Biologically Plausible Deep Learning: A Critical Review of Guerguiev *et al.* (2017) [1]

Robert Tjarko Lange[2]
rtl17@ic.ac.uk
www.rob-lange.com

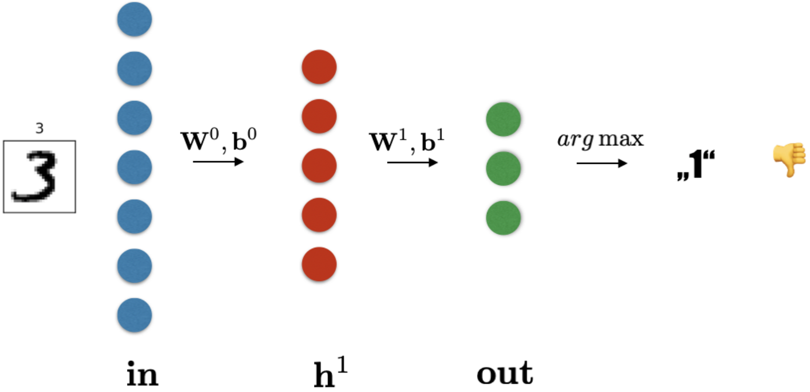
Einstein Center for Neurosciences Berlin

February 10, 2019

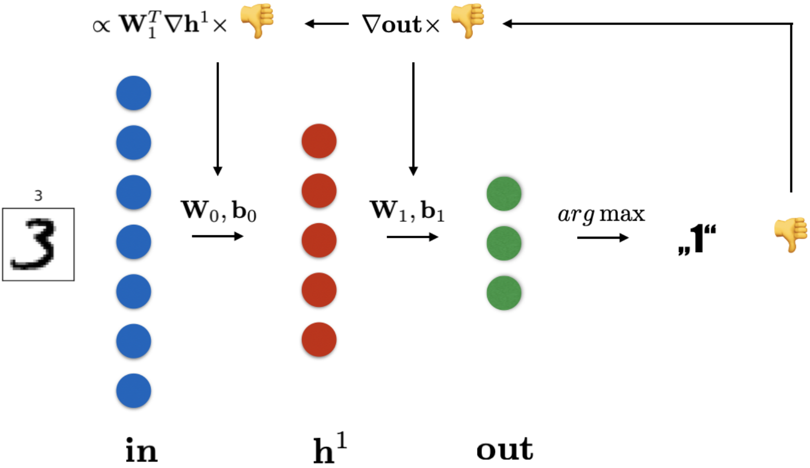
[1] Guerguiev, J., Lillicrap, T. P., & Richards, B. A. (2017). Towards deep learning with segregated dendrites. ELife, 6, e22901.

[2] Code: github.com/RobertTLange/Bio-Plausible-DeepLearning

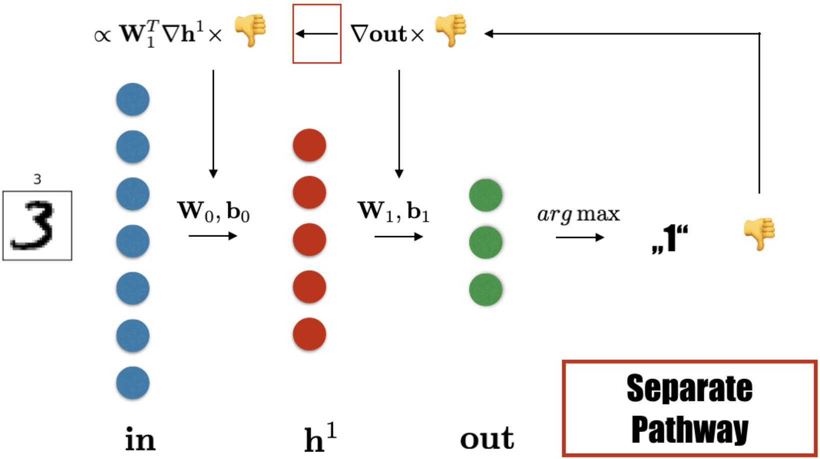
Backprop ... and its problems



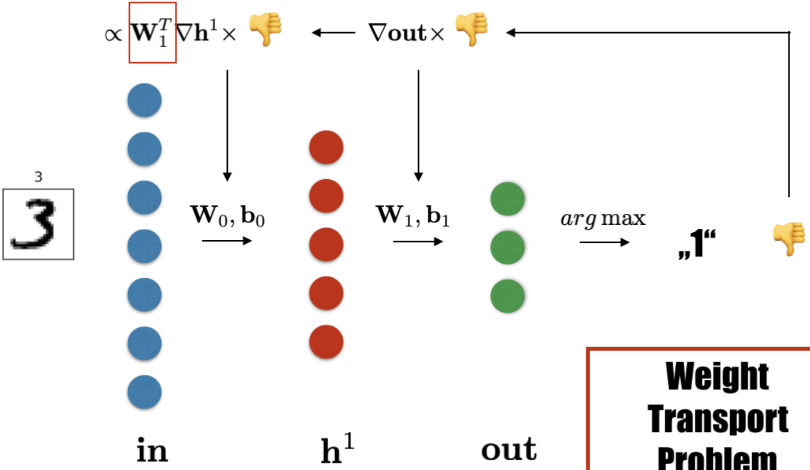
Backprop ... and its problems



Backprop ... and its problems

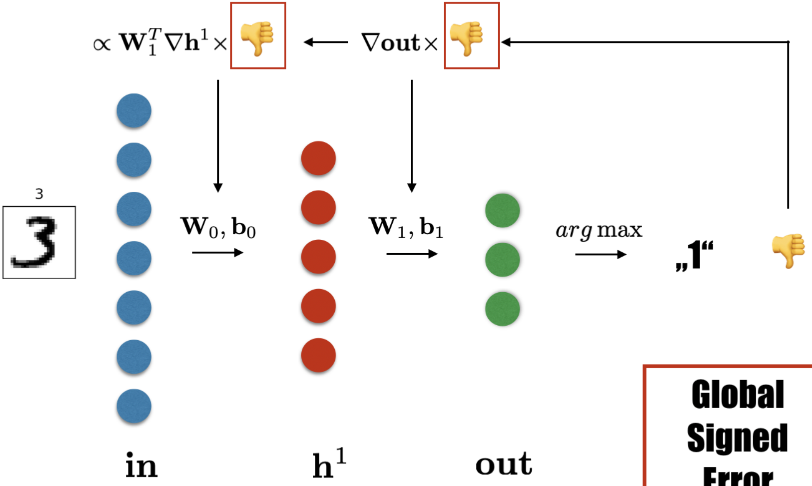


Backprop ... and its problems



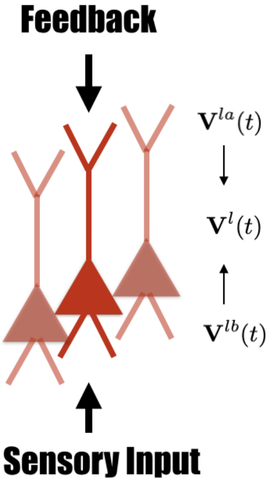
Weight Transport Problem

Backprop ... and its problems

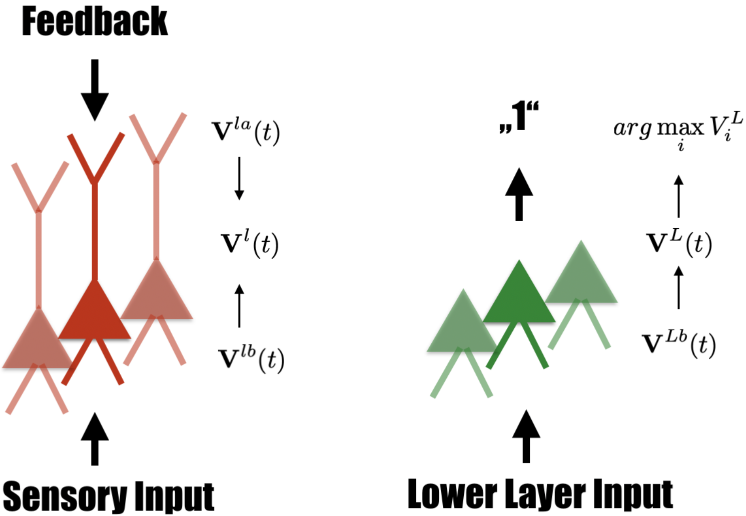


Global Signed Error

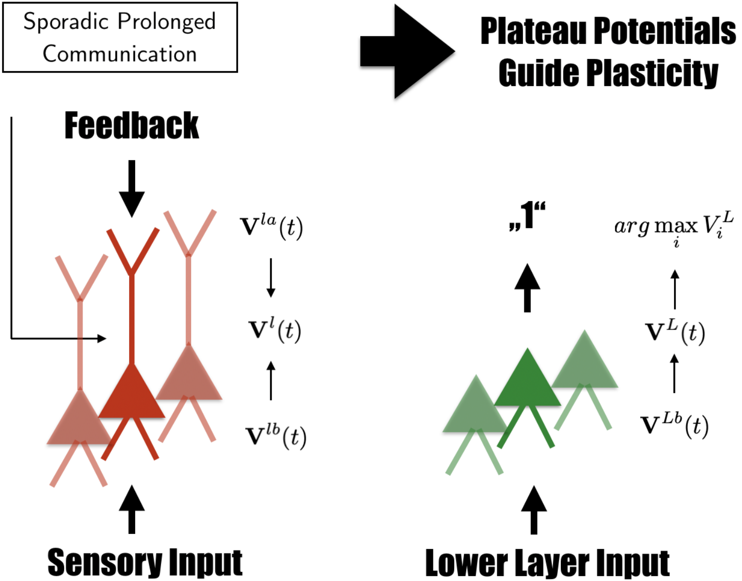
A Solution - Electrical Segregation of ↓ and ↑ Info



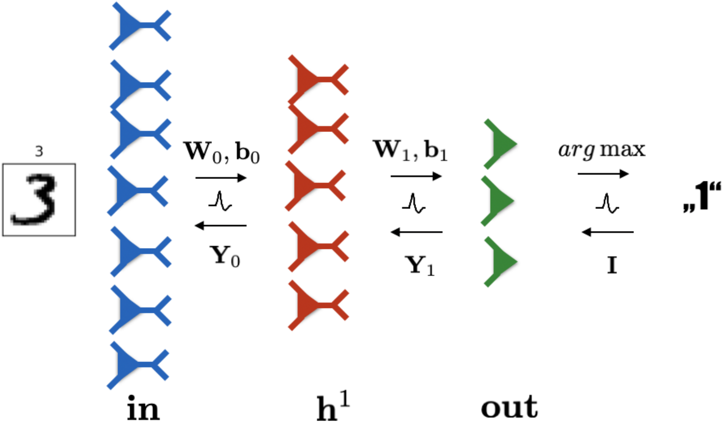
A Solution - Electrical Segregation of ↓ and ↑ Info



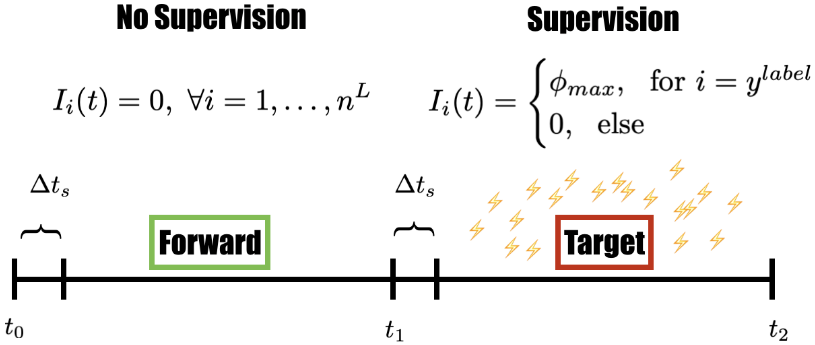
A Solution - Electrical Segregation of ↓ and ↑ Info



No Separate Pathway via Segregation



Credit Assignment Signals



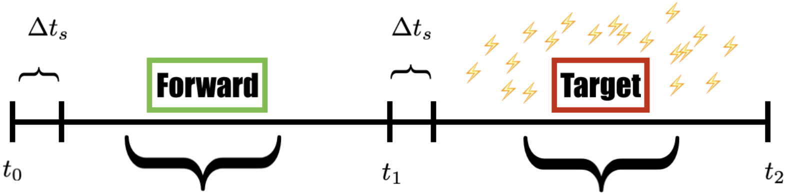
Credit Assignment Signals

No Supervision

$$I_i(t) = 0, \ \forall i = 1, \dots, n^L$$

Supervision

$$I_i(t) = \begin{cases} \phi_{max}, & \text{for } i = y^{label} \\ 0, & \text{else} \end{cases}$$



$$\sigma\left(\frac{1}{\Delta t_1}\int_{t_1-\Delta t_1}^{t_1} V_i^{la}(t)dt\right) := \alpha^{lf}(t)$$

$$\sigma\left(\frac{1}{\Delta t_2}\int_{t_2-\Delta t_2}^{t_2} V_i^{la}(t)dt\right) := \alpha^{lt}(t)$$

Forward Plateau Potential

Target Plateau Potential

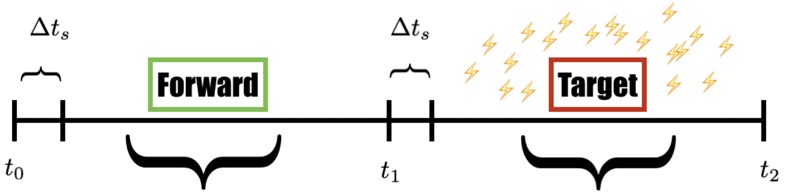
Credit Assignment Signals

No Supervision

$$I_i(t) = 0, \ \forall i = 1, \ldots, n^L$$

Supervision

$$I_i(t) = \begin{cases} \phi_{max}, & \text{for } i = y^{label} \\ 0, & \text{else} \end{cases}$$



$$\sigma\left(\frac{1}{\Delta t_1}\int_{t_1-\Delta t_1}^{t_1} V_i^{la}(t)dt\right)$$

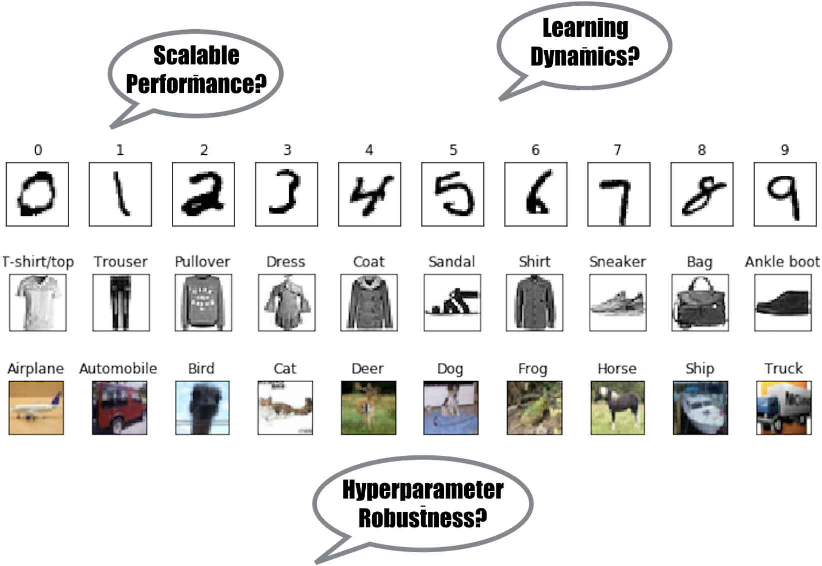
$$:= \alpha^{lf}(t)$$

$$\sigma\left(\frac{1}{\Delta t_2}\int_{t_2-\Delta t_2}^{t_2} V_i^{la}(t)dt\right)$$

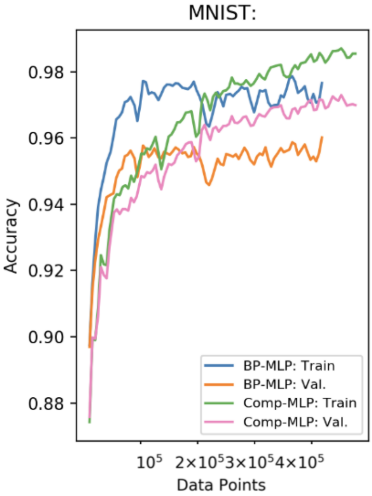
$$:= \alpha^{lt}(t)$$

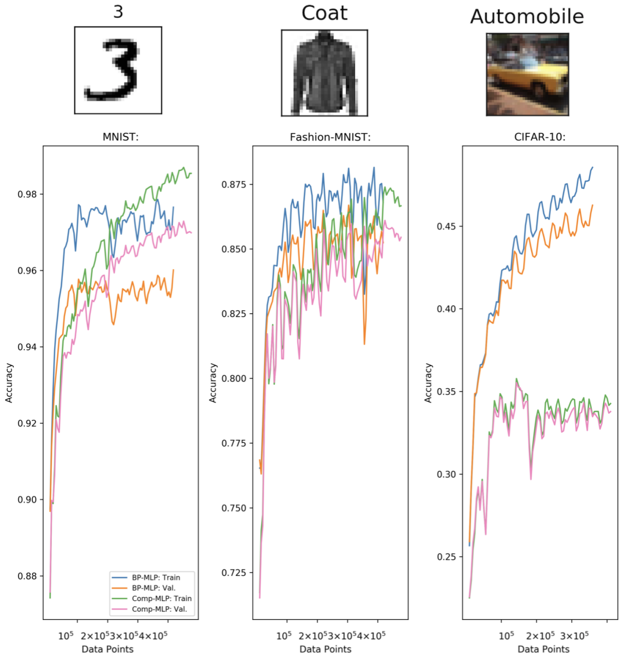
Forward Plateau Potential ⟷ Target Plateau Potential

Empirical Investigations

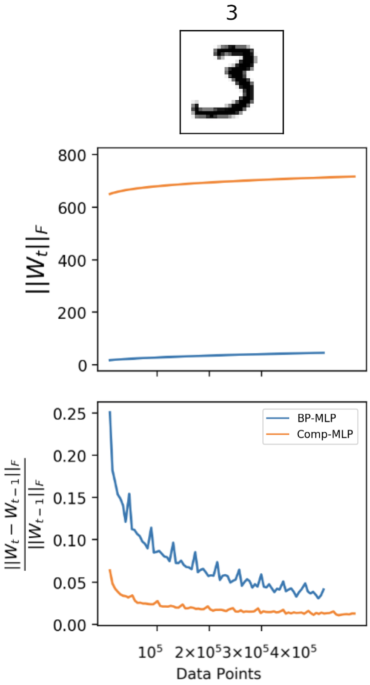


3

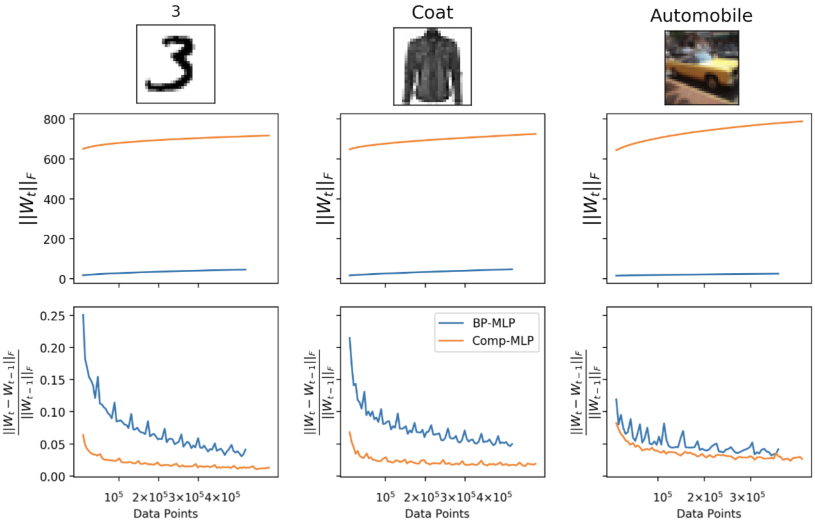

MNIST:
Accuracy
0.98
0.96
0.94
0.92
0.90
0.88
10^5
2×10^5
3×10^5
4×10^5
Data Points
BP-MLP: Train
BP-MLP: Val.
Comp-MLP: Train
Comp-MLP: Val.

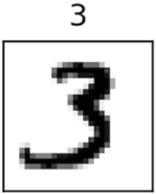


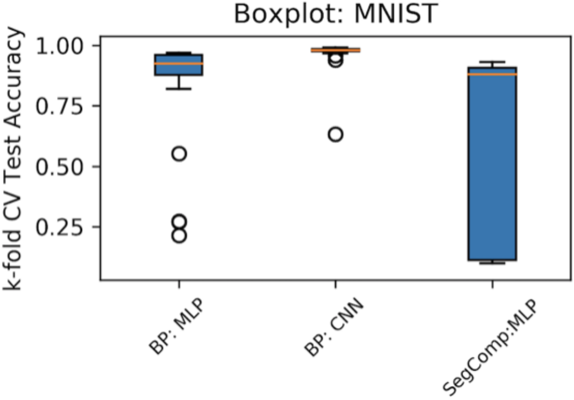
3
Coat
Automobile
MNIST:
Fashion-MNIST:
CIFAR-10:
Accuracy
Data Points
BP-MLP: Train
BP-MLP: Val.
Comp-MLP: Train
Comp-MLP: Val.

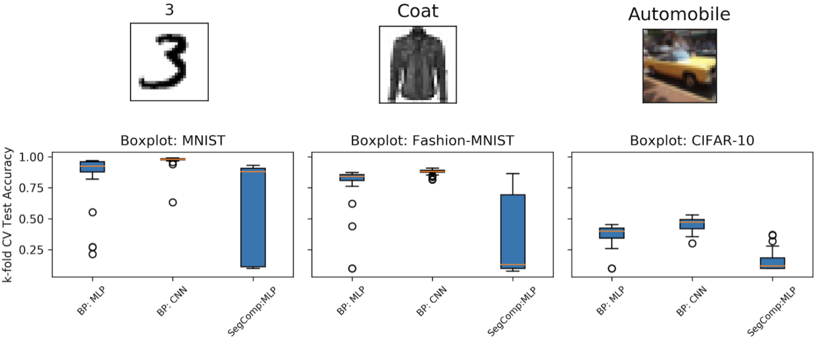


3
800
600
400
200
0
$||W_t||_F$
0.25
0.20
0.15
0.10
0.05
0.00
$\frac{||W_t - W_{t-1}||_F}{||W_{t-1}||_F}$
BP-MLP
Comp-MLP
10^5
2×10^5
3×10^5
4×10^5
Data Points


3
Coat
Automobile
$||W_t||_F$
800
600
400
200
0
0.25
0.20
0.15
0.10
0.05
0.00
$\frac{||W_t - W_{t-1}||_F}{||W_{t-1}||_F}$
BP-MLP
Comp-MLP
10^5 2×10^5 3×10^5 4×10^5
Data Points


3


Boxplot: MNIST
k-fold CV Test Accuracy
1.00
0.75
0.50
0.25
BP: MLP
BP: CNN
SegComp:MLP


3
Coat
Automobile
Boxplot: MNIST
Boxplot: Fashion-MNIST
Boxplot: CIFAR-10
k-fold CV Test Accuracy
1.00
0.75
0.50
0.25
BP: MLP
BP: CNN
SegComp:MLP

Summary ... and where to go?

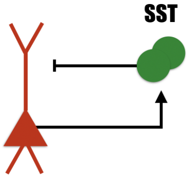
- ☑ Electrical segregation $\Rightarrow$ Near-continuous time
- ☒ **Computational** and **Physiological** Problems

Summary ... and where to go?

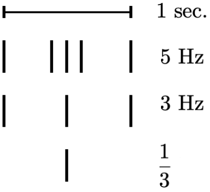
☑ Electrical segregation ⇒ Near-continuous time

☒ **Computational** and **Physiological** Problems

→ Sacramento *et al.* (2018): Local error from mismatch with local **interneurons**



→ Naud & Sprekeler (2018): **Multiplexing**



References I

Guerguiev, Jordan, Lillicrap, Timothy P, & Richards, Blake A. 2017. Towards deep learning with segregated dendrites. *ELife*, **6**, e22901.

Naud, Richard, & Sprekeler, Henning. 2018. Sparse bursts optimize information transmission in a multiplexed neural code. *Proceedings of the National Academy of Sciences*, 201720995.

Sacramento, João, Costa, Rui Ponte, Bengio, Yoshua, & Senn, Walter. 2018. Dendritic cortical microcircuits approximate the backpropagation algorithm. *Pages 8735–8746 of: Advances in Neural Information Processing Systems*.

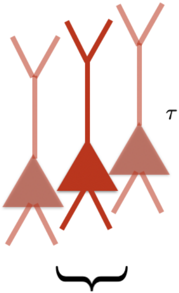
Biologically Implausible Backpropagation

→ MLP: $h_l := f(h_{l-1}; \theta_l) = \sigma_l(W_l h_{l-1} + b_l), \ \ l = 1, \dots, L$

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \theta_l} &= \left(\frac{dh_l}{d\theta_l}\right)^T \frac{\partial \mathcal{L}}{\partial h_l} = \left(\frac{dh_l}{d\theta_l}\right)^T \left(\frac{dh_{l+1}}{dh_l}\right)^T \frac{\partial \mathcal{L}}{\partial h_{l+1}} \\
&= \left(\frac{dh_l}{d\theta_l}\right)^T \underbrace{\left(\boxed{W_{l+1}} diag\left(\sigma'_{l+1}(\boxed{W_{l+1}} h_l + b_{l+1})\right)\right)^T}_{:=\delta_{l+1}} \frac{\partial \mathcal{L}}{\partial h_{l+1}} \\
&= \left(\frac{dh_l}{d\theta_l}\right)^T \left(\prod_{i=l+1}^{L} \delta_i\right) \frac{\partial \mathcal{L}}{\partial h_L}
\end{aligned}
$$

Weight Transport Problem

Guerguiev *et al.* (2017) - Hidden Layer Structure



n^l

$$\mathbf{V}^{la}(t) \leftarrow \mathbf{Y}^l\mathbf{s}^{l+1} \longleftrightarrow s^{l+1} \sim Poisson\left(\phi_{max}\sigma(V^{l+1}(t))\right)$$

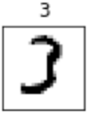
$$\downarrow$$

$$\tau\frac{dV_i^l(t)}{dt} = -V_i^l(t) + \frac{g_b}{g_{leak}}\left(V_i^{lb}(t) - V_i^l(t)\right) + \frac{g_a}{g_{leak}}\left(V_i^{la}(t) - V_i^l(t)\right)$$

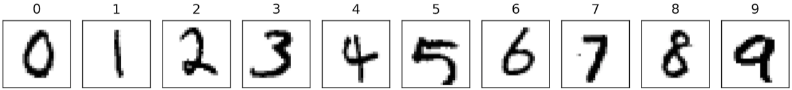
$$\uparrow$$

$$\mathbf{V}^{lb}(t) \leftarrow \mathbf{W}^l\mathbf{s}^{l-1} + \mathbf{b}^l$$

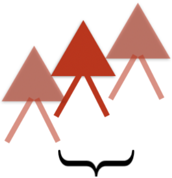
$$s_0 :\sim Poisson\left(\mathbf{x}\right)$$



Guerguiev *et al.* (2017) - Output Layer Structure



$$\phi_i^L(t) = \phi_{max}\sigma(V_i^L(t)) \text{ with } I_i(t) = 0, \; \forall \; i = 1, \ldots, n_L$$



n^L

$$\tau \frac{dV_i^L(t)}{dt} = -V_i^L(t) + \frac{g_d}{g_{leak}}\left(V_i^{Lb}(t) - V_i^L(t)\right) + I_i(t)$$

$$\mathbf{V}^{Lb}(t) \;\longleftarrow\; W^L s^{L-1} + b^L$$

Guerguiev *et al.* (2017) - Learning

Forward phase activity $\Leftrightarrow$ Target phase activity

$\Rightarrow$ Output Layer:

- Target firing rates: $\phi_i^{L\star} = \frac{1}{\Delta t_2} \int_{t_1+\Delta t_s}^{t_2} \phi_i^L(t)dt$
- Loss function:

$$L^L = ||\phi^{L\star} - \bar{\phi}^{Lf}||_2^2 = ||\frac{1}{\Delta t_2} \int_{t_1+\Delta t_s}^{t_2} \phi^L(t)dt - \frac{1}{\Delta t_1} \int_{t_0+\Delta t_s}^{t_1} \phi^L(t)dt||_2^2$$

$\Rightarrow$ Hidden Layer:

- Target firing rates: $\phi_i^{l\star} = \bar{\phi}_i^{lf} + \alpha_i^{lt} - \alpha_i^{lf}$
- Loss function:

$$L^l = ||\phi^{l\star} - \bar{\phi}^{lf}||_2^2 = ||\alpha^{lt} - \alpha^{lf}||_2^2$$

$\Rightarrow$ Local error minimization via gradient descent