《一阶复本对称破缺理论》

PART V Chap. 19

分享人：孟凡辉

中山大学 COIN Lab

2023 年 03 月 01 日


國立中山大學

Background

一阶复本对称破缺 (one-step replica symmetry breaking, 1RSB) 理论的由来: 解决 "负熵" 危机.

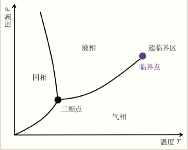


图1 水的相图

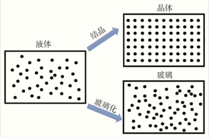


图2 结晶与玻璃化

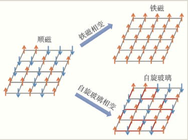


图3 铁磁相变和自旋玻璃相变。灰色的边代表铁磁相互作用 $J_{ij}=1$，红色的边代表反铁磁相互作用 $J_{ij}=-1$

$$[\log Z] = \lim_{n \to 0} \frac{[Z^n] - 1}{n}$$

$$q_{\alpha\beta} = \frac{1}{N} \sum_{i=1}^{N} \left[m_i^{\alpha} m_i^{\beta} \right]$$

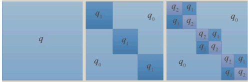


图7 复本交叠矩阵。从左到右分别为复本对称、一阶复本对称破缺、二阶复本对称破缺

表1 不同阶复本对称破缺理论得到的零温熵 $S(0)$[7, 8, 10]

K	$S(0)$
0	−0.16
1	−0.01
2	−0.004
∞	0

Background

消息传递 (message passing) 算法的局限性.

The effectiveness of belief propagation depends on one basic assumption: when a function node is pruned from the factor graph, the adjacent variables become weakly correlated with respect to the resulting distribution. This hypothesis may break down either because of the existence of small loops in the factor graph or because variables are correlated at large distances. In factor graphs with a locally tree-like structure, the second scenario is responsible for the failure of BP. The emergence of such long-range correlations is a signature of a phase transition separating a 'weakly correlated' and a 'highly correlated' phase. The latter is often characterized by the decomposition of the (Boltzmann) probability distribution into well-separated 'lumps' (pure Gibbs states).

Despite these complications, physicists have developed a non-rigorous approach to deal with this phenomenon: the 'one-step replica symmetry breaking' (1RSB) cavity method. This method postulates a few properties of the pure-state decomposition, and, on this basis, allows one to derive a number of quantitative predictions ('conjectures' from a mathematical point of view). Examples include the satisfiability threshold for random K-SAT and for other random constraint satisfaction problems.

The method is rich enough to allow some self-consistency checks of such assumptions. In several cases in which the 1RSB cavity method has passed this test, its predictions have been confirmed by rigorous arguments (and there is no case in which they have been falsified so far). These successes encourage the quest for a mathematical theory of Gibbs states on sparse random graphs.

Introduce a Boltzmann distribution over Bethe measures, write it in the form of a graphical model, and use BP to study this model.


國立中山大學

BP equations

$$\nu_{j\to a}^{(t+1)}(x_j) \cong \prod_{b\in\partial j\setminus a} \widehat{\nu}_{b\to j}^{(t)}(x_j)\,, \tag{14.14}$$

$$\widehat{\nu}_{a\to j}^{(t)}(x_j) \cong \sum_{\underline{x}_{\partial a\setminus j}} \psi_a(\underline{x}_{\partial a}) \prod_{k\in\partial a\setminus j} \nu_{k\to a}^{(t)}(x_k)\,. \tag{14.15}$$

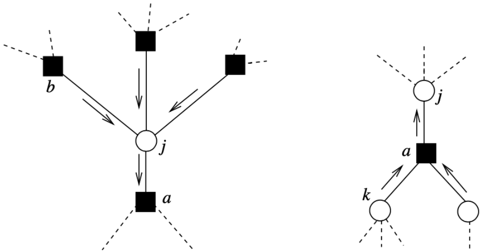


Fig. 14.3 *Left*: the portion of the factor graph involved in the computation of $\nu_{j\to a}^{(t+1)}(x_j)$. This message is a function of the 'incoming messages' $\widehat{\nu}_{b\to j}^{(t)}(x_j)$, with $b \neq a$. *Right*: the portion of the factor graph involved in the computation of $\widehat{\nu}_{a\to j}^{(t)}(x_j)$. This message is a function of the 'incoming messages' $\nu_{k\to a}^{(t)}(x_k)$, with $k \neq j$.

Bethe measures

As in Chapter 14, we consider a factor graph $G = (V, F, E)$, with variable nodes $V = \{1, \ldots, N\}$, factor nodes $F = \{1, \ldots, M\}$, and edges E. The joint probability distribution over the variables $\underline{x} = (x_1, \ldots, x_N) \in \mathcal{X}^N$ takes the form

$$\mu(\underline{x}) = \frac{1}{Z} \prod_{a=1}^{M} \psi_a(\underline{x}_{\partial a}) \,. \tag{19.1}$$

Given a subset of variable nodes $U \subseteq V$ (which we shall call a 'cavity'), the **induced subgraph** $G_U = (U, F_U, E_U)$ is defined as the factor graph that includes all of the factor nodes a such that $\partial a \subseteq U$, and the adjacent edges. We also write $(i, a) \in \partial U$ if $i \in U$ and $a \in F \setminus F_U$. Finally, a **set of messages** $\{\widehat{\nu}_{a \to i}\}$ is a set of probability distributions over $\mathcal{X}$, indexed by directed edges $a \to i$ in E with $a \in F$, $i \in V$.

Definition 19.1. (Informal) *The probability distribution μ is a **Bethe measure** (or **Bethe state**) if there exists a set of messages $\{\widehat{\nu}_{a \to i}\}$ such that, for 'almost all' of the 'finite-size' cavities U, the distribution $\mu_U(\,\cdot\,)$ of the variables in U can be approximated as*

$$\mu_U(\underline{x}_U) \cong \prod_{a \in F_U} \psi_a(\underline{x}_{\partial a}) \prod_{(ia) \in \partial U} \widehat{\nu}_{a \to i}(x_i) \;+\; \mathsf{err}(\underline{x}_U) \,, \tag{19.2}$$

where $\mathsf{err}(\underline{x}_U)$ is a 'small' error term, and $\cong$ denotes, as usual, equality up to a normalization.

Example 1

Exercise 19.1 Assume that $G = (V, F, E)$ has girth larger than 2, and that $\mu(\cdot)$ is a Bethe measure with respect to the message set $\{\widehat{\nu}_{a\to i}\}$, where $\widehat{\nu}_{a\to i}(x_i) > 0$ for any $(i, a) \in E$, and $\psi_a(\underline{x}_{\partial a}) > 0$ for any $a \in F$. For $U \subseteq V$ and $(i, a) \in \partial U$, define a new subset of variable nodes as $W = U \cup \partial a$ (see Fig. 19.1).

Applying eqn (19.2) to the subsets of variables U and W, show that a message must satisfy (up to an error term of the same order as $\mathsf{err}(\cdot)$)

$$\widehat{\nu}_{a\to i}(x_i) \cong \sum_{\underline{x}_{\partial a\setminus i}} \psi_a(\underline{x}_{\partial a}) \prod_{j\in\partial a\setminus i} \Big\{ \prod_{b\in\partial j\setminus a} \widehat{\nu}_{b\to j}(x_j) \Big\}. \tag{19.3}$$

Show that these equations are equivalent to the BP equations (14.14) and (14.15). [Hint: Define, for $k \in V$, $c \in F$ and $(k, c) \in E$, $\nu_{k\to c}(x_k) \cong \prod_{d\in\partial k\setminus c} \widehat{\nu}_{d\to k}(x_k)$.]

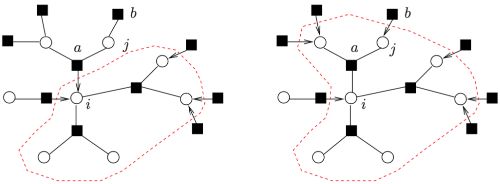


Fig. 19.1 Two examples of cavities. The *right*-hand one is obtained by adding the extra function node a. The consistency of the Bethe measure in these two cavities implies the BP equation for $\widehat{\nu}_{a\to i}$; see Exercise 19.1.

Example 2

It would be pleasant if the converse was true, i.e. if each quasi-solution of the BP equations corresponded to a distinct Bethe measure. In fact, such a relation will be at the heart of the assumptions of the 1RSB method. However, one should keep in mind that this is not always true, as the following example shows.

Example 19.2 Let G be a factor graph with the same degree $K \geq 3$ at both factor and variable nodes. Consider binary variables ($\mathcal{X} = \{0, 1\}$) and, for each $a \in F$, let

$$\psi_a(x_{i_1(a)}, \dots, x_{i_K(a)}) = \mathbb{I}(x_{i_1(a)} \oplus \cdots \oplus x_{i_K(a)} = 0)\,. \tag{19.4}$$

Given a perfect matching $\mathsf{M} \subseteq E$, a solution of the BP equations can be constructed as follows. If $(i, a) \in \mathsf{M}$, then let $\widehat{\nu}_{a\to i}(x_i) = \mathbb{I}(x_i = 0)$ and $\nu_{i\to a}(0) = \nu_{i\to a}(1) = 1/2$. If, on the other hand, $(i, a) \notin \mathsf{M}$, then let $\widehat{\nu}_{a\to i}(0) = \widehat{\nu}_{a\to i}(1) = 1/2$ and $\nu_{i\to a}(0) = \mathbb{I}(x_i = 0)$.

Check that this is a solution of the BP equations and that all the resulting local marginals coincide with the ones of the measure $\mu(\underline{x}) \cong \mathbb{I}(\underline{x} = \underline{0})$, independently of M. If we take, for instance, G to be a random regular graph with degree $K \geq 3$, both at factor nodes and at variable nodes, then the number of perfect matchings of G is, with high probability, exponential in the number of nodes. Therefore we have constructed an exponential number of solutions of the BP equations that describe the same Bethe measure.

Generic scenarios

RS (replica-symmetric). This is the simplest possible scenario: the distribution $\mu(\,\cdot\,)$ is a Bethe measure. A slightly more complicated situation (which we still place in the 'replica-symmetric' family) arises when $\mu(\,\cdot\,)$ decomposes into a finite set of Bethe measures related by 'global symmetries', as in the Ising ferromagnet discussed in Section 17.3.

d1RSB (dynamic one-step replica symmetry breaking). There exist an exponentially large (in the system size N) number of Bethe measures. The measure μ decomposes into a convex combination of these Bethe measures, i.e.

$$\mu(\underline{x}) = \sum_n w_n \, \mu^n(\underline{x}) \,, \tag{19.5}$$

with weights w_n exponentially small in N. Furthermore $\mu(\,\cdot\,)$ is itself a Bethe measure.

s1RSB (static one-step replica symmetry breaking). As in the **d1RSB** case, there exist an exponential number of Bethe measures, and μ decomposes into a convex combination of such states. However, a finite number of the weights w_n are of order 1 as $N \to \infty$, and (unlike in the previous case) μ is not itself a Bethe measure.

Three assumptions

Assumption 1 *There exist exponentially many quasi-solutions of the BP equations. The number of such solutions with free entropy $\mathbb{F}(\underline{\nu}^n) \approx N\phi$ is (to leading exponential order) $\exp\{N\Sigma(\phi)\}$, where $\Sigma(\cdot)$ is the* ***complexity*** *function.*[1]

Assumption 2 *The 'canonical' measure μ, defined as in eqn (19.1), can be written as a convex combination of extremal Bethe measures*

$$\mu(\underline{x}) = \sum_{n \in \mathsf{E}} w_n \, \mu^n(\underline{x}) \,, \tag{19.6}$$

with weights related to the Bethe free entropies $w_n = \mathrm{e}^{\mathbb{F}_n}/\Xi$, $\Xi \equiv \sum_{n \in \mathsf{E}} \mathrm{e}^{\mathbb{F}_n}$.

Assumption 3 *To leading exponential order, the number of extremal Bethe measures equals the number of quasi-solutions of the BP equations: the number of extremal Bethe measures with free entropy $\approx N\phi$ is also given by $\exp\{N\Sigma(\phi)\}$.*

Generalized partition function

Within the three assumptions described above, the complexity function $\Sigma(\phi)$ provides basic information on how the measure μ decomposes into Bethe measures. Since the number of extremal Bethe measures with a given free-entropy density is exponential in the system size, it is natural to treat them within a statistical-physics formalism. The BP messages of the original problem will be the new variables, and Bethe measures will be the new configurations. This is what 1RSB is about.

We introduce the auxiliary statistical-physics problem through the definition of a canonical distribution over extremal Bethe measures: we assign to a measure $n \in \mathsf{E}$ the probability $w_n(\mathtt{x}) = e^{\mathtt{x}\mathbb{F}_n}/\Xi(\mathtt{x})$. Here $\mathtt{x}$ plays the role of an inverse temperature (and is often called the **Parisi 1RSB parameter**). [2] The partition function of this generalized problem is

$$\Xi(\mathtt{x}) = \sum_{n \in \mathsf{E}} e^{\mathtt{x}\mathbb{F}_n} \doteq \int e^{N[\mathtt{x}\phi + \Sigma(\phi)]}\, d\phi\,. \qquad (19.7)$$

Free energy landscape

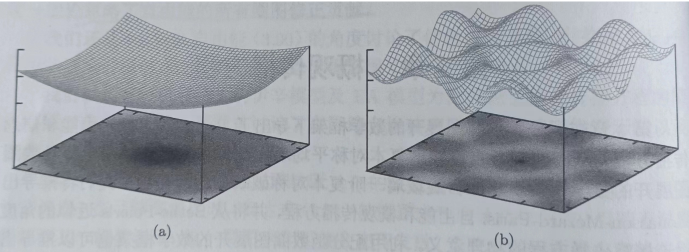
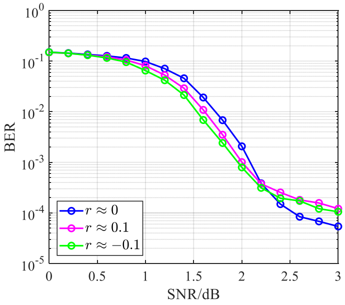
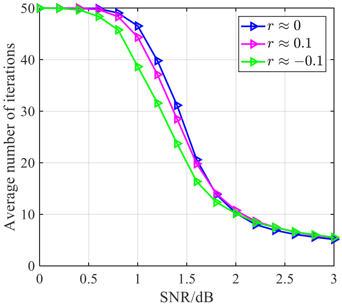


图 4.1 自旋玻璃自由能图景随温度变化可能发生定性改变

在高温时（a），系统只有一个宏观态，自由能图景作为一组恰当选取的序参量的函数只有一个极小值；在低温时（b），系统有许多宏观态，自由能图景作为序参量的函数有许多极小点和鞍点。该图引自文献 [123]


BER
SNR/dB
$r \approx 0$
$r \approx 0.1$
$r \approx -0.1$


Average number of iterations
SNR/dB
$r \approx 0$
$r \approx 0.1$
$r \approx -0.1$

Thanks!