北京大学
PEKING UNIVERSITY

北京大学计算机学院
School of Computer Science

Global Credit Assignment via Dynamical Criticality

Wentao Wang

Research direction

北京大学 PEKING UNIVERSITY

Understanding the structure to function mechanism

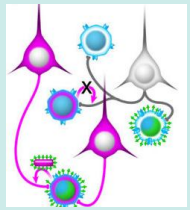
How?

Information processing

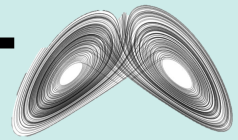
Brain and brain-inspired systems

Computer

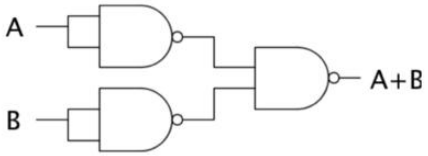
Structure



Dynamics

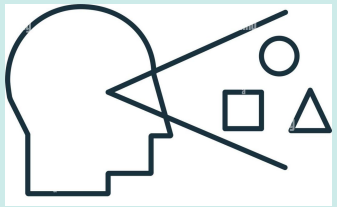


Physiology



Electronic circuit

Function



Cognitive computation, intelligence

$$\begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}$$

Arithmetic computation

Wu et al, Science Advances, 2025
Bi et al, iScience, 2025
Huang et al, NIPS, 2025
Liu et al, AAAI NeuroAI workshop,2026

Gao et al, ICLR, 2026
Ye et al, ICLR, 2026
Zhou et al, ICML, 2026
Wang et al, ICML, 2026

Chen, Scherr, Maass, Science Advances, 2022
Chen & Gong, Science Advances, 2022
Chen, Qu, Gong, Neural Networks, 2022
Chen & Gong, Nature Communications, 2019

Credit assignment

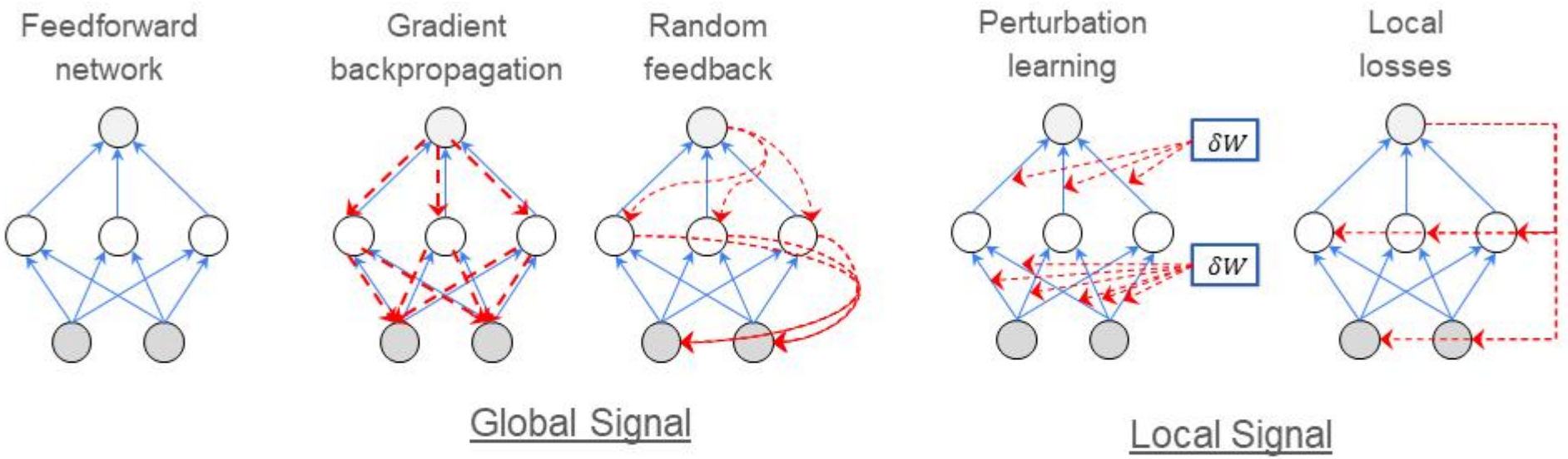
北京大学 PEKING UNIVERSITY

Question

How to set synaptic weights to achieve more ideal output?

Global learning

Local learning



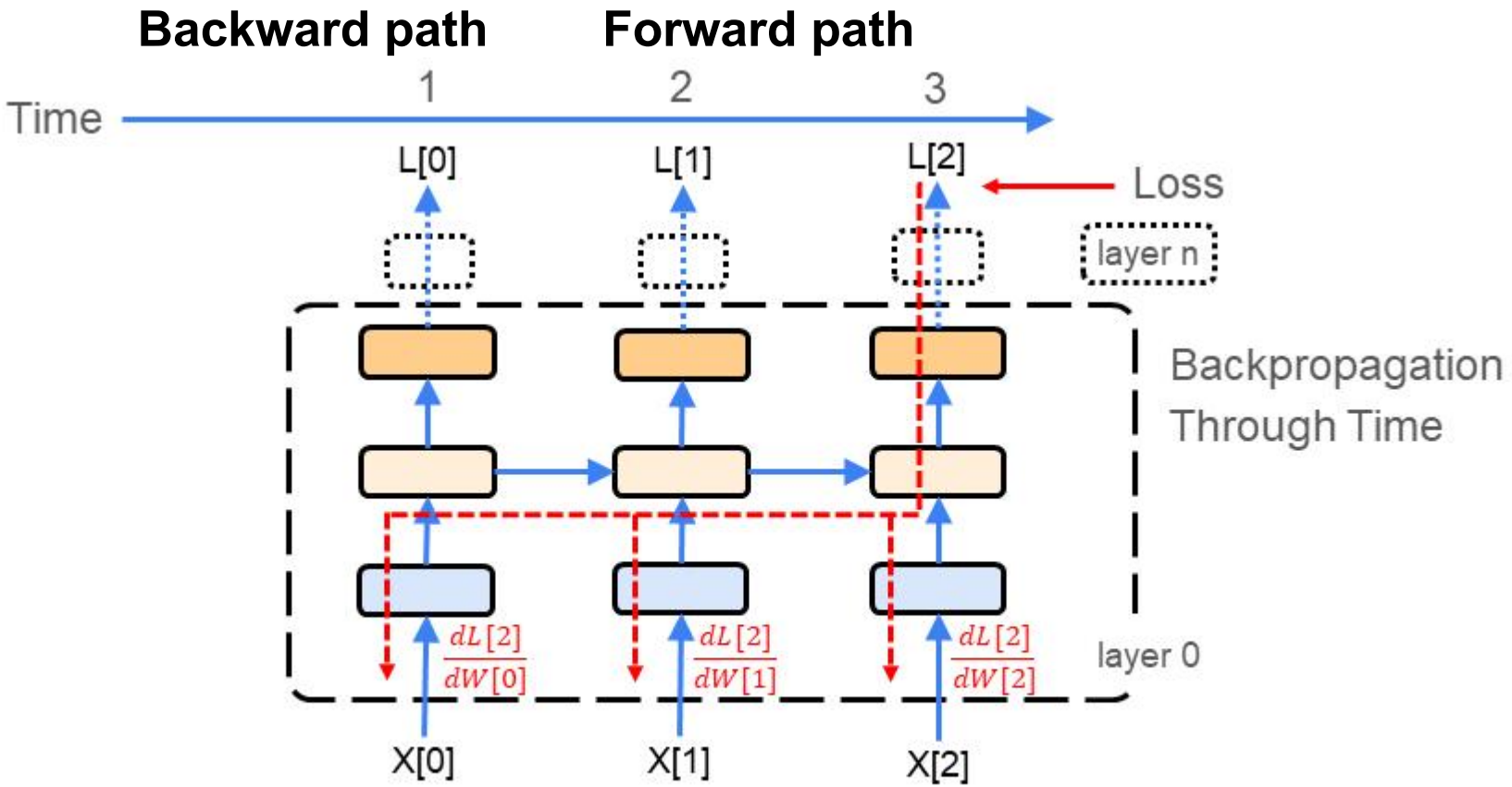
https://amassivek.github.io/sigprop

Temporal credit assignment

北京大学 PEKING UNIVERSITY

Question

Long backward path causes gradient unstable



https://amassivek.github.io/sigprop

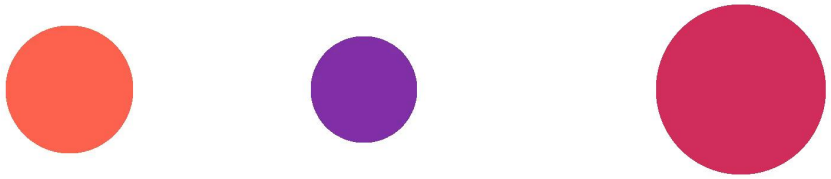
Feedforward vs recurrent: parallel and serial

北京大学 PEKING UNIVERSITY

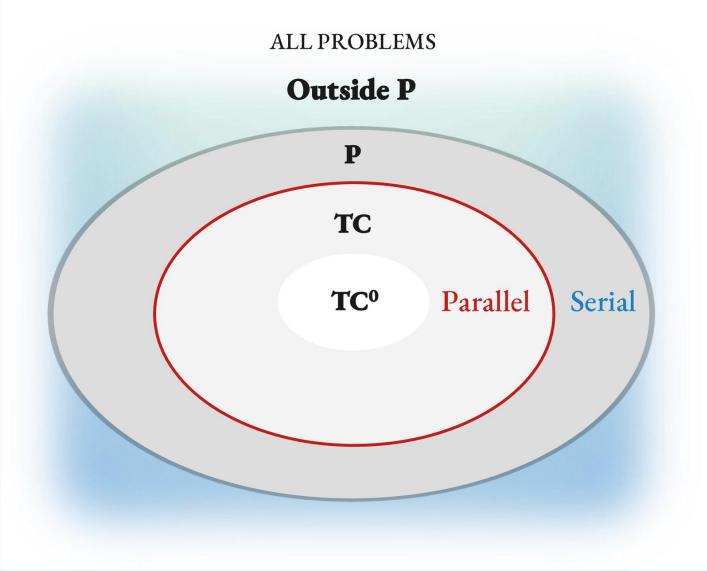
Importance

Not all problems can be solved in parallel

you can add more cooks to the kitchen, but the soup still needs time to simmer



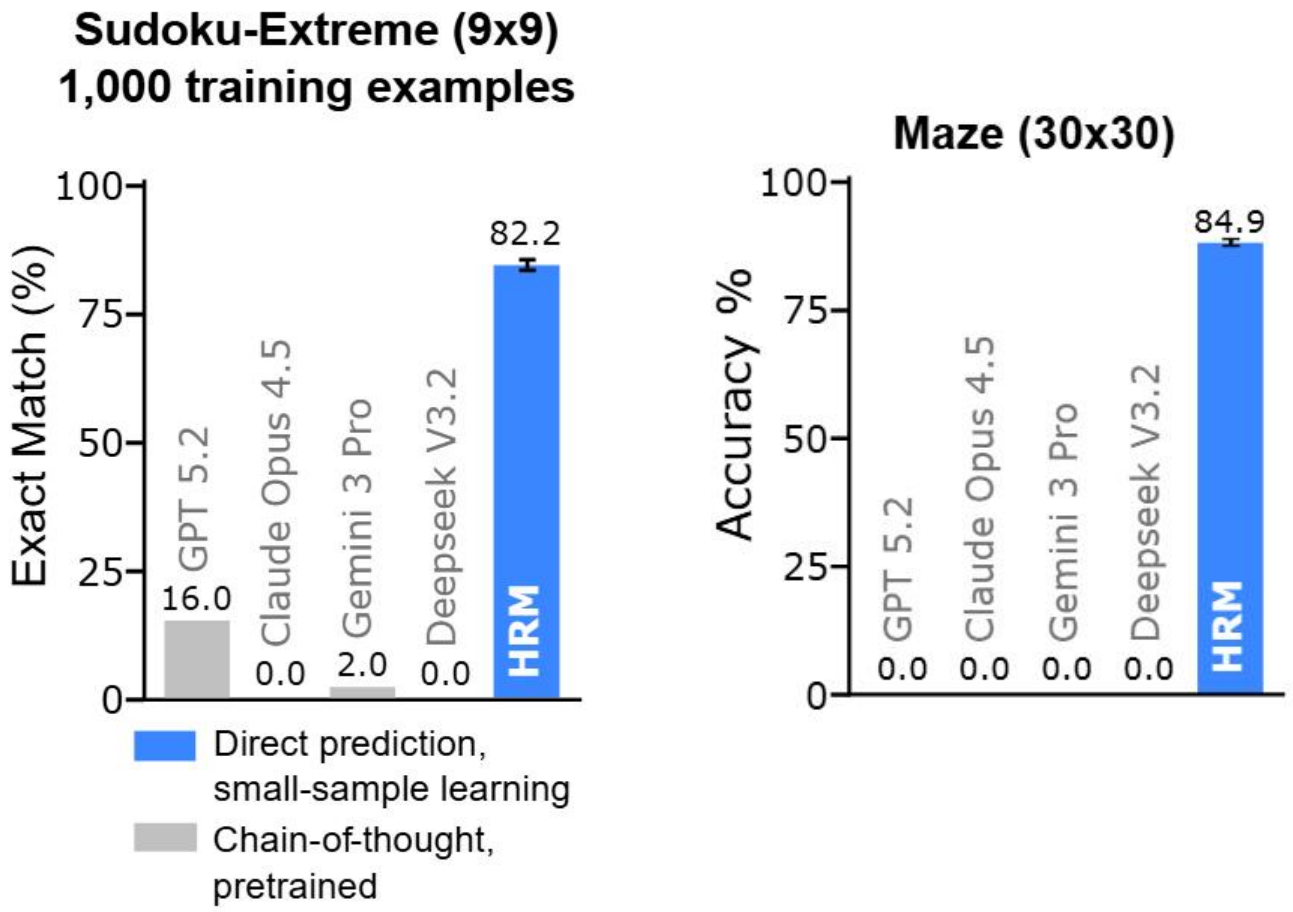
Even a handful of freely moving objects becomes genuinely hard to predict after a few seconds: every collision re-routes the entire future trajectory, and the state at any later time is the endpoint of a chain of dependent events.



Complexity theory demonstrates that a fixed- depth model cannot do better than TC^0, no matter how wide you make it.

Liu et al, ICLR, 2026

Recurrent model outperforms feedforward model in reasoning



Model size:
Deepseek R1 —— 671B
HRM —— 0.027B (1/25000)

Wang et al, under review

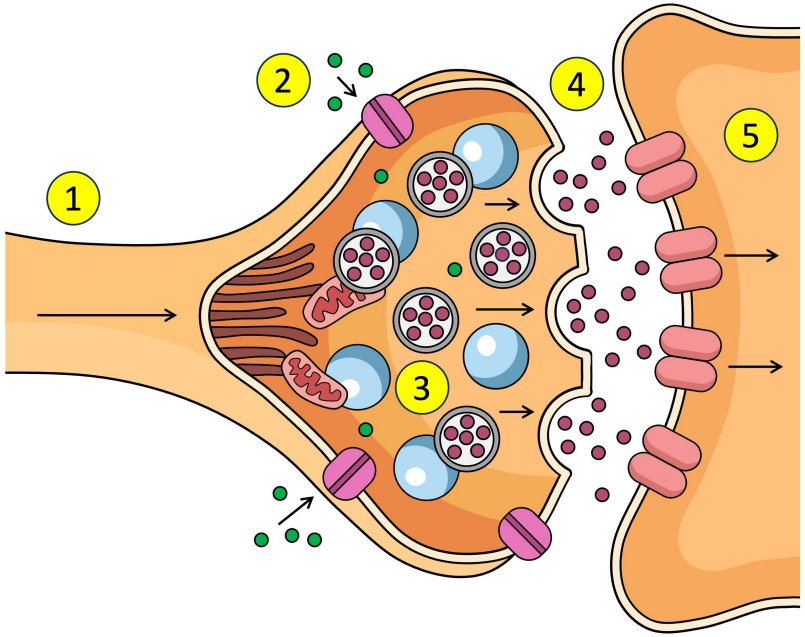
Important for the brain and brain-inspired AI

北京大学 PEKING UNIVERSITY

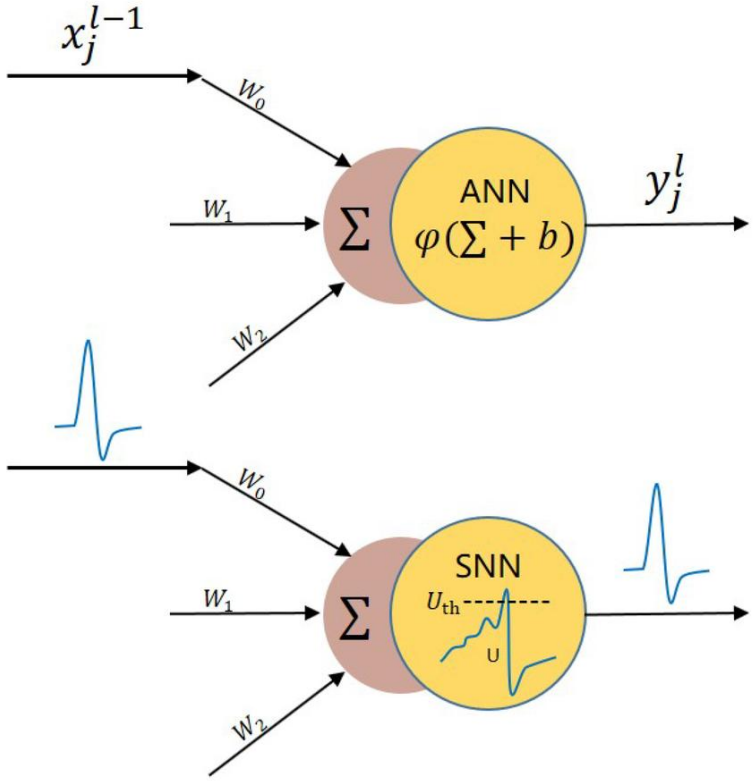
Question

The brain is doing BPTT?

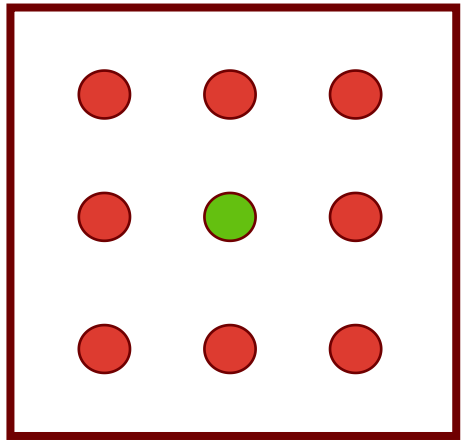
No way to do exact backpropagation



Strong nonlinearity is harmful



Dilemma of existing methods

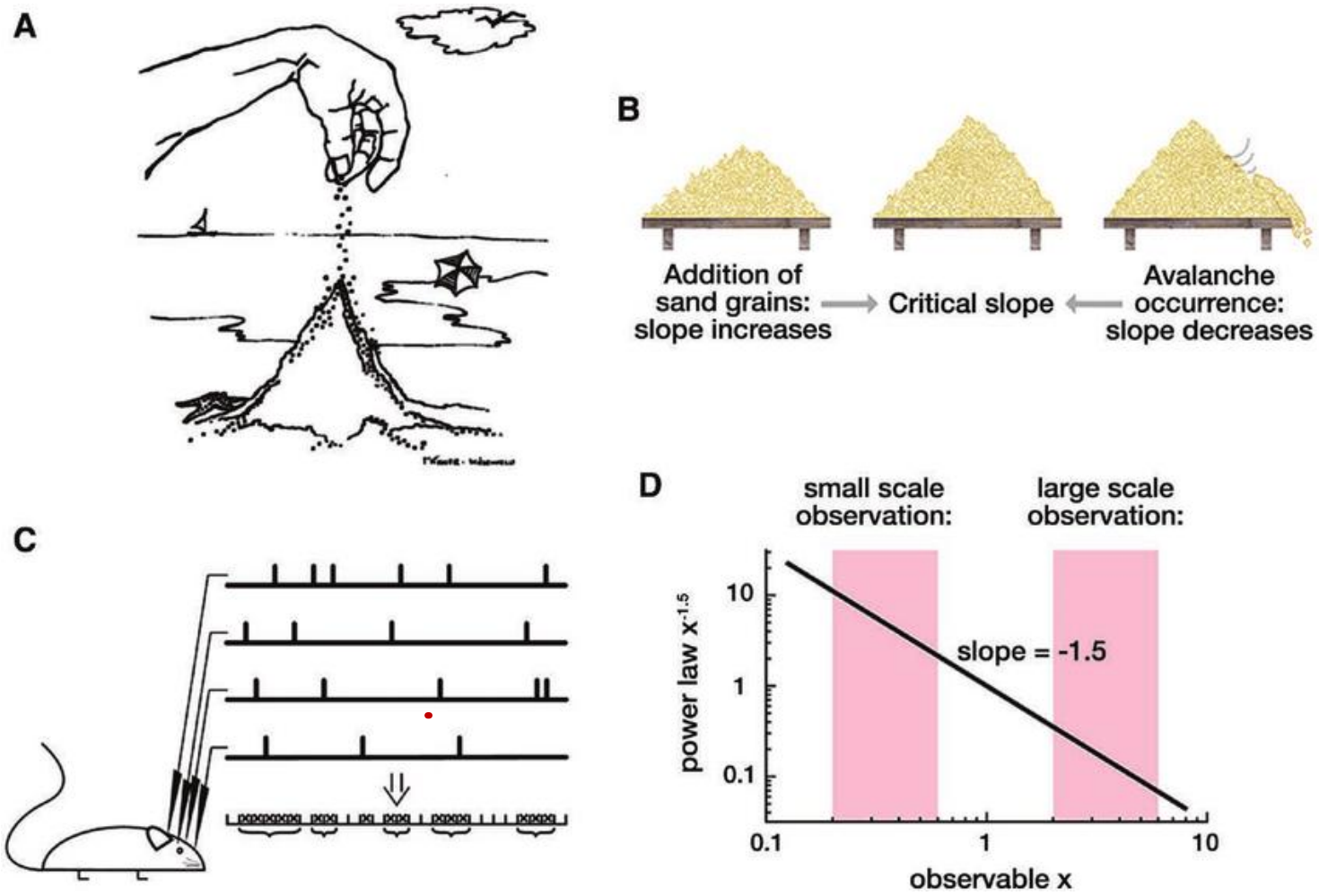


Backpropagation through time

The Dilemma of Existing Methods:

- BPTT: Precise gradients but requires unrolling the computational graph; memory scales linearly with sequence length (O(T)), prone to vanishing/exploding gradients.
- Online Learning: Computationally efficient (O(1) memory) but performing far worse than BPTT.
- Local Learning: Computationally efficient (good for chips) but performing far worse than BPTT.

Background: Criticality



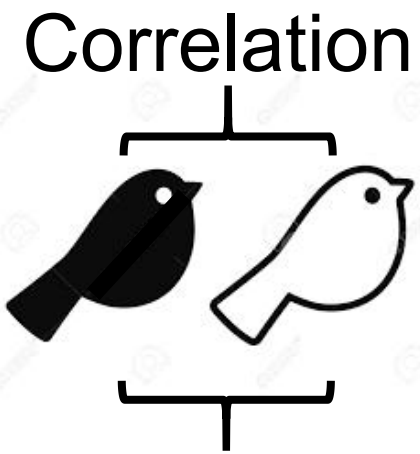
Lee and Mashour, Anesthesilogy, 2018

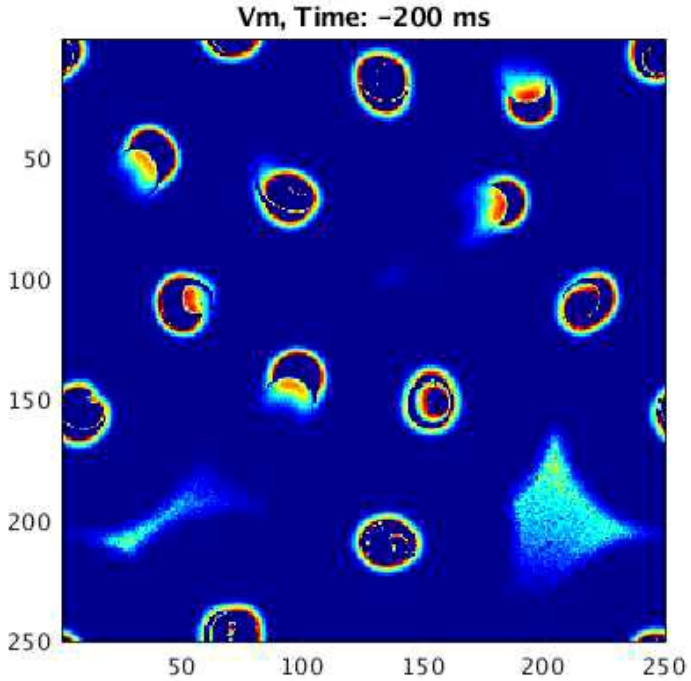
Scale-free property of power law

$$f(x) \propto x^{\alpha}$$

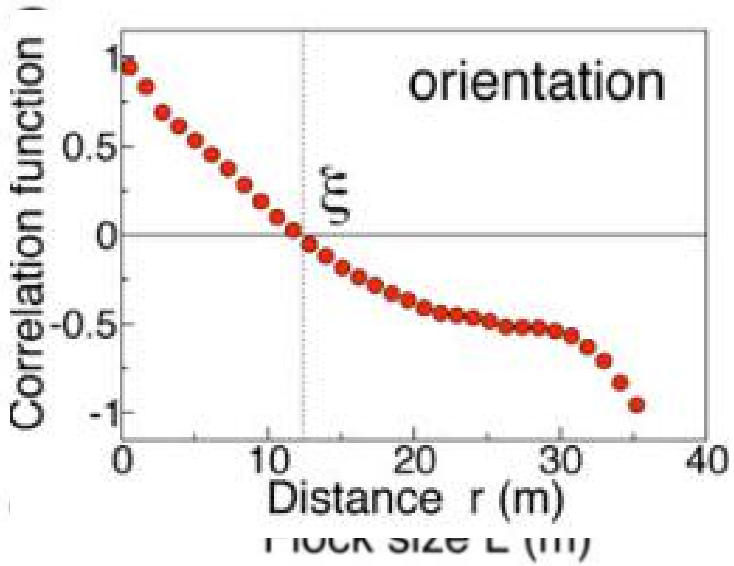
$$f(nx) \propto x^{\alpha}$$

Background: Long range correlation





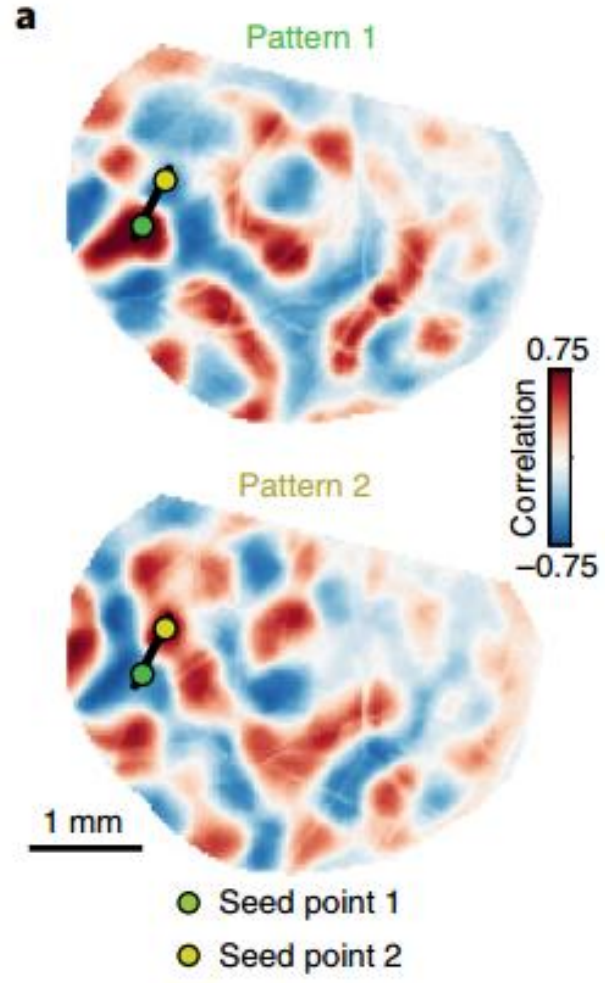
Chen & Gong, Nature Communications, 2019



Carvagna et al including Parisis, PNAS, 2010

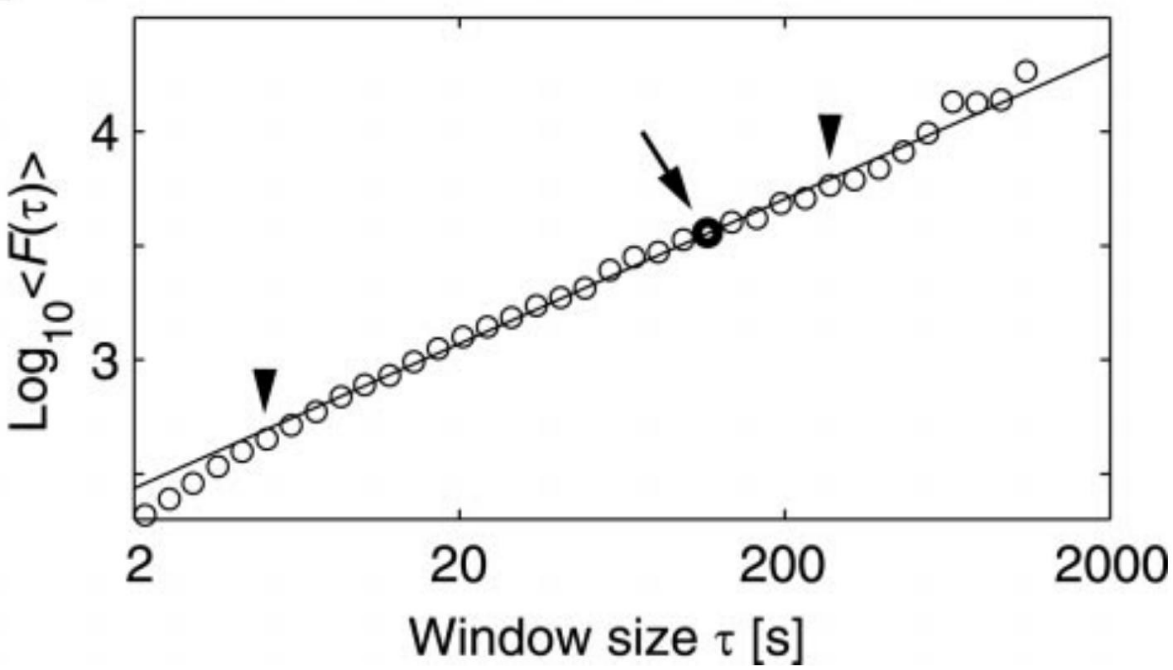
Background: Long range correlation

Calcium imaging of mouse V1
Long-range spatial correlation



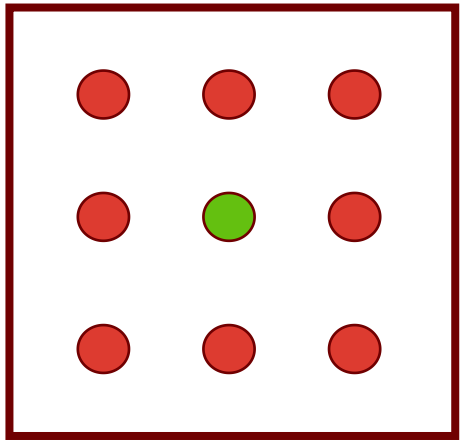
Smith et al., Nat. Neurosci., 2019

MEG of human brain
DFA
Long-range temporal correlation



Linkenkaer-Hansen et al, Journal of Neuroscience, 2001

Callback: Dilemma of existing methods

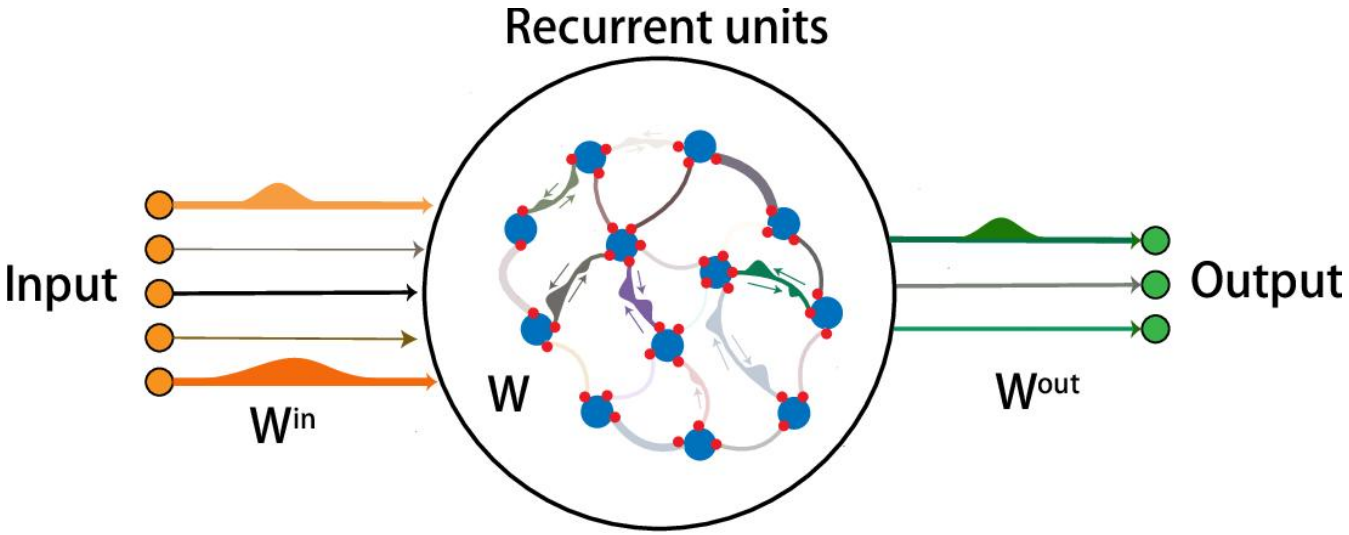


Backpropagation through time

The Dilemma of Existing Methods:

- BPTT: Precise gradients but requires unrolling the computational graph; memory scales linearly with sequence length (O(T)), prone to vanishing/exploding gradients.
- Online Learning: Computationally efficient (O(1) memory) but performing far worse than BPTT.
- Local Learning: Computationally efficient (good for chips) but performing far worse than BPTT.

Mathematical deduction by a vanilla RNN



$$\mathbf{x}_t = \mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{W}_{xh}\mathbf{u}_t + \mathbf{b}_h,$$

$$\mathbf{h}_t = \phi(\mathbf{x}_t),$$

$$\mathbf{o}_t = \mathbf{W}_{hy}\mathbf{h}_t + \mathbf{b}_y,$$

$$\mathbf{p}_t = \mathrm{softmax}(\mathbf{o}_t),$$

$$\mathcal{L}(\theta) = \sum_{t=1}^{T} \mathcal{L}_t = \sum_{t=1}^{T} \bar{w}_t \, \ell(\mathbf{o}_t, \mathbf{y}_t^{\star}),$$

Mathematical deduction by a vanilla RNN

北京大学 PEKING UNIVERSITY

Quantify criticality by Laypunov exponent

$$\mathbf{J}_t \triangleq \frac{\partial \mathbf{h}_t}{\partial \mathbf{h}_{t-1}} = \mathrm{diag}\left(\frac{\partial \mathbf{h}_t}{\partial \mathbf{x}_t}\right)\mathbf{W}_{hh}. \qquad \lambda_{\max} \approx \frac{1}{T}\log \sigma_{\max}\left(\prod_{t=1}^{T}\mathbf{J}_t\right)$$

Initialization with g to control Chaos

$$\mathbf{W}_{hh}(g) = g\,\widetilde{\mathbf{W}}_{hh},$$
$$\widetilde{\mathbf{W}}_{hh,ij} \sim \mathcal{N}\left(0, \frac{1}{H}\right) \text{ (i.i.d.).}$$

Two Core Hypotheses: Under the critical state (edge of chaos), neural signals exhibit two key properties

Assumption 1: long-range temporal correlations: On short time scales in a stable near-critical regime, the BPTT pre-activation gradient $\delta_t^{(i)}$ continues approximately by a per-unit scalar **(breaking backpropagation through time->online)**

$$\delta_{t+1}^{(i)} \approx \alpha_i \delta_t^{(i)}$$

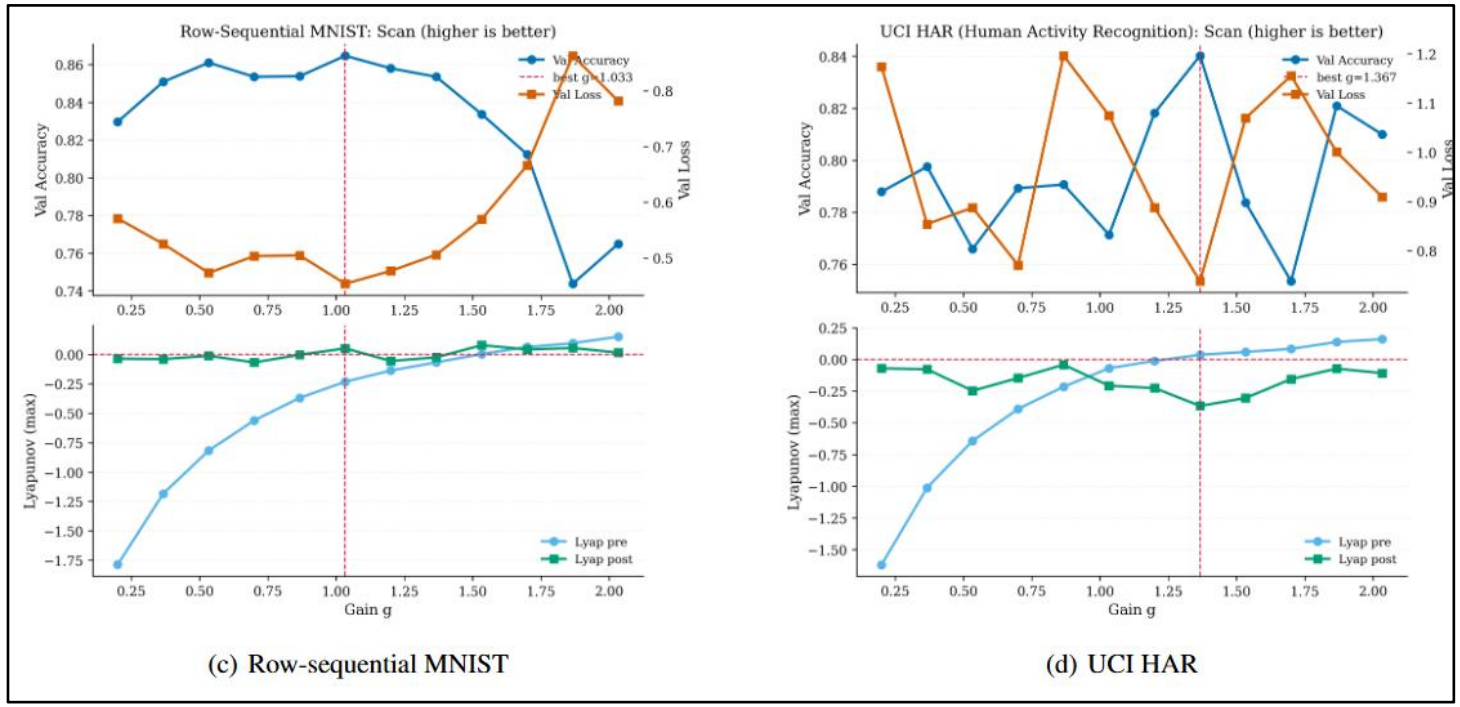
Assumption 2: long-range spatial correlations: In a stable near-critical regime, the within-step teaching signal is assumed to be spatially coherent. Specifically, for a fixed unit i, the error signal of any unit j is approximately proportional to that of unit i **(breaking backpropagation through space->local)** :

$$\delta_t^{(j)} \approx \beta_{ji} \delta_t^{(i)}$$

$\delta_t, \delta_t^{(i)}$	$\delta_t \equiv \partial \mathcal{L}_{\text{total}}/\partial \mathbf{x}_t$: The true BPTT gradient

Hypothesis Validation

Laypunov exponents demonstrate near edge-of-chao criticality



(c) Row-sequential MNIST (d) UCI HAR

Long-range spatiotemporal correlation is validated

Quantity	Low-rank	IID
Temporal AR(1)+bias one-step R^2	0.9944 ± 0.0094	0.8089 ± 0.0334
Teaching-signal EVR_1	0.9999 ± 0.0001	0.9836 ± 0.0112
Teaching-signal effective rank	1.0001 ± 0.0002	1.0339 ± 0.0237

Temporal AR(1)+bias one-step R^2 on $\boldsymbol{\delta}_t^{\text{true}}$

Fit $\delta_{t'}^{(i)} \approx \alpha_i(t)\delta_{t'-1}^{(i)} + \beta_i(t)$ on short windows of length W, then report the per-unit one-step R^2 for predicting $\delta_t^{(i)}$ from $\delta_{t-1}^{(i)}$ using the fitted $(\alpha_i(t), \beta_i(t))$ (averaged over units). Higher indicates stronger short-window temporal persistence (H1).

Mathematical Derivation: From BPTT to COLA

ϕ, ϕ'	Activation function and its first derivative.
$\mathbf{W}_{xh}, \mathbf{W}_{hh}, \mathbf{W}_{hy}$	Input-to-hidden, recurrent, and hidden-to-output weight matrices.

1) The original BPTT form

$$\boldsymbol{\delta}_t = \frac{\partial \mathcal{L}_t}{\partial \mathbf{x}_t} + \mathbf{U}_t \mathbf{W}_{hh}^\top \boldsymbol{\delta}_{t+1}. \qquad \mathbf{U}_t \equiv \mathrm{diag}\big(\phi'(x_t^{(1)}), \ldots, \phi'(x_t^{(H)})\big)$$

2) By using assumption 1: $\delta_{t+1}^{(i)} \approx \alpha_i \delta_t^{(i)}$, we get

$$\boldsymbol{\delta}_{t+1} = \mathbf{D}_\alpha \boldsymbol{\delta}_t \qquad \mathbf{D}_\alpha \equiv \mathrm{diag}(\alpha_1, \ldots, \alpha_H)$$

3) Combine 1) and 2)

$$\boldsymbol{\delta}_t \approx \frac{\partial \mathcal{L}_t}{\partial \mathbf{x}_t} + \mathbf{A}_t \boldsymbol{\delta}_t, \qquad \mathbf{A}_t \equiv \mathbf{U}_t \mathbf{W}_{hh}^\top \mathbf{D}_\alpha.$$

No propagation through time now

4) For a fixed unit i, the i-th component reads

$$\delta_t^{(i)} \approx \frac{\partial \mathcal{L}_t}{\partial x_t^{(i)}} + \frac{\partial h_t^{(i)}}{\partial x_t^{(i)}} \sum_{j=1}^{H} W_{hh}^{ji} \, \alpha_j \, \delta_t^{(j)}.$$

Mathematical Derivation: From BPTT to COLA

4) $$\delta_t^{(i)} \approx \frac{\partial \mathcal{L}_t}{\partial x_t^{(i)}} + \frac{\partial h_t^{(i)}}{\partial x_t^{(i)}} \sum_{j=1}^{H} W_{hh}^{ji}\, \alpha_j\, \delta_t^{(j)}.$$

5) By using assumption 2: $\delta_t^{(j)} \approx \beta_{ji}\delta_t^{(i)}$, we get

$$\sum_{j=1}^{H} W_{hh}^{ji}\, \alpha_j\, \delta_t^{(j)} \approx \sum_{j=1}^{H} W_{hh}^{ji}\, \alpha_j\, \left(\beta_{ji}(t)\, \delta_t^{(i)}\right) = \left(\sum_{j=1}^{H} W_{hh}^{ji}\, \alpha_j\, \beta_{ji}(t)\right) \delta_t^{(i)}.$$

No spatial propagation now

6) Define μ_i

$$\mu_t^{(i)} \approx \sum_{j=1}^{H} W_{hh}^{ji}\, \alpha_j\, \beta_{ji}(t)$$

7) From 4), 5) and 6)

$$\delta_t^{(i)} \approx \frac{\partial \mathcal{L}_t}{\partial x_t^{(i)}} + \frac{\partial h_t^{(i)}}{\partial x_t^{(i)}}\, \mu_t^{(i)}\, \delta_t^{(i)}$$

8) By transposing terms in 7)

$$\widetilde{\delta}_t^{(i)} \approx \frac{\partial \mathcal{L}_t / \partial x_t^{(i)}}{1 - \mu_i \frac{\partial h_t^{(i)}}{\partial x_t^{(i)}}}$$

Closed-form online estimation

Estimation of α_i

To estimate α_i, the teaching signal is treated as a local autoregressive (AR(1)) model, $\widetilde{\delta}_t^{(i)} \approx \alpha_i \widetilde{\delta}_{t-1}^{(i)}$, and use the least-squares solution

$$\alpha_i^\star = \frac{\mathbb{E}[\widetilde{\delta}_t^{(i)} \widetilde{\delta}_{t-1}^{(i)}]}{\mathbb{E}[(\widetilde{\delta}_{t-1}^{(i)})^2]},$$

Estimation of μ_i

$$\mu_i^{(t)} = \frac{\alpha_i \frac{\partial \mathcal{L}_{t-1}}{\partial x_{t-1}^{(i)}} - \frac{\partial \mathcal{L}_t}{\partial x_t^{(i)}}}{\alpha_i \phi'(x_t^{(i)}) \frac{\partial \mathcal{L}_{t-1}}{\partial x_{t-1}^{(i)}} - \phi'(x_{t-1}^{(i)}) \frac{\partial \mathcal{L}_t}{\partial x_t^{(i)}}}$$

How to update

$$\Delta W_{hh}^{ij} = -\eta \widetilde{\delta}_t^{(i)} h_{t-1}^{(j)},$$
$$\Delta W_{xh}^{ij} = -\eta \widetilde{\delta}_t^{(i)} u_t^{(j)},$$
$$\Delta b_h^{(i)} = -\eta \widetilde{\delta}_t^{(i)}.$$

$$\Delta W_{hy}^{ij} = -\eta \left(\frac{\partial \mathcal{L}_t}{\partial o_t^{(i)}} \right) h_t^{(j)},$$
$$\Delta b_y^{(i)} = -\eta \frac{\partial \mathcal{L}_t}{\partial o_t^{(i)}}.$$

- Gradient estimation (error signal) multiplying presynaptic neural activity
- Similar to BTSP, interestingly

Performance

北京大学 PEKING UNIVERSITY

Comparable to BPTT and its variants on both simple and hard tasks

Table 2. RNN benchmarks (mean ± std over seeds 42/43/44). Best among the methods shown is bold.

Task	BPTT	TBPTT-1	TBPTT-10	e-prop	FPTT	UORO	SnAp-1	**COLA**
Adding	0.1426	0.0863	0.0991	0.0813	0.1356	0.1444	0.0860	**0.0059**
MSE↓	± 0.0326	± 0.0031	± 0.0013	± 0.0031	± 0.0336	± 0.0150	± 0.0034	± **0.0026**
Lorenz	0.6841	0.2019	0.6749	0.3602	0.8038	0.1395	0.2025	**0.0058**
MSE↓	± 0.0873	± 0.0845	± 0.0163	± 0.0201	± 0.0250	± 0.1072	± 0.0676	± **0.0003**
PTB-Char	2.8695	1.8316	2.2272	2.0852	2.3184	2.2279	**1.8279**	1.8352
Loss↓	± 0.0203	± 0.0040	± 0.0029	± 0.0013	± 0.0057	± 0.0063	± **0.0041**	± 0.0056
WikiText-2	3.3620	2.2833	2.8780	2.5851	3.0499	2.6225	**2.2793**	2.4174
Loss↓	± 0.0165	± 0.0058	± 0.0201	± 0.0073	± 0.0307	± 0.0653	± **0.0054**	± 0.0035
Row-MNIST	0.9713	0.9343	**0.9734**	0.9565	0.9481	0.3060	0.9403	0.9544
Acc↑	± 0.0007	± 0.0011	± **0.0020**	± 0.0005	± 0.0037	± 0.0193	± 0.0047	± 0.0036
Row-CIFAR10	0.4545	0.3894	**0.4759**	0.3836	0.4021	0.2185	0.3825	0.4418
Acc↑	± 0.0108	± 0.0162	± **0.0130**	± 0.0026	± 0.0064	± 0.0057	± 0.0070	± 0.0083
UCI HAR	**0.8877**	0.8416	0.8799	0.8314	0.8367	0.5594	0.8420	0.8690
Acc↑	± **0.0454**	± 0.0291	± 0.0390	± 0.0446	± 0.0171	± 0.0359	± 0.0134	± 0.0257

Performance

北京大学 PEKING UNIVERSITY

COLA works on multiple types of neural network models

Table 9. ConvRNN refinement benchmarks (Acc↑). Best is bold; second best is underlined.

Task	BPTT	TBPTT-1	TBPTT-10	e-prop	FPTT	COLA
Fashion-MNIST	**0.8726**	0.8281	<u>0.8619</u>	0.8398	0.7815	0.8367
Permuted MNIST	**0.9418**	0.9071	<u>0.9317</u>	0.9109	0.8734	0.9188
DVS-Gesture	**0.6667**	0.6136	<u>0.6439</u>	0.6098	0.5492	0.6174
DVS-CIFAR10	**0.4820**	0.4480	<u>0.4650</u>	0.4290	0.3960	0.4440

Table 3. Spiking recurrent benchmarks on directly comparable cell-task pairings. Best within each row is bold.

Dataset	Cell	BPTT	pp-prop	COLA
N-MNIST	RLIF	98.33% ± 0.04%	98.25% ± 0.03%	**98.54% ± 0.03%**
N-MNIST	RadLIF	98.29% ± 0.02%	**98.40% ± 0.03%**	93.56% ± 0.12%
SHD	RLIF	94.01% ± 0.12%	93.93% ± 0.28%	**94.07% ± 0.45%**
SHD	RadLIF	94.72% ± 1.06%	95.33% ± 0.11%	**95.37% ± 0.13%**
DVS-Gesture	EGRU	**97.45% ± 0.27%**	97.29% ± 0.16%	97.19% ± 0.32%

Performance

北京大学 PEKING UNIVERSITY

COLA beats BPTT when the sequence is long

Table 19. Long-horizon gesture stress test (accuracy).

	ASRNN		LTC-SNN	
Seq. length	BPTT	COLA	BPTT	COLA
20	79.5%	81.6%	89.9%	86.1%
100	77.8%	86.8%	44.1%	86.8%
500	70.8%	85.4%	Diverged	85.8%
1000	66.3%	86.5%	Diverged	86.5%

Cost

北京大学 PEKING UNIVERSITY

COLA is light weight

Table 1. Asymptotic costs for a vanilla RNN. Here, T denotes the sequence length, H the number of hidden units, D the input dimension, and K the output dimension. $N_\theta = \Theta\left(H^2 + HD + KH\right)$ represents the total number of trainable parameters. "Activation memory" refers to the storage required for hidden states that scales with the sequence length (excluding parameters).

Method	Time (per sequence)	Activation memory	Extra state (excluding parameters)
Full BPTT (Werbos, 2002)	$\mathcal{O}(TN_\theta)$	$\mathcal{O}(TH)$	none
TBPTT (truncation τ)	$\mathcal{O}(TN_\theta)$	$\mathcal{O}(\tau H)$	none
RTRL (Williams & Zipser, 1989)	$\mathcal{O}(TH^4)$	$\mathcal{O}(1)$	$\mathcal{O}(H^3)$
UORO (Tallec & Ollivier, 2017)	$\mathcal{O}(TN_\theta)$	$\mathcal{O}(1)$	rank-one factors $\mathcal{O}(N_\theta + H)$
SnAp-1 (Menick et al., 2021)	$\mathcal{O}(TN_\theta)$	$\mathcal{O}(1)$	one-step sparse influences $\mathcal{O}(N_\theta + H)$
e-prop (Bellec et al., 2020)	$\mathcal{O}(TN_\theta)$	$\mathcal{O}(1)$	eligibility traces $\mathcal{O}(N_\theta)$
FPTT (Kag & Saligrama, 2021)	$\mathcal{O}(TN_\theta)$	$\mathcal{O}(1)$	shadow/dual variables $\mathcal{O}(N_\theta)$
Ours (COLA)	$\mathcal{O}(TN_\theta)$	$\mathcal{O}(1)$	α, μ stats $\mathcal{O}(H)$

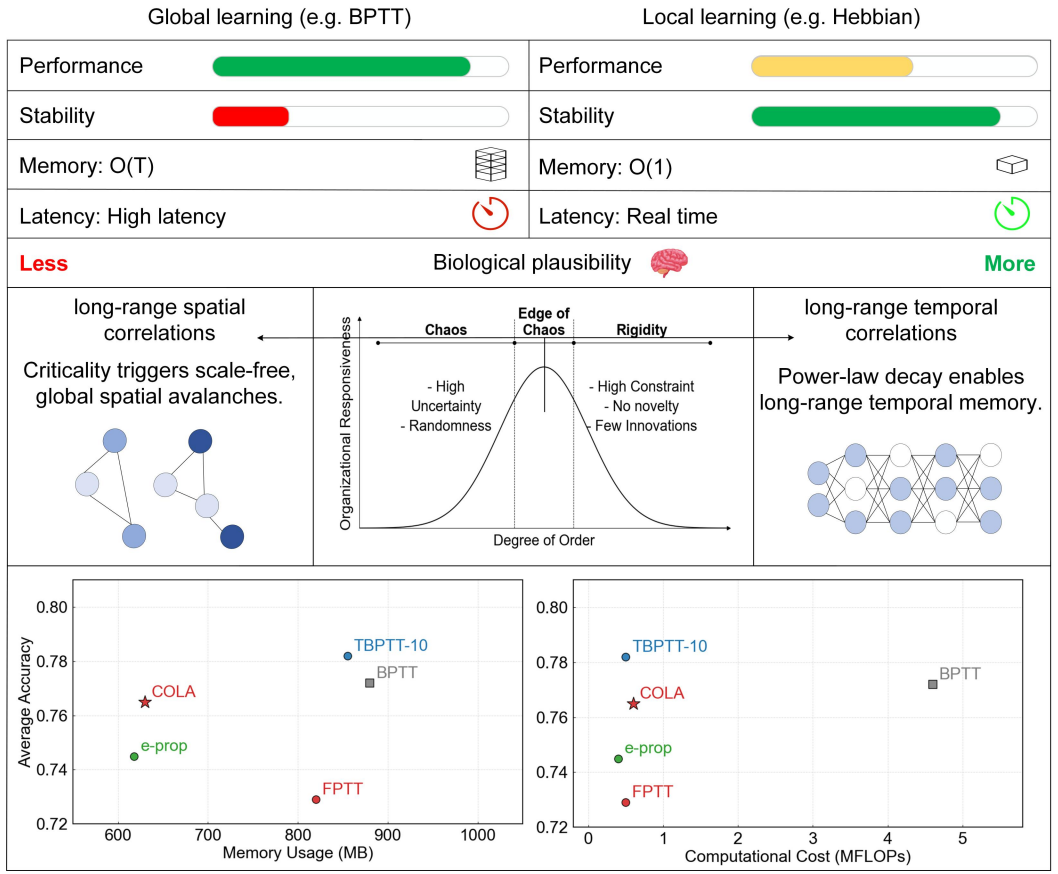
Take home messages

COLA (Criticality-driven Online Local Alignment): A rigorous online local learning rule.

Mechanism: Uses Dynamical Criticality as an enabler to transform BPTT's temporal recursion into a local approximation at the same timestep.

Key Advantages:

- Constant Memory: Requires only O(H) auxiliary states; activation memory is O(1).
- High Performance: Matches BPTT in benchmarks and shows superior stability in long-period tasks.
- Linear Time Complexity: O(TP) temporal overhead.



Thank you!