Generalized Criteria for Uniqueness of Gibbs Measures

Hongbo Zhou[1,2] *and Zhiming Zheng*[1]
1:Key Laboratory of Mathematics, Informatics and Behavioral Semantics, Ministry of Education, China.
2:Computer Science Department, Southern Illinois University Carbondale.
zhouhongbo@ss.buaa.edu.cn, zzheng@pku.edu.cn

Abstract

We study the uniqueness of Gibbs measures constructed on infinite graphs, in which the vertexes admit certain markov properties related to the connectivity properties of the graphs. Inspired by Weitz and Winkler's work, we generalize Dobrushin's influence of site to site via sites to a more general form of block to block via blocks by path coupling technique. Our condition for uniqueness of Gibbs measure, is a mathematical statement of the spatial correlation decay conjecture.

1 Introduction

We are getting more and more involved in dealing with super-scale systems, such as the ever-growing internet, the soma and synapse networks in human brain, as well as large protein interaction networks. More often, we can only get partial information within small parts of the whole systems. Comparing with these very limited known-parts, the super-scale systems can be viewed as infinite systems. A fundamental question is how to build well-defined probability distributions for certain parameters in these infinite systems, and this will be crucial for using many model-based pattern recognition methods. By well-define, we mean the definition of parameters within these finite known-parts should be consistent with the system as a whole. And, Gibbs Measure is a mathematical exact framework to describe such consistencies.

As soon as we have defined proper parameters for an infinite system, we may want to know the dynamics (regarding the parameters) within our interesting parts. This can be stated by the uniqueness of Gibbs measures. Additionally, there are at least two interesting issues involving the uniqueness of Gibbs measures. The first aspect is the hunting for the phase transition boundaries, varying from the classic Ising model in statistical physics to the modern research of Constrained Satisfiability Problems built on various graphs. The second application involves the calculation of partition functions in a myriad of statistical inference systems. Within the correlation decay boundary, the marginal probabilities can be computed efficiently by message-passing algorithms, such as cavity method or sum-product algorithms [Tatikonda S. and Jordan M., 2002]. Further, we can readily recover the partition functions by these marginal probabilities, and hence almost every interesting parameter such as free-energy and entropy can be determined. Again, the correlation decay boundary, which is an alternative interpretation for borderline of uniqueness of Gibbs measures, plays a key role in determining the feasibility and efficiency of the marginal probabilities calculation.

The aim of our work is to supply such a mathematical exact reference for this spatial correlation decay conjecture, which states that a system of weakly interacted in each (spatial) part will have a better chance to maintain only one macroscopic state. As related work, Dobrushin developed criteria of uniqueness involving a measurement of influence of *site to site via sites* in his seminal papers [Dobrushin R.L., 1968; Dobrushin R.L., 1970]. Weitz D. (2005) generalized Dobrushin's condition by introducing intermediate block update rules, and the measurement of influence can be interpreted as *site to site via blocks*. Here we propose a more refined measurement of influence as *block to block via blocks* based on Winkler S. and Tatikonda S. (2007).

2 Preliminaries, Specification and Gibbs Measure

2.1 Preliminaries

Let $G = (V, \mathbb{E})$ be a finite graph with vertex set V and edge set $\mathbb{E}$. Let $\mathbb{S} = \{S \in V, S \neq \Phi\}$ be the set of all nonempty finite subset of V. Associated with each vertex $i \in V$ there is a random variable X_i taking values in a finite set $\mathcal{X}_i$. Then $X = \{X_i, i \in V\}$ will

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

be the set of all random variables. The product space $\mathcal{X} := \prod_{i\in V} \mathcal{X}_i$ equipped with product σ-field $\mathcal{F}$, for which $\mathbb{P}\{X \in E\} = \mu E$ for all events $E \in \mathcal{F}$. $x, y \in \mathcal{X}$ are called configurations. For arbitrary $S \in \mathbb{S}$, let $\mathcal{X}_S = \prod_{i\in S} \mathcal{X}_i$ be the projection of $\mathcal{X}$ onto S. For any $S \in \mathbb{S}$, $S^c := V \setminus S$ is the set of vertexes outside of S, and $N(S) := \{i \in S^c | \exists j \in S, s.t. \{i,j\} \in \mathbb{E}\}$ is called the boundary of S.

A *clique* C is defined as $C := \{S \in \mathbb{S} | \exists \{i,j\} \in \mathbb{E}, \forall i,j \in S\}$. We denote the collection of all cliques in G as $\mathbb{C}$. For each clique $C \in \mathbb{C}$, we can associate a non-negative function $\Psi_c : \mathcal{X} \times \ldots \times \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ on $\mathcal{X}_c$. In most case, we call Ψ_c as *potential* which is derived from statistical physics. Again, the set of all potentials is denoted by Ψ.

By Hammersley Clifford theorem [Hammersley J. M. and Clifford P., 1971], we can define Gibbs measure with respect to any finite graph G equipped with cliques $\mathbb{C}$, or equivalently, any finite factor graph F equipped with factors $\mathbb{F}$,

$$\mu(x) = \frac{1}{Z} \prod_{a\in\mathbb{F}} \Psi_a(x), \qquad (1)$$

where *partition function* $Z = \sum_{x\in\mathcal{X}} \prod_{a\in\mathbb{F}} \Psi_a(x)$ is a proper normalizing factor. The potential set Ψ is called permissible, if and only if there exists at least one $x \in \mathcal{X}$ which guarantees $Z \neq 0$. We assume all potentials in this paper are permissible.

2.2 Specification and Infinite Gibbs Measure

Here we consider infinite Gibbs measure built on factor graph F, where V, $\mathcal{X}$ and $\mathbb{F}$ are all countable infinite set. Thus, it is always the case that $Z = \sum_{x\in\mathcal{X}} \prod_{a\in\mathbb{F}} \Psi_a(x)$ will be *zero* or *infinite*. To avoid such embarrassment, Dobrushin introduced a family of finite dimensional conditional distributions for each finite subset $\Lambda \in V$,

$$p^x_\Lambda(x_\Lambda | x_{N(\Lambda)}) = \frac{1}{Z_\Lambda(x_{N(\Lambda)})} \prod_{a\in\partial\Lambda} \Psi_a(x), \qquad (2)$$

where $\partial\Lambda$ is the set of factor nodes incident on variable node $i \in \Lambda$, and $Z_\Lambda(x_{N(\Lambda)}) := \sum_{x_\Lambda} \prod_{a\in\partial\Lambda} \Psi_a(x)$. For all finite dimensional region $\Lambda \in V$, conditional distribution $\mathcal{P} = \{p^x_\Lambda(x_\Lambda | x_{N(\Lambda)})\}$ is called a specification.

Definition 1 [Weitz D., 2005]: A probability measure μ over the subset of feasible configurations is called a Gibbs measure for the specification p if, for every finite region Λ and μ-almost every configuration ω,

$$\mu(\cdot | \omega_{\Lambda^c}) = p^\omega_\Lambda. \qquad (3)$$

The physical interpretation is that macroscopic equilibrium can be viewed as certain stationarity of p^ω_Λ over their boundaries.

The collection of Gibbs measures for the specification $\mathcal{P}$ is denoted by $\mathcal{G}(\mathcal{P})$. According to a compactness argument, there exists at least one Gibbs measure for every specification[Georgii H.O., 1988]. Thus, the core issue is to hunt for the boundaries of uniqueness of Gibbs measures.

2.3 Measure of Influence

To compare the distance between two distributions, we need to define some metrics on $\mathcal{X} = \prod_{i\in V} \mathcal{X}_i$, where V is a countable index set and each corordinate space $\mathcal{X}_i$ is finite. Given ρ_i on each $\mathcal{X}_i$, $x_i \in \mathcal{X}_i$ and countable subset $S \in V$, we equip $\mathcal{X}_S$ with the metrics,

$$\rho_S(x_S, y_S) := \sum_{i\in S} \rho_i(x_i, y_i). \qquad (4)$$

Along with Weitz D.'s definition[Weitz D., 2005], the metrics $\{\rho_i\}$ is summable if and only if $\sup_{x,y} \rho(x,y) < \infty$, and bounded if and only if $\sup_{i\in V} \max_{x,y} \rho_i(x_i, y_i) < \infty$. If V is finite, then $\{\rho_i\}$ will be summable and bounded. More specifically, if $\rho_i(x_i, y_i) := |\{x_i \neq y_i\}|$, then $\rho(x,y)$ will be the Hamming distance between x and y.

Similarly, for each $S \subset V$ we can define

$$\overline{\rho}_S(x_S, y_S) := \max_{i\in S} \rho_i(x_i, y_i). \qquad (5)$$

The summable and bounded of metrics $\overline{\rho}$ is the same as that defined for ρ.

Another metrics usually adopted is the total variation distance, which is defined for probability distributions μ and ν as follows:

$$||\mu(f) - \nu(f)||_{TV} := \frac{1}{2} \sum_{x\in V} |\mu(x) - \nu(x)|. \qquad (6)$$

For each $S \subset V$, let $\mathcal{Q}_S$ denote the space of probability measures on $\mathcal{X}_S$, equipped with product field $\mathcal{F}_S$. A coupling of two probability distribution μ_S, ν_S is a probability measure on the product space $\mathcal{X}_S \times \mathcal{X}_S$, equipped with product σ-field $\mathcal{F}_s \otimes \mathcal{F}_s$. we denote the set of all couplings of μ_S and ν_S by $\mathcal{K}(\mu_S, \nu_S)$.

To facilitate the tools of coupling, we define Wasserstein distance between two probability distributions μ and ν w.r.t. a given metrics d,

$$W_d(\mu, \nu) := \inf_{k\in\mathcal{K}(\mu,\nu)} d(k) \qquad (7)$$

where the minimum is obtained by some optimal couplings k.

2.4 Uniqueness Criteria

From now on, we call each finite non-empty subset $B \subset V$ as a block, and Let $\mathbb{B}$ be a finite cover for V. Apparently, each node(or site) $i \in V$ is contained in at least one and at most finitely many blocks.

Dobrushin R.L.(1968) defines the influence of *site j on site i via sites* w.r.t. the total variation distance(Wasserstein distance is adopted in his later work [Dobrushin R.L., 1970]) by,

$$C_{ij} := \sup_{x_j \neq y_j, x_{V\setminus j} = y_{V \setminus j}} ||p_i(\cdot|x) - p_i(\cdot|y)||_{TV}. \quad (8)$$

Proposition 1 [Dobrushin R.L., 1968]: if $\sup_{i \in V} \sum_{j \in N(i)} C_{ij} < 1$, the Gibbs measure for the specification $\mathcal{P}$ is unique.

Further, Weitz D.(2005) introduces the influence of *site j on site i via blocks* as following,

$$I_{i \leftarrow j} := \sup_{x_j \neq y_j, x_{V\setminus j} = y_{V \setminus j}} \sum_{B \in \mathbb{B} \bigcap N(j)} \lambda_B \frac{p_B(\rho_i|x,y)}{\rho_j(x_j, y_j)}, \quad (9)$$

where $p_B(\cdot|x,y)$ is a coupling of $p_B(\cdot|x)$ and $p_B(\cdot|y)$ and λ_B is the mass put on B w.r.t. measure λ. Thus, the influence on site i via blocks is,

$$I_{i \leftarrow} := \sum_{j \in V} I_{i \leftarrow j}. \quad (10)$$

Again, the influence of site j via blocks can be defined as,

$$I_{\leftarrow j} := \sup_{x_j \neq y_j, x_{V\setminus j} = y_{V \setminus j}} \sum_{B \in N(j)} \lambda_B \frac{p_B(\rho_B|x,y)}{\rho_j(x_j, y_j)}. \quad (11)$$

Weitz D.(2005) proves the uniqueness criteria for infinite Gibbs measures as follows,

Proposition 2 [Weitz D., 2005]: If $\{\rho_i\}$ is bounded and $\inf_p \sup_{i \in V} \frac{1}{\lambda_i} I_{i\leftarrow} < 1$, then the Gibbs measure is unique. if $\inf_{i\in V} \lambda_i > 0$,$\sup_{i \in V} \lambda_i < \infty$, $\{\rho_i\}$ is summable and $\inf_p \sup_{j \in V} \frac{1}{\lambda_j} I_{\leftarrow j} < 1$, then the Gibbs measure is unique.

A most recently progress is due to Winkler S. and Tatikonda S. (2007), who propose a notion of influence on block B impacted by the change of single site j w.r.t. metrics ρ:

$$C_{Bj}(\rho) := \sup_{x_j \neq y_j, x_{V\setminus j} = y_{V \setminus j}} \frac{W_\rho(p_B(\cdot|x), p_B(\cdot|y))}{\rho_j(x_j, y_j)}. \quad (12)$$

Proposition 3 [Winkler S. and Tatikonda S., 2007]: If $\{\rho_i\}$ is bounded and $\sup_{B \in \mathbb{B}} \sum_{j \in N(B)} C_{Bj}(\overline{\rho}) < 1$, then the Gibbs measure is unique. if $\sup_{i \in V} \lambda_i < \infty$, $\{\rho_i\}$ is summable and $\sup_{j \in V} \frac{1}{\lambda_j} \sum_{B \in N(j)} \lambda_B C_{Bj}(\rho) < 1$, then the Gibbs measure is unique.

3 Generalized Uniqueness Criteria

3.1 Influence of Block to Block via Blocks

To define the influence of block to block, we first prove that we only need to consider the boundary blocks. Note that the Gibbs measure μ is invariant for every block $B \in \mathbb{B}$ under the prescribed specification $\mathcal{P} = \{p_B(\cdot|x)\}$, which depends on x only through the neighbor set $N(B)$. More mathematically, this can be interpreted by the markov properties of μ as follows,

$$\begin{aligned} \mu(x_B|x_{V\setminus B}) &= \frac{\mu(x_{V\setminus N(B)}|x_{N(B)})}{\mu(x_{V\setminus (B \cup N(B))}|x_{N(B)})} \\ &= \frac{\mu(x_{V\setminus (B \cup N(B))}|x_{N(B)})\mu(x_B|x_{N(B)})}{\mu(x_{V\setminus (B \cup N(B))}|x_{N(B)})} \\ &= \mu(x_B|x_{N(B)}). \end{aligned} \quad (13)$$

Here we give out a larger set $\mathbb{N}(B)$, $N(B) \subset \mathbb{N}(B) \in \mathbb{B}$. More specifically, $\mathbb{N}(B) = N(N(B))$ will be enough. So, it is clear that only the boundary block $B^{'} \subset \mathbb{N}(B)$ is involved instead of all blocks. Intuitively, we will show that the influence of *block to block* is a more refined criterion than the previous known *site to block* and *site to site* w.r.t. metrics d as follows:

$$\begin{aligned} &\underbrace{\sup_{x_{N(B)}, y_{N(B)}} ||p_{B,B'}(\cdot|x_{B'}) - p_{B,B'}(\cdot|y_{B'})||_d}_{\texttt{our notion of block to block}} \\ &\le \underbrace{\sup_{x_i, y_i} ||p_{B,i}(\cdot|x_i) - p_{B,i}(\cdot|y_i)||_d}_{\texttt{site to block by Winkler S.}} \\ &\le \underbrace{\sup_{x_i, y_i} ||p_{j,i}(\cdot|x_i) - p_{j,i}(\cdot|y_i)||_d,}_{\texttt{site to site by Dobushin R.L.}} \end{aligned} \quad (14)$$

where $p_{B,B'}$ is the $\mathcal{X}_{B'}$ marginal probabilities on the block $B^{'} \subset \mathbb{N}(B)$ determined by the specification P, $i \in B^{'}$ and $j \in B$.

In (14), if we take $B^{'} = \{i | i \in N(B)\}$ as all singletons, then we get conditions as Winkler S. and Tatikonda S. (2007); Also, if we take $B^{'} = \{i|i \in N(B)\}$ as all singletons and $B = \{i|i \in B\}$ as all singletons as well, then we have Dobrushin R.L. (1968) uniqueness conditions.

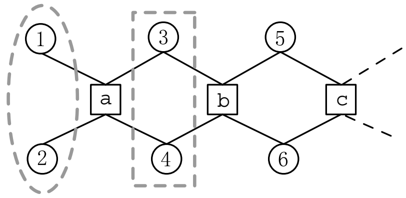

Figure 1. block maginal probability

Example: In Figure 1, We take $B = \{1,2\}$,$B^{'} = \{3,4\}$ and $N(B) = \{3,4\}$. So, we have

$$\begin{aligned} &\sup_{x_2,x_3,x_4} ||p_{B,B'}(x_1,x_2|x_3 \cap x_4)||_d \\ &\quad \le \sup_{x_2,x_3} ||p_{B,\{x_3\}}(x_1,x_2|x_3)||_d \\ &= \sup_{x_2,x_3} ||\sum_{x_2} P_{1,2}(x_1|x_2)P_{2,3}(x_2|x_3)||_d \\ &\quad \le \sup_{x_2} ||P_{1,2}(x_1|x_2)||_d = C_{12}, \qquad (15) \end{aligned}$$

where C_{12} denotes the Dobrushin R. L. (1968)'s influence.

Here we need to define the disagreement of two configurations $x_{B'}$ and $y_{B'}$ on a block $B^{'}$ by $x_{-B'} = y_{-B'}$, in other words, $x_i = y_i, i \in V \setminus B^{'}$ and $x_{B'} \neq y_{B'}$. We introduce the influence of block to block w.r.t. a metrics ρ by

$$C_{B \leftarrow B'}(\rho) := \sup_{x_{-B'}=y_{-B'}} \frac{W_\rho(p_B(\cdot|x), p_B(\cdot|y))}{\rho_{B'}(x_{B'}, y_{B'})}, \qquad (16)$$

where the $W_\rho(\mu, \nu)$ is the Wasserstein distance which can be attained by some optimal couplings(7).

Thus, the influence on block B via blocks is,

$$I_{B \leftarrow B'} := \sup_{B \in \mathbb{B}} \sum_{B' \subset \mathbb{N}(B)} C_{B \leftarrow B'}(\overline{\rho}), \qquad (17)$$

Where $\overline{\rho}$ is defined by(5). Again, the influence of block B via blocks can be defined as,

$$I_{\leftarrow B} := \sup_{B \in \mathbb{B}} \frac{1}{\lambda_B} \sum_{B' \subset \mathbb{N}(B)} C_{B \leftarrow B'}(\rho). \qquad (18)$$

Based on our definitions above, we give out the following two criteria regarding the uniqueness of Gibbs measure,

Proposition 4 : If $\{\rho_i\}$ is bounded and $I_{B \leftarrow B'} < 1$, then the Gibbs measure is unique.

Proposition 5: if $\sup_{B \in V} \lambda_B < \infty$, $\{\rho_i\}$ is summable and $I_{\leftarrow B} < 1$, then the Gibbs measure is unique.

The proofs for Proposition 4 and Proposition 5 can be found at: http://www.cs.siu.edu/∼hbzhou/icpr08/.

4 Conclusions

In this paper, we introduce generalized criteria regarding the uniqueness of infinite Gibbs measures by presenting the influence of *block to block via blocks*. Our criteria completely encompass Dobrushin R.L. (1968), Weitz D. (2005) and Winkler S. and Tatikonda S. (2007) conditions, and we give out more refined results. All notions of influence in their work can be recovered by simplifying our generalized notion of influence.

References

[1] Bubley R. and Dyer M.E., 1997. Path coupling: A technique for proving rapid mixing in markov chains, in proceedings of the 38th annual symposium on foundations of computer science, vol.38. IEEE Computer Society,pp223-233.

[2] Dobrushin R.L., 1968. The problem of uniqueness of a Gibbsian random field and the problem of phase transitions. Functional Analysis and its Applications, 2:302-312.

[3] Dobrushin R.L., 1970. Prescribing a system of random variables by conditional distributions. Theory of Probability and its Applications, 15:458-486.

[4] Georgii H.O., 1988. Gibbs measures and phase transitions, de Gruyter, Berlin.

[5] Hammersley J. M. and Clifford P., 1971: Markov fields on finite graphs and lattices. Unpublished.

[6] Weitz D., 2005. Combinatorial criteria for uniqueness of Gibbs measures. Random Structures Algorithms, 27(4):445-475.

[7] Winkler S. and Tatikonda S., 2006. Criteria for Rapid Mixing of Gibbs Samplers and Uniqueness of Gibbs Measures, 44th Allerton Conference on Communication, Control, and Computing, September 2006 (Monticello, IL).

[8] Tatikonda S. and Jordan M., 2002. Loopy belief propagation and Gibbs measures. In Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence (UAI '02), pages 493-500.