Biologically Plausible Deep Learning: A Critical Review

Robert Tjarko Lange[1]
rtl17@ic.ac.uk
www.rob-lange.com

Einstein Center for Neurosciences Berlin

January 21, 2019

[1] https://github.com/RobertTLange/Bio-Plausible-DeepLearning

Motivation - Backpropagation

→ MLP: Composition of layers $\{h_l\}_{l=1}^{L}$, $h_0 = x$, $\theta_l = \{W_l, b_l\}$ and "Learn" synaptic weights, $\Theta = \{\theta_l\}_{l=1}^{L}$ iteratively.

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

$$\min_{\theta} \mathcal{L}(h_L|\Theta) = -\sum_{y} q(y|x) \log p(y|h_L; \Theta)$$

Motivation - Backpropagation

→ MLP: Composition of layers $\{h_l\}_{l=1}^{L}$, $h_0 = x$, $\theta_l = \{W_l, b_l\}$ and "Learn" synaptic weights, $\Theta = \{\theta_l\}_{l=1}^{L}$ iteratively.

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

$$\min_{\theta} \mathcal{L}(h_L|\Theta) = -\sum_{y} q(y|x) \log p(y|h_L; \Theta)$$

→ Backpropagation:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \theta_l} &= \left(\frac{dh_l}{d\theta_l}\right)^T \frac{\partial \mathcal{L}}{\partial h_l} = \left(\frac{dh_l}{d\theta_l}\right)^T \left(\frac{dh_{l+1}}{dh_l}\right)^T \frac{\partial \mathcal{L}}{\partial h_{l+1}} \\
&= \left(\frac{dh_l}{d\theta_l}\right)^T \underbrace{\left(W_{l+1} diag\left(\sigma'_{l+1}(W_{l+1}h_l + b_{l+1})\right)\right)^T}_{:=\delta_{l+1}} \frac{\partial \mathcal{L}}{\partial h_{l+1}} \\
&= \left(\frac{dh_l}{d\theta_l}\right)^T \delta_{l+1}\delta_{l+2} \frac{\partial \mathcal{L}}{\partial h_{l+2}} = \dots \\
&= \left(\frac{dh_l}{d\theta_l}\right)^T \left(\prod_{i=l+1}^{L} \delta_i\right) \frac{\partial \mathcal{L}}{\partial h_L}
\end{aligned}$$

Motivation - Problems with Backpropagation

$$\frac{\partial \mathcal{L}}{\partial \theta_l} = \left(\frac{dh_l}{d\theta_l}\right)^T \left(\prod_{i=l+1}^{L} \delta_i\right) \frac{\partial \mathcal{L}}{\partial h_L}$$

✗ **Weight Transport Problem**

Motivation - Problems with Backpropagation

$$\frac{\partial \mathcal{L}}{\partial \theta_l} = \left(\frac{dh_l}{d\theta_l}\right)^T \left(\prod_{i=l+1}^{L} \delta_i\right) \frac{\partial \mathcal{L}}{\partial h_L}$$

- ☒ **Weight Transport Problem**
- ☒ **Global signed error signal**

Motivation - Problems with Backpropagation

$$\frac{\partial \mathcal{L}}{\partial \theta_l} = \left(\frac{dh_l}{d\theta_l}\right)^T \left(\prod_{i=l+1}^{L} \delta_i\right) \frac{\partial \mathcal{L}}{\partial h_L}$$

- ✗ **Weight Transport Problem**
- ✗ **Global signed error signal**
- ✗ **Computationally Expensive Matrix Transposition**

Motivation - Problems with Backpropagation

$$\frac{\partial \mathcal{L}}{\partial \theta_l} = \left(\frac{dh_l}{d\theta_l}\right)^T \left(\prod_{i=l+1}^{L} \delta_i\right) \frac{\partial \mathcal{L}}{\partial h_L}$$

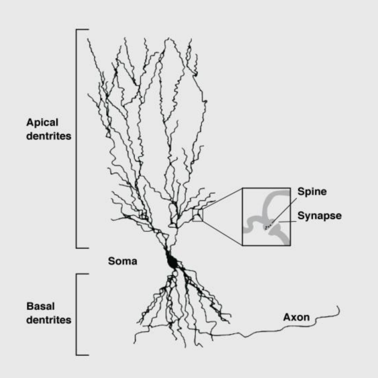
- ☒ **Weight Transport Problem**
- ☒ **Global signed error signal**
- ☒ **Computationally Expensive Matrix Transposition**
- ☐ Alternatives:
  - Feedback Alignment (Lillicrap *et al.* , 2016)
  - Target Propagation (Lee *et al.* , 2014)

Motivation - Problems with Backpropagation

$$\frac{\partial \mathcal{L}}{\partial \theta_l} = \left(\frac{dh_l}{d\theta_l}\right)^T \left(\prod_{i=l+1}^{L} \delta_i\right) \frac{\partial \mathcal{L}}{\partial h_L}$$

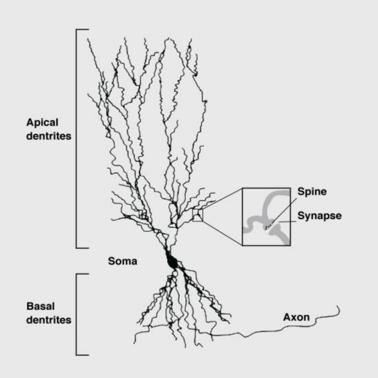
- ☒ **Weight Transport Problem**
- ☒ **Global signed error signal**
- ☒ **Computationally Expensive Matrix Transposition**
- ☐ Alternatives:
  - Feedback Alignment (Lillicrap *et al.* , 2016)
  - Target Propagation (Lee *et al.* , 2014)
- ☐ Problems with Alternatives:
  - Need form of info transmission to determine local errors
  - Not possible in single compartment neurons without feedback pathway

Motivation - Electrical Segregation of ↓ and ↑ Info



- → Körding & König (2001): Local error computation via electrical segregation
- → Multi-compartmental segregation avoids need for feedback pathway
- → Apical dendrites ($\Downarrow$)
- → Basal dendrites ($\Uparrow$)

Motivation - Electrical Segregation of ↓ and ↑ Info



→ Körding & König (2001): Local error computation via electrical segregation

→ Multi-compartmental segregation avoids need for feedback pathway

→ Apical dendrites (⇓)

→ Basal dendrites (⇑)

→ **Plateau Potentials**:

- Apical ⇒ soma via voltage-gate Ca^{2+} channels
- Prolonged upswing in MP due to events in apical shaft

⇒ Can guide plasticity in pyramidal neurons

Guerguiev *et al.* (2017) - Neuron and Network Model

→ 3 Compartment Hidden Layer: $\mathbf{V}^{0a}(t), \mathbf{V}^{0b}(t), \mathbf{V}^{0}(t) \in \mathbb{R}^m$

$$\tau \frac{dV_i^0(t)}{dt} = -V_i^0(t) + \frac{g_b}{g_l}\left(V_i^{0b}(t) - V_i^0(t)\right) + \frac{g_a}{g_l}\left(V_i^{0a}(t) - V_i^0(t)\right)$$

$$V_i^{0b} = \sum_{j=1}^{l} W_{ij}^0 s_j^{input}(t) + b_i^0 \quad \text{and} \quad V_i^{0a} = \sum_{j=1}^{n} Y_{ij} s_j^1(t)$$

Guerguiev *et al.* (2017) - Neuron and Network Model

$\rightarrow$ 3 Compartment Hidden Layer: $\mathbf{V}^{0a}(t), \mathbf{V}^{0b}(t), \mathbf{V}^{0}(t) \in \mathbb{R}^m$

$$\tau \frac{dV_i^0(t)}{dt} = -V_i^0(t) + \frac{g_b}{g_l}\left(V_i^{0b}(t) - V_i^0(t)\right) + \frac{g_a}{g_l}\left(V_i^{0a}(t) - V_i^0(t)\right)$$

$$V_i^{0b} = \sum_{j=1}^{l} W_{ij}^0 s_j^{input}(t) + b_i^0 \quad \text{and} \quad V_i^{0a} = \sum_{j=1}^{n} Y_{ij} s_j^1(t)$$

$\rightarrow$ 2 Compartment Output Layer: $\mathbf{V}^{1b}(t), \mathbf{V}^{1}(t) \in \mathbb{R}^n$

$$\tau \frac{dV_i^1(t)}{dt} = -V_i^1(t) + \frac{g_d}{g_l}\left(V_i^{1b}(t) - V_i^1(t)\right) + I_i(t)$$

$$V_i^{1b} = \sum_{j=1}^{l} W_{ij}^1 s_j^0(t) + b_i^1$$

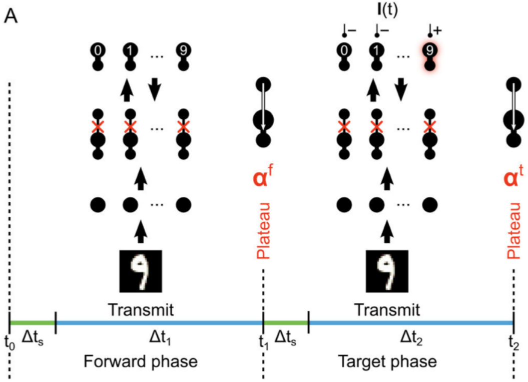
$\rightarrow$ $s_j^{input}(t) = \sum_k \kappa(t - t_{jk}^{input})$ with κ response kernel

Guerguiev *et al.* (2017) - Credit Assignment Signals

- $\rightarrow$ **Forward** $(t_0 + \Delta t_s \rightarrow t_1)$: $I_i(t) = 0, \ \forall i = 1, \ldots, n$
  - At t_1: $\alpha_i^f = \sigma\left(\frac{1}{\Delta t_1}\int_{t_1-\Delta t_1}^{t_1} V_i^{0a}(t)dt\right)$
- $\rightarrow$ **Target** $(t_1 + \Delta t_s \rightarrow t_2)$: $I_k(t) = \phi_{max}$ for $y_{sample} = k$
  - At t_2: $\alpha_i^t = \sigma\left(\frac{1}{\Delta t_2}\int_{t_2-\Delta t_2}^{t_2} V_i^{0a}(t)dt\right)$

Guerguiev *et al.* (2017) - Credit Assignment Signals

- → **Forward** $(t_0 + \Delta t_s \rightarrow t_1)$: $I_i(t) = 0, \ \forall i = 1, \dots, n$
  - At t_1: $\alpha_i^f = \sigma\left(\frac{1}{\Delta t_1}\int_{t_1 - \Delta t_1}^{t_1} V_i^{0a}(t)dt\right)$
- → **Target** $(t_1 + \Delta t_s \rightarrow t_2)$: $I_k(t) = \phi_{max}$ for $y_{sample} = k$
  - At t_2: $\alpha_i^t = \sigma\left(\frac{1}{\Delta t_2}\int_{t_2 - \Delta t_2}^{t_2} V_i^{0a}(t)dt\right)$



Guerguiev *et al.* (2017) - Learning

$\rightarrow$ How is learning defined in such a model?

Forward phase dynamics $\Leftrightarrow$ Target phase dynamics

Guerguiev *et al.* (2017) - Learning

$\rightarrow$ How is learning defined in such a model?

Forward phase dynamics $\Leftrightarrow$ Target phase dynamics

$\Rightarrow$ Somatic compartments generate Poisson process spikes:

- Rates-of-fire: $\phi_i^l(t) = \phi_{max}\sigma(V_i^l(t))$

Guerguiev *et al.* (2017) - Learning

→ How is learning defined in such a model?

> Forward phase dynamics ⇔ Target phase dynamics

⇒ Somatic compartments generate Poisson process spikes:

- Rates-of-fire: $\phi_i^l(t) = \phi_{max}\sigma(V_i^l(t))$

⇒ Output Layer:

- Target firing rates: $\phi_i^{1\star} = \frac{1}{\Delta t_2}\int_{t_1+\Delta t_s}^{t_2} \phi_i^1(t)dt$
- Loss function: $L^1 = ||\phi_i^{1\star} - \bar{\phi}_i^{1f}||_2^2 =$
  $||\frac{1}{\Delta t_2}\int_{t_1+\Delta t_s}^{t_2} \phi_i^1(t)dt - \frac{1}{\Delta t_1}\int_{t_0+\Delta t_s}^{t_1} \phi_i^1(t)dt||_2^2$

Guerguiev *et al.* (2017) - Learning

→ How is learning defined in such a model?

Forward phase dynamics ⇔ Target phase dynamics

⇒ Somatic compartments generate Poisson process spikes:

- Rates-of-fire: $\phi_i^l(t) = \phi_{max}\sigma(V_i^l(t))$

⇒ Output Layer:

- Target firing rates: $\phi_i^{1\star} = \frac{1}{\Delta t_2}\int_{t_1+\Delta t_s}^{t_2} \phi_i^1(t)dt$
- Loss function: $L^1 = ||\phi_i^{1\star} - \bar{\phi}_i^{1f}||_2^2 =$ $||\frac{1}{\Delta t_2}\int_{t_1+\Delta t_s}^{t_2} \phi_i^1(t)dt - \frac{1}{\Delta t_1}\int_{t_0+\Delta t_s}^{t_1} \phi_i^1(t)dt||_2^2$

⇒ Hidden Layer:

- Target firing rates: $\phi_i^{0\star} = \bar{\phi}_i^{0f} + \alpha_i^t - \alpha_i^f$
- Loss function: $L^0 = ||\phi_i^{0\star} - \bar{\phi}_i^{0f}||_2^2 = ||\alpha^t - \alpha^f||_2^2$

Guerguiev *et al.* (2017) - Learning

$\rightarrow$ How is learning defined in such a model?

Forward phase dynamics $\Leftrightarrow$ Target phase dynamics

$\Rightarrow$ Somatic compartments generate Poisson process spikes:

- Rates-of-fire: $\phi_i^l(t) = \phi_{max}\sigma(V_i^l(t))$

$\Rightarrow$ Output Layer:

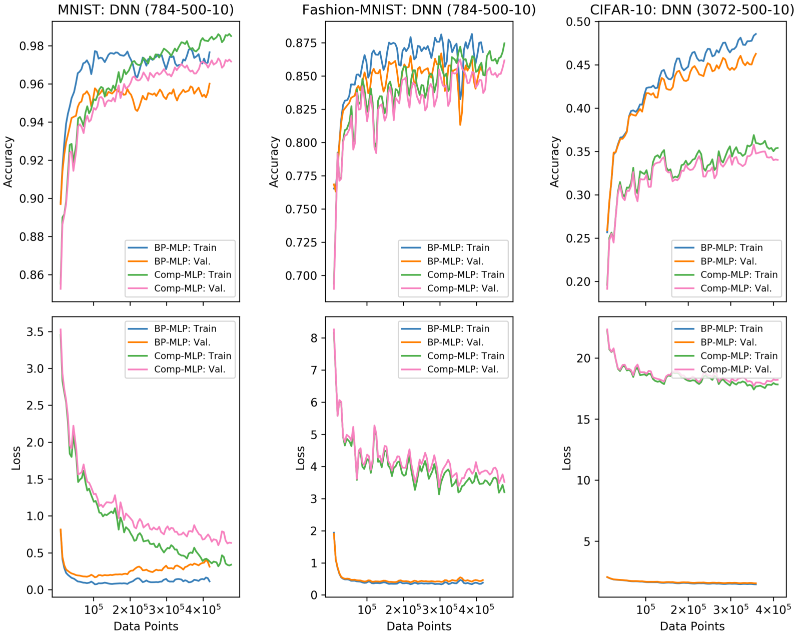
- Target firing rates: $\phi_i^{1\star} = \frac{1}{\Delta t_2}\int_{t_1+\Delta t_s}^{t_2} \phi_i^1(t)dt$
- Loss function: $L^1 = ||\phi_i^{1\star} - \bar{\phi}_i^{1f}||_2^2 =$
  $||\frac{1}{\Delta t_2}\int_{t_1+\Delta t_s}^{t_2} \phi_i^1(t)dt - \frac{1}{\Delta t_1}\int_{t_0+\Delta t_s}^{t_1} \phi_i^1(t)dt||_2^2$

$\Rightarrow$ Hidden Layer:

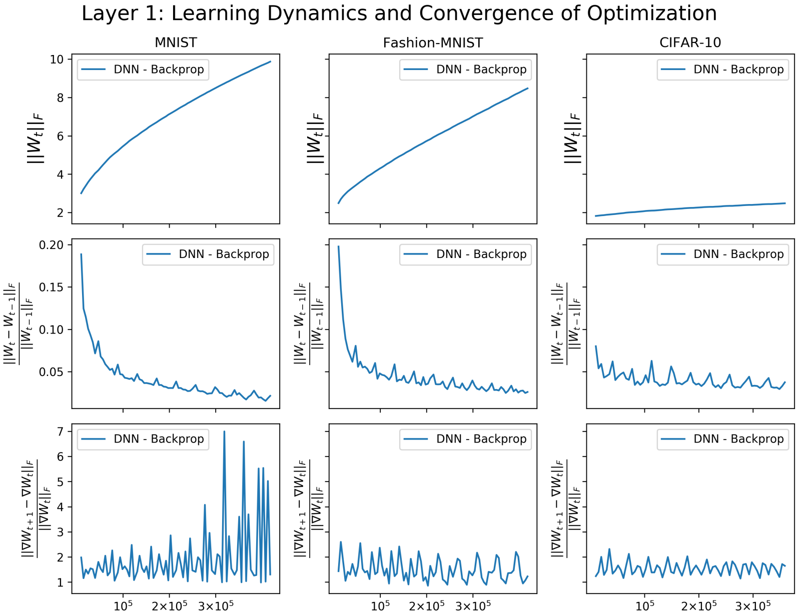
- Target firing rates: $\phi_i^{0\star} = \bar{\phi}_i^{0f} + \alpha_i^t - \alpha_i^f$
- Loss function: $L^0 = ||\phi_i^{0\star} - \bar{\phi}_i^{0f}||_2^2 = ||\alpha^t - \alpha^f||_2^2$

$\Rightarrow$ Local error minimization via SGD

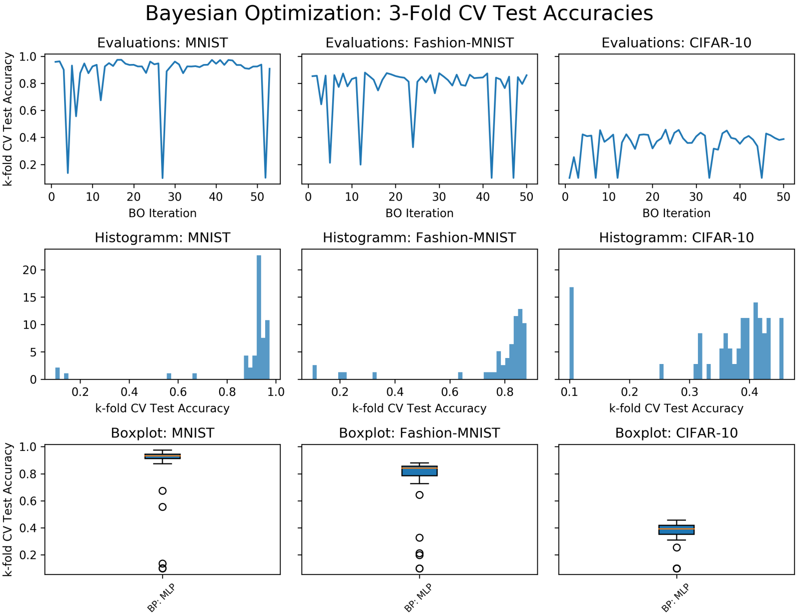
Experiments - Learning Dynamics: Performance



Experiments - Learning Dynamics: Dynamics



Experiments - Learning Dynamics: Robustness



Guerguiev *et al.* (2017) - Accomplishments/Problems

- ☑ Segregated compartments generate local targets that act as credit assignment signals in a physiologically plausible manner
- ☑ Signal can be used to exploit depth in near-continuous time
- ☒ **Computational** Problems
  - $\rightarrow$ Huge hyperparameter space $\rightarrow$ most likely not robust!
- ☒ **Physiological** Problems
  - $\rightarrow$ How is the teaching signal internally generated?
  - $\rightarrow$ 2 global phases? - Length sampled from inverse Gaussian
  - $\rightarrow$ Stoch. gen. of plateau potentials - apical calcium spikes

$\Rightarrow$ Sacramento *et al.* (2018): Neocortical micro-circuits and inhibitory interneurons might act synchronizing.

Literature Review

	Backprop (Rummelhart et al., 1986)	Feedback Alignment (Lillicrap et al., 2016)	Target Propagation (LeCun, 1986)	Difference TP (Lee et al., 2015)	Simplified DTP (Bartunov et al., 2018)	Segregated Compartments (Guergiev et al., 2017)	Microcircuits (Sacramento et al., 2018)
Exact Gradients	✅	❌	❌	❌	❌	❌	✅ (In Limit)
No Weight Transport	❌	✅	✅/❌ (Final Layer)	✅/❌ (Final Layer)	✅	✅	✅
No Separate Pathways	✅	❌	❌	❌	❌	✅	✅
Dendritic Integration	-	❌	❌	❌	❌	✅	✅
Separate Weights Learned	-	❌	✅	✅	✅	❌	✅/❌
Linear Stabilization	-	-	❌	✅	✅	-	-
Explicit Error Representation	✅	✅	✅	✅	✅	❌	✅

What do I want to analyze? What's next?

Literature Review

- ☑ Feedback Alignment (Lillicrap *et al.* , 2016)
- ☑ Target Propagation (Lee *et al.* , 2014; Bartunov *et al.* , 2018),
- ☑ Segregated Compartments (Guerguiev *et al.* , 2017; Sacramento *et al.* , 2018)

Implement different models/learning rules

- ☑ Standard Backprop MLP, CNN in PyTorch
- ☑ Segregated Compartment MLP in Numpy
- ☐ k-fold CV pipeline for SC MLP

Analyze learning dynamics

- ☐ $||W_t||_F$ - Overfitting? $||\Delta W_t||_F$ - Convergence? $||\Delta \nabla W_t||_F$ - Local optima? Feedback alignment - Inverse Jacobian
- ☐ Different SGD variants: Momentum, Adam, RMSprop

Analyze Hyperparameter/Dataset Robustness

- ☑ Different datasets (MNIST, Fashion, CIFAR-10)
- ☑ Bayesian Optimization for Hyperparam. Search Backprop
- ☐ Bayesian Optimization for Hyperparam. Search SC

References I

Bartunov, Sergey, Santoro, Adam, Richards, Blake, Marris, Luke, Hinton, Geoffrey E, & Lillicrap, Timothy. 2018. Assessing the scalability of biologically-motivated deep learning algorithms and architectures. *Pages 9389–9399 of: Advances in Neural Information Processing Systems*.

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

Körding, Konrad P, & König, Peter. 2001. Supervised and unsupervised learning with two sites of synaptic integration. *Journal of computational neuroscience*, **11**(3), 207–215.

Lee, Dong-Hyun, Zhang, Saizheng, Biard, Antoine, & Bengio, Yoshua. 2014. Target Propagation. *CoRR*, **abs/1412.7525**.

Lillicrap, Timothy P, Cownden, Daniel, Tweed, Douglas B, & Akerman, Colin J. 2016. Random synaptic feedback weights support error backpropagation for deep learning. *Nature communications*, **7**, 13276.

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*.