Bypassing correlation decay for matchings with an application to XORSAT

Marc Lelarge
INRIA-ENS
marc.lelarge@ens.fr

***Abstract*—Many combinatorial optimization problems on sparse graphs do not exhibit the correlation decay property. In such cases, the cavity method remains a sophisticated heuristic with no rigorous proof. In this paper, we consider the maximum matching problem which is one of the simplest such example. We show that monotonicity properties of the problem allows us to define solutions for the cavity equations. More importantly, we are able to identify the 'right' solution of these equations and then to compute the asymptotics for the size of a maximum matching. The results for finite graphs are self-contained. We give references to recent extensions making use of the notion of local weak convergence for graphs and the theory of unimodular networks.**

As an application, we consider the random XORSAT problem which according to the physics literature has a 'one-step replica symmetry breaking' (1RSB) glass phase. We derive new bounds on the satisfiability threshold valid for general graphs (and conjectured to be tight).

I. INTRODUCTION

Belief Propagation (BP) is a message-passing heuristic for solving optimization problems in the context of sparse graphs. Despite the apparent empirical success of the BP algorithm for solving a variety of problems, theoretical understanding of BP is far from complete. The effectiveness of BP depends on one basic assumption: absence of long-range correlations. Physicists have developed a non-rigorous approach to deal with the emergence of long-range correlations: the one-step replica symmetry breaking (1RSB) cavity method. We will present some rigorous results confirming predictions made by this method. In this paper, we study the matching number of sparse graphs. Some of the predictions of the cavity method [1] have been rigorously proved in [2]. This problem is particularly interesting as for some random graphs long-range correlations appear at zero temperature. Indeed we will also derive partial results for random XORSAT which is known to exhibit a 1RSB phase [3], [4].

The main purpose of this paper is to present recent contributions to a rigorous formalization of the cavity method [2], [5], [6] and [7]. We will concentrate on the finite graph case and give a self-contained presentation of the computation of the matching number (Section II). Transferring these results to infinite graphs is rather standard (Section III) and requires two main mathematical tools: the notion of local weak convergence for graphs and the theory of unimodular networks [8]. We will not present these tools and we refer to the references cited above for the detailed proofs. In Section III-B, we will explain why basic methods based on the correlation decay fail at zero temperature. Finally, we give some applications to irregular XORSAT in Section IV. In particular, we construct a random instance where each variable and each clause has degree at least 3; with more variables than clauses and still the instance is not satisfiable with high probability.

II. MATCHINGS ON FINITE GRAPHS

We consider a graph $G=(V,E)$. We denote by the same symbol ∂v the set of neighbors of node $v\in V$ and the set of edges incident to v. A matching is encoded by a binary vector $\mathbf{B}=(B_e,\ e\in E)\in\{0,1\}^E$ defined by $B_e=1$ if and only if the edge e belongs to the matching. We have for all $v\in V$, $\sum_{e\in\partial v}B_e\leq 1$. The size of the matching is given by $\sum_e B_e$. We introduce the family of probability distributions on the set of matchings parametrised by a parameter $z>0$:

$$\mu_G^z(\mathbf{B})=\frac{z^{\sum_e B_e}}{P_G(z)},\tag{1}$$

where $P_G(z)=\sum_{\mathbf{B}}z^{\sum_e B_e}\prod_{v\in V}\mathbf{1}\left(\sum_{e\in\partial v}B_e\leq 1\right)$. For a finite graph G, we define the matching number of G as $\nu(G)=\max\{\sum_e B_e\}$ where the maximum is taken over matchings of G. For any finite graph, when z tends to infinity, the distribution μ_G^z converges to the uniform distribution over maximum matchings so that we have

$$\nu(G)=\lim_{z\to\infty}\sum_{e\in E}\mu_G^z(B_e=1).\tag{2}$$

A. Associated BP message passing

We introduce the set $\overrightarrow{E}$ of directed edges of G comprising two directed edges $u\to v$ and $v\to u$ for each undirected edge $uv\in E$. For $\overrightarrow{e}\in\overrightarrow{E}$, we denote by $-\overrightarrow{e}$ the edge with opposite direction. With a slight abuse of notation, we denote by ∂v the set of incident edges to $v\in V$ directed towards v. A set of messages $\mathbf{X}$ is an assignement of numbers $X_{u\to v}\geq 0$ to every oriented edges in $\overrightarrow{E}$. Given a set of messages $\mathbf{X}$, we define a new set of messages $\mathbf{Y}$ by:

$$Y_{u\to v}=\frac{1}{1+\sum_{w\in\partial u\backslash v}X_{w\to u}},\tag{3}$$

with the convention that the sum over the empty set equals zero. We denote by $\mathcal{R}_G$ the mapping sending $\mathbf{X}\in[0,\infty)^{\overrightarrow{E}}$ to $\mathbf{Y}=\mathcal{R}_G(\mathbf{X})$. We also denote by $\mathcal{R}_{\overrightarrow{e}}$ the local update rule (3): $Y_{\overrightarrow{e}}=\mathcal{R}_{\overrightarrow{e}}(\mathbf{X})$.

Proposition 1. (i) *For any finite graph G and $z>0$, the fixed point equation:*

$$\mathbf{X}=z\mathcal{R}_G(\mathbf{X})\tag{4}$$

has a unique attractive solution denoted $\mathbf{Y}(z)$.

(ii) *The function* $z \mapsto \mathbf{Y}(z)$ *is non-decreasing and the function* $z \mapsto \frac{\mathbf{Y}(z)}{z}$ *is non-increasing for* $z > 0$.

(iii) *If in addition, G is a finite tree, then for all* $e \in E$, *the law of* B_e *under* μ_G^z *is a Bernoulli distribution with*

$$\mu_G^z(B_e = 1) = \frac{Y_{\overrightarrow{e}}(z)\mathcal{R}_{-\overrightarrow{e}}(\mathbf{Y}(z))}{1 + Y_{\overrightarrow{e}}(z)\mathcal{R}_{-\overrightarrow{e}}(\mathbf{Y}(z))}. \quad (5)$$

Comparisons between vectors are always componentwise. Note that the right-hand side of (5) does not depend on the choice of orientation of the edge e as $\mathbf{Y}(z)$ solves (4). Before proving this proposition, let define for all $v \in V$, the following function of the messages $(Y_{\overrightarrow{e}},\ \overrightarrow{e} \in \partial v)$,

$$\begin{aligned}\mathcal{D}_v(\mathbf{Y}) &= \sum_{\overrightarrow{e} \in \partial v} \frac{Y_{\overrightarrow{e}}\mathcal{R}_{-\overrightarrow{e}}(\mathbf{Y})}{1 + Y_{\overrightarrow{e}}\mathcal{R}_{-\overrightarrow{e}}(\mathbf{Y})} & (6)\\ &= \frac{\sum_{\overrightarrow{e} \in \partial v} Y_{\overrightarrow{e}}}{1 + \sum_{\overrightarrow{e} \in \partial v} Y_{\overrightarrow{e}}}. & (7)\end{aligned}$$

In view of point (iii) of Proposition 1, we see that if the graph G is a tree, $\mathcal{D}_v(\mathbf{Y}(z))$ is simply the probability for vertex v to be covered by a matching distributed according to μ_G^z. In particular, when G is a tree, we can rewrite (2) as

$$\nu(G) = \lim_{z\to\infty} \frac{1}{2} \sum_{v \in V} \mathcal{D}_v(\mathbf{Y}(z)). \quad (8)$$

Proof: For the first point, we follow the proof of Theorem 3 in [5]. Let $z > 0$ and define the sequence of messages: $\mathbf{X}^0(z) = 0$ and for $t \geq 0$,

$$X^{t+1}_{u\to v}(z) = \frac{z}{1 + \sum_{w \in \partial u \backslash v} X^t_{w\to u}(z)}. \quad (9)$$

The sequence $\mathbf{X}^{2t}(z)$ (resp. $\mathbf{X}^{2t+1}(z)$) is non-decreasing (resp. non-increasing). We define $\lim_{t\to\infty} \uparrow \mathbf{X}^{2t}(z) = \mathbf{X}^-(z)$ and $\lim_{t\to\infty} \downarrow \mathbf{X}^{2t+1}(z) = \mathbf{X}^+(z)$. For any $\mathbf{Y}(z)$ fixed point of (4), a simple induction shows that

$$0 \leq \mathbf{X}^{2t}(z) \leq \mathbf{X}^-(z) \leq \mathbf{Y}(z) \leq \mathbf{X}^+(z) \leq \mathbf{X}^{2t+1}(z) \leq z.$$

We now prove that $\mathbf{X}^-(z) = \mathbf{X}^+(z)$ finishing the proof of the first point. Note that we have $\mathbf{X}^+(z) = z\mathcal{R}_G(\mathbf{X}^-(z))$ and $\mathbf{X}^-(z) = z\mathcal{R}_G(\mathbf{X}^+(z))$. In particular for any $z > 0$, we have $X^+_{\overrightarrow{e}}(z)\mathcal{R}_{-\overrightarrow{e}}(\mathbf{X}^+(z)) = X^-_{-\overrightarrow{e}}(z)\mathcal{R}_{\overrightarrow{e}}(\mathbf{X}^-(z))$ so that in view of (6), we have

$$\sum_{v \in V} \mathcal{D}_v(\mathbf{X}^+(z)) = \sum_{v \in V} \mathcal{D}_v(\mathbf{X}^-(z)). \quad (10)$$

We see from (7) that for each $v \in V$, $\mathcal{D}_v$ is an increasing function of the $(X_{\overrightarrow{e}},\ \overrightarrow{e} \in \partial v)$, so that (10) together with $\mathbf{X}^-(z) \leq \mathbf{X}^+(z)$ imply the desired result.

We now prove that $z \mapsto \frac{\mathbf{X}^t(z)}{z}$ and $z \mapsto \mathbf{X}^t(z)$ are respectively non-increasing and non-decreasing, this implies point (ii). We prove it by induction on t: consider $z \leq z'$ if $\mathbf{X}^t(z) \leq \mathbf{X}^t(z')$ then by (9) we have $\frac{\mathbf{X}^{t+1}(z)}{z} \geq \frac{\mathbf{X}^{t+1}(z')}{z'}$ and if $\frac{\mathbf{X}^t(z)}{z} \geq \frac{\mathbf{X}^t(z')}{z'}$ then again by (9), we have $\mathbf{X}^{t+1}(z) \leq \mathbf{X}^{t+1}(z')$.

We consider now the case where G is a tree. For any directed edge $u \to v$, we define $T_{u\to v}$ as the subtree containing u and v and obtained from G by removing all incident edges to v except the edge uv. A simple computation shows that

$$\frac{\mu^z_{T_{u\to v}}(B_{uv=1})}{\mu^z_{T_{u\to v}}(B_{uv=0})} = \frac{z}{1 + \sum_{w \in \partial u \backslash v} \frac{\mu^z_{T_{w\to u}}(B_{wu=1})}{\mu^z_{T_{w\to u}}(B_{wu=0})}}.$$

This directly implies that for a finite tree, $Y_{u\to v}(z) = \frac{\mu_{T^z_{u\to v}(B_{uv=1})}}{\mu_{T^z_{u\to v}(B_{uv=0})}}$. Then a simple computation shows that

$$\begin{aligned}\frac{\mu^z_G(B_{uv=1})}{\mu^z_G(B_{uv=0})} &= z\left(1 + \sum_{w \in \partial u \backslash v} Y_{w\to u}(z)\right)^{-1}\\ &\quad \times \left(1 + \sum_{w' \in \partial v \backslash u} Y_{w'\to u}(z)\right)^{-1}\\ &= \frac{Y_{u\to v}(z)Y_{v\to u}(z)}{z}\\ &= Y_{u\to v}(z)\mathcal{R}_{v\to u}(\mathbf{Y}(z)),\end{aligned}$$

which directly implies (5). ■

B. The zero temperature limit

In order to compute the matching number, we must let z tend to infinity in $\mathbf{Y}(z) = z\mathcal{R}_G(\mathbf{Y}(z))$. Iterating once this recursion, we get $\mathbf{Y}(z) = z\mathcal{R}_G(z\mathcal{R}_G(\mathbf{Y}(z)))$. Note that we have for any $z > 0$,

$$z\mathcal{R}_{u\to v}(zX) = \frac{1}{z^{-1} + \sum_{w \in \partial u \backslash v} X_{w\to u}}$$

Hence we can define for any $\mathbf{X} \in (0,1]^{\overrightarrow{E}}$, $\mathcal{Q}_G(\mathbf{X}) = \lim_{z\to\infty} \uparrow z\mathcal{R}_G(z\mathbf{X}) \in (0,\infty]^{\overrightarrow{E}}$ by its local update rule:

$$\mathcal{Q}_{u\to v}(\mathbf{X}) = \frac{1}{\sum_{w \in \partial u \backslash v} X_{w\to u}}, \quad (11)$$

with the conventions $1/0 = \infty$ and the sum over the empty set equals zero (in particular, if u is a leaf of the graph G, then $\mathcal{Q}_{u\to v}(\mathbf{X}) = \infty$).

By point (ii) of Proposition 1, we can define $\lim_{z\to\infty} \uparrow \mathbf{Y}(z) = \mathbf{Y} \in [0,\infty]^{\overrightarrow{E}}$ and $\lim_{z\to\infty} \downarrow \frac{\mathbf{Y}(z)}{z} = \mathbf{X} \in [0,1]^{\overrightarrow{E}}$. Then, we have

$$\mathbf{X} = \mathcal{R}_G(\mathbf{Y}) \text{ and, } \mathbf{Y} = \mathcal{Q}_G(\mathbf{X}), \quad (12)$$

provided we can extend the maps $\mathcal{R}_G$ and $\mathcal{Q}_G$ continuously from their respective domains $[0,\infty)^{\overrightarrow{E}}$ and $(0,1]^{\overrightarrow{E}}$ to their compactifications $[0,\infty]^{\overrightarrow{E}}$ and $[0,1]^{\overrightarrow{E}}$ respectively. This can be done easily as follows: if there exists $w \in \partial u \backslash v$ with $Y_{w\to u} = \infty$, then we set $\mathcal{R}_{u\to v}(\mathbf{Y}) = 0$; and if $X_{w\to u} = 0$ for all $w \in \partial u \backslash v$, then we set $\mathcal{Q}_{u\to v}(\mathbf{X}) = \infty$.

Lemma 1. *Let* $\lim_{z\to\infty} \uparrow \mathbf{Y}(z) = \mathbf{Y} \in [0,\infty]^{\overrightarrow{E}}$. *Then* $\mathbf{Y}$ *is the smallest solution to the fixed point equation* $\mathbf{Y} = \mathcal{Q}_G \circ \mathcal{R}_G(\mathbf{Y})$.

Proof: Let $\mathbf{Z} = \mathcal{Q}_G \circ \mathcal{R}_G(\mathbf{Z})$. For any $z > 0$, we have for any $\mathbf{X} \in [0,1]^{\overrightarrow{E}}$, $z\mathcal{R}_G(z\mathbf{X}) \leq \mathcal{Q}_G(\mathbf{X})$ so that an easy induction implies that $\mathbf{X}^{2t}(z) \leq \mathbf{Z}$ where $\mathbf{X}^{2t}(z)$ is the sequence defined in the proof of Proposition 1. Letting first t and then z tend to infinity, allows us to conclude. ■

Note that thanks to (7), we can extend the functions $\mathcal{D}_v(\mathbf{Y})$ continuously on $[0,\infty]^{\overrightarrow{E}}$ by setting $\mathcal{D}_v(\mathbf{Y}) = 1$ as soon as there exists $Y_{\overrightarrow{e}} = \infty$ for $\overrightarrow{e} \in \partial v$. To summarize, we have for each $v \in V$,

$$\lim_{z\to\infty} \mathcal{D}_v(\mathbf{Y}(z)) = \mathcal{D}_v(\mathbf{Y}) \leq 1, \tag{13}$$

where $\mathbf{Y}$ is the smallest solution to the fixed point equation $\mathbf{Y} = \mathcal{Q}_G \circ \mathcal{R}_G(\mathbf{Y})$ that can be written as:

$$Y_{u\to v} = \frac{1}{\sum_{w\in\partial u\backslash v} \frac{1}{1+\sum_{w'\in\partial w\backslash u} Y_{w'\to w}}}, \tag{14}$$

with the conventions $1/0 = \infty$ and $1/\infty = 0$ and the sum over the empty set equals zero.

Lemma 2. *We have for any $\mathbf{Y} \in [0,\infty]^{\overrightarrow{E}}$ and $v \in V$,*

$$\begin{aligned} \mathcal{D}_v(\mathbf{Y}) &= \sum_{\overrightarrow{e}\in\partial v} \frac{Y_{\overrightarrow{e}}\mathcal{R}_{-\overrightarrow{e}}(\mathbf{Y})}{1+Y_{\overrightarrow{e}}\mathcal{R}_{-\overrightarrow{e}}(\mathbf{Y})} \mathbf{1}\left(Y_{\overrightarrow{e}} < \infty\right) \\ &\quad + \mathbf{1}\left(\exists \overrightarrow{e} \in \partial v, Y_{\overrightarrow{e}} = \infty\right), \end{aligned} \tag{15}$$

where the first sum on the right-hand side should be understood as a sum over $\overrightarrow{e} \in \partial v$ with $Y_{\overrightarrow{e}} < \infty$.

Proof: We only need to conside the case where there exists $\overrightarrow{e} \in \partial v$ such that $Y_{\overrightarrow{e}} = \infty$. By the discussion before the lemma, we have in this case $\mathcal{D}_v(\mathbf{Y}) = 1$. Hence we need to prove that the first term in the right-hand side of (15) vanishes. This follows form the following fact: let $\overrightarrow{e}' \in \partial v \backslash \overrightarrow{e}$, then $Y_{\overrightarrow{e}} = \infty$ implies that $\mathcal{R}_{-\overrightarrow{e}'}(\mathbf{Y}) = 0$. ■

Given a set of $\{0,1\}$-valued messages $\mathbf{I}$, we define a new set of $\{0,1\}$-valued messages by:

$$J_{u\to v} = \mathbf{1}\left(\sum_{\ell\in\partial u\backslash v} I_{w\to u} = 0\right),$$

with the convention that the sum over the empty set equals zero. We denote by $\mathcal{P}_G$ the mapping sending $\mathbf{I}$ to $\mathbf{J} = \mathcal{P}_G(\mathbf{I})$ and as above, $\mathcal{P}_{\overrightarrow{e}}$ denotes the local update rule. For the messages $\mathbf{Y} \in [0,\infty]^{\overrightarrow{E}}$ (resp. $\mathbf{X} \in [0,1]^{\overrightarrow{E}}$) defined in (12), we define the $\{0,1\}$-valued messages $\mathbf{I}^Y$ (resp. $\mathbf{I}^X$) by $I^Y_{u\to v} = \mathbf{1}(Y_{u\to v} = \infty)$ (resp. $I^X_{u\to v} = \mathbf{1}(X_{u\to v} > 0)$. It follows directly from (12) and the definitions above that

$$\mathbf{I}^Y = \mathcal{P}_G(\mathbf{I}^X), \text{ and, } \mathbf{I}^X = \mathcal{P}_G(\mathbf{I}^Y). \tag{16}$$

We now show that for any finite graph G, the right-hand term in (8) is a function of $\mathbf{I}^X$ and $\mathbf{I}^Y$ only. Note that the equality in (8) is only valid for finite trees and that in this case, the fixed point equation (16) has a unique solution. However, we will see in Section III-B that this is not anymore true for infinite trees and this multiplicity of solutions is related to the absence of correlation decay. In the rest of this section, we will consider any finite graph G (i.e. the analysis is not restricted to trees). In such case, the fixed point equation (16) has always a solution and might have several solutions (as in the simple case of the cycle C_3).

For any $Y \in [0,\infty]$, we define $I(Y) = \mathbf{1}(Y = \infty)$ and still denote by I the function acting similarly on vectors componentwise, i.e. if $\mathbf{I} = I(\mathbf{Y})$ then $I_{\overrightarrow{e}} = I(Y_{\overrightarrow{e}})$. We define for each $v \in V$ and $\mathbf{I} \in \{0,1\}^{\overrightarrow{E}}$,

$$F_v(\mathbf{I}) = 1 \wedge \sum_{u\in\partial v} I_{u\to v} + \mathbf{1}\left(\sum_{u\in\partial v} \mathcal{P}_{u\to v}(\mathbf{I}) \geq 2\right), \tag{17}$$

where $a \wedge b = \min(a,b)$.

Lemma 3. *For $\mathbf{Y} \in [0,\infty]^{\overrightarrow{E}}$, we define $\mathbf{Y}' = \mathcal{Q}_G \circ \mathcal{R}_G(\mathbf{Y})$. If $\mathbf{Y}' \leq$ (resp. $\geq$) $\mathbf{Y}$, then*

$$\sum_{v\in V} \mathcal{D}_v(\mathbf{Y}) \geq \text{ (resp. } \leq\text{) } \sum_{v\in V} F_v(I(\mathbf{Y})).$$

Proof: Suppose $\mathbf{Y}' \leq \mathbf{Y}$, then using Lemma 2, we get

$$\begin{aligned} \sum_v \mathcal{D}_v(\mathbf{Y}) &\geq \underbrace{\sum_{\overrightarrow{e}\in\overrightarrow{E}} \frac{Y'_{\overrightarrow{e}}\mathcal{R}_{-\overrightarrow{e}}(\mathbf{Y})}{1+Y'_{\overrightarrow{e}}\mathcal{R}_{-\overrightarrow{e}}(\mathbf{Y})} \mathbf{1}\left(Y'_{\overrightarrow{e}} < \infty\right)}_{A} \\ &\quad + \sum_{v\in V} \mathbf{1}\left(\exists \overrightarrow{e} \in \partial v, Y_{\overrightarrow{e}} = \infty\right). \end{aligned}$$

For the first term A, denote $\mathbf{X} = \mathcal{R}_G(\mathbf{Y})$ so that $\mathbf{Y}' = \mathcal{Q}_G(\mathbf{X})$.Then we have

$$\begin{aligned} A &= \sum_{\overrightarrow{e}\in\overrightarrow{E}} \frac{\mathcal{Q}_{\overrightarrow{e}}(\mathbf{X})X_{-\overrightarrow{e}}}{1+\mathcal{Q}_{\overrightarrow{e}}(\mathbf{X})X_{-\overrightarrow{e}}} \mathbf{1}\left(\mathcal{Q}_{\overrightarrow{e}}(\mathbf{X}) < \infty\right) \\ &= \sum_{\overrightarrow{e}\in\overrightarrow{E}} \frac{\mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X})X_{\overrightarrow{e}}}{1+\mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X})X_{\overrightarrow{e}}} \mathbf{1}\left(\mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X}) < \infty\right) \\ &= \sum_{v\in V} \underbrace{\sum_{\overrightarrow{e}\in\partial v} \frac{X_{\overrightarrow{e}}\mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X})}{1+X_{\overrightarrow{e}}\mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X})} \mathbf{1}\left(\mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X}) < \infty\right)}_{B_v}. \end{aligned}$$

We now prove that

$$B_v = \mathbf{1}\left(\sum_{\overrightarrow{e}\in\partial v} \mathbf{1}(X_{\overrightarrow{e}} > 0) \geq 2\right). \tag{18}$$

Indeed if $\exists w \neq w'$ both in ∂v with $X_{w\to v}X_{w'\to v} > 0$, then we have $0 < \mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X}) < \infty$ for all $\overrightarrow{e} \in \partial v$, so that in this case we have

$$B_v = \sum_{\overrightarrow{e}\in\partial v} \frac{X_{\overrightarrow{e}}}{\mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X})^{-1} + X_{\overrightarrow{e}}} = 1.$$

Note now that if $B_v > 0$, there must exists $\overrightarrow{e} \in \partial v$ such that $X_{\overrightarrow{e}} > 0$ and $\mathcal{Q}_{-\overrightarrow{e}}(\mathbf{X}) < \infty$ and this last constraint implies that there exists $\overrightarrow{e}' \neq \overrightarrow{e} \in \partial v$ with $X_{\overrightarrow{e}'} > 0$ and we finished the proof of (18). Hence we obtain

$$\begin{aligned} \sum_v \mathcal{D}_v(\mathbf{Y}) &\geq \sum_{v\in V} \mathbf{1}\left(\sum_{\overrightarrow{e}\in\partial v} \mathbf{1}(X_{\overrightarrow{e}} > 0) \geq 2\right) \\ &\quad + \sum_{v\in V} \left(1 \wedge \sum_{\overrightarrow{e}\in\partial v} I(\mathbf{Y}_{\overrightarrow{e}})\right). \end{aligned}$$

■

We are now ready to state our first main result for finite graphs:

Proposition 2. *For any finite graph G, we have*

$$\sum_{v\in V}\mathcal{D}_v(\mathbf{Y})=\lim_{z\to\infty}\sum_{v\in V}\mathcal{D}_v(\mathbf{Y}(z))=\inf_{\mathbf{I}}\sum_{v\in V}F_v(\mathbf{I}),$$

where the infimum is over the solutions of $\mathbf{I}=\mathcal{P}_G\circ\mathcal{P}_G(\mathbf{I})$.

Proof: Let $\mathbf{Y}=\lim_{z\to\infty}\uparrow\mathbf{Y}(z)$ and recall that we denoted $\mathbf{I}^Y=I(\mathbf{Y})$ so that $\mathbf{I}^Y=\mathcal{P}_G\circ\mathcal{P}_G(\mathbf{I}^Y)$ by (16). By Lemma 3 and (13), we have

$$\lim_{z\to\infty}\sum_{v\in V}\mathcal{D}_v(\mathbf{Y}(z))=\sum_{v\in V}\mathcal{D}_v(\mathbf{Y})=\sum_{v\in V}F_v(\mathbf{I}^Y).$$

We need to prove thah if $\mathbf{I}=\mathcal{P}_G\circ\mathcal{P}_G(\mathbf{I})$ then we have $\sum_v F_v(\mathbf{I})\geq\sum_{v\in V}\mathcal{D}_v(\mathbf{Y})$. For any such $\mathbf{I}$, we define $\mathbf{W}^0$ as follows:

$$W^0_{\overrightarrow{e}}=\begin{cases}\infty & \text{if } I_{\overrightarrow{e}}=1\\ 0 & \text{otherwise.}\end{cases}$$

Then let $\mathbf{W}^{k+1}=\mathcal{Q}_G\circ\mathcal{R}_G(\mathbf{W}^k)$ for $k\geq 0$. A simple induction shows that $I(\mathbf{W}^{k+1})=\mathcal{P}_G\circ\mathcal{P}_G(I(\mathbf{W}^k))=\mathbf{I}$ for all $k\geq 0$. In particular, $\mathbf{W}^0\leq\mathbf{W}^1$ and again by induction, we see that the sequence $\{\mathbf{W}^k\}_k$ is non-decreasing and we denote by $\mathbf{W}^{\mathbf{I}}$ its limit. Applying Lemma 3 to $\mathbf{W}^k$, we get

$$\sum_{v\in V}\mathcal{D}_v(\mathbf{W}^k)\leq\sum_{v\in V}F_v(\mathbf{I}).$$

Taking the limit $k\to\infty$, we obtain

$$\sum_{v\in V}\mathcal{D}_v(\mathbf{W}^{\mathbf{I}})\leq\sum_{v\in V}F_v(\mathbf{I})$$

Moreover $\mathbf{Y}$ being the smallest solution to the fixed point equation $\mathbf{Y}=\mathcal{R}_G\circ\mathcal{R}_G(\mathbf{Y})$, we have $\mathbf{Y}\leq\mathbf{W}^{\mathbf{I}}$ and using the fact that $\mathcal{D}_v$ is increasing, we get

$$\sum_{v\in V}\mathcal{D}_v(\mathbf{Y})\leq\sum_{v\in V}\mathcal{D}_v(\mathbf{W}^{\mathbf{I}})\leq\sum_{v\in V}F_v(\mathbf{I}),$$

which concludes the proof. ■

C. From trees to general graphs

As far as matching number is concerned, our results so far allow us to compute it for trees only by combining (8) and Proposition 2. Indeed if G is a finite tree, it is simple to see that the solution to $\mathbf{I}=\mathcal{P}_G\circ\mathcal{P}_G(\mathbf{I})$ is the unique solution to $\mathbf{I}=\mathcal{P}_G(\mathbf{I})$ so that we finally have:

$$\nu(G)=\sum_{v\in V}\mathbf{1}\left(\sum_{u\in\partial v}I_{u\to v}\geq 2\right)+\frac{1}{2}\mathbf{1}\left(\sum_{u\in\partial v}I_{u\to v}\right),$$

where $\mathbf{I}$ is the unique solution to $\mathbf{I}=\mathcal{P}_G(\mathbf{I})$. We now consider the more general case of bipartite graphs. Adapting the proof of Lemma 3 allows us to get the first part of the following proposition. The second part of the proposition does not follow directly from our analysis and its proof relies on König-Hall min-max Theorem and can be found in [9] (see Theorem 1).

Proposition 3. *For any finite bipartite graph $G=(U,V,E)$, we have*

$$\begin{aligned}\sum_{u\in U}\mathcal{D}_u(\mathbf{Y}) \;=\; \inf_{\mathbf{I}=\mathcal{P}_G(\mathbf{I})}&\left\{\sum_{u\in U}\mathbf{1}\left(\sum_{v\in\partial u}I_{v\to u}\geq 1\right)\right.\\ &\left.+\sum_{v\in V}\mathbf{1}\left(\sum_{u\in\partial v}\mathcal{P}_{u\to v}(\mathbf{I})\geq 2\right)\right\},\quad(19)\end{aligned}$$

and $\nu(G)=\sum_{u\in U}\mathcal{D}_u(\mathbf{Y})$.

We end this section by a result of Godsil [10] allowing to reduce the computation of the matching number of any graph to computations on trees (these results are not needed for the sequel but help to put them into perspective). We recall Godsil's notion of the *path-tree* associated with a rooted graph G: if G is any rooted graph with root v, we define its path-tree T_G as the rooted tree whose vertex-set consists of all finite simple paths starting at the root v; whose edges are the pairs $\{P,P'\}$ of the form $P=v_1\dots v_n$, $P'=v_1\dots v_nv_{n+1}(n\geq 1)$; whose root is the single-vertex path v. By a *finite simple path*, we mean here a finite sequence of distinct vertices $v_1\dots v_n$ ($n\geq 1$) such that $v_iv_{i+1}\in E$ for all $1\leq i<n$. It is well-known since Godsil's result [10] that path-trees capture considerable information about matchings in general graph and are easier to work with than the graph itself. For a rooted graph $[G,v]$, let $T_{[G,v]}$ be the associated path-tree. The event v is uncovered is equal to $\sum_{e\in\partial v}B_e=0$.

Proposition 4. *For any finite graph G, we have for any $z>0$,*

$$\mu_G^z(v\ \textit{uncovered})=\frac{1}{1+\sum_{u\in\partial v}Y^v_{u\to v}(z)},$$

where $\mathbf{Y}^v=z\mathcal{R}_{T_{[G,v]}}\mathbf{Y}^v$. As a consequence, we have $\nu(G)=\sum_{v\in V}\frac{1-x_v}{2}$, with $x_v=\left(1+\sum_{u\in\partial v}Y^v_{u\to v}\right)^{-1}$ and $\mathbf{Y}^v$ is the smallest solution to $\mathbf{Y}=\mathcal{Q}_{T_{[G,v]}}\circ\mathcal{R}_{T_{[G,v]}}(\mathbf{Y})$.

Note that our results in previous section show that all quantities are well-defined in the statement of this proposition. Such results have been used to obtain counting algorithm for matchings [11] and sublinear-time algorithms [12]. Note however that Proposition 2 does not apply as each $\mathbf{Y}^v$ is computed on a different tree.

III. Matchings on infinite trees

As explained above, an analysis on (path-)trees allows us to capture most of the information about matchings. In the rest of this paper, we will deal with sequences of graphs with size diverging to infinity and compute the asymptotics for their matching numbers. As shown in [2], [5], this computation can be done using the local weak convergence of graphs and then interpreting the Gibbs distribution (1) on infinite trees. As explained in [6], the analysis made in previous section extends to *unimodular* trees [8]. Note in particular that the operators $\mathcal{R}_G$, $\mathcal{Q}_G$ or $\mathcal{P}_G$ extend to infinite graphs. Given the applications we have in mind, we will here restrict ourselves to multi-type Galton-Watson trees (GWT).

A. Random bipartite graphs and Multi-type Galton-Watson trees

We start by describing a simple ensemble of bipartite graphs between variable and function nodes, which we call

$\mathbb{G}_N = G(N, \Lambda, \Gamma)$, where N is the number of variable nodes, $\Lambda(x) = \sum_{d\geq 0} \Lambda_d x^d$ is the variable-node degree distribution and $\Gamma(x) = \sum_{d\geq 0} \Gamma_d x^d$ is the function-node degree distribution. The number of function nodes is $M = N\alpha$ with $\alpha = \frac{\Lambda'(1)}{\Gamma'(1)}$. We refer to [13] Chapter 9.2 for more details.

The following result was first proved in [2].

Proposition 5. *For a sequence of graphs $\mathbb{G}_N = G(N, \Lambda, \Gamma)$, where Λ and Γ are fixed and with M function nodes, we have*

$$\frac{1}{M}\nu\left(\mathbb{G}_N\right) \to \min_{x\in[0,1]} F(x),$$

where

$$\begin{aligned} F(x) &= 1 - \Gamma\left(1 - \frac{\Lambda'(1-x)}{\Lambda'(1)}\right) \\ &\quad + \frac{\Gamma'(1)}{\Lambda'(1)}\left(1 - \Lambda(1-x) - x\Lambda'(1-x)\right). \end{aligned}$$

We will not give a full proof of this result and refer to [6] for a proof which extends the analysis made in Section II. The general idea is first to show that the random graphs considered are locally tree-like and converge to multi-type Galton-Watson trees. Using the branching property of the GWT, the recursion (16) simplifies into a recursive distributional equation (RDE):

$$I \overset{d}{=} \mathbf{1}\left(\sum_{i=1}^{N^{\hat{\Lambda}}} J_i\right), \qquad J \overset{d}{=} \mathbf{1}\left(\sum_{i=1}^{N^{\hat{\Gamma}}} I_i\right), \tag{20}$$

where $I, I_1, \dots$ are i.i.d. Bernoulli random variables with parameter p_I, $J, J_1, \dots$ are i.i.d. Bernoulli random variables with parameter p_J and $N^{\widehat{\Lambda}}$, $N^{\widehat{\Gamma}}$ are independent random variables with the edge-perspective degree profiles: $\widehat{\Lambda}(x) = \frac{\Lambda'(x)}{\Lambda'(1)}$ and $\widehat{\Gamma}(x) = \frac{\Gamma'(x)}{\Gamma'(1)}$. Taking expectation in (20), we get

$$p_I = \widehat{\Lambda}(1-p_J), \qquad p_J = \widehat{\Gamma}(1-p_I). \tag{21}$$

Now in order to compute the limit $\frac{1}{|U|}\sum_{u\in U}\mathcal{D}_u(\mathbf{Y})$ in (19) (where U is the set of function nodes and $|U| = \alpha N$), we replace the operations $\frac{1}{|U|}\sum_{u\in U}(.)$ and $\frac{1}{|V|}\sum_{v\in V}(.)$ by expectations so that we get:

$$\begin{aligned} \lim_{N\to\infty}\frac{1}{|U|}\sum_{u\in U}\mathbf{1}\left(\sum_{v\in\partial u} I_{v\to u} \geq 1\right) &= \mathbb{P}\left(\sum_{i=1}^{N^\Gamma} I_i \geq 1\right) \\ &= 1 - \Gamma(1-p_I), \end{aligned}$$

where N^Γ is distributed as Γ and $I_1, I_2, \dots$ are i.i.d. Bernoulli r.v. with parameter p_I; and

$$\begin{aligned} &\lim_{N\to\infty}\frac{1}{|U|}\sum_{v\in V}\mathbf{1}\left(\sum_{u\in\partial v}\mathcal{P}_{u\to v}(\mathbf{I}) \geq 2\right) \\ &\quad = \frac{1}{\alpha}\mathbb{P}\left(\sum_{i=1}^{N^\Lambda} J_i \geq 2\right) = 1 - \Lambda(1-p_J) - p_J\Lambda'(1-p_J), \end{aligned}$$

where N^Λ is distributed as Λ and $J_1, J_2, \dots$ are i.i.d. Bernoulli r.v. with parameter p_J. Summing these two terms and using the expression $p_I = \widehat{\Lambda}(1-p_J)$, we obtain $F(p_J)$ as defined in Proposition 5 with p_J solution of the fixed point equation (21). Now differentiating $F(x)$, we obtain:

$$F'(x) = -\frac{\Gamma'(1)}{\Lambda'(1)}\Lambda''(1-x)\left(\widehat{\Gamma}\left(1 - \widehat{\Lambda}(1-x)\right) - x\right).$$

We see that any local minimum of F must satisfy the fixed point equation (21) so that the expression simplifies to the one given in Proposition 5.

B. Correlation decay for $z \to \infty$

In this section, we show that the uniform distribution over maximum matchings will not have the correlation decay property even for simple GWT and explain why standard techniques cannot be applied.In this section, the sequence $(G_n = (V_n, E_n))_{n\in\mathbb{N}}$ is a sequence of finite graphs whose local weak limit is a GWT with degree distribution Λ with finite first moment (i.e. there is no type $\Gamma = \Lambda$).

For any finite graph G, the leaf removal algorithm proceeds as follows: start with the empty matching and then as long as there is a pendant edge $e = (u, v)$ with u of degree one, add this edge to the matching and remove the edge e and all its adjacent edges from the graph. The algorithm stops when there is no more pendant edge. The graph G is thus simplified into a sub-graph with only isolated vertices, matched pairs and a so-called core denoted by $C(G)$ with minimum degree at least 2. Let $LR(G)$ be the number of isolated vertices produced by the leaf-removal algorithm on G. As explained in [14], there exists a maximum matching containing the matched pairs produced by the leaf-removal algorithm. Hence in this maximum matching, the $LR(G)$ isolated vertices will be exposed and we get the bounds:

$$LR(G) \leq 1 - 2\nu(G) \leq LR(G) + |C(G)|.$$

It is clear that these bounds will be tight if we can prove that $|C(G_n)|/|V_n| \to 0$. We now relate this condition to a simple RDE. The analysis of the leaf-removal algorithm has been done in [14, Section 4] (see also [15, Proposition 15] for a more closely related framework). The idea is to analyze the leaf-removal step by step where in one step, all the pendant edges of the current graph are removed. We denote by G_k the graph obtained after k steps. We now put labels on the vertices. First, all isolated vertices of G are of type L. After $k \geq 0$ steps, for all the pendant edges $e = (u, v)$ of G_k with u of degree one in G_k and v of degree at least 2, we say that u is of type L (a leaf of G_k) and v is of type N (v will be covered, i.e. not exposed). All the pendant edges $e = (u, v)$ of G_k with both u and v of degree one in G_k are of type P (they are paired). Let $L_k(G)$ (resp. $N_k(G), P_k(G)$) denote the sets of vertices of type L (resp. N, P) after k steps. Then the number of isolated vertices produced by the leaf-removal algorithm after k steps is given by $LR_k(G) = |L_k(G)| - |N_k(G)|$ and we have the bounds:

$$|C(G)| \leq |V| - |L_k(G)| - |N_k(G)| - |P_k(G)|. \tag{22}$$

The computations of the limits $\lim_n \frac{|L_k(G_n)|}{|V_n|}$, $\lim_n \frac{|N_k(G_n)|}{|V_n|}$ and $\lim_n \frac{P_k(G_n)}{|V_n|}$ can be done thanks to a simple analysis on the limiting tree. Consider a GWT; for any v children of the root $\circ$, let p_k (resp. q_k) be the probability that v is of type L (resp. N) after k-steps of the leaf-removal algorithm. By construction v

is of type N after k steps if and only if one of its children is of type L after k steps, hence we have $q_k = 1 - \widehat{\Lambda}(1 - p_k)$. Similarly, v is of type L after k steps if and only if all its children are of type N after $k-1$ steps, hence $p_k = \widehat{\Lambda}(q_{k-1})$. Hence for all $k \geq 1$, we have $p_k = \widehat{\Lambda}(1 - \widehat{\Lambda}(1 - p_{k-1}))$ and $p_0 = 0$. Since $x \mapsto \widehat{\Lambda}(1-\widehat{\Lambda}(1-x))$ is non-decreasing, p_k converges to p, the smallest solution to the fixed point equation $x = \widehat{\Lambda}(1 - \widehat{\Lambda}(1 - x))$ and q_k converges to $q = 1 - \widehat{\Lambda}(1 - p)$. A careful analysis (done in [15, Proposition 15]) shows that

$$\begin{aligned} \mathbb{P}(\circ \in L_k) &= \Lambda(q_{k-1}) + (1 - q_{k-1} - p_k)\Lambda'(q_{k-1}), \\ \mathbb{P}(\circ \in N_k) &= 1 - \Lambda(1 - p_k) - p_k \Lambda'(q_{k-1}), \\ \mathbb{P}(\circ \in P_k) &= p_k \Lambda'(q_{k-1}), \end{aligned}$$

and a simple coupling argument shows that these quantities correspond to $\lim_n \frac{|L_k(G_n)|}{|V_n|}$, $\lim_n \frac{|N_k(G_n)|}{|V_n|}$ and $\lim_n \frac{|P_k(G_n)|}{|V_n|}$ respectively.

We consider the case where the fixed point equation $x = \widehat{\Lambda}(1-\widehat{\Lambda}(1-x))$ has a unique solution, namely $p = p^*$ where p^* is the unique solution to the fixed point equation $x = \widehat{\Lambda}(1-x)$. In this case, we have $q = 1 - p$ so that by (22), we get

$$\lim_{n\to\infty} \frac{\nu(G_n)}{n} = 1 - \Lambda(1-p^*) - \frac{p^*}{2}\Lambda'(1-p^*). \tag{23}$$

Indeed if $p = p^*$, then p^* is the unique minimum of the function F defined in Proposition 5 and (23) is in accordance with Proposition 5. In words, the leaf-removal algorithm leaves a core of size $o(n)$ and produces a maximum matching on the complementary part of the core.

We now consider the case, where $p < p^*$. In this case, we need to consider the RDE associated to the fixed point equation $\mathbf{Y} = \mathcal{Q}_G \circ \mathcal{R}_G(\mathbf{Y})$ see (14). This RDE has been solved in [15, Theorem 8]. In particular, if $\min_{x\in[0,1]} F(x) = F(p)$ where F is defined in Proposition 5, then this RDE has a unique solution. In this case, the correlation decay still holds and a standard coupling argument similar as described above is sufficient to compute the limit of the matching number. In words, in this case, although the size of the core is macroscopic, we see that the number of uncovered vertices on the core is $o(n)$.

As soon as $\min_{x\in[0,1]} F(x) < F(p)$, there exists no (almost) perfect matching on the core and the correlation decay property fails. The introduction of the measure (1) allows us to find the right solution to the RDE by letting z tend to infinity.

IV. Application: Irregular XORSAT

In this section, we consider the XORSAT decision problem (see Chapter 18 in [13]). An instance is given by a pair $(\mathbb{H}, b)$, where $\mathbb{H}$ is a $M \times N$ binary matrix and b is a binary vector of size M. The XORSAT decision problem requires to answer the question: does there exist a solution to the linear system $\mathbb{H}x = b$, i.e. does b belongs to the image of $\mathbb{H}$?

The following lemma will allow us to make a connection between this problem and our result for matchings. We interpret the matrix $\mathbb{H}$ as the biadjacency matrix of a bipartite graph with M function nodes and N variable nodes and an edge between variable node i and function node a iff $H_{ia} = 1$. We define the binary rank $\mathrm{rk}_2(\mathbb{H})$ as the rank calculated over $GF(2)$.

Lemma 4. *Let $\mathbb{G}$ be the bipartite graph associated to $\mathbb{H}$, then we have $\mathrm{rk}_2(\mathbb{H}) \leq \nu(\mathbb{G})$.*

Proof: Note that for any $k \times k$ submatrices S of $\mathbb{H}$, we have in $GF(2)$, $\det(S) = \sum_{\sigma\in\Sigma_k} \prod_{i=1}^k S_{i,\sigma(i)}$, so that we can have $\det(S) > 0$ only if there exists a perfect matching in the subgraph corresponding to S. The result follows from the fact that there exists a $\mathrm{rk}_2(\mathbb{H}) \times \mathrm{rk}_2(\mathbb{H})$ non-singular submatrix of $\mathbb{H}$. ■

We will consider random large instance of the XORSAT problem. We choose b uniformly at random so that for a given $\mathbb{H}$, the probability that b belongs to the image of $\mathbb{H}$ is simply $2^{\mathrm{rk}_2 \mathbb{H} - M}$. We denote by $\mathbb{H}(N, \Lambda, \Gamma)$ the biadjacency matrix of the graph $\mathbb{G}(N, \Lambda, \Gamma)$.

Corollary 1. *Consider a random XORSAT instance $(\mathbb{H}(N,\Lambda,\Gamma), b)$ then the probability for this instance to be satisfiable goes to zero as N tends to infinity as soon as $\max_{x\in[0,1]} H(x) > 0$, where*

$$\begin{aligned} H(x) &= \Gamma\left(1 - \frac{\Lambda'(1-x)}{\Lambda'(1)}\right) \\ &\quad - \frac{\Gamma'(1)}{\Lambda'(1)}\left(1 - \Lambda(1-x) - x\Lambda'(1-x)\right) \end{aligned}$$

We actually conjecture that

Conjecture 1. *We have as N tends to infinity*

$$\mathrm{rk}_2(\mathbb{H}(N,\Lambda,\Gamma)) - \nu(\mathbb{G}(N,\Lambda,\Gamma)) = o(N). \tag{24}$$

This conjecture is known to hold in the particular of k-XORSAT, when Λ is a Poisson distribution and Γ is deterministic equals to k [16], [17] (see also [3] and [4]). Indeed it is easy to see that when G is a finite tree, there is equality in Lemma 4 and (24) has been proved if one replaces rk_2 by the (regular) rank rk [15].

We end our paper with a numerical example illustrating Corollary 1. We consider a case where all function nodes or variable nodes can have degrees 3 or 15 only. More precisely, the variable-node degree distribution is $\Lambda(x) = \frac{4}{5}x^3 + \frac{1}{5}x^{15}$ and the function-node degree distribution is $\Gamma(x) = bx^3 + (1-b)x^{15}$ with $b = 5/4 - 9/(20\alpha)$ where $\alpha = M/N$ is the ratio of the number of clauses to the number of variables. As α increases, the number of constraints increases and the XORSAT problem becomes less likely to be satisfiable. Applying Corollary 1, we find that for $\alpha > \alpha^*$, the random XORSAT instance becomes non-satisfiable with $0.963025298 < \alpha^* < 0.963025299$ (see Figure 1). It is interesting to note that $\alpha^* < 1$ so that the number of variables is (much) larger than the number of clauses but still the instance is not satisfiable. If Conjecture 1 is true, then α^* should be the threshold for satisfiability of this random XORSAT problem.

A. 1RSB computation at zero temperature

In this section, we follow the non-rigorous approach made in Section 19.3 of [13]. Indeed points (a) and (b) of Exercise

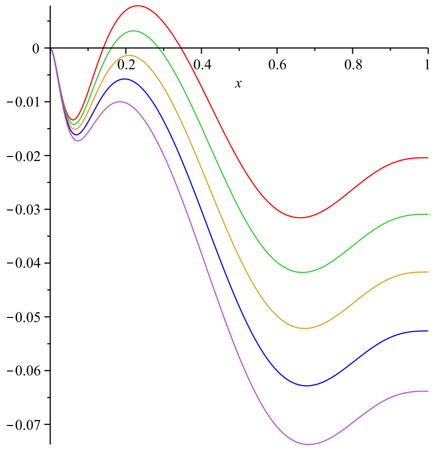

Fig. 1. Function $x \mapsto H(x)$ for $\alpha = 0.98; 0.97; 0.96; 0.95; 0.94$

19.4 is an easy extension of the computation made in the section 19.3 and tells us that the complexity is given by:

$$\Sigma_{\text{tot}} = 1 - \frac{\Lambda'(1)}{\Gamma'(1)}\Gamma\left(1 - \frac{\Lambda'(1-x)}{\Lambda'(1)}\right) - \Lambda(1-x) - x\Lambda'(1-x)$$

where x solves $x = \widehat{\Gamma}\left(1 - \widehat{\Lambda}(1-x)\right)$. It is claimed in Exercise 19.4 [13], that the relevant solution is $x = 1$ but this claim cannot be correct in such generality. If this was the case, we would have $\Sigma_{\text{tot}} = 1 - \frac{\Lambda'(1)}{\Gamma'(1)} = 1 - \alpha$ and then XORSAT would be satisfiable with high probability as long as $\alpha < 1$ contradicting Corollary 1. This suggests a picture quite different from what is described in [13]. We describe it for our example with degrees 3 and 15. First assuming that our conjecture is correct, the satisfiability threshold should be $\alpha_S = \alpha^*$ so that

$$\Sigma_{\text{tot}} = \begin{cases} 1-\alpha & \text{for } \alpha < \alpha^* \\ 0 & \text{otherwise.} \end{cases}$$

Note in particular, that the complexity Σ_{tot} is 'discontinuous' at α^*.

Acknowledgement

The author thanks Guilhem Semerjian for interesting discussions about the 1RSB computation. The author acknowledges the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-11-JS02-005-01 (GAP project).

References

[1] L. Zdeborová and M. Mézard, "The number of matchings in random graph," *Journal of Statistical Mechanics*, vol. 2006, no. 5, p. P05003, 2006.

[2] C. Bordenave, M. Lelarge, and J. Salez, "Matchings on infinite graphs," *Probability Theory and Related Fields*, pp. 1–26, 2012.

[3] S. Cocco, O. Dubois, J. Mandler, and R. Monasson, "Rigorous decimation-based construction of ground pure states for spin-glass models on random lattices," *Physical review letters*, vol. 90, no. 4, p. 047205, 2003.

[4] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina, "Two solutions to diluted p-spin models and XORSAT problems," *J. Statist. Phys.*, vol. 111, no. 3-4, pp. 505–533, 2003.

[5] J. Salez, "Weighted enumeration of spanning subgraphs in locally tree-like graphs," *Random Structures & Algorithms*, 2012.

[6] M. Lelarge, "A new approach to the orientation of random hypergraphs," in *SODA*, Y. Rabani, Ed. SIAM, 2012, pp. 251–264.

[7] M. Leconte, M. Lelarge, and L. Massoulié, "Convergence of multivariate belief propagation, with applications to cuckoo hashing and load balancing," in *SODA*, S. Khanna, Ed. SIAM, 2013, pp. 35–46.

[8] D. Aldous and R. Lyons, "Processes on unimodular random networks," *Electronic Journal of Probability*, vol. 12, pp. 1454–1508, 2007.

[9] M. Leconte, M. Lelarge, and L. Massoulié, "Bipartite graph structures for efficient balancing of heterogeneous loads," in *SIGMETRICS*, P. G. Harrison, M. F. Arlitt, and G. Casale, Eds. ACM, 2012, pp. 41–52.

[10] C. D. Godsil, "Matchings and walks in graphs," *J. Graph Theory*, vol. 5, no. 3, pp. 285–297, 1981.

[11] M. Bayati, D. Gamarnik, D. Katz, C. Nair, and P. Tetali, "Simple deterministic approximation algorithms for counting matchings," in *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*. ACM, 2007, p. 127.

[12] M. Lelarge and H. Zhou, "Sublinear-time algorithms for monomer-dimer systems on bounded degree graphs," *arXiv preprint arXiv:1208.3629*, 2012.

[13] M. Mézard and A. Montanari, *Information, physics, and computation*, ser. Oxford Graduate Texts. Oxford: Oxford University Press, 2009.

[14] R. Karp and M. Sipser, "Maximum matchings in sparse random graphs," *Proc. of the Twenty-second Annual Symposium on Foundations of Computer Science*, vol. IEEE, pp. 364–375, 1981.

[15] C. Bordenave, M. Lelarge, and J. Salez, "The rank of diluted random graphs," *Ann. Probab.*, vol. 39, no. 3, pp. 1097–1121, 2011.

[16] O. Dubois and J. Mandler, "The 3-XORSAT Threshold," in *FOCS '02: Proceedings of the 43rd Symposium on Foundations of Computer Science*. Washington, DC, USA: IEEE Computer Society, 2002, pp. 769–778.

[17] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, and M. Rink, "Tight thresholds for cuckoo hashing via xorsat," in *ICALP (1)*, ser. Lecture Notes in Computer Science, S. Abramsky, C. Gavoille, C. Kirchner, F. M. auf der Heide, and P. G. Spirakis, Eds., vol. 6198. Springer, 2010, pp. 213–225.