Stochastic Ising model with plastic interactions

Eugene Pechersky[a,b], Guillem Via[a], Anatoly Yambartsev[a]

[a]Institute of Mathematics and Statistics, University of São Paulo, Brazil.
[b]Institute for Information Transmission Problem, Russian Academy of Science, Russia.

2nd Workshop NeuroMat, November 25, 2016

Phenomenon: Strengthening of the synapses between co-active neurons

This phenomenon is known today as Hebbian plasticity and is a form of Long-Term Potentiation (LTP) and of activity-dependent plasticity. Martin et al. (2000) and Neves et al. (2008) review some of the experimental evidence showing that the memory attractors are formed by means of LTP.

Models for memory attractors: Hopfield (1982) proposed a model to study the dynamics of attractors and the storage capacity of neural networks by means of the Ising model (for review about results from the model see Brunel et al. (2013)). Each neuron is represented by a spin whose up and down states correspond to a high and a low ring rates, respectively. Then the cell assembly would be represented by the set of vertices in the network, the engram by the connectivity matrix and the attractor by the stable spin configurations. Hopfield gave a mathematical expression for the connectivity matrix that supports a given set of attractors chosen a priori. However, the learning phase in which such connectivity is built through synaptic plasticity mechanisms is not been considered within its framework.

To the best of our knowledge, analytical results on neural networks with plastic synapses where the learning phase is been considered are restricted to models of binary neurons with binary synapses. We could not find any analytical result on neural networks with non-binary synapses or using the Ising model with plastic interactions in the literature.

Model: We present a model of a network of binary point neurons also based on the Ising model. However, in our case we consider the connections between neurons to be plastic so that their strengths change as a result of neural activity. In particular, the form of the transitions for the coupling constants resemble a basic Hebbian plasticity rule, as described by Gerstner and Kistler (2002). Therefore, it represents a mathematically treatable model capable of reproducing several features from learning and memory in neural networks.

The model combines the stochastic dynamics of spins on a finite graph together with the dynamics of the coupling constants between adjacent sites. The dynamics are described by a non-stationary continuous-time Markov process.

Let $G = (V; E)$ be a finite undirected graph without self-loops. For each vertex $v \in V$ we associate a *spin* $\sigma_v \in \{-1,1\}$ and, for each edge $e = (v, v') \in E$ we associate a *coupling constant* $J_e \equiv J_{vv'} \in \mathbb{Z}$. These constants are often called exchange energy constants. Here we will also use the term *strength* for the coupling constants.
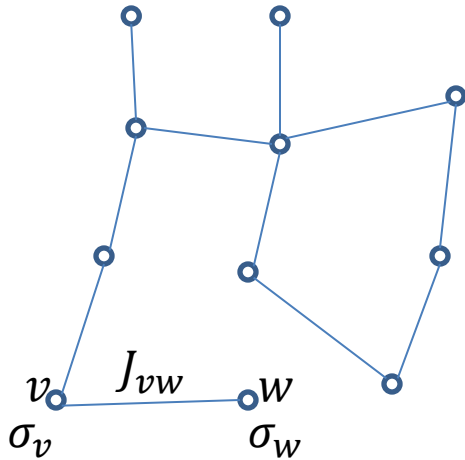
configuration of spins $\qquad \boldsymbol{\sigma} = (\sigma_v, v \in V) \in \{-1,1\}^V$
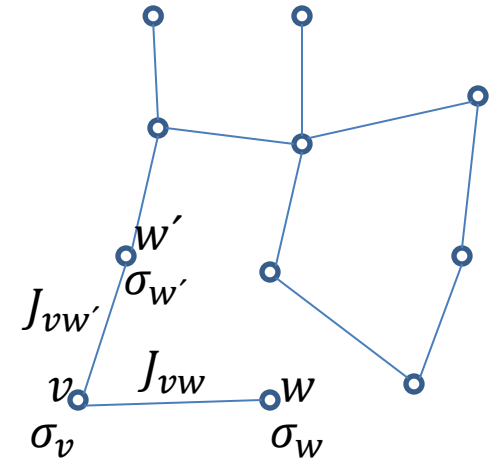configuration of strengths $\qquad \boldsymbol{J} = (J_e, e \in E) \in \mathbb{Z}^V$
state space $\mathcal{A}$ is the set of all possible pairs of configurations
of spins and strengths $\qquad \mathcal{A} = \{-1,1\}^V \times \mathbb{Z}^V$

The following functions will play a key role
in further definitions: *weight for sign flip*

$$\eta_v(\boldsymbol{\sigma}, \boldsymbol{J}) = \sigma_v \sum_{v':v'\sim v} J_{vv'}\sigma_{v'}$$

$$\eta_v(\boldsymbol{\sigma}, \boldsymbol{J}) = J_{vw}\sigma_v\sigma_w + J_{vw'}\sigma_v\sigma_{w'}$$

configuration of spins $\quad\quad\quad \boldsymbol{\sigma} = (\sigma_v, v \in V) \in \{-1,1\}^V$

configuration of strengths $\quad\quad \boldsymbol{J} = (J_e, e \in E) \in \mathbb{Z}^V$

state space $\mathcal{A}$ is the set of all possible pairs of configurations
of spins and strengths $\quad\quad\quad \mathcal{A} = \{-1,1\}^V \times \mathbb{Z}^V$

weight for sign flip

$$\eta_v(\boldsymbol{\sigma}, \boldsymbol{J}) = \sigma_v \sum_{v':v'\sim v} J_{vv'}\sigma_{v'}$$

Transitions rates: $\quad$ for given state $(\boldsymbol{\sigma}, \boldsymbol{J}) \in \mathcal{A}$

spin flip $\quad \sigma_v \to -\sigma_v$ occurs with rate $\quad c_v(\boldsymbol{\sigma}, \boldsymbol{J}) = \dfrac{1}{1+\exp(2\eta_v(\boldsymbol{\sigma}, \boldsymbol{J}))}$

strength change $J_{vv'} \to J_{vv'} + \sigma_v \sigma_{v'}$ occurs with constant rate $v_{vv'}(\boldsymbol{\sigma}, \boldsymbol{J}) \equiv v$

Continuum time Markov chain: $\xi(t) = (\boldsymbol{\sigma}(t), \boldsymbol{J}(t))$

Discrete time Markov chain: $\xi_m = (\hat{\boldsymbol{\sigma}}(t), \hat{\boldsymbol{J}}(t))$ embedded Markov chain with transitions

spin flip $\quad \sigma_v \to -\sigma_v$ occurs with probability $\dfrac{c_v(\boldsymbol{\sigma}, \boldsymbol{J})}{D(\boldsymbol{\sigma}, \boldsymbol{J})}$

strength change $J_{vv'} \to J_{vv'} + \sigma_v \sigma_{v'}$ occurs with probability $\dfrac{v}{D(\boldsymbol{\sigma}, \boldsymbol{J})}$ where

$$D(\boldsymbol{\sigma}, \boldsymbol{J}) = |E|v + \sum_{v \in V} c_v(\boldsymbol{\sigma}, \boldsymbol{J})$$

for given state $(\boldsymbol{\sigma}, \boldsymbol{J}) \in \mathcal{A}$

$$|E|v < D(\boldsymbol{\sigma}, \boldsymbol{J}) \leq |E|v + |V|$$

Theorem 1: The Markov chain ξ_m ($\xi(t)$) is transient

Lyapunov function criteria for transience Fayolle et al. (1995), Menshikov et al. (2017)

For a discrete-time Markov chain $\mathcal{L} = (\zeta_m, m \in \mathbb{N})$ with state space Σ to be transient it is necessary and sufficient that there exists a measurable positive function (the Lyapunov function) $f(\alpha)$, on the state space, $\alpha \in \Sigma$, and a non-empty set $A \subset \Sigma$, such that the following inequalities hold true

(L1) $\mathbb{E}[f(\zeta_{m+1}) - f(\zeta_m) \mid \zeta_m = \alpha] \leq 0$, for any $\alpha \notin A$,
(L2) there exists $\alpha \notin A$ such that $f(\alpha) < \inf_{\beta \in A} f(\beta)$.

Moreover, for any initial $\alpha \notin A$

$$\mathbb{P}(\tau_A < \infty \mid \zeta_0 = \alpha) \leq \frac{f(\alpha)}{\inf_{\beta \in A} f(\beta)}$$

Theorem 1: The Markov chain ξ_m ($\xi(t)$) is transient.

Proof of Theorem:

choose N such that $e^{2N}v \geq |V|(N+1)$ then the Lyapunov function will be defined as

$$f(\boldsymbol{\sigma},\boldsymbol{J}) = \begin{cases} \sum_{v \in V} \dfrac{1}{\eta_v}, & \text{if } \eta_v > N, \text{for all } v \in V, \\[2ex] \dfrac{|V|}{N}, & \text{otherwise.} \end{cases}$$

and the set A will be defined as

$$A = \{(\boldsymbol{\sigma},\boldsymbol{J}) \in \mathcal{A}: \ \min_{v \in V} \eta_v(\boldsymbol{\sigma},\boldsymbol{J}) \leq N\}$$
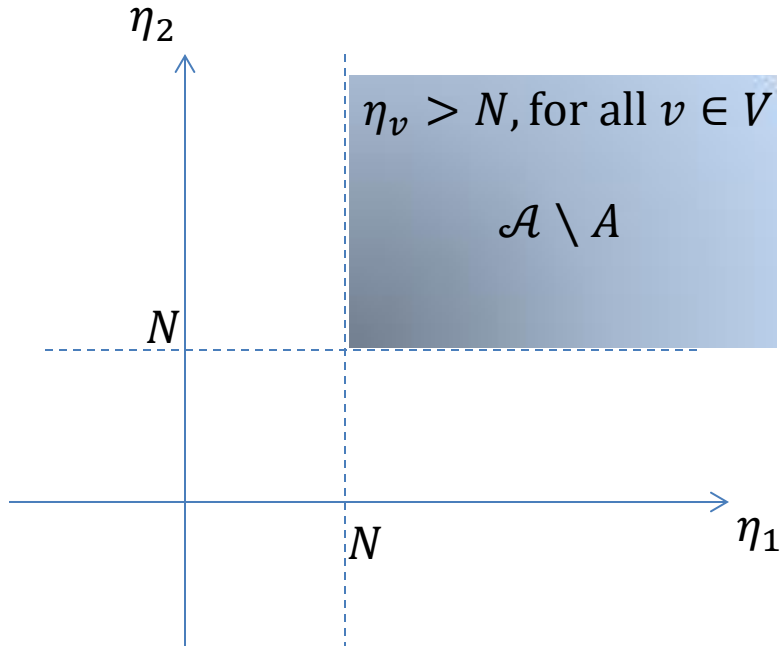
Then (L1) and (L2) hold true

Let τ be the *freezing time*

$\tau := \max\{m \geq 1: \; \hat{\sigma}(m-1) \neq \hat{\sigma}(m)\}$ assuming $\max\{\emptyset\} = 0$.

Theorem 2: $\mathbb{P}(\tau < \infty) = 1$

Proof of Theorem 2

$$f(\boldsymbol{\sigma},\boldsymbol{J}) = \begin{cases} \displaystyle\sum_{v\in V}\frac{1}{\eta_v}, & \text{if } \eta_v > N, \text{for all } v \in V, \\[2em] \displaystyle\frac{|V|}{N}, & \text{otherwise.} \end{cases}$$

the A was defined as

$$A = \{(\boldsymbol{\sigma},\boldsymbol{J}) \in \mathcal{A}: \ \min_{v\in V}\eta_v(\boldsymbol{\sigma},\boldsymbol{J}) \leq N\}$$

Moreover for any $(\boldsymbol{\sigma},\boldsymbol{J}) \in \mathcal{A} \setminus A$

$$\mathbb{P}(\tau_A < \infty \mid \xi_0 = (\boldsymbol{\sigma},\boldsymbol{J})) \leq \frac{f(\boldsymbol{\sigma},\boldsymbol{J})}{\displaystyle\inf_{\beta\in A} f(\beta)} = \frac{\sum_{v\in V}\eta_v^{-1}(\boldsymbol{\sigma},\boldsymbol{J})}{\dfrac{|V|}{N}} \leq \frac{N}{\displaystyle\min_{v\in V}\eta_v(\boldsymbol{\sigma},\boldsymbol{J})} \leq \frac{N}{N+1} < 1$$
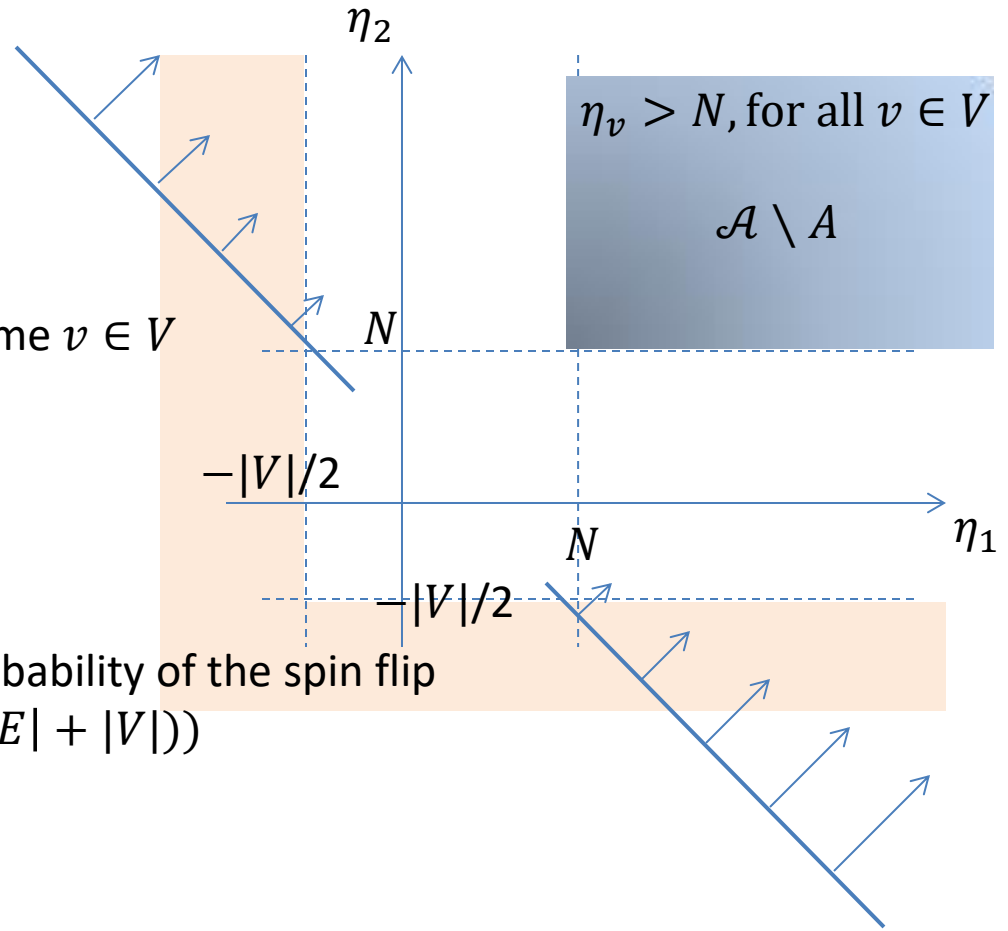
$$\mathbb{P}(\tau_A = \infty \mid \xi_0 = (\boldsymbol{\sigma},\boldsymbol{J})) \geq \frac{1}{N+1}$$

Proof of Theorem 2

$$s(\sigma, J) = \sum_{v \in V} \eta_v(\sigma, J)$$

If (σ, J) such that $\eta_v(\sigma, J) < -|V|/2$ for some $v \in V$
$$\mathbb{E}[s(\xi_{m+1}) - s(\xi_m) \mid \xi_m = (\sigma, J)] \geq 1/2$$

$\eta_v > N$, for all $v \in V$

$\mathcal{A} \setminus A$

If (σ, J) such that $\eta_v(\sigma, J) < 0$ then the probability of the spin flip
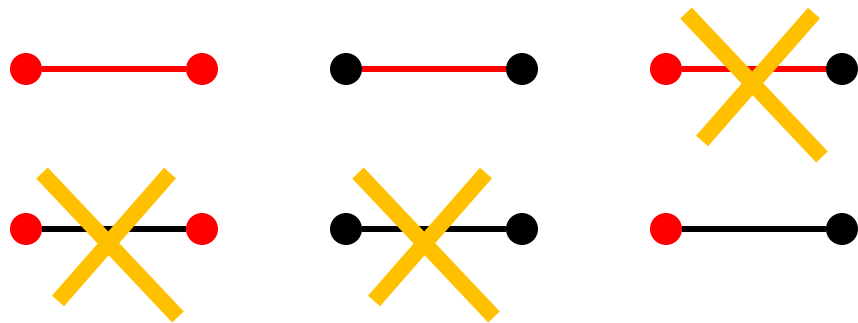$\eta_v \to -\eta_v$ is at least $1/2D(\sigma, J) > 1/(2(v|E| + |V|))$

Let $\mathcal{B} = \left\{(\sigma, J) : \min_{v \in V} \eta_v(\sigma, J) < -\frac{|V|}{2}\right\} \subset A$. Thus if initial $(\sigma, J) \in \mathcal{B}$ then
$$\mathbb{P}\left(\tau_{\mathcal{A} \setminus \mathcal{B}} < \infty \mid \xi_0 = (\sigma, J)\right) = 1$$

Theorem 3: As a consequence of Theorem 2, for any $e \in E$ almost surely

$$\lim_{m \to \infty} \frac{\left|\hat{J}_e(m)\right|}{m} = \frac{1}{|E|}, \qquad \lim_{t \to \infty} \frac{|J_e(t)|}{t} = v.$$

Theorem 4: $j_{vv'}^{\infty} = \sigma_v^{\infty} \sigma_{v'}^{\infty}$ for any $(v, v') \in E$

References:

1. Martin, S., Grimwood, P., Morris, R., 2000. Synaptic plasticity and memory: an evaluation of the hypothesis. Annual review of neuroscience 23 (1), 649-711.

2. Neves, G., Cooke, S. F., Bliss, T. V., 2008. Synaptic plasticity, memory and the hippocampus: a neural network approach to causality. Nature Reviews Neuroscience 9 (1), 65-75.

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

4. Brunel, N., del Giudice, 160 P., Fusi, S., Parisi, G., Tsodyks, M., 2013. Selected Papers of Daniel Amit (1938-2007). World Scientic Publishing Co., Inc.

5. Gerstner, W., Kistler, W. M., 2002. Spiking neuron models: Single neurons, populations, plasticity. Cambridge university press.

6. Fayolle, G., Malyshev, V. A., Menshikov, M. V., 1995. Topics in the constructive theory of countable Markov chains. Cambridge university press.

7. Menshikov, M., Popov, S., Wade, A., 2017. Non-homogeneous random walks - Lyapunov function methods for near-critical stochastic systems. Cambridge university press, to appear.