Boltzmann Machines on the Nishimori line

Francesco Camilli
University of Bologna

Bologna, 28/04/2021

Outline

Wigner Spiked model (WS)

We face the task of recovering the components of a binary ± 1 signal $(\sigma_i^*)_{i\in\Lambda}$, with $|\Lambda| = N$ set of indices, *iid* with probabilities $1/2$ through the observations

$$y_{ij}(\sigma^*) = \sqrt{\frac{\mu}{2N}}\sigma_i^*\sigma_j^* + z_{ij}\,, \quad \tilde{y}_i(\sigma^*) = \sqrt{h}\sigma_i^* + \tilde{z}_i\,, \quad z_{ij}\,, \tilde{z}_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$$

The problem is high dimensional since we have a large number of components N to recover. The Bayes posterior measure is

$$P(\sigma^* = \sigma | y, \tilde{y}) = \exp\left[-\frac{1}{2}\sum_{i,j\in\Lambda}\left(y_{ij} - \sqrt{\frac{\mu}{2N}}\sigma_i\sigma_j\right)^2 - \frac{1}{2}\sum_{i\in\Lambda}\left(\tilde{y}_i - \sqrt{h}\sigma_i\right)^2\right]\frac{1}{Z(y,\tilde{y})}$$
$$= \frac{\exp\left[-H_N(\sigma,\sigma^*,z)\right]}{Z(\sigma,\sigma^*,z)}$$

The posterior can thus be rewritten as a Gibbs random measure, the randomness being in the quenched variables σ^*, z. We can treat this problem as a disordered model.

The Hamiltonian is

$$H_N(\sigma,\sigma^*,z) = -\sum_{i,j\in\Lambda}\left[\sqrt{\frac{\mu}{2N}}z_{ij}\sigma_i\sigma_j + \frac{\mu}{2N}\sigma_i\sigma_j\sigma_i^*\sigma_j^*\right] - \sum_{i\in\Lambda}\left[\sqrt{h}\tilde{z}_i\sigma_i + h\sigma_i\sigma_i^*\right]$$

This Hamiltonian has a $\mathbb{Z}_2$-gauge symmetry:

$$z_{ij} \mapsto z_{ij}\sigma_i^*\sigma_j^*\,, \quad \sigma_i\sigma_j \mapsto \sigma_i\sigma_i^*\sigma_j\sigma_j^*$$

that allows us to get rid of the ground truth in the Hamiltonian:

$$\begin{aligned} H_N(\sigma,\sigma^*,z) &= -\sum_{i,j\in\Lambda}\left[\sqrt{\frac{\mu}{2N}}z_{ij}\sigma_i\sigma_j + \frac{\mu}{2N}\sigma_i\sigma_j\right] - \sum_{i\in\Lambda}\left[\sqrt{h}\tilde{z}_i\sigma_i + h\sigma_i\right] \\ &\overset{\mathrm{D}}{=} -\sum_{i,j\in\Lambda}J_{ij}\sigma_i\sigma_j - \sum_{i\in\Lambda}h_i\sigma_i\,, \quad J_{ij}\overset{\mathrm{iid}}{\sim}\mathcal{N}\left(\frac{\mu}{2N},\frac{\mu}{2N}\right), h_i\overset{\mathrm{iid}}{\sim}\mathcal{N}(h,h)\,. \end{aligned}$$

The inference is performed in the Baysian Optimal Setting, meaning that the prior distribution of the spins is known, and this gives rise to the Nishimori identities that ease computations.

Thermodynamic quantities

Let f be a function of 2 *replicas*, then we define the Boltzmann-Gibbs (BG) replicated average for a disordered model as

$$\langle f \rangle_N = \frac{\sum_{\sigma,\tau\in\Sigma_N} f(\sigma,\tau) e^{-H_N(\sigma)-H_N(\tau)}}{\left(\sum_\lambda e^{-H_N(\lambda)}\right)^2}$$

Notice that σ and τ are sampled independently w.r.t. to the BG measure, but the disorder in H_N is the same. The goal is generally the computation of the limiting average quenched pressure per particle

$$\lim_{N\to\infty} p_N = \lim_{N\to\infty} \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} e^{-H_N(\sigma)} \overset{\text{a.s.}}{=} \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log \sum_{\sigma\in\Sigma_N} e^{-H_N(\sigma)}$$

if it exists. The a.s. convergence is guaranteed by the fact that p_N is a Lipschitz function of standard Gaussian r.vs.

$-p_N$ is the mutual information $I(\sigma^*, (y, \tilde{y}))/N$ of the WS model up to a constant.

The Nishimori line (NL)

For a spin-glass with couplings $J_{ij} \overset{\text{iid}}{\sim} \mathcal{N}(\mu_{ij}, \Delta^2_{ij})$ and biases $h_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu_i, \Delta^2_i)$ the Nishimori line is the subregion of the phase space $(\mu_{ij}, \Delta^2_{ij}; \mu_i, \Delta^2_i)$ where

$$\mu_{ij} = \Delta^2_{ij}\,, \quad \mu_i = \Delta^2_i \quad \forall\, i, j \in \Lambda\,.$$

Therefore we have just seen that the WS model corresponds to a spin-glass model called SK model on the NL

$$H_N(\sigma) = -\sum_{i,j\in\Lambda} J_{ij}\sigma_i\sigma_j - \sum_{i\in\Lambda} h_i\sigma_i\,, \quad J_{ij} \overset{\text{iid}}{\sim} \mathcal{N}\left(\frac{\mu}{2N}, \frac{\mu}{2N}\right), h_i \overset{\text{iid}}{\sim} \mathcal{N}(h, h)\,.$$

As it happens in inference, a model on the Nishimori line (or in the Bayesian Optimal Setting) is greatly constrained and some identities on the correlation functions arise, for instance:

$$\mathbb{E}\langle\sigma_i\rangle^2_N = \mathbb{E}\langle\sigma_i\tau_i\rangle_N = \mathbb{E}\langle\sigma_i\rangle_N\,, \quad \mathbb{E}\langle\sigma_i\sigma_j\rangle^2_N = \mathbb{E}\langle\sigma_i\sigma_j\rangle_N\,.$$

Type I and II correlation inequalities on the NL

For a model on the NL we have:

Theorem (General case by Contucci, Morita, Nishimori '05)

$$\frac{\partial \mathbb{E} p_N}{\partial \mu_i} = \frac{1}{2}\left[1 + \mathbb{E}\langle \sigma_i \rangle_N\right] > 0$$

$$\frac{\partial^2 \mathbb{E} p_N}{\partial \mu_i \partial \mu_j} = \frac{1}{2}\frac{\partial}{\partial \mu_j}\mathbb{E}\langle \sigma_i \rangle_N = \mathbb{E}[(\langle \sigma_i \sigma_j \rangle_N - \langle \sigma_i \rangle_N \langle \sigma_j \rangle_N)^2] \geq 0\,.$$

Analogous results hold for J_{ij}-derivatives and mixed μ_i, J_{jk}-derivatives.
Fundamental observation:

$\mathbb{E}\langle \sigma_i \rangle_N$ is non decreasing in the parameters μ_j.

Correlation inequalities and the Nishimori identities together are sufficient to force replica symmetry in our model.

WS and SK compared

Both models can be solved comparing them with decoupled systems by means of an *adaptive interpolation*.

Theorem (J. Barbier, N. Macris '19)

Let the signal prior P_0 be bounded with second moment ρ. The mutual information for the spiked Wigner model verifies

$$\lim_{n\to+\infty}\frac{1}{n}I(\boldsymbol{x}^*;\boldsymbol{Y})=\inf_{q\in[0,\rho]}\Big\{\frac{\mu}{4}(q^2+\rho^2)+ \\ -\mathbb{E}\log\int dP_0(x)\exp\Big[\mu q x X^*+\sqrt{\mu q}Zx-\frac{\mu q}{2}x^2\Big]\Big\},\quad Z\sim\mathcal{N}(0,1)\,.$$

Theorem

The quenched pressure of the planted SK model in the thermodynamic limit is:

$$\lim_{N\to\infty}\mathbb{E}p_N=\sup_{q\in[0,1]}\Big\{\mu\frac{(1-q)^2}{4}-\mu\frac{q^2}{2}+\mathbb{E}\log 2\cosh\left(z\sqrt{\mu q}+\mu q\right)\Big\},\ z\sim\mathcal{N}(0,1)$$

Adding structure

The model defined by the Hamiltonian

$$H_N(\sigma) = -\sum_{i,j\in\Lambda} J_{ij}\sigma_i\sigma_j - \sum_{i\in\Lambda} h_i\sigma_i\,, \quad J_{ij} \overset{\text{iid}}{\sim} \mathcal{N}\left(\frac{\mu}{2N}, \frac{\mu}{2N}\right), h_i \overset{\text{iid}}{\sim} \mathcal{N}(h, h)$$

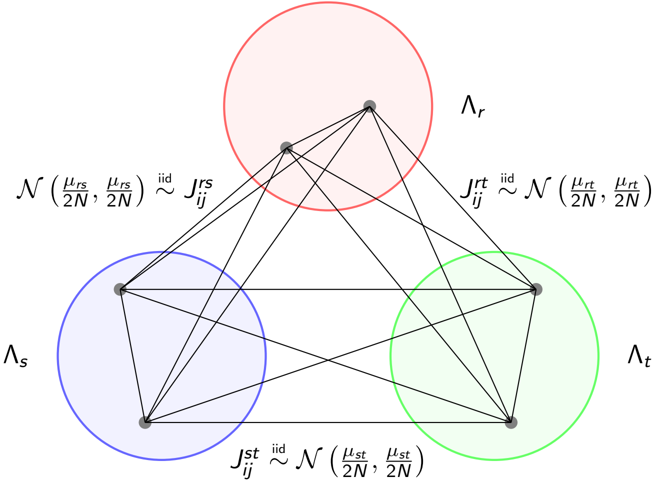
fulfills a complete permutation symmetry in distribution among the spins.
Let us partition Λ into K disjoint subsets: $\Lambda = \uplus_{r=1}^K \Lambda_r$, $|\Lambda_r| = N_r = \alpha_r N$. Now we can set the distribution of the J_{ij}, h_i according to which partition i, j belong to:

$$H_N(\sigma) := -\sum_{r,s=1}^K \sum_{(i,j)\in\Lambda_r\times\Lambda_s} J_{ij}^{rs}\sigma_i\sigma_j - \sum_{r=1}^K\sum_{i\in\Lambda_r} h_i^r\sigma_i,$$

$$J_{ij}^{rs} \overset{\text{iid}}{\sim} \mathcal{N}\left(\frac{\mu_{rs}}{2N}, \frac{\mu_{rs}}{2N}\right), \quad h_i^r \overset{\text{iid}}{\sim} \mathcal{N}(h_r, h_r)\,, \quad (i,j)\in\Lambda_r\times\Lambda_s$$

with $\mu_{rs} = \mu_{sr}$. The previous is the Hamiltonian of the so called multi-species SK model.

The permutation symmetry is now preserved only inside each partition!



This generalization has added structure to the model making it able to capture the thermodynamic ($N \to \infty$) behaviour of a system where different populations of particles interact.

The convex case: the replica symmetric formula

Surprisingly, this disordered model is replica symmetric.

Theorem (Alberici, C., Contucci, Mingione '20)

Let $\alpha = diag(\alpha_1, \dots, \alpha_K)$ *and* $\Delta := \alpha\mu\alpha$ *with* $\mu \geq 0$*. The thermodynamic limit of the pressure* $\bar{p}(\mu, h) := \lim_{N\to\infty} \bar{p}_N(\mu, h)$ *exists and:*

$$\bar{p}(\mu, h) = \sup_{\boldsymbol{x}\in\mathbb{R}^K_{\geq 0}} \bar{p}(\mu, h; \boldsymbol{x})$$

where

$$\bar{p}(\mu, h; \boldsymbol{x}) := \frac{(1-\boldsymbol{x}, \Delta(1-\boldsymbol{x}))}{4} - \frac{(\boldsymbol{x}, \Delta\boldsymbol{x})}{2} + \sum_{r=1}^{K} \alpha_r \psi((\mu\alpha\boldsymbol{x} + \boldsymbol{h})_r)$$

$$\psi(h) = \mathbb{E}\log\cosh(z\sqrt{h} + h)$$

with the following stationary condition:

$$\boldsymbol{x} - \mathbb{E}_z \tanh\left(z\sqrt{\mu\alpha\boldsymbol{x} + \boldsymbol{h}} + \mu\alpha\boldsymbol{x} + \boldsymbol{h}\right) \in Ker\Delta\,, \quad z \sim \mathcal{N}(0,1)$$

Phase transition

Proposition (Alberici, C., Contucci, Mingione '20)

Let $\mu > 0$. Denote by $\rho(A)$ the spectral radius of a matrix A and by $\mathscr{H}_{\boldsymbol{x}}\bar{p}$ the Hessian matrix of $\bar{p}$. The following implication holds:

$$\rho(\mu\alpha) = \rho(\alpha^{-1}\Delta) < 1 \quad \Rightarrow \quad \mathscr{H}_{\boldsymbol{x}}\bar{p}(\mu, h; \boldsymbol{x}) < 0,\ \forall \boldsymbol{x} \in \mathbb{R}^K_{\geq 0}$$

or equivalently $\bar{p}(\mu, h; \boldsymbol{x})$ is strictly concave w.r.t. $\boldsymbol{x}$.

Furthermore

$$\mathscr{H}_{\boldsymbol{x}}\bar{p}(\mu, 0; 0) = \frac{1}{2}\Delta^{1/2}\left[-\mathbb{1} + \Delta^{1/2}\alpha^{-1}\Delta^{1/2}\right]\Delta^{1/2}$$

therefore whenever $\boldsymbol{h} = 0$ and $\rho(\mu\alpha) < 1$, $\boldsymbol{x} = 0$ is the unique maximizer. $\boldsymbol{x} = 0$ becomes unstable as soon as $\rho(\mu\alpha) > 1$.

Critical exponents for the monopartite case

Proposition (Alberici, C., Contucci, Mingione '20)

Define:

$$\bar{p}_{var}(x;\mu,h)=\mu\frac{(1-x)^2}{4}-\frac{\mu x^2}{2}+\mathbb{E}\log 2\cosh\left(z\sqrt{\mu x}+\mu x\right)\ .$$

The following hold:

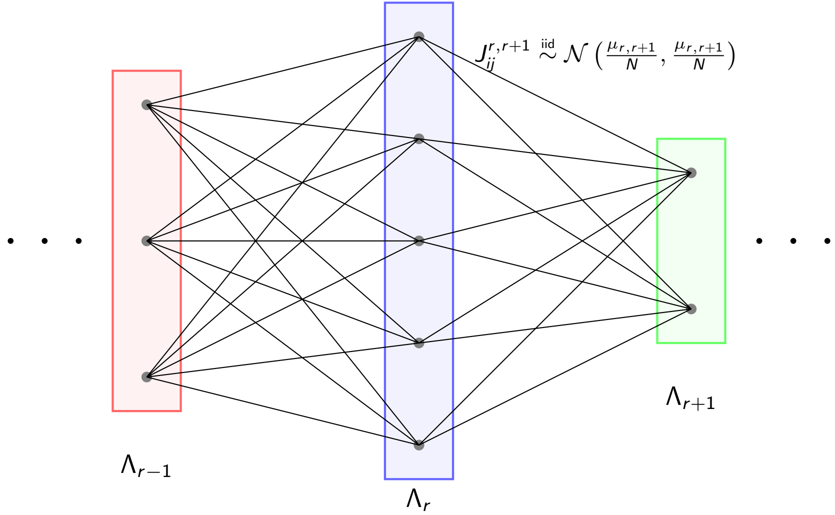
1. *if* $\mu<1$ *then* $\bar{p}_{var}$ *is concave in* x *and if further* $h=0$ *then* $x=0$ *the unique maximum point;*
2. *denoting the maximum point by* $\bar{x}(\mu,h)$*:* $\lim_{(\mu,h)\to(1,0)}\bar{x}(\mu,h)=0=\bar{x}(1,0)$*;*
3. *for fixed* $h=0$ *we have* $\bar{x}=(1+o(1))\frac{\mu-1}{\mu^2}$ *where* $o(1)$ *goes to* 0 *when* $\mu\to1_+$*. Therefore the critical exponent* $\beta=1$*.*
4. *For fixed* $\mu=1$ *and* $h\to0_+$ *the magnetization behaves as* $\bar{x}^2=h(1+o(1))$ *where* $o(1)\to0$ *when* $h\to0_+$*. Therefore we have a critical exponent* $\delta=2$*.*

Deep restricted Boltzmann Machine on the NL

In this model the partitions are rearranged in a consecutive way and only inter-partition interactions are allowed (restricted). It corresponds to the choice

$$\mu = \begin{pmatrix} 0 & \mu_{12} & 0 & \cdots & 0 \\ \mu_{21} & 0 & \mu_{23} & \cdots & 0 \\ 0 & \mu_{32} & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \mu_{K-1,K} \\ 0 & 0 & 0 & \mu_{K,K-1} & 0 \end{pmatrix} .$$

μ has eigenvalues with alternating sign (symmetric w.r.t. 0). For centered interactions the solution to this problem is unknown though we have some bounds.



Even the simplest $K = 2$ bipartite case for centered interactions is still unsolved.

Replica symmetric formula for the DBM

Theorem (Alberici, C, Contucci, Mingione '20)

Consider $J_{ij}^{rs} \overset{\text{iid}}{\sim} \mathcal{N}\left(\frac{\mu_{rs}}{2N}, \frac{\mu_{rs}}{2N}\right)$, $h_i^r \overset{\text{iid}}{\sim} \mathcal{N}(h_r, h_r)$ *for* $(i,j) \in \Lambda_r \times \Lambda_s$. *Denote by* $\boldsymbol{x}_o$ *the odd components of a vector* $\boldsymbol{x}$. *Similarly for* $\boldsymbol{x}_e$.

$$\lim_{N\to\infty} p_N \overset{\text{a.s.}}{=} \lim_{N\to\infty} \mathbb{E} p_N = \sup_{\boldsymbol{x}_o} \inf_{\boldsymbol{x}_e} \bar{p}(\boldsymbol{x}; \mu, \boldsymbol{h}) \,,$$

$$\bar{p}(\boldsymbol{x}; \mu, \boldsymbol{h}) := \sum_{r=1}^{K} \alpha_r \, \psi\left((\mu\alpha\boldsymbol{x})_r + h_r\right) + \sum_{r=1}^{K} \frac{\Delta_{r,r+1}}{2} \left[(1 - x_r)(1 - x_{r+1}) - 2x_r x_{r+1}\right]$$

where

$$\psi(x) := \mathbb{E}_z \log 2\cosh\left(z\sqrt{x} + x\right) \,, \quad z \sim \mathcal{N}(0,1) \,.$$

Phase transition

Denote by $A^{(oo)}$ the submatrix of A obtained by erasing its even rows and even columns.

Theorem

Let K be even and $\boldsymbol{h} = 0$. If $\rho([(\mu\alpha)^2]^{(oo)}) < 1$ then $\boldsymbol{x} = 0$ is the unique solution to the variational problem. Conversely, if $\rho([(\mu\alpha)^2]^{(oo)}) > 1$ then the solution is a vector $\boldsymbol{x} = \bar{\boldsymbol{x}}(\mu\alpha)$ with strictly positive components satisfying the consistency equation:

$$x_r = \mathbb{E}_z \tanh\left(z\sqrt{(\mu\alpha\boldsymbol{x})_r} + (\mu\alpha\boldsymbol{x})_r\right) \quad \forall\, r = 1, \ldots, K\,.$$

Geometry and phase transition

Proposition

The maximum of the spectral radius of $[(\mu\alpha)^2]^{(oo)}$ *over* $\alpha_1, \dots, \alpha_K \geq 0$, $\sum_r \alpha_r = 1$, *equals* $\frac{1}{4}\max_r \mu^2_{r,r+1}$ *and is reached if and only if:*

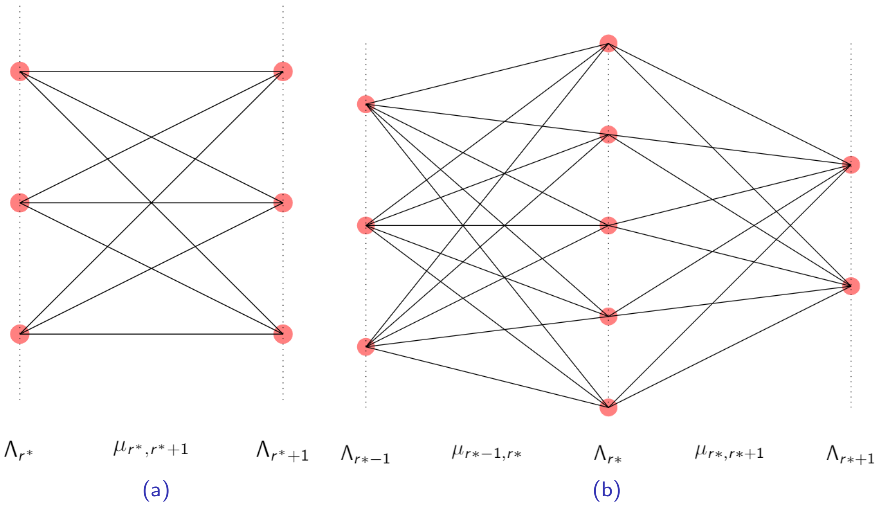
$$(a) \quad \alpha_{r^*} = \alpha_{r^*+1} = \frac{1}{2} \qquad \text{for } r^* \in \arg\max_r \Delta_{r,r+1}, \text{ or:}$$

$$(b) \quad \alpha_{r^*-1} + \alpha_{r^*+1} = \alpha_{r^*} = \frac{1}{2} \qquad \text{for } r^*,\ r^*-1 \in \arg\max_r \Delta_{r,r+1}.$$

Remark

$$\mu_{r,r+1} < 2\ \forall r = 1, \dots, K-1 \quad \Rightarrow \quad \bar{\boldsymbol{x}}(\mu\alpha) = \lim_{N\to\infty} \mathbb{E}\langle \boldsymbol{m} \rangle_N = \lim_{N\to\infty} \mathbb{E}\langle \boldsymbol{q}_{12} \rangle_N = 0$$

since there is no way to rearrange the spins into the partitions (α_r's) to obtain $\rho([(\mu\alpha)^2]^{(oo)}) > 1$.


Λ_{r^*}
μ_{r^*,r^*+1}
Λ_{r^*+1}
(a)
Λ_{r*-1}
$\mu_{r*-1,r*}$
Λ_{r*}
$\mu_{r*,r*+1}$
Λ_{r*+1}
(b)