Inferred Learning Rules From IT Cortex Are Optimal for Memory Storage and Lead to Graded and Time Varying Neural Representations

Ulises Pereira
ulises@uchicago.edu

Nicolas Brunel
nbrunel@uchicago.edu

The attractor neural network (ANN) scenario has been a popular scenario for memory storage in association cortex, but there is still a large gap between these models and experimental data. Three issues have however recently received some attention: (1) whether the distribution of the learned patterns is compatible with data; (2) whether the learning rules used in such models are compatible with data; (3) whether the temporal evolution of single neuron activity is compatible with data. A new study [1] has found in IT cortex a distribution of neuronal responses close to lognormal, at odds with bimodal distributions of firing rates used in the vast majority of theoretical studies; and a Hebbian learning rule dominated by depression with a non-linear dependence on postsynaptic firing rate. On the other hand, multiple studies have shown high temporal variability in single units during delay periods in delay match to sample (DMS) experiments. This is at odds with the classical ANN's view of the delay period, where the network state converges to a static attractor correlated with one of the memories but where activities of neurons do not vary in time. To investigate these questions, we consider a network of N neurons with firing rates represented by a vector of analog variables $\vec{r}$, where patterns $\{\vec{\xi}^k\}_{k=1}^p$ are imprinted in the connectivity matrix as the corresponding firing rates elicited by the patterns, neglecting contributions of the recurrent connections (i.e. $\eta_i^k \equiv F(\xi_i^k)$). The resulting connectivity matrix is given by

$$J_{ij} = \frac{c_{ij}}{cN} \sum_{k=1}^{p} f(\eta_i^k) g(\eta_j^k) \tag{1}$$

where the learned patterns are distributed as $\xi_i^k \overset{iid}{\sim} \mathcal{N}(0,1)$, and c_{ij} is a sparse directed Erdős-Rényi structural connectivity where each synapse is present with probability c. The firing rates $r_i(t)$ of each neuron evolve according to $\dot{r}_i = -r_i + F\left(I_i + \sum_{i\neq j}^{N} J_{ij} r_j\right)$. The pair of functions f and g define together the learning rule (LR), where f can be inferred from data [1]. This is a generalization of the classical Hebbian learning rule that allows a nonlinear dependence of the synaptic strength with the pre and post synaptic activity, but maintains the separability of the rule in pre and post-synaptic rates. We further assume that $\mathbb{E}_\xi(g(F(\xi)) = 0$ so that the average change in connection strength due to learning of a single pattern is zero. We perform a mean field analysis in the limit $p, N \to \infty$ and $c \ll 1$, maintaining the number of patterns stored by synapse or load ($\alpha \equiv p/(Nc)$) finite. We found self-consistent equations for the covariance between the steady state of the network and a nonlinear transformation of one of patterns learned (the *overlap* $m = \mathbb{E}_{\xi,\xi^1}(rg(F(\xi)))$) and for the second moment of the steady state of the network ($M = \mathbb{E}_{\xi,\xi^1}\left(r^2\right)$). Here $\mathbb{E}_{\xi,\xi^1}(\cdot)$ denotes an average over the quenched disorder of all the learned patterns ξ but the first which is the initial condition of the dynamics ξ^1. From our mean field theory (MFT) we conclude that both order parameters depend heavily on the specific form of f and g. We investigate which choice of the f and g functions lead to the maximal possible storage capacity (i.e. $\max_{f,g} \alpha_c$). We find that in optimal learning rules, f and g are strongly non-linear (close to being step functions), and characterized by identical thresholds for f and g. Furthermore, the optimal rule has thresholds between potentiation and depression that are a rate that is much higher than the mean firing rate of the patterns learned, and the post-synaptic function f is balanced towards depression, consistent with rules extracted from data. We derive the distribution of the steady state firing rates r for both 'novel' patterns (patterns that are not stored in the connectivity matrix) and 'familiar' patterns (that are stored in the connectivity matrix). We find a large region of parameters both distributions are unimodal with low mean firing rates, similar to the graded firing rate responses during the delay period in DMS experiments. Moreover, we found that transfer functions (i.e. $F(\cdot)$) and learning rules inferred from IT cortex *in vivo* recordings lead to storage capacities (α_c) close to the optimal (i.e. $\max_{f,g} \alpha_c$), and match well the empirical distributions. Lastly, we found that for loads close to maximal capacity, our model presents a chaotic phase with associative memory properties. In this parameter region, fixed point attractors become chaotic attractors correlated with one of the stored patterns but not with the others (i.e. finite overlap). We analyze this chaotic phase using a dynamical MFT. Our theory makes quantitative predictions about the boundaries of the chaotic phase. This chaotic phase with associative memory properties can explain intra and across trials variability reported during delay periods, reconciling ANNs theory with the large fluctuations observed during delay periods. Numerical simulations of large networks match well our analytical results.

References

[1] Lim et al, "Inferring learning rules from distributions of firing rates in cortical neurons" *Nature Neurosci.* (2015)

Ulises Pereira

Contact Information

The University of Chicago
Department of Statistics
George Herbert Jones Laboratory
5747 S. Ellis Avenue
Chicago, IL 60637

773-702-8335
ulises@uchicago.edu

Research Interests

Theoretical Neuroscience and Machine Learning.

Education

The University of Chicago, Chicago, IL

Ph.D., Statistics, *Expected:* Fall 2018

- Thesis Disertation: *Unsupervised Learning of Spatiotemporal Attractors In Neural Networks*
- Advisor: Nicolas Brunel, Ph.D

M.S., Statistics, Aug 2015

Universidad de Chile, Santiago, Chile.

Magister, Physics, September 2013

- Thesis Topic: *Nonlinear Dynamics*
- Advisor: Enrique Tirapegui, Ph.D
- *W*ith Highest Distinction.

Ingeniería, Molecular Biotechnology, June 2013

- *W*ith Highest Distinction. Ranked first out of eight graduates (1/8).

Licenciatura, Physics, June 2012

- *W*ith Distinction. Ranked first out of six graduates (1/6).

Refereed Journal Publications

1. **Pereira U**, Coullet P. and Tirapegui E. The Bogdanov-Takens Normal Form: A Minimal Model for Single Neuron Dynamics. Entropy. 2015.
2. Vera J., Pezzoli M., **Pereira U.**, Bacigalupo J. and Sanhueza M. Electrical Resonance in the θ Frequency Range in Olfactory Amygdala Neurons. Plos One. 2014.
3. Contreras D., **Pereira U.**, Hernández V., Reynaert B. and Letelier J.C. A loop conjecture for metabolic closure. *Advances in Artificial Life, ECAL 2011.* MIT press. 2011. Selected one of the ten best papers of ECAL 2011.
4. Jaramillo S., Honorato-Zimmer R., **Pereira U.** , Contreras D., Reynaert B., Hernández V., Soto-Andrade J., Cárdenas M.L., Cornish-Bowden A. and Letelier J.C. (M,R) Systems and RAF Sets: Common Ideas, Tools and Projections. *Artificial life XII.* MIT press. 2010.

Papers in Preparation

1. **U. Pereira** and N. Brunel. Inferred Learning Rules From IT Cortex Are Optimal for Memory Storage and Lead to Graded and Time Varying Neural Representations.
2. **U. Pereira** and N. Brunel. Unsupervised Learning of Persistent and Sequential Activity.

Honors and Awards

- **Doctoral Becas-Chile Scholarship**. Chilean Government. *Commission of Research in Science and Technology of the Chilean Government (CONICYT). 2013*. Scholarship declined.
- **Doctoral Fulbright Fellowship**. U.S. Government. *Fulbright commission . 2012*.
- **Best Physics Student of Class 2011**. Universidad de Chile. *Department of Physics. 2011*.
- **CONICYT Master Fellowship**. Chilean Government. *Commission of Research in Science and Technology of the Chilean Government (CONICYT). 2011*. Ranked 5/1584 at national level.
- **Bicentenario Scholarship for Undergraduate Studies**. Chilean Government. *Ministry of Education. 2004*.
- **Scholarship for Outstanding Score in PSU**. *PSU (Spanish acronym) is the national Chilean university selection test.* Pontifical Catholic University of Chile. *2004*. Scholarship declined.

Conference Posters

Pereira U. and Brunel N. Optimal Unsupervised Hebbian Learning Rules For Attractor Neural Networks. COSYNE Poster Presentation. Salt Lake City, EEUU. February, 2017.

Pereira U. and Brunel N. Unsupervised Learning of Persistent and Sequential Activity. COSYNE Poster Presentation. Salt Lake City, EEUU. February, 2016.

Pereira U. and Brunel N. Unsupervised Learning of Sequential Activity. XV International Workshop on Instabilities and Nonequilibrium Structures. Valparaíso, Chile. December, 2015.

Vera J., **Pereira U.**, Reynaert B., Bacigalupo J. and Sanhueza M. Modulation of frequency preference by changes in input resistance. 44th Annual Meeting Society for Neuroscience. Washington D.C., USA. November, 2014.

Vera J., **Pereira U.**, Reynaert B., Deichler A., Astudillo D., Bacigalupo J., and Sanhueza M. A biological context for theta-frequency neuronal resonance: a comparative study between cortical amygdala and hippocampal neurons. X Annual meeting of the Chilean Society for Neuroscience. October, 2014. Valdivia, Chile. **Awarded for best panel presentation.**

Pereira U., Tirapegui E. Una Ecuación Universal Para la Dinámica Neuronal. In Proceedings of the XVII Conference on Nonequilibrium Statistical Mechanics and Nonlinear Physics. Santiago, Chile. December 2012.

Pereira U., Tirapegui E. Una Ecuación Universal Para la Dinámica Neuronal. In Proceedings of the XVIII Simposio Chileno de Física. La Serena, Chile. November, 2012.

Pereira U., Vera J., Pezzoli M., Bacigalupo J. and Sanhueza M. A computational conductance-based model that reproduce theta resonance dynamics in olfactory amygdala neurons.*41st Annual meeting of the Society for Neuroscience.* Washington DC, EEUU. November, 2011.

Vera J, **Pereira U.**, Pezzoli M., Bacigalupo J. and Sanhueza M. Sub and supra-threshold dynamics of resonant neurons in the olfactory amygdala.*41st Annual meeting of the Society for Neuroscience.* Washington DC, EEUU. November, 2011.

Contreras D, **Pereira U.**, Hernández V., Reynaert B. and Letelier J.C..A loop conjecture for metabolic closure.*Eleventh European Conference on the Synthesis and Simulation of Living Systems.* Paris, France. August, 2011.

Pereira U., Pezzoli M., Bacigalupo J., Sanhueza M.. A computational conductance-based model of electrical resonance in the theta frequency range in olfactory amygdala neurons. VI meeting of the Chilean Society of Neuroscience. Valdivia, Chile. September, 2010.

Jaramillo S., Honorato-Zimmer R., **Pereira U.**, Contreras D., Reynaert B., Hernández V., Soto-Andrade J., Cárdenas M.L., Cornish-Bowden A. and Letelier J.C. (M,R) Systems and RAF Sets: Common Ideas, Tools and Projections. *XII Artificial life Conference*. Odense, Denmark. August, 2010.

Teaching Experience

Instructor

- **Statistical Models and Methods**. The University of Chicago. *Winter 2015*.

Teaching Assistant

- **Theoretical Neuroscience: Network Dynamics and Computation**. The University of Chicago. *Winter 2017*
- **Statistical Methods and Applications**. The University of Chicago. *Fall 2016*.
- **Statistical Methods and Applications**. The University of Chicago. *Fall 2015*.
- **Statistical Models and Methods**. The University of Chicago. *Spring 2015*.
- **Elementary Statistic**. The University of Chicago. *Fall 2014*.
- **Statistical Methods and Applications**. The University of Chicago. *Spring 2014*.
- **Theoretical Neuroscience: Network Dynamics and Computation**. The University of Chicago. *Winter 2013*
- **General Physiology**. Universidad de Chile. *Autumn Semester 2010*
- **Biological Instrumentation**. Universidad de Chile. *Spring Semester 2008*.

Courses

- **Latin American Summer School in Computational Neuroscience**. Institute of Complex Systems. *Valparaíso, Chile. 11 to 29 January, 2010*.
- **VI Summer School of Complex Systems**. Institute of Complex Systems. *Valparaíso, Chile. 7 to 11 of January, 2008*.