

Notes on Homeostatic Adaptation and Error Driven Adaptation in ESNs

Fabian Schubert

September 29, 2019

1 Introduction

Strategies for the optimization of ESN hyperparameters can be divided in two categories: supervised and unsupervised methods, where the first one utilizes an error signal, while the latter only uses information contained within the network dynamics.

In the first part of our research, we investigated the possibility of defining an unsupervised homeostatic mechanism that controls the mean and variance of neuronal firing in such a way that the network acts in a regime that yields good performance in sequence learning tasks. This mechanism acts on two sets of parameters, biases b_i and neural gain factors a_i . It should be emphasized that we did not attempt to define an arbitrarily complex measure that would be most suitable for optimization, e.g. from a machine learning perspective. Rather, we restricted ourselves to adhere to ‘biologically plausible’ mechanisms. While no exact definition of this term exists, it embraced two aspects in our work:

- The dynamics of all variables must be local, i.e., they are bound to a specific neuron and may only access other variables that are locally accessible. In a strict sense, this means all other dynamic variables of the neuron itself and information about the activity of adjacent neurons.
- We use a time-discrete model where the state of a variable in the next step may only be determined by states of the previous step. This means that information about past states must be integrated dynamically.

Our approach was based on the assumption that network performance is optimal when the spectral radius of the effective recurrent connectivity, given by $a_i W_{ij}$, is close to, but slightly below 1. We attempted to transfer this non-local measure into a condition that could be implemented in a biologically plausible way.

2 Model

2.1 Network dynamics

$$y_i(t) = \tanh(a_i x_i(t) - b_i) \quad (1)$$

$$x_i(t) = \sum_{j=1}^N W_{ij} y_j(t-1) + \sum_{j=1}^{D_{\text{in}}} W_{ij}^u u_j(t) \quad (2)$$

$$o_i(t) = o_i^0 + \sum_{j=1}^{D_{\text{out}}} W_{ij}^o y_j(t) \quad (3)$$

where $\mathbf{y}, \mathbf{x}, \mathbf{a}, \mathbf{b} \in \mathbb{R}^N$, $W \in \mathbb{R}^{N \times N}$, $\mathbf{u} \in \mathbb{R}^{D_{\text{in}}}$, $W^u \in \mathbb{R}^{N \times D_{\text{in}}}$, $\mathbf{o}, \mathbf{o}^0 \in \mathbb{R}^{D_{\text{out}}}$ and $W^o \in \mathbb{R}^{D_{\text{out}} \times N}$.

Furthermore

$$p(W_{ij} = x) = \begin{cases} \delta(x) & i = j \\ p_r \mathcal{N}(x, \mu = 0, \sigma = \sigma_w / \sqrt{N p_r}) + (1 - p_r) \delta(x) & \text{else} \end{cases} \quad (4)$$

$$p(W_{ij}^u = x) = \mathcal{N}(x, \mu = 0, \sigma = \sigma_e) \quad (5)$$

Initially, we chose $N = 1000$ as the network size, however, due to computational complexity, the results presented here are generated with a network of size $N = 500$, unless stated otherwise. See Table 1 for the standard network parameters.

Table 1: Standard network parameters.

N	D_{in}	D_{out}	p_r	σ_w	σ_e
500	1	1	0.1	1	variable

2.2 Homeostatic Adaptation

For our homeostatic update mechanism, we use the following dynamics:

$$b_i(t) = b_i(t-1) + \epsilon_b [y_i(t) - \mu_i^t] \quad (6)$$

$$\mu_i^y(t) = [1 - \epsilon_\mu] \mu_i^y(t-1) + \epsilon_\mu y_i(t) \quad (7)$$

$$a_i(t) = a_i(t-1) + \epsilon_a [\sigma_i^t{}^2 - (y_i(t) - \mu_i^y(t))^2] . \quad (8)$$

See Table 2 for the standard values.

Table 2: Standard homeostasis parameters.

ϵ_b	ϵ_μ	ϵ_a	μ_i^t	σ_i^t
10^{-3}	10^{-4}	10^{-3}	0.05	variable

455.24411pt