Synaptic learning rule from optimization of information transmission and energy cost: a theoretical calculation[1] [2]

Summary

From a top-down perspective of function optimization and structural adaptation, this work aims to design synaptic learning rule theoretically that can account for organizational features in many biological neural networks. The function optimization includes maximizing the mutual information between neuron pairs and minimizing the wiring cost as well as the firing cost. The structural adaptation endows synaptic plasticity which is reliant on the neural activities in the network. The idealized neural network here is described by the linear firing rate model. The derived learning rule based on gradient ascent can lead to small-world structure in the network.

The learning rule in a firing rate network

Consider a linear firing rate neural network with Gaussian external input (Equation 1). r_i is the firing rate of i-th neuron. w_{ij} is the weight of synapse from j-th neuron to i-th neuron. s_i is the external Gaussian input to i-th neuron. τ_r is time scale of the neurons' firing dynamics.

$$\tau_r \frac{dr_i}{dt} = -r_i + \Sigma_{j,j\neq i} w_{ij} r_j + s_i, (i = 1, 2, ..., n) \tag{1}$$

Use mutual information between each pair of neuron to quantify information transmission between them. The need for communication can be described by a functional adjacency matrix $\mathbf{W}_{\text{info}}$. The network wide information transmission I in the network is thus the circle product of functional adjacency matrix and mutual information matrix $\mathbf{I_m}$ (Equation 2).

$$I = \sum_{i,j} I_m(r_i, r_j) = \mathbf{W}_{\text{info}} \circ \mathbf{I_m} \tag{2}$$

Firing cost is linear with firing rate of all the neurons since single action potential of neurons is constant approximately. Wiring cost is proportional to the power of synaptic weight. γ is the material cost efficient. The total energy cost E is shown in Equation 3.

$$E = \sum_{i,j} w_{ij}^{\gamma} + \sum_i r_i \tag{3}$$

The objective function is Equation 4. ζ is the ratio of information portion to energy portion in the objective function. I_0 and E_0 are initial information and energy respectively before any optimization.

$$H = I/I_0 - \zeta E/E_0 \tag{4}$$

To optimize H, a **learning rule** (Equation 5) is derived such that $dH \geq 0$ as the network evolves.

$$\Delta w_{ij} = \alpha(\Delta w_{ij})_1 / I_0 - \alpha(\Delta w_{ij})_2 / E_0 \tag{5}$$

where

$$(\Delta w_{ij})_1 =$$

$$\sum_{s,t} \left((w_{sl} + \sum_l w_{sl} w_{li}) \frac{\Sigma_{ss}\Sigma_{jt} - \Sigma_{st}\Sigma_{js}}{\Sigma_{ss}} + (w_{ti} + \sum_l w_{tl} w_{li}) \frac{\Sigma_{tt}\Sigma_{js} - \Sigma_{st}\Sigma_{jt}}{\Sigma_{tt}} \right) \left(\Sigma_{ss}\Sigma_{tt} - \Sigma_{st}^2\right) \Sigma_{st}$$

and

$$(\Delta w_{ij})_2 = r_j \sum_p \left(w_{pi} + \sum_t w_{pt} w_{ti} \right) + r_j \sum_p \left(w_{pi} + \sum_t w_{pt} w_{ti} \right)$$

Small-world structures resulted from the learning rule

Numerical implementation of the neural network is performed with customized Python codes. The termination of simulation is determined by the change of objective function at each time step. To quantify the final network structure, 'small-world' parameter ω borrowed from network science is applied here. The closeness of ω to 0 suggests the 'small-worldness' of the network structure.

[1] This manuscript summarizes my senior thesis work under the supervision of Dr. Songting Li at Shanghai Jiao Tong University, China. Brief mention of 'Acknowledgement' part: Thank Dr. Li for his patient guidance and insightful advice during this work. Thank Dr. Yanyang Xiao and Qianning Hou for helpful discussions. Thank people in Dr. Douglas Zhou's group for warm host during my stay at SJTU. I treasure my undergraduate study at University of Science and Technology of China. Finally thank my parents for the support along the way.

[2] Yuru Song (yus005@ucsd.edu) Nov-05-2019.

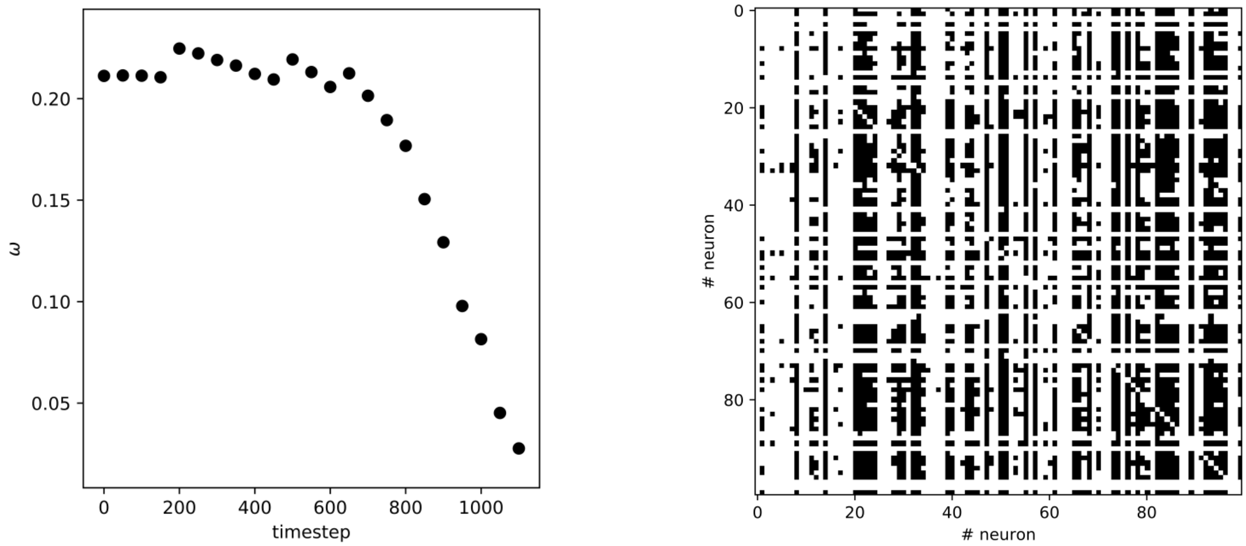

Figure 1: Optimized structure as small-world network. Left: Small-world metric ω decreases to 0 as network adapts. Right: Connection matrix of optimized structure. Hub nodes can be identified.

Step-by-step calculation of the learning rule

A biologically plausible assumption is made that the external inputs into each neuron are independent Gaussian random variables, i.e.

$$\vec{s} = (s_1, s_2, \ldots, s_n)^T \sim \mathbf{N}(\vec{u}_s, \mathbf{\Pi})$$

which will be convenient for calculation.

Calculation of $(\Delta w_{ij})_1$

Since we are interested in the adaptation of synaptic weights, instead of temporal dynamics of neurons' firing rate, which has much shorter time scale than the former, Equation 1 will be assigned 0 here. Then follows the firing rates at steady state,

$$\vec{r} \sim \boldsymbol{N}((\mathbf{I} - \mathbf{W})^{-\mathbf{1}}\vec{u}_s, (\mathbf{I} - \mathbf{W})^{-\mathbf{1}}\mathbf{\Pi}(\mathbf{I} - \mathbf{W})^{-\mathbf{T}})$$

One good property of Gaussian random variables is that their mutual information can be explicitly written out in a simple form as

$$\mathrm{I} = -\frac{1}{2}\log\left(1 - \rho^2\right)$$

ρ is the correlation coefficient. In our case,

$$I_m\left(r_s, r_t\right) = -\frac{1}{2}\log\left(1 - \rho_{st}^2\right), \rho_{st} = \frac{\Sigma_{st}}{\sqrt{\Sigma_{ss}\Sigma_{tt}}}$$

We continue to calculate

$$\frac{\partial I_m\left(r_s, r_t\right)}{\partial w_{ij}}$$

$$\partial I_m\left(r_s, r_t\right) = -\frac{1}{2\left(\Sigma_{ss}\Sigma_{tt} - \Sigma_{st}^2\right)}\Sigma_{ss}\partial\Sigma_{tt} + \Sigma_{tt}\partial\Sigma_{ss} - 2\Sigma_{st}\partial\Sigma_{st} + \frac{1}{2\Sigma_{ss}}\partial\Sigma_{ss} + \frac{1}{2\Sigma_{tt}}\partial\Sigma_{tt}$$

Since

$$\mathbf{\Sigma} = (\mathbf{I} - \mathbf{W})^{-\mathbf{1}}\mathbf{\Pi}(\mathbf{I} - \mathbf{W})^{-\mathbf{1}}$$

it can be approximated as

$$\left(\mathbf{W} + \mathbf{W^2}\right)\mathbf{\Pi}\left(\mathbf{W} + \mathbf{W^2}\right)^{\mathbf{T}}$$

if the spectral radius of $\mathbf{W}$ is less than 1. We can get

$$\Sigma_{st} \approx \sum_p \left(w_{sp} + \sum_l w_{sl}w_{lp}\right)\Pi_{pp}\left(w_{pt} + \sum_l w_{pl}w_{lt}\right)$$

and thus

$$\left(\frac{\partial \Sigma_{pq}}{\partial w_{ij}}\right) = \left(\frac{\partial \mathbf{\Sigma}}{\partial w_{ij}}\right)_{pq} = \left((\mathbf{I}-\mathbf{W})^{-\mathbf{1}}\mathbf{J^{ij}}\mathbf{\Sigma}\right)_{pq} + \left((\mathbf{I}-\mathbf{W})^{-\mathbf{1}}\mathbf{J^{ij}}\mathbf{\Sigma}\right)_{qp}$$

Substituting all the partial derivatives using above relationship, we have

$$\frac{\partial I_m\left(r_s, r_t\right)}{\partial w_{ij}} =$$

$$\left(\left(w_{si} + \sum_l w_{sl} w_{li}\right)\left(\frac{\Sigma_{ss}\Sigma_{jt} - \Sigma_{st}\Sigma_{js}}{\Sigma_{ss}}\right) + \left(w_{ti} + \sum_l w_{tl} w_{li}\right)\frac{\Sigma_{tt}\Sigma_{js} - \Sigma_{st}\Sigma_{jt}}{\Sigma_{tt}}\right)\frac{\Sigma_{st}}{\Sigma_{ss}\Sigma_{tt} - \Sigma_{st}^2}$$

sum it over s, t, we can get the formulation of $(\Delta w_{ij})_1$ in Equation 5.

Calculation of $(\Delta w_{ij})_2$

Using similar approximation, we have

$$r_i = \sum_j \left(w_{ij} + \sum_l w_{il} w_{lj}\right) s_j$$

insert it into following equation

$$\partial E/\partial w_{ij} = \sum_i \partial r_i/\partial w_{ij} + \sum_{ij} \gamma w_{ij}^{\gamma-1}$$

we can get $(\Delta w_{ij})_2$ in Equation 5.