}_{j'} \in \mathcal{A}_{j'} \\ a_{j',i} = 0}} \prod_{\substack{i' \in \mathcal{I}_{j'} \setminus i \\ a_{j',i'} = 1}} \overrightarrow{\mathrm{V}}^{(t-1)}_{i',j'} \right).$$

[1] A very similar approach was taken by Mooij and Kappen to establish necessary conditions for the global convergence of the SPA. In particular, [21, Corollary 1] analyzes the SPA in the LLR domain and gives a sufficient condition for global convergence that is independent of the local evidence. This result does not imply our result because cycle codes cannot, in general, be represented by pairwise models with binary variables. It is fair to say, however, that our result is the natural extension of this result when one focuses on the special case of cycle codes. In the same paper, Mooij and Kappen also provide sufficient conditions for general graphical models with strictly positive factors. These conditions are extended to some cases with non-negative factors by the appropriate use of limits. In this case, we determined that the expression in [21, Eq. (52)] evaluates to 1 when each zero in the factors is replaced by ϵ and the limit is taken as $\epsilon \to 0$. This establishes that the SPA is non-expansive. In contrast, our result shows that the SPA is contractive on any bounded subset of the domain that is mapped inside itself. Other similar results in the literature (*e.g.* by Ihler and by Tatikonda and Jordan) are either restricted to pairwise models or do not provide any explicit bounds that can be applied to the cycle code case.

[2] Note that the max-norm of a matrix $\boldsymbol{S} = (s_{\ell,\ell'})_{\ell,\ell'}$ is $\max_{\ell'} \sum_{\ell} |s_{\ell,\ell'}|$, whereby we assume multiplication on the left-hand side of $\boldsymbol{S}$ by row vectors.

Note that all summands on the left-hand side and all summands on the right-hand side are non-negative. Then, because for every summand on the left-hand side there is at least one summand on the right-hand side (after multiplying out), we see that the lemma statement is indeed true. ■

We say that the SPA converges to the all-zero codeword if $\lim_{t\to\infty} \overrightarrow{\lambda}_{i,j}^{(t)} = +\infty$ for all $i \in \mathcal{I}$, $j \in \mathcal{J}_i$. This condition is equivalent to $\lim_{t\to\infty} \overrightarrow{\mathbf{V}}^{(t)} \to \mathbf{0}$, where we have defined $\overrightarrow{\mathbf{V}}^{(t)} \triangleq \big(\overrightarrow{\mathsf{V}}_{i,j}^{(t)}\big)_{i\in\mathcal{I},j\in\mathcal{J}_i}$.

Theorem 12 *Let $\rho_{\boldsymbol{\gamma}}$ be the spectral radius of $\boldsymbol{Z}_{\boldsymbol{\gamma}}$ (cf. (1)). The SPA converges to the all-zero codeword if and only if $\rho_{\boldsymbol{\gamma}} < 1$.*

Proof: (Sketch.) From Lemma 11 it follows that one can write $\overrightarrow{\mathbf{V}}^{(t)} \leqslant \overrightarrow{\mathbf{V}}^{(t-1)} \cdot \boldsymbol{Z}_{\boldsymbol{\gamma}}$, where the inequality is to be understood component-wise. Clearly, if $\rho_{\boldsymbol{\gamma}} < 1$ then $\lim_{t\to\infty} \overrightarrow{\mathbf{V}}^{(t)} = \mathbf{0}$. However, for $\rho_{\boldsymbol{\gamma}} > 1$ the SPA does not converge to the all-zero codeword because of Corollary 9 and Theorem 10. ■

Theorem 13 *The SPA converges to a point that achieves the (global) minimum of the induced Bethe free energy function.*

Proof: The key ingredients of this proof are Theorems 7, 10, and 12 and Corollary 9, along with the SPA/BFE theorem by Yedidia, Freeman, and Weiss. The details are omitted. ■

V. Implications

Corollary 14 *For cycle codes, SPA decoding after infinitely many iterations is equivalent to symbolwise graph-cover decoding (SGCD).*

Proof: (Sketch.) Follows from the definition of SGCD in [6, Section IX] and Theorem 13. ■

The above corollary justifies the use of graph-cover pseudo-codewords to characterize the behavior of SPA decoding of cycle codes. In that spirit, the upcoming Corollary 15, Definition 16, and Theorem 17 show how finite graph covers, along with the edge zeta function, can be used to characterize the set of $\boldsymbol{\gamma}$ for which the SPA converges to the all-zero codeword.

Corollary 15 *The SPA converges to the all-zero codeword if and only if $\mathbf{U}$ is in the region of convergence of the edge zeta function $\zeta_{\mathsf{N}}(\mathbf{U})$, where $\mathsf{U}_i = \exp(-\gamma_i)$, $i \in \mathcal{I}$.*

Proof: (Sketch.) Follows from Theorems 5 and 12. ■

The following definition generalizes the concept of the exact weight generating function of $\mathcal{C}_{\mathrm{ch}}$, which is defined to be $\mathrm{EWGF}(\mathbf{U}) \triangleq \sum_{\boldsymbol{c}\in\mathcal{C}_{\mathrm{ch}}} \mathbf{U}^{\boldsymbol{c}} = \sum_{\boldsymbol{c}\in\mathcal{C}_{\mathrm{ch}}} \mathsf{U}_1^{c_1} \cdots \mathsf{U}_n^{c_n}$.

Definition 16 *For any positive integer M, the M-cover average exact weight generating function is defined to be*

$$\overline{\mathrm{EWGF}}_M(\mathbf{U}) \triangleq \sum_{\substack{\text{unscaled} \\ \text{pseudo-codeword } \boldsymbol{\omega}}} \begin{pmatrix} \text{average number of codewords} \\ \text{per } M\text{-cover that map to } \boldsymbol{\omega} \end{pmatrix} \cdot \mathbf{U}^{\boldsymbol{\omega}}.$$

□

Note that $\overline{\mathrm{EWGF}}_1(\mathbf{U}) = \mathrm{EWGF}(\mathbf{U})$. In a sense that is made more precise in the following theorem, $\overline{\mathrm{EWGF}}_M(\mathbf{U})$ better and better approximates $\zeta_{\mathsf{N}}(\mathbf{U})$ as $M \to \infty$. In particular, the approximation is good enough to extract the region of convergence of $\zeta_{\mathsf{N}}(\mathbf{U})$, and with that the set of $\boldsymbol{\gamma}$ for which the SPA converges to the all-zero codeword.

Theorem 17 *Consider a cover degree M and consider an unscaled pseudo-codeword $\boldsymbol{\omega}$. For $1 \ll \ell \ll M$ we have*

$$\mathrm{coeff}\left(\overline{\mathrm{EWGF}}_M(\mathbf{U}), \ell\boldsymbol{\omega}\right) \doteq \mathrm{coeff}\left(\zeta_{\mathsf{N}}(\mathbf{U}), \ell\boldsymbol{\omega}\right).$$

Here, $f(\ell) \doteq g(\ell)$ means that $f(\ell) = \exp\big(o(\ell)\big) \cdot g(\ell)$, with constants independent of M.

Proof: Omitted. ■

References

[1] S. Lin and D. J. Costello, Jr., *Error Control Coding*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2004.

[2] T. Richardson and R. Urbanke, *Modern Coding Theory*. New York, NY: Cambridge University Press, 2008.

[3] S. L. Hakimi and J. Bredeson, "Graph theoretic error-correcting codes," *IEEE Trans. Inf. Theory*, vol. 14, no. 4, pp. 584–591, Jul. 1968.

[4] R. Koetter and P. O. Vontobel, "Graph covers and iterative decoding of finite-length codes," in *Proc. 3rd Intern. Symp. on Turbo Codes and Related Topics*, Brest, France, Sep. 1–5 2003, pp. 75–82.

[5] 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.

[6] P. O. Vontobel, "Counting in graph covers: a combinatorial characterization of the Bethe entropy function," *accepted for publication in IEEE Trans. Inf. Theory, available online under* `http://arxiv.org/abs/1012.0065` (ver. 2), Jan. 2013.

[7] N. Wiberg, "Codes and decoding on general graphs," Ph.D. dissertation, Department of Electrical Engineering, Linköping University, Sweden, 1996.

[8] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[9] G. D. Forney, Jr., "Codes on graphs: normal realizations," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001.

[10] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Sig. Proc. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.

[11] H. M. Stark and A. A. Terras, "Zeta functions of finite graphs and coverings," *Adv. in Math.*, vol. 121, no. 1, pp. 124–165, Jul. 1996.

[12] R. Koetter, W.-C. W. Li, P. O. Vontobel, and J. L. Walker, "Characterizations of pseudo-codewords of (low-density) parity-check codes," *Adv. in Math.*, vol. 213, no. 1, pp. 205–229, Aug. 2007.

[13] P. O. Vontobel, "Connecting the Bethe entropy and the edge zeta function of a cycle code," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 13–18 2010, pp. 704–708.

[14] Y. Watanabe and K. Fukumizu, "Graph zeta function in the Bethe free energy and loopy belief propagation," in *Proc. Neural Information Processing Systems Conference*, Vancouver, Canada, Dec. 7–12 2009.

[15] Y. Watanabe, "Discrete geometric analysis of message passing algorithm on graphs," Ph.D. dissertation, Dept. of Statistical Science, Graduate University for Advanced Studies, Tokyo, Japan, 2010.

[16] Y.-Y. Jian and H. D. Pfister, "Convergence of weighted min-sum decoding via dynamic programming on trees," *submitted to IEEE Trans. Inf. Theory, available online under* `http://arxiv.org/abs/1107.3177`, Jul. 2011.

[17] N. Halabi and G. Even, "On decoding irregular Tanner codes with local-optimality guaranties," *submitted to IEEE Trans. Inf. Theory, available online under* `http://arxiv.org/abs/1107.2677`, Nov. 2011.

[18] N. Axvig and D. Dreher, "Graphical characterizations of linear programming pseudocodewords for cycle codes," *accepted for publication in IEEE Trans. Inf. Theory*, May 2013.

[19] K. Hashimoto, "Zeta functions of finite graphs and representations of p-adic groups," in *Automorphic Forms and Geometry of Arithmetic Varieties*, ser. Adv. Stud. Pure Math. Boston, MA: Academic Press, 1989, vol. 15, pp. 211–280.

[20] P. O. Vontobel, "The Bethe permanent of a non-negative matrix," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1866–1901, Mar. 2013.

[21] J. M. Mooij and H. J. Kappen, "Sufficient conditions for convergence of the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4422–4437, Dec. 2007.