The Annealed Pressure of Non-Pairwise Interaction Spin-Glass Models on Random Regular Graphs

Yizhong Richard Hu

Honors' Thesis Advisor: Kavita Ramanan

Post-doc Mentor: Sarath Yasodharan

Graduate Student Mentor: I-Hsun Chen

April 29, 2024

Abstract

Spin glasses, a probability model on random graphs, was introduced as a statistical mechanics model to study disordered magnetic alloys, but they have since become a versatile tool for complex, disordered systems arising in many fields. Previous work has focused on dense random graphs and lattices. Interactions on sparse random graphs, such as random regular graphs, remains to be understood. In these systems, a quantity known as the *quenched pressure* (QP) is of particular interest. In the "replica symmetric phase" of these systems, physicists have proposed heuristics for the quenched pressure, called the *Bethe prediction*. It has been proven to be correct for first-neighbor pairwise interaction models such as the Ising and Potts models on random regular graphs[2, 3], but no rigorous proofs have been provided for non-first-neighbor interactions. In this thesis, we made crucial steps in non-first-neighbor interactions. Using the Next Nearest Neighbor Ising (NNNI) model as a motivating example, we reformulated pairwise next-nearest neighbor interactions as a more general, non-pairwise first-neighborhood interaction. We made the important step of proving that the *annealed pressure* (AP), a quantity closely related to the quenched pressure, is equal to the heuristics given by the Bethe prediction on non-pairwise first-neighborhood interaction models on random regular graphs. It remains to be shown if AP is equal to QP, an important step that will be included in our future work. This work sets the foundation for understanding ranged-interactions in sparse random graph structures.

Contents

1 Introduction

Spin glasses, initially introduced to describe the exotic phase behavior of certain disordered magnetic alloys, but have since become a versatile tool for complex, disordered systems arising in a variety of fields including social sciences, genetics, and combinatorics . They describe large collections of particles that interact locally with respect to an underlying random network: a graph with the edges of the graph representing direct dependencies between particle states, and their randomness representing the disordered state. Early work focused on dense random graphs and the case where the graph is a d-dimensional lattice. Subsequently, to study amorphous materials and the statistical mechanics of disordered media spin glasses on sparse random graphs were investigated.

Mathematically, given a graph, the spin glasses on it is specified by a probability distribution over the configuration space of the states (or "spins") of the particles. The probability of a config is proportional to its energy, and the normalizing constant that makes it a probability distribution is known as the partition function. Specifically, much attention is dedicated to calculating the quenched pressure or limiting free energy density, defined as the limit of the average which is defined as the expected log partition function as the number of particles approaches infinity. The variation in quenched pressure under changes in model parameters provides insight into the phase transition properties, or microscopically, the self-organizing behavior of a large number of particles.

In the "replica symmetric phase" of the system, physicists have provided a non-rigorous heuristic formula for the quenched pressure known as the replica symmetry (RS) solution. The RS solution expresses the quenched pressure in terms of the supremum of the so-called Bethe free energy functional over probability measures that satisfy a fixed-point equation of a *cavity map* on probability measures. The fixed point equation is called the cavity equation. The RS solution is exact on trees, but it has been conjectured that it is also true for certain graphs with loops in the replica symmetric phase. This prediction is known as the Bethe prediction.

Rigorous proofs for the correctness of the cavity equation have focused on first-neighbor pairwise interactions with discrete spin values. Dembo et al. proved that for the Potts model[3] and the Ising model[2] on random regular graphs, the quenched pressure in a certain temperature range is equal to the RS solution. The work considered the more general *factor model*, which includes the Potts model as a special case, and proved that the quenched pressure matches the maximum Bethe free energy over fixed point probability distributions of a *belief propagation* algorithm. This belief propagation fixed point is equivalent to the fixed point of the cavity map for pairwise interactions and is also referred to as the Bethe prediction. The method employed in Dembo et al.'s work considers another quantity, the annealed pressure. The annealed pressure is the limiting average log expectation of the partition function, in contrast to the quenched pressure, which is the limit if the log of the average expectation of the partition function. That is, the order of log and expectations are switched for two quantities. They first showed that the annealed pressure matches the RS solution for all temperature ranges, and then that the annealed pressure is equal to quenched pressure in the replica symmetric regime. Their work used a combinatorial method to show that the annealed pressure can be expressed in terms of an optimization problem, the stationary points of which are characterized by the fixed points of the cavity equation, leading to a formulation of annealed pressure as the maximum of the Bethe free energy functional.

Motivated by non-first-neighbor pairwise interaction models such as the Next Nearest Neighbor Ising (NNNI) model, our work aims to extend this equivalence between the RS solution and quenched pressure to more general models between on random regular graphs. In particular, we first show that non-first neighbor pairwise interaction models can be reformulated as a more complicated, non-pairwise interacting first-neighborhood model. Then, for a class of non-pairwise interacting first neighborhood models, we focus on the first step outlined in Dembo et al.'s work, namely proving the correspondence that annealed pressure is equal to the RS solution. To achieve the latter, we employ a large deviation result [5, 1] to rewrite the annealed pressure as the supremum of some functional. To solve the optimization problem, we find the stationary points using Lagrange multipliers and then relate it to the fixed points of the associated cavity map.

Section 2 presents some basic definitions and known results that expresses the annealed pressure as the supremum of some functional over probability measures on marked stars. We then introduce the RS solution, including the cavity map, the Bethe free-energy functional, and belief propagation. In section 3, we state the main theorem of this thesis, outline the key steps of the proof, and discuss the relation between the fixed points of the cavity map and the belief propagation fixed point. In section 4, we provide the necessary lemmas and the proof of the theorem.

2 Preliminaries

We will mostly be using the notation in Ramanan and Sarath's work [5].

2.1 General notation

Let $\mathbb{N}$ denote the natural numbers $\{1, 2, 3, \cdots\}$ with $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Let $\mathbb{R}$ denote the real numbers, and $\mathbb{R}_+$ denote $(0, \infty)$. $\mathbf{1}\{\cdot\}$ denotes the indicator function. Let $\mathcal{P}(\mathcal{S})$ denote the space of probability measures on $\mathcal{S}$. For a probability measure $\mu \in \mathcal{P}(\mathcal{S})$, we denote its pmf as $\mu(s)$ for each $s \in \mathcal{S}$.

For $\mathcal{S}$ nonempty finite set and $\mu, \nu \in \mathcal{P}(\mathcal{S})$, the relative entropy of μ with respect to ν is

$$H(\mu \| \nu) = \sum_{s \in \mathcal{S}} \mu(s) \log \frac{\mu(s)}{\nu(s)} \tag{2.1}$$

using the convention that $a \log(a/0) = \infty$ for all $a > 0$ and $0 \log(0/0) = 0$.

We denote multisets with square brackets, $[1, 1, 2, 3, \cdots]$. For some nonempty finite set $\mathcal{X}$, we denote its cardinality by $|\mathcal{X}|$. We denote the set of finite multisets where each number comes from $\mathcal{X}$ as $[\mathcal{X}]^*$. Let $\kappa \in \mathbb{N}_0$, we denote the set of multisets with κ elements as $[\mathcal{X}]^\kappa$. For a multiset $\underline{\mathrm{x}} \in [\mathcal{X}]^*$ and $x \in \mathcal{X}$, we use $\underline{\mathrm{x}} \oplus x$ to denote the multiset where x is inserted into $\underline{\mathrm{x}}$

2.2 The space of rooted stars

Let $G = (V, E)$ denote a graph with vertices V and edges E. A tree T is a graph with no loops. For a vertex $v \in V$, $\mathcal{N}_v(G) = \{u \in V : \{u, v\} \in E\}$ denotes the set of neighbors of v. The degree of a vertex v is the number of neighbors of v, that is, $\deg_G(v) = |\mathcal{N}_v(G)|$. A vertex whose degree is 1 in a tree is called a leaf. Let $u, v \in V$, the distance between u and v is denoted $d_G(u, v)$, with disconnected vertices having infinite distance. The vertex set V is assumed to be nonempty and finite.

For a tree $T = (V, E)$ and a vertex $o \in V$, (T, o) is called a rooted tree, with o being the root. For a vertex $v \in V$, its depth is $d_T(o, v)$. The depth of a rooted tree T is the maximum depth of all vertices. Two rooted trees (T_1, o_1) and (T_2, o_2) with $T_1 = (V_1, E_1)$ and $T_2 = (V_2, E_2)$ are considered isomorphic if there exists some bijective $\varphi : V_1 \mapsto V_2$ such that $\varphi(o_1) = o_2$ and $\{u, v\} \in E_1$ if and only if $\{\varphi(u), \varphi(v)\} \in E_2$. Such a φ is called a graph isomorphism. Denote $(T_1, o_1) \simeq (T_2, o_2)$. Let $\mathcal{T}_*$ denote the equivalence classes of rooted trees under isomorphism. $\mathcal{T}_{*,1}$ denotes the set of rooted graphs with depth at most 1. Since such a tree has a root directly connected to leaves with degree 1, it is also called a star. Under the equivalence class, each star is uniquely identified by the degree of the root.

we define regular graphs. For $\kappa \in \mathbb{N}$, a κ-regular graph is a graph where all vertices have degree κ. The set of κ-regular graphs is denoted by $\mathcal{G}^\kappa$. The set of κ-regular graphs with n vertices is denoted $\mathcal{G}_n^\kappa$. A κ-regular tree is a tree where all vertices are either a leaf, or has degree κ. The set of κ-stars is denoted $\mathcal{T}_{*,1}^\kappa$ (see Figure 1).

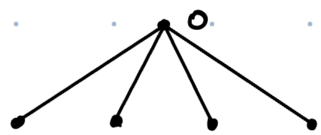


Figure 1: *A star with root degree 4*

Let T be a tree with vertex set V, and $\mathcal{X}$ a nonempty finite set. Mark each vertex in T with $\underline{x} = \{x_v \in \mathcal{X}\}_{v \in V}$. We call $(T, \underline{x})$ a marked graph. Similarly, for a rooted tree (T, o) with a set of marks $\underline{x}$ on T, $(T, o, \underline{x})$ is a rooted marked tree. For some $U \subseteq V$, we denote $\underline{x}_U = \{x_v\}_{v \in U}$. For two trees $T_1 = (V_1, E_1)$ and $T_2 = (V_2, E_2)$, the rooted marked trees $(T_1, o_1, \underline{x}^1)$ and $(T_2, o_2, \underline{x}^2)$ are called isomorphic if there exists some graph isomorphism $\varphi : V_1 \mapsto V_2$ such that $\underline{x}^1_v = \underline{x}^2_{\varphi(v)}$ for all $v \in V_1$.

Let $\mathcal{T}_*[\mathcal{X}]$ denote the set of rooted marked trees with marks $\mathcal{X}$. For some $(T, o, \underline{x}) \in \mathcal{T}_*[\mathcal{X}]$, we denote its root mark as $(T, o, \underline{x})_o = x_o$. We call $\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$ the set of marked κ-stars. Notice that all unmarked κ-stars are isomorphic to each other, the only feature differentiating the marked κ-stars are their marks. It is therefore possible to unique describe a marked κ-star based on with its marks: $(\tau_o, [\tau_v : v \in \mathcal{N}_o(\tau)])$, where $[\cdots]$ denotes a multiset.

We will denote elements of $\mathcal{T}_*[\mathcal{X}]$ with τ, τ', etc. Random variables on the set will be denoted $\boldsymbol{\tau}$, $\boldsymbol{\tau}'$, etc.

2.3 Shortening, cutting and joining rooted marked trees

There are multiple operations we can perform on the trees. Given some $\tau \in \mathcal{T}_*[\mathcal{X}]$, let τ_h denote the tree where we only keep vertices whose distance to o is less than or equal to h. All vertices that are more than h away from o are removed (see Figure 2).

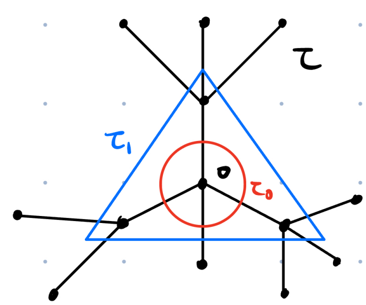


Figure 2: *The tree τ can be shortened to τ_1. Shortening it to height* 0 *gives a tree with only its root with mark τ_o.*

Given some $\tau \in \mathcal{T}_*[\mathcal{X}]$ and an edge $\{u, v\} \in \tau$. We use $\tau(u \backslash v)$ to denote the branch cut of the tree. Namely, we remove the edge $\{u, v\}$ and keep the subtree containing u, making u the new root (see Figure 3 left).

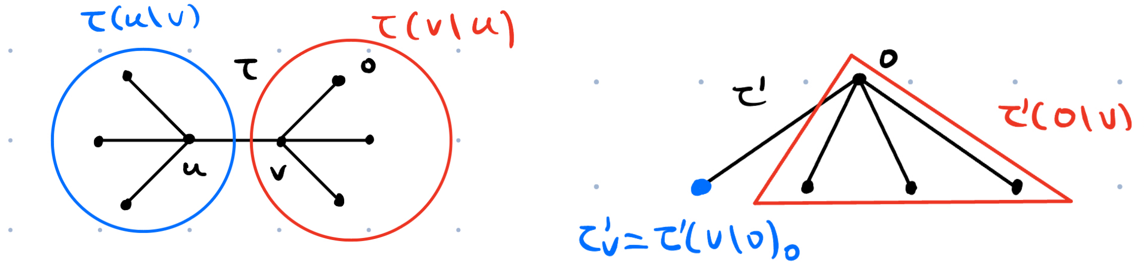


Figure 3: *A tree τ can be cut at $\{u, v\}$ and obtain two subtrees $\tau(u\backslash v)$ and $\tau(v\backslash u)$. A star τ' can be cut at $\{o, v\}$ to obtain the vertex.*

For a star $\tau \in \mathcal{T}_{*,1}[\mathcal{X}]$ with some leaf $v \in \mathcal{N}_o(\tau)$, we denote the mark on vertex v by $\tau_v = \tau(v\backslash o)_o$. Additionally, similar to how we can remove a leaf and obtain $\tau(o\backslash v)$, we can also add a leaf to a star. For some $\tau \in \mathcal{T}_{*,1}^{\kappa-1}[\mathcal{X}]$ and $x \in \mathcal{X}$, we use $\tau \oplus x$ to denote the κ-regular star where an edge is added between some vertex with mark x and the root mark o. Consequently, one can say that for some $\tau \in \mathcal{T}_{*,1}[\mathcal{X}]$ and $v \in \mathcal{N}_o(\tau)$, $\tau = \tau(o\backslash v)_1 \oplus \tau_v$ (see Figure 3 right).

2.4 Factor Graphs

A graph $G = (V, E)$ is bipartite if there exists a bipartition of V, i.e., some $X, Y \subseteq V$ such that $V = X \sqcup Y$, such that all edges connect vertices in one set to another. For a factor graph, we call one set of vertices variable nodes, and the other function nodes. We denote variable nodes with circles and function notes with squares (see Figure 4).

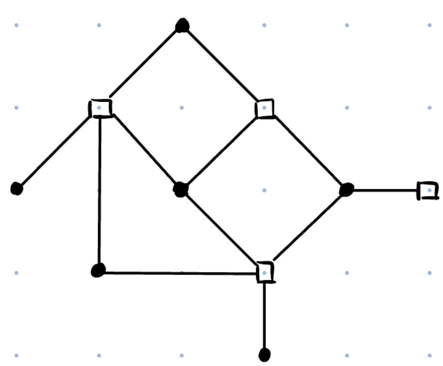


Figure 4: *Example of a factor graph*

A factor graph $G = (V \sqcup F, E)$ is called regular if all the variable nodes have the same degree, and all the function nodes have the same degree. If the variable nodes have degree κ and function nodes have degree λ, we call it a κ,λ-regular factor graph.

2.5 Combinatorics on marked stars

Before defining probability measures, we first need to introduce some quantities to measure some combinatorial properties of marked stars.

Definition 2.1. *Let $\mathcal{X}$ be a nonempty finite set, and $\kappa \in \mathbb{N}$. Let $\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$ and let $x, x' \in \mathcal{X}$. Define*

$$E_1(x, x')(\tau) = |\{v \in \mathcal{N}_o(\tau) : \tau_v = x, \tau_o = x'\}|. \tag{2.2}$$

An equivalent definition can be done with index functions, which we will use frequently:

$$E_1(x, x')(\tau) = \sum_{v \in \mathcal{N}_o(\tau)} \mathbf{1}\{\tau_v = x, \tau_o = x'\} \tag{2.3}$$

For convenience, we will denote $n_x(\tau) = E_1(x, \tau_o)(\tau)$, the number of first-generation vertices with mark x. Denote $\mathbf{n}(\tau) = \{n_x(\tau)\}_{x \in \mathcal{X}}$. Notice that for any $\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$, $\sum_x n_x(\tau) = \kappa$

These quantities, especially $\mathbf{n}(\tau)$, enable us to talk about isomorphic marked stars. Given two distinct leaves u and v. Swapping the marks of u and v will give a star isomorphic to the original, but if u and v have different marks, the set of marks $\underline{\mathrm{x}}$ will be different. This means that multiple mark sets correspond to the same marked star. How many distinct mark sets of a star give isomorphic graphs? This is given by the multinomial coefficient:

Definition 2.2 (Multinomial Coefficient)**.** *Let $\mathcal{X}$ be a nonempty finite set, and $\kappa \in \mathbb{N}$. Let $\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$. Define the multinomial coefficient*

$$\binom{\kappa}{\mathbf{n}(\tau)} = \frac{\kappa!}{\prod_{x \in \mathcal{X}} n_x(\tau)!}, \tag{2.4}$$

where $n!$ represents the factorial of $n \in \mathbb{N}_0$:

$$n! = \begin{cases} 1, & n = 0 \\ n \cdot (n-1)!, & \textit{otherwise} \end{cases} \tag{2.5}$$

Indeed, the above claim is correct:

Lemma 2.3 (Multinomial coefficient is the multiplicity of mark sets)**.** *Given a marked κ-star τ. There are $\binom{\kappa}{\mathbf{n}(\tau)}$ distinct mark sets of τ that results in a tree isomorphic to τ.*

We will also use this fact about multinomial coefficients

Lemma 2.4 (Multiplicity of joined trees)**.** *Let $\tau \in \mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]$, $x \in \mathcal{X}$, the multiplicity*

$$\binom{\kappa}{\mathbf{n}(\tau \oplus x)} = \frac{\kappa}{n_x(\tau \oplus x)} \binom{\kappa - 1}{\mathbf{n}(\tau)} \tag{2.6}$$

Similar to trees, whose multiplicity depends on the leaf marks, represented by a multiset, we can also define multiplicity for a multiset. Given some nonempty finite set $\mathcal{X}$ and $\kappa \in \mathbb{N}_0$, consider some multiset $\underline{\mathrm{x}} \in [\mathcal{X}]^{\kappa}$. For each $x \in \mathcal{X}$, we denote the number of occurences of x in $\underline{\mathrm{x}}$ as $n_x(\underline{\mathrm{x}})$. We denote $\mathbf{n}(\underline{\mathrm{x}}) = \{n_x(\underline{\mathrm{x}})\}_{x \in \mathcal{X}}$. The multinomial coefficient for $\underline{\mathrm{x}}$ is given by

$$\binom{\kappa}{\mathbf{n}(\underline{\mathrm{x}})} = \frac{\kappa!}{\prod_{x \in \mathcal{X}} n_x(\underline{\mathrm{x}})!}. \tag{2.7}$$

2.6 Probability measures on marked stars

First, we define a uniform distribution on marked stars. Let $\mathcal{X}$ be a nonempty finite set, and $\kappa \in \mathbb{N}$. We will use ρ_1 to denote a probability measure on $\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$. Denote the uniform probability measure η_1, where the mark sets are sampled independently and uniformly from $\mathcal{X}$. The probability that some $\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$ will be sampled is

$$\eta_1(\tau) = \frac{1}{|\mathcal{X}|^{\kappa+1}} \binom{\kappa}{\mathbf{n}(\tau)} \tag{2.8}$$

Given some $\rho_1 \in \mathcal{P}(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])$, we sample τ from ρ_1 and a leaf v of τ uniformly and independently. The resulting probability measure on (τ_v, τ_o) is denoted π_{ρ_1}.

Lemma 2.5. *Let $\mathcal{X}$ be a nonempty finite set, $\kappa \in \mathbb{N}$, $\rho_1 \in \mathcal{P}(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])$. The probability measure π_{ρ_1} has pmf*

$$\pi_{\rho_1}(x, x') = \mathbb{E}_{\rho_1}\left[\frac{1}{\kappa} E_1(x, x')(\boldsymbol{\tau})\right], \tag{2.9}$$

and

$$\pi_{\eta_1}(x, x') = \frac{1}{|\mathcal{X}|^2}. \tag{2.10}$$

For some $\rho_1 \in \mathcal{P}(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])$, π_{ρ_1} is called symmetric if $\pi_{\rho_1}(x, x') = \pi_{\rho_1}(x, x')$ for all $x, x' \in \mathcal{X}$. ρ_1 is called admissible if π_{ρ_1} is symmetric. The set of symmetric probability measures on $\mathcal{X}^2$ is denoted $\mathcal{P}_s(\mathcal{X}^2)$. The set of admissible probability measures on $\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$ is denoted $\mathcal{P}_s(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])$.

2.7 Spin glass models

The main topic of discussion in this thesis is the first-neighborhood interaction model.

Definition 2.6 (First Neighborhood Interaction Model)**.** *Let $\mathcal{X}$ be a nonempty finite set, and Hamiltonian $\mathcal{H} : \mathcal{T}_{*,1}[\mathcal{X}] \mapsto \mathbb{R}$, for any finite graph $G = (V, E)$, the Gibbs' measure on the set of marks $\underline{\mathrm{x}} = \{x_v \in \mathcal{X}\}_{v \in V}$ with Hamiltonian $\mathcal{H}$ is given by*

$$\mu_G(\underline{\mathrm{x}}) = \frac{1}{Z_G(\mathcal{H})} \exp\left[\sum_{v \in V} \mathcal{H}\left(x_v, \underline{\mathrm{x}}_{\mathcal{N}_v(G)}\right)\right], \tag{2.11}$$

where the normalization factor $Z_G(\mathcal{H})$, also known as the partition function, is given by

$$Z_G(\mathcal{H}) = \sum_{\underline{\mathrm{x}} \in \mathcal{X}^{|V|}} \exp\left[\sum_{v \in V} \mathcal{H}\left(x_v, \underline{\mathrm{x}}_{\mathcal{N}_v(G)}\right)\right]. \tag{2.12}$$

There are several models from statistical physics that fall within this framework. We provide two examples below.

Example 2.7 (Ising model). *Let $\beta, B \in \mathbb{R}$. With $\mathcal{X} = \{+1, -1\}$, and $\mathcal{H}(\tau) = \sum_{v \in \mathcal{N}_o(\tau)} \frac{\beta}{2}\tau_o\tau_v + B\tau_o$, we obtain the Ising model, with Gibbs' measure [4]*

$$\mu_G(\underline{\mathrm{x}}) = \frac{1}{Z_G(\beta, B)} \exp\left[\sum_{\{u,v\}\in E} \beta x_u x_v + \sum_{v \in V} B x_v\right], \tag{2.13}$$

with $Z_G(\beta, B) = \sum_{\underline{\mathrm{X}} \in \mathcal{X}^{|V|}} \exp\left[\sum_{\{u,v\}\in E} \beta x_u x_v + \sum_{v\in V} B x_v\right]$.

There has also been interest in studying spin models beyond first neighborhood interactions, see [6]. In our framework, this model can be viewed as a first neighborhood interaction model.

Example 2.8 (Next Nearest Neighbor Ising Model). *Let $\mathcal{X} = \{+1, -1\}$, $\lambda \in \mathbb{R}$ and $\mathcal{H}(\tau) = \sum_{v \in \mathcal{N}_o(\tau)} \frac{\lambda}{2}\tau_o\tau_v - \sum_{\{u,v\}\subseteq \mathcal{N}_o(\tau)} \tau_u\tau_v$, we have the Next Nearest Neighbor Ising (NNNI) model, with Gibbs' measure*

$$\mu_G(\underline{\mathrm{x}}) = \frac{1}{Z_G(\lambda)} \exp\left[\sum_{\{u,v\}\in E} \lambda x_u x_v - \sum_{\{u,v\}\subseteq V} \mathfrak{n}(u,v) x_u x_v\right], \tag{2.14}$$

where $\mathfrak{n}(u,v) = |w \in V : \{u,w\}, \{w,v\} \in E|$ denotes the number of paths of length 2 between u and v, and $Z_G(\lambda) = \sum_{\underline{\mathrm{X}} \in \mathcal{X}^{|V|}} \exp\left[\sum_{\{u,v\}\in E} \lambda x_u x_v - \sum_{\{u,v\}\subseteq V} \mathfrak{n}(u,v) x_u x_v\right]$.

In other words, pairwise interactions happen between both nearest neighbors and next-nearest neighbors, hence the name of the model.

For the rest of this thesis, we will only discuss regular graphs and regular trees. We will denote the degree with $\kappa \in \mathbb{N}$. In this case, the Hamiltonian density $\mathcal{H}$ will only need to be a function on marked κ-stars.

Definition 2.9 (Pressure). *Let $\mathcal{X}$ be a nonempty finite set, $\kappa \in \mathbb{N}$, $\mathcal{H} : \mathcal{T}_{*,1}^{\kappa}[\mathcal{X}] \mapsto \mathbb{R}$. Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of graphs where each G_n is sampled independently and uniformly from $\mathcal{G}_n^{\kappa}$. For a first neighbor interaction model with marks $\mathcal{X}$ and Hamiltonian $\mathcal{H}$, the quenched pressure (QP) on random κ-regular graphs is defined by*

$$QP: \quad \Phi(\mathcal{H}) = \lim_{n\to\infty} \frac{1}{n} \mathbb{E} \log[Z_{G_n}(\mathcal{H})]. \tag{2.15}$$

The annealed pressure (AP) on random κ-regular graphs is defined by

$$AP: \quad \Psi(\mathcal{H}) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[Z_{G_n}(\mathcal{H})]. \tag{2.16}$$

Note that since $Z_{G_n}(\mathcal{H})$ is only dependent on the Hamiltonian and the underlying regular graph G_n, the expectation in (2.15) and (2.16) are over the randomness in G_n.

It is known that (see [1, 5]) the annealed pressure for the first-neighborhood interaction model is given by

$$\Psi(\mathcal{H}) = \sup_{\rho_1 \in \mathcal{P}_s(\mathcal{T}_{*,1}^{\kappa}[\mathcal{X}])} \left\{\mathbb{E}_{\rho_1}[\mathcal{H}(\boldsymbol{\tau})] - \left[H(\rho_1\|\eta_1) - \frac{\kappa}{2} H(\pi_{\rho_1}\|\pi_{\eta_1})\right]\right\} + \log|\mathcal{X}|. \tag{2.17}$$

For simplicity, define

$$V_{\mathcal{H}}(\rho_1) = \mathbb{E}_{\rho_1}[\mathcal{H}(\boldsymbol{\tau})] - \left[H(\rho_1\|\eta_1) - \frac{\kappa}{2}H(\pi_{\rho_1}\|\pi_{\eta_1})\right] \tag{2.18}$$

The quenched pressure is predicted to follow the Bethe prediction, which says that the quantity is one of the stationary points of the cavity equation:

Definition 2.10 (Cavity Equation)**.** *Let $\mathcal{X}$ be a nonempty finite set, $\kappa \in \mathbb{N}$, $\mathcal{H} : \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}] \mapsto \mathbb{R}$. The cavity equation of a first neighbor interaction model with Hamiltonian density $\mathcal{H}$ on $\mathcal{X}$ for κ-regular graphs is a functional $\Gamma : \mathcal{P}(\mathcal{X}^2) \mapsto \mathcal{P}(\mathcal{X}^2)$ that sends $h \in \mathcal{P}(\mathcal{X}^2)$ to Γh, where Γh is a measure with pmf given, for all $x, x' \in \mathcal{X}$, by*

$$(\Gamma h)(x, x') = \frac{1}{z_h} \sum_{\tau \in \mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]} \mathbf{1}\{\tau_o = x\} \binom{\kappa - 1}{\mathbf{n}(\tau)} \exp\left[\mathcal{H}(\tau \oplus x')\right] \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, x), \tag{2.19}$$

where

$$z_h = \sum_{x, x' \in \mathcal{X}} \sum_{\tau \in \mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]} \mathbf{1}\{\tau_o = x\} \binom{\kappa - 1}{\mathbf{n}(\tau)} \exp\left[\mathcal{H}(\tau \oplus x')\right] \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, x). \tag{2.20}$$

The fixed-point equation $h = \Gamma h$ is called the cavity equation.

Note that $(\Gamma h)(x, x')$ is a non-negative function with normalization constant z_h, thus making it a valid pmf.

The quenched pressure is predicted to be equal to the Bethe free-entropy density[6].

Definition 2.11 (Bethe free-energy functional)**.** *Assume that $h \in \mathcal{P}(\mathcal{X}^2)$ satisfies $h = \Gamma h$. The Bethe free-energy of a first neighborhood interaction model on κ-regular graphs is*

$$\Phi_{BF}(\mathcal{H}, h) = \log\left[\sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \binom{\kappa}{\mathbf{n}(\tau)} \exp \mathcal{H}(\tau) \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, \tau_o)\right] - \frac{\kappa}{2} \log\left[\sum_{x, x' \in \mathbf{x}} h(x, x')h(x', x)\right]. \tag{2.21}$$

The RS solution says that for certain temperatures of the system, $\Phi(\mathcal{H}) = \sup_{\substack{h \in \mathcal{P}(\mathcal{X}^2) \\ h = h\Gamma}} \Phi_{\mathrm{BF}}(\mathcal{H}, h)$. A non-rigorous statement would be that when the number of vertices $n \to \infty$, the free energy for κ-regular graphs $\mathbb{F} = \mathbb{E} \log Z_{G_n}(\mathcal{H}) \simeq n\Phi_{\mathrm{BF}}(\mathcal{H}, h)$ for some h that satisfies $h = \Gamma h$. In other words, the free energy is roughly

$$\mathbb{F} = |V| \log\left[\sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \binom{\kappa}{\mathbf{n}(\tau)} \exp \mathcal{H}(\tau) \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, \tau_o)\right] - |E| \log\left[\sum_{x, x' \in \mathbf{x}} h(x, x')h(x', x)\right]. \tag{2.22}$$

Using similar methods, we can obtain a similar formula for regular factor graphs.

2.8 Factor Models

Factor model is a generalization of first neighborhood interactions, described by Mézard and Montanari [4]. For factor models, interaction is associated with each factor node. Its neighbors, which are all variable nodes, contain some value that are constrained by factor nodes with a compatibility function

Definition 2.12 (Factor model)**.** *Let $\mathcal{X}$ be a nonempty finite set, and compatibility function $\psi : [\mathcal{X}]^* \mapsto \mathbb{R}_+$, for any finite factor graph $G = (V \sqcup F, E)$, the Gibb's measure on the set of marks $\underline{x} = \{x_v \in \mathcal{X}\}_{v \in V}$ with compatibility function ψ is given by*

$$\mu_G(\underline{x}) = \frac{1}{Z_G(\psi)} \prod_{a \in F} \psi\left(\underline{x}_{\mathcal{N}_a(G)}\right), \tag{2.23}$$

where the normalization factor $Z_G(\psi)$, also known as the partition function, can be expressed as

$$Z_G(\psi) = \sum_{\underline{x} \in \mathcal{X}^{|V|}} \prod_{a \in F} \psi\left(\underline{x}_{\mathcal{N}_a(G)}\right). \tag{2.24}$$

It generalizes first neighborhood interaction models in the following way. Assume some first neighborhood interaction model with nonempty finite $\mathcal{X}$ and Hamiltonian $\mathcal{H}$ on some graph $G = (V, E)$. Construct a factor graph based on G by creating a factor node for each interaction between a vertex $v \in V$ and its neighbors $\mathcal{N}_v(G)$. In other words, for each $v \in V$, create a factor node a that connects to v and $\mathcal{N}_v(G)$. The compatibility function at such a factor node can be specified as

$$\psi(\underline{x}_{\mathcal{N}_a(G)}) = \psi(x_v, \underline{x}_{\mathcal{N}_v(G)}) = \exp[\mathcal{H}(x_v, \underline{x}_{\mathcal{N}_v(G)})]. \tag{2.25}$$

This results in the same Gibb's measure and partition function as the first neighborhood interaction model.

The free energy prediction for factor graphs[4] is given by an algorithm called belief propagation, also known as the message-passing algorithm. Similar to the cavity equation, it is a fixed-point equation. Consider the set of probability measures called "messages": $\underline{\nu} = \{\nu, \hat{\nu} \in \mathcal{P}(\mathcal{X})\}$.

Definition 2.13 (Belief propagation)**.** *Let $\mathcal{X}$ be a nonempty finite set, $\kappa, \lambda \in \mathbb{N}$, and compatibility function $\psi : [\mathcal{X}]^* \mapsto \mathbb{R}_+$. The belief propagation of the factor model with compatibility function ψ on κ, λ-regular factor graphs is a functional BP mapping some message $\underline{\nu}$ to some other message $BP\underline{\nu}$. For each $x \in \mathcal{X}$*

$$\begin{aligned}(BP\nu)(x) &= \frac{1}{z}\hat{\nu}(x)^{\kappa-1} \\ (BP\hat{\nu})(x) &= \frac{1}{\hat{z}} \sum_{\underline{x} \in [\mathcal{X}]^{\lambda-1}} \binom{\lambda-1}{\mathbf{n}(\underline{x})} \psi(\underline{x} \oplus x) \prod_{x' \in \underline{x}} \nu(x')\end{aligned} \tag{2.26}$$

where

$$\begin{aligned} z &= \sum_{x\in\mathcal{X}} \widehat{\nu}(x)^{\kappa-1} \\ \widehat{z} &= \sum_{x\in\mathcal{X}} \sum_{\underline{\mathrm{X}}\in[\mathcal{X}]^{\lambda-1}} \binom{\lambda-1}{\mathbf{n}(\underline{\mathrm{x}})} \psi(\underline{\mathrm{x}}\oplus x) \prod_{x'\in\underline{\mathrm{X}}} \nu(x') \end{aligned} \tag{2.27}$$

Given some message $\underline{\nu}$, the Bethe free energy of factor models on κ,λ-regular factor graphs are given by

$$\mathbb{F} = |F| \log\left[\sum_{\underline{\mathrm{X}}\in[\mathcal{X}]^{\lambda}} \binom{\lambda}{\mathbf{n}(\underline{\mathrm{x}})} \psi(\underline{\mathrm{x}}) \prod_{x\in\underline{\mathrm{X}}} \nu(x)\right] + |V| \log\left[\sum_{x\in\mathcal{X}} \widehat{\nu}(x)^{\kappa}\right] - |E| \log\left[\sum_{x\in\mathcal{X}} \nu(x)\widehat{\nu}(x)\right]. \tag{2.28}$$

3 Results

3.1 Annealed Pressure

Consider ρ_1 and π_{ρ_1} as real vectors over $\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$ and $\mathcal{X}^2$ respectively. For $\rho_1 \in \mathbb{R}^{|\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]|}$ and $\pi_{\rho_1} \in \mathbb{R}^{|\mathcal{X}^2|}$, denote

$$\widetilde{V}_{\mathcal{H}}(\rho_1, \pi_{\rho_1}) = \mathbb{E}_{\rho_1}[\mathcal{H}(\boldsymbol{\tau})] - \left[H(\rho_1 \| \eta_1) - \frac{\kappa}{2} H(\pi_{\rho_1} \| \pi_{\eta_1})\right]. \tag{3.1}$$

We can transform the problem into a constraint optimization problem on ρ_1 and π_{ρ_1} bound by (2.9). The optimization problem becomes

$$\Psi(\mathcal{H}) = \sup_{\substack{\rho_1 \in \mathcal{P}_s(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]),\ \pi_{\rho_1} \in \mathcal{P}_s(\mathcal{X}^2) \\ \mathbb{E}_{\rho_1}[E_1(x,x')(\boldsymbol{\tau})] = \kappa \pi_{\rho_1}(x,x')}} \widetilde{V}_{\mathcal{H}}(\rho_1, \pi_{\rho_1}) + \log |\mathcal{X}|. \tag{3.2}$$

Theorem 3.1 (Annealed pressure is equal to maximum Bethe free-energy)**.** *Assume that in the optimization problem in (3.2) is always attained in the interior point of the probability simplex, then*

1. *the annealed pressure of first-neighborhood interation model with Hamiltonian $\mathcal{H}$ on random κ-regular graphs is*

$$\Psi(\mathcal{H}) = \sup_{\rho_1 \in \mathcal{P}_s(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])} V_{\mathcal{H}}(\rho_1) + \log |\mathcal{X}| = \sup_{\substack{h \in \mathcal{P}(\mathcal{X}^2) \\ h = \Gamma h}} \Phi_{BF}(\mathcal{H}, h), \tag{3.3}$$

2. *for any $h^* \in \mathcal{P}(\mathcal{X}^2)$ such that $h^* = \Gamma h^*$ and $\Psi(\mathcal{H}) = \Phi_{BF}(\mathcal{H}, h^*)$, the exists some $\rho_1^* \in \mathcal{P}_s(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])$ such that $\Psi(\mathcal{H}) = V_{\mathcal{H}}(\rho_1^*)$, with its pmf given by, for all $\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$*

$$\rho_1^*(\tau) = \frac{1}{z_{\rho_1^*}} \binom{\kappa}{\mathbf{n}(\tau)} \exp[\mathcal{H}(\tau)] \prod_{v \in \mathcal{N}_o(\tau)} h^*(\tau_v, \tau_o) \tag{3.4}$$

   with normalizing constant

$$z_{\rho_1^*} = \sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \binom{\kappa}{\mathbf{n}(\tau)} \exp[\mathcal{H}(\tau)] \prod_{v \in \mathcal{N}_o(\tau)} h^*(\tau_v, \tau_o), \tag{3.5}$$

3. *and for any $\rho_1^* \in \mathcal{P}_s(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])$ such that $\Psi(\mathcal{H}) = V_{\mathcal{H}}(\rho_1^*)$, there exists some $h^* \in \mathcal{P}(\mathcal{X}^2)$ such that $h^* = \Gamma h^*$, $\Psi = \Phi_{BF}(\mathcal{H}, h^*)$, with ρ_1^* given by the above equation.*

We now provide an outline of the proof. Use Lagrange multiplier to find the stationary points of $\widetilde{V}_{\mathcal{H}}(\rho_1, \pi_{\rho_1})$ in the domain. We need the following constraints:

$$\sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \rho_1(\tau) = 1$$

$$\rho_1(\tau) \geqslant 0 \qquad \forall \tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$$

$$\sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \rho_1(\tau) \left[\frac{1}{\kappa} E_1(x, x')(\tau)\right] = \pi_{\rho_1}(x, x') \qquad \forall x, x' \in \mathcal{X}$$

$$\pi_{\rho_1}(x, x') = \pi_{\rho_1}(x', x) \qquad \forall x, x' \in \mathcal{X}$$

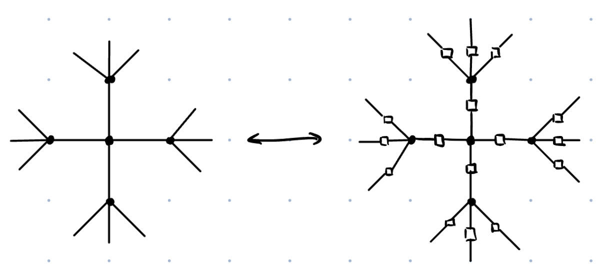
$$\sum_{x, x' \in \mathcal{X}} \pi_{\rho_1}(x, x') = 1$$

Note that this guarantees that π_{ρ_1} is a symmetric probability measure on $\mathcal{X}^2$. It is non-negative since E_1 is non-negative.

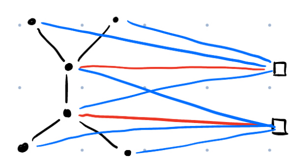
Solving the Lagrange multiplier problem allows us to find the cavity equation, an important functional that underlies all stationary points of $\widetilde{V}_{\mathcal{H}}$. We can show that $h^* = \Gamma h^*$ at stationary points, and taking the max over all stationary points gives the optimal value.

3.2 Relation to Belief Propagation

From the previous section, we showed that factor graph is a generalization of first-neighborhood interaction models. Indeed, for first-neighborhood pairwise interactions like the Ising model, we can associate a factor node with each edge (see Figure 5a), and the resulting belief propagation, under some transformations, would match the cavity equation.



(a) *The factor graph corresponding to a first-neighborhood pairwise interaction on a tree is still a tree.*



(b) *The factor graph for first-neighborhood non-pairwise interaction on a 3-regular graph with radius 3. Red edges represent connections to the center vertex, and blue edges represent connections to its neighbors*

One can observe various formal similarities between the two stationary point equations, as similar ideas have been used to arrive at these conclusions. However, it is not apparent for general first-neighborhood models, the cavity equation as proven, is equivalent to the more general belief propagation. Why is this the case? Belief propagation and cavity equation are both based on the idea that the underlying graph is a tree. Belief propagation is proven to be exact on a factor tree [4]. While the factor graph for Ising model on a tree is still a tree, factor graph corresponding to general first-neighborhood interactions usually contain short loops even if the underlying graph is a tree (see Figure 5b). This means that for the cases where our cavity equation can be proven to be exact, the corresponding factor graph may not be.

Additionally, using the factor graph model as a generalization for first-neighborhood interactions means that the center of the interaction, or the root of the κ-regular tree, needs to be treated as distinct from the rest. The factor node has a "special" connection to their center of interaction, (marked in red in Figure 5b), and thus requires a different message. The resulting iterative formula is not immediately obvious to interpret, especially in a factor graph with loops. And it is unknown if the properties that come with having identical neighbor connections, one of the more important properties we leveraged, still applies to factor models.

4 Proofs

For the rest of this section, assume that $\mathcal{X}$ is a nonempty finite set, and $\kappa \in \mathbb{N}$.

4.1 Preliminary Results

Proof of Lemma 2.3. Since the tree is isomorphic under permutation of the leafs, the number of leaves with some leaf mark x, given by $\mathbf{n}(\tau)$, along with the root mark τ_o, uniquely identifies a marked star under isomorphism. We therefore only need to count the number of distinct ways to label the leaves without changing $\mathbf{n}(\tau)$. Given any label $\underline{\mathrm{x}}$, it has $\kappa!$ permutations. Some of them are considered equal. Namely, permutations that change the order of leaves with the same labels does not change $\underline{\mathrm{x}}$. For each label $x \in \mathcal{X}$, there are $n_x(\tau)$ ways to permute their order, for each permutation, there are $\prod_x n_x(\tau)!$ ways to permute the leaves without changing the label $\underline{\mathrm{x}}$. This gives the multiplicity $\binom{\kappa}{\mathbf{n}(\tau)}$ □

Proof of Lemma 2.4. To do this, notice that for all $x' \in \mathcal{X}$

$$n_{x'}(\tau \oplus x) = \begin{cases} n_{x'}(\tau) & x' \neq x \\ n_x(\tau) + 1 & x' = x \end{cases}$$

So we can separate the term in the product

$$\begin{aligned}
\binom{\kappa}{\mathbf{n}(\tau \oplus x)} &= \frac{\kappa!}{\prod_{x' \in \mathcal{X}} n_{x'}(\tau \oplus x)!} \\
&= \frac{1}{n_x(\tau) + 1!} \frac{\kappa!}{\prod_{x' \neq x} n_{x'}(\tau)!} \\
&= \frac{1}{(n_x(\tau) + 1)\ n_x(\tau)!} \frac{\kappa(\kappa - 1)!}{\prod_{x' \neq x} n_{x'}(\tau)!} \\
&= \frac{\kappa}{n_x(\tau) + 1} \frac{(\kappa - 1)!}{\prod_{x' \in \mathcal{X}} n_{x'}(\tau)!}
\end{aligned}$$

Using the definitions again,

$$= \frac{\kappa}{n_x(\tau \oplus x)} \binom{\kappa - 1}{\mathbf{n}(\tau)}$$

And thus we obtain the result. □

Proof of Lemma 2.5. Let τ be a marked star sampled from ρ_1. The probability of choosing any $v \in \mathcal{N}_o(\tau)$ is $1/\kappa$. Given that the vertex chosen is v, we have that the probability that $(\tau_v, \tau_o) = (x, x')$ is $\mathbf{1}\{\tau_v = x, \tau_o = x'\}$. So

$$\pi_{\rho_1}(x, x') = \sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \rho_1(\tau) \sum_{v \in \mathcal{N}_o(\tau)} \frac{1}{\kappa} \mathbf{1}\{\tau_v = x, \tau_o = x'\} = \mathbb{E}_{\rho_1}\left[\frac{1}{\kappa} E_1(x, x')(\boldsymbol{\tau})\right].$$

Similarly, we know that once v is chosen, τ_v and τ_o are uniformly distributed with probability $1/|\mathcal{X}|$. This therefore gives $\pi_{\eta_1}(x,x') = \sum_{v\in\mathcal{N}_o(\tau)} \frac{1}{\kappa}\frac{1}{|\mathcal{X}|^2} = \frac{1}{|\mathcal{X}|^2}$. □

Lemma 4.1. *For $f : \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}] \mapsto \mathbb{R}$ and $g : \mathcal{X}^2 \mapsto \mathbb{R}$, let*

$$\pi_f(x,x') = \sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau)\frac{1}{\kappa}E_1(x,x')(\tau)$$

for all $x, x' \in \mathcal{X}$. We have

$$\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau) \sum_{v\in\mathcal{N}_o(\tau)} g(\tau_v,\tau_o) = \sum_{x,x'\in\mathcal{X}} \pi_f(x,x')\kappa g(x,x') \tag{4.1}$$

As a result of this, notice that for probabilities measures ρ_1 on κ-regular stars, the above statement can be converted to an expected value statement:

$$\mathbb{E}_{\rho_1}\left[\sum_{v\in\mathcal{N}_o(\boldsymbol{\tau})} g(\boldsymbol{\tau}_v,\boldsymbol{\tau}_o)\right] = \kappa\mathbb{E}_{\pi_{\rho_1}} g(\mathbf{x},\mathbf{x}') \tag{4.2}$$

Proof. Notice that

$$\sum_{x,x'\in\mathcal{X}} \pi_f(x,x')\kappa g(x,x')$$

$$= \sum_{x,x'\in\mathcal{X}} \sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau)\frac{1}{\kappa}E_1(x,x')(\tau)\kappa g(x,x')$$

The κ's cancel

$$= \sum_{x,x'\in\mathcal{X}} \sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau)E_1(x,x')(\tau)g(x,x')$$

Expand definition of E_1

$$\begin{aligned}
&= \sum_{x,x'\in\mathcal{X}} \sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau) \sum_{v\in\mathcal{N}_o(\tau)} \mathbf{1}\{\tau_v = x, \tau_o = x'\}g(x,x') \\
&= \sum_{x,x'\in\mathcal{X}} \sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau) \sum_{v\in\mathcal{N}_o(\tau)} \mathbf{1}\{\tau_v = x, \tau_o = x'\}g(\tau_v,\tau_o) \\
&= \sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau) \sum_{v\in\mathcal{N}_o(\tau)} \sum_{x,x'\in\mathcal{X}} \mathbf{1}\{\tau_v = x, \tau_o = x'\}g(\tau_v,\tau_o) \\
&= \sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau) \sum_{v\in\mathcal{N}_o(\tau)} g(\tau_v,\tau_o).
\end{aligned}$$

And the expected value is a sum that can be rewritten in the above form. □

Lemma 4.2. *For $f : \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}] \mapsto \mathbb{R}$, let $\pi_f(x,x') = \sum_\tau f(\tau)E_1(x,x')(\tau)$, have*

$$\sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau) = \sum_{x,x' \in \mathcal{X}} f(x,x')$$

Proof. Apply (4.1) with $g(x,x') = 1/\kappa$

$$\sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau) = \sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} f(\tau) \sum_{v \in \mathcal{N}_o(\tau)} \frac{1}{\kappa} = \sum_{x,x' \in \mathcal{X}} \pi_f(x,x') \ \kappa \ \frac{1}{\kappa} = \sum_{x,x' \in \mathcal{X}} \pi_f(x,x')$$

□

This means that the constraint $\sum_\tau \rho_1(\tau) = 1$ is redundant. It is satisfied if $\sum_{x,x'} \pi_{\rho_1}(x,x') = 1$.

4.2 Proof of Theorem 3.1

Proof of Theorem 3.1. Since $\log|\mathcal{X}|$ us a constant and it does not play a role in the optimization problem, in the rest of the proof, we focus on the optimization problem:

$$\sup_{\rho_1 \in \mathcal{P}_s(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])} \left\{ \mathbb{E}_{\rho_1}[\mathcal{H}(\boldsymbol{\tau})] - \left[H(\rho_1 \| \eta_1) - \frac{\kappa}{2} H(\pi_{\rho_1} \| \pi_{\eta_1}) \right] \right\}$$

Assume that ρ_1 is not related to π_{ρ_1}, we can define π_{ρ_1} with ρ_1

$$= \sup_{\rho_1 \in \mathcal{P}_s(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}])} \sup_{\substack{\pi_{\rho_1} \in \mathcal{P}_s(\mathcal{X}^2) \\ \kappa \pi_{\rho_1}(x,x') = \mathbb{E}_{\rho_1}[E_1(x,x')(\boldsymbol{\tau})]}} \left\{ \mathbb{E}_{\rho_1}[\mathcal{H}(\boldsymbol{\tau})] - \left[H(\rho_1 \| \eta_1) - \frac{\kappa}{2} H(\pi_{\rho_1} \| \pi_{\eta_1}) \right] \right\}$$

which is equivalent to optimizing over both variables with the constraint $\pi_{\rho_1}(x,x') = \mathbb{E}_{\rho_1}\left[\frac{1}{\kappa} E_1(x,x')(\boldsymbol{\tau})\right]$

$$= \sup_{\substack{\rho_1 \in \mathcal{P}_s(\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]), \pi_{\rho_1} \in \mathcal{P}_s(\mathcal{X}^2) \\ \kappa \pi_{\rho_1}(x,x') = \mathbb{E}_{\rho_1}[E_1(x,x')(\boldsymbol{\tau})]}} \left\{ \mathbb{E}_{\rho_1}[\mathcal{H}(\boldsymbol{\tau})] - \left[H(\rho_1 \| \eta_1) - \frac{\kappa}{2} H(\pi_{\rho_1} \| \pi_{\eta_1}) \right] \right\}.$$

Now we view the pmf's ρ_1 and π_{ρ_1} as real vectors with normality constraints, saying that pmf's are non-negative and sum to 1.

$$\begin{aligned} \rho_1(\tau) &\geqslant 0 & \forall \tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}] \\ \sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \rho_1(\tau) &= 1 \\ \pi_{\rho_1}(x,x') &\geqslant 0 & \forall x, x' \in \mathcal{X} \\ \sum_{x,x'} \pi_{\rho_1}(\tau) &= 1 \end{aligned}$$

We denote each component of ρ_1 and π_{ρ_1} as $\rho_1(\tau)$ and $\pi_{\rho_1}(x,x')$ respectively.

Additionally, we have the constraints that π_{ρ_1} is symmetric and definition of π_{ρ_1}

$$\sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \rho_1(\tau) \left[\frac{1}{\kappa} E_1(x,x')(\boldsymbol{\tau})\right] = \pi_{\rho_1}(x,x') \qquad \forall x, x' \in \mathcal{X}$$

$$\pi_{\rho_1}(x,x') = \pi_{\rho_1}(x',x) \qquad \forall x, x' \in \mathcal{X}$$

Notice that $\pi_{\rho_1}(x,x') = \mathbb{E}_{\rho_1}\left[\frac{1}{\kappa} E_1(x,x')(\boldsymbol{\tau})\right]$ and $\rho_1(x,x') \geqslant 0$ implies that $\pi_{\rho_1}(\tau) \geqslant 0$. $\sum_{x,x'} \pi_{\rho_1}(\tau) = 1$ implies $\sum_{\tau} \rho_1(\tau) = 1$ by Lemma 4.2.

We denote the unconstrained optimization target as

$$\widetilde{V}_{\mathcal{H}}(\rho_1, \pi_{\rho_1}) = \mathbb{E}_{\rho_1}[\mathcal{H}(\boldsymbol{\tau})] - \left[H(\rho_1 \| \eta_1) - \frac{\kappa}{2} H(\pi_{\rho_1} \| \pi_{\eta_1})\right] \tag{4.3}$$

The goal is to find the stationary points of this function given the above constraints.

Before we start doing Lagrange multipliers to find the stationary points, we first examine the optimization space. Without any constraints, the vector space is some $\mathbb{R}^{|\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]|+|\mathcal{X}^2|}$. Consider only the equality constraints, which are all affine linear on ρ_1 and π_{ρ_1}, we obtain an affine subspace of this $\mathbb{R}^{|\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]|+|\mathcal{X}^2|}$. The points where $\rho_1(\tau) = 0$ for some $\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$, we will call the boundaries of this space. Since we made the assumption that the optimum is not on the boundary of the probability simplex, examining the stationary points of this function is sufficient.

Assigning each remaining constraint a Lagrange multiplier:

$$\sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \rho_1(\tau) \left[\frac{1}{\kappa} E_1(x,x')(\boldsymbol{\tau})\right] = \pi_{\rho_1}(x,x') \qquad \alpha(x,x')$$

$$\pi_{\rho_1}(x,x') = \pi_{\rho_1}(x',x) \qquad \beta(x,x')$$

$$\sum_{x,x' \in \mathcal{X}} \pi_{\rho_1}(x,x') = 1 \qquad \gamma$$

$$-\rho_1(\tau) \leqslant 0 \qquad \delta(\tau),$$

and $\delta(\tau) \geqslant 0$, $\rho_1(\tau)\delta(\tau) = 0$ for all $\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$. We have

$$\begin{aligned}\mathcal{L}(\rho_1, \pi_{\rho_1}, \alpha, \beta, \gamma, \delta) &= \mathbb{E}_{\rho_1}\mathcal{H}(\boldsymbol{\tau}) - \left[H(\rho_1 \| \eta_1) - \frac{\kappa}{2} H(\pi_{\rho_1} \| \pi_{\eta_1})\right] \\ &+ \sum_{x,x' \in \mathcal{X}} \alpha(x,x') \left\{\mathbb{E}_{\rho_1}\left[\frac{1}{\kappa} E_1(x,x')(\boldsymbol{\tau})\right] - \pi_{\rho_1}(x,x')\right\} + \sum_{x,x' \in \mathcal{X}} \beta(x,x') \left[\pi_{\rho_1}(x,x') - \pi_{\rho_1}(x',x)\right] \\ &+ \gamma \left[\sum_{x,x' \in \mathcal{X}} \pi_{\rho_1}(x,x') - 1\right] - \sum_{\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \delta(\tau)\rho_1(\tau).\end{aligned}$$

Taking derivative with respect to ρ_1 and π_{ρ_1}. Obtain, for all $\tau \in \mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]$

$$\frac{\partial \mathcal{L}(\rho_1, \pi_{\rho_1}, \alpha, \beta, \gamma, \delta)}{\partial \rho_1(\tau)} = \mathcal{H}(\tau) - \left[\log \frac{\rho_1(\tau)}{\eta_1(\tau)} - 1\right] + \sum_{x,x' \in \mathcal{X}} \alpha(x,x') \frac{1}{\kappa} E_1(x,x')(\tau) - \delta(\tau) = 0, \tag{4.4}$$

and for all $x, x' \in \mathcal{X}$

$$\frac{\partial \mathcal{L}(\rho_1, \pi_{\rho_1}, \alpha, \beta, \gamma, \delta)}{\partial \pi_{\rho_1}(x, x')} = \frac{\kappa}{2}\left[\log \frac{\pi_{\rho_1}(x, x')}{\pi_{\eta_1}(x, x')}\right] - \alpha(x, x') + \beta(x, x') - \beta(x', x) + \gamma = 0. \quad (4.5)$$

We first examine (4.4):

$$\begin{aligned}
\sum_{x,x' \in \mathcal{X}} \alpha(x, x')\frac{1}{\kappa}E_1(x, x')(\tau) &= \sum_{x,x' \in \mathcal{X}} \alpha(x, x')\frac{1}{\kappa} \sum_{v \in \mathcal{N}_o(\tau)} \mathbf{1}\{\tau_v = x, \tau_o = x'\} \\
&= \sum_{v \in \mathcal{N}_o(\tau)} \sum_{x,x' \in \mathcal{X}} \alpha(x, x')\frac{1}{\kappa}\mathbf{1}\{\tau_v = x, \tau_o = x'\} \\
&= \sum_{v \in \mathcal{N}_o(\tau)} \sum_{x,x' \in \mathcal{X}} \alpha(\tau_v, \tau_o)\frac{1}{\kappa}\mathbf{1}\{\tau_v = x, \tau_o = x'\} \\
&= \sum_{v \in \mathcal{N}_o(\tau)} \frac{1}{\kappa}\alpha(\tau_v, \tau_o),
\end{aligned}$$

so

$$\begin{aligned}
\log \frac{\rho_1(\tau)}{\eta_1(\tau)} - 1 &= \mathcal{H}(\tau) + \sum_{v \in \mathcal{N}_o(\tau)} \frac{1}{\kappa}\alpha(\tau_v, \tau_o) - \delta(\tau) \\
\log \frac{\rho_1(\tau)}{\eta_1(\tau)} &= \mathcal{H}(\tau) + \sum_{v \in \mathcal{N}_o(\tau)} \left[\frac{1}{\kappa}\alpha(\tau_v, \tau_o) + \frac{1}{\kappa}\right] - \delta(\tau) \\
\frac{\rho_1(\tau)}{\eta_1(\tau)} &= \exp\left\{\mathcal{H}(\tau) + \sum_{v \in \mathcal{N}_o(\tau)} \left[\frac{1}{\kappa}\alpha(\tau_v, \tau_o) + \frac{1}{\kappa}\right] - \delta(\tau)\right\}
\end{aligned}$$

Define the probability measure h on $\mathcal{X}^2$ with the following pmf:

$$h(x, x') = \frac{1}{\overline{z_h}} \exp\left[\frac{1}{\kappa}\alpha(x, x') + \frac{1}{\kappa}\right] \quad (4.6)$$

so

$$\rho_1(\tau) = \eta_1(\tau)\overline{z_h}^{\kappa} \exp[\mathcal{H}(\tau) - \delta(\tau)] \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, \tau_o)$$

Notice that for all τ, the $\psi(\tau) > 0$, and the exponential is also strictly positive, meaning that $\rho_1(\tau) > 0$, and thus $\delta(\tau) = 0$, so

$$\rho_1(\tau) = \eta_1(\tau)\overline{z_h}^{\kappa} \exp \mathcal{H}(\tau) \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, \tau_o)$$

Notice that since $\eta_1(\tau) = \binom{\kappa}{\mathbf{n}(\tau)}/|\mathcal{X}|^{\kappa+1}$

$$\rho_1(\tau) = \frac{\overline{z_h}^{\kappa}}{|\mathcal{X}|^{\kappa+1}} \binom{\kappa}{\mathbf{n}(\tau)} \exp \mathcal{H}(\tau) \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, \tau_o).$$

And

$$\rho_1(\tau) = \frac{1}{z_{\rho_1}} \binom{\kappa}{\mathbf{n}(\tau)} \exp \mathcal{H}(\tau) \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, \tau_o), \tag{4.7}$$

with

$$z_{\rho_1} = \frac{|\mathcal{X}|^{\kappa+1}}{\overline{z_h}^{\kappa}}. \tag{4.8}$$

Turning our attention to (4.5):

$$\frac{\kappa}{2}\left[\log \frac{\pi_{\rho_1}(x,x')}{\pi_{\eta_1}(x,x')} - 1\right] = \alpha(x,x') - \beta(x,x') + \beta(x',x) - \gamma$$

Since both π_{ρ_1} and π_{η_1} are symmetric,

$$\frac{\kappa}{2}\left[\log \frac{\pi_{\rho_1}(x,x')}{\pi_{\eta_1}(x,x')} - 1\right] = \frac{\kappa}{2}\left[\log \frac{\pi_{\rho_1}(x',x)}{\pi_{\eta_1}(x',x)} - 1\right].$$

Thus

$$\begin{aligned}
\alpha(x,x') - \beta(x,x') + \beta(x',x) - \gamma &= \alpha(x',x) - \beta(x',x) + \beta(x,x') - \gamma \\
-2\beta(x,x') + 2\beta(x',x) &= \alpha(x',x) - \alpha(x,x') \\
-\beta(x,x') + \beta(x',x) &= \frac{1}{2}\alpha(x',x) - \frac{1}{2}\alpha(x,x').
\end{aligned}$$

And

$$\begin{aligned}
\frac{\kappa}{2}\left[\log \frac{\pi_{\rho_1}(x,x')}{\pi_{\eta_1}(x,x')} - 1\right] &= \alpha(x,x') + \frac{1}{2}\alpha(x',x) - \frac{1}{2}\alpha(x,x') - \gamma \\
&= \frac{1}{2}\alpha(x,x') + \frac{1}{2}\alpha(x',x) - \gamma \\
\log \frac{\pi_{\rho_1}(x,x')}{\pi_{\eta_1}(x,x')} &= \frac{1}{\kappa}\alpha(x,x') + \frac{1}{\kappa}\alpha(x',x) - \frac{2\gamma}{\kappa} + 1 \\
\frac{\pi_{\rho_1}(x,x')}{\pi_{\eta_1}(x,x')} &= \exp\left[\frac{1}{\kappa}\alpha(x,x') + \frac{1}{\kappa}\alpha(x',x) - \frac{2\gamma}{\kappa} + 1\right]
\end{aligned}$$

By $\pi_{\eta_1}(x,x') = 1/|\mathcal{X}|^2$,

$$\pi_{\rho_1}(x,x') = \frac{1}{|\mathcal{X}|^2} \exp\left[\frac{1}{\kappa}\alpha(x,x') + \frac{1}{\kappa}\alpha(x',x) - \frac{2\gamma}{\kappa} + 1\right]$$

Similar to ρ_1, we rewrite α in terms of h

$$= \frac{1}{|\mathcal{X}|^2} \exp\left[\left(\frac{1}{\kappa}\alpha(x,x') + \frac{1}{\kappa}\right) + \left(\frac{1}{\kappa}\alpha(x',x) + \frac{1}{\kappa}\right) - \frac{2\gamma}{\kappa} + 1 - \frac{2}{\kappa}\right]$$

By definition of h in (4.6),

$$\begin{aligned} &= \frac{1}{|\mathcal{X}|^2}\overline{z_h}h(x,x')\overline{z_h}h(x',x)\exp\left[-\frac{2\gamma}{\kappa} + 1 - \frac{2}{\kappa}\right] \\ &= \frac{\overline{z_h}^2}{|\mathcal{X}|^2}\exp\left[-\frac{2\gamma}{\kappa} + 1 - \frac{2}{\kappa}\right] h(x,x')h(x',x). \end{aligned}$$

Combine the constants by defining $z_{\pi_{\rho_1}} = |\mathcal{X}|^2/\overline{z_h}^2 \exp\left[1 - \frac{2}{\kappa}(\gamma+1)\right]$

$$\pi_{\rho_1}(x,x') = \frac{1}{z_{\pi_{\rho_1}}} h(x,x')h(x',x), \tag{4.9}$$

Observe that π_{ρ_1} is symmetric. However, we still have not satisfied the other constraints. The normalization constraint on π_{ρ_1} is relatively simple to satisfy, we simply set

$$z_{\pi_{\rho_1}} = \sum_{x,x'\in\mathcal{X}} h(x,x')h(x',x). \tag{4.10}$$

For the remaining constraint,

$$\begin{aligned} &\mathbb{E}_{\rho_1}\left[\frac{1}{\kappa}E_1(x,x')(\boldsymbol{\tau})\right] \\ =&\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \rho_1(\tau)\frac{1}{\kappa}E_1(x,x')(\tau) \\ =&\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \frac{1}{\kappa}E_1(x,x')(\tau)\frac{1}{z_{\rho_1}}\binom{\kappa}{\mathbf{n}(\tau)}\exp\mathcal{H}(\tau)\prod_{v\in\mathcal{N}_o(\tau)} h(\tau_v,\tau_o) \\ =&\frac{1}{z_{\rho_1}}\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]}\frac{1}{\kappa}\sum_{w\in\mathcal{N}_o(\tau)}\mathbf{1}\{\tau_w = x, \tau_o = x'\}\binom{\kappa}{\mathbf{n}(\tau)}\exp\mathcal{H}(\tau)\prod_{v\in\mathcal{N}_o(\tau)} h(\tau_v,\tau_o) \end{aligned}$$

Introduce $\tau' = \tau(w\backslash o)_1$z

$$\begin{aligned} =&\frac{1}{z_{\rho_1}}\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]}\frac{1}{\kappa}\sum_{w\in\mathcal{N}_o(\tau)}\sum_{\tau'\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau(w\backslash o)_1 = \tau'\}\mathbf{1}\{\tau_w = x, \tau_o = x'\}\binom{\kappa}{\mathbf{n}(\tau)} \\ &\exp\mathcal{H}(\tau)\prod_{v\in\mathcal{N}_o(\tau)} h(\tau_v,\tau_o) \end{aligned}$$

Rewrite everything in terms of τ'

$$=\frac{1}{z_{\rho_1}}\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]}\frac{1}{\kappa}\sum_{w\in\mathcal{N}_o(\tau)}\sum_{\tau'\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau = \tau'\oplus x\}\mathbf{1}\{\tau_w = x, \tau'_o = x'\}\binom{\kappa}{\mathbf{n}(\tau'\oplus x)}$$

$$\exp\mathcal{H}(\tau'\oplus x)h(x,x')\prod_{v\in\mathcal{N}_o(\tau')}h(\tau'_v,x')$$

Collect all terms still written in terms of τ

$$=\frac{1}{z_{\rho_1}}\sum_{\tau'\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau=\tau'\oplus x,\tau'_o=x'\}\frac{1}{\kappa}\sum_{w\in\mathcal{N}_o(\tau)}\mathbf{1}\{\tau_w=x\}\exp\mathcal{H}(\tau'\oplus x)h(x,x')\prod_{v\in\mathcal{N}_o(\tau')}h(\tau'_v,x')$$

Notice that $\sum_{w\in\mathcal{N}_o(\tau)}\mathbf{1}\{\tau_w=x\}=n_x(\tau)$

$$=\frac{1}{z_{\rho_1}}\sum_{\tau'\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau=\tau'\oplus x,\tau'_o=x'\}\frac{n_x(\tau'\oplus x)}{\kappa}\binom{\kappa}{\mathbf{n}(\tau'\oplus x)}\exp\mathcal{H}(\tau'\oplus x)h(x,x')\prod_{v\in\mathcal{N}_o(\tau')}h(\tau'_v,x')$$

So now we can remove τ

$$=\frac{1}{z_{\rho_1}}\sum_{\tau'\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau'_o=x'\}\frac{n_x(\tau'\oplus x)}{\kappa}\binom{\kappa}{\mathbf{n}(\tau'\oplus x)}\exp\mathcal{H}(\tau'\oplus x)h(x,x')\prod_{v\in\mathcal{N}_o(\tau')}h(\tau'_v,x')$$

Using (2.6),

$$=\frac{1}{z_{\rho_1}}\sum_{\tau\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau_o=x'\}\binom{\kappa}{\mathbf{n}(\tau)}\exp\mathcal{H}(\tau)h(x,x')\prod_{v\in\mathcal{N}_o(\tau)}h(\tau_v,\tau_o)$$
$$=\frac{1}{z_{\rho_1}}h(x,x')\sum_{\tau\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau_o=x'\}\binom{\kappa}{\mathbf{n}(\tau)}\exp\mathcal{H}(\tau)h(x,x')\prod_{v\in\mathcal{N}_o(\tau)}h(\tau_v,\tau_o)$$
$$=\pi_{\rho_1}(x,x')$$
$$=\frac{1}{z_{\pi_{\rho_1}}}h(x,x')h(x',x).$$

Thus it follows that

$$\frac{1}{z_{\pi_{\rho_1}}}h(x',x)=\frac{1}{z_{\rho_1}}\sum_{\tau\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau_o=x'\}\binom{\kappa}{\mathbf{n}(\tau)}\exp\mathcal{H}(\tau)\prod_{v\in\mathcal{N}_o(\tau)}h(\tau_v,x')$$
$$h(x',x)=\frac{z_{\pi_{\rho_1}}}{z_{\rho_1}}\sum_{\tau\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau_o=x'\}\binom{\kappa}{\mathbf{n}(\tau)}\exp\mathcal{H}(\tau)\prod_{v\in\mathcal{N}_o(\tau)}h(\tau_v,x')$$

Set

$$z_h=\frac{z_{\rho_1}}{z_{\pi_{\rho_1}}},\tag{4.11}$$

we obtain the fixed-point cavity equation

$$h(x,x')=\frac{1}{z_h}\sum_{\tau\in\mathcal{T}^{\kappa-1}_{*,1}[\mathcal{X}]}\mathbf{1}\{\tau_o=x\}\binom{\kappa}{\mathbf{n}(\tau)}\exp\mathcal{H}(\tau)\prod_{v\in\mathcal{N}_o(\tau)}h(\tau_v,x).\tag{4.12}$$

Additionally, since h is a probability measure, we know that [1]

$$z_h = \sum_{x,x' \in \mathcal{X}} \sum_{\tau \in \mathcal{T}_{*,1}^{\kappa-1}[\mathcal{X}]} \mathbf{1}\{\tau_o = x\} \binom{\kappa - 1}{\mathbf{n}(\tau)} \exp \mathcal{H}(\tau) \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, x). \tag{4.13}$$

By our definition, have $h = \Gamma h$.

To summarize, we found that the stationary points of $\mathcal{L}(\rho_1, \pi_{\rho_1}, \alpha, \beta, \gamma)$ to be

$$\begin{aligned} \rho_1(\tau) &= \frac{1}{z_{\rho_1}} \binom{\kappa}{\mathbf{n}(\tau)} \exp \mathcal{H}(\tau) \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, \tau_o) \\ \pi_{\rho_1}(x, x') &= \frac{1}{z_{\pi_{\rho_1}}} h(x, x') h(x', x), \end{aligned}$$

where $h = \Gamma h$. These should also be the stationary points of $\widetilde{V}_{\mathcal{H}}$ on the specified domain.

The function value at these points can also be calculated:

$$\begin{aligned} H(\rho_1 \| \eta_1) &= \sum_{\tau \in \mathcal{T}_{*,1}^{\kappa}[\mathcal{X}]} \rho_1(\tau) \log \frac{\rho_1(\tau)}{\eta_1(\tau)} \\ &= \sum_{\tau \in \mathcal{T}_{*,1}^{\kappa}[\mathcal{X}]} \rho_1(\tau) \log \frac{\frac{1}{z_{\rho_1}} \binom{\kappa}{\mathbf{n}(\tau)} \exp \mathcal{H}(\tau) \prod_{v \in \mathcal{N}_o(\tau)} h(\tau_v, \tau_o)}{\frac{1}{|\mathcal{X}|^{\kappa+1}} \binom{\kappa}{\mathbf{n}(\tau)}} \\ &= -\log z_{\rho_1} + (\kappa + 1) \log |\mathcal{X}| + \mathbb{E}_{\rho_1} \mathcal{H}(\boldsymbol{\tau}) + \mathbb{E}_{\rho_1} \left[\sum_{v \in \mathcal{N}_o(\tau)} \log h(\boldsymbol{\tau}_v, \boldsymbol{\tau}_o) \right] \end{aligned}$$

Apply (4.2) with $g(x, x') = \log h(x, x')$,

$$= -\log z_{\rho_1} + (\kappa + 1) \log |\mathcal{X}| + \mathbb{E}_{\rho_1} \mathcal{H}(\boldsymbol{\tau}) + \kappa \mathbb{E}_{\pi_{\rho_1}} \log h(\mathbf{x}, \mathbf{x}').$$

On the other hand,

$$\begin{aligned} H(\pi_{\rho_1} \| \pi_{\eta_1}) &= \sum_{x,x' \in \mathcal{X}} \pi_{\rho_1}(\tau) \log \frac{\pi_{\rho_1}(\tau)}{\pi_{\eta_1}(\tau)} \\ &= \sum_{x,x' \in \mathcal{X}} \pi_{\rho_1}(\tau) \log \frac{\frac{1}{z_{\pi_{\rho_1}}} h(x, x') h(x', x)}{1/|\mathcal{X}|^2} \\ &= -\log z_{\pi_{\rho_1}} + 2 \log |\mathcal{X}| + \mathbb{E}_{\pi_{\rho_1}} \log h(\mathbf{x}, \mathbf{x}') + \mathbb{E}_{\pi_{\rho_1}} \log h(\mathbf{x}', \mathbf{x}) \end{aligned}$$

By π_{ρ_1} symmetric,

$$= -\log z_{\pi_{\rho_1}} + 2 \log |\mathcal{X}| + 2 \mathbb{E}_{\pi_{\rho_1}} [\log h(\mathbf{x}, \mathbf{x}')].$$

[1] We can technically prove it with (4.11), but I have not found a way to do that yet.

Plugging into $\widetilde{V}_{\mathcal{H}}(\rho_1, \pi_{\rho_1})$,

$$
\begin{aligned}
&\widetilde{V}_{\mathcal{H}}(\rho_1, \pi_{\rho_1})\\
=&\mathbb{E}_{\rho_1}\mathcal{H}(\boldsymbol{\tau}) - \Big[H(\rho_1\|\eta_1) - \frac{\kappa}{2}H(\pi_{\rho_1}\|\pi_{\eta_1})\Big]\\
=&\mathbb{E}_{\rho_1}\mathcal{H}(\boldsymbol{\tau}) + \log z_{\rho_1} - (\kappa+1)\log|\mathcal{X}| - \mathbb{E}_{\rho_1}\mathcal{H}(\boldsymbol{\tau}) - \kappa\mathbb{E}_{\pi_{\rho_1}}\log h(x,x')\\
&-\frac{\kappa}{2}\log z_{\pi_{\rho_1}} + \kappa\log|\mathcal{X}| + \kappa\mathbb{E}_{\pi_{\rho_1}}\log h(x,x')
\end{aligned}
$$

everything cancels, and

$$
= \log z_{\rho_1} - \frac{\kappa}{2}\log z_{\pi_{\rho_1}} - \log|\mathcal{X}|
$$

This is the value of $\widetilde{V}_{\mathcal{H}}$ for each stationary point. Since the maxima cannot be on a boundary (where some values of the pmf is zero), the maxima over all stationary points gives the maximum. Knowing this, we can state the annealed pressure as the following

$$
\begin{aligned}
\Psi(\mathcal{H}) &= \sup_{\substack{h\in\mathcal{P}(\mathcal{X}^2)\\ h=\Gamma h}} \left\{ \log\left[\sum_{\tau\in\mathcal{T}^{\kappa}_{*,1}[\mathcal{X}]} \binom{\kappa}{\mathbf{n}(\tau)} \exp\mathcal{H}(\tau) \prod_{v\in\mathcal{N}_o(\tau)} h(\tau_v,\tau_o)\right] - \frac{\kappa}{2}\log\left[\sum_{x,x'\in\mathcal{X}} h(x,x')h(x',x)\right]\right\}\\
&= \sup_{\substack{h\in\mathcal{P}(\mathcal{X}^2)\\ h=\Gamma h}} \Phi_{\mathrm{BF}}(\mathcal{H}, h)
\end{aligned}
$$

Thus annealed pressure is equal to maximum Bethe free energy density. □

5 Conclusion

In this thesis, we have made significant progress towards proving the Bethe prediction for non-pairwise interactions on random κ-regular graphs, proving that annealed pressure is equal to the RS solution for all temperature regimes. Through a large deviation result, we rewrote annealed pressure as the supremum of a functional over probability measures on marked κ-stars. Solving for the stationary points of this functional with Lagrange multipliers, we were able to recover both the cavity equation and the Bethe free energy functional, leading to the above conclusion.

It is not immediately obvious whether the cavity equation recovered is equivalent to the more general belief propagation that Mézard and Montanari [4] proposed, or in the case that they are not equivalent, whether they give the same prediction as the number of particles go to infinity. However, for non-pairwise interactions, there can be great intuition offered in this cavity equation, that is not apparent in the corresponding belief propagation formulation.

This work offers insight into a more general form of calculating annealed pressure problems in spin-glass models. In contrast to the work done by Dembo et al. [3], the large deviation result is more general than the combinatorial argument to obtain the annealed pressure functional, easily applied to non-pairwise interactions like mentioned, and even continuous-state spin-glass systems. Using a similar method such as this one, we can expand the equivalence between the RS solution and annealed pressure to many other models, including those with longer-range interactions, such as a three-neighbor equivalent of the NNNI model, or models with continuous spin-states, such as the XY model or the Heisenberg model.

Another important piece of future work involves the next step outlined by Mézard and Montanari: proving that annealed pressure is equal to quenched pressure. We wish to show this for the NNNI model, and expand the current knowledge about Bethe prediction to interactions in the first-neighborhood.

References

[1] I-Hsun Chen, Kavita Ramanan, and Sarath Yasodharan. *Variational formulae for the annealed pressure for statistical physics models with discrete and continuous spins*. 2024. Forthcoming.

[2] Amir Dembo and Andrea Montanari. *Ising models on locally tree-like graphs*. Sept. 2010. URL: https://arxiv.org/abs/0804.4726.

[3] Amir Dembo et al. "The replica symmetric solution for Potts models on D-regular graphs." In: *Communications in Mathematical Physics* 327.2 (Mar. 2014), pp. 551–575. DOI: 10.1007/s00220-014-1956-6.

[4] Marc Mezard and Andrea Montanari. *Information, physics and computation*. Oxford University Press, 2017.

[5] Kavita Ramanan and Sarath Yasodharan. *On the large deviation rate function for marked sparse random graphs*. 2023. arXiv: 2312.16103 [math.PR].

[6] Jack Raymond and K Y Michael Wong. "Next nearest neighbour Ising models on random graphs." In: *Journal of Statistical Mechanics: Theory and Experiment* 2012.09 (Sept. 2012). DOI: 10.1088/1742-5468/2012/09/p09007.