Limit Theorems of estimators in the Tensor Curie Weiss Potts Model

Sanchayan Bhowal¹

(joint work with S. Mukherjee²)

arXiv:2307.01052

¹Statistics and Mathematics Unit Indian Statistical Institute, Bangalore

²Department of Statistics and Data Science National University of Singapore, Singapore

Bangalore Probability Seminar, 26 February, 2024

Thank you Nishant, Siva, and Mathew for the opportunity to present at the Bangalore Probability Seminar.

()





- 3 Asymptotics of Magnetization vector
- Asymptotics of the Maximum Likelihood Estimates

()

Spin Glass Model



Figure: Magnetic Spins in a graph with black vertices denoting -1 and white vertices denoting ± 1

 $\sigma \in \{-1,1\}^n$ denotes the magnetic spin vector in the graph.

$$\mathbb{P}(\sigma) \propto \exp(\beta \sum_{j=1}^n J_{ij}\sigma_i\sigma_j)$$

where J: interaction matrix, $\beta > 0$: inverse temperature.

Potts Model



Figure: Coloring with 3 colors in a graph(may not be proper)

 $X \in [q]^N$ denotes the coloring in the graph.

$$\mathbb{P}(X) \propto \exp\left(eta \sum_{j=1}^n J_{ij} \mathbb{1}_{X_i = X_j}
ight)$$

Tensor Potts Model



Figure: Introducing peer group interactions

 $X \in [q]^n$ denotes the coloring in the hypergraph.

$$\mathbb{P}(X) \propto \exp\left(\beta \sum_{i_1, \dots, i_p} J_{i_1, \dots, i_p} \mathbb{1}_{X_{i_1} = \dots = X_{i_p}} + h \sum_{i=1}^N \mathbb{1}_{X_i = 1}\right)$$

where, $h \ge 0$: magnetic field.

• Consider the Curie-Weiss setting where $J_{i_1,...,i_p} = \frac{1}{N^{p-1}}$.

▲御▶ ▲屋▶ ▲屋▶

- Consider the Curie-Weiss setting where $J_{i_1,...,i_p} = \frac{1}{N^{p-1}}$.
- This is an exponential family with paramters $(\beta, h) \in (0, \infty) \times [0, \infty)$.

• • = • • = •

- Consider the Curie-Weiss setting where $J_{i_1,...,i_p} = \frac{1}{N^{p-1}}$.
- This is an exponential family with paramters
 (β, h) ∈ (0,∞) × [0,∞).
- Define $\bar{X}_r = rac{1}{N}\sum_{i=1}^N \mathbbm{1}_{X_i=r}$ then,

$$\mathbb{P}_{\beta,h,N}(\boldsymbol{X}) := \frac{1}{q^N Z_N(\beta,h)} \exp\left(\beta N \sum_{r=1}^q \bar{X}_r^p + Nh\bar{X}_{\cdot 1}\right) \quad (1.1)$$

• • = • • = •

- Consider the Curie-Weiss setting where $J_{i_1,...,i_p} = \frac{1}{N^{p-1}}$.
- This is an exponential family with paramters
 (β, h) ∈ (0,∞) × [0,∞).
- Define $\bar{X}_r = rac{1}{N}\sum_{i=1}^N \mathbbm{1}_{X_i=r}$ then,

$$\mathbb{P}_{\beta,h,N}(\boldsymbol{X}) := \frac{1}{q^N Z_N(\beta,h)} \exp\left(\beta N \sum_{r=1}^q \bar{X}_r^p + Nh\bar{X}_1\right) \quad (1.1)$$

• Magnetization Vector, $\bar{\boldsymbol{X}}_N := (\bar{X}_r)_{r=1}^q$





3 Asymptotics of Magnetization vector

Asymptotics of the Maximum Likelihood Estimates

()

- Asymptotics of the maximum likelihood (ML) estimates of the parameters β and h.
- Asymptotics of the magnetization vector (\bar{X}_N) .

A B M A B M







4 Asymptotics of the Maximum Likelihood Estimates

< E > < E >

 \bar{X}_N concentrates around global maximizers of $H_{\beta,h}$:

$$H_{\beta,h}(\boldsymbol{t}) := \beta \sum_{r=1}^{q} t_r^p + ht_1 - \sum_{r=1}^{q} t_r \log t_r$$

Theorem (B., Mukherjee (2023))

Let $\beta_N \to \beta$ and $h_N \to h$. Then, under $\mathbb{P}_{\beta_N,h_N,N}$, the empirical magnetization $\bar{\mathbf{X}}_N$ satisfies a large deviation principle with speed N and rate function $-H_{\beta,h} + \sup H_{\beta,h}$.

• • = • • = •

• Regular: if the function $H_{\beta,h}$ has a unique global maximizer \boldsymbol{m}_* and the $\boldsymbol{Q}_{\boldsymbol{m}_*,\beta} = Hess(H_{\beta,h})$ is negative definite at \boldsymbol{m}_* on $\mathcal{H}_q := \{ \boldsymbol{t} \in \mathbb{R}^q : \sum_{r=1}^q t_r = 0 \}.$

- Regular: if the function $H_{\beta,h}$ has a unique global maximizer \boldsymbol{m}_* and the $\boldsymbol{Q}_{\boldsymbol{m}_*,\beta} = Hess(H_{\beta,h})$ is negative definite at \boldsymbol{m}_* on $\mathcal{H}_q := \{ \boldsymbol{t} \in \mathbb{R}^q : \sum_{r=1}^q t_r = 0 \}.$
- **2** Critical: if $H_{\beta,h}$ has more than one global maximizer.

- Regular: if the function H_{β,h} has a unique global maximizer
 *m*_{*} and the *Q_{m*,β}* = Hess(H_{β,h}) is negative definite at *m*_{*} on
 H_q := {*t* ∈ ℝ^q : Σ^q_{r=1} *t_r* = 0}.
- **2** Critical: if $H_{\beta,h}$ has more than one global maximizer.
- Special: if H_{β,h} has a unique global maximizer m_{*} and Q_{m*,β} is singular on H_q.

Lemma

The maximizers
$$\boldsymbol{m}$$
 are of the form $\left(\frac{1+(q-1)s}{q}, \frac{1-s}{q}, \cdots, \frac{1-s}{q}\right)$

$$f_{eta,h}(s) := (q-1)k\left(rac{1-s}{q}
ight) + k\left(rac{1+(q-1)s}{q}
ight) + \left(rac{1+(q-1)s}{q}
ight) \cdot h,$$

where $k(x) = k_{\beta,p}(x) := \beta x^p - x \log x$.

▲御▶ ▲ 国▶ ▲ 国▶

3

Lemma

The maximizers
$$\boldsymbol{m}$$
 are of the form $\left(\frac{1+(q-1)s}{q}, \frac{1-s}{q}, \cdots, \frac{1-s}{q}\right)$

$$f_{eta,h}(s) := (q-1)k\left(rac{1-s}{q}
ight) + k\left(rac{1+(q-1)s}{q}
ight) + \left(rac{1+(q-1)s}{q}
ight) \cdot h,$$

where $k(x) = k_{\beta,p}(x) := \beta x^p - x \log x$. Special:

type-I, if
$$f_{\beta,h}^{(4)}(s) < 0$$
.
 type-II, if $f_{\beta,h}^{(4)}(s) = 0$ (Not observed in Ising Model)



Figure: Partition of the paramter space, for p=7,q=5

< 注入 < 注入 →



Figure: Partition of the paramter space, for p=4,q=2

・ロト ・回ト ・ヨト ・ヨト

CLT of \bar{X}_N under perturbation

Theorem (B., Mukherjee(2023))

9 Regular: for
$$m{X} \sim \mathbb{P}_{eta + N^{-rac{1}{2}}ar{eta}, h+N^{-rac{1}{2}}ar{h}, p}$$
 for some $ar{eta}, ar{h} \in \mathbb{R}$,

$$N^{rac{1}{2}}\left(ar{m{X}}_{N}-m{m}_{*}
ight) \xrightarrow{D} \mathcal{N}_{q}\left(\Sigma(ar{eta}
hom{m}_{*}^{p-1}+ar{h}m{e}_{1}),\Sigma
ight),$$

CLT of \bar{X}_N under perturbation

Theorem (B., Mukherjee(2023))

2

) Regular: for
$$m{X} \sim \mathbb{P}_{eta + m{N}^{-rac{1}{2}}ar{eta}, m{h} + m{N}^{-rac{1}{2}}ar{m{h}}, m{p}}$$
 for some $ar{eta}, ar{m{h}} \in \mathbb{R}$,

$$\mathcal{N}^{\frac{1}{2}}\left(ar{m{X}}_{N}-m{m}_{*}
ight) \xrightarrow{D} \mathcal{N}_{q}\left(\Sigma(ar{eta}
hom{m}_{*}^{p-1}+ar{h}m{e}_{1}),\Sigma
ight),$$

Critical:
• for
$$\mathbf{X} \sim \mathbb{P}_{\beta,h,p}$$
, as $N \to \infty$, we have:
 $\bar{\mathbf{X}}_N \xrightarrow{P} \sum_{k=1}^{K} p_k \delta_{m_k}$,
• under $\mathbb{P}_{\beta_N,h_N} \left(\cdot \left| \bar{\mathbf{X}}_N \in B(\mathbf{m}_i, \varepsilon) \right) :$
 $\sqrt{N} \left(\bar{\mathbf{X}}_N - \mathbf{m}_i \right) \xrightarrow{D} \mathcal{N}_q(\Sigma'(\bar{\beta}p\mathbf{m}_i^{p-1} + \bar{h}\mathbf{e}_1), \Sigma')$,

CLT of \bar{X}_N under perturbation

Theorem ((contd.))

3

• Special: Let
$$\mathbf{u} =:= (1 - q, 1, \dots, 1)$$
,
• Type-I: Under $\mathbb{P}_{\beta + N^{-\frac{3}{4}}\overline{\beta}, h + N^{-\frac{3}{4}}\overline{h}, p}$, as $N \to \infty$,

$$N^{\frac{1}{4}}(\bar{\boldsymbol{X}}_{N}-\boldsymbol{m}_{*})\xrightarrow{D}T_{\bar{\beta},\bar{h}}\boldsymbol{u},$$

where $T_{\bar{\beta},\bar{h}}$ has density proportional to,

$$\exp\left(\frac{x^4}{24}q^4f^{(4)}_{\beta,h}(s)+(\bar{\beta}p\langle \boldsymbol{m}^{p-1}_*,\boldsymbol{u}\rangle+\bar{h}(1-q))x\right).$$

・ロト ・回ト ・ヨト ・ヨト

CLT of $ar{m{X}}_N$ under perturbation

Theorem ((contd.))

Special: Let $\mathbf{u} =:= (1 - q, 1, \dots, 1)$,
• Type-I: Under $\mathbb{P}_{\beta+N^{-\frac{3}{4}}\bar{\beta},h+N^{-\frac{3}{4}}\bar{h},p}$, as $N \to \infty$,

$$N^{\frac{1}{4}}(\bar{\boldsymbol{X}}_{N}-\boldsymbol{m}_{*})\xrightarrow{D}T_{\bar{\beta},\bar{h}}\boldsymbol{u},$$

where $T_{ar{eta},ar{h}}$ has density proportional to,

$$\exp\left(\frac{x^4}{24}q^4f^{(4)}_{\beta,h}(s)+(\bar{\beta}p\langle\boldsymbol{m}^{p-1}_*,\boldsymbol{u}\rangle+\bar{h}(1-q))x\right).$$

• Type-II: Under $\mathbb{P}_{\beta+N^{-\frac{5}{6}}\bar{\beta},h+N^{-\frac{5}{6}}\bar{h},p}$

$$N^{\frac{1}{6}}\left(\bar{\boldsymbol{X}}_{N}-\boldsymbol{m}_{*}\right)\xrightarrow{D}F_{\bar{h}}\boldsymbol{u}.$$

where $F_{\bar{h}}$ has density proportional to $\exp\left(-\frac{32}{15}x^6 - \bar{h}x\right)$.

э

Proof Sketch

- Denote $\boldsymbol{W}_N := \sqrt{N} \left(\boldsymbol{\bar{X}}_N \boldsymbol{m}_* \right).$
- $g: \mathbb{R}^q
 ightarrow \mathbb{R}$ a bounded, continuous function .

伺 と く ヨ と く ヨ と …

Proof Sketch

- Denote $\boldsymbol{W}_N := \sqrt{N} \left(\boldsymbol{\bar{X}}_N \boldsymbol{m}_* \right).$
- $\bullet~g:\mathbb{R}^q\to\mathbb{R}$ a bounded, continuous function .
- Weak Convergence:

$$q^{N} Z_{N}(\beta_{N}, h_{N}) \mathbb{E}_{\beta_{N}, h_{N}, N} \left[g(\boldsymbol{W}_{N}) \mathbb{1}_{\|\boldsymbol{W}_{N}\| \leq M} \right]$$
$$= \sum g(\boldsymbol{w}(\boldsymbol{v})) \mathbb{1}_{\|\boldsymbol{w}(\boldsymbol{v})\| \leq M} q^{N} Z_{N}(\beta_{N}, h_{N}) \mathbb{P}_{\beta_{N}, h_{N}, N}(\bar{\boldsymbol{X}}_{N} = \boldsymbol{v})$$

• • = • • = •

Proof Sketch

- Denote $\boldsymbol{W}_N := \sqrt{N} \left(\boldsymbol{\bar{X}}_N \boldsymbol{m}_* \right).$
- $\bullet~g:\mathbb{R}^q\to\mathbb{R}$ a bounded, continuous function .
- Weak Convergence:

$$q^{N}Z_{N}(\beta_{N},h_{N})\mathbb{E}_{\beta_{N},h_{N},N}\left[g(\boldsymbol{W}_{N})\mathbb{1}_{||\boldsymbol{W}_{N}||\leq M}\right]$$
$$=\sum g(\boldsymbol{w}(\boldsymbol{v}))\mathbb{1}_{||\boldsymbol{w}(\boldsymbol{v})||\leq M}q^{N}Z_{N}(\beta_{N},h_{N})\mathbb{P}_{\beta_{N},h_{N},N}(\bar{\boldsymbol{X}}_{N}=\boldsymbol{v})$$

• A small lemma,

$$q^{N}Z_{N}(\beta,h)\mathbb{P}_{\beta,h,N}(\bar{\boldsymbol{X}}_{N}=\boldsymbol{v}) \sim N^{-\frac{q-1}{2}}A(\boldsymbol{v})e^{NH_{\beta,h}(\boldsymbol{v})}$$

where $A(\boldsymbol{v}) := (2\pi)^{-(q-1)/2}\prod_{r=1}^{q}v_{r}^{-1/2}.$

• • = • • = •

э

Taylor expand H_{β_N,h_N} to get,

$$q^{N}Z_{N}(\beta_{N},h_{N})\mathbb{E}_{\beta_{N},h_{N},N}\left[g(\boldsymbol{W}_{N})\mathbb{1}_{\|\boldsymbol{W}_{N}\|\leq M}\right]$$

★ E ► < E ►</p>

Taylor expand H_{β_N,h_N} to get,

$$q^{N}Z_{N}(\beta_{N},h_{N})\mathbb{E}_{\beta_{N},h_{N},N}\left[g(\boldsymbol{W}_{N})\mathbb{1}_{\|\boldsymbol{W}_{N}\|\leq M}\right]$$

$$\sim \frac{A(\boldsymbol{m}_{*})}{N^{\frac{q-1}{2}}}e^{NH_{\beta_{N},h_{N}}(\boldsymbol{m}_{*})}\sum g(\boldsymbol{w})\mathbb{1}_{\|\boldsymbol{w}\|\leq M}e^{\langle\bar{\beta}p\boldsymbol{m}_{*}^{p-1}+\bar{h}\boldsymbol{e}_{1},\boldsymbol{w}\rangle+\frac{1}{2}\boldsymbol{Q}_{\boldsymbol{m}_{*},\beta}(\boldsymbol{w})$$

★ E ► < E ►</p>

Taylor expand H_{β_N,h_N} to get,

$$q^{N}Z_{N}(\beta_{N},h_{N})\mathbb{E}_{\beta_{N},h_{N},N}\left[g(\boldsymbol{W}_{N})\mathbb{1}_{||\boldsymbol{W}_{N}||\leq M}\right]$$

$$\sim \frac{A(\boldsymbol{m}_{*})}{N^{\frac{q-1}{2}}}e^{NH_{\beta_{N},h_{N}}(\boldsymbol{m}_{*})}\sum g(\boldsymbol{w})\mathbb{1}_{||\boldsymbol{w}||\leq M}e^{\langle\bar{\beta}p\boldsymbol{m}_{*}^{p-1}+\bar{h}\boldsymbol{e}_{1},\boldsymbol{w}\rangle+\frac{1}{2}\boldsymbol{Q}_{\boldsymbol{m}_{*},\beta}(\boldsymbol{w})}$$

$$\sim A(\boldsymbol{m}_{*})e^{NH_{\beta_{N},h_{N}}(\boldsymbol{m}_{*})}\int_{\mathcal{H}_{q}\bigcap B(\boldsymbol{0},M)}g(\boldsymbol{w})e^{\langle\bar{\beta}p\boldsymbol{m}_{*}^{p-1}+\bar{h}\boldsymbol{e}_{1},\boldsymbol{w}\rangle+\frac{1}{2}\boldsymbol{Q}_{\boldsymbol{m}_{*},\beta}(\boldsymbol{w})}$$

< E > < E >

Hence, under $\mathbb{P}_{\beta_N,h_N,N}$, W_N conditioned on $||W_N|| \leq M$ converges weakly to the density proportional to

$$\mathbf{w} \mapsto e^{\langle ar{eta} p \mathbf{m}_*^{p-1} + ar{h} \mathbf{e}_1, \mathbf{w}
angle + rac{1}{2} \mathbf{Q}_{\mathbf{m}_*, \beta}(\mathbf{w})}$$

Conclude the proof by uniform integrability of $\{X_N\}$.





3 Asymptotics of Magnetization vector



Asymptotics of the Maximum Likelihood Estimates

* 3 > < 3</p>

We interested to esimate:

- **1** β with *h* known,
- **2** *h* with β known.

MLE of the parameters

Define $u_{N,p}$ and $u_{N,1}$ as,

 $u_{N,p}(\beta, h, p) := \mathbb{E}_{\beta,h,p}(\|\bar{\boldsymbol{X}}_N\|_p^p) \quad \text{and} \quad u_{N,1}(\beta, h, p) := \mathbb{E}_{\beta,h,p}(\bar{X}_1).$

From classical results of exponential family the ML estimate $\hat{\beta}$ satisfies the equation for fixed *h*:

$$u_{N,p}(eta,h,p) = \|ar{oldsymbol{X}}_N\|_p^p$$

and for fixed β , the ML estimate \hat{h} satisfies the equation:

$$u_{N,1}(eta,h,p)=ar{X}_1.$$

Lemma

 $u_{N,p}(\beta, h)$ and $u_{N,1}(\beta, h)$ are increasing in β and h respectively.

Proof Sketch: Define $F_N := \log q^N Z_N(\beta, h)$. Then note that $\frac{\partial}{\partial \beta} F_N = u_{N,p}(\beta, h)$ and $\frac{\partial}{\partial h} F_N = u_{N,1}(\beta, h)$. From holder's inequality F_N is convex in β, h .

A B F A B F

Limit Theorems of MLE

Theorem (B., Mukherjee(2023))

For regular points we have:

$$N^{rac{1}{2}}\left(\hat{h}_N-h
ight) \stackrel{D}{
ightarrow} \mathcal{N}\left(0,-rac{q^2}{(q-1)^2}f_{eta,h}''(s)
ight)$$

Proof.

$$\mathbb{P}_{\beta,h,p}\left(N^{\frac{1}{2}}\left(\hat{h}_{N}-h\right)\leq t\right)$$

$$=\mathbb{P}_{\beta,h,p}\left(\hat{h}_{N}\leq h+\frac{t}{N^{\frac{1}{2}}}\right)$$

$$=\mathbb{P}_{\beta,h,p}\left(u_{N,1}\left(\beta,\hat{h}_{N},p\right)\leq u_{N,1}\left(\beta,h+\frac{t}{N^{\frac{1}{2}}},p\right)\right)$$

() 《 문 》 《

Proof(contd.)

$$\begin{split} &= \mathbb{P}_{\beta,h,p} \left(\bar{X}_1 \leq \mathbb{E}_{\beta,h+N^{-\frac{1}{2}}t,p} \left(\bar{X}_1 \right) \right) \\ &= \mathbb{P}_{\beta,h,p} \left(N^{\frac{1}{2}} \left(\bar{X}_1 - m_1 \right) \leq \mathbb{E}_{\beta,h+N^{-\frac{1}{2}}t,p} \left(N^{\frac{1}{2}} \left(\bar{X}_1 - m_1 \right) \right) \right) \\ &\rightarrow \mathbb{P}_{\beta,h,p} \left(\mathcal{N} \left(0, -\frac{(q-1)^2}{q^2 f_{\beta,h}''(s)} \right) \leq -\frac{t(q-1)^2}{q^2 f_{\beta,h}''(s)} \right) \\ &= \mathbb{P}_{\beta,h,p} \left(\mathcal{N} \left(0, -\frac{q^2 f_{\beta,h}''(s)}{(q-1)^2} \right) \leq t \right). \end{split}$$

▲□ ▶ ▲ □ ▶ ▲ □ ▶ ...



Figure: Regular: $\hat{\beta}_N$ is \sqrt{N} consistent except at the red line. \hat{h}_N is \sqrt{N} consistent



Figure: Strongly Critical: $\hat{\beta}_N$ is \sqrt{N} consistent. \hat{h}_N is \sqrt{N} consistent



Figure: Weakly Critical: $\hat{\beta}_N$ is \sqrt{N} inconsistent. \hat{h}_N is \sqrt{N} consistent

э



Figure: Transition Point: $\hat{\beta}_N$ is \sqrt{N} inconsistent. \hat{h}_N is \sqrt{N} consistent

э

э



Figure: Special Point: $\hat{\beta}_N$ is $N^{3/4}$ consistent except for (p,q) = (2,2), (3,2). \hat{h}_N is $N^{3/4}$ consistent.

Special Case of p = 4, q = 2



Figure: Special Point: $\hat{\beta}_N$ is $N^{5/6}$ in-consistent. \hat{h}_N is $N^{5/6}$ consistent.

э

・ 同 ト ・ ヨ ト ・ ヨ ト

• Phase transitions in the asymptotics of the magnetization vector.

< ∃ >

- Phase transitions in the asymptotics of the magnetization vector.
- Asymptotics depend largely on the number of maximizers of a certain function $H_{\beta,h}$ and (p,q).

- Phase transitions in the asymptotics of the magnetization vector.
- Asymptotics depend largely on the number of maximizers of a certain function $H_{\beta,h}$ and (p,q).
- For "almost every" points, both $\hat{\beta}_N$ and \hat{h}_N are \sqrt{N} -consistent and asymptotically normal.

- Phase transitions in the asymptotics of the magnetization vector.
- Asymptotics depend largely on the number of maximizers of a certain function $H_{\beta,h}$ and (p,q).
- For "almost every" points, both $\hat{\beta}_N$ and \hat{h}_N are \sqrt{N} -consistent and asymptotically normal.
- At one point, both $\hat{\beta}_N$ and \hat{h}_N can be superefficient, converging at rate $N^{3/4}$ to non-Gaussian distributions.

- Phase transitions in the asymptotics of the magnetization vector.
- Asymptotics depend largely on the number of maximizers of a certain function $H_{\beta,h}$ and (p,q).
- For "almost every" points, both $\hat{\beta}_N$ and \hat{h}_N are \sqrt{N} -consistent and asymptotically normal.
- At one point, both $\hat{\beta}_N$ and \hat{h}_N can be superefficient, converging at rate $N^{3/4}$ to non-Gaussian distributions.
- There exists a continuous curve in the interior of the parameter plane, on which the MLEs have mixture distributions. The components are normal/half-normal and point masses.

- Phase transitions in the asymptotics of the magnetization vector.
- Asymptotics depend largely on the number of maximizers of a certain function $H_{\beta,h}$ and (p,q).
- For "almost every" points, both $\hat{\beta}_N$ and \hat{h}_N are \sqrt{N} -consistent and asymptotically normal.
- At one point, both $\hat{\beta}_N$ and \hat{h}_N can be superefficient, converging at rate $N^{3/4}$ to non-Gaussian distributions.
- There exists a continuous curve in the interior of the parameter plane, on which the MLEs have mixture distributions. The components are normal/half-normal and point masses.
- The curious case of (p,q) = (4,2), where \hat{h}_N can converge at rate of $N^{5/6}$ to a non-Gaussian distribution.

Thank You!