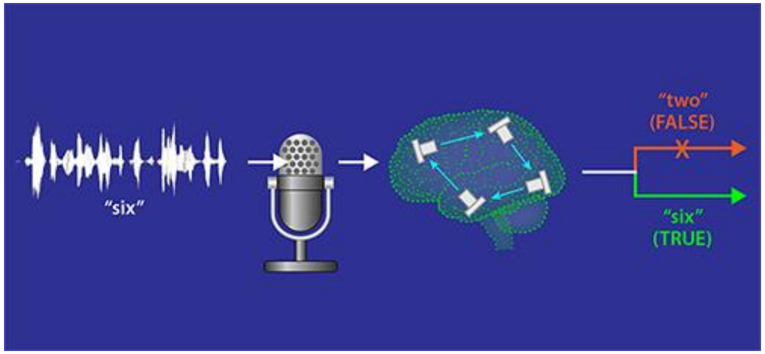
Photonic hardware implementation of time-multiplexed reservoir computing and extreme-learning machine

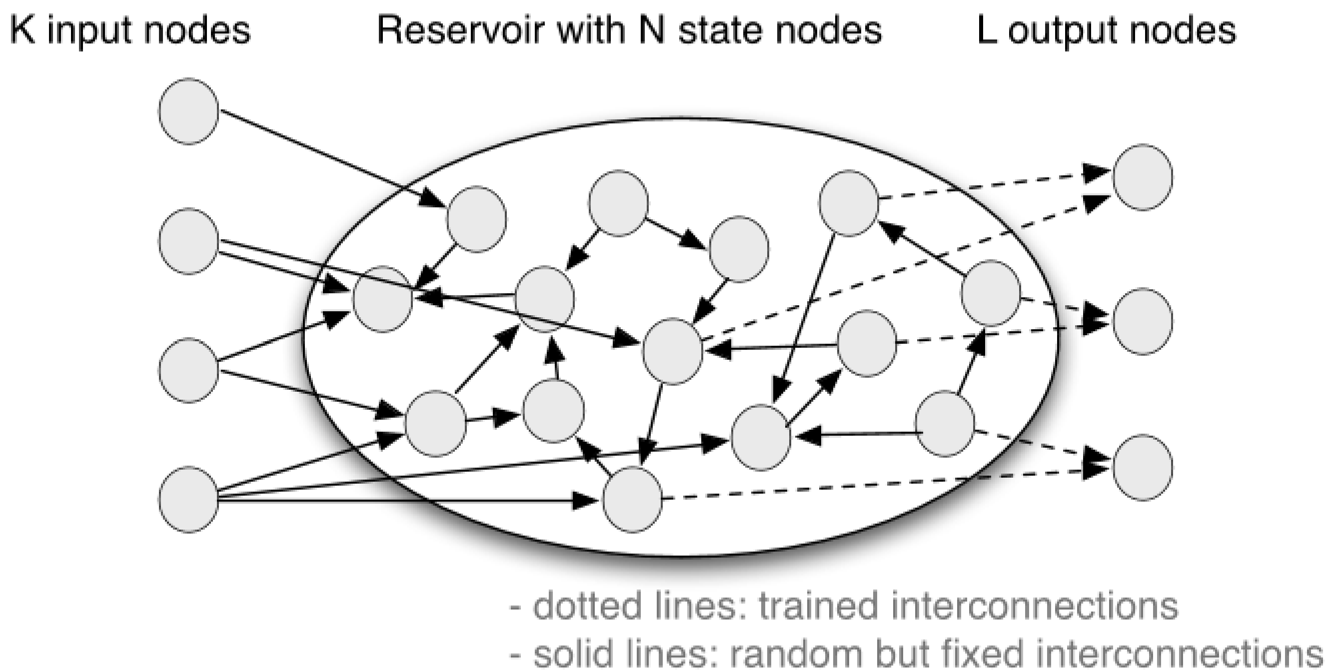
MIGUEL C. SORIANO



IFISC
AGENCIA ESTATAL DE INVESTIGACIÓN
UNIT OF EXCELLENCE MARÍA DE MAEZTU

ERC International Workshop
Photonic Reservoir Computing and Information Processing in Complex Networks

Reservoir Computing: combination of simplicity and power of concept, very easy training, inherent memory - ability to process sequential data streams



Untrained reservoirs implies suitability for **hardware implementations** -> Photonic RC

Exploit advantages of photonics hardware: speed, parallelism and **nonlinearities**

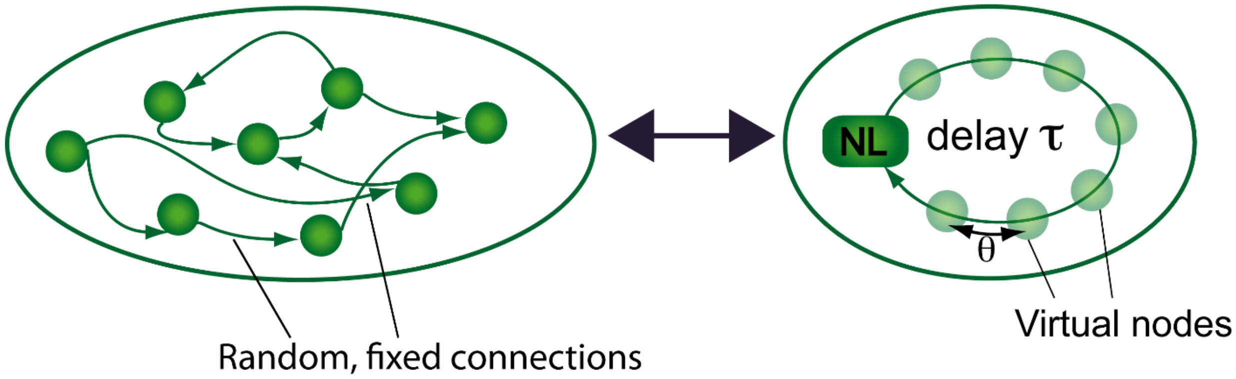
K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, and J. Van Campenhout, “Toward optical signal processing using Photonic Reservoir Computing”, Optics Express 16, 11182-11192 (2008).

Minimal approach with time-multiplexed RC

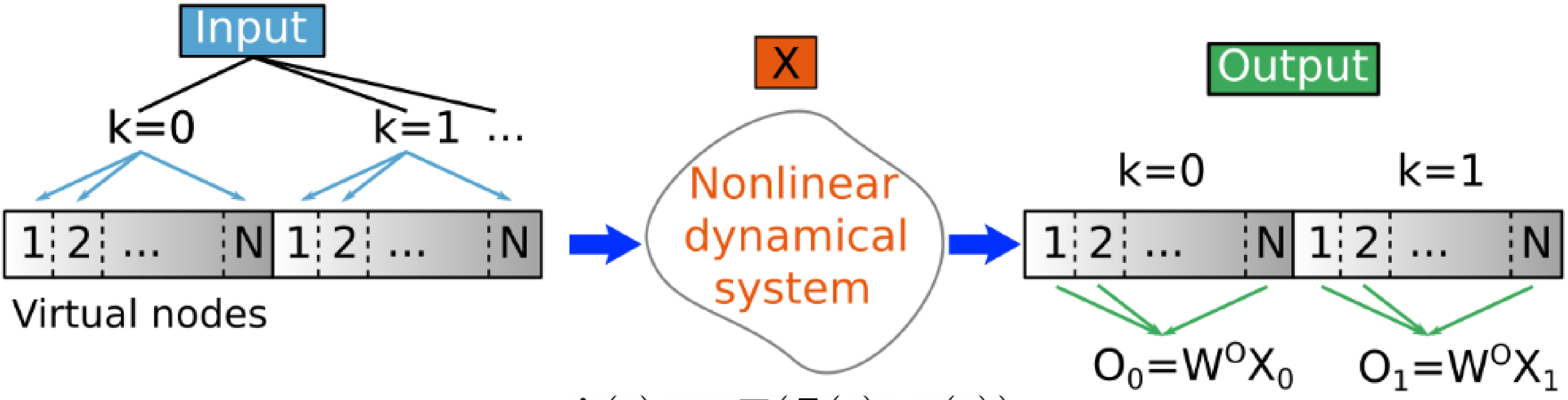
Minimal ingredients for RC

- Non-linear transformation
- High-dimensional mapping
- Fading memory

Traditional Reservoir | Time-multiplexed Reservoir



L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, “Information processing using a single dynamical node as complex system”, Nature Communications 2, 468 (2011).



$$\dot{\mathbf{x}}(t) = F(\mathbf{J}(t), \mathbf{x}(t))$$

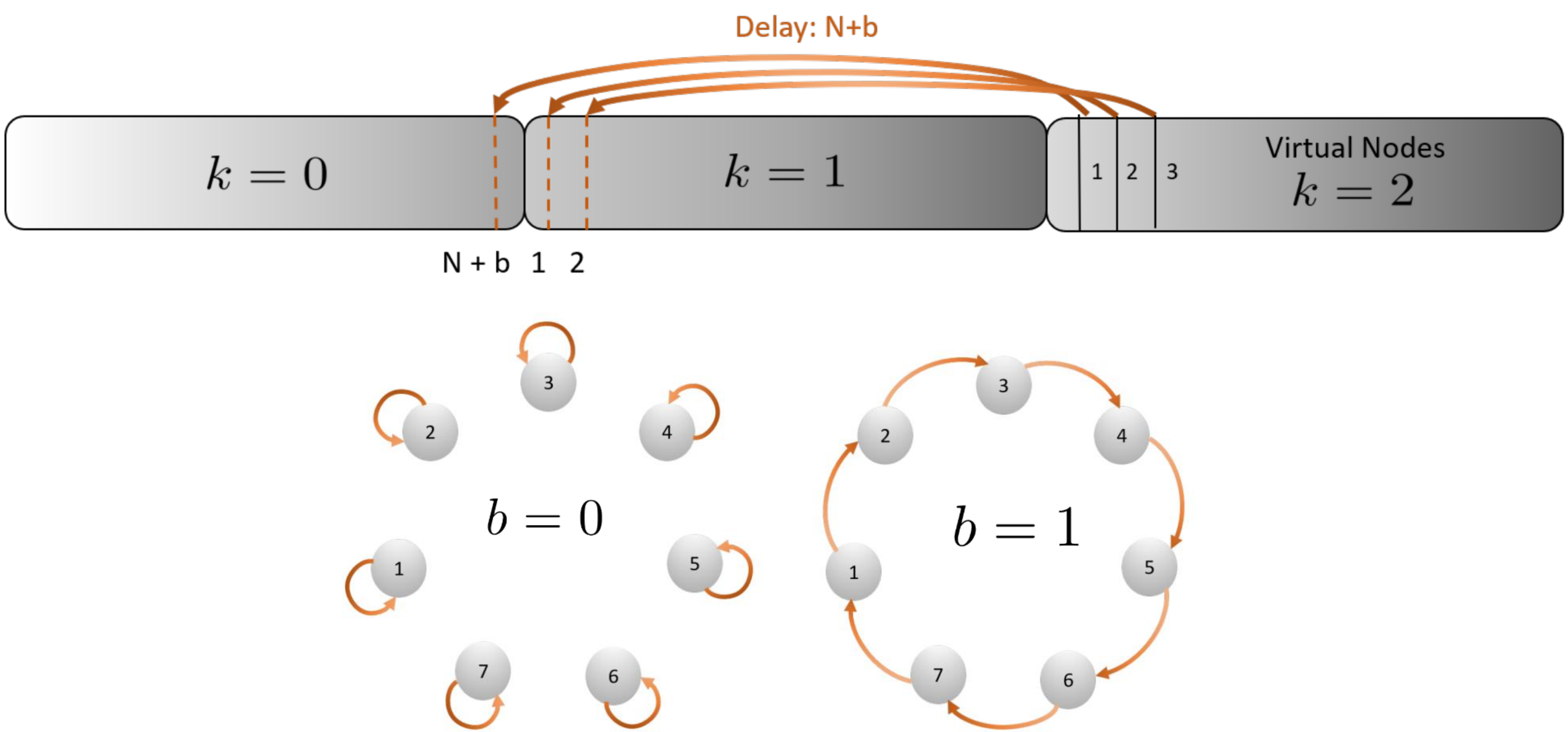
Time-multiplexed input with random scaling allocated to different temporal slots

Sequential input drives the nonlinear dynamical system

Output is formed by a weighted sum of the virtual node states with trained weights

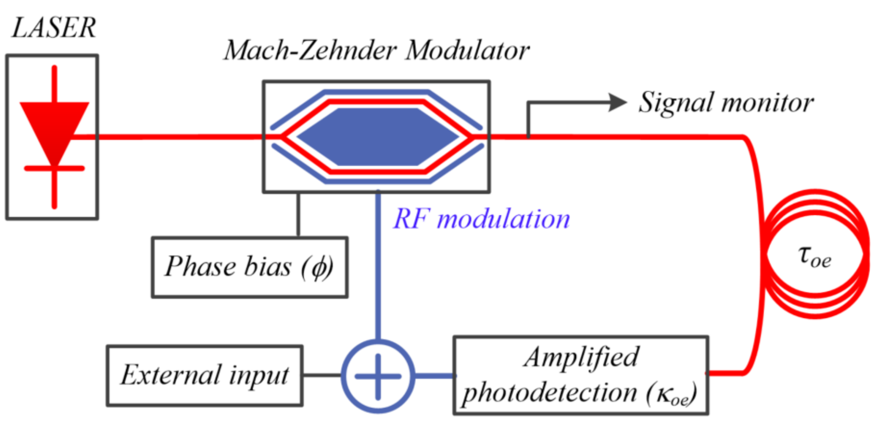
Time-multiplexed Input → Delay

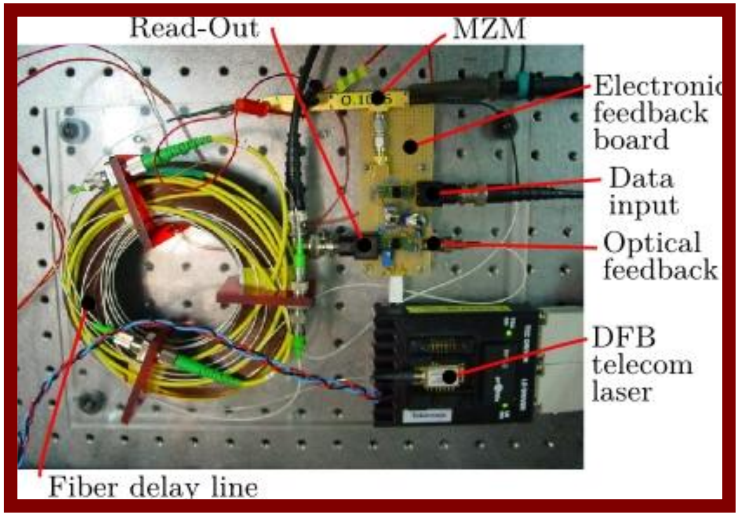
$$\mathbf{x}(t) = F\left(\gamma \mathbf{J}(t) + \beta \mathbf{x}(t - \tau) + \phi\right)$$

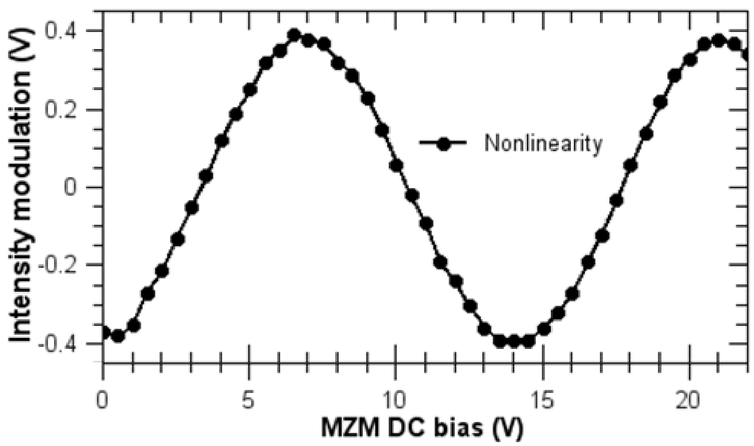


$F(x) = sin^2(x)$ ⟶ Experiments on opto-electronic RC implementation

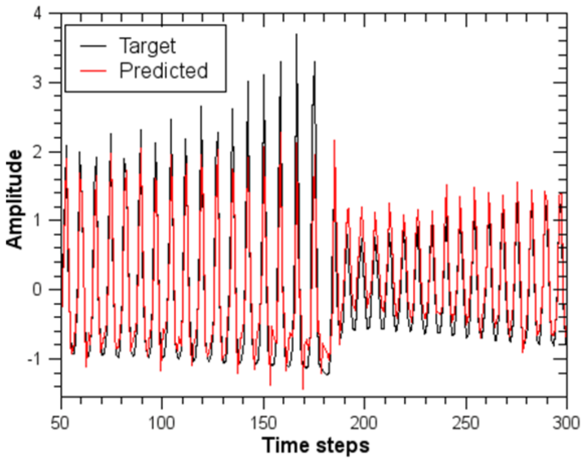
Opto-electronic Reservoir Computer







Chaotic time series prediction becomes possible:
- 1 hardware node
- 400 virtual nodes

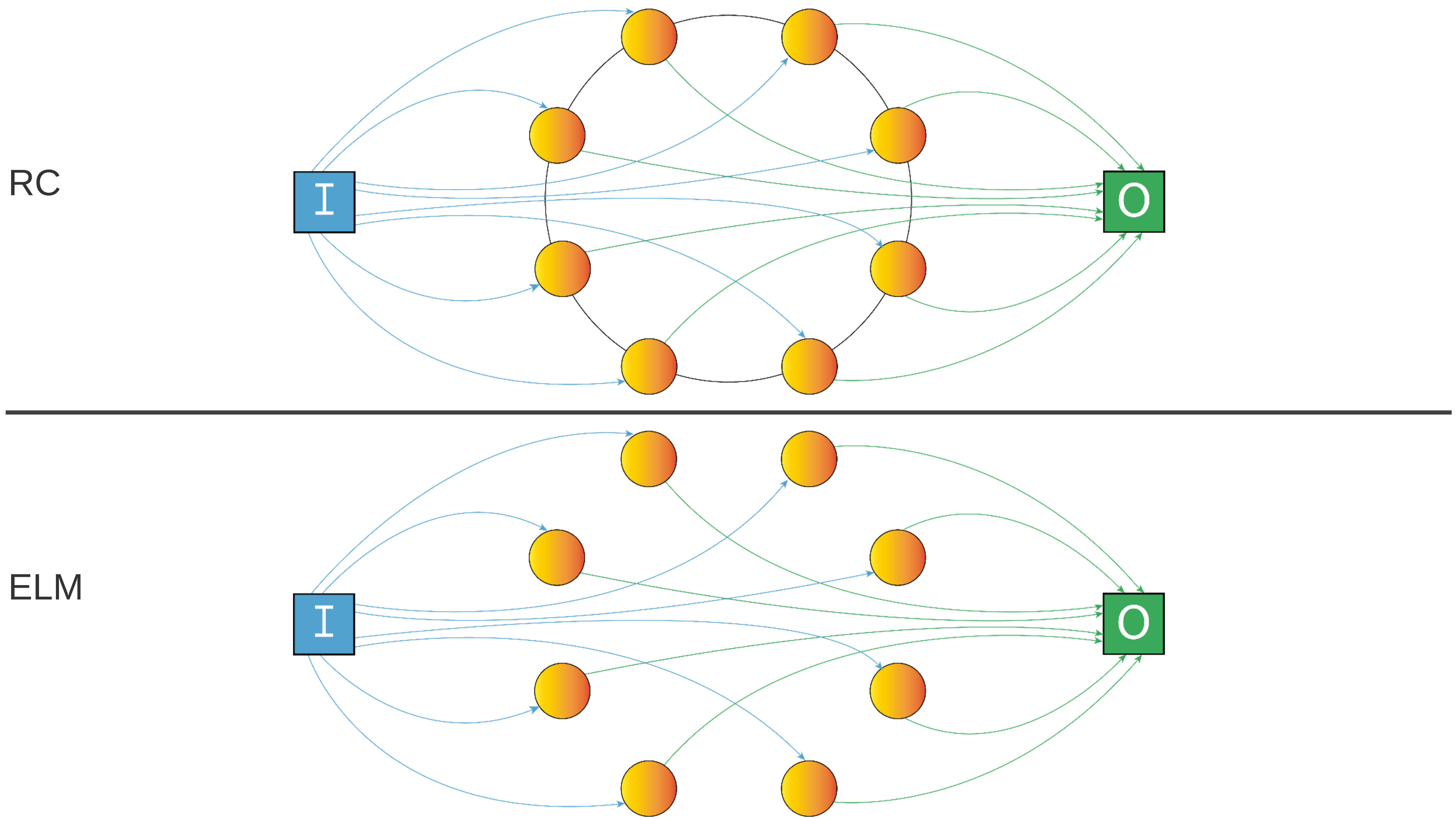


L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing, Optics Express 20, 3241-3249 (2012).

Outline

- Overview of Time-Multiplexed (Opto-electronic) Reservoir Computing
- Comparison of RC and Extreme Learning Machines (ELMs)
- Evaluation of experimental implementation
- Summary and conclusions

Comparison of RC and ELM schemes



G. B. Huang, Q. Y. Zhu, and C. K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70, 489-501 (2006).

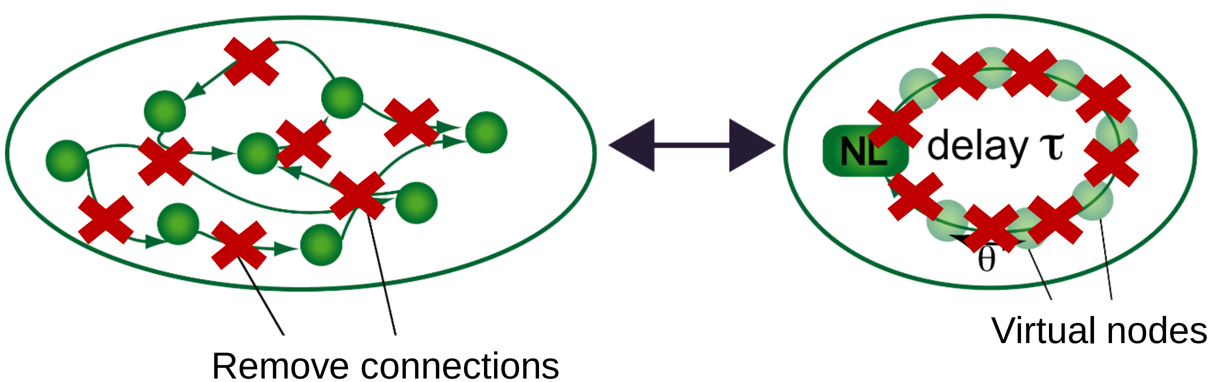
Y. H. Pao, G. H. Park, and D. J. Sobajic. Learning and generalization characteristics of the random vector functional-link net, Neurocomputing 6, 163-180 (1994).

Minimal approach with time-multiplexed ELM

Minimal ingredients for ELM

- Non-linear transformation
- High-dimensional mapping
- ~~Fading memory~~

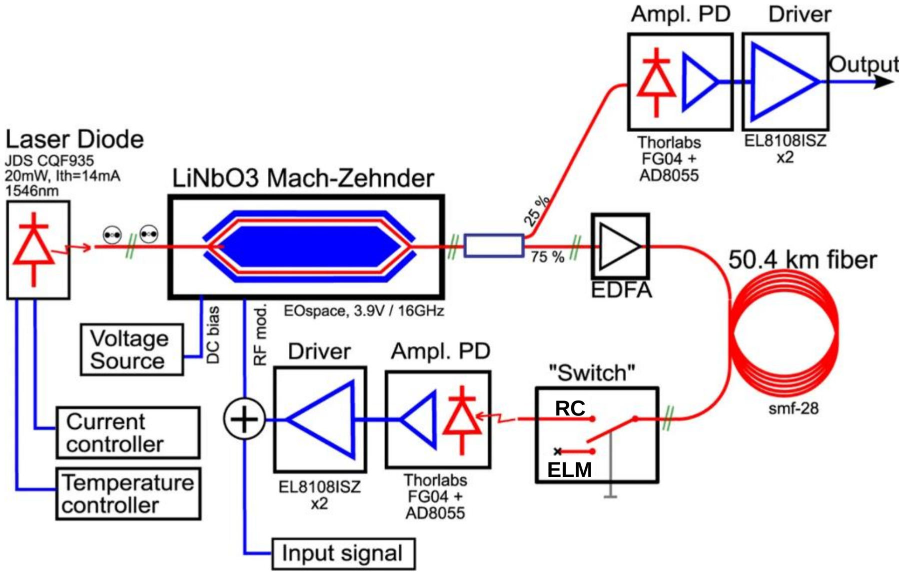
Traditional ELM | Time-multiplexed ELM

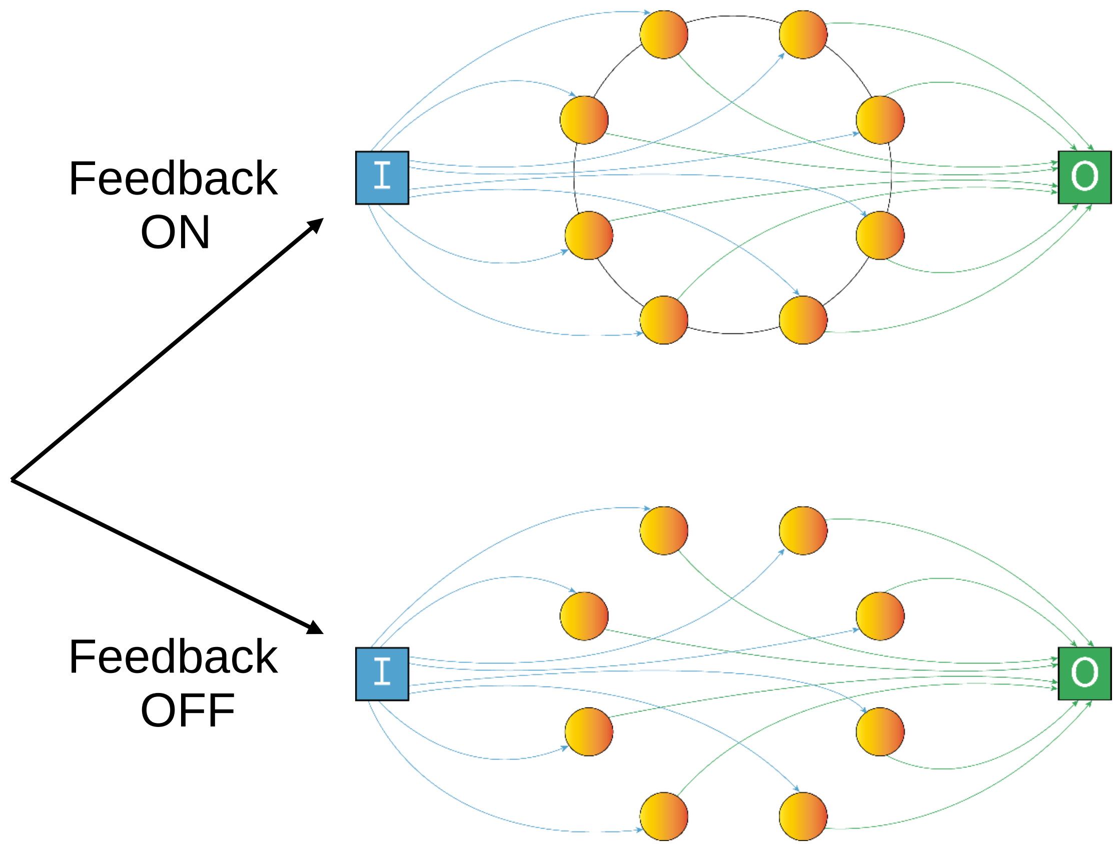


S. Ortín, M. C. Soriano, L. Pesquera, D. Brunner, D. San-Martín, I. Fischer, C. R. Mirasso, and J. M. Gutiérrez, “A unified framework for reservoir computing and extreme learning machines based on a single time-delayed neuron”, Scientific Reports 5, 14945 (2015).

Dual implementation with time-multiplexing

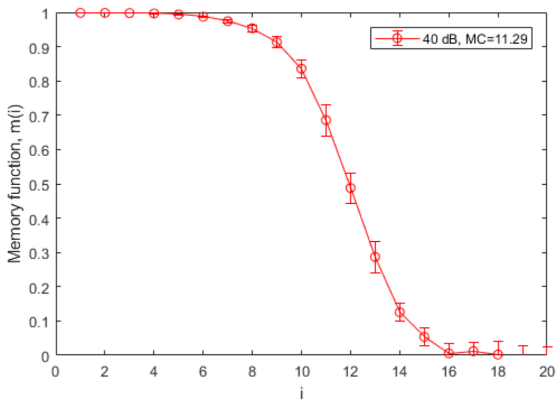
Opto-electronic RC / ELM





S. Ortín, M. C. Soriano, L. Pesquera, D. Brunner, D. San-Martín, I. Fischer, C. R. Mirasso, and J. M. Gutiérrez, “A unified framework for reservoir computing and extreme learning machines based on a single time-delayed neuron”, Scientific Reports 5, 14945 (2015).

Memory Capacity



Visualize fading memory

$$y(n) = u(n-i)$$

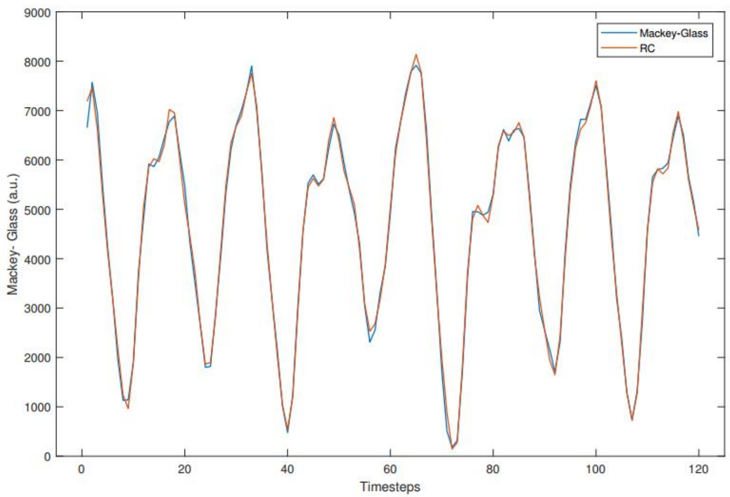
u(n) random time series

$$m(i) = corr[o_i(n), y_i(n)]$$

$$MC = \sum_i^{\infty} m(i)$$

MC quantifies the capability to recall previous inputs from the current state of the reservoir

Chaotic Time series prediction



One-step ahead prediction

$$y(n) = u(n+1)$$

MG chaotic time-series (MG17)

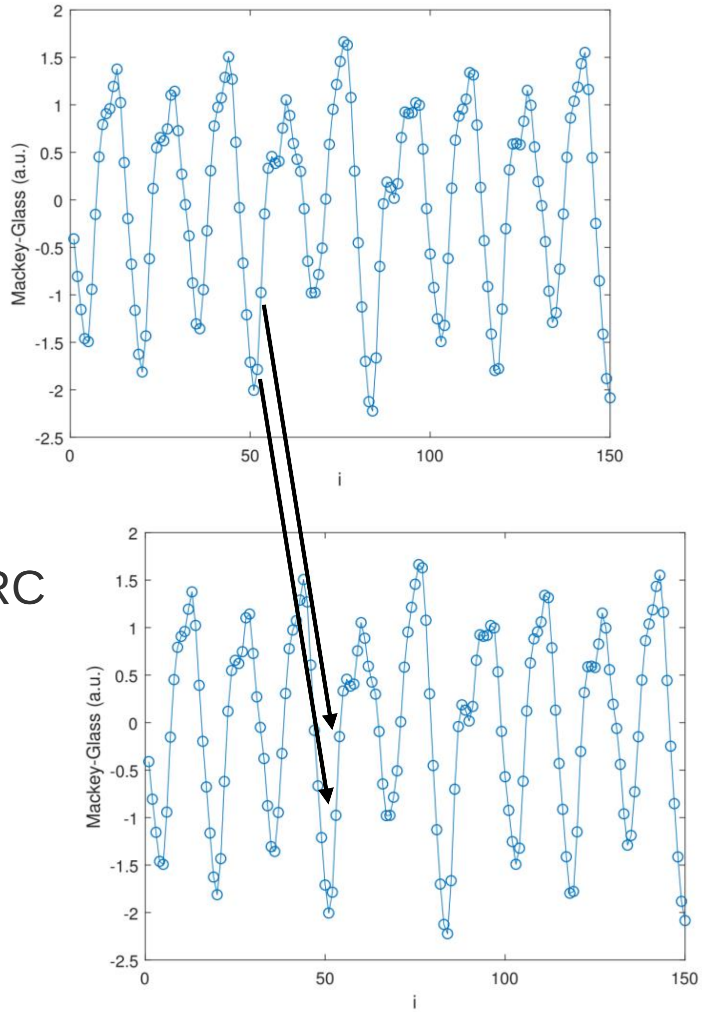
$$\dot{z}(t) = \frac{az(t-\tau)}{1+z^{10}(t-\tau)} - bz(t)$$

$$u(n) = z(nT)$$

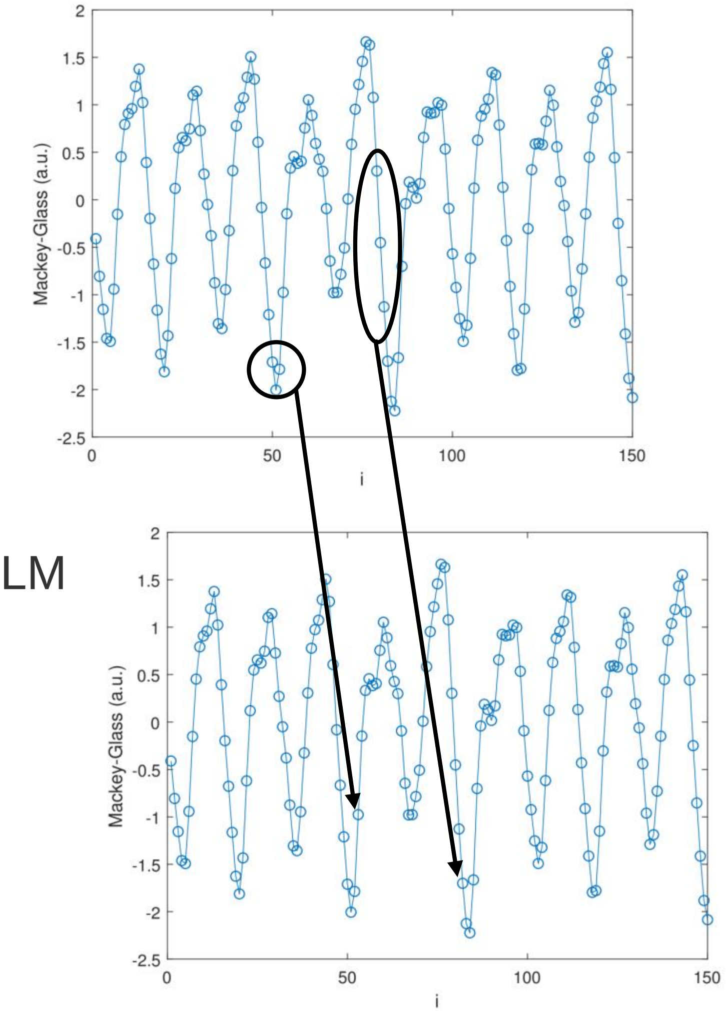
This task requires memory of previous inputs and the capability to approximate nonlinear functions

Example: Chaotic time series, one step-ahead prediction

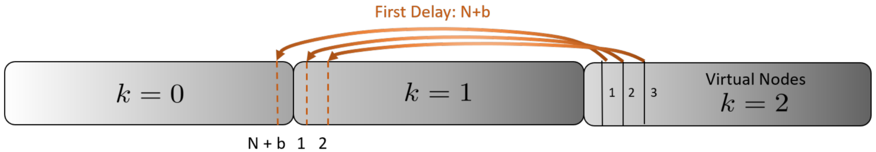
RC

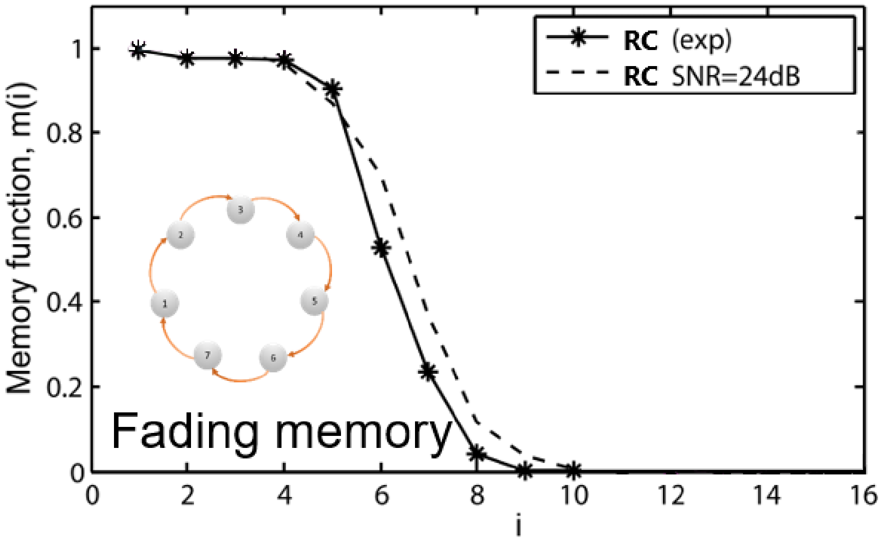


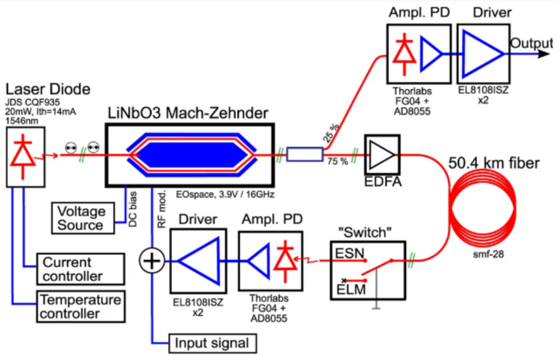
ELM



- Sequential input feeding is required in RC (not in ELM)
- External memory is needed in ELM (not in RC)

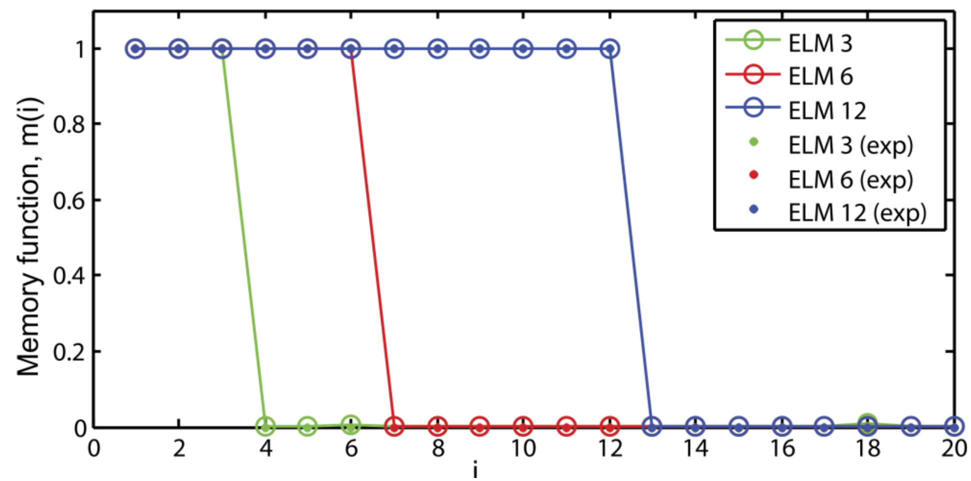






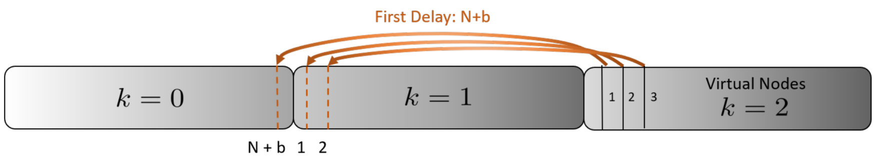
Recurrent connectivity creates intrinsic memory in RC

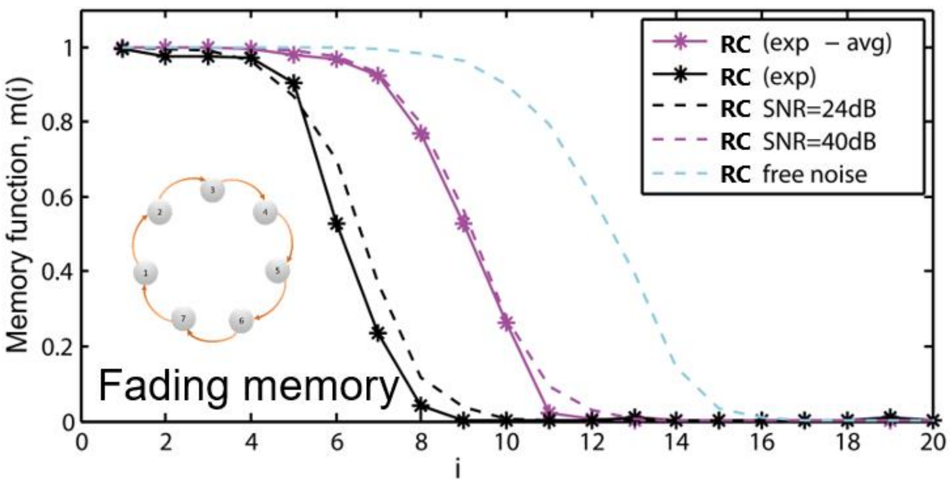
Memory in ELM needs to be explicitly given



* Numerical simulations consider system noise and quantization (7-bits)

Memory Capacity in RC (role of noise)

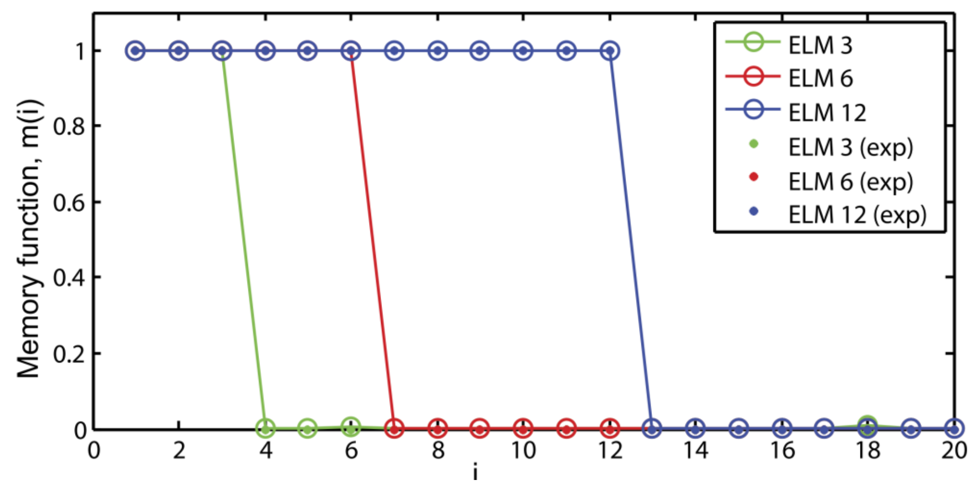




Recurrent connectivity creates intrinsic memory in RC

Finite SNR reduces the memory capacity in RC

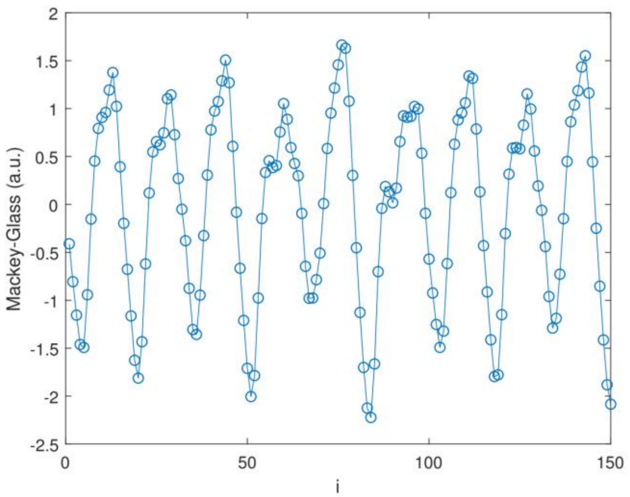
Memory in ELM needs to be explicitly given



Comparison with experiments and numerical simulations:
- Signal-to-noise ratio needs to be considered

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\mathrm{RMS}^2_{signal}}{\mathrm{RMS}^2_{noise}}\right)$$

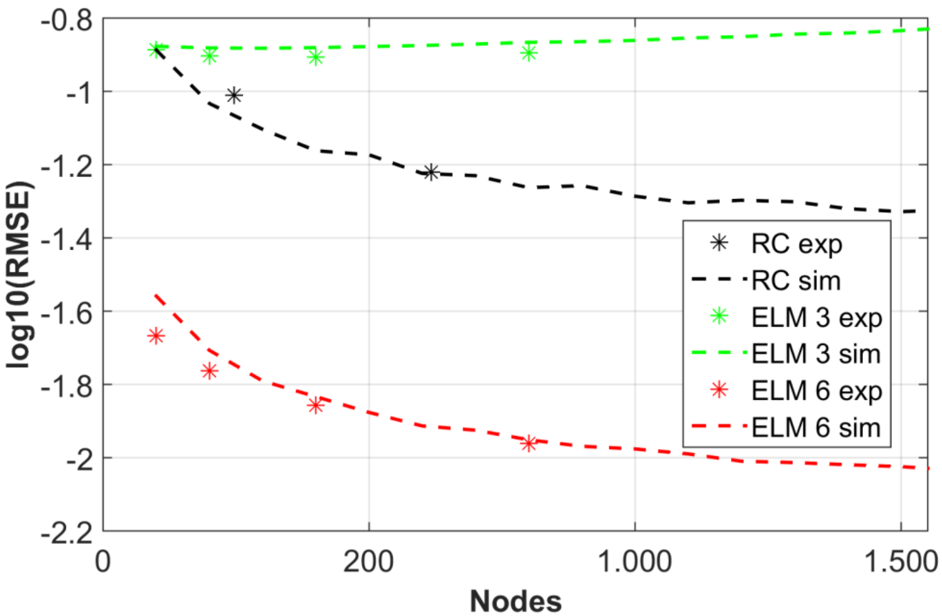
Experimental and numerical* results on the chaotic time series prediction



The quality of the prediction is quantified by the RMSE

Reservoir Computing with optimized parameters

ELM with optimized parameters and different amounts of external memory



RC can perform the chaotic time-series prediction task with good accuracy

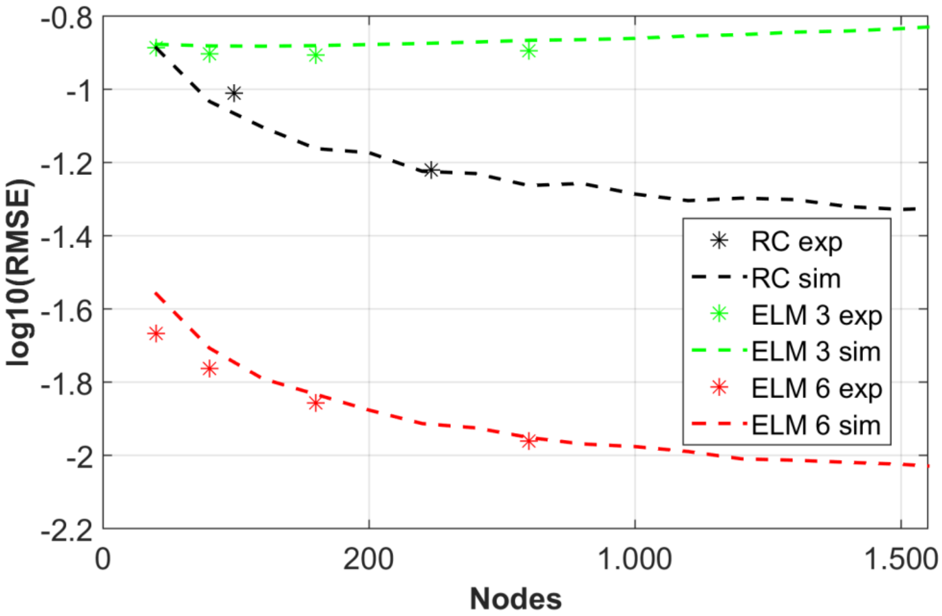
ELM with an external memory of 3 past inputs performs the task with poor accuracy

ELM with an external memory of 6 past inputs performs the task with high accuracy

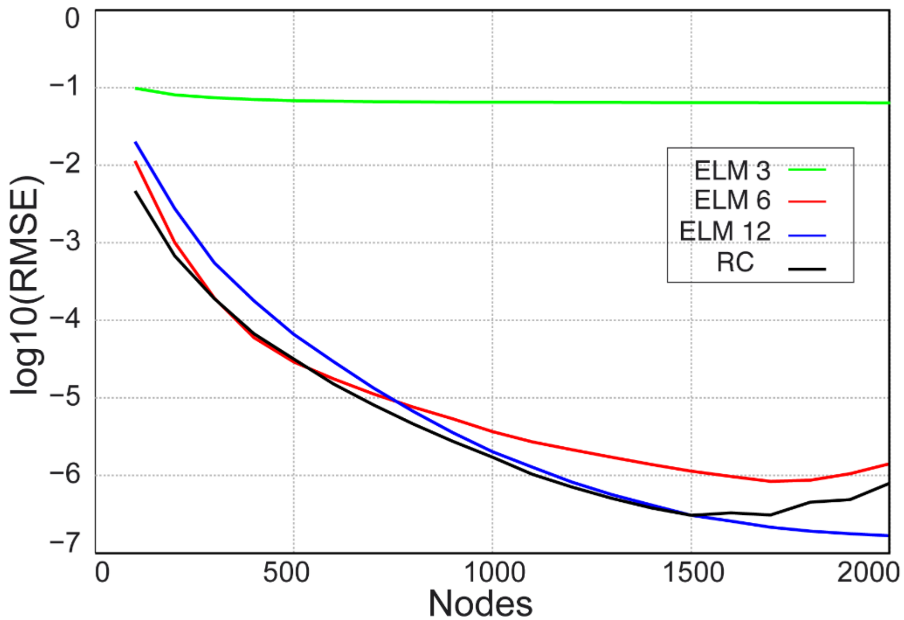
* Numerical simulations consider system noise and quantization (7-bits)

Performance comparison ELM and RC (role of noise)

Experimental and numerical results with system noise and quantization

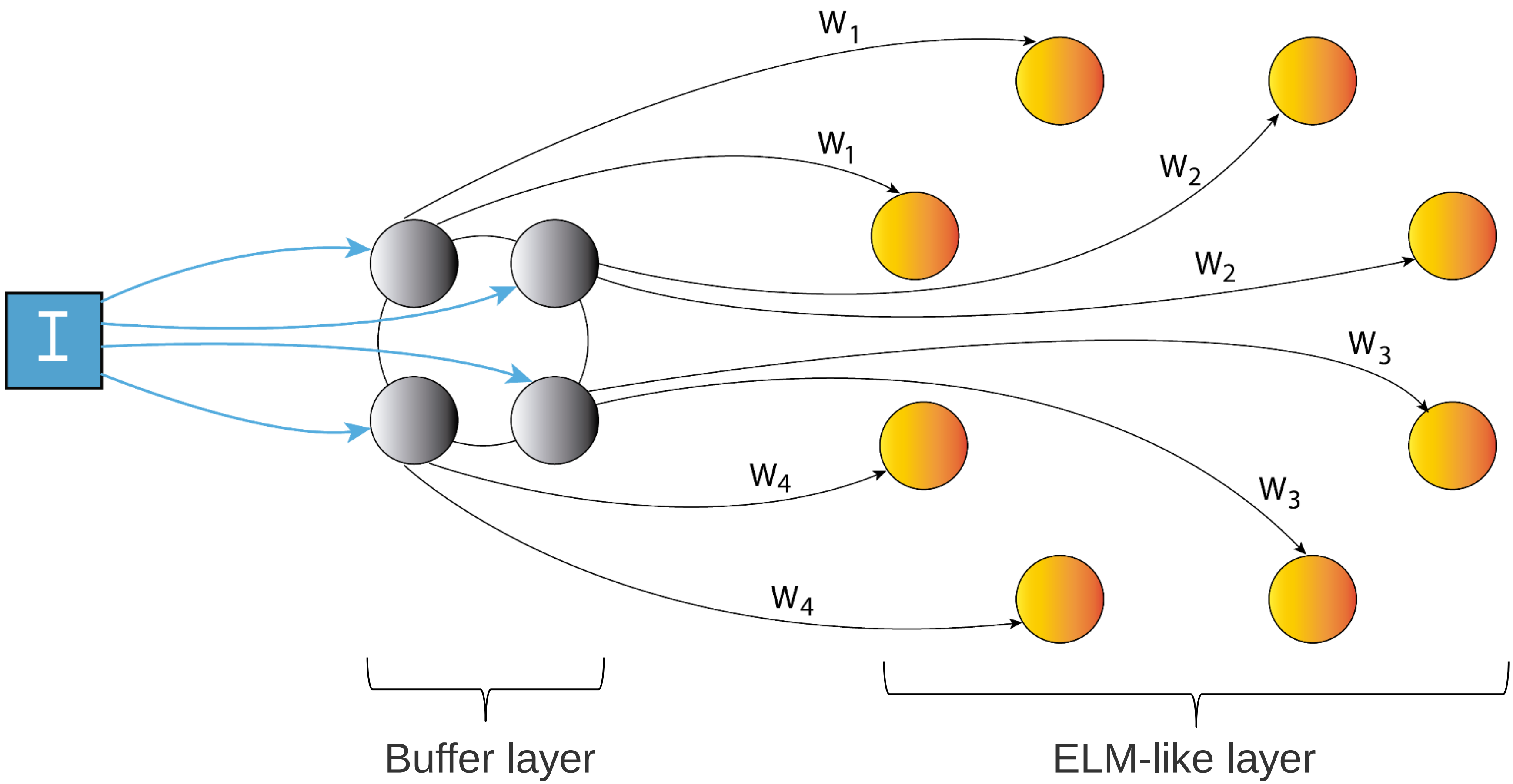


Numerical results without noise



Experimentalists need to worry about bounded performance in presence of noise
Simulators need to worry about transferability of numerical results to experimental platforms

Extension to multi-layer reservoirs

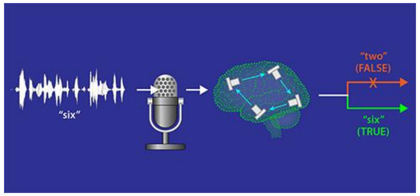


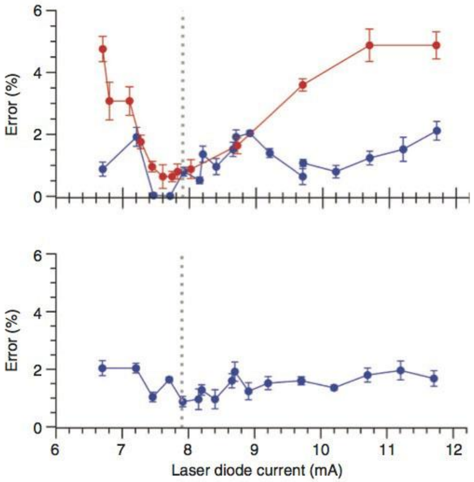
Hybrid approaches may prove to be beneficial in hardware implementations

complexphotonics@gmail.com

- Computation with a photonic system @ Gbyte/s

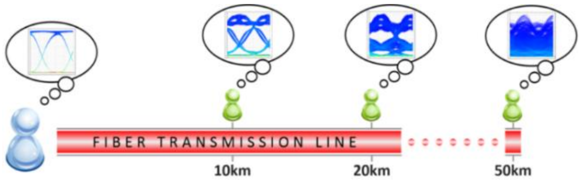
1) Spoken digit recognition (perfect classification of 1Mwords/s)

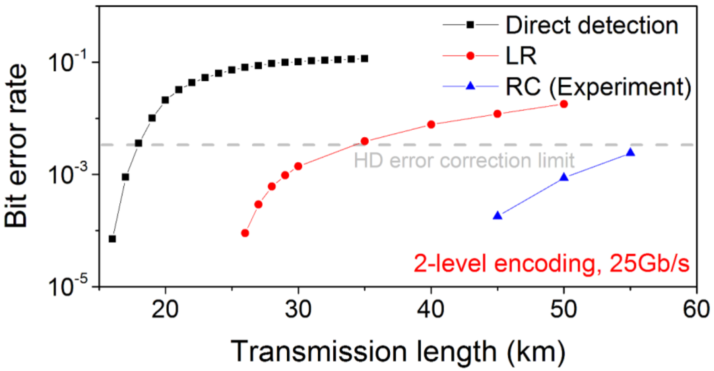




D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, Parallel photonic information processing at gigabyte per second data rates using transient states, Nature Communications 4, 1364(2013).

2) Recovery of optical communication signals (improvement by 1 to 3 orders of magnitude)





A. Argyris, J. Bueno, and I. Fischer, Photonic machine learning implementation for signal recovery in optical communications, Scientific Reports 8, 8487 (2018).

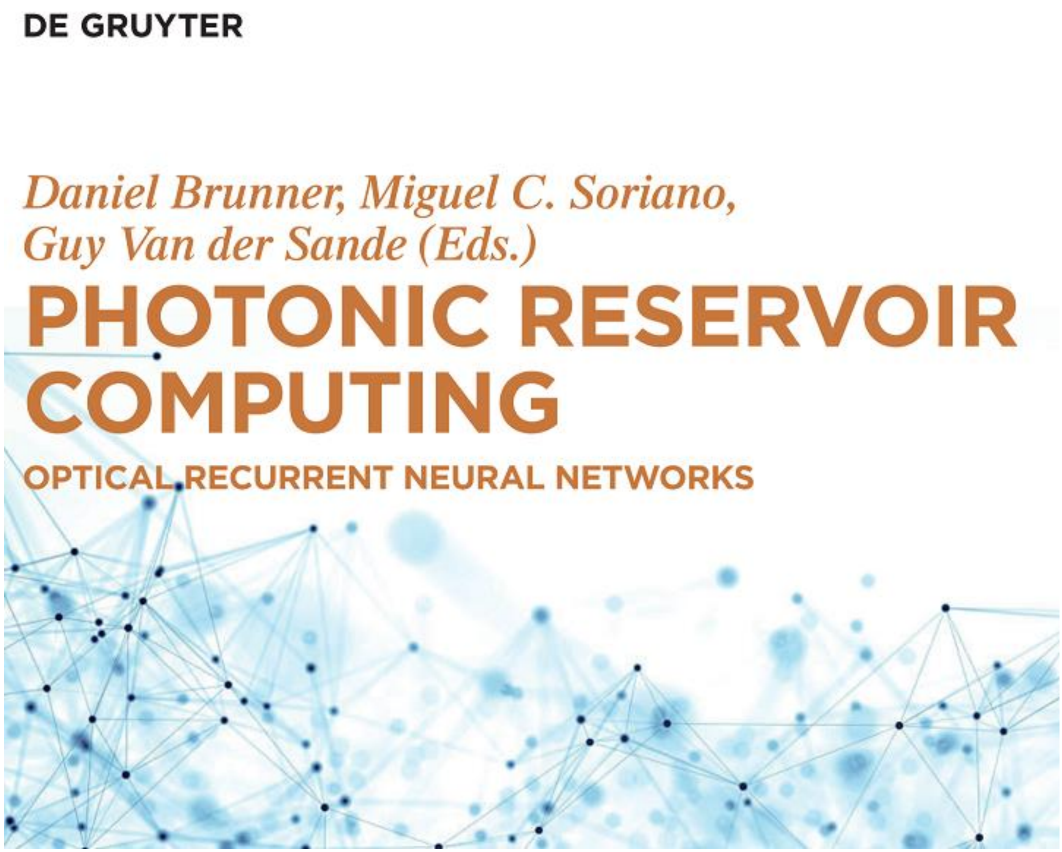
• Machine learning methods with **random mappings** find a natural home in hardware implementations

• The consideration of a **time-multiplexed** input relaxes the hardware requirements of **physical machine learning** implementations

• Time-multiplexed **photonic/optoelectronic systems** have proven to be a **versatile platform** for both ELM and RC

• Extension to **multi-layer** approaches allows for an **extended functionality**

G. Van der Sande, D. Brunner, and M. C. Soriano, “Advances in photonic reservoir computing”, Nanophotonics 6, 561-576 (2017).



Special issue 'Trends in Reservoir Computing' of the journal 'Cognitive Computation'

Paper submission deadline: 31 January 2020

THANK YOU

MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD
AGENCIA ESTATAL DE INVESTIGACIÓN
UNIÓN EUROPEA
FONDO EUROPEO DE DESARROLLO REGIONAL
"Una manera de hacer Europa"
IFISC