Fluctuations in the mean-field Ising model

Nabarun Deb

(Joint work with Dr. Sumit Mukherjee, Department of Statistics, Columbia University)

2019 Saint-Flour Probability Summer School.

Structure

Structure

The Ising Model

Consider,

$$\mathbb{P}^{N}_{\beta,B}(\sigma=\tau) := \frac{1}{Z_N(\beta,B)}\exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^{N}\tau_i\right) \quad (1)$$

for $\tau \in \{-1,1\}^N$.

The Ising Model

Consider,

$$\mathbb{P}^N_{\beta,B}(\sigma = \tau) := \frac{1}{Z_N(\beta, B)} \exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^N \tau_i\right) \quad (1)$$

for $\tau \in \{-1, 1\}^N$.

1. $A(G_N)$ – Adjacency matrix of an underlying graph (may be random), think of social networks.

The Ising Model

Consider,

$$\mathbb{P}^N_{\beta,B}(\sigma = \tau) := \frac{1}{Z_N(\beta, B)} \exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^{N}\tau_i\right) \quad (1)$$

for $\tau \in \{-1, 1\}^N$.

1. $A(G_N)$ – Adjacency matrix of an underlying graph (may be random), think of social networks.
2. $\bar{d}_N$ – Average degree of G_N, i.e., $\bar{d}_N = (1/N)\sum_i d_i$.

The Ising Model

Consider,

$$\mathbb{P}^N_{\beta,B}(\sigma = \tau) := \frac{1}{Z_N(\beta, B)} \exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^{N} \tau_i\right) \tag{1}$$

for $\tau \in \{-1, 1\}^N$.

1. $A(G_N)$ – Adjacency matrix of an underlying graph (may be random), think of social networks.
2. $\bar{d}_N$ – Average degree of G_N, i.e., $\bar{d}_N = (1/N)\sum_i d_i$.
3. (β, B) – parameters in $[0, \infty) \times (-\infty, \infty)$.
4. We observe one copy of (spins) σ.

The Ising Model (continued)

Consider,

$$\mathbb{P}^{N}_{\beta,B}(\sigma = \tau) := \frac{1}{Z_N(\beta, B)} \exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^{N}\tau_i\right)$$

for $\tau \in \{-1, 1\}^N$.

The Ising Model (continued)

Consider,

$$\mathbb{P}^N_{\beta,B}(\sigma = \tau) := \frac{1}{Z_N(\beta, B)} \exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^{N}\tau_i\right)$$

for $\tau \in \{-1, 1\}^N$.

1. $\beta = 0$ implies X_i's are i.i.d.
2. $B > 0$ implies the spin sites incline towards $+1$.
3. $B < 0$ implies the spin sites incline towards -1.

The Ising Model (continued)

Consider,

$$\mathbb{P}^N_{\beta,B}(\sigma = \tau) := \frac{1}{Z_N(\beta, B)} \exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^{N}\tau_i\right)$$

for $\tau \in \{-1, 1\}^N$.

1. $\beta = 0$ implies X_i's are i.i.d.
2. $B > 0$ implies the spin sites incline towards $+1$.
3. $B < 0$ implies the spin sites incline towards -1.
4. $B = 0$ implies the absence of external influence.

The Ising Model (continued)

Consider,

$$\mathbb{P}^N_{\beta,B}(\sigma=\tau) := \frac{1}{Z_N(\beta,B)}\exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^N \tau_i\right)$$

for $\tau \in \{-1,1\}^N$.

1. $\beta = 0$ implies X_i's are i.i.d.
2. $B > 0$ implies the spin sites incline towards $+1$.
3. $B < 0$ implies the spin sites incline towards -1.
4. $B = 0$ implies the absence of external influence.
5. $(A(G_N))_{ij} > 0$ implies sites i and j are inclined to align in the same direction.

Assumptions

(A1) (Bounded row sums):

$$\limsup_{N\to\infty} \max_{1\le i\le N} \frac{d_i}{\overline{d}_N} < \infty.$$

(A2) (Well connectedness):

$$\limsup_{N\to\infty} \lambda_2\left(\frac{A(G_N)}{\overline{d}_N}\right) < 1.$$

Commonly studied examples

$A(G_N)$ is the adjacency matrix of:

- Lattice graph in d dimensions: Original paper by Ising (1925) considered the one dimensional lattice. It was proposed by Lenz to model sharp phase transition of ferromagnetic properties (at Curie temperature).
- Complete graphs: Curie-Weiss model.
- Random graphs: Erdős Rényi, random regular, etc.

Applications

1. Modeling dynamics of interactive systems –e.g., atomic motion in lattice gas, earthquake dynamics, image segmentation, etc.

Applications

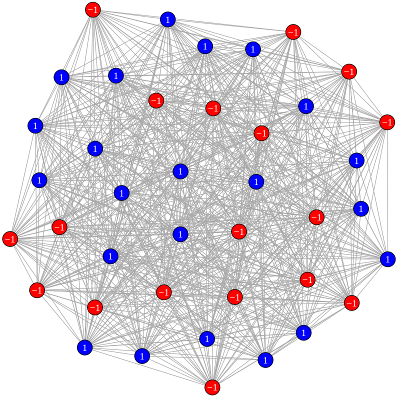
1. Modeling dynamics of interactive systems –e.g., atomic motion in lattice gas, earthquake dynamics, image segmentation, etc.

2. In social networks, to study trends in opinions (voting choices), where G_N could be determined by "friendships" within the network.

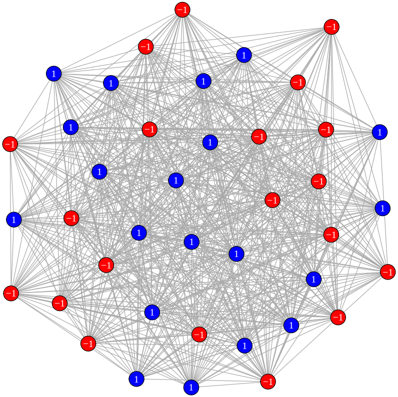
Applications

1. Modeling dynamics of interactive systems –e.g., atomic motion in lattice gas, earthquake dynamics, image segmentation, etc.
2. In social networks, to study trends in opinions (voting choices), where G_N could be determined by "friendships" within the network.
3. Modeling ferromagnetic properties (i.e., **sharp** change in magnetic properties of magnetic materials when heated beyond a certain (**Curie**) temperature).

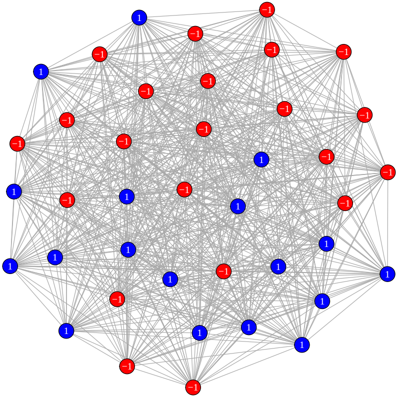
$N = 40$, $p = 0.9$, $\beta = 0.5$, $B = 0$, seed=7.11



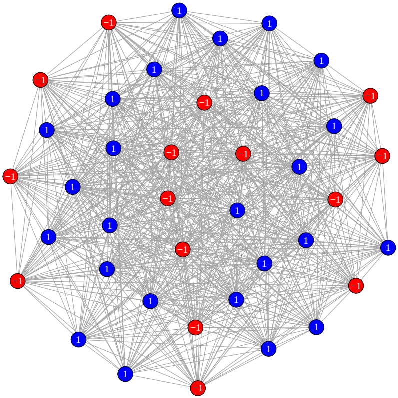
$N = 40,\ p = 0.9,\ \beta = 0.5,\ B = 0$



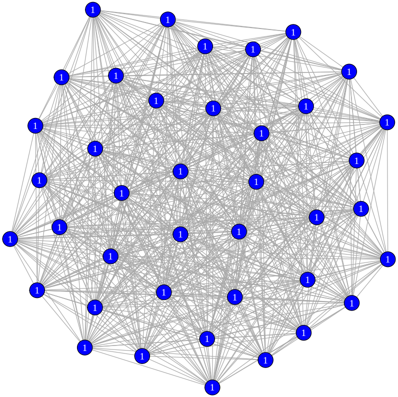
$N = 40,\ p = 0.9,\ \beta = 0.5,\ B = 0$



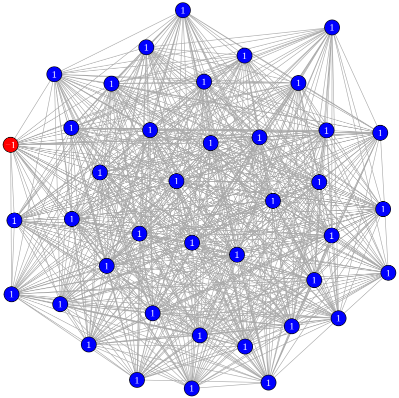
$N = 40,\ p = 0.9,\ \beta = 0.5,\ B = 0$



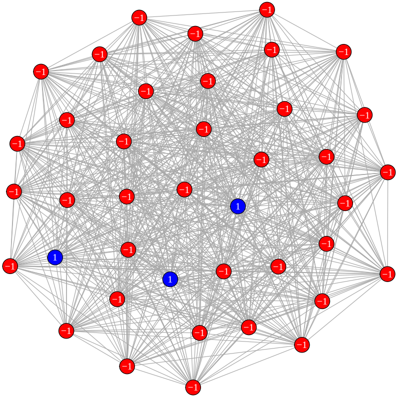
$N = 40$, $p = 0.9$, $\beta = 1.5$, $B = 0$, seed=7.11



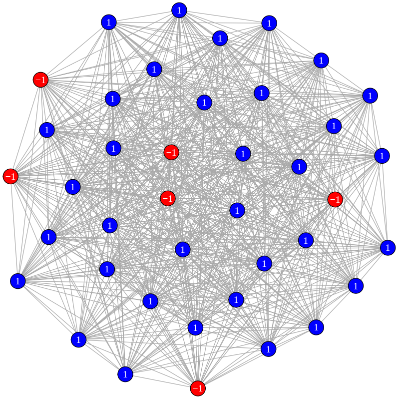
$N = 40$, $p = 0.9$, $\beta = 1.5$, $B = 0$



$N = 40,\ p = 0.9,\ \beta = 1.5,\ B = 0$



$N = 40,\ p = 0.9,\ \beta = 1.5,\ B = 0$



Applications continued...

- Physics: Ising [1925], Fytas et al. [2019].
- Social Sciences: Marsman et al. [2018], Schelling [1971].
- Neuroscience: Yuste [2015], Fraiman et al. [2009].
- Image analysis: Besag [1986], Sun et al. [2012].
- Machine Learning: Hopfield [1982], Zhang et al. [2015].
- Earth Sciences: Siegel et al. [2018], Raymond et al. [2016].
- Theoretical (goodness of fit): Martín del Campo et al. [2017], Daskalakis et al. [2018].

Statistical motivation

Motivating Theorem, Barndorff [2014]

Suppose $\alpha \in \mathbb{R}$, $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ and $\mathcal{P}_\mu(\mathbb{R}^n)$: prob. measures on $\mathbb{R}^n$, $\ll \mu$. Then,

$$\max_{\mathcal{P}_\mu(\mathbb{R}^n)} -\int p(x) \log p(x)\, d\mu(x) \text{ subject to } \mathbb{E}_p \phi(\cdot) = \alpha$$

is attained at uniquely at $p^*_\theta(x) \propto \exp(\theta\phi(x))$ where $\mathbb{E}_{p^*_\theta}\phi(\cdot) = \alpha$.

Statistical motivation

Motivating Theorem, Barndorff [2014]

Suppose $\alpha \in \mathbb{R}$, $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ and $\mathcal{P}_\mu(\mathbb{R}^n)$: prob. measures on $\mathbb{R}^n$, $\ll \mu$. Then,

$$\max_{\mathcal{P}_\mu(\mathbb{R}^n)} -\int p(x) \log p(x)\, d\mu(x) \text{ subject to } \mathbb{E}_p \phi(\cdot) = \alpha$$

is attained at uniquely at $p^*_\theta(x) \propto \exp\left(\theta\phi(x)\right)$ where $\mathbb{E}_{p^*_\theta}\phi(\cdot) = \alpha$.

Put constraints on pairwise weighted correlations (simple interaction): $\sum_{i,j} w_{ij}\sigma_i\sigma_j$ (from the data). Then $p^*(\sigma) \propto \exp(\beta \sum_{i,j} w_{ij}\sigma_i\sigma_j)$ ((1) with $B = 0$).

Structure

Literature Review

Recall, for $\tau \in \{-1, 1\}^N$,

$$\mathbb{P}^N_{\beta,B}(\sigma = \tau) := \frac{1}{Z_N(\beta, B)} \exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^{N}\tau_i\right)$$

Literature Review

Recall, for $\tau \in \{-1,1\}^N$,

$$\mathbb{P}^N_{\beta,B}(\sigma = \tau) := \frac{1}{Z_N(\beta,B)} \exp\left(\frac{\beta}{2\bar{d}_N}\tau' A(G_N)\tau + B\sum_{i=1}^{N}\tau_i\right)$$

- (Main motivation): Basak and Mukherjee [2017] show that,

$$\frac{1}{N}\log Z_N(\beta,B) \tag{2}$$

has an "universal limit" (free of the underlying graph sequence) as long as G_N is "approximately" regular and $\bar{d}_N$ diverges to ∞ with N.

Regular graphs

- $A_N := A(G_N)/\overline{d}_N$.
- Note that $(A_N)_{ij} = 1/\overline{d}_N$ if i and j are neighbors in G_N.

Regular graphs

- $A_N := A(G_N)/\overline{d}_N$.
- Note that $(A_N)_{ij} = 1/\overline{d}_N$ if i and j are neighbors in G_N.
- $\sum_j (A_N)_{ij} = d_i/\overline{d}_N = 1$ for all $i \in [N]$ (equivalently $\mathbf{1}'A_N = \mathbf{1}'$) : **(P1)**.
- By Perron-Frobenius, $\lambda_1(A_N) = 1$: **(P2)**.

An example of "Approximately regular" graphs

- Suppose $G_N \sim$ Erdős Rényi (N, p_N) graph, with $p_N \gg \log N/N$.

An example of "Approximately regular" graphs

- Suppose $G_N \sim$ Erdős Rényi (N, p_N) graph, with $p_N \gg \log N/N$.
- In this case, $\sum_{i=1}^N (d_i/\overline{d}_N - 1)^2 = o_{\mathbb{P}}(N/\log N) = o_{\mathbb{P}}(N)$ (rather weak notion of regularity).

An example of "Approximately regular" graphs

- Suppose $G_N \sim$ Erdős Rényi (N, p_N) graph, with $p_N \gg \log N/N$.
- In this case, $\sum_{i=1}^N (d_i/\overline{d}_N - 1)^2 = o_{\mathbb{P}}(N/\log N) = o_{\mathbb{P}}(N)$ (rather weak notion of regularity).
- Our methods also work under random graphs, such as Erdős Rényi above.
- Lower bounds on the "extent of regularity".

Variational Approach to "universality"

- Given $\mathbb{Q}_N(\mathbf{x})$ be a discrete measure supported on $\{-1,1\}^N$. Then $D(\mathbb{Q}_N \| \mathbb{P}_N)$ equals,

$$\sum_{\mathbf{x}\in\{-1,1\}^N} \mathbb{Q}_N(\mathbf{x}) \log \frac{\mathbb{Q}_N(\mathbf{x})}{\mathbb{P}_N(\mathbf{x})}$$

$$= \sum_{\mathbf{x}\in\{-1,1\}^N} \mathbb{Q}_N(\mathbf{x}) \log \mathbb{Q}_N(\mathbf{x}) - \frac{\beta}{2} \sum_{\mathbf{x}\in\{-1,1\}^N} \mathbb{Q}_N(\mathbf{x}) \mathbf{x}' A_N \mathbf{x}$$

$$- B \sum_{\mathbf{x}\in\{-1,1\}^N} \mathbb{Q}_N(\mathbf{x}) \mathbf{1}'\mathbf{x} + \log Z_N(\beta, B)$$

Variational Approach towards "universality"

- As $D(\cdot\|\cdot) \geq 0$,

$$\log Z_N(\beta, B) \geq \mathbb{E}_{\mathbb{Q}_N}\left\{\frac{\beta}{2}\mathbf{x}' A_N \mathbf{x} + B\mathbf{1}'\mathbf{x} - \log \mathbb{Q}_N(\mathbf{x})\right\}.$$

for any $\mathbb{Q}_N(\cdot)$.

Variational Approach towards "universality"

- As $D(\cdot\|\cdot) \geq 0$,

$$\log Z_N(\beta, B) \geq \mathbb{E}_{\mathbb{Q}_N}\left\{\frac{\beta}{2}\mathbf{x}' A_N \mathbf{x} + B\mathbf{1}'\mathbf{x} - \log \mathbb{Q}_N(\mathbf{x})\right\}.$$

  for any $\mathbb{Q}_N(\cdot)$.

- Choose $\mathbb{Q}_N(\mathbf{x})$ as the measure induced by independent ± 1 rv's with mean vector $\mathbf{m} := (m_1, m_2, \ldots, m_N)$.

$$\log Z_N(\beta, B) \geq \frac{\beta}{2}\mathbf{m}' A_N \mathbf{m} + \sum_{i=1}^{N}(Bm_i - I(m_i))$$

  for any $\mathbf{m}$, where $I(x) := \frac{1+x}{2}\log\frac{1+x}{2} + \frac{1-x}{2}\log\frac{1-x}{2}$ is the binary entropy function.

Variational Approach towards "universality"

- We are interested in

$$\sup_{\mathbf{m}} \frac{\beta}{2}\mathbf{m}' A_N \mathbf{m} + \sum_{i=1}^{N} (Bm_i - I(m_i))$$

which (as of now) depends on the underlying graph.

Variational Approach towards "universality"

- We are interested in

$$\sup_{\mathbf{m}} \frac{\beta}{2}\mathbf{m}' A_N \mathbf{m} + \sum_{i=1}^{N} (Bm_i - I(m_i))$$

which (as of now) depends on the underlying graph.

- Now restrict to "regular graphs". By **(P2)**,

$$\sup_{\mathbf{m}} \left\{ \frac{\beta}{2}\mathbf{m}' A_N \mathbf{m} + \sum_{i=1}^{N} (Bm_i - I(m_i)) \right\}$$

$$\leq \sup_{\mathbf{m}} \sum_{i=1}^{N} \left\{ \frac{\beta}{2} m_i^2 + Bm_i - I(m_i) \right\}$$

which is decoupled.

Variational Approach towards “universality”

- The supremum on RHS is attained at $\mathbf{m} = t\mathbf{1}'$ where $t > 0$ satisfies $t = \tanh(\beta t + B)$. Therefore,

$$\sup_{\mathbf{m}} \left\{ \frac{\beta}{2}\mathbf{m}' A_N \mathbf{m} + \sum_{i=1}^{N} (Bm_i - I(m_i)) \right\}$$

$$\leq N \left\{ \frac{\beta t^2}{2} + Bt - I(t) \right\}$$

Variational Approach towards "universality"

- The supremum on RHS is attained at $\mathbf{m} = t\mathbf{1}'$ where $t > 0$ satisfies $t = \tanh(\beta t + B)$. Therefore,

$$\sup_{\mathbf{m}} \left\{ \frac{\beta}{2}\mathbf{m}' A_N \mathbf{m} + \sum_{i=1}^{N} (Bm_i - I(m_i)) \right\} \leq N \left\{ \frac{\beta t^2}{2} + Bt - I(t) \right\}$$

- By **(P1)**, $\mathbf{m}' A_N \mathbf{m} = Nt^2$ at $\mathbf{m} = t\mathbf{1}'$. So,

$$\sup_{\mathbf{m}} \left\{ \frac{\beta}{2}\mathbf{m}' A_N \mathbf{m} + \sum_{i=1}^{N} (Bm_i - I(m_i)) \right\} = N \left\{ \frac{\beta t^2}{2} + Bt - I(t) \right\}.$$

Variational Approach towards "universality"

- Combining above observations,

$$\frac{\log Z_N(\beta, B)}{N} \geq \frac{\beta t^2}{2} + Bt - I(t) \tag{3}$$

for regular graphs. The RHS above is often called "mean-field prediction".

- Note that the right hand side no longer depends on A_N ("universality") for the whole spectrum of (β, B).

Actual Theorem

Theorem in Basak and Mukherjee [2017] (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \xrightarrow{N\to\infty} \infty$, then:

$$\frac{1}{N}\log Z_N(\beta, B) \xrightarrow{N\to\infty} \sup_x \left(\frac{\beta}{2}x^2 + Bx - I(x)\right).$$

Actual Theorem

Theorem in Basak and Mukherjee [2017] (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \stackrel{N\to\infty}{\longrightarrow} \infty$, then:

$$\frac{1}{N} \log Z_N(\beta, B) \stackrel{N\to\infty}{\longrightarrow} \sup_x \left(\frac{\beta}{2} x^2 + Bx - I(x) \right).$$

- (Lower bound) There are bounded degree graph sequences for which the "mean-field prediction" does not hold.

Statistic of interest

A statistic of interest: $T_N = \frac{1}{N}\sum_i \sigma_i$.

Statistic of interest

A statistic of interest: $T_N = \frac{1}{N}\sum_i \sigma_i$.

- Average alignment of magnetic spins.
- Who wins the vote?

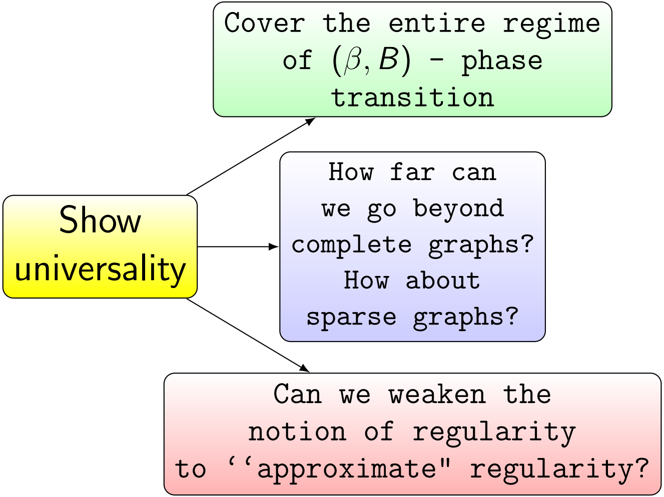
Statistic of interest

A statistic of interest: $T_N = \frac{1}{N}\sum_i \sigma_i$.

- Average alignment of magnetic spins.
- Who wins the vote?

Is the behaviour of T_N universal?

Our goal


Show universality
Cover the entire regime of (β, B) - phase transition
How far can we go beyond complete graphs? How about sparse graphs?
Can we weaken the notion of regularity to ''approximate'' regularity?

Weak Limits

Proposition (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \xrightarrow{N\to\infty} \infty$, then:

Weak Limits

Proposition (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \xrightarrow{N\to\infty} \infty$, then:

1. If $\beta \leq 1, B = 0$ or $B > 0$, then $T_N \xrightarrow{\mathbb{P}^{\mathbb{N}}_{\beta,\mathbb{B}}} t$.

Weak Limits

Proposition (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \xrightarrow{N\to\infty} \infty$, then:

1. If $\beta \leq 1, B = 0$ or $B > 0$, then $T_N \xrightarrow{\mathbb{P}^{\mathbb{N}}_{\beta,\mathbb{B}}} t$.
2. If $\beta > 1, B = 0$ and G_N, then

$$T_N \xrightarrow{\mathbb{P}^{\mathbb{N}}_{\beta,\mathbb{B}}} \begin{cases} t & \text{w.p. } 0.5 \\ -t & \text{w.p. } 0.5 \end{cases}.$$

Here $t \geq 0$ is the maximizer of the RHS in (3). For non-uniqueness, **(A2)** is required.

A natural question

Does universality extend beyond $Z_N(\beta, B)$ and the weak limits for T_N?

$$N^{?}(T_N - t) \xrightarrow{w} \text{ universal limit distributions}$$

(perhaps after conditioning).

Our results

Theorem 1 (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \gg \sqrt{N}$, then:

Our results

Theorem 1 (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \gg \sqrt{N}$, then:

(i). (Uniqueness) If $\beta < 1, B = 0$ or $B > 0$,

$$\sqrt{N}(T_N - t) \xrightarrow{d} \mathcal{N}\left(0, \frac{1-t^2}{1-\beta(1-t^2)}\right).$$

Our results

Theorem 1 (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \gg \sqrt{N}$, then:

(i). (Uniqueness) If $\beta < 1, B = 0$ or $B > 0$,

$$\sqrt{N}(T_N - t) \xrightarrow{d} \mathcal{N}\left(0, \frac{1-t^2}{1-\beta(1-t^2)}\right).$$

(ii). (Non-uniqueness) If $\beta > 1, B = 0$ (and **(A2)** holds),

$$\sqrt{N}(T_N - t)\big| T_N > 0 \xrightarrow{d} \mathcal{N}\left(0, \frac{1-t^2}{1-\beta(1-t^2)}\right).$$

Our results

Theorem 1 (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \gg \sqrt{N}$, then:

(i). (Uniqueness) If $\beta < 1, B = 0$ or $B > 0$,
$$\sqrt{N}(T_N - t) \xrightarrow{d} \mathcal{N}\left(0, \frac{1-t^2}{1-\beta(1-t^2)}\right).$$

(ii). (Non-uniqueness) If $\beta > 1, B = 0$ (and **(A2)** holds),
$$\sqrt{N}(T_N - t)\Big| T_N > 0 \xrightarrow{d} \mathcal{N}\left(0, \frac{1-t^2}{1-\beta(1-t^2)}\right).$$

(iii). (Critical) if $B = 0, \beta = 1$ (and **(A2)** holds),
$$N^{1/4} T_N \xrightarrow{d} W(\text{ density } \propto \exp(-x^4/12)).$$

Is $\sqrt{N}$ an artifact?

Can we have universal behaviour for $\bar{d}_N \leq \sqrt{N}$?

Is $\sqrt{N}$ an artifact?

Can we have universal behaviour for $\bar{d}_N \leq \sqrt{N}$?
No! ($\sqrt{N}$ is tight).

Is $\sqrt{N}$ an artifact?

Can we have universal behaviour for $\bar{d}_N \le \sqrt{N}$?
No! ($\sqrt{N}$ is tight).

Eg. Consider $\bar{d}_N = \lambda\sqrt{N}$ and G_N as a disjoint union of $N/\bar{d}_N$ complete graphs, each of size $\bar{d}_N$, then we prove (for $B > 0$):

$$\sqrt{N}(T_N - t) \xrightarrow{w} \mathcal{N}\left(\frac{\mu(\beta, B)}{\lambda}, \frac{1-t^2}{1-\beta(1-t^2)}\right).$$

$\mu(\beta, B) \neq 0$ and free of λ.

Some background

- The same problem of fluctuations was studied in Ellis and Newman [1978] when G_N is a complete graph. Our result strengthens it considerably.

Some background

- The same problem of fluctuations was studied in Ellis and Newman [1978] when G_N is a complete graph. Our result strengthens it considerably.
- In Löwe et al. [2018] and Berthet et al. [2016], the authors study fluctuations in "block-spin Ising models".

No universal behaviour for fluctuations?

Does this mean, at the fluctuation level, the entire asymptotic regime (from Basak and Mukherjee [2017]) cannot be reproduced?

No universal behaviour for fluctuations?

Does this mean, at the fluctuation level, the entire asymptotic regime (from Basak and Mukherjee [2017]) cannot be reproduced?

Under model (1),

$$\mathbb{E}[\sigma_i|\sigma_1,\ldots,\sigma_{i-1},\sigma_{i+1},\ldots,\sigma_N] = \tanh(\beta m_i + B)$$

where $m_i = \sum_j a_{ij}\sigma_j$ (average response from neighbours).

No universal behaviour for fluctuations?

Does this mean, at the fluctuation level, the entire asymptotic regime (from Basak and Mukherjee [2017]) cannot be reproduced?

Under model (1),

$$\mathbb{E}[\sigma_i|\sigma_1,\ldots,\sigma_{i-1},\sigma_{i+1},\ldots,\sigma_N] = \tanh(\beta m_i + B)$$

where $m_i = \sum_j a_{ij}\sigma_j$ (average response from neighbours).

Look at $S_n := \frac{1}{N}\sum_i(\sigma_i - \tanh(\beta m_i + B))$. For bounded $\bar{d}_N$, S_N was analysed in Comets and Janžura [1998].

Conditionally centered fluctuation

Theorem 2 (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \xrightarrow{N\to\infty} \infty$, then for any (β, B):

$$\sqrt{N}S_N = \frac{1}{\sqrt{N}}\sum_i (\sigma_i - \tanh(\beta m_i + B)) \xrightarrow{d} \mathcal{N}(0, \tau^2).$$

Here $\tau^2 = (1-t^2)(1-\beta(1-t^2))$. For the non-uniqueness regime, **(A2)** is required.

Conditionally centered fluctuation

Theorem 2 (universality)

Under model (1) (and **(A1)**), if G_N is a sequence of "approximately" regular graphs, with $\bar{d}_N \xrightarrow{N\to\infty} \infty$, then for any (β, B):

$$\sqrt{N} S_N = \frac{1}{\sqrt{N}} \sum_i (\sigma_i - \tanh(\beta m_i + B)) \xrightarrow{d} \mathcal{N}(0, \tau^2).$$

Here $\tau^2 = (1-t^2)(1-\beta(1-t^2))$. For the non-uniqueness regime, **(A2)** is required.

When $\beta = 1, B = 0$ however, $\tau = 0$. We expect in that case, $N^{3/4} S_n = O_p(1)$.

"Ising" on the cake?

For Theorem 1, we actually get Berry-Esseen equivalents for the weak-convergence, e.g., if $\beta < 1, B = 0$ or $B > 0$, then:

$$\sup_x \left|\mathbb{P}(\sqrt{N}(T_N - t) \leq x) - \mathbb{P}(\sigma Z \leq x)\right| \lesssim \sqrt{N}/\bar{d}_N.$$

Here σ^2 is the asymptotic variance from Theorem 1. Same rates hold for the other regimes of (β, B).

"Ising" on the cake?

For Theorem 1, we actually get Berry-Esseen equivalents for the weak-convergence, e.g., if $\beta < 1, B = 0$ or $B > 0$, then:

$$\sup_x |\mathbb{P}(\sqrt{N}(T_N - t) \leq x) - \mathbb{P}(\sigma Z \leq x)| \lesssim \sqrt{N}/\bar{d}_N.$$

Here σ^2 is the asymptotic variance from Theorem 1. Same rates hold for the other regimes of (β, B).

1. If $\bar{d}_N \approx N$ (dense graphs), then RHS$\lesssim N^{-1/2}$.
2. If $\bar{d}_N \approx \sqrt{N}$, then RHS$\lesssim O(1)$ (we know that the result is not true in this regime).

More implications

1. In terms of parameter estimation (β, B), we improve the existing results from Ghosal and Mukherjee [2018] and Bhattacharya and Mukherjee [2018] (e.g., Extending impossibility regimes).

More implications

1. In terms of parameter estimation (β, B), we improve the existing results from Ghosal and Mukherjee [2018] and Bhattacharya and Mukherjee [2018] (e.g., Extending impossibility regimes).
2. We prove that, marginally,

$$\sqrt{N}(\tilde{\beta} - \beta) \xrightarrow{d} \mathcal{N}\left(0, \frac{1 - \beta(1 - t^2)}{1 - t^2}\right)$$

$$\sqrt{N}(\tilde{B} - B) \xrightarrow{d} \mathcal{N}\left(0, \frac{1 - \beta(1 - t^2)}{1 - t^2}\right)$$

for certain regimes. $\tilde{\beta}$ represents pseudo-likelihood estimator (for B known), see Besag [1986].

Implications continued

Our results extend to graphs which are "approximately" regular (with counterexamples showing tightness for the extent of regularity). In addition to deterministic graphs, we also cover:

1. Random $d-$regular graphs.
2. $G_N :=$ Erdős-Rényi graph with parameters (N, p_N) if $p_N \gg (\log N)N^{-1}$.
3. $G_N :=$ stochastic block model with balanced, possibly growing, block sizes.
4. $W-$random graphons (where $W(\cdot)$ is regular).


The End

References I

O. Barndorff, Nielsen. *Information and exponential families: in statistical theory*. John Wiley & Sons, 2014.

A. Basak and S. Mukherjee. Universality of the mean-field for the potts model. *Probability Theory and Related Fields*, 168(3-4):557–600, 2017.

Q. Berthet, P. Rigollet, and P. Srivastava. Exact recovery in the ising blockmodel. *arXiv preprint arXiv:1612.03880*, 2016.

J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(3):259–279, 1986.

B. Bhattacharya and S. Mukherjee. Inference in ising models. *Bernoulli*, 24(1):493–525, 2018.

References II

F. Comets and M. Janžura. A central limit theorem for conditionally centred random fields with an application to markov fields. *Journal of applied probability*, 35(3): 608–621, 1998.

C. Daskalakis, N. Dikkala, and G. Kamath. Testing ising models. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1989–2007. Society for Industrial and Applied Mathematics, 2018.

R. S. Ellis and C. M. Newman. The statistics of curie-weiss models. *Journal of Statistical Physics*, 19(2):149–161, 1978.

References III

D. Fraiman, P. Balenzuela, J. Foss, and D. R. Chialvo. Ising-like dynamics in large-scale functional brain networks. *Physical Review E*, 79(6):061922, 2009.

N. G. Fytas, V. Martin-Mayor, G. Parisi, M. Picco, and N. Sourlas. Evidence for supersymmetry in the random-field ising model at d= 5. *arXiv preprint arXiv:1901.08473*, 2019.

P. Ghosal and S. Mukherjee. Joint estimation of parameters in ising model. *arXiv preprint arXiv:1801.06570*, 2018.

J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.

References IV

E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, Feb 1925. ISSN 0044-3328. doi: 10.1007/BF02980577. URL https://doi.org/10.1007/BF02980577.

M. Löwe, K. Schubert, et al. Fluctuations for block spin ising models. *Electronic Communications in Probability*, 23, 2018.

M. Marsman, D. Borsboom, J. Kruis, S. Epskamp, R. van Bork, L. Waldorp, H. v. d. Maas, and G. Maris. An introduction to network psychometrics: Relating ising network models to item response theory models. *Multivariate behavioral research*, 53(1):15–35, 2018.

References V

A. Martín del Campo, S. Cepeda, and C. Uhler. Exact goodness-of-fit testing for the ising model. *Scandinavian Journal of Statistics*, 44(2):285–306, 2017.

J. Raymond, S. Yarkoni, and E. Andriyash. Global warming: Temperature estimation in annealers. *Frontiers in ICT*, 3:23, 2016.

T. C. Schelling. Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186, 1971.

C. Siegel, J. Campos, and P. Toledo. Thermodynamics of a double-couple fault plane model by spin-lattice montecarlo simulations. In *AGU Fall Meeting Abstracts*, 2018.

References VI

D. Sun, E. B. Sudderth, and M. J. Black. Layered segmentation and optical flow estimation over time. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1768–1775. IEEE, 2012.

R. Yuste. From the neuron doctrine to neural networks. *Nature reviews neuroscience*, 16(8):487, 2015.

S. Zhang, Y. Yu, and H. Wang. Mittag-leffler stability of fractional-order hopfield neural networks. *Nonlinear Analysis: Hybrid Systems*, 16:104–121, 2015.