Scaling up Echo-State Networks with multiple light scattering

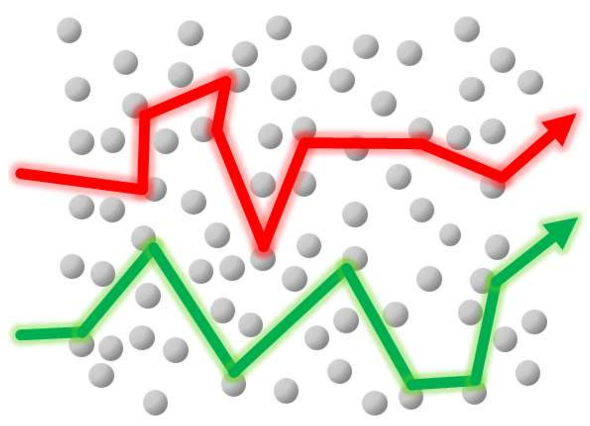
Jonathan Dong[1,2,4], Florent Krzakala[2], Gilles Wainrib[3], Sylvain Gigan[1]

[1] Laboratoire Kastler Brossel, CNRS UMR 8552 & Ecole Normale Supérieure, Paris 75005, France
[2] Laboratoire de Physique Statistique, CNRS UMR 8550 & Ecole Normale Supérieure, Paris 75005, France
[3] Département d'Informatique, Ecole Normale Supérieure, Paris 75005, France
[4] LightOn, 2 rue de la Bourse, Paris 75002, France

We present an optical device that performs random projections using the physical properties of multiple coherent scattering of light in random media. These efficient optical random projections are used to iterate an Echo-State Network, a Recurrent Neural Network with fixed internal weights. This new method is fast, power efficient and easily scalable to very large networks: we reach sizes that exceed the RAM memory limit.

Random projections using light scattering

- After propagation in a complex medium, light forms a random speckle pattern

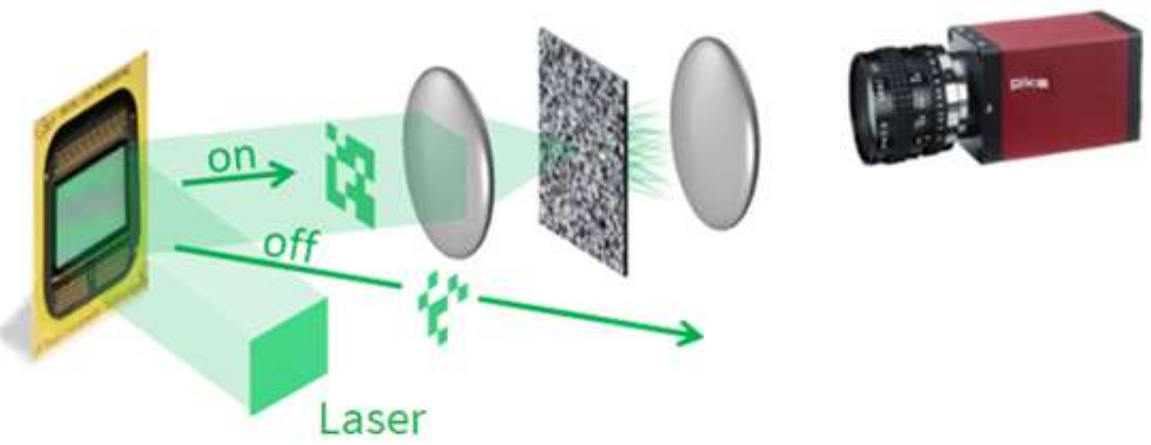


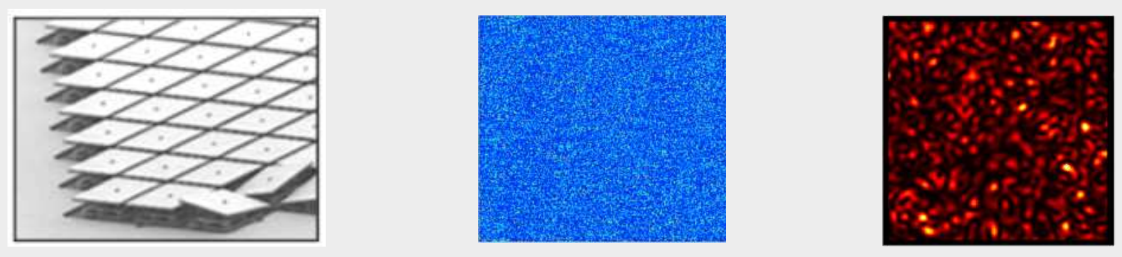
Light multiple scattering

Examples
White paint, milk, biological tissues

Applications
Microscopy, fiber communication, quantum information

- Before the complex medium, input light is modulated by a Digital Micromirror Device (DMD).





Binary input field — Transmission matrix H — Speckle pattern

- After the complex medium, the output speckle image is a random projection of the DMD image.

$$I = |HE_{in}|^2$$

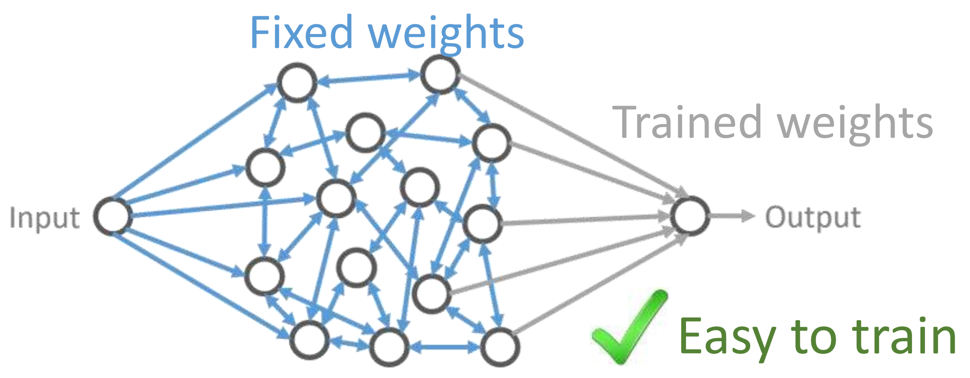
Transmission matrix
Fixed dense random matrix

Key features
- Input dimension: 10^5
- Output dimension: 10^5
- Speed: 300 Hz

Ref: Popoff et al, PRL 2010

Optical Reservoir Computing

- Echo-State Network: Recurrent Neural Network with fixed weights to bypass complicated training
- Reservoir Computing: Generalization to any dynamical system.



Ref: Jaeger, GMD Report 2001
Lukosevicius et al, Computer Science Review 2009

- The update equation involves a multiplication by a random weight matrix

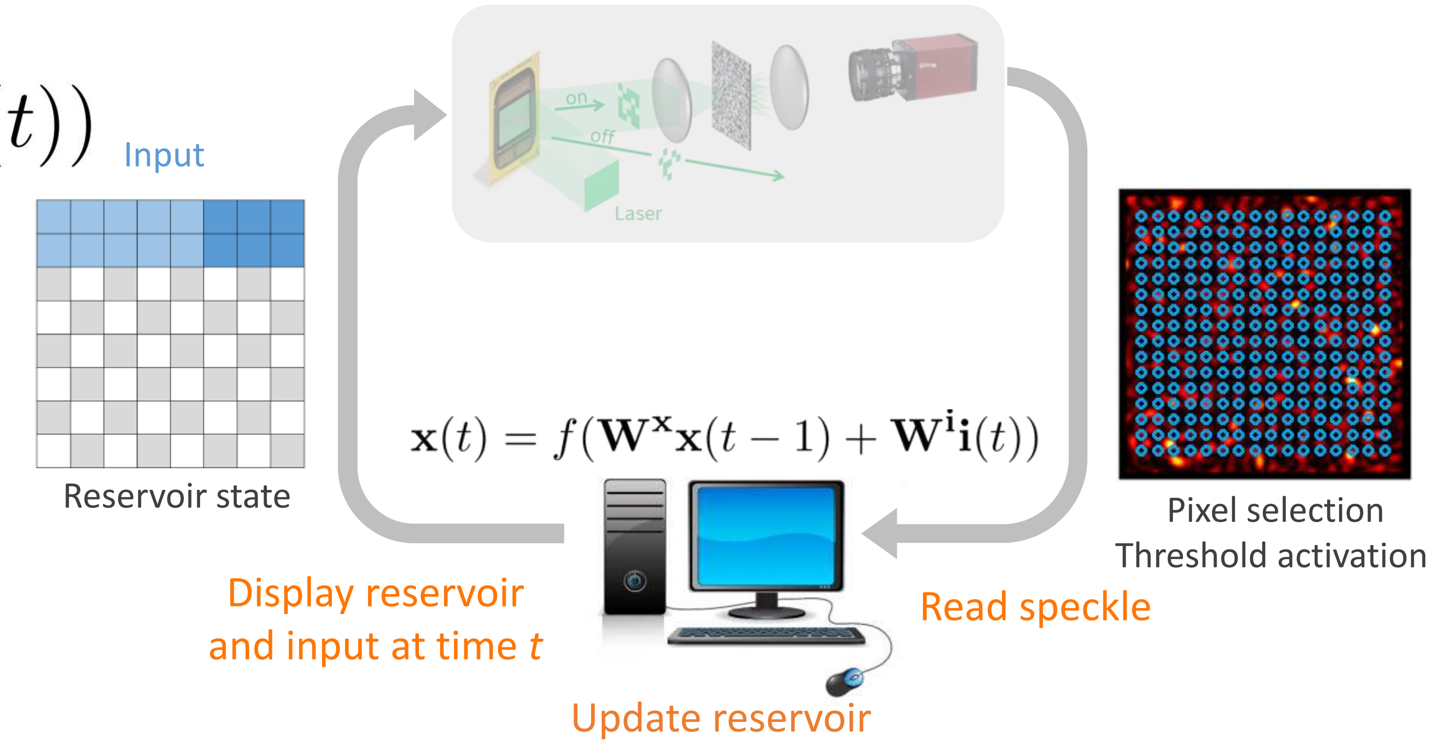
$$x(t+1) = f(Wx(t) + Vi(t))$$

Random matrices

Notations
- $x(t)$ network state at time t
- $i(t)$ input at time t
- f non-linear activation function

Optical Echo-State Network
- Binary neurons
- Dense connections
- Large dimension

- We use the optical implementation above to compute this update equation



Implementation details and results

- First lab implementation followed by a high-performance one by LightOn

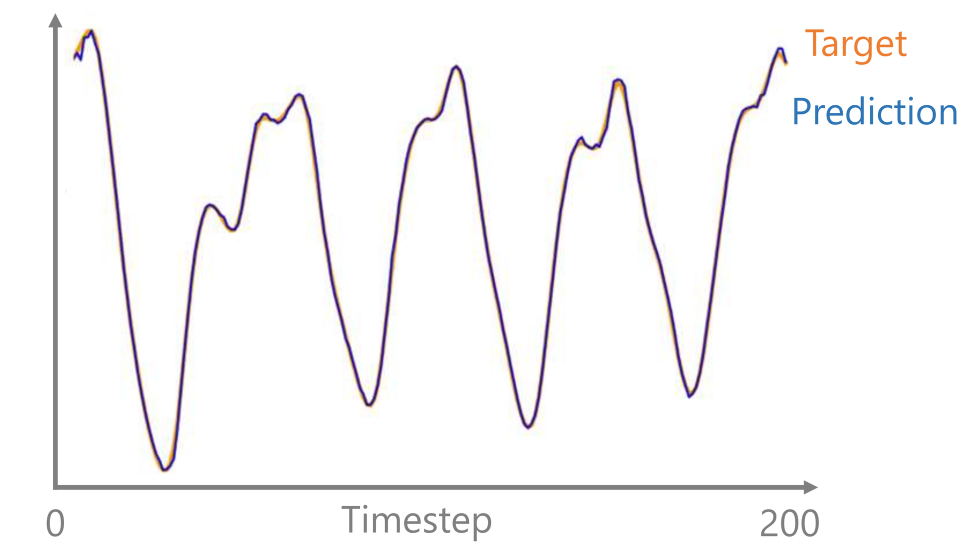
Lab implementation at LKB	Implementation at LightOn
10 Hz	300 Hz
20'000 neurons	100'000 neurons
September 2016	February 2018
arXiv v1	arXiv v2

Ref: Dong et al, IEEE SSP 2018

- Tested on non-linear chaotic time series prediction

Mackey-Glass time series prediction (with 100'000 neurons)

$$\frac{dx}{dt} = \beta \frac{x_\tau}{1 + {x_\tau}^n} - \gamma x, \quad \gamma, \beta, n > 0$$



- This optical implementation is fast and scalable to very large dimension

	CPU Intel Xeon E5-2690v3 on Microsoft Azure	Optical implementation
Complexity	$O(N^2)$	$O(1)$
Maximal size	RAM limit 50'000 neurons (56 GB)	Resolution limit 100'000 neurons (> 1M possible)
Time per 1000 iterations (for 50'000 neurons)	720 s	3.2 s

225 times faster
at size 50'000 (RAM limit)

We successfully trained a large-scale binary Optical Echo-State Networks on non-linear time series prediction. Based on multiple light scattering, fast optical random projections are very efficient to compute update iterations in Reservoir Computing. Future research directions include the study of the asymptotic behavior of Reservoir Computing, the impact of the binarization scheme and the large-scale linear regression.