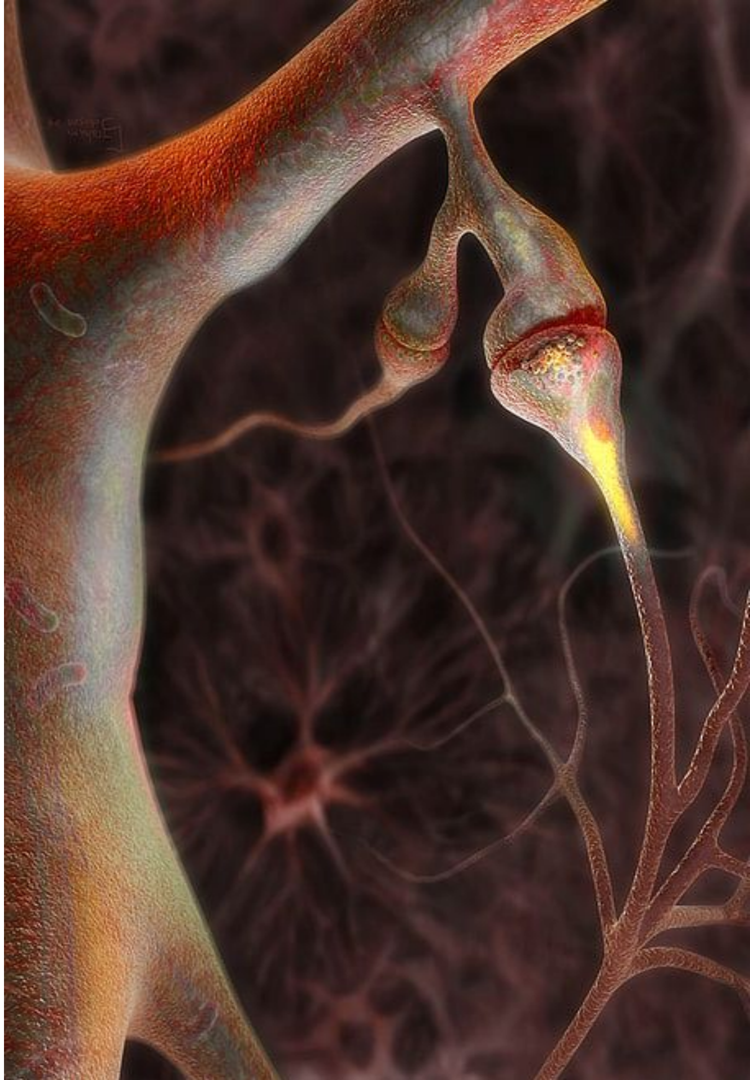
Minimizing Control for Credit Assignment with Strong Feedback

Alexander Meulemans*, Matilde Tristany Farinha*, Maria R. Cervera*,
João Sacramento, Benjamin F. Grewe
(* equal contribution)

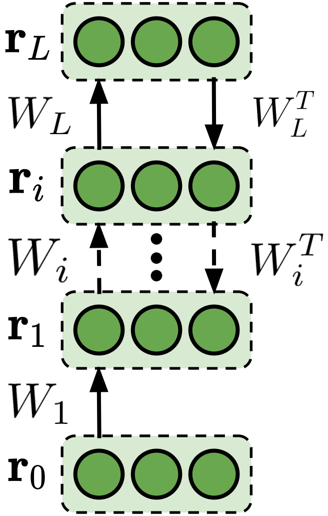
Outline

“ How does the strength of a synapse need to be changed to improve the system’s global behaviour?”

Spatial Credit Assignment



Spatial credit assignment: backpropagation



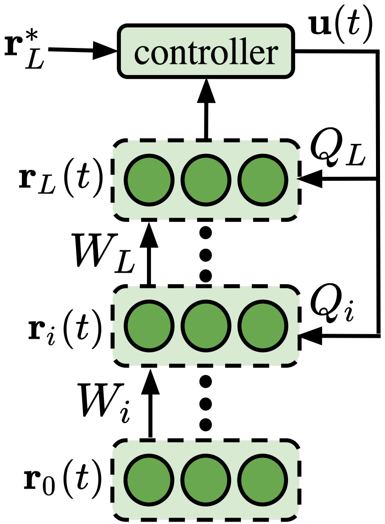
Some biological issues:

- Weight transport
- Feedback does not influence neural activity only synaptic strength

Research Question

"Is gradient-based **credit assignment** possible **without** the need for the **weak-feedback assumption**?"

Strong-DFC: intuition



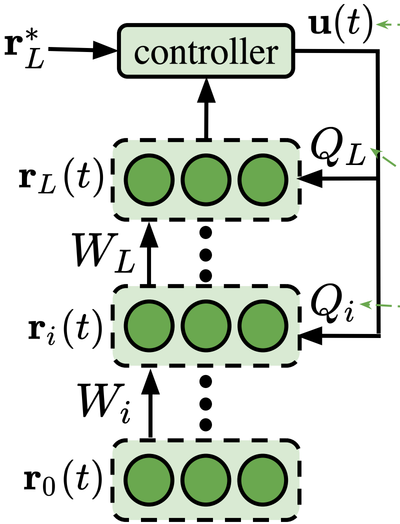
Usual approach:

- Use the error only for synaptic plasticity

New approach:

- Error adjusts the neural activations and drives the network's output to its supervised target

Strong-DFC: dynamics



Controller signal (feedback signal):

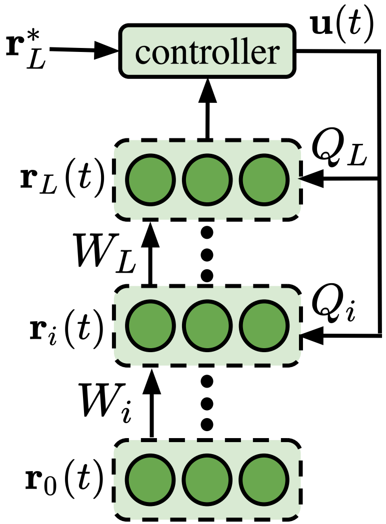
integrates the error signal

$$\mathbf{e}(t) = \mathbf{r}_L^* - \mathbf{r}_L(t), \quad \mathbf{r}_L^* = \mathbf{r}_L^{\text{true}}$$

Feedback weights

flexible feedback connectivity constraints

Strong-DFC: minimizing control

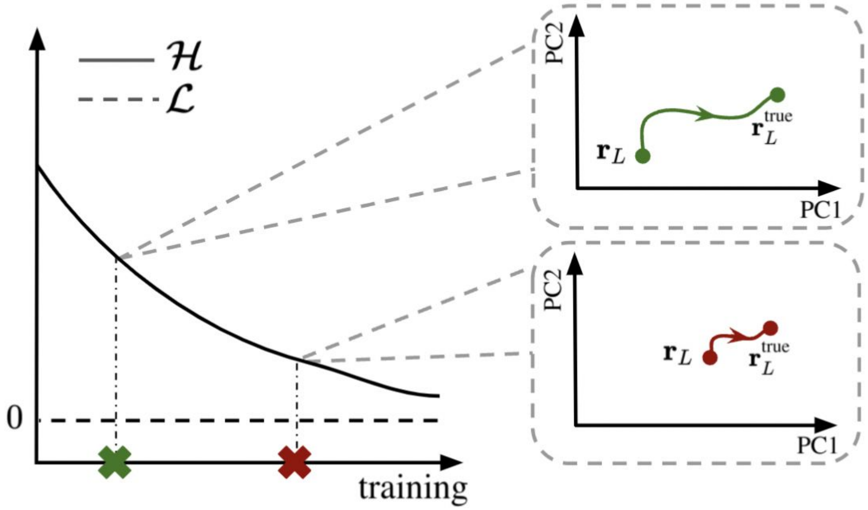


Amount of control: $\mathcal{H} = \|Q\mathbf{u}_{\text{ss}}\|_2^2$

"Every neuron tries to minimize its feedback input"

$$\Delta W_i \propto -\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}W_i}$$

Strong-DFC: minimizing control



Goal of training: update the forward weights to reduce the amount of control needed to reach $\mathbf{r}_L^{\text{true}}$

Minimizing $\mathcal{H}$ leads to minimizing $\mathcal{L}$:

$$\mathcal{H} = \|Q\mathbf{u}_{\text{ss}}\|_2^2 = 0 \iff \mathcal{L}(\mathbf{r}_L, \mathbf{r}_L^{\text{true}}) = 0$$

Strong-DFC: noise robustness

	MNIST (%)	Fashion MNIST (%)
BP	1.83 ± 0.11	10.60 ± 0.44
Strong-DFC (noise)	2.19 ± 0.05	12.07 ± 0.16
DFC (noise)	15.15 ± 0.44	16.29 ± 0.41
Strong-DFC (no noise)	1.98 ± 0.05	11.36 ± 0.17
DFC (no noise)	2.09 ± 0.10	11.31 ± 0.14

Conclusions

- **Principled credit assignment** with **strong feedback**
- **Flexible constraints** on the **feedback** mappings
- Enables **simultaneous learning** of **feedback** and **feedforward** weights
- **Novel** view on **optimization** as **minimizing control**
- It **works**!

Conclusions

- **Principled credit assignment** with **strong feedback**
- **Flexible constraints** on the **feedback** mappings
- Enables **simultaneous learning** of **feedback** and **feedforward** weights
- **Novel** view on **optimization** as **minimizing control**
- It **works**!

Thank you!

Discussion?

At ICML in person! :)
(also: mtristany@ethz.ch)