On the Thermodynamic Equivalence between Hopfield Networks and Hybrid Boltzmann Machines

Enrica Santucci

On the equivalence of Hopfield Networks and Boltzmann Machines

(A. Barra, A. Bernacchia, E. Santucci, P. Contucci, Neural Networks 34 (2012) 1-9)

UNIVERSITY OF CAGLIARI
PraLab - Department of Electrical and Electronic Engineering

4 novembre 2016

Parte I

Description of the models

Spin glass: Sherrington Kirkpartick (SK) - 1975

Spin system whose low temperature state appears as a disordered one rather than the uniform or periodic structure pattern that one use to find in conventional Ising magnets

K. H. Fischer, J. A. Hertz (1991) - **M. Mezard, G. Parisi, M. Virasoro** (1987)

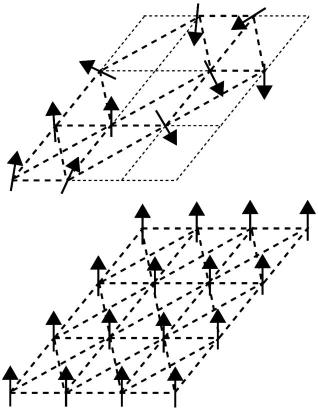


Figure 1: Schematic representation of a spin glass structure versus a ferromagnet one

SK Hamiltonian

- N particles (where N is very large)
- $\sigma_i \in \{-1, +1\}$ Ising spin related to the i-th particle ($i = 1, \ldots, N$)
- $J_{ij} \sim N(0, 1)$ interaction matrix between the lattice particles
- T temperature of the system ($\beta = 1/T$)

Hamiltonian

$$H_{sg}(\sigma, J) = -\frac{\beta}{\sqrt{N}} \sum_{1 \leq i,j \leq N} J_{ij}\sigma_i\sigma_j \tag{1}$$

- *frustration*: we cannot simultaneously minimize all the Hamiltonian terms because the interactions J_{ij} are random variables

SK Phase Diagram

- Partition Function $Z_N(\beta) = \sum_\sigma \exp(-H_N(\sigma, J))$
- Average over the interactions $\mathbb{E}(F(J)) = \int d\mu(J) F(J)$

Free Energy

$$f_N(\beta) = -\frac{1}{\beta N} \mathbb{E} \ln Z_N(\beta)$$

- Order Parameters $m = \frac{1}{N} \sum_{i=1}^N \sigma_i \quad q_{ab} = \frac{1}{N} \sum_{i=1}^N \sigma_i^{(a)} \sigma_i^{(b)}$
- Free Energy for $N \to \infty$
- Minimization of the free energy with respect to the order parameters
- Self-consistence equations

Gaussian spin glass (A. Barra, G. Genovese, F. Guerra - 2012)

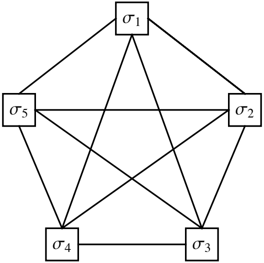
- $z_i,\ i = 1, \ldots, N \ \sim \ N(0,1)$
- $J_{ij},\ i,j = 1, \ldots, N \ \sim \ N(0,1)$

Hamiltonian

$$H_N(z,J) = -\frac{\beta}{\sqrt{N}} \sum_{1 \le i < j \le N} J_{ij} z_i z_j \tag{2}$$

- Order Parameter $\quad q_{ab} = \frac{1}{N} \sum_{i=1}^{N} z_i^{(a)} z_i^{(b)}$
- Free Energy for $N \ \rightarrow \ \infty$
- Minimization of the free energy with respect to the order parameters
- Self-consistence equation

Hopfield Model (HM) - 1982



- Stored patterns: $\xi^\mu = (\xi_1^\mu, \ldots, \xi_N^\mu), \quad \mu = 1, \ldots, P \quad \xi_i^\mu \in \{-1, +1\}$
- Digital units (activation levels): $\sigma = (\sigma_1, \ldots, \sigma_N), \ \ \sigma_i \in \{-1, +1\}$
- Activation function: Sign function
- Two-way information flow
- Symmetric synapses ($J_{ij} = J_{ji}$)

HM Hamiltonian

Hamiltonian

$$H_{hop}(\sigma, J) = -\frac{\beta}{N} \sum_{1 \leq i,j \leq N} J_{ij} \sigma_i \sigma_j \tag{3}$$

Hebbian learning rule

$$J_{ij} = \sum_{\mu=1}^{P} \xi_i^{\mu} \xi_j^{\mu} \quad \forall i.j = 1, \ldots, N$$

- Order parameters $m^{\mu} = \frac{1}{N} \sum_{i=1}^{N} \xi_i^{\mu} \sigma_i \quad q_{ab} = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^{(a)} \sigma_i^{(b)}$
- Free energy for $N \rightarrow \infty$
- Minimization of the free energy with respect to the order parameters
- Self-consistence equations

Analogy between Sherrington Kirkpatrick and Hopfield models

- N particles $\longleftrightarrow$ neurons
- σ_i Ising spin $\longleftrightarrow$ neuronal activation level
- J_{ij} spin interactions $\longleftrightarrow$ synapses
- T temperature $\longleftrightarrow$ noise level

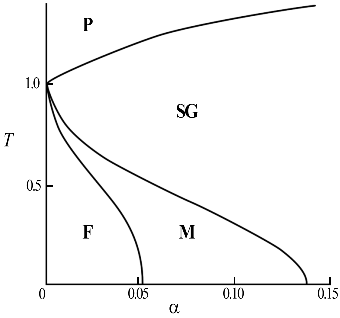
$$P \to \infty$$

$$\Downarrow$$

$$\text{Sherrington-Kirkpartick} \iff \text{Hopfield}$$

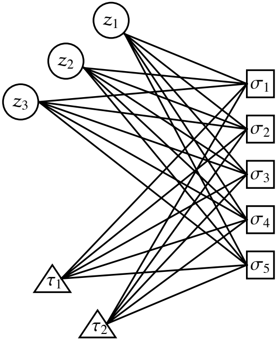
HM Phase Diagram

- $\alpha = \lim_{N \to \infty} \frac{P}{N}$ control parameter (*high storage regime*)
- T temperature



- Retrieval Phase **F** $(0 < \alpha \leq 0.05)$
- Mixed phase **M** $(0.05 < \alpha \leq 0.14)$
- Spin glass phase **SG** $(\alpha > 0.14)$
- Paramagnetic phase **P**

Boltzmann Machine (G. E. Hinton, T. J. Sejnowski - 1983)



- Digital visible layer: $\sigma_i \in \{+1, -1\}$ $(i = 1, \dots, N)$
- Two analog hidden layers: $z_\mu, \tau_\nu \sim N(0,1)$ $\mu = 1, \dots, P$ $\nu = 1, \dots, K$
- Activation function: sigmoidal function
- Two-way information flow
- Symmetric synaptic weights ξ_i^μ η_i^ν

Restricted and Hybrid version of the Boltzmann Machine (RHBM)

Assumptions

- *hybrid*: one digital layer of visible units and two analog layers of hidden units
- *restricted*: no connections between the hidden layers

Hamiltonian

$$H_{rhbm}(\beta,\sigma,z,\tau;\xi,\eta) = \frac{1}{2}\sum_{\mu=1}^{P} z_\mu^2 + \frac{1}{2}\sum_{\nu=1}^{K}\tau_\nu^2 - \sqrt{\frac{\beta}{N}}\Big(\sum_{i,\mu=1}^{N,P}\sigma_i\xi_i^\mu z_\mu + \sum_{i,\nu=1}^{N,K}\sigma_i\eta_i^\nu\tau_\nu\Big) \quad (4)$$

Parte II

Results

Dynamics of the hidden layers

Ornstein-Uhlembeck Diffusion Process

$$D\frac{dz_\mu}{dt} = -z_\mu(t) + \sum_{i=1}^{N} \xi_i^\mu \sigma_i + \sqrt{\frac{2D}{\beta}}\,\zeta_\mu(t)$$

$$D^*\frac{d\tau_\nu}{dt} = -\tau_\nu(t) + \sum_{i=1}^{N} \eta_i^\nu \sigma_i + \sqrt{\frac{2D^*}{\beta}}\,\rho_\nu(t)$$

- ζ, ρ white Gaussian noises
- D, D^* quantifiers of the timescale of the dynamics
- β measure of the strength of the fluctuations

Probability distribution of the hidden variables

$$Pr(z_\mu|\sigma) = \sqrt{\frac{\beta}{2\pi}} \exp\left[-\frac{\beta}{2}\left(z_\mu - \sum_{i=1}^{N} \xi_i^\mu \sigma_i\right)^2\right]$$

$$Pr(\tau_\nu|\sigma) = \sqrt{\frac{\beta}{2\pi}} \exp\left[-\frac{\beta}{2}\left(\tau_\nu - \sum_{i=1}^{N} \eta_i^\nu \sigma_i\right)^2\right]$$

for $\mu = 1, \ldots, P$ and $\nu = 1, \ldots, K$

Dynamics of the visible layer

$$\sigma_i(t+1) = \text{sign}\Big[\sum_{i=1}^{N}\Big(\sum_{\mu=1}^{P}\xi_\mu^i\sigma_i(t) + \sum_{\nu=1}^{K}\eta_\nu^i\sigma_i(t)\Big) - T_i\Big]$$

- t discrete time unit
- T_i threshold potential

Probability distribution of the visible units (Glauber dynamics)

$$Pr(\sigma_i|z) = \frac{\exp[\beta\sigma_i\sum_{\mu=1}^{P}\xi_i^\mu z_\mu]}{\exp[\beta\sum_{\mu=1}^{P}\xi_i^\mu z_\mu] + \exp[-\beta\sum_{\mu=1}^{P}\xi_i^\mu z_\mu]}$$

$$Pr(\sigma_i|\tau) = \frac{\exp[\beta\sigma_i\sum_{\nu=1}^{K}\eta_i^\nu \tau_\nu]}{\exp[\beta\sum_{\nu=1}^{K}\eta_i^\nu \tau_\nu] + \exp[-\beta\sum_{\nu=1}^{K}\eta_i^\nu \tau_\nu]}$$

$$Pr(z|\sigma) = \prod_{\mu=1}^{P} Pr(z_\mu|\sigma) \qquad Pr(\tau|\sigma) = \prod_{\nu=1}^{K} Pr(\tau_\nu|\sigma)$$

$$Pr(\sigma|z) = \prod_{i=1}^{N} Pr(\sigma_i|z) \qquad Pr(\sigma|\tau) = \prod_{i=1}^{N} Pr(\sigma_i|\tau)$$

Statistical equivalence between Hopfield network and Boltzmann machine

$$Pr(\sigma, z, \tau) \propto \exp\left[-H_{rhbm}(\sigma, z, \tau)\right]$$

$$\Downarrow$$

$$Pr(\sigma) \propto \exp\left[\frac{\beta}{2N}\sum_{i,j=1}^{N}\left(\sum_{\mu=1}^{P}\xi_i^{\mu}\xi_j^{\mu} + \sum_{\nu=1}^{K}\eta_i^{\nu}\eta_j^{\nu}\right)\sigma_i\sigma_j\right] = \exp\left[-H_{hop}(\sigma)\right]$$

- Thermodynamics of the visible units in a RHBM is equivalent to the one of a Hopfield network
- The dynamics of a Hopfield network, requiring the update of N neurons and the storage of N^2 synapses, can be simulated by a RHBM, requiring the update of $N + P$ neurons but the storage of only NP synapses

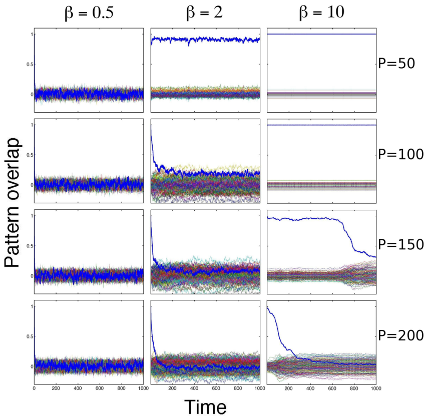
Counterpart of the HM Phase Diagram in a RHBM

- N number of neurons $\longleftrightarrow$ number of visible units
- P, K number of stored patterns $\longleftrightarrow$ number of hidden units
- ξ, η stored patterns $\longleftrightarrow$ synaptic weights

Hopfield model $\Longleftrightarrow$ *Boltzmann Machine*

- Retrieval Phase $\longleftrightarrow$ Few hidden units
- Spin Glass Phase $\longleftrightarrow$ Too many hidden units

Numerical simulations of the RHBM with a single hidden layer for different values of the parameters β $(= 1/T)$ and P



- $\beta = 0.5$ (high T) no retrieval is possible regardless of the number of hidden units P
- $\beta = 2$ (intermediate T) retrieval is possible provided that the number of hidden units is not too large
- $\beta = 10$ (low T) retrieval is maintained up to large values of P

Noise Source (I): Connection between the hidden layers

$$\tilde{H}_{rhbm}(\sigma, z, \tau; \xi, \eta) = \frac{1}{2}\sum_{\mu=1}^{P} z_\mu^2 + \frac{1}{2}\sum_{\nu=1}^{K}\tau_\nu^2 - \sqrt{\frac{\beta}{N}}\Big(\sum_{i,\mu}^{N,P}\xi_i^\mu \sigma_i z_\mu + \sum_{i,\nu}^{N,K}\eta_i^\nu \sigma_i \tau_\nu + \epsilon \sum_{\mu,\nu}^{P,K}\zeta_\mu^\nu z_\mu \tau_\nu\Big)$$

$$\Downarrow$$

Integration in z_μ e τ_ν

$$\Downarrow$$

$$\tilde{H}_{hop}(\sigma; \xi, \eta) = -\frac{\beta}{2N}\sum_{i,j=1}^{N}\Big[\sum_{\mu}^{\alpha N}\xi_i^\mu \xi_j^\mu \Big(1 - \epsilon\frac{\beta\gamma}{4}\Big) + \sum_{\nu}^{\gamma N}\eta_i^\nu \eta_j^\nu \Big(1 - \epsilon\frac{\beta\alpha}{4}\Big)\Big]\sigma_i\sigma_j$$

Noise Source (II): System subjected to an external field

$$z_\mu, \tau_\nu \;\sim\; \mathcal{N}(0,1) \quad\longrightarrow\quad z_\mu \;\sim\; \mathcal{N}(z_0,1),\;\; \tau_\nu \;\sim\; \mathcal{N}(\tau_0,1)$$

$$\tilde{H}_{rhbm}(\sigma,z,\tau;\xi,\eta) = \frac{1}{2}\sum_{\mu=1}^{P}(z_\mu - z_0)^2 + \frac{1}{2}\sum_{\nu=1}^{K}(\tau_\nu - \tau_0)^2 - \sqrt{\frac{\beta}{N}}\Big(\sum_{i,\mu}^{N,P}\xi_i^\mu \sigma_i z_\mu + \sum_{i,\nu}^{N,K}\eta_i^\nu \sigma_i \tau_\nu\Big)$$

$$\Downarrow$$

$$\tilde{H}_{hop}(\sigma) \quad\longrightarrow\quad H_{hop}(\sigma) + \sqrt{\beta} z_0 \sum_{i=1}^{N}\chi_i\sigma_i + \sqrt{\beta}\tau_0 \sum_{i=1}^{N}\psi_i\sigma_i$$

- $\chi_i = \frac{1}{\sqrt{P}}\sum_{\mu=1}^{P}\xi_i^\mu, \qquad \psi_i = \frac{1}{\sqrt{K}}\sum_{\nu=1}^{K}\eta_i^\nu$ external random fields