# Photonic spiking neural networks using silicon microring resonators

Alessio Lugnan, Stefano Biasi and Lorenzo Pavesi

24 November 2023

UNIVERSITÀ DI TRENTO

NanoLab Department of Physics

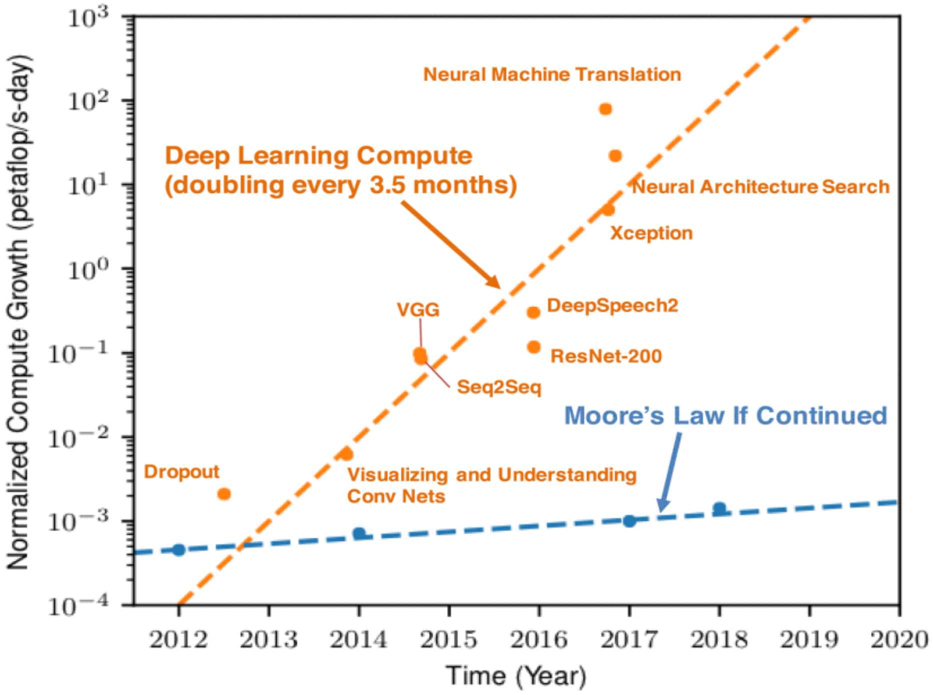
ALPACA Workshop, Trento

# Outline

- **Why** neuromorphic hardware?
- **Why** neuromorphic photonics?
- **Why** silicon microring resonators (MRRs)?
- **Applications**: MMRs as synapses
- **Applications**: MMR-based ANNs
- **Applications**: phase change materials (PCMs)
- **Scalability** of training hardware ANNs
- **Self-adaptive** plasticity with MRRs + PCMs
- **Reservoir computing**: a hardware-friendly approach
- **Applications**: MRRs for reservoir computing
- MRR as a **spiking** neuron
- **Scalability** potential
- Conclusions

# Why neuromorphic hardware?

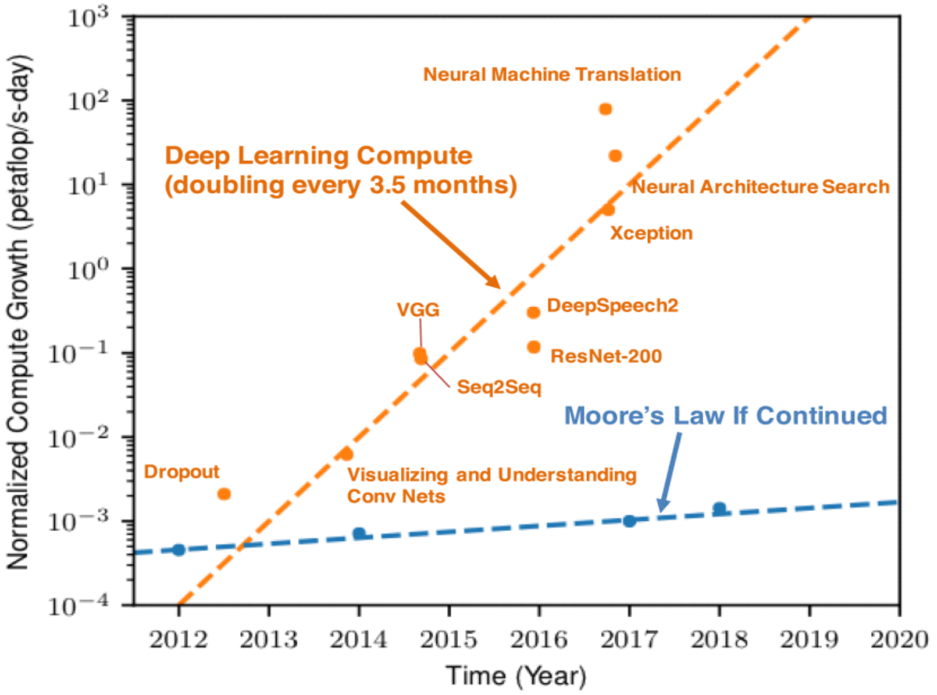
Deep learning is quickly saturating available computing growth




De Lima, Thomas Ferreira, et al. "Machine learning with neuromorphic photonics." Journal of Lightwave Technology 37.5 (2019)

# Why neuromorphic hardware?

Deep learning is quickly saturating available computing growth




De Lima, Thomas Ferreira, et al. "Machine learning with neuromorphic photonics." Journal of Lightwave Technology 37.5 (2019)

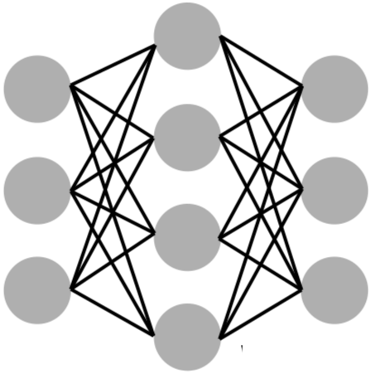
Projected contribution to global economy by 2030 from AI: **$15.7tr** (www.pwc.com)

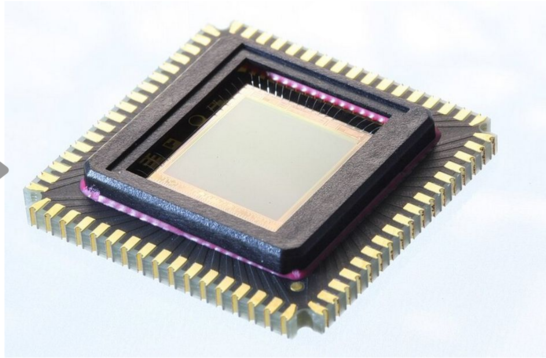
Nowadays: AI is in data centres (cloud computing)

**Current growth is far from sustainable**

Solution: ~~Von Neuman architecture~~ → neuromorphic

- machine learning accelerators
- neuromorphic processors for edge computing





https://commons.wikimedia.org/wiki/File:PMD_Imager.jpg

UNIVERSITÀ DI TRENTO

NanoLab Department of Physics

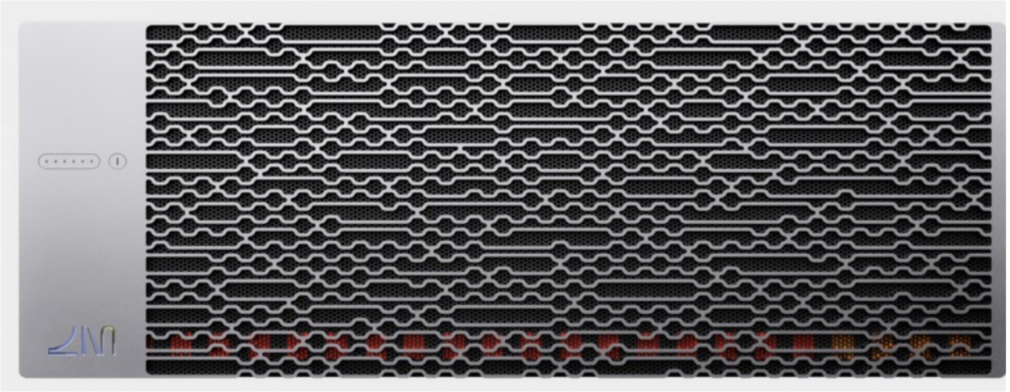
# Why neuromorphic photonics?

Wavelength division multiplexing, low power dissipation,
no parasitic capacitance, ...
→ parallel, energy efficient, low-latency **linear operations**

High-tech industry is now interested in photonics

AI accelerators: perform general matrix multiplications

Sartup Lightmatter (>100M funding to accelerate ANNs)



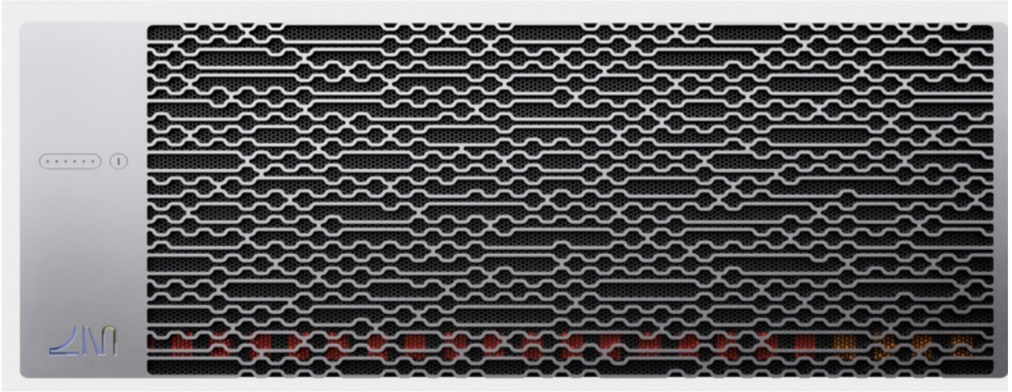
# Why neuromorphic photonics?

Wavelength division multiplexing, low power dissipation, no parasitic capacitance, ...
→ parallel, energy efficient, low-latency **linear operations**

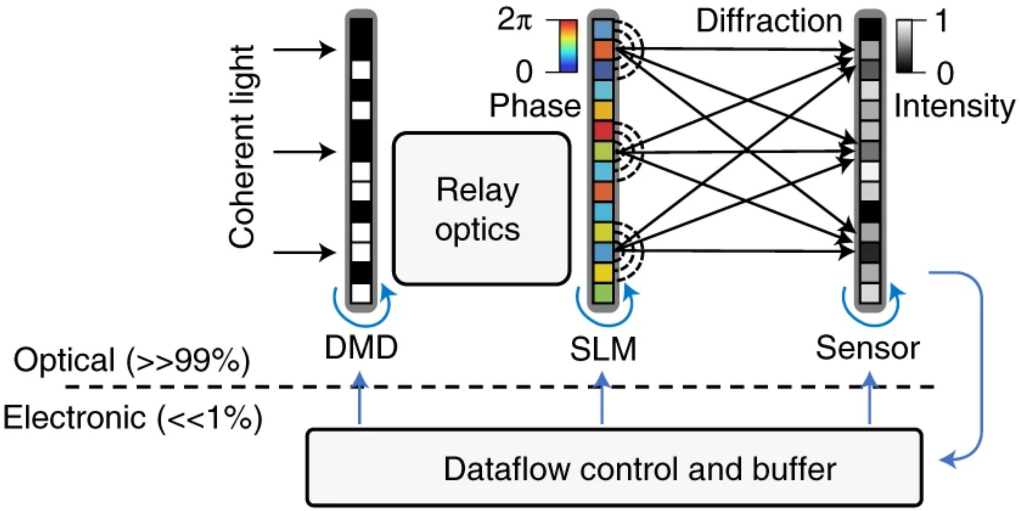
High-tech industry is now interested in photonics

AI accelerators: perform general matrix multiplications

Sartup Lightmatter (>100M funding to accelerate ANNs)



Photonics-based ANN outperforms cutting-edge electronics in energy efficiency and speed



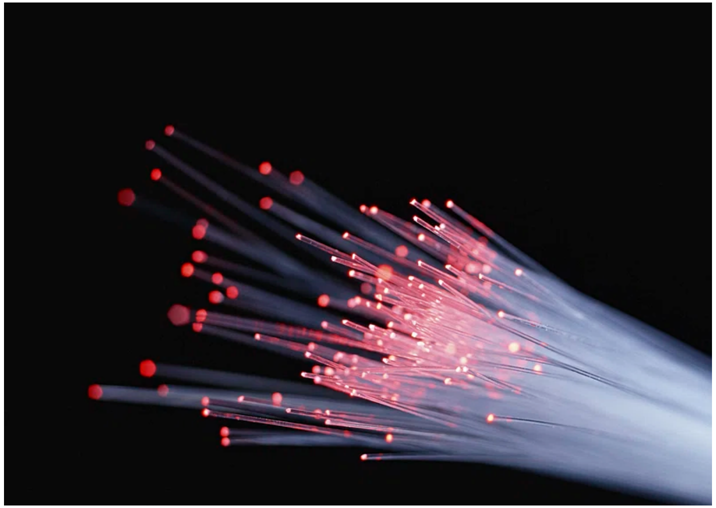

Optoelectronic implementation

Zhou, Tiankuang, et al. "Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit." Nature Photonics 15.5 (2021)

# Why neuromorphic photonics?

Advantage in processing **signals originally in the optical domain** → no Optical-Electrical conversion:

- internet data through optical fiber
- fiber sensing applications
- optics-based biomedical sensing
- ...



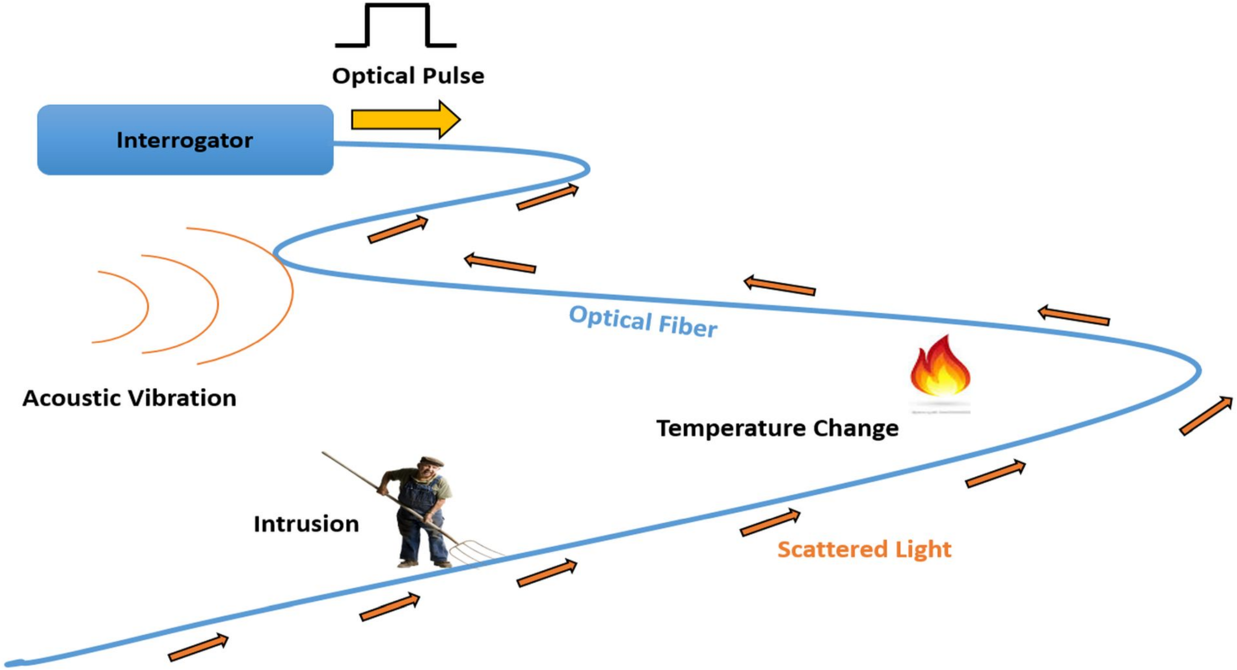
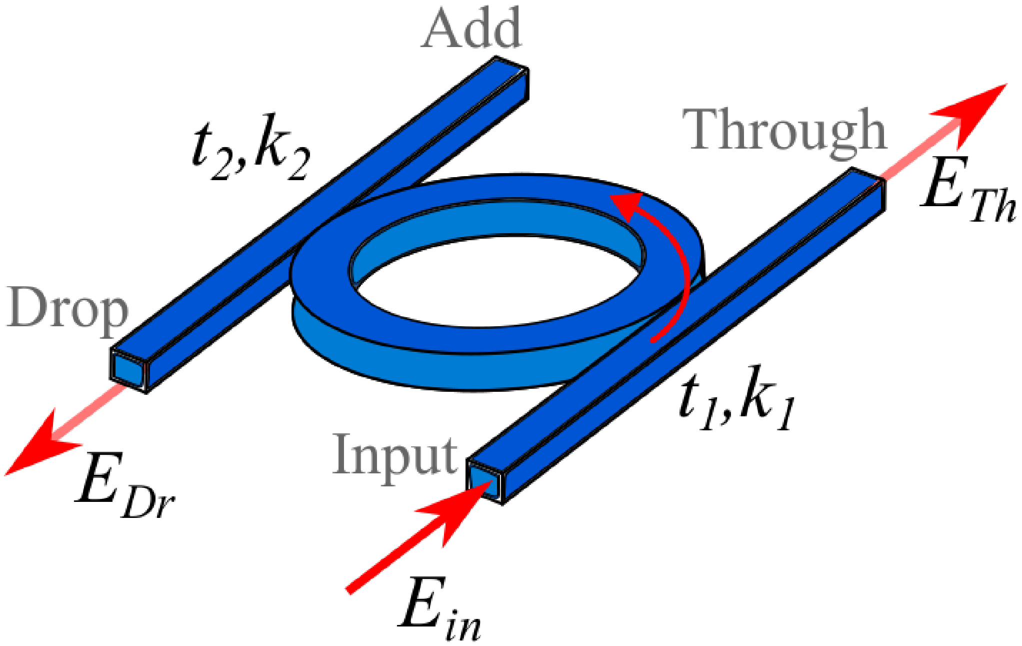



Image from website of Polytechnique Fédérale de Lausanne (EPFL, Switzerland)
https://www.epfl.ch/

# Why silicon microring resonators (MRRs)?

Infrared light (telecom wavelength: 1550 nm)




Optical resonance:

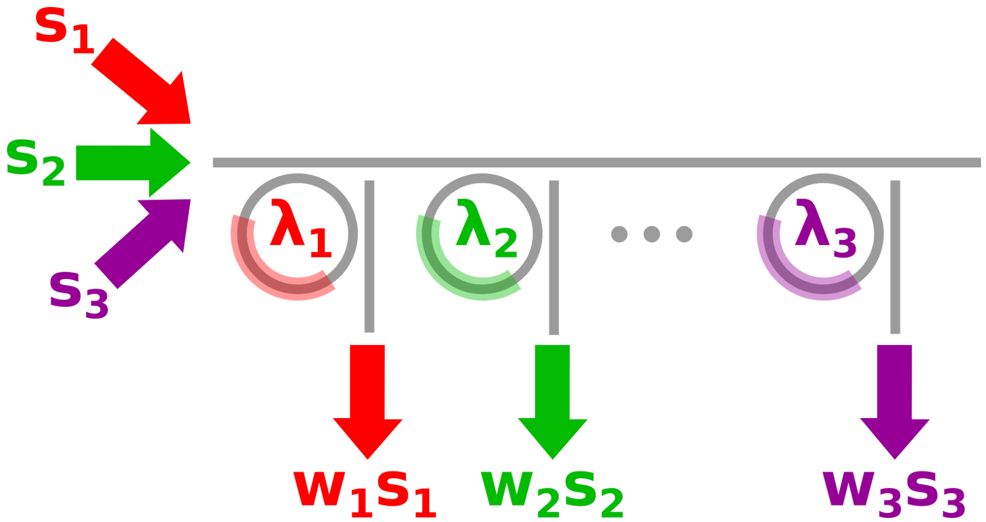
- only specific wavelengths enter the ring
- light power accumulates
- enhanced sensitivity to light path perturbations

Image from:
Biasi S, Donati G, Lugnan A, Mancinelli M, Staffoli E, Pavesi L., “Photonic neural networks based on integrated silicon microresonators.” arXiv preprint (2023)

# Why silicon microring resonators (MRRs)?

MRRs can be tuned to **separate and weight multiple wavelengths** in the same channel (WDM, short for *wavelength division multiplexing*)

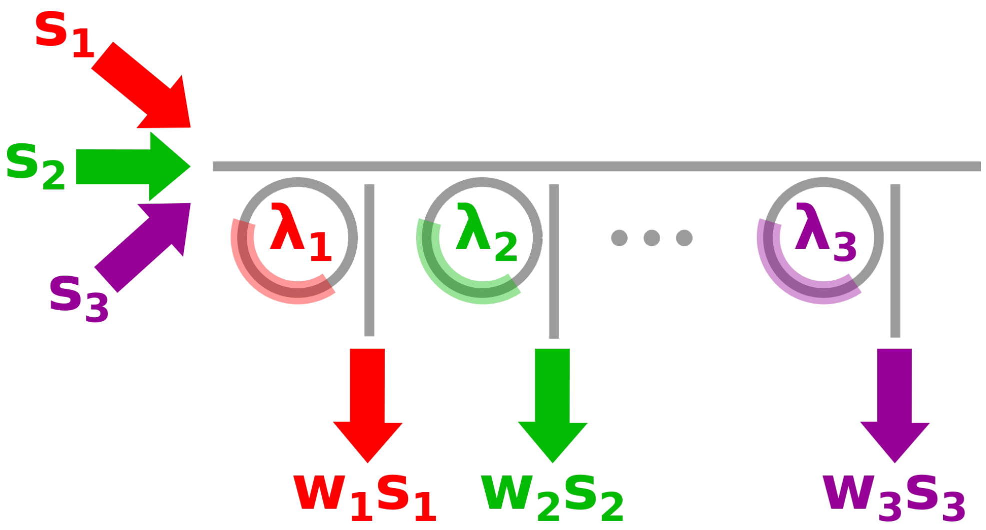
→ **artificial synapse**



# Why silicon microring resonators (MRRs)?

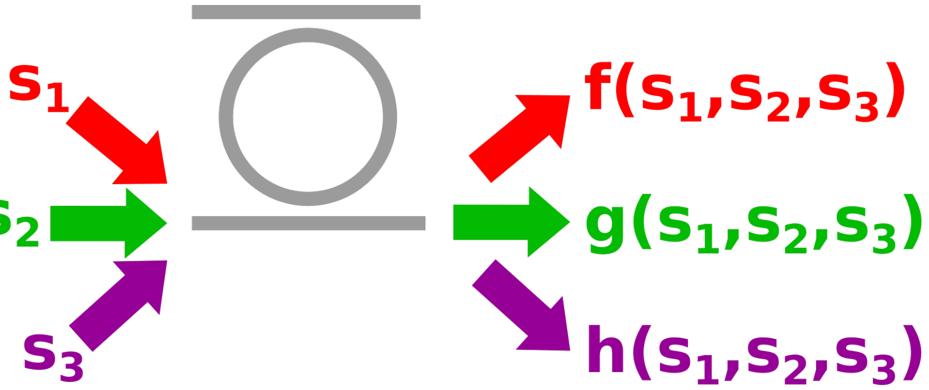
MRRs can be tuned to **separate and weight multiple wavelengths** in the same channel (WDM, short for *wavelength division multiplexing*)

→ **artificial synapse**




A MRR can **nonlinearly couple** multiple wavelengths

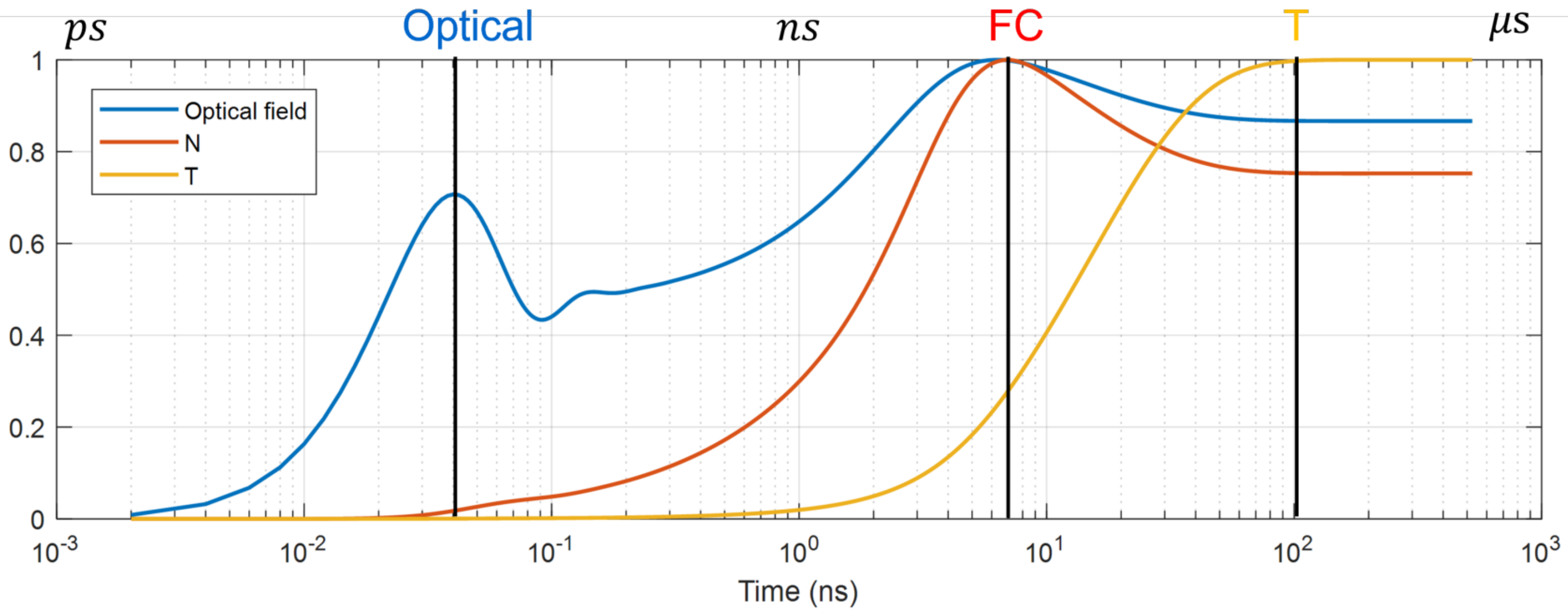
→ **artificial neuron**




Few mW power

# Why silicon microring resonators (MRRs)?

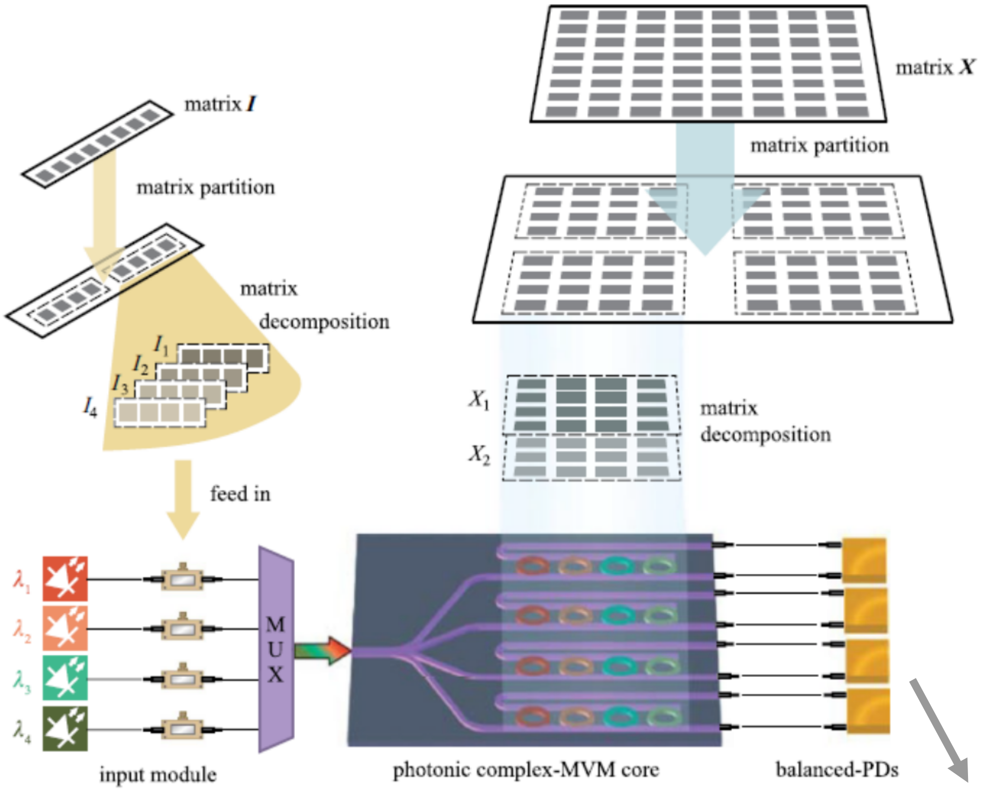
Multiscale volatile memory → **short- and long- term plasticity**




Biasi S, Donati G, Lugnan A, Mancinelli M, Staffoli E, Pavesi L., “Photonic neural networks based on integrated silicon microresonators.” arXiv preprint (2023)

# Applications: MMRs as synapses

Hardware **accelerators** for large ANNs (like TPUs)
→ “only” linear **matrix-vector multiplications**



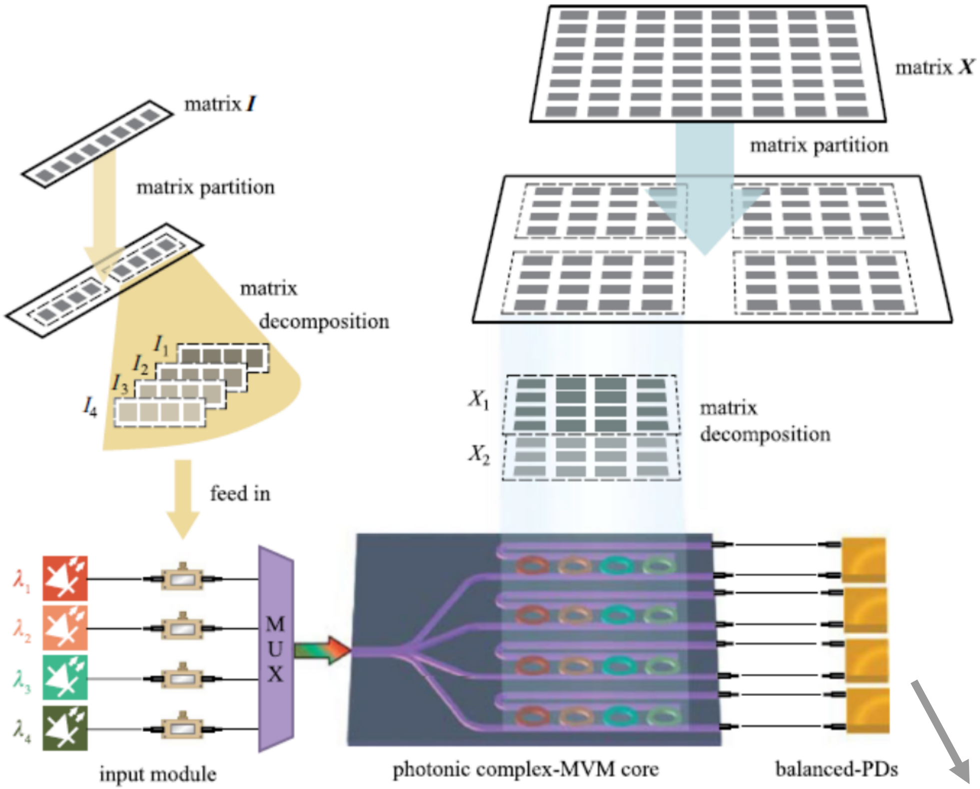

UNIVERSITÀ DI TRENTO

NanoLab
Department of Physics

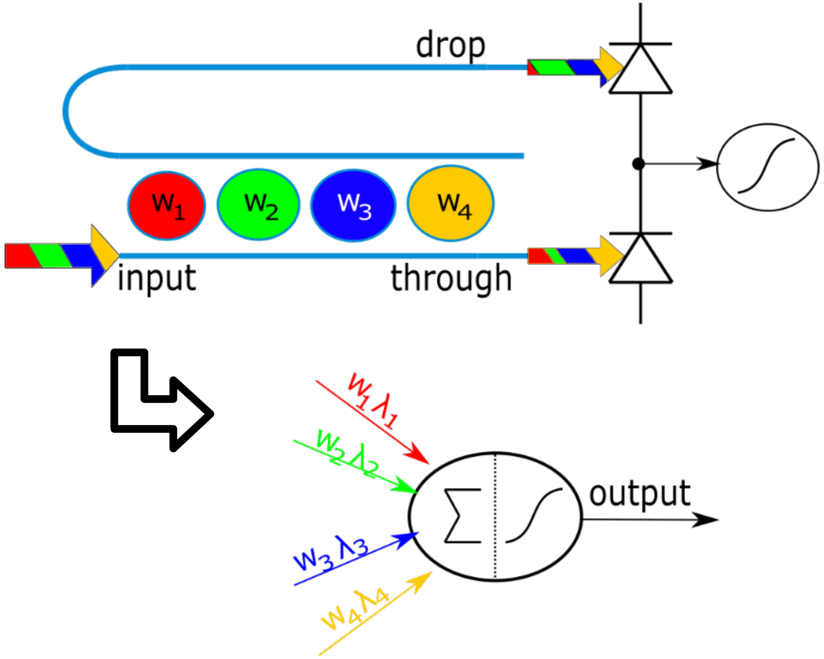
J. Cheng et al., “A small microring array that performs large complex-valued matrix-vector multiplication,” Frontiers of Optoelectronics, (2022).

# Applications: MMRs as synapses

Hardware **accelerators** for large ANNs (like TPUs)
→ “only” linear **matrix-vector multiplications**




**Photonic synapses + neuron**



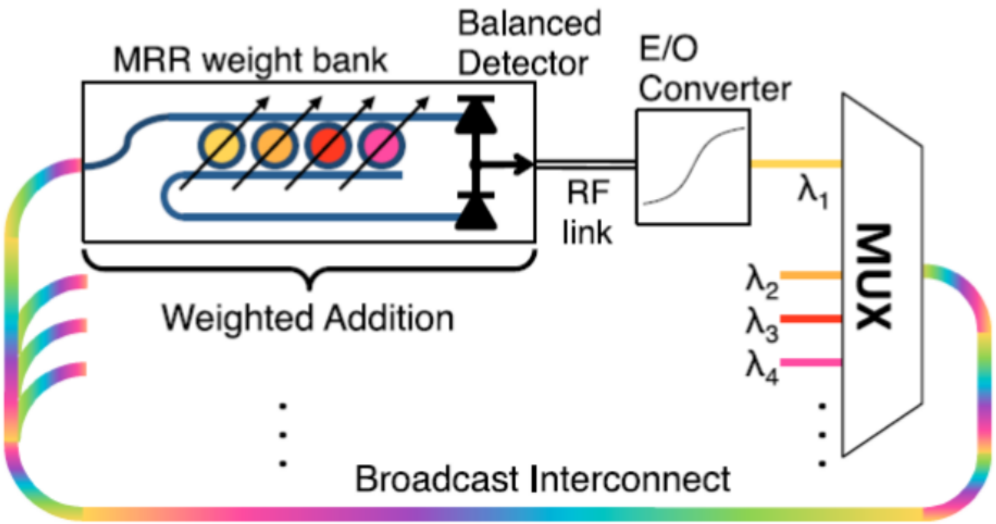

A. Tait et al., “Balanced wdm weight banks for analog optical processing and networking in silicon,” IEEE SUM, (2015)

J. Cheng et al., “A small microring array that performs large complex-valued matrix-vector multiplication,” Frontiers of Optoelectronics, (2022).

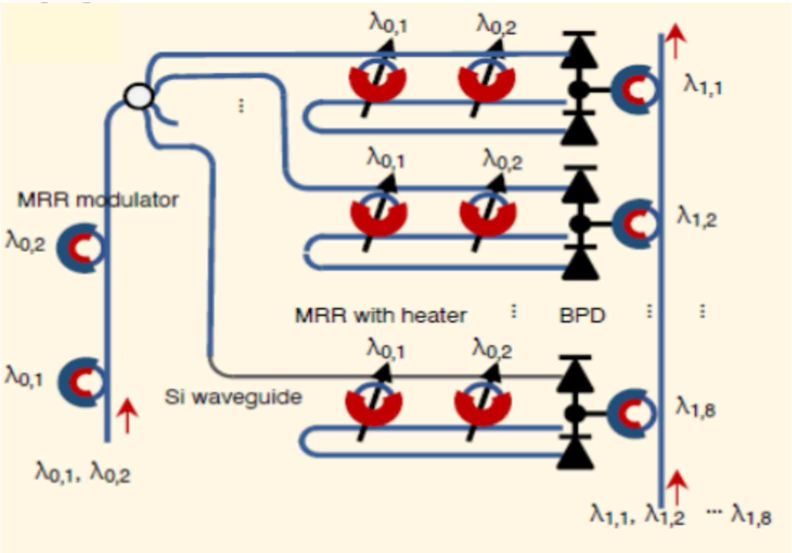
# Applications: MMR-based ANNs

On-chip photonic neural network:

- high-speed processing
- multiple input and outputs in one waveguide
- potentially high energy efficiency



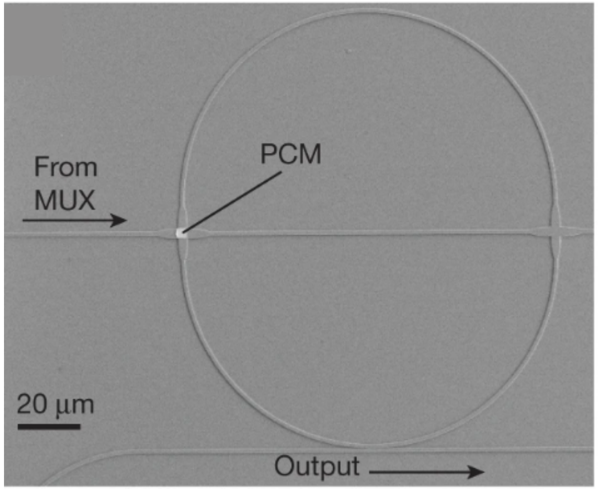

Tait, A. N., et al. "Microring weight banks." IEEE Journal of Selected Topics in Quantum Electronics, (2016).



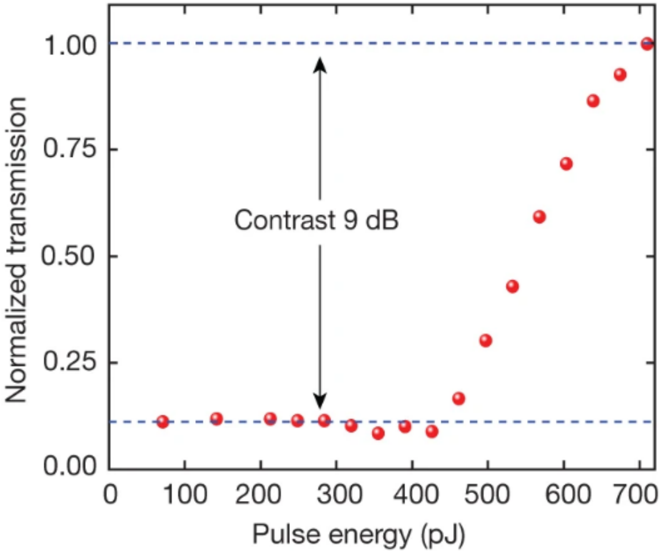

Huang, Chaoran, et al. "A silicon photonic–electronic neural network for fibre nonlinearity compensation." Nature Electronics, (2021).

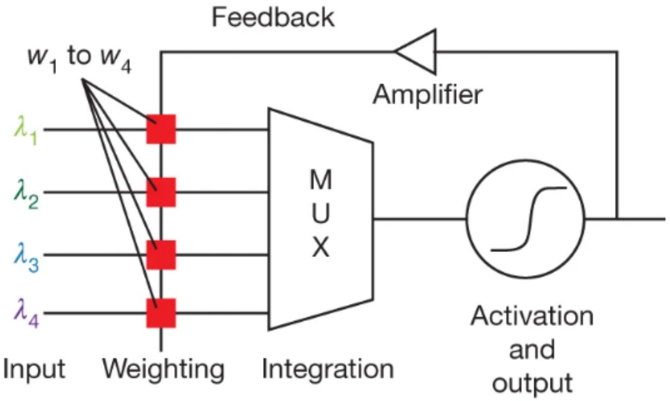
# Applications: phase change materials (PCMs)

Tuning MRRs with PCMs: integrated **non-volatile photonic memory** instead of heaters
→ much higher energy efficiency
→ no thermal corss-talk




All-optical activation function




Feldmann, J., et al. "All-optical spiking neurosynaptic networks with self-learning capabilities." Nature, (2019).

Spike time dependent plasticity (**STDP**) for unsupervised on-chip self-learning

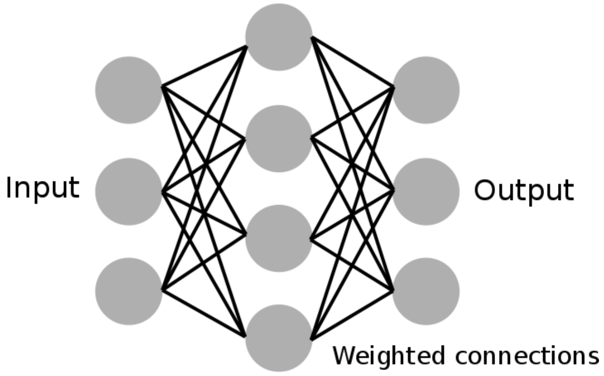
Still **difficult to scale up!**

# Scalability of training hardware ANNs

Nowadays, neuromorphic computing systems are mainly **trained externally**, running **backpropagation** and **gradient descent** on a computer, requiring:

- control of parameters (weights)
- neuron states observability

↳ Not biologically plausible and **not scalable**!

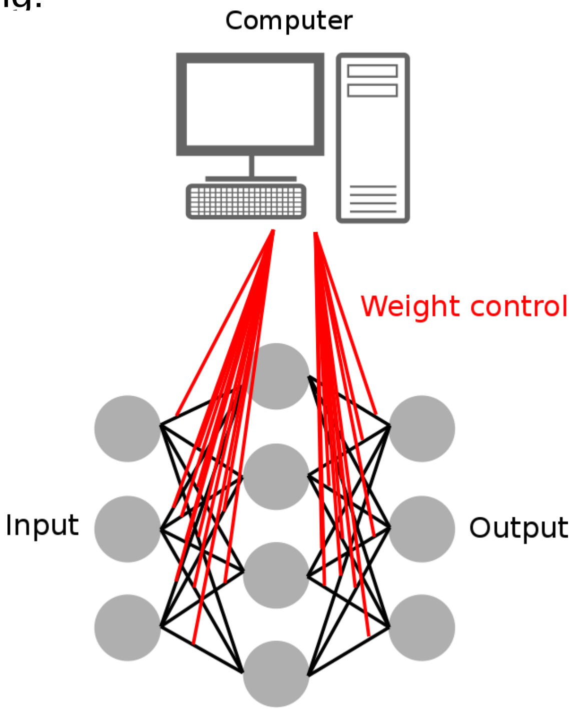


# Scalability of training hardware ANNs

Nowadays, neuromorphic computing systems are mainly **trained externally**, running **backpropagation** and **gradient descent** on a computer, requiring:

- control of parameters (weights)
- neuron states observability

↳ Not biologically plausible and **not scalable**!

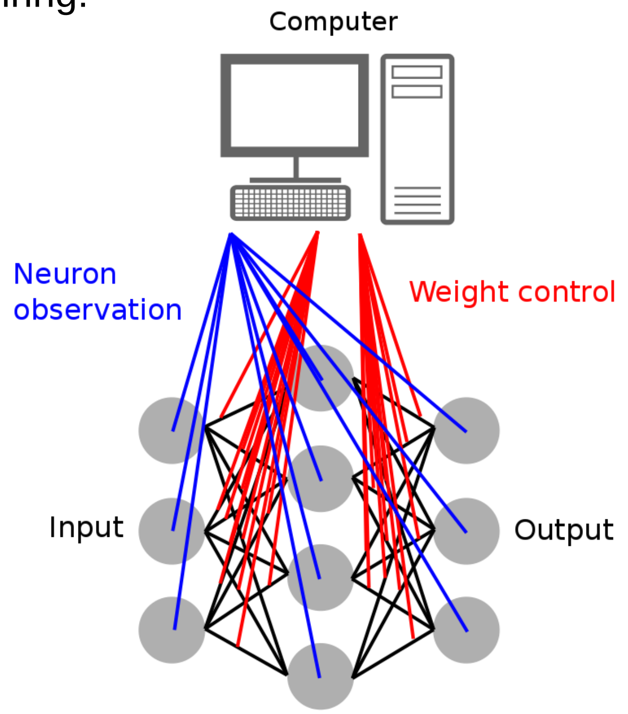


# Scalability of training hardware ANNs

Nowadays, neuromorphic computing systems are mainly **trained externally**, running **backpropagation** and **gradient descent** on a computer, requiring:

- control of parameters (weights)
- neuron states observability

↳ Not biologically plausible and **not scalable**!

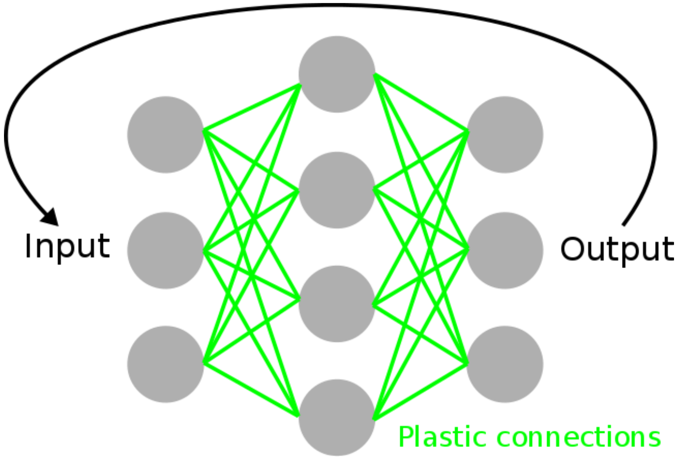


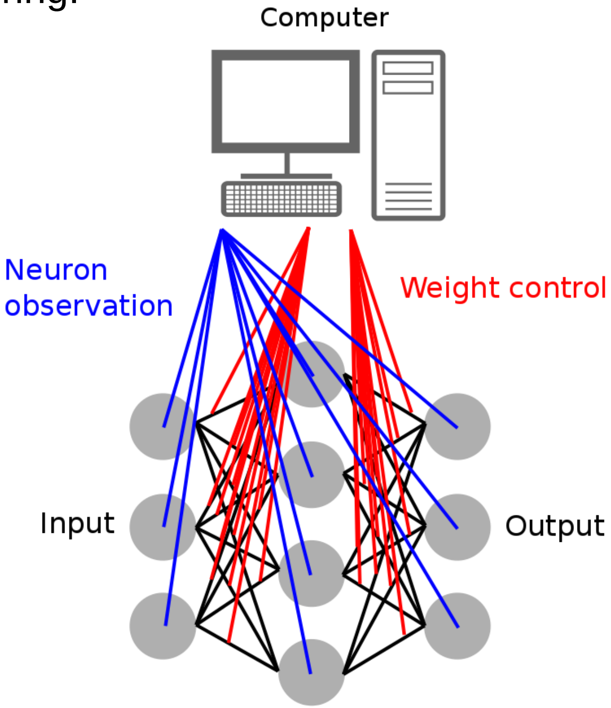
# Scalability of training hardware ANNs

Nowadays, neuromorphic computing systems are mainly **trained externally**, running **backpropagation** and **gradient descent** on a computer, requiring:

- control of parameters (weights)
- neuron states observability

↳ Not biologically plausible and **not scalable**!



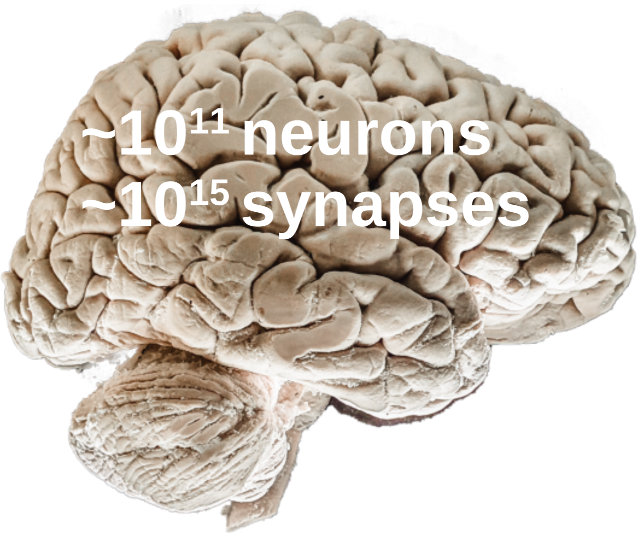

Learning via **emergent self-adaptation** is the holy grail!

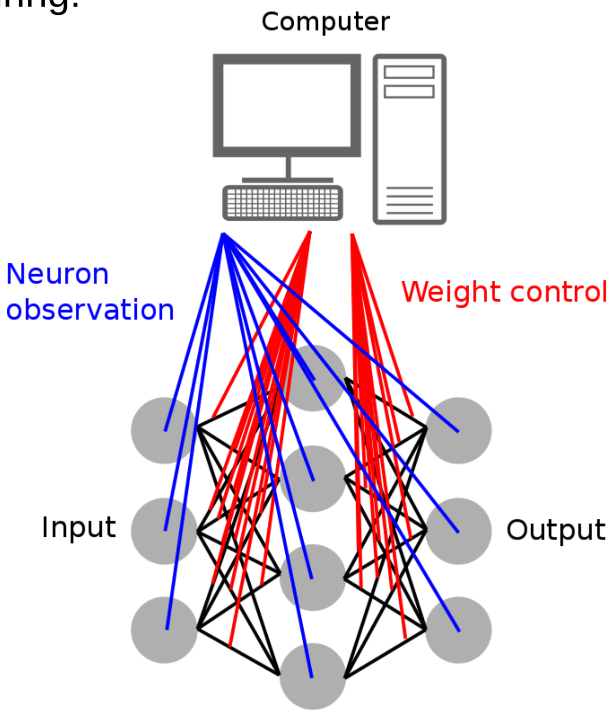
# Scalability of training hardware ANNs

Nowadays, neuromorphic computing systems are mainly **trained externally**, running **backpropagation** and **gradient descent** on a computer, requiring:

- control of parameters (weights)
- neuron states observability

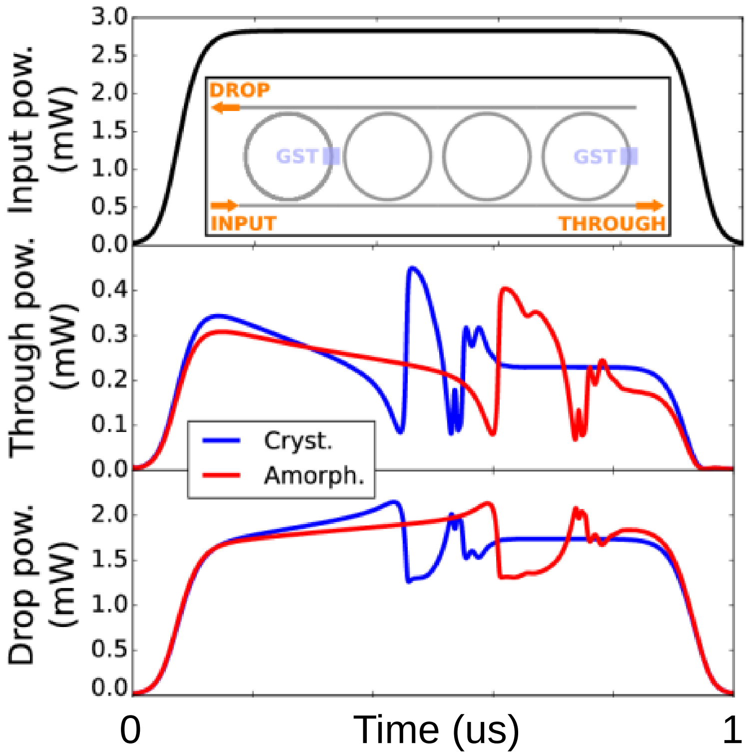
↳ Not biologically plausible and **not scalable**!



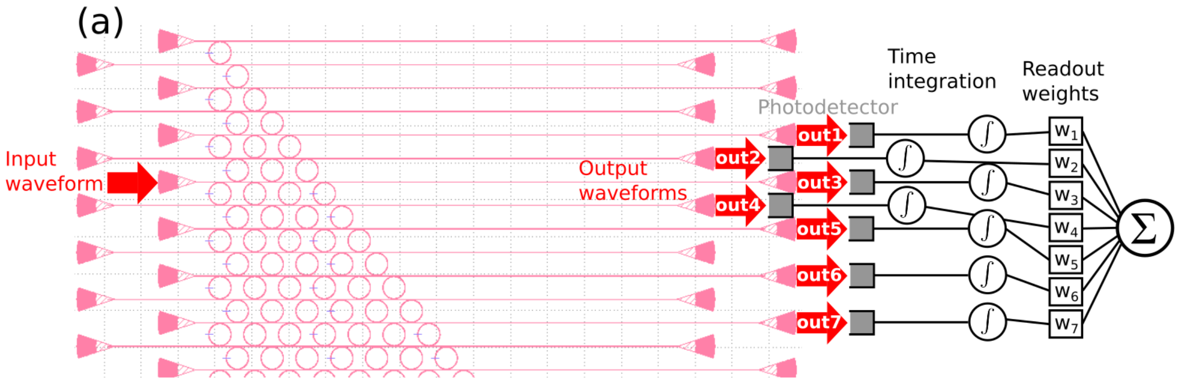



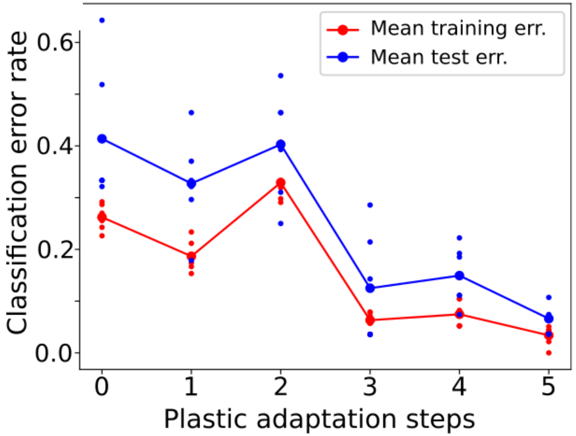

Learning via **emergent self-adaptation** is the holy grail!

# Self-adaptive plasticity with MRRs + PCMs




A. Lugnan, S. G.-C. Carrillo, C. D. Wright, and P. Bienstman, "Rigorous dynamic model of a silicon ring resonator with phase change material for a neuromorphic node," Optics Express, (2022)

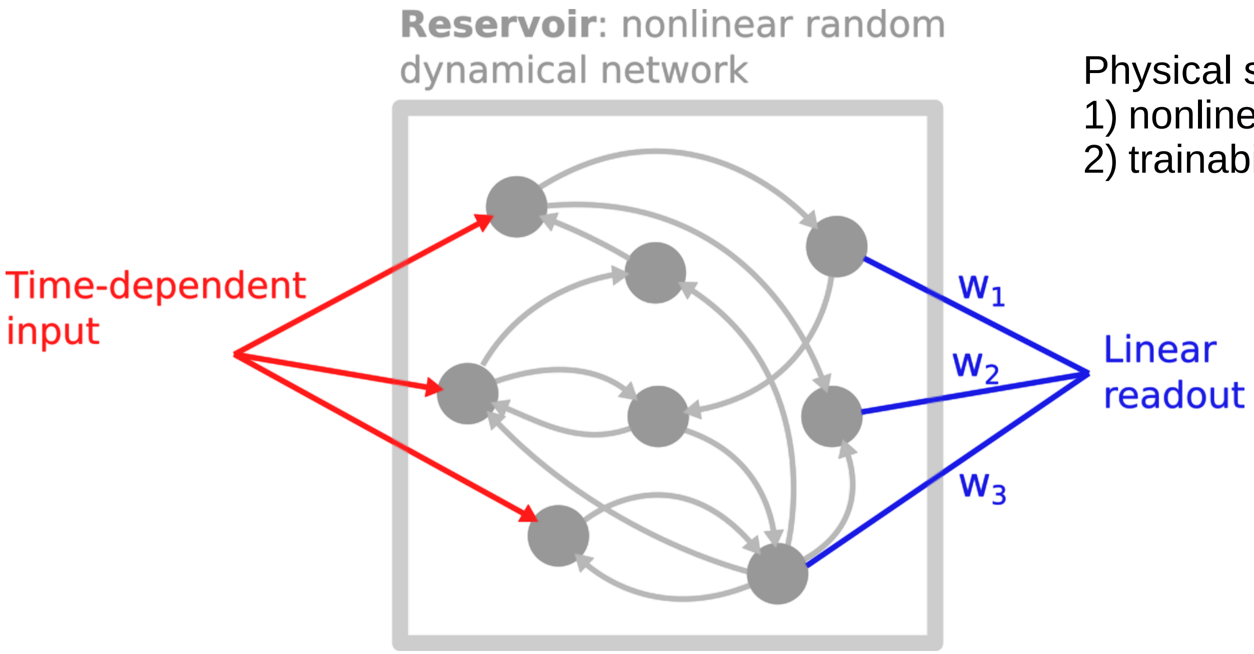



There is no suitable theory on self-learning based on plasticity, yet

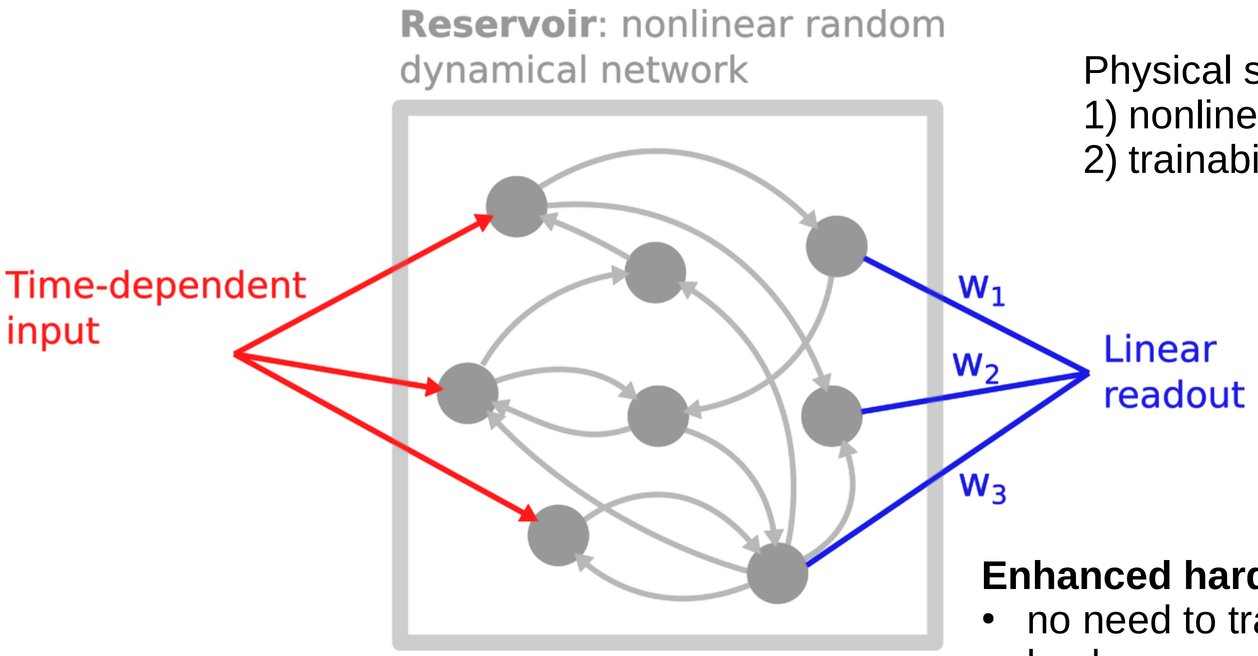
Lugnan A, Aggarwal S, Brückerhoff-Plückelmann F, Pernice WH, Bhaskaran H, Bienstman P. "Performance enhancement via synaptic plasticity in an integrated photonic recurrent neural network with phase-change materials". European Quantum Electronics Conference, (2023).

# Reservoir computing: a hardware-friendly approach




Physical separation of ANN tasks:
1) nonlinearity + dynamics → reservoir (physical)
2) trainability → single-layer readout (software)

# Reservoir computing: a hardware-friendly approach

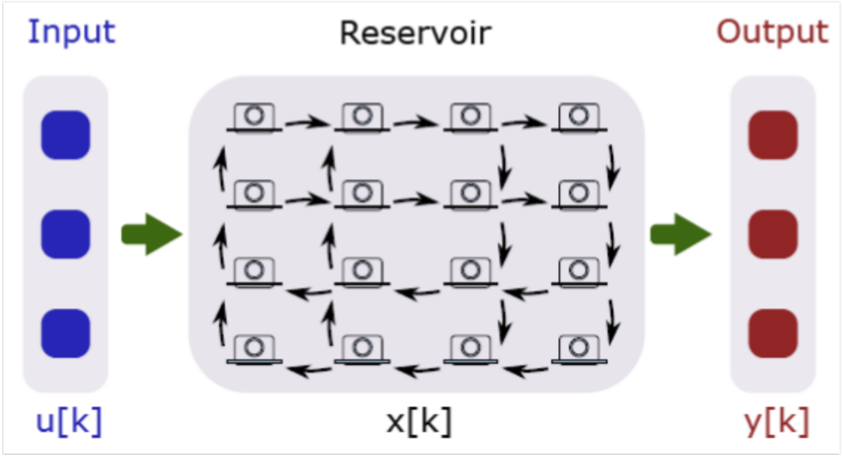



Physical separation of ANN tasks:
1) nonlinearity + dynamics → reservoir (physical)
2) trainability → single-layer readout (software)

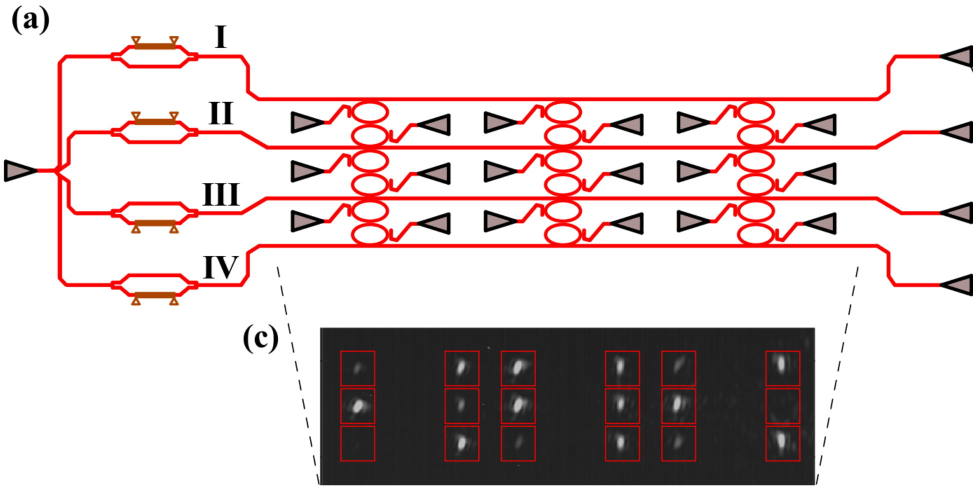
**Enhanced hardware scalability**:
- no need to train the reservoir parameters
- backpropagation-free (no need to know error gradients)

… but limited computational power because of lack of learnable hierarchy

# Applications: MRRs for reservoir computing

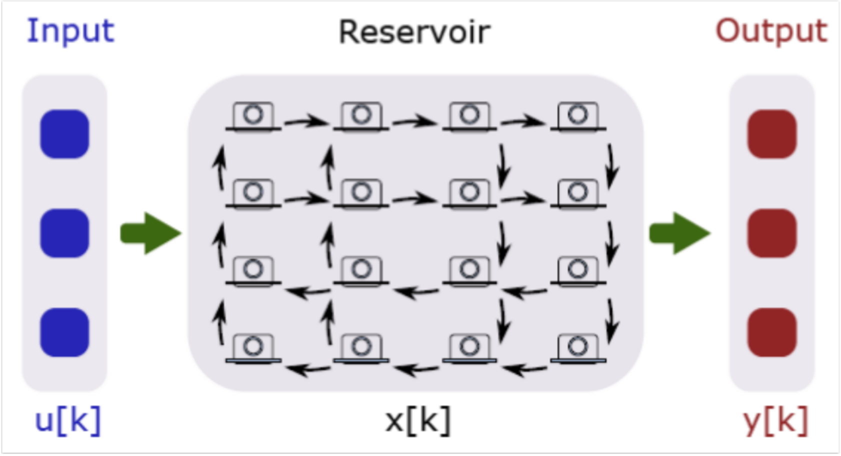



F. Denis-Le Coarer et al., “All-optical reservoir computing on a photonic chip using silicon-based ring resonators,” IEEE Journal of Selected Topics in Quantum Electronics, (2018)

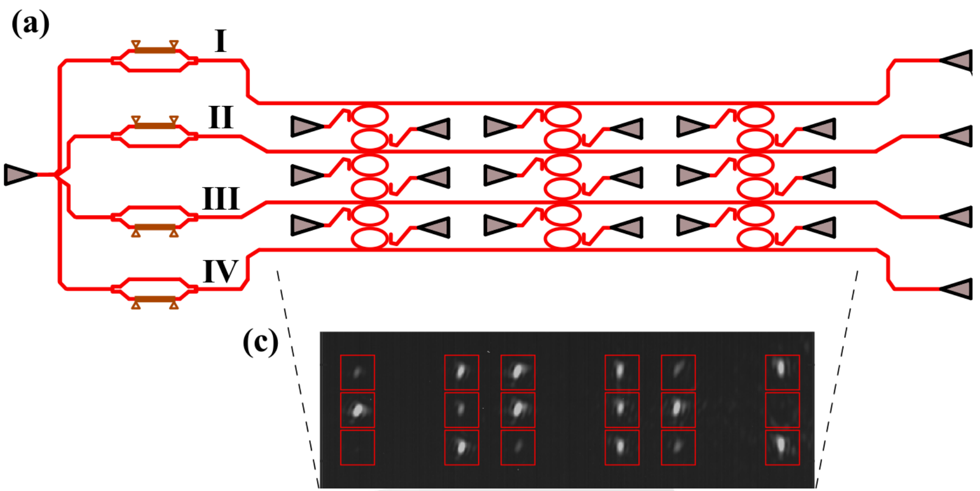



S. Biasi, R. Franchi, L. Cerini, L. Pavesi; “An array of microresonators as a photonic extreme learning machine”. APL Photonics, (2023).

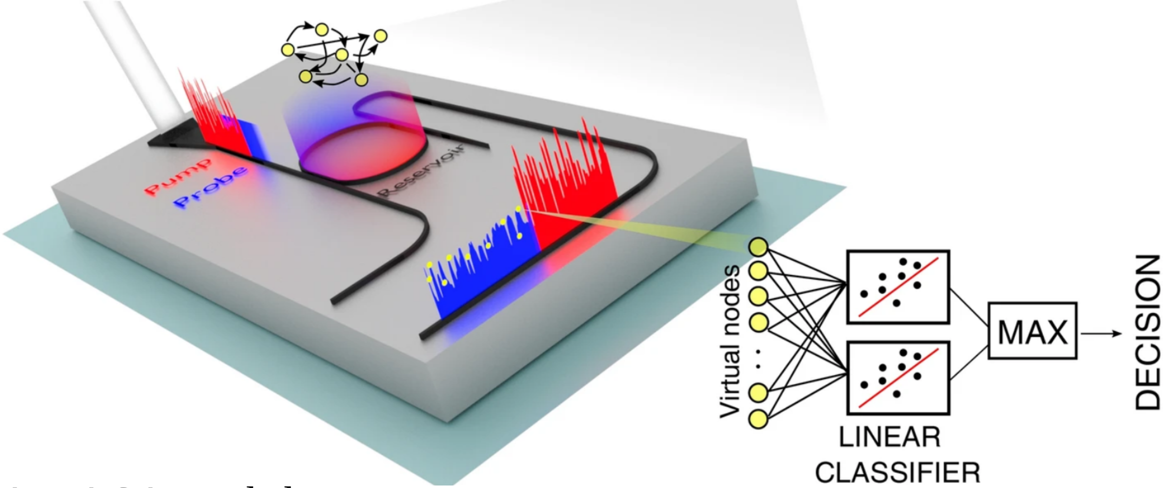
# Applications: MRRs for reservoir computing




F. Denis-Le Coarer et al., “All-optical reservoir computing on a photonic chip using silicon-based ring resonators,” IEEE Journal of Selected Topics in Quantum Electronics, (2018)



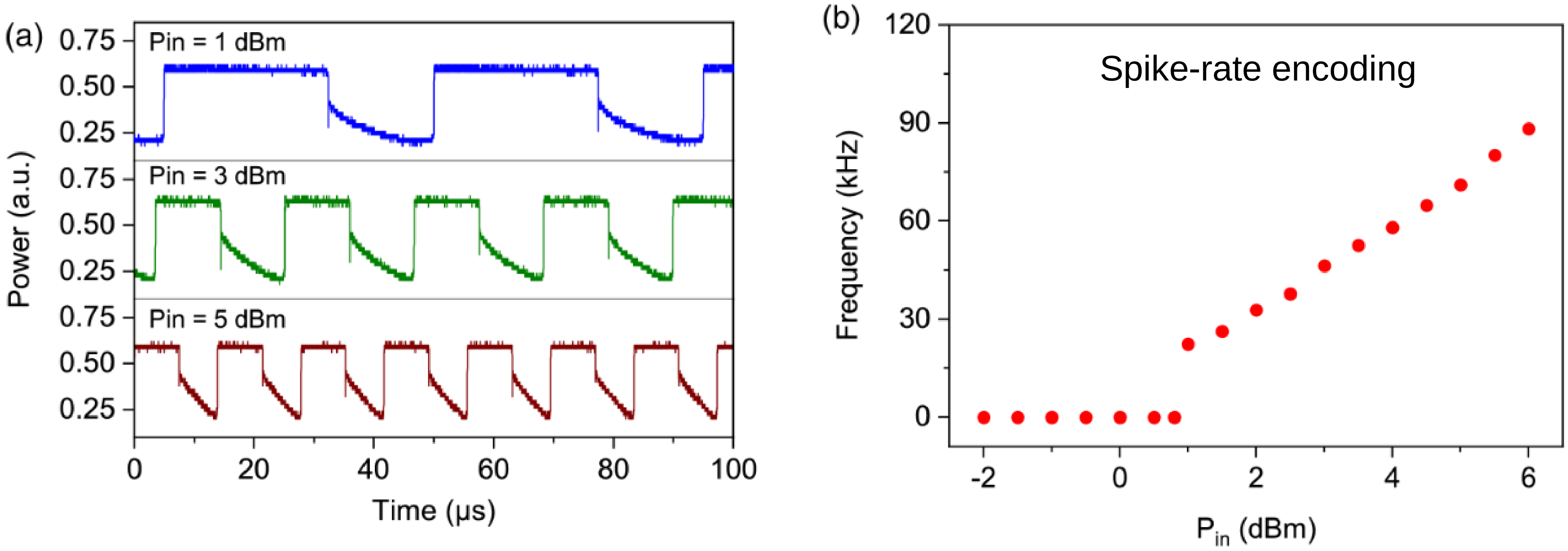

Still, relatively **small networks**




Borghi, M., Biasi, S., and Pavesi, L.. "Reservoir computing based on a silicon microring and time multiplexing for binary and analog operations." Scientific Reports, (2021).

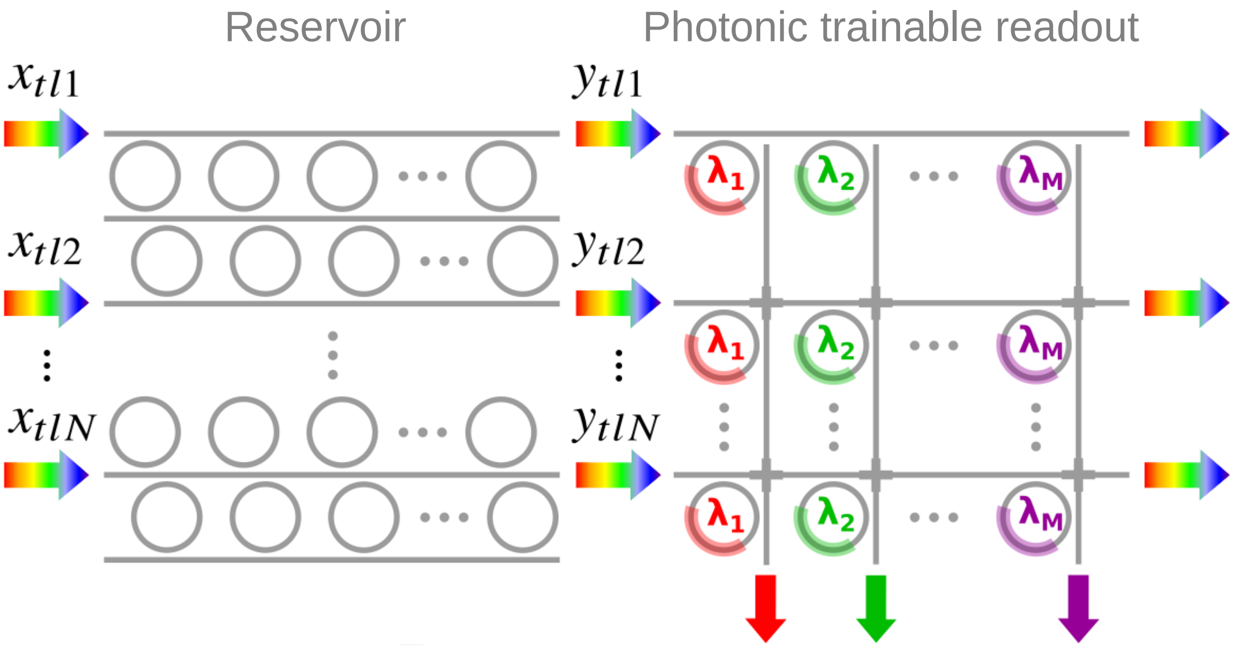
# MRR as a spiking neuron

Self-pulsing regime: input constant power → output pulses (**spiking**)




Jinlong Xiang, Yujia Zhang, Yaotian Zhao, Xuhan Guo, and Yikai Su, "All-optical silicon microring spiking neuron," Photon. Res., (2022)

# Scalability potential




Biasi, S., Donati, G., Lugnan, A., Mancinelli, M., Staffoli, E., & Pavesi, L., “Photonic neural networks based on integrated silicon microresonators”. arXiv preprint. Accepted in *Intelligent Computing SPJ*. (2023).

Input and output encoded and nonlinearly coupled in:

- **space** (physical network)
- **time** (nonlinear dynamics)
- **wavelength** (resonances of MRRs)

→ input-output neurons: ~ $10^4$ x $10^4$ per 100 ns per 0.5 mm$^2$ chip area

- throughput:
  $10^{15}$ (MACs+NLOs)/s/mm$^2$
- energy efficiency:
  $10^{16}$ (MACs+NLOs)/J

# Conclusion

*Silicon microring resonators* are very **versatile** and promising components for efficient and **scalable** integrated photonics neural networks.

# Conclusion

*Silicon microring resonators* are very **versatile** and promising components for efficient and **scalable** integrated photonic neural networks.

Thank you for your attention!

UNIVERSITÀ DI TRENTO

NanoLab Department of Physics