1ST CHAPTER ISING MODEL

N BINARY SPIN VARIABLES $(s_i)_{i=1,\dots,N}$

MAGNETIC FIELDS $(h_i)_i$

COUPLINGS MATRIX $(J_{ij})_{ij}$, $J_{ij} = J_{ji}$

HAMILTONIAN $H(s) = -\sum_i h_i s_i - \sum_{i<j} J_{ij} s_i s_j$

BOLTZMANN MEASURE $P(s) = \frac{1}{Z} \exp[-H(s)]$

PARTITION FUNCTION $Z(J,h) = \sum_s \exp[-H(s)]$

SPIN VARIABLE: σ_i. REALIZATION: s_i.

THERMAL AVERAGE $\langle Q(\sigma)\rangle = \sum_s P(s)\, Q(s)$.

"FIRST MOMENTS" $m_i = \langle \sigma_i \rangle$ [MAGNETISATIONS]

"SECOND MOMENTS" $\chi_{ij} = \langle \sigma_i \sigma_j \rangle$ [CORRELATIONS]

FACT: $\dfrac{\partial \log Z}{\partial h_i} = \langle \sigma_i \rangle$. $\dfrac{\partial \log Z}{\partial J_{ij}} = \langle \sigma_i \sigma_j \rangle$.

"HELMOLTZ FREE ENERGY" $F(J,h) = -\log Z(J,h)$

FACT: $F(J,h)$ IS CONCAVE AND $-\dfrac{\partial F}{\partial h_i}(J,h) = m_i$, $-\dfrac{\partial F}{\partial J_{ij}}(J,h) = \chi_{ij}$

FACT: "FUNCTIONAL MAXIMU ENTROPY PRINCIPLE" I

$$F(J, h) = \min_{q \in P_{bi}} \left\{ \langle H \rangle_q (J, h) - \langle \log q \rangle_q \right\}$$

$$= \min_{q \in P_{bi}} \left\{ \mathcal{U}[q] - \mathcal{S}[q] \right\}.$$

$$= \min_{q \in P_{bi}} \left\{ \mathcal{F}[q] \right\}$$

P_{bi} IS THE SET OF BINARY DISTRIBUTIONS ON SPINS.

WE CALL $\mathcal{F}[q]$ THE "FUNCTIONAL HELMHOLTZ FREE ENERGY"

2^{ND} CHAPTER INVERSE PROBLEM

[AT EQUILIBRIUM]

WE SEE N iid SAMPLES $D = \{\underline{s}^\mu\}$, $\mu = 1, \dots, M$

I MODEL THE REAL DISTRIBUTION WITH A BOLTZMANN,
I WANT TO INFER J, h.

[I WANT TO INFER A SPARSE J]

$$\langle Q \rangle^D = \frac{1}{M} \sum_\mu Q(\underline{s}^\mu)$$

WHY DO WE ASSUME A BOLTZMANN?

MAXIMU ENTROPY DISTRIBUTION

$$\max_{P \in P_{bi}} - \sum_{s} P(s_i) \log P(s_i)$$

SUCH THAT $\sum_{s} P(s_i) = 1$

$$\sum_{s} P(s) s_i = m_i^*$$

$$\sum_{s} P(s) s_i s_j = \chi_{ij}^*$$

IMPLIES THAT $P(s) = \frac{1}{Z} \exp\left[-\sum_{i} h_i s_i - \sum_{i<j} J_{ij} s_i s_j\right]$

WITH J, h THAT MAKE $m_i^* = \langle \sigma_i \rangle$

$\chi_{ij}^* = \langle \sigma_i \sigma_j \rangle$

$$\max_{P \in \mathcal{P}_{\sigma_i}} -\sum_{s} P(s_i) \log P(s_i)$$

$$= \max_{J,h} -\sum_{s} P(s) \left[+\sum_i h_i s_i + \sum_{i<j} J_{ij} s_i s_j - \log Z(J,h) \right]$$

$$= \max_{J,h} -\sum_i h_i m_i^{*} - \sum_{i<j} J_{ij} \chi_{ij}^{*} + \log Z(J,h)$$

MAX IS GIVEN WITH J, h SUCH THAT $m_i^{*} = \langle \sigma_i \rangle$

$\chi_{ij}^{*} = \langle \sigma_i \sigma_j \rangle$

FACT: THIS IS EQUIVALENT TO SEARCHING FOR THE MAX. LIKELIHOOD DISTRIBUTION AMONG BOLTZMANN FAMILY GIVEN SAMPLES

$$\mathcal{L}_D(J,h) = \frac{1}{M}\log P(D\,|\,J,h)$$

$$= \sum_{i<j} J_{ij}\frac{1}{M}\sum_{\mu} s_i^{\mu} s_j^{\mu} + \sum_i h_i \frac{1}{M}\sum_{\mu} s_i^{\mu} - \log Z(J,h)$$

$$= \sum_{i<j} J_{ij}\langle\sigma_i\sigma_j\rangle^D + \sum_i h_i\langle\sigma_i\rangle^D - \log Z(J,h).$$

FACT: THIS IS EQUIVALENT TO SEARCHING FOR THE DISTRIB. THAT MINIMIZES $D_{KL}(\cdot\,||\,P_D)$ AMONG BOLTZMANN FAMILY

CRUDE MAXIMUM LIKELIHOOD

$$\{J^{ML}, h^{ML}\} = \text{ARGMAX } \mathcal{L}_D(J,h)$$

FACT:

$$\frac{\partial \mathcal{L}_D}{\partial h_i}(J,h) = \langle \sigma_i \rangle^D - \langle \sigma_i \rangle^{J,h}$$

$$\frac{\partial \mathcal{L}_D}{\partial J_{ij}}(J,h) = \langle \sigma_i \sigma_j \rangle^D - \langle \sigma_i \sigma_j \rangle^{J,h}$$

SINCE $\mathcal{L}_D(h, J)$ IS CONCAVE $\rightsquigarrow$ GRADIENT ASCENT [BML].

$$\boxed{\begin{aligned} h_i^{n+1} &= h_i^n + \eta \frac{\partial \mathcal{L}_D}{\partial h_i}(J^n, h^n) \\ J_{ij}^{n+1} &= J_{ij}^n + \eta \frac{\partial \mathcal{L}_D}{\partial J_{ij}}(J^n, h^n) \end{aligned}}$$

BUT COMPUTING $\langle \sigma_i \rangle^n$, $\langle \sigma_i \sigma_j \rangle^n$ IS INFEASIBLE

MAXIMUM LIKELIHOOD + MCMC

APPROXIMATE $\langle \sigma_i \rangle^{\sim} \approx \frac{1}{S} \sum_{s=1}^{S} \sigma_i^{s}$, $\langle \sigma_i \sigma_j \rangle^{\sim} = \frac{1}{S} \sum_{s=1}^{S} \sigma_i^{s} \sigma_j^{s}$

GENERATED BY MCMC.

GLAUBER

METROPOLIS HASTINGS

WOLFF

SIMULATED ANNEALING

TEMPERING

LANGEVIN...

MAXIMUM LIKELIHOOD + MCMC + BAYES

SAME AS BEFORE, BUT ADDING A PRIOR ON J, h.

$\pi(J,h \mid D) \propto \pi(J,h)\, P(D \mid J,h)$.

INSTEAD OF MAXIMIZING $\mathcal{L}_D(J,h)$, I MAXIMIZE

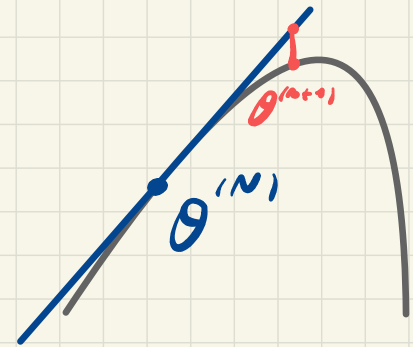
$$\mathcal{L}_D(J,h) + \log \pi(J,h)$$

[YOU CAN GET SPARSE REGULARIZATION:

$$\mathcal{L}_D^{L_1} = \mathcal{L}_D - \gamma \sum_{i<j} |J_{ij}| \qquad \mathcal{L}_D^{L_2} = \mathcal{L}_D - \gamma \sum_{i<j} J_{ij}^2 \;]$$

MAXIMUM LIKELIHOOD + NEWTON

IN CLASSIC GRADIENT ASCENT,



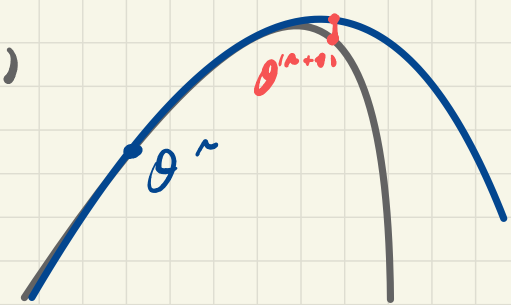
$$f(\theta^{(n)}+\delta\theta^{(n)}) \approx f(\theta^{(n)}) + \nabla f(\theta^{(n)})(\delta\theta^{(n)})$$

$$\delta\theta^{(n)} = \eta \nabla f(\theta^{(n)})$$

$$\theta^{(n+1)} = \theta^{(n)} + \delta\theta^{(n)}$$

IN NEWTON,

JACOBIAN.



$$\nabla f(\theta^{(n)}+\delta\theta^{(n)}) \approx \nabla f(\theta^{n}) + J_f(\theta^{(n)})(\delta\theta^{(n)})$$

$$\nabla f(\theta^{(n)}+\delta\theta^{(n)}) = 0,$$

$$\delta\theta^{(n)} = -[J_f(\theta^{(n)})]^{-1} \nabla f(\theta^{(n)})$$

$$\theta^{(n+1)} = \theta^{(n)} + \delta\theta^{(n)}$$

EFFICIENT ALGORITHM TO EVALUATE $[H_f(\theta^{(n)})]^{-1}$: BFGS

MAXIMUM LIKELIHOOD + MOMENTUM

IT HELPS GAINING VELOCITY IN THE FIRST STEPS OF BML.

CLASSIC MOMENTUM: $\delta\theta^{(n)} = \eta\nabla f(\theta^{(n)} + \beta\delta\theta^{(n-1)}) + \beta\delta\theta^{(n-1)}$

NESTEROV MOMENTUM: $\delta\theta^{(n)} = \eta\nabla f(\theta^{(n)} + \beta\delta\theta^{(n-1)}) + \beta\delta\theta^{(n-1)}$.

SCORE MATCHING

IN ABSTRACT: DATA COMES FROM $g(x)\,d\lambda(x)$ [UNKNOWN]

WE MODEL IT WITH $f(x,\theta)\,d\lambda(x)$ [FOR US BOLTZMANN]:

HEURISTICS: $\theta^{SM} = \underset{\theta}{\text{ARGMIN}}\ \mathbb{E}_g \left[\| \nabla_x \log g(x) - \nabla_x \log f(x,\theta) \|^2 \right]$

AFTER INTEGRATING BY PARTS:

$$\theta^{SM} = \underset{\theta}{\text{ARGMIN}}\ \mathbb{E}_g \left[(\nabla_x \log f(x,\theta))^2 + 2 \nabla_x^2 \log f(x,\theta) \right]$$

$$\approx \underset{\theta}{\text{ARGMIN}}\ \mathbb{E}_D \left[(\nabla_x \log f(x,\theta))^2 + 2 \nabla_x^2 \log f(x,\theta) \right]$$

$$\approx \underset{\theta}{\text{ARGMIN}}\ \mathbb{E}_D \left[(\nabla_x H)^2 - 2 \Delta_x H \right]$$ DOES NOT REQUIRE Z.

PROBLEM [IN MY OPINION]: ∇_x DOES NOT MAKE SENSE FOR BOLTZMANN.

WASSERSTEIN VIEW

FACT: IF $X_0 \sim \nu_0$ AND $\dot{X}_t = v_t(X_t)$, THEN $\nu_t = \mathrm{LAW}(X_t)$

SATISFIES $\partial_t \nu_t + \nabla \cdot (v_t \nu_t) = 0$ [IN WEAK SENSE]

GIVEN $\mathcal{F} : \mathcal{P}_2 \longrightarrow \mathbb{R}$ I CAN DEFINE $\nabla_{W_2} \mathcal{F}[\nu](x) : \mathbb{R}^d \longrightarrow \mathbb{R}$

$\nu \longmapsto \mathcal{F}[\nu]$ $\qquad$ $x \longmapsto \nabla \delta \mathcal{F}[\nu](x)$

IDEA: $v_t = -\nabla_{W_2} \mathcal{F}[\nu_t]$.

$\partial_t \nu_t - \nabla \cdot (\nabla_{W_2} \mathcal{F}[\nu_t] \nu_t) = 0$ IS THE "WGF" OF $\mathcal{F}$.

FACT: IF $\mathcal{F}$ IS (STRICTLY) CONVEX ALONG (WASSERSTEIN) GEODESICS, THEN ν_t GOES ESPONENTIALLY FAST AND MONOTONICALLY TOWARDS $\nu^* = \operatorname{ARGMIN}_{\nu} \mathcal{F}[\nu]$ POLYAK THEOREM.

FACT: IF $d\tilde{\pi}_\theta = f_\theta \, dx$, $f_\theta = \frac{1}{Z} e^{-H_\theta}$, $d\nu_g = g \, dx$

THEN. $\nabla_{W_2} D_{KL}(\nu_g \| \tilde{\pi}_\theta)(x) = \nabla_x H_\theta(x) + \nabla_x \log g(x)$

FACT: $D_{KL}(\cdot \| \tilde{\pi}_\theta)$ IS STRICTLY CONVEX ALONG WASSERSTEIN GEODESICS, PROVIDED H_θ IS STRICTLY CONVEX.

IDEA LOOK AT WGF OF $D_{KL}(\cdot \| \tilde{\pi}_\theta)$ WHERE $\tilde{\pi}_\theta$ IS BOLTZMANN, STARTING FROM THE REAL DISTRIBUTION ν_g

IF $\tilde{\pi}_\theta \simeq \nu_g$, THEN THE FLOW IS ALMOST STATIONARY SO THE (WASSERSTEIN) GRADIENT IS SMALL .

I WANT THE PARTICLES TO NOT MOVE MUCH

I WANT $E_{\mathcal{D}}[|-\nabla_x H_\theta - \nabla_x \log g|^2]$ TO BE SMALL

$$= E_{\mathcal{D}}[\|\nabla_x \log f(x,\theta) - \nabla_x \log g(x)\|^2].$$

MINIMUM PROBABILITY FLOW

SIMILAR HEURISTIC, FOR DISCRETE DOMAINS (LIKE ISING).

DATA COMES FROM $P^{(0)}$. $P_i^{(0)} = P^{(0)}(s_i)$

I WANT TO APPROXIMATE IT WITH BOLTZMANN $P_\theta^{(\infty)}$. $(P_\theta^{(\infty)})_i = P_\theta^{(\infty)}(s_i)$

I CREATE A DYNAMICS THAT CONVERGES TO $P_\theta^{(\infty)}$.

[I HAVE TO IMPOSE DETAILED BALANCE: $(\Gamma_\theta)_{ji} (P_\theta^{(\infty)})_i = (\Gamma_\theta)_{ij} (P_\theta^{(\infty)})_j$

WITH $(\Gamma_\theta)_{ij} = \mathbb{P}[j \to i]$]

I CHOOSE $(\Gamma_\theta)_{ij} = g_{ij} \, \mathrm{EXP}\left[\frac{1}{2}(H_\theta(s_j) - H_\theta(s_i))\right]$

$g_{ij} = 1$ IF AND ONLY IF s_i, s_j DIFFERS BY ONE FLIP.

GLAUBER

I FOLLOW THE DYNAMICS STARTING FROM $p^{(0)}$.

$$\dot{p}_i^{(n)} = \sum_{j\neq i} (\Gamma_\theta)_{ij}\, p_j^{(n)} - \sum_{j\neq i} (\Gamma_\theta)_{ji}\, p_i^{(n)}$$

IDEA: $p^{(0)} \approx p_\theta^{(\infty)} \Longrightarrow D_{KL}\left(p^{(0)} \| p_\theta^{(\varepsilon)}\right) \approx 0.$

$$D_{KL}\left(p^{(0)} \| p^{(\varepsilon)}\right) \approx D_{KL}\left(p^{(0)} \| p^{(t)}\right)\Big|_{t=0} + \varepsilon \left.\frac{\partial D_{KL}\left(p^{(0)} \| p_\theta^{(t)}\right)}{\partial t}\right|_{t=0}$$

$$= \varepsilon \sum_{i\notin D} \dot{p}_i^{(0)}$$

$$= \frac{\varepsilon}{M} \sum_{i\notin D} \sum_{j\in D} (\Gamma_\theta)_{ij}$$

$$= \frac{\varepsilon}{M} \sum_{j\in D} \sum_{i\notin D} g_{ij} \exp\left[\tfrac{1}{2}\left[H_\theta(s_j) - H_\theta(s_i)\right]\right]$$

MINIMIZE THIS AND FIND θ^{MPF}.

VARIATIONAL PRINCIPLES PRELUDE

THE LEGENDRE TRANSFORM OF $F(J,h)$ IS

$$S(\chi, m) = \min_{J,h} \left[-\sum_i h_i m_i - \sum_{i<j} J_{ij}\chi_{ij} - F(J,h) \right] \quad \text{CONCAVE}$$

FACT: $S(\chi, m)$ IS THE ENTROPY OF A BOLTZMANN WITH FIXED $\langle \sigma_i \rangle = m_i$, $\langle \sigma_i \sigma_j \rangle = \chi_{ij}$. [UP TO THE SIGN]

FACT: $-\frac{\partial S}{\partial m_i}(\chi, m) = h_i$, $\quad \frac{\partial^2 S}{\partial m_i \partial m_j}(\chi, m) = J_{ij}$ [J, h FROM CORRESP. BOLT.]

THE GIBBS FREE ENERGY IS

$$G(J, m) = \max_h \left[\sum_i h_i m_i + F(J, h) \right] \quad \text{CONCAVE}$$

$$\frac{\partial G}{\partial m_i}(J, m) = h_i \qquad \frac{\partial^2 G}{\partial m_i \partial m_j}(J, m) = (C^{-1})_{ij}$$
$$= \left[(\chi_{ij} - m_i m_j)^{-1} \right]_{ij}$$

FACT: "FUNCTIONAL MAXIMUM ENTROPY PRINCIPLE" II.

$$S(X, m) = \max_{q \in \underline{S}} \{ \mathcal{S}[q] \}$$

WITH $\underline{S} = P_{bi} \cap \{ q : \langle \sigma_i \rangle_q = m_i ,\ \langle \sigma_i \sigma_j \rangle_q = X_{ij} \}$

$$G(J, m) = \min_{q \in \underline{G}} \{ -\sum_{i<j} J_{ij} \langle \sigma_i \sigma_j \rangle_q - \mathcal{S}[q] \}$$

$$= \min_{q \in \underline{G}} \{ \mathcal{G}[q] \}$$

WITH $\underline{G} = P_{bi} \cap \{ q : \langle \sigma_i \rangle_q = m_i \}$.

AND $\mathcal{G}[q]$ IS SAID AS THE FUNCTIONAL GIBBS FREE ENERGY

VARIATIONAL PRINCIPLES + MEAN FIELD

IDEA [VERY GENERAL]

$$G(J,m) = \min_{q \in \underline{G}} \left\{ -\sum_{i<j} J_{ij} \langle \sigma_i \sigma_j \rangle_q - S[q] \right\}$$

$$\approx \min_{q \in \underline{G}^*} \left\{ -\sum_{i<j} J_{ij} \langle \sigma_i \sigma_j \rangle_q - S[q] \right\}$$

$\underline{G}^* = \underline{G} \cap [\dots]$

FIND A CLOSED FORM FOR $G(J,m)$ UNDER THIS RESTRICTION.

USE $\frac{\partial G}{\partial m_i}(J, m^D) = h_i$, $\frac{\partial^2 G}{\partial m_i \partial m_j}(J, m^D) = (C^{-1})^D_{ij}$

AND FIND J^{MF}, h^{MF}. [UNDERLYING PRINCIPLE: MAX. ENTROPY $\longleftrightarrow$ ML]

IN THIS CASE $[\dots] = \left\{ P^{MF} : \; P^{MF}(s_i) = \prod_i \frac{1+\tilde{m} s_i}{2} \right\}$

THIS FORCES $m_i = \tilde{m}_i$ AND SO $|\hat{\mathcal{G}}^{MF}| = |\mathcal{G} \cap [\dots]| = 1$.

$$\hat{G}(m,J) = \hat{G}^{MF}(m,J)$$

$$= -\sum_{i<j} J_{ij} m_i m_j + \sum_i \left[\frac{1+m_i}{2} \log \frac{1+m_i}{2} + \frac{1-m_i}{2} \log \frac{1-m_i}{2} \right]$$

WHICH ENDS UP IN:

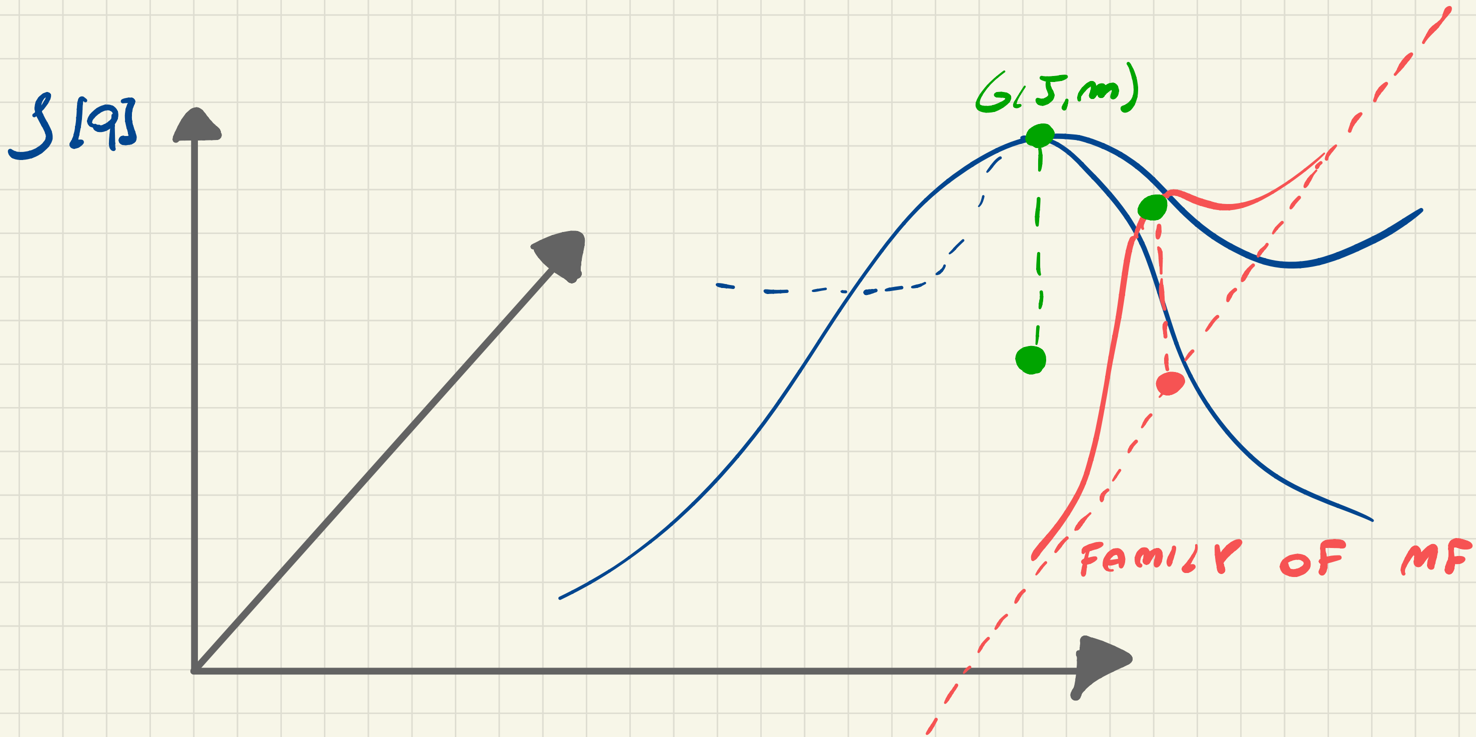
$$(C^{-1})_{ij}^{D} = -J_{ij}^{MF} \quad (i \neq j)$$

$$h_i^{MF} = -\sum_{j \neq i} J_{ij}^{MF} + \operatorname{ARCTANH} m_i^{D}$$

A BIT FORCED

[IN MY OPINION]

[YOU STILL NEED TO DO MATRIX INVERSION, POLYTIME].


S[q]
G(J,m)
FAMILY OF MF

VARIATIONAL PRINCIPLES + ONSAGER

SAME AS BEFORE BUT YOU ADD

$$G^{TAP}(J,m) = G^{MF}(J,m) - \frac{1}{2}\sum_{i<j} J_{ij}^2 (1-m_i^2)(1-m_j^2)$$

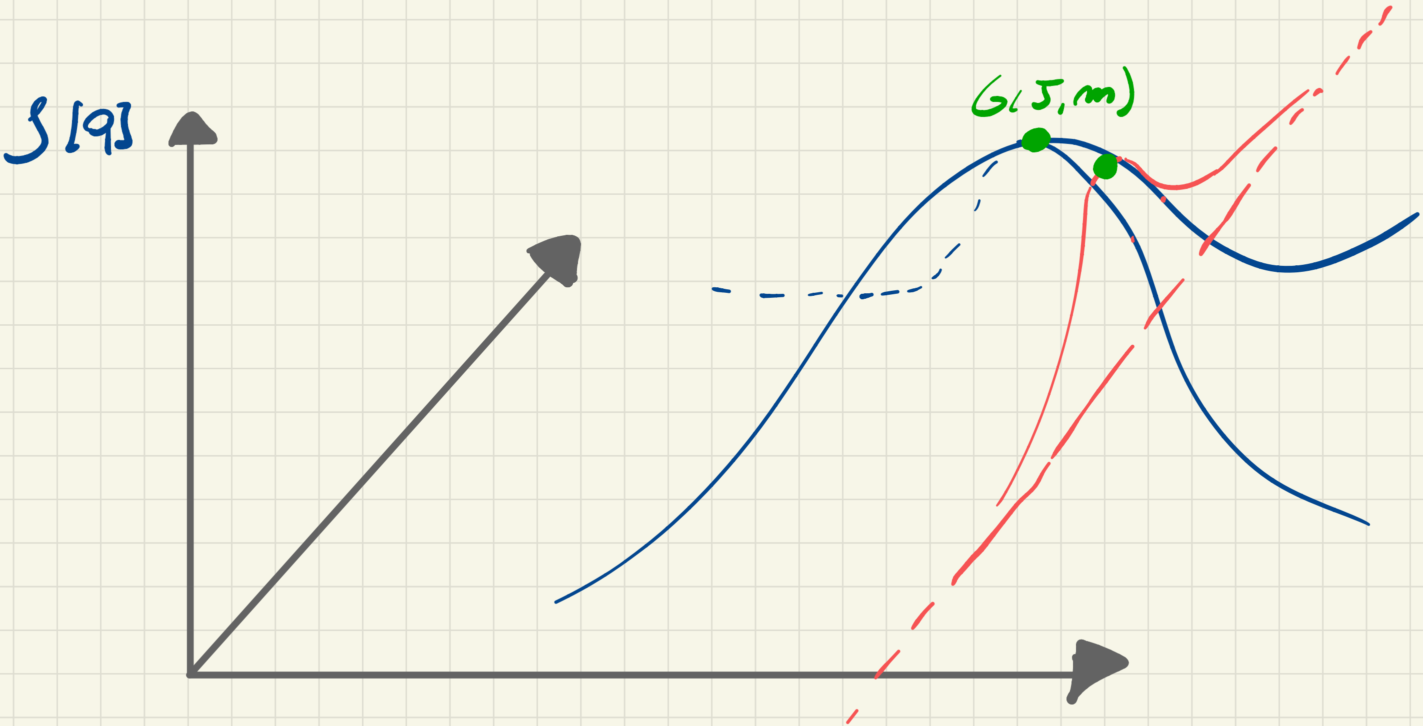
GIVING NOW

$$(C^{-1})^D_{ij} = -J^{TAP}_{ij} - 2\, J^{TAP}_{ij}\, m_i^D m_j^D \,.$$

$$h_i^{TAP} = \text{ARCTAN } m_i^D - \sum_{j\neq i} J^{TAP}_{ij} m_j^D + m_i^D \sum_{j\neq i} \left(J^{TAP}_{ij}\right)^2 (1-(m_j^D)^2).$$

WHICH ALSO DESCRIBES "FLUCTUATIONS"

[IT'S A BETTER APPROX. WHY? LATER].



S[g]
G(J,m)

TREE ISING

IDEA: IF J DESCRIBES A TREE [OR FOREST], i.e. IF THERE ARE NO LOOPS, THEN:

$$P_T = \prod_{i \in V_T} P_i(s_i) \prod_{(i,j) \in E_T} \frac{P_{ij}(s_i, s_j)}{P_i(s_i)\, P_j(s_j)}$$

$$= \prod_{(i,j) \in E_T} P_{ij}(s_i, s_j) \prod_{i \in V_T} \left(P_i(s_i)\right)^{1-|\partial_i|} \qquad |\partial_i| = \#\{ j \mid J_{ij} \neq 0 \}$$

FACT $D_{KL}(P^D \| P)$ IS MINIMIZED WHEN $P_i = P_i^D$, $P_{ij} = P_{ij}^D$

SO: $$\min_{P_i, P_{ij}} D_{KL}(P^D \| P) = -H^D + \sum_{i \in V^T} H_i - \sum_{(i,j) \in E_T} I_{ij}$$

$$I_{ij} = \sum_{s_i, s_j} P_{ij}^D(s_i, s_j) \log \frac{P_{ij}^D(s_i, s_j)}{P_i^D(s_i)\, P_j^D(s_j)}$$

IN D_{KL}, WE MINIMIZE OVER THE LAST TERM AMONG TREE TOPOLOGIES.

$$E_{T_{OPT}} = \underset{E_T}{\mathrm{ARGMIN}} \; -\sum_{(i,j) \in E_T} I_T.$$

SINCE $I_{i,j} \geq 0$, I WILL GET A CONNECTED GRAPH

I USE KRUSKAL \ PRIM FOR MST POLYTIME.

ONCE I HAVE T_{OPT} IT IS EASY TO GET $P_i, P_{i,j}$.

WILSON...

VARIATIONAL METHODS + BP

SAME STEPS OF MF, BUT EVEN LESS FORMAL.

DEFINE $\tilde{p}_i(s_i) = \dfrac{1+\tilde{m}_i s_i}{2}$

$$\tilde{p}_{ij}(s_i, s_j) = \frac{(1+\tilde{m}_i s_i)(1+\tilde{m}_j s_j) + \hat{c}_{ij} s_i s_j}{4}$$

WHERE: $-1 \leq \tilde{m}_i \leq 1$

$$-1 + |\tilde{m}_i + \tilde{m}_j| \leq \hat{c}_{ij} + \tilde{m}_i \tilde{m}_j \leq 1 - |\tilde{m}_i - \tilde{m}_j|$$

$$\left\{ p^{BP}(s_i) = \prod_i \tilde{p}_i(s_i) \prod_{(i,j) \in E} \frac{\tilde{p}_{ij}(s_i, s_j)}{\tilde{p}_i(s_i)\, \tilde{p}_j(s_j)} \right\} = BP$$

$$\underline{G}^* = BP \cap \{ q : \langle \sigma_i \rangle = m_i \} \longrightarrow$$ GET CLOSED FORM, TAKE DERIVATIVE

p^{BP} IS NORMALIZED ONLY ON TREES.

CAVITY METHODS

IT IS A TRICK TO CALCULATE $Z(J,h)$, $\langle\sigma_i\rangle$, $\langle\sigma_i\sigma_j\rangle$ EFFICIENTLY ON TREES

CAN BE EXTENDED TO GRAPHS. [SUBSCEPTIBILITY]

$Z_{i\to j}(s_i)$ IS THE PARTITION FUNCTION OF THE PART OF THE SYSTEM CONTAINING i WHERE (i,j) IS BROKEN. AND i IS CONSTRAINED TO BE s_i.

$$Z_{i\to j}(s_i) = e^{h_i s_i} \prod_{k\in\partial i\setminus j}\left[\sum_{s_k} Z_{k\to i}(s_k)\, e^{J_{ik} s_i s_k}\right]$$

FAST IF YOU START FROM LEAVES.

$$Z_i(s_i) = e^{h_i s_i} \prod_{k\in\partial i}\left[\sum_{s_k} Z_{k\to i}(s_k)\, e^{J_{ik} s_i s_k}\right]$$

PARTITION FUNCTION WITH i CONSTRAINED ON s_i.

YOU CAN COMPUTE $P_i(s_i)$ BY NORMALIZING $P_i(s_i) \propto Z_i(s_i)$

ADAPTIVE CLUSTER EXPANSION

APPROXIMATE THE ENTROPY $S(\chi, m)$ AND TAKE DERIV.

IF WE HAD A SINGLE SPIN, $S_{(1)}(m_i) = \sum_{s_i} \frac{1+m_i s_i}{2} \log \frac{1+m_i s_i}{2}$

IF WE HAD TWO SPINS,

$$S_{(2)}(m_i, m_j, \chi_{ij}) = \sum_{s_i, s_j} \frac{1+m_i s_i + m_j s_j + \chi_{ij} s_i s_j}{4} + \log \frac{1+m_i s_i + m_j s_j + \chi_{ij} s_i s_j}{4}$$

IF $i \perp j$, $S_{(2)} = S_{(1)}(m_i) + S_{(1)}(m_j)$

$$\Delta S_{(2)}(m_i, m_j, \chi_{ij}) := S_{(2)} - S_{(1)}(m_i) - S_{(1)}(m_j)$$

DEFINE ACCORDINGLY $\Delta S_{(1)} = S_{(1)}$

$\Delta S_{(3)} = [\text{TRIPLET} - \text{COUPLES} + \text{SINGLETONS}]$

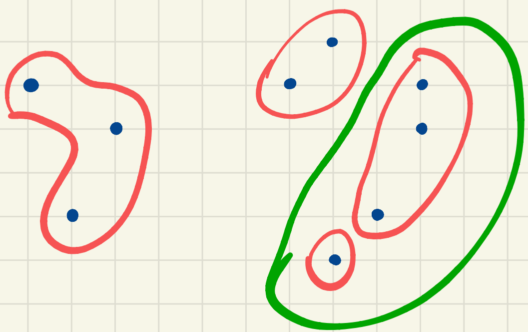
IDEA $S(X, m) \approx \sum_i \Delta S_{(1)}(m_i) + \sum_{(i,j)} \Delta S_{(2)}(m_i, m_j, X_{ij}) + \dots$

FIX A THRESHOLD δ.

ADD A CLUSTER IF ITS $\Delta S \geq \delta$.

STOP.

YOU GET $\tilde{S}$, TAKE DERIVATIVES, FIND $\tilde{J}$, $\tilde{h}$.



PLEFKA'S EXPANSION

SIMILAR TO PERTURBATION TRICKS

LET $H(s) = H_0(s) + \lambda V(s)$

$$H_0(s) = -\sum_i h_i s_i$$

$$V(s) = -\sum_{i<j} J_{ij} s_i s_j$$

$$F^{\lambda}(J,h) = F^{(0)}(h) - \lambda \langle V \rangle_0 + \frac{\lambda^2}{2}\left[\langle V^2 \rangle_0 - \langle V \rangle_0^2\right] + \cdots$$

$\hookrightarrow$ TAYLOR EXPANSION OF $F(\lambda)$ AROUND ZERO.

$$F^{(0)}(h) = \sum_i 2\cosh h_i$$

LEGENDRE TRANSFORM TO GET

$$G^{\lambda} = G^{(0)} + \lambda G^{(1)} + \cdots + \lambda^2 G^{(2)} \quad \left[\, G^{\lambda} = \sum_i h_i m_i + F^{\lambda}(J,h), \quad m_i = -\frac{\partial F^{\lambda}}{\partial h_i}(J,h)\right]$$

USE $h^{\lambda} = h^{(0)} + \lambda h^{(1)} + \dots + \lambda^2 h^{(2)} + \dots$

INTO $m_i = -\dfrac{\partial F^{\lambda}}{\partial h_i}(J, h)$ TO FIND $h^{(i)}$ AS A FUNCTION OF m, J.

$$\rightsquigarrow \quad G^{(0)}(J, m) = \sum_i \frac{1+m_i}{2} \log \frac{1+m_i}{2} + \frac{1-m_i}{2} \log \frac{1-m_i}{2}$$

$$G^{(1)}(J, m) = -\sum_{i<j} J_{ij} m_i m_j$$

$$G^{(2)}(J, m) = \sum_{i<j} J_{ij}^2 (1-m_i^2)(1-m_j^2)$$

IMPOSE OBSERVATIONS AND TAKE DERIVATIVES.

PSEUDO LIKELIHOOD I

$$H(\underline{s}) = H_i(s_i) + H_{\setminus i}(\underline{s}\setminus s_i)$$

$$= -h_i s_i - s_i \sum_{j\neq i} J_{ij} s_j + H_{\setminus i}(\underline{s}\setminus s_i)$$

$$Z(J,h) = \sum_{\underline{s}\setminus s_i} 2\cosh\Big[h_i + \sum_{j\neq i} J_{ij} s_j\Big] e^{-H_{\setminus i}(\underline{s}\setminus s_i)}$$

$$\langle \sigma_i \rangle = \Big\langle TH\Big[h_i + \sum_{k\neq i} J_{ik}\sigma_k\Big]\Big\rangle$$

$$\langle \sigma_i \sigma_j \rangle = \Big\langle \sigma_j\, TH\Big[h_i + \sum_{k\neq i} J_{ik}\sigma_k\Big]\Big\rangle$$

IDEA: IMPOSE

$$\langle \sigma_i \rangle^D = \Big\langle TH\Big[h_i^{PL} + \sum_{k\neq i} J_{ik}^{PL}\sigma_k\Big]\Big\rangle^D$$

$$\langle \sigma_i \sigma_j \rangle^D = \Big\langle \sigma_j\, TH\Big[h_i^{PL} + \sum_{k\neq i} J_{ik}^{PL}\sigma_k\Big]\Big\rangle^D$$

$\ast$ GRADIENT DESCENT

SOLVE FOR h_i^{PL}, J_{ij}^{PL}, J_{ji}^{PL}

$$J_{ij}^{PL} \longleftarrow \frac{1}{2}(J_{ij}^{PL} + J_{ji}^{PL}).$$

PSEUDOLIKELIHOOD II

NOTICE THAT $P(s_i \mid (s_j)_{j\neq i}) = \frac{1}{2}\left[1 + s_i \, TH\left[h_i + \sum_{j\neq i} J_{ij} s_j\right]\right]$

$$= \frac{1}{1 + EXP\left[h_i + \sum_{j\neq i} J_{ij} s_j\right]}$$

$$\mathcal{L}_D^i(J_{i*}, h_i) = \frac{1}{M}\sum_\mu \log P\left[s_i^\mu \mid (s_j^\mu)_{j\neq i}\right]$$

$$= \frac{1}{M}\sum_\mu \log \frac{1}{2}\left[1 + s_i^\mu \, TH\left[h_i + \sum_{j\neq i} J_{ij} s_j^\mu\right]\right]$$

SETTING DERIVATIVES TO ZERO RECOVERS *

BUT WE CAN WRITE $\mathcal{L}^{PL}(J, h) = \sum_i \mathcal{L}_D^i(J_{i*}, h_i)$

AND WE CAN MAXIMIZE OVER J, h.

IF WE FORCE J SYMMETRIC, HARD.

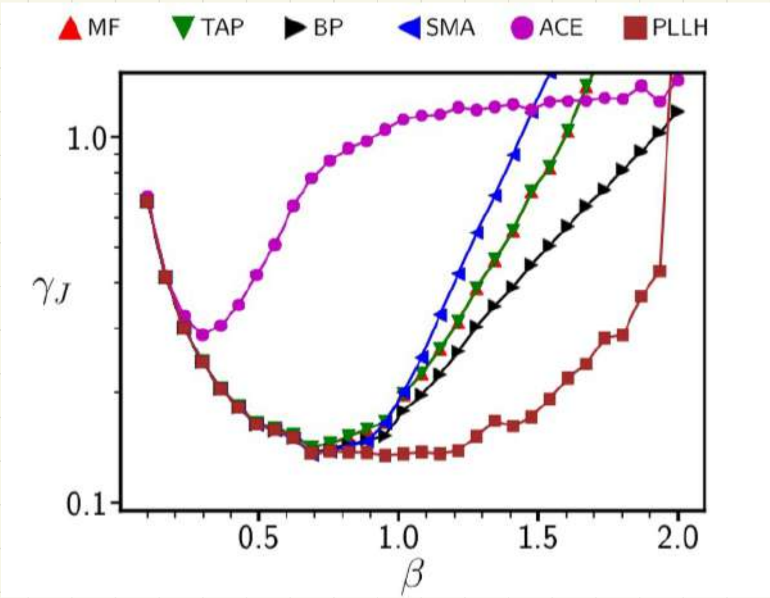
RESULTS

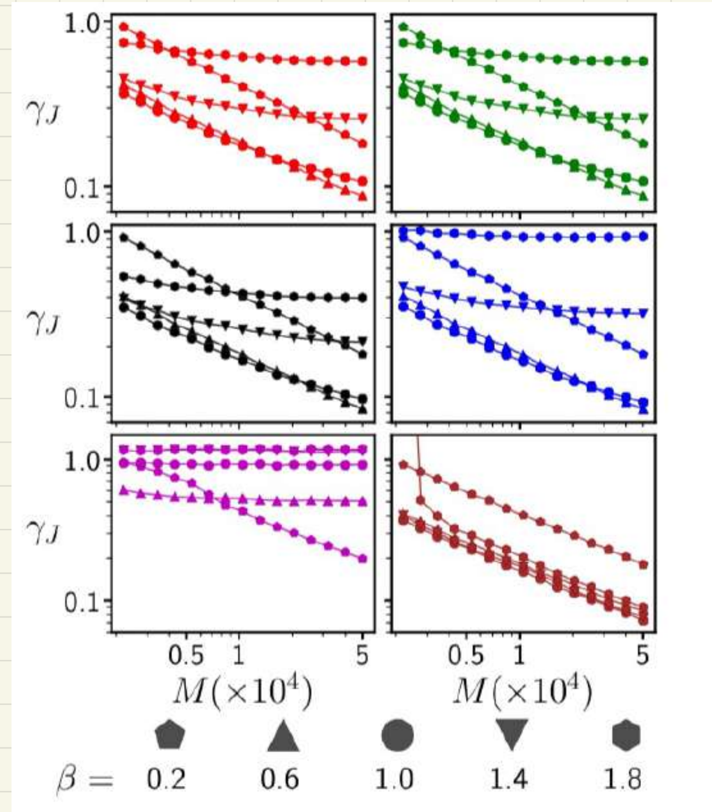
GROUND TRUTH

- FULLY CONNECTED GRAPH $\quad J^0_{ij} \sim \mathcal{N}(0, \beta/\sqrt{N})$ I
- LONG LOOPS [RANDOM GRAPH, FIXED DEGREE] $\quad J^0_{ij} \sim \mathrm{UNIF}[-\beta, \beta]$ II
- SHORT LOOPS [2D SQUARE LATTICE] $\quad J^0_{ij} \sim \mathrm{UNIF}[-\beta, \beta]$. III

$h^0_i \sim \mathrm{UNIF}[-0.3\beta, 0.3\beta]$

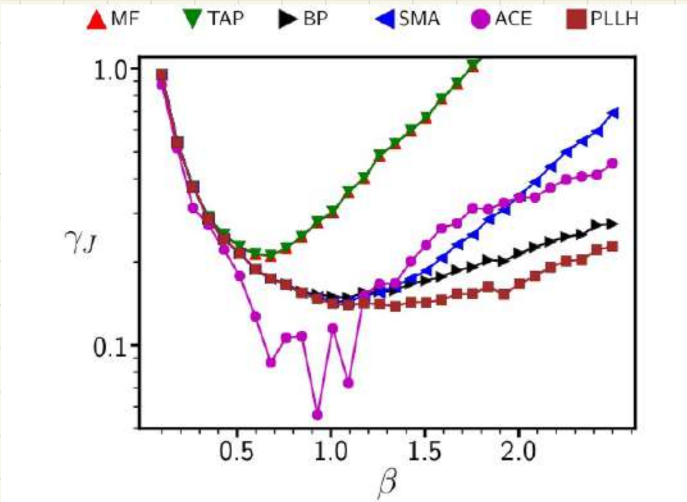
$\boxed{\text{I}}$



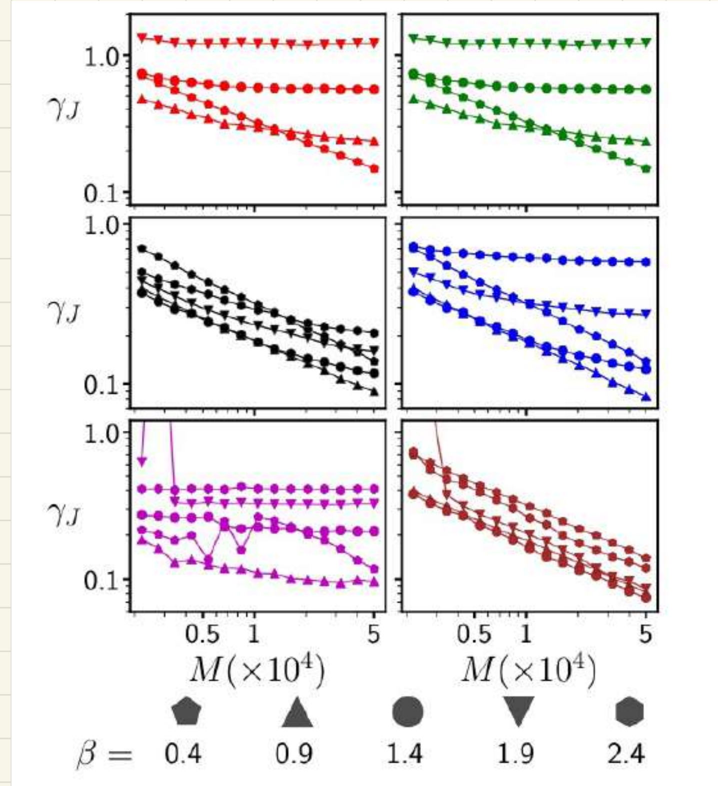


$$\gamma_J = \sqrt{\frac{\sum_{i<j} (J_{ij} - J^0_{ij})^2}{\sum_{i<j} (J^0_{ij})^2}}$$

II

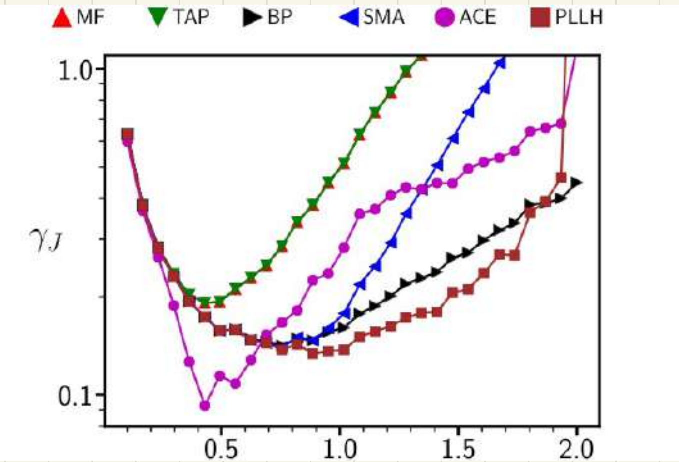


MF
TAP
BP
SMA
ACE
PLLH
1.0
0.1
γ_J
0.5
1.0
1.5
2.0
β

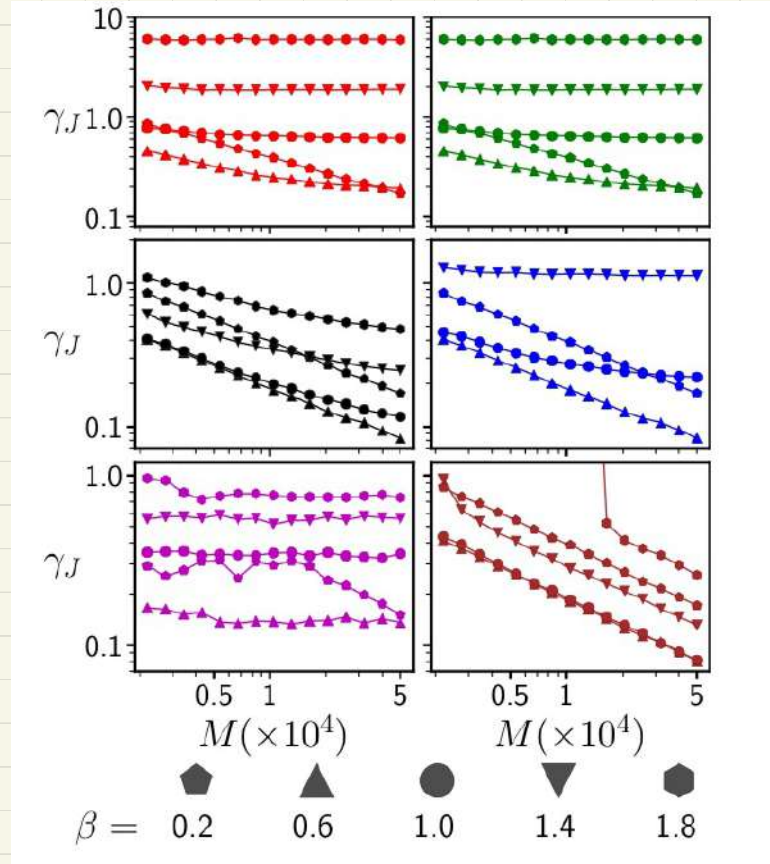


1.0
0.1
γ_J
0.5
1
5
$M(\times 10^4)$
$\beta =$
0.4
0.9
1.4
1.9
2.4

III



MF
TAP
BP
SMA
ACE
PLLH
1.0
0.1
γ_J
0.5
1.0
1.5
2.0



10
1.0
0.1
γ_J
0.5
1
5
$M(\times 10^4)$
$\beta =$
0.2
0.6
1.0
1.4
1.8

3RD CHAPTER

INVERSE PROBLEM [NOT EQUILIBRIUM]

IN A BOLTZMANN, $J_{ij} = J_{ji}$.

I CAN START A GLAUBER THAT SATISFIES DETAILED BALANCE $\Gamma_{ji} P_i = \Gamma_{ij} P_j$

CONVERGES TO EQUILIBRIUM $(\dot{P}_i)^{(t)} = 0$.

IF $J_{ij} \neq J_{ji}$, GLAUBER WILL LEAD TO A NESS:

$$(\dot{P}_i)^{(t)} = 0, \quad \Gamma_{ji} P_i \neq \Gamma_{ij} P_j$$

GLAUBER I

- PICK RANDOMLY i.
- $P(s_i(t+1) \mid s(t)) = \dfrac{\exp[s_i(t+1)\theta_i(t)]}{2\cosh \theta_i(t)}$, $\theta_i(t) = \sum_j J_{ij} s_j(t) + h_i$.

IF $J_{ij} = J_{ji}$ WE CONVERGE TO BOLTZMANN [THERE VELOCITY RESULTS].

IF $J_{ij} \neq J_{ji}$ WE CONVERGE TO A NESS. DIFFICULT TO CHARACTERIZE (NO CLOSED FORM.)

IF I PICK i, $m_i(t+1)_{s(t)} = \langle \sigma_i(t+1) \rangle_{s(t)} = \mathrm{TH}(\theta_i(t))$.

IF I DO NOT PICK i, $m_i(t+1)_{s(t)} = s_i(t)$.

IN NESS

$$m_i = \langle \sigma_i \rangle = \langle TH\ \theta_i \rangle = \langle TH(\sum_j J_{ij} \sigma_j + h_i) \rangle$$

$$\chi_{ij} = \langle \sigma_i \sigma_j \rangle = \frac{1}{2} \langle \sigma_i TH(\theta_j) \rangle + \frac{1}{2} \langle \sigma_j TH(\theta_i) \rangle$$

$$\phi_{ij} = \langle \sigma_i(t+1) \sigma_j(t) \rangle = \frac{1}{N} \langle TH(\theta_i) \sigma_j \rangle + \frac{N-1}{N} \chi_{ij}$$

GLAUBER II

- ALL SPIN VARIABLES CAN CHANGE AT $t+1$:
- $$P(\mathbf{s}(t+1)\,|\,\mathbf{s}(t)) = \frac{\exp\left[\sum_i s_i(t+1)\,\theta_i(t)\right]}{\prod_i 2\cosh\theta_i(t)}.$$

IN NESS

$$m_i = \langle \mathrm{TH}\,\theta_i \rangle$$
$$\chi_{ij} = \langle \mathrm{TH}\,\theta_i \rangle$$
$$\phi_{ij} = \langle \mathrm{TH}(\theta_i)\,\sigma_j \rangle$$

[NESS II $\neq$ NESS I.]

TIMES SERIES LIKELIHOOD

FOR GLAUBER II,

$$\mathcal{L}_D(J,h) = \frac{1}{M}\sum_{t=1}^{M-1}\sum_i \left\{ s_i(t+1)\,\theta_i(t) - \log 2\cosh\theta_i(t) \right\}$$

$$\hookrightarrow \frac{\partial \mathcal{L}_D}{\partial h_i}(J,h) = \frac{1}{M}\sum_{t=1}^{M-1}\left[s_i(t+1) - \mathrm{TH}\,\theta_i(t) \right]$$

$$\frac{\partial \mathcal{L}_D}{\partial J_{ij}}(J,h) = \frac{1}{M}\sum_{t=1}^{M-1}\left[s_i(t+1)\,s_j(t) - \mathrm{TH}\,\theta_i(t)\,s_j(t) \right]$$

CAN BE EVALUATE IN POLYNOMIAL TIME [GRADIENT ASCENT BML]

MEAN FIELD IN NESS

P IS A NESS OF A GLAUBER II.

LET SUPPOSE WE ARE ALREADY IN NESS P.

LET $q(\sigma) = \prod_i \frac{1 + m_i \sigma_i}{2}$ BE CLOSE TO P..

q IS THE NESS OF A DIFFERENT GLAUBER II:

$$h_i^q = \text{ARCTANH}\, m_i$$

$$J_{ij}^q = 0.$$

$$m_i = \langle \text{TH}(\sum_j J_{ij} \sigma_j + h_i) \rangle_P$$

IDEA: PERTURB THESE AND SEE HOW GLAUBER CHANGES

[BY TAKING DERIVATIVES OF $m_i = \langle TH(\sum_j J_{ij} \sigma_j + h_i) \rangle$]

$$\Delta m_i = \sum_j \frac{\partial \langle s_i \rangle}{\partial h_j}\bigg|_q \Delta h_j + \sum_{k,j} \frac{\partial \langle s_i \rangle}{\partial J_{kj}}\bigg|_q \Delta J_{kj} + \cdots$$

$$= (1-m_i^2)\Delta h_i + (1-m_i^2) \sum_j \Delta J_{ij} m_j + \cdots$$

AND IMPOSE $\Delta h_i = h_i - h_i^q$

$$\Delta J_{ij} = J_{ij} - J_{ij}^q = J_{ij}$$

IMPOSE $\Delta m_i = 0$ [WE WERE ALREADY CLOSE]

$$h_i^q = h_i + \sum_j J_{ij} m_j \rightsquigarrow m_i = TH(h_i + \sum_j J_{ij} m_j)$$

THEN, DO THE SAME PROCEDURE FOR ϕ_{ij}

$$\rightsquigarrow \phi_{ij} = m_i m_j + (1 - m_i^2) \sum_{l} J_{il} (\chi_{lj} - m_l m_j) \quad \circledast\circledast$$

IN $\circledast$, $\circledast\circledast$ IMPOSE m AND χ ARE THE SAMPLE AVERAGES ON NESS...

SOLVE: h^{MF}, J^{MF}.

RESULTS

GROUND TRUTH

$J_{ij}^{0} \sim N(0, \beta/\sqrt{N})$

$J_{ji}^{0} \sim N(0, \beta/\sqrt{N})$

$J_{ij}^{0} \perp J_{ji}^{0}$

$h_{i}^{0} \sim \text{UNIF}[-0.3\beta, 0.3\beta]$.