A rigorous proof of the cavity method for counting matchings

Mohsen Bayati and Chandra Nair

***Abstract*— In this paper we rigorously prove the validity of the cavity method for the problem of counting the number of matchings in graphs with large girth. Cavity method is an important heuristic developed by statistical physicists that has lead to the development of faster distributed algorithms for problems in various combinatorial optimization problems. The validity of the approach has been supported mostly by numerical simulations. In this paper we prove the validity of cavity method for the problem of counting matchings using rigorous techniques. We hope that these rigorous approaches will finally help us establish the validity of the cavity method in general.**

I. INTRODUCTION

A. Motivation

Distributed message passing algorithms like belief propagation have been around for over a decade now [14], [20], [18], [1]. Recently some important problems in combinatorial optimization have seen faster distributed algorithms - motivated using a heuristic technique in statistical physics called the cavity method - that seem to solve problem instances much larger than what was previously feasible. This method has also led to analytical predictions about some threshold phenomenon in various problems of cross-disciplinary interest. Some examples of its application include the satisfiability threshold for random constraint satisfaction [11], [12] and the corresponding survey-propagation algorithm , iterative decoding algorithms in multi-user CDMA [9], etc.

However, very few rigorous results are known concerning the validity of the cavity method and the convergence of the algorithms. In this paper we wish to add to the body of rigorous results [2], [3], [6], [15], [16] supporting the predictions of the cavity method by showing its correctness for the problem of counting matchings in large sparse graphs. We borrow some of the techniques from Gamarnik et.al. [6] but we hope that some newer lines of the argument (e.g. showing validity of the free energy shifts) could lead to useful insights into the validity of the method for other instances in which the method was applied.

The algorithms generated using this method in some instances resemble the naive believe propagation equations, whereas in some other instances they resemble two-layered belief propagation (or survey propagation) equations, and in few other cases are significantly more involved. In the problem of counting matchings, the equations generated using the cavity method resemble the naive belief propagation equations. In this paper we show the convergence of the cavity equations and the uniqueness of the fixed points for arbitrary graphs G. In general such convergence results are not known except for trees or graphs with exactly one cycle [19].

This work was supported by Microsoft research.

M. Bayati is with the Department of Electrical Engineering, Stanford University, 350 Serra Mall, Stanford, CA 94305, USA bayati@stanford.edu

C. Nair is with Theory group, Microsoft Research, One Microsoft Way, Redmond, WA 98052 cnair@microsoft.com

B. The matching problem

Counting the number and size of matchings on various types of random graphs has been a classical problem in graph theory. This problem has been intensively studied for a long time by mathematicians and computer scientists [13]. Very recently Zdebrová and Mézard [21] used the cavity method to solve this problem. They believed that the results obtained by this heuristic are exact for the matching problem and using this approach: a) they derived an algorithm that computes the entropy for arbitrary graphs with girth that diverges in the large size limit and b) derived analytical results for regular and Erdös-Rényi random graph ensembles.

We first define the problem of finding the number of perfect matchings in a simple graph and then describe the cavity equations for solving it.

1) Problem setup and notation: Consider a graph $G = (V, E)$ with n vertices V, and edge-set E. Throughout this paper we will always assume that G is simple (i.e. G has no multi-edge or self-loops) and undirected. The *girth* of a graph G is defined as the length of the shortest cycle. A *matching* is a subset of edges $M \subset E$ such that no two edges of M have a common endpoint. Let $|M|$ denote the size of matching M and let M^* be a matching of maximum size. If $|M^*| = n/2$ then M^* is called a *perfect* matching.

Counting the number of perfect matchings in a graph G is shown to be #-P complete [17], (i.e. in general no polynomial-time algorithm can find the exact number of perfect matchings of G unless $P = NP$). Since it is widely believed that $P \neq NP$, many approaches have been focused on finding polynomial-time algorithms for approximately counting the number of perfect matchings [7], [8], [5], [4].

Let $a, b, \ldots$ denote the vertices of G and $i, j, \ldots$ denote the edges. For every vertex $a \in G$ let $N(a)$ denote the vertex-neighborhood of vertex a, i.e. $N(a) = \{b : (a, b) \in E\}$, and let $E(a)$ denote the edge-neighborhood of vertex a, i.e. the set of edges in E that have the vertex a as an endpoint.

Describe a matching M by variables $s_i = s_{a,b} \in \{0, 1\}$ assigned to each edge $i = (a, b) \in E$ with

$$s_i = \begin{cases} 1 & \text{edge } i \in M \\ 0 & \text{edge } i \notin M \end{cases}$$

Since M is a matching, it follows that for any vertex $a \in V$:

$$\sum_{b \in N(a)} s_{a,b} \leq 1$$

For every matching $M \subset E$ define its energy to be its number of unmatched vertices:

$$E_G(M) = \sum_{a \in V} E_a(M) = n - 2|M|$$

where $E_a(M) = 1 - \sum_{b \in N(a)} s_{a,b}$.

This induces a probability distribution (called the *Gibb's* distribution) on the set of all matchings, $\mathcal{M}(G)$, of the graph G, defined by:

$$P_{G,\beta}(M) = \frac{1}{Z_G(\beta)} e^{-\beta E_G(M)}$$

where β is a positive number and is called *inverse temperature*. The normalizing term $Z_G(\beta) = \sum_M e^{-\beta E_G(M)}$ is called *partition function*.

The partition function of an empty graph is defined as 1. To simplify notation we will omit the dependence on β and write P_G instead of $P_{G,\beta}$ and Z_G instead of $Z_G(\beta)$.

Let a configuration denote a collection of edges $S \subset E$. Observe that P_G can also be represented on the set of all configurations, as below. Let for all $a \in V$

$$\psi_a(S) = \mathbf{1}_{\{\sum_{b \in N(a)} s_{a,b} \leq 1\}} e^{-\beta(1 - \sum_{b \in N(a)} s_{a,b})}.$$

Then it follows that

$$P_G(S) = \frac{1}{Z_G} \prod_{a \in V} \psi_a(S).$$

Note that $E_G(M)$ is minimized when M is a maximum size matching and for large values of β the partition function is dominated by the terms corresponding to maximum size matchings of G. Hence for $\beta \gg 0$

$$F_G \triangleq -\frac{1}{\beta} \log Z_G \approx \frac{-\log(N_G)}{\beta} + E_G(M^*)$$

where N_G is number of maximum size matchings in G. $F_G(\beta)$ is defined as the free-energy of the system. Observe that for large β, we have

$$\begin{aligned} E_G(M^*) &\approx \frac{\partial}{\partial \beta}(\beta F_G) \\ \log(N_G) &\approx \beta \left(\frac{\partial}{\partial \beta}(\beta F_G) - F_G \right), \end{aligned}$$

with the approximation becoming exact as $\beta \to \infty$.

The cavity method of statistical physics is a heuristic that is used to evaluate the partition function. In the next section we will state the equations derived in [21] using the cavity approach and then prove that these equations compute the partition function exactly for sparse graphs.

Remark 1: In the first few sections the term sparse graph is used loosely to mean graphs that have no short loops. The precise dependence needed between the length of the shortest loop and the size of the graphs (measured in terms of the number of vertices and number of edges) will be spelt out in the final section.

C. The cavity-claims for the matching problem

Let $h^{i \to a} : E \times V \to \mathbb{R}$ be the 'message' that edge $i = (a, b)$ conveys to vertex a, one of its end-points. Note that there are $2|E|$ messages as there are two messages for each edge.

The following two claims form the algorithmic and analytical crux of the cavity method for this problem.

Claim 1 (Zdeborová-Mézard): Consider the iterative equation defined by

$$h^{i \to a}(t+1) = -\frac{1}{\beta} \log \left[e^{-\beta} + \sum_{j \in E(b) \setminus i} e^{\beta h^{j \to \beta}(t)} \right]. \quad (1)$$

These iterative equations converge to a unique fixed point for a large sparse graph whose girth diverges with the size of the graph.

Let $h^{i \to a}$ be the unique fixed points of the system of equations in the above claim, i.e.

$$h^{i \to a} = -\frac{1}{\beta} \log \left[e^{-\beta} + \sum_{j \in E(b) \setminus i} e^{\beta h^{j \to \beta}} \right]. \quad (2)$$

Claim 2 (Zdeborová-Mézard): The free energy for a single large sparse graph is given by

$$F_G = \sum_a \Delta F_a - \sum_i \Delta F_i$$

where

$$e^{-\beta \Delta F_a} = e^{-\beta} + \sum_{i \in E(a)} e^{\beta h^{i \to a}}, \quad e^{-\beta \Delta F_i} = 1 + e^{\beta(h^{i \to a} + h^{i \to b})}.$$

ΔF_a is often called the *free energy shift* corresponding to the removal of a vertex a and its associated edges, and ΔF_i the *free energy shift* corresponding to the removal of the edge i.

Remark 2: Our outline of the proof is as follows: first we will prove the cavity-claims for the case when G is a tree, and then we will proceed to establish this for a large sparse graph. Validity of Claim 1 is well-known for the case of a tree and the main result in the next section is to show that Claim 2 is exact as well.

II. The validity of the cavity-claims on a tree

In this section, we shall assume that the graph G is a tree. Note that removing any edge $i = (a, b)$ from G splits the graph into two subgraphs: $G^{i,a}$ containing the vertex a, and $G^{i,b}$ containing the vertex b.

Remark 3: For any graph $G = (V, E)$ with $a \in V$ and $i \in E$, let G_i denote the graph with edge i removed. Further, let G_a denote the graph G with vertex a and with all edges adjacent to a removed. In Figure 1, observe that $G_a^{i,a}$ is formed from $G^{i,a}$ by removing vertex a and all its adjacent edges.

Observe that,

$$\frac{Z_{G_i}}{Z_G} = \frac{Z_{G^{i,a}} Z_{G^{i,b}}}{Z_{G^{i,a}} Z_{G^{i,b}} + Z_{G_a^{i,a}} Z_{G_b^{i,b}}}$$

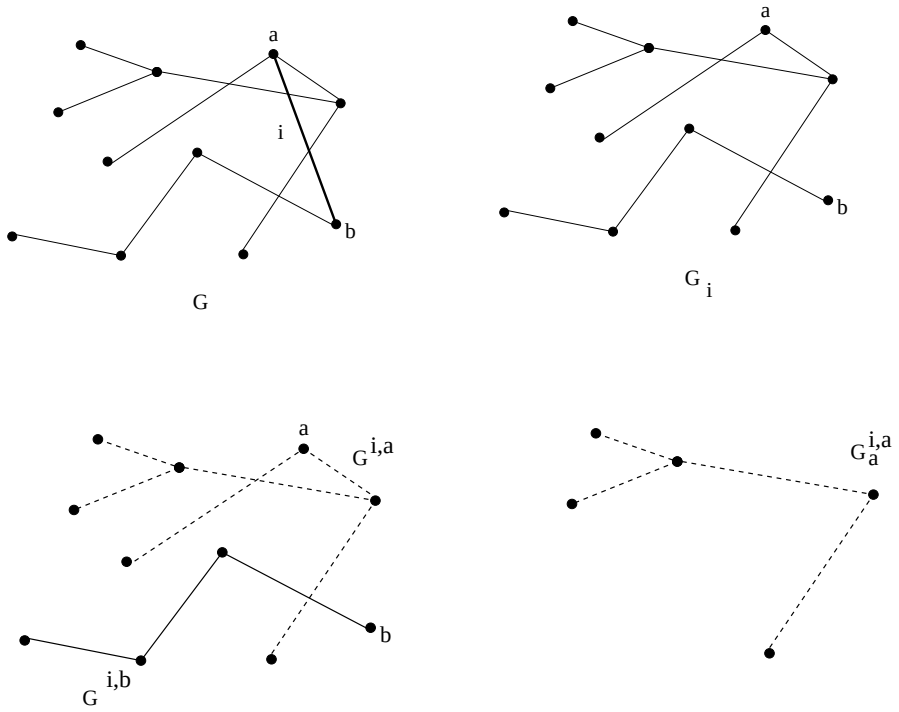

Fig. 1. The graph G and the various sub-graphs

Defining

$$e^{\beta h^{i\to b}} = \frac{Z_{G_a^{i,a}}}{Z_{G^{i,a}}}, \qquad e^{\beta h^{i\to a}} = \frac{Z_{G_b^{i,b}}}{Z_{G^{i,b}}}, \tag{3}$$

we see that

$$e^{-\beta\Delta F_i} \triangleq \frac{Z_G}{Z_{G_i}} = 1 + e^{\beta(h^{i\to a}+h^{i\to b})}. \tag{4}$$

Note that

$$\frac{Z_{G_a}}{Z_G} = \frac{\prod_{i\in E(a)} Z_{G^{i,b_i}}}{\prod_{i\in E(a)} Z_{G^{i,b_i}}(e^{-}\beta + \sum_{i\in E(a)} \frac{Z_{G_{b_i}^{i,b_i}}}{Z_{G^{i,b_i}}})}.$$

Thus,

$$e^{-\beta\Delta F_a} \triangleq \frac{Z_G}{Z_{G_a}} = e^{-\beta} + \sum_{i\in E(a)} e^{\beta h^{i\to a}}. \tag{5}$$

Lemma 1: Free energy F_G, can be expressed as the sum of free energy shifts when G is a tree, i.e.

$$F_G = \sum_a \Delta F_a - \sum_i \Delta F_i$$

Proof: Observe that

$$\begin{aligned}\prod_{a\in V} e^{-\beta\Delta F_a} \prod_{i\in E} e^{\beta\Delta F_i} &= \prod_{a\in V} \frac{Z_G}{Z_{G_a}} \prod_{i\in E} \frac{Z_{G_i}}{Z_G} \\ &= \prod_{a\in V} \frac{Z_G}{\prod_{i\in E(a)} Z_{G^{i,b_i}}} \prod_{i\in E} \frac{Z_{G^{i,a}} Z_{G^{i,b}}}{Z_G} \\ &\overset{(a)}{=} \frac{Z_G^{|V|}}{Z_G^{|E|}} \quad \overset{(b)}{=} Z_G = e^{-\beta F_G}\end{aligned} \tag{6}$$

Here (a) follows form the fact that

$$\prod_{a\in V} \prod_{i\in E(a)} Z_{G^{i,b_i}} = \prod_{i\in E} Z_{G^{i,a}} Z_{G^{i,b}}$$

and (b) follows form the fact that in a tree $|E| = |V| - 1$. ■

To complete the proof of Claim 2 for the case that G is a tree we need to show the following:

(i) The variables $h^{i\to a}$ defined in (3) satisfy equation (2).

(ii) The equations (2) have a unique fixed point.

Remark 4: The fact that the equations (2) have a unique fixed point is a well-known fact for the case when G is a tree. For general graphs G, the convergence and the uniqueness is not known. One of the main technical ingredients in this paper is to establish the convergence and the uniqueness for general graphs as well.

Lemma 2: The variables $h^{i\to a}$ defined in (3) satisfy equation (2).

Proof: We need to show that

$$h^{i\to a} = -\frac{1}{\beta}\log\left[e^{-\beta} + \sum_{j\in E(b)\setminus i} e^{\beta h^{j\to\beta}}\right].$$

This is equivalent to showing that

$$e^{-\beta h^{i\to a}} + e^{\beta h^{i\to b}} = e^{-\beta} + \sum_{j\in E(b)} e^{\beta h^{j\to b}}.$$

Now using the equations (5), (3) we see that this reduces to showing

$$\frac{Z_{G^{i,b}}}{Z_{G_b^{i,b}}} + \frac{Z_{G_a^{i,a}}}{Z_{G^{i,a}}} = \frac{Z_G}{Z_{G_b}}.$$

Observe that this follows from the following: $Z_G = Z_{G^{i,b}} Z_{G^{i,a}} + Z_{G_b^{i,b}} Z_{G_a^{i,a}}$ and $Z_{G_b} = Z_{G^{i,a}} Z_{G_b^{i,b}}$.

■

III. Convergence of the iterative equations

Consider any simple graph G and consider the iterative equations defined on it according to (1). The proof for the convergence of the iterative equations is based on the following lemma:

Lemma 3: Let $f : \mathbb{R}^{rs} \longrightarrow \mathbb{R}$ be a real valued function defined as follows:

$$f(\mathbf{x}) = \frac{-1}{\beta} \log \left(e^{-\beta} + \sum_{k=1}^{r} \frac{1}{e^{-\beta} + \sum_{\ell=1}^{s} e^{\beta x_{k\ell}}} \right) \tag{7}$$

then for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{rs}$:

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq \frac{r}{r + e^{-2\beta}} \|\mathbf{x} - \mathbf{y}\|_\infty. \tag{8}$$

Proof: Since $f(\mathbf{x})$ is differentiable, multi-variable version of the mean value theorem implies that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{rs}$ there exist a point $\mathbf{z}$ on the line-segment connecting $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^{rs}$ such that:

$$f(\mathbf{x}) - f(\mathbf{y}) = \nabla f(\mathbf{z}) \cdot (\mathbf{x} - \mathbf{y})$$

From Hölder's inequality it follows that

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq \|\nabla f(\mathbf{z})\|_1 \|\mathbf{x} - \mathbf{y}\|_\infty$$

In order to show (8) it suffices to show that $\|\nabla f(\mathbf{z})\|_1 \leq \frac{r}{r+e^{-2\beta}}$. Note that,

$$\begin{aligned} \|\nabla f(\mathbf{z})\|_1 &= \sum_{k,\ell} |\frac{\partial f}{\partial z_{k\ell}}| = \sum_{k,\ell} \frac{\frac{e^{\beta z_{k\ell}}}{(e^{-\beta} + \sum_{q=1}^{s} e^{\beta z_{kq}})^2}}{e^{-\beta} + \sum_{p=1}^{r} \frac{1}{e^{-\beta} + \sum_{q=1}^{s} e^{\beta z_{pq}}}} \\ &= \sum_{k=1}^{r} \frac{\frac{\sum_{\ell=1}^{s} e^{\beta z_{k\ell}}}{(e^{-\beta} + \sum_{q=1}^{s} e^{\beta z_{kq}})^2}}{e^{-\beta} + \sum_{p=1}^{r} \frac{1}{e^{-\beta} + \sum_{q=1}^{s} e^{\beta z_{pq}}}}. \end{aligned}$$

For simplicity of notation let $A_k = \frac{1}{e^{-\beta} + \sum_{\ell=1}^{s} e^{\beta z_{k\ell}}}$, then one obtains

$$\begin{aligned} \|\nabla f(\mathbf{z})\|_1 &= \frac{\sum_{k=1}^{r} (1 - e^{-\beta} A_k) A_k}{e^{-\beta} + \sum_{k=1}^{r} A_k} \\ &= 1 - \frac{e^{-\beta} + e^{-\beta} \sum_{k=1}^{r} A_k^2}{e^{-\beta} + \sum_{k=1}^{r} A_k}. \end{aligned}$$

Now using $0 \leq A_k \leq e^{\beta}$:

$$\|\nabla f(\mathbf{z})\|_1 \quad \leq \quad 1 - \frac{e^{-\beta}}{e^{-\beta} + re^{\beta}} = \frac{r}{r + e^{-2\beta}}$$

This completes the proof of Lemma 3. ■

Next we will use Lemma 3 to prove convergence of the cavity equations (1) for any graph G.

Theorem 1: For any graph G the set of cavity equations (1) converges to a unique fixed point independent of its initial conditions.

Proof: Consider an arbitrary initial condition $\{h^{i\to a}(0)\}$. The iterative equations in (1) states that

$$h^{i\to a}(t+1) = -\frac{1}{\beta} \log \left[e^{-\beta} + \sum_{j \in E(b) \setminus i} e^{\beta h^{j\to\beta}(t)} \right].$$

Define $F(\mathbf{x})$ to be the multi-valued function from $\mathbb{R}^{2|E|}$ to $\mathbb{R}^{2|E|}$ such that

$$\{h^{i\to a}(t+1)\} = F(\{h^{i\to a}(t)\}).$$

Consider the two-iterate of function F, i.e. let $F^2 = F \circ F$. Let $F^2 = (f_1(\mathbf{x}), \ldots, f_{2|E|}(\mathbf{x}))$ where each f_i is a real valued function on $\mathbb{R}^{2|E|}$.

Observe that each function $f_i(\mathbf{x})$ can be written in the form (7) where $r, s \leq \max_{a \in V}(\deg(a))$. Let $\Delta = \max_{a \in V}(\deg(a))$, the maximum degree of a vertex in G. Now using Lemma 3 for any $t \geq 2$, we have:

$$\begin{aligned} &\|\{h^{i\to a}(t+2)\} - \{h^{i\to a}(t)\}\|_\infty \\ &\quad = \|F^2(\{h^{i\to a}\}(t)) - F^2(\{h^{i\to a}(t-2)\})\|_\infty \\ &\quad = \max_{1 \leq k \leq 2|E|} (|f_k(\{h^{i\to a}(t)\}) - f_k(\{g^{i\to a}(t-2)\}|) \\ &\quad \leq \frac{\Delta}{e^{-2\beta} + \Delta} \|\{h^{i\to a}(t)\} - \{h^{i\to a}(t-2)\}\|_\infty \\ &\quad \leq \left(\frac{\Delta}{e^{-2\beta} + \Delta} \right)^{t/2} \|\{h^{i\to a}(2)\} - \{h^{i\to a}(0)\}\|_\infty. \end{aligned} \tag{9}$$

Thus the sequence $\{h^{i\to a}(t)\}$ is Cauchy and hence converges to a point $\{h^{i\to a}\} \in \mathbb{R}^{2|E|}$. This shows that the equations (1) converge for any graph G.

To show uniqueness consider two different initial conditions $\{h^{i\to a}(0)\}$ and $\{g^{i\to a}(0)\}$. Using the same argument as in equation (9) one has:

$$\begin{aligned} &\|\{h^{i\to a}(t)\} - \{g^{i\to a}(t)\}\|_\infty \\ &\quad \leq \quad \left(\frac{\Delta}{e^{-2\beta} + \Delta} \right)^{t/2} \|\{h^{i\to a}(0)\} - \{g^{i\to a}(0)\}\|_\infty \end{aligned} \tag{10}$$

so both sequences $\{h^{i\to a}(t)\}$ and $\{g^{i\to a}(t)\}$ converge to the same point and this contradicts the existence of multiple fixed points.

This completes the proof of the Theorem 1 and shows that the equations (1) converge to a unique fixed point for any graph. ■

In the next section we show that validity of the cavity-claims when the graph G has a large girth.

IV. Validity of the Cavity-Claims for graphs with large girth

Theorem 1 proves the validity of the Claim 1 for arbitrary graphs and in particular for graphs with large girth. Therefore it suffices to show the validity of Claim 2 for graphs with large girth. Before starting the proof, we note the following bounds on the values of the fixed points of the equation (1). The proof of this lemma is straightforward and is omitted.

Lemma 4: Let $\{h^{i\to a}\}$ be the the unique fixed points of the iterative equations on any graph G with maximum degree Δ. Then

$$-\frac{1}{\beta}\log\left[e^{-\beta}+(\Delta-1)e^{\beta}\right]\leq h^{i\to a}\leq 1.$$

Consider a fixed graph G of size n. Let r_a denote the maximum distance such that the subgraph, $G(a;r_a)$, formed using the vertices that are within a distance r_a from the vertex a is a tree. Let the vertices in $G(a;r_a)$ be denoted by $V(a;r_a)$. Consider the set of edges, C, that connect $V(a;r_a)$ and $V(a;r_a)^c$. Further let the subgraph formed by the vertices $V(a;r_a)^c$ be denoted as $H(a;r_a)$.

This decomposes the original graph G into three parts: the subgraph $G(a;r_a)$, the set of edges C, and the subgraph $H(a;r_a)$. Pick any subset of the edges $D\subset C$. Let $\mathcal{M}_D\subset\mathcal{M}(G)$ denote the set of matchings in G that use precisely the subset of edges D in C. Let $V_C\subset V(a;r_a)$ denote the set of vertices in $G(a;r_a)$ that are the endpoints of the edges C. Let $V_D\subset V_C$ denote the set of vertices in $G(a;r_a)$ that are the endpoints of the edges D. Further, let U_D denote the set of vertices in $H(a;r_a)$ that are the endpoints of the edges D.

Denote $G(a;r_a-1)$ as the sub-graph formed using the vertices that are within distance r_a-1 from vertex a. Observe that

$$G(a;r_a)\supset G_{V_D}(a;r_a)\supset G(a;r_a-1)$$

Lemma 5: Let D_1 and D_2 be two different subsets of C. Let $\{h^{j\to b}\},\{g^{j\to b}\}$ be the unique fixed points of the iterative equations on the graphs $G_{V_{D_1}}$ and $G_{V_{D_2}}$. Then

$$|h^{i\to a}-g^{i\to a}|\leq\left(\frac{\Delta}{e^{-2\beta}+\Delta}\right)^{(r_a-1)/2}\frac{1}{\beta}\log\left[1+(\Delta-1)e^{2\beta}\right]$$

Proof: Since $G_{V_{D_1}}(a;r_a),G_{V_{D_2}}(a;r_a)\supset G(a;r_a-1)$ we can set $\{h^{j\to b}(0)\}=\{g^{j\to b}(0)\}$ for the messages in the sub-graph $G(a;r_a-2)$.

To bound the difference in the messages at the boundary depending on choice of D, observe the following: Let v be any vertex at distance r_a-1 from a and let u be any neighbor of v that is at distance r_a-2. Let $0\leq\delta_1,\delta_2\leq\Delta$ denote the degree of the vertex v in the graphs $G_{V_{D_1}}(a;r_a),G_{V_{D_2}}(a;r_a)$ respectively. Let e denote the edge joining v to u. Then from Lemma 4 we see that

$$|h^{e\to u}-g^{e\to u}|\leq\frac{1}{\beta}\log\left[1+(\Delta-1)e^{2\beta}\right].$$

We can use Lemma 3 to determine the propagation of this difference to the messages at a. It is not difficult to see from repeated use of Lemma 3 that

$$|h^{i\to a}-g^{i\to a}|\leq\left(\frac{\Delta}{e^{-2\beta}+\Delta}\right)^{(r_a-1)/2}\frac{1}{\beta}\log\left[1+(\Delta-1)e^{2\beta}\right]$$

■

For simplicity of notation let us denote

$$\nu=\frac{1}{\beta}\log\left[1+(\Delta-1)e^{2\beta}\right],\quad K=\frac{\Delta}{e^{-2\beta}+\Delta}.$$

It is easy to see that $\nu\leq 3$ for large enough β.

Let i be an edge that is connected to the vertex a. We will show that the free-energy shift ΔF_i for the graph G can be approximated by the free energy shift corresponding to the tree $G(a,r_a)$.

Lemma 6: Let $\delta=\nu K^{(r_a-1)/2}$. Then,

$$e^{-2\beta\delta}\frac{Z(G(a;r_a))}{Z(G_i(a;r_a))}\leq\frac{Z(G)}{Z(G_i)}\leq e^{2\beta\delta}\frac{Z(G(a;r_a))}{Z(G_i(a;r_a))}$$

Proof: Observe that

$$Z(G)=\sum_D Z(G_{V_D}(a;r_a))Z(H_{U_D}(a;r_a))$$

Similarly for G_i, obtained by removing edge i in G, we obtain

$$Z(G_i)=\sum_D Z(G_{V_D,i}(a;r_a))Z(H_{U_D}(a;r_a))$$

Let $h_{\emptyset}^{i\to a}$ be the converged values of the iterative equations for the case when $D=\emptyset$. Combining the result for trees and our bounds on the converged values of $h^{i\to a}$ for different initial conditions in Lemma 5 we know that

$$\begin{aligned}1+e^{\beta(h_\emptyset^{i\to a}+h_\emptyset^{i\to b}-2\delta)}&\leq\frac{Z(G_{V_D}(a;r_a))}{Z(G_{V_D,i}(a;r_a))}\\&\leq 1+e^{\beta(h_\emptyset^{i\to a}+h_\emptyset^{i\to b}+2\delta)}.\end{aligned}$$

Therefore for all choices of D the following holds

$$\frac{Z(G(a;r_a))}{Z(G_i(a;r_a))}e^{-2\beta\delta}\leq\frac{Z(G_{V_D}(a;r_a))}{Z(G_{V_D,i}(a;r_a))}\leq\frac{Z(G(a;r_a))}{Z(G_i(a;r_a))}e^{2\beta\delta}$$

Using this result and the decompositions of $Z(G)$ and $Z(G_i)$ presented above we obtain that

$$\frac{Z(G(a;r_a))}{Z(G_i(a;r_a))}e^{-2\beta\delta}\leq\frac{Z(G)}{Z(G_i)}\leq\frac{Z(G(a;r_a))}{Z(G_i(a;r_a))}e^{2\beta\delta}$$

■

Observing that the boundary conditions do not influence the value of $h^{i\to a}$, the fixed points of the iterative equations for the graph G, we can infer that

$$\begin{aligned}\left(1+e^{\beta(hi\to a+hi\to b)}\right)e^{-2\beta\delta}&\leq\frac{Z(G)}{Z(G_i)}\\&\leq\left(1+e^{\beta(hi\to a+hi\to b)}\right)e^{2\beta\delta}.\end{aligned}$$

In a very similar manner to the removal of an edge, we can also show that for some $\tilde{\delta}\leq\log(\Delta)+\delta$

$$\frac{Z(G(a;r_a))}{Z(G_a(a;r_a))}e^{-\beta\tilde{\delta}}\leq\frac{Z(G)}{Z(G_a)}\leq\frac{Z(G(a;r_a))}{Z(G_a(a;r_a))}e^{\beta\tilde{\delta}}.\quad(11)$$

Similar to the case of the removal of the edge, the above equation implies that

$$\left(e^{-\beta}+\sum_{i\in E(a)}e^{\beta h^{i\to a}}\right)e^{-\beta\tilde{\delta}}\le\frac{Z(G)}{Z(G_a)}\le\left(e^{-\beta}+\sum_{i\in E(a)}e^{\beta h^{i\to a}}\right)e^{\beta\tilde{\delta}}.$$

A. Proof of Claim 2 for graphs with large girth

Let G be a graph that satisfies the *girth condition*.

$$\text{girth}(G)>\frac{8(\log\nu+\log\beta+2\log m+\log\Delta+\log\frac{1}{\log(1+\epsilon)})}{\log\frac{1}{K}} \tag{12}$$

where $m=|V|+|E|$.

Theorem 2: For any $\epsilon,\beta>0$. Let G be a graph satisfying the girth condition. Then the free energy shifts approximates Z_G within factor $1+\epsilon$, i.e.

$$Z_G(1+\epsilon)^{-1}\le X_G\triangleq\frac{\prod_a\frac{Z_G}{Z_{G_a}}}{\prod_i\frac{Z_G}{Z_{G_i}}}\le Z_G(1+\epsilon). \tag{13}$$

Note that when $E=\emptyset$, then $X_G=1$ and so (13) holds. Let $E=\{i_1,i_2,\ldots,i_{|E|}\}$ be an ordering of edges and $g=\lfloor\text{girth}(G)/4-1\rfloor$. The following lemma is crucial to prove Theorem 2.

Lemma 7: For all $1\le r\le|E|$:

$$\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}e^{-\nu\Delta\beta mK^{g/2}}\le\frac{X_{G_{E_r}}}{X_{G_{E_{r-1}}}}\le\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}e^{\nu\Delta\beta mK^{g/2}}$$

where $E_r=\{i_1,\ldots,i_r\}$, $K=\frac{\Delta}{\Delta+e^{-2\beta}}$ and $m=n+|E|$.

Before proving Lemma 7 we will show how it can be used to prove Theorem 2.

Proof: [Proof of Theorem 2]

Assuming Lemma 7 and taking the telescopic product of $r=1$ to $|E|$, we obtain

$$e^{-\nu\Delta\beta m^2K^{g/2}}\prod_{r=1}^{|E|}\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}\le\prod_{r=1}^{|E|}\frac{X_{G_{E_r}}}{X_{G_{E_{r-1}}}}\le e^{\nu\Delta\beta m^2K^{g/2}}\prod_{r=1}^{|E|}\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}$$

and hence

$$e^{-\nu\Delta\beta m^2\left(\frac{\Delta}{\Delta+e^{-2\beta}}\right)^{g/2}}Z_G\le X_G\le e^{\nu\Delta\beta m^2\left(\frac{\Delta}{\Delta+e^{-2\beta}}\right)^{g/2}}Z_G.$$

The assumptions on the girth of the graph in (12) implies that $e^{\nu\Delta\beta m^2K^{g/2}}\le1+\epsilon$ and this completes the proof. ■

Therefore, all that remains is to prove Lemma 7.

Proof: [Proof of Lemma 7]

We need to show that

$$\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}e^{-\nu\Delta\beta mK^{g/2}}\le\frac{X_{G_{E_r}}}{X_{G_{E_{r-1}}}}\le\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}e^{\nu\Delta\beta mK^{g/2}}.$$

Observe that,

$$\frac{X_{G_{E_r}}}{X_{G_{E_{r-1}}}}=\frac{\prod_{a\in V}\frac{Z_{G_{E_r}}}{Z_{G_{E_r,a}}}}{\prod_{i\in E\setminus E_r}\frac{Z_{G_{E_r}}}{Z_{G_{E_r,i}}}}\left(\frac{\prod_{a\in V}\frac{Z_{G_{E_{r-1}}}}{Z_{G_{E_{r-1},a}}}}{\prod_{i\in E\setminus E_{r-1}}\frac{Z_{G_{E_{r-1}}}}{Z_{G_{E_{r-1},i}}}}\right)^{-1}=\frac{\prod_{a\in V}\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}\left(\frac{Z_{G_{E_r,a}}}{Z_{G_{E_{r-1},a}}}\right)^{-1}}{\prod_{i\in E\setminus E_r}\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}\left(\frac{Z_{G_{E_r,i}}}{Z_{G_{E_{r-1},i}}}\right)^{-1}}. \tag{14}$$

Let a_r be an endpoint of the edge i_r. We will estimate the product

$$\frac{\prod_{a\in V}\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}\left(\frac{Z_{G_{E_r,a}}}{Z_{G_{E_{r-1},a}}}\right)^{-1}}{\prod_{i\in E\setminus E_r}\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}\left(\frac{Z_{G_{E_r,i}}}{Z_{G_{E_{r-1},i}}}\right)^{-1}}$$

by partitioning the vertices and edges into two groups; those that are in $G(a_r,g)$ and those that are outside $G(a_r,g)$.

Using equation (11) whenever $a\notin G_{E_{r-1}}(a_r;g)$, we have

$$e^{-\nu\Delta\beta K^{g/2}}\le\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}\left(\frac{Z_{G_{E_r,a}}}{Z_{G_{E_{r-1},a}}}\right)^{-1}\le e^{\nu\Delta\beta K^{g/2}}. \tag{15}$$

Similarly whenever $i\notin G_{E_{r-1}}(a_r;g)$ then from Lemma 6 we have

$$e^{-\nu\Delta\beta K^{g/2}}\le\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}\left(\frac{Z_{G_{E_r,i}}}{Z_{G_{E_{r-1},i}}}\right)^{-1}\le e^{\nu\Delta\beta K^{g/2}}. \tag{16}$$

Now we consider the case when $a,i\in G(a_r,g)$.

Since $G_{E_{r-1}}(a_r;g)\subset G_{E_{r-1}}(a_r;2g)$ and both graphs are trees, we can use (11) for $a\in G_{E_{r-1}}(a_r;g)$ to obtain

$$e^{-\nu\Delta\beta K^{g/2}}\frac{Z_{G_{E_r}}}{Z_{G_{E_r;a}}}\le\frac{Z_{G_{E_r}(a_r;g)}}{Z_{G_{E_r,a}(a_r;g)}}\le\frac{Z_{G_{E_r}}}{Z_{G_{E_r,a}}}e^{\nu\Delta\beta K^{g/2}}$$

$$e^{-\nu\Delta\beta K^{g/2}}\frac{Z_{G_{E_{r-1}}}}{Z_{G_{E_{r-1},a}}}\le\frac{Z_{G_{E_{r-1}}(a_r;g)}}{Z_{G_{E_{r-1},a}(a_r;g)}}\le\frac{Z_{G_{E_{r-1}}}}{Z_{G_{E_{r-1},a}}}e^{\nu\Delta\beta K^{g/2}}. \tag{17}$$

Similarly, for $i\in G_{E_{r-1}}(a_r;g)$ using Lemma 6 we have

$$e^{-\nu\Delta\beta K^{g/2}}\frac{Z_{G_{E_r}}}{Z_{G_{E_r,i}}}\le\frac{Z_{G_{E_r}(a_r,g)}}{Z_{G_{E_r,i}(a_r,g)}}\le\frac{Z_{G_{E_r}}}{Z_{G_{E_r,i}}}e^{\nu\Delta\beta K^{g/2}}$$

$$e^{-\nu\Delta\beta K^{g/2}}\frac{Z_{G_{E_{r-1}}}}{Z_{G_{E_{r-1},i}}}\le\frac{Z_{G_{E_{r-1}}(a_r,g)}}{Z_{G_{E_{r-1},i}(a_r,g)}}\le\frac{Z_{G_{E_{r-1}}}}{Z_{G_{E_{r-1},i}}}e^{\nu\Delta\beta K^{g/2}}. \tag{18}$$

Note that,

$$\frac{X_{G_{E_r}}}{X_{G_{E_{r-1}}}} = \frac{\prod_{a\notin G(a_r;g)} \frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}} \left(\frac{Z_{G_{E_r,a}}}{Z_{G_{E_{r-1},a}}}\right)^{-1}}{\prod_{i\notin G(a_r;g)} \frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}} \left(\frac{Z_{G_{E_r,i}}}{Z_{G_{E_{r-1},i}}}\right)^{-1}} \times \frac{\prod_{a\in G(a_r;g)} \frac{Z_{G_{E_r}}}{Z_{G_{E_r,a}}}}{\prod_{i\in G(a_r;g)} \frac{Z_{G_{E_r}}}{Z_{G_{E_r,i}}}} \left(\frac{\prod_{a\in G(a_r;g)} \frac{Z_{G_{E_{r-1}}}}{Z_{G_{E_{r-1},a}}}}{\prod_{i\in G(a_r;g)} \frac{Z_{G_{E_{r-1}}}}{Z_{G_{E_{r-1},i}}}}\right)^{-1}. \quad (19)$$

Using equations (15), (16), in the first product and equations (17), (18) in the second product (and from the definition of X_G), we obtain

$$e^{-\nu\Delta\beta m K^{g/2}} \frac{X_{G_{E_r}(a_r;g)}}{X_{G_{E_{r-1}}(a_r;g)}} \leq \frac{X_{G_{E_r}}}{X_{G_{E_{r-1}}}} \leq e^{\nu\Delta\beta m K^{g/2}} \frac{X_{G_{E_r}(a_r;g)}}{X_{G_{E_{r-1}}(a_r;g)}}.$$

Lemma 1 implies that $X_G = Z_G$ when G is a tree. Therefore, since $G_{E_r}(a_r;g)$, $G_{E_{r-1}}(a_r;g)$ are trees we can replace $\frac{X_{G_{E_r}(a_r;g)}}{X_{G_{E_{r-1}}(a_r;g)}}$ with $\frac{Z_{G_{E_r}(a_r;g)}}{Z_{G_{E_{r-1}}(a_r;g)}}$ and subsequently using Lemma 6 to replace $\frac{Z_{G_{E_r}(a_r;g)}}{Z_{G_{E_{r-1}}(a_r;g)}}$ with $\frac{Z_{G_{E_r}}}{Z_{G_{E_{r-1}}}}$, we complete the proof of Lemma 7. ■

V. CONCLUSIONS AND FUTURE WORKS

In this paper we show the validity of the cavity method for the problem of counting the number of matchings for graphs with large girth. The girth condition we have in this paper is quite restrictive and several graphs of practical relevance do not meet this condition. However we hope that the methods presented here can be extended in a straightforward manner to random regular graphs and Erdös-Rényi graphs. This would lead to, as observed in [21], tighter estimates for counting the number of matchings in such graphs.

We also demonstrate the convergence and uniqueness of the iterative equations for arbitrary graphs. The techniques used in this paper do not heavily depend on the nature of the problem and therefore there is a good possibility of these techniques having a wider interest and applicability to other important problems were cavity method has been applied.

REFERENCES

[1] S. M. Aji and R. J. McEliece, "The Generalized Distributive Law," *IEEE Trans. Inform. Theory*, Vol. 46, pp. 325-343, 2000.

[2] D. Aldous, "The $\zeta(2)$ limit in the random assignment problem", *Rand. Struct. and Algo.*, 18, 381-418, 2001.

[3] M. Bayati, D. Shah and M. Sharma, "Maximum weight matching via Max-Product Belief Propagation", *Proceedings of the International Symposium on Information Theory*, 2005.

[4] S. Chien, "A determinant-based algorithm for counting matchings in a general graph", *Proceedings of the fifteenth annual ACM-SIAM Symposium On Discrete Algorithms*, 728-735, 2004.

[5] M. Fürer and S Kasiviswanathan, "Approximately Counting Perfect Matchings in General Graphs", *SIAM proceedings of ALENEX/ANALCO 2005*, 2005.

[6] D.Gamarnik, R. Nowicki and G. Swirscsz, "Maximum Weight Independent Sets and Matchings in Sparse Random Graphs. Exact Results using the Local Weak Convergence Method", Lecture Notes on Computer Science, 3122, 357 [math.PR/0309441]

[7] M. Jerrum and A. Sinclair, "The Markov chain Monte Carlo method: an approach to approximate counting and integration, In Approximation algorithms for NP-hard problems (D. Hochbaum, ed.)", PWS Publishing Company, Boston MA, 482-520, 1996.

[8] M. Jerrum, A. Sinclair and E. Vigoda, "A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries", *Journal of the ACM* 51 no. 5 671-697, 2004.

[9] Y. Kabashima, "A CDMA multiuser detection algorithm based on survey propagation," cs/0506062, 2005.

[10] R. Karp and M. Sipser, "Maximum matchings in sparse random graphs", *Proceedings of the 22nd Annual IEEE Symposium on Foundations of Computer Science*, 364-375, 1981.

[11] M. Mézard, G. Parisi and R. Zecchina, "Analytic and Algorithmic Solution of Random Satisfiability Problems", *Science*, 297 p 812, 2002.

[12] M. Mézard and R. Zecchina, "The random K-Satisfiability problem: from an analytic solution to an efficient algorithm", *Phys. Rev.*, E 66 056126, 2002.

[13] S. Micali and V. Vazirani, "An O($|E|\sqrt{|V|}$) algorithm for finding maximum matching in general graphs", *Proceedings of 21st IEEE Symposium on Foundations of Computer Science*, 17-27, 1980.

[14] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference," San Francisco, CA: Morgan Kaufmann, 1988.

[15] M. Talagrand, "Spin glasses: A challenge for mathematicians", Springer, 2003.

[16] T. Richardson and R. Urbanke, "The Capacity of Low-Density Parity Check Codes under Message-Passing Decoding," *IEEE Trans. Info. Theory*, Vol. 47, pp 599-618, 2001.

[17] L. Valiant, "The complexity of computing the permanent", *Theoretical Computer Science* 8, 189-201, 1979.

[18] M. Wainwright, M. Jordan, "Graphical models, exponential families, and variational inference," Dept. of Stat., University of Cal., Berkeley, CA, Tech. Report, 2003.

[19] Y. Weiss, "Belief propagation and revision in networks with loops," MIT AI Lab., Tech. Rep. 1616, 1997.

[20] J. Yedidia, W. Freeman and Y. Weiss, "Understanding Belief Propagation and its Generalizations," Mitsubishi Elect. Res. Lab., TR-2001-22, 2000.

[21] L. Zdeborová and M. Mézard, "The number of matchings in random graphs", J. Stat. Mech., P05003, 2006.