Exploiting Multiple Scattering of Light for Computing

Sylvain Gigan

Optica Technical Group Webinar

June 8th 2023

Laboratoire Kastler Brossel
Physique quantique et applications

ENS
SORBONNE UNIVERSITÉ
COLLÈGE DE FRANCE
cnrs

Team « Complex Media Optics »

Our Mission statement :
Understand and exploit the complexity of light propagation in complex media

PIs:
Prof. Sylvain GIGAN (Sorbonne U.)
Dr. Hilton BARBOSA DE AGUIAR (CNRS)

Postdocs
Dr. YoonSeok BAEK
Dr. Fei XIA
Dr. Gianni IACCUCI
Dr. Raj PANDYA
Dr. Clémence GENTNER
Dr. Tengfei WU
Dr. Caio RIMOLI
Dr. Jianqi HU

PhD Students
Louis DELLOYE
Léa CHIBANI
Adrian MAKOWSKI
Shupeng ZHAO
Francesco MAZZONCONI
Baptiste COURME
Malo JOLY
Hao WANG

Alumni

PhDs:
Sébastien Popoff (CNRS)
Daria Andreoli
Pierre Bondareff
Tthomas Chaigne (CNRS)
Hugo Defienne (CNRS)
S. Leedumrongwatthanakun
Jonathan Dong
Alexandra D’arco
Mickael Mounaix
Baptiste Blochet
Tom Sperber
Julien Guilbert
Antoine Boniface
Louisiane Devaud

Postdocs:
P. Boucher
D.Martina
R. Savo
G. Volpe (UCL)
J.Bertolotti (U.Exeter)
O.Katz (HUJI)
T. Juffman (U. Vienna)
L. Valzania
M. Rafayelan (U. Yerevan)
M. Dabrowski
C. Moretti
B. Rauer
F. Soldevila

Main Collaborations
L. Bourdieu (IBENS)
C. Conti (Sapienza, Roma)
M. Guillon (U. Paris)
L. Novotny (ETHz)
E. Charbon (EPFL)
F..Krzakala (EPFL)
R. Lapkiewicz (Warsaw)
S. Rotter (TU Wien)
H. Cao (Yale)
Peter McMahon (Cornell)

Fundings

The deep Learning revolution

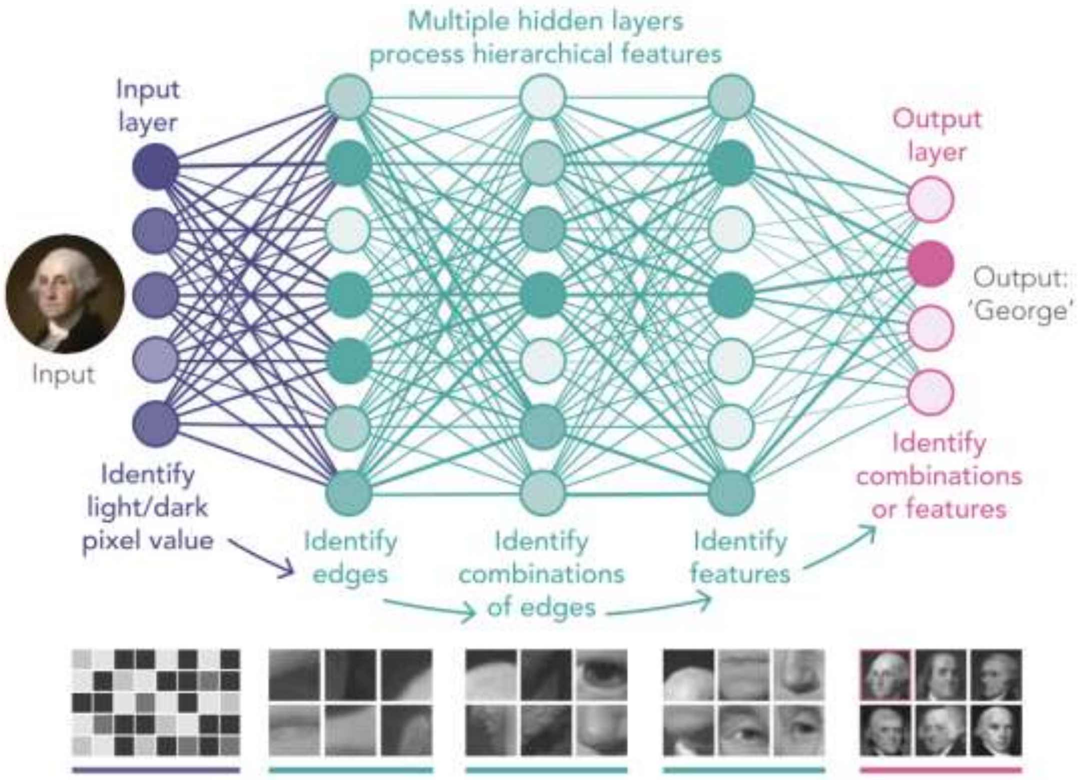


Image: Waldrop, PNAS (2019)

State of the art in image recognition, speech recognition, text generation, driving …

- Many (10s to 100s) Layers
- Each layer = a matrix multiplication
- 10s BILLIONS weights / parameters
- Huge (size+dimensions) datasets
- Training (Backpropagation) and inference are extremely demanding

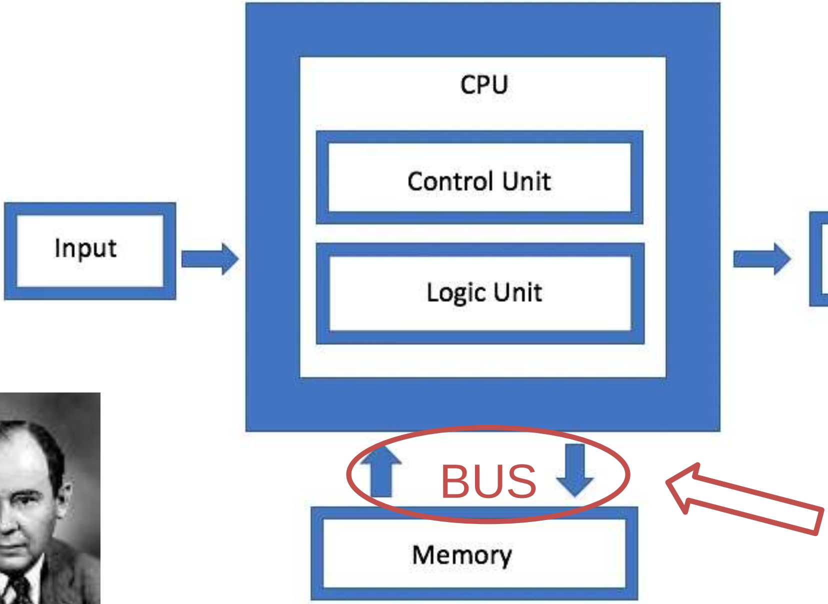
See: Deep learning, LeCun, Bengio & Hinton, Nature 521, 436 (2015)

AI and Compute

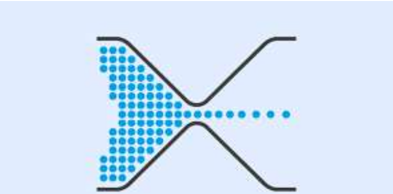
x300.000
(Moore Law: x7)

https://openai.com/blog/ai-and-compute/

The Von Neumann architecture

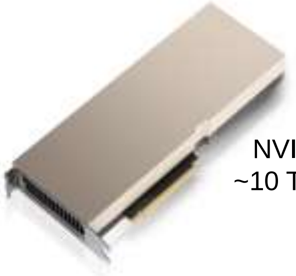


John VON NEUMANN
(1903-1957)



Memory access:
The Von Neumann
bottleneck

CPU
Desktop PC
~100 GigaFLOPS
300W



GPU
NVIDIA A100
~10 TeraFLOPS
300W

TPU (Google)

~4 T-Ops
75W

Versatile …. *But not adapted for AI :* ***slow, power inefficient***

« …how hardware chooses which ideas succeed and which fail. »
(and vice-versa)

advocates for « (…) joint collaboration between hardware, software and machine learning communities. »

https://hardwarelottery.github.io
arXiv:2009.06489

Machine Learning typically require low precision

Analog computing is interesting

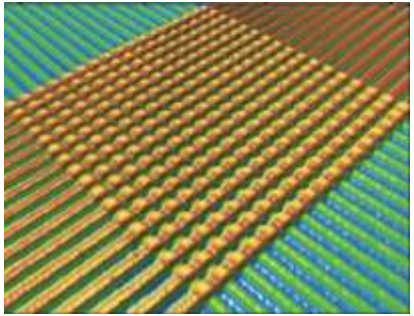
Emulating neural network architecture directly in a physical system?

« brain inspired »

Neuromorphic computing

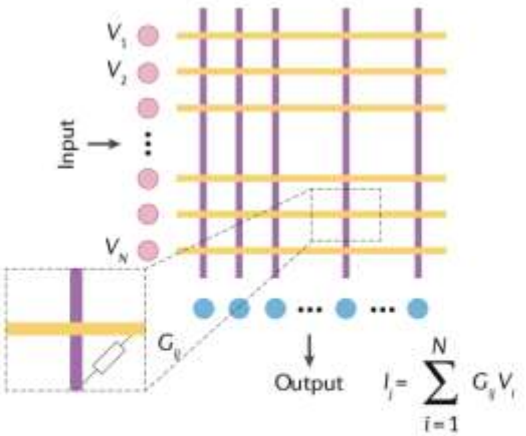
- « In-memory » computing
- Speed
- Energy efficiency

« Non-Von-Neumann »



e.g. Memristive & CMOS crossbar array

Source: Nature Materials 18, 309(2019)



Comes in many flavor :

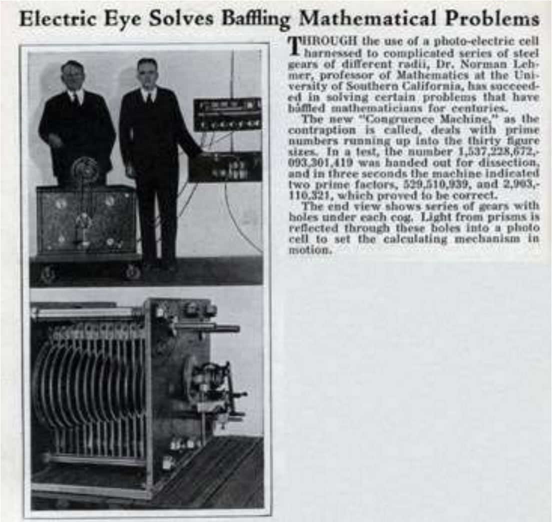
- CMOS
- Memristors
- Phase-change materials
- Spintronic
- Superconductive synapses

Physics for neuromorphic computing

Danijela Marković, Alice Mizrahi, Damien Querlioz and Julie Grollier

Nature Reviews Physics volume 2, pages 499–510 (2020)

From Sieves ... to Fourier Transforms … all the way to Neural Networks

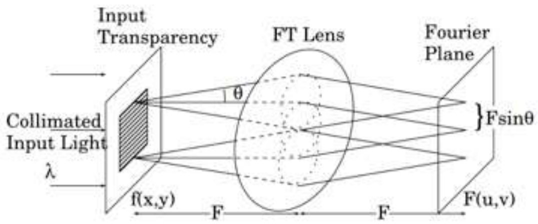


Electric Eye Solves Baffling Mathematical Problems

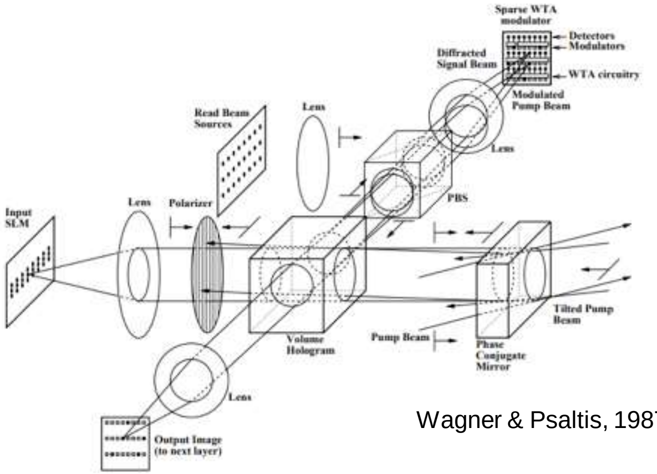
THROUGH the use of a photo-electric cell harnessed to complicated series of steel gears of different radii, Dr. Norman Lehmer, professor of Mathematics at the University of Southern California, has succeeded in solving certain problems that have baffled mathematicians for centuries.

The new "Congruence Machine," as the contraption is called, deals with prime numbers running up into the thirty figure sizes. In a test, the number 1,537,228,672,093,301,419 was handed out for dissection, and in three seconds the machine indicated two prime factors, 529,510,939, and 2,903,110,321, which proved to be correct.

The end view shows series of gears with holes under each cog. Light from prisms is reflected through these holes into a photo cell to set the calculating mechanism in motion.



$$F(u,v) = \int\int f(x,y) e^{i\frac{2\pi}{\lambda F}(xx'+yy')} dxdy$$



Wagner & Psaltis, 1987

1930's → 1960's → 1980's

2000’ :

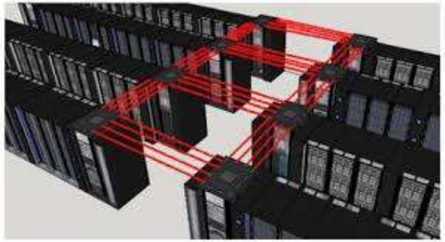
A winter for optical computing

…and a new era in optical communication

optical interconnects

Long distance fiber links

In data centers



(virtually) any communication beyond a few meters is optical:

- Low power consumption
- high bandwidth

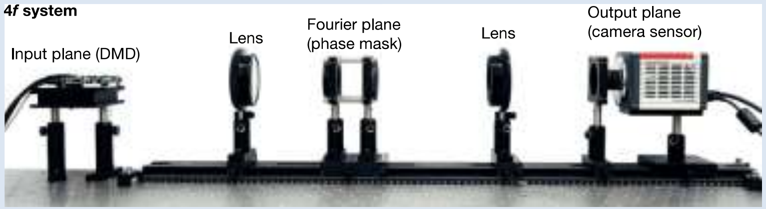
2015 – now : Neuromorphic computing with light – redux

Optics has distinct advantages…

- Low energy consumption
- Easy interconnect – Multiplexing
- Low latency and blazzing speed

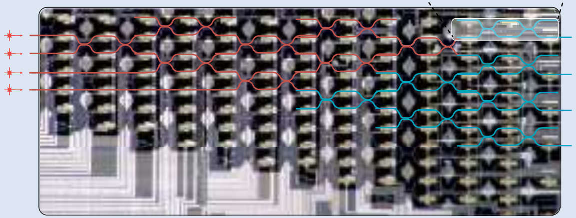
… but also some disadvantages:

- rather bulky
- Tricky non-linearities / storage



Sci. Rep. **8**, 12324 (2018).

Free space
vs
Integrated optics



Sci. Adv. **5**, eaay6946 (2019).

Nature Photonics 15, 10 (2021)

Photonics for artificial intelligence and neuromorphic computing

REVIEW ARTICLE | FOCUS
https://doi.org/10.1038/s41566-020-00754-y

Bhavin J. Shastri[1,2,7], Alexander N. Tait[2,3,7], T. Ferreira de Lima[2], Wolfram H. P. Pernice[4], Harish Bhaskaran[5], C. D. Wright[6] and Paul R. Prucnal[2]

Nature 588, 39 (2020)

Perspective

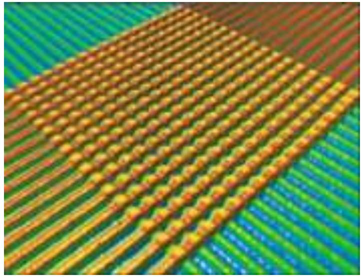
Inference in artificial intelligence with deep optics and photonics

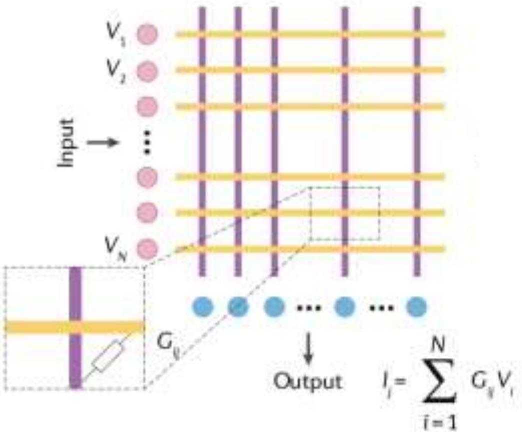
https://doi.org/10.1038/s41586-020-2973-6
Received: 28 November 2019

Gordon Wetzstein[1], Aydogan Ozcan[2], Sylvain Gigan[3], Shanhui Fan[1], Dirk Englund[4], Marin Soljačić[4], Cornelia Denz[5], David A. B. Miller[1] & Demetri Psaltis[6]

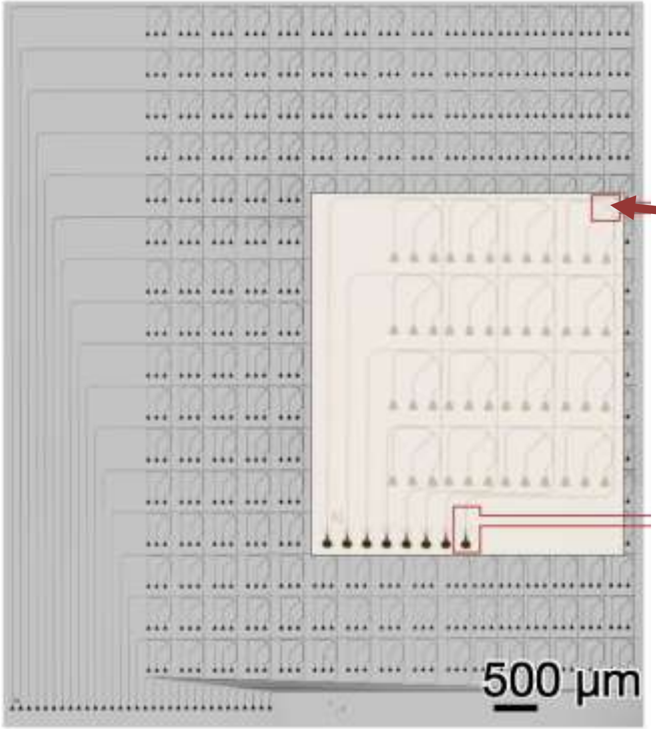
Several startups : Optalysys, LIGHTELLIGENCE, LIGHTMATTER, LightOn

The cross-bar array – Optical implementation





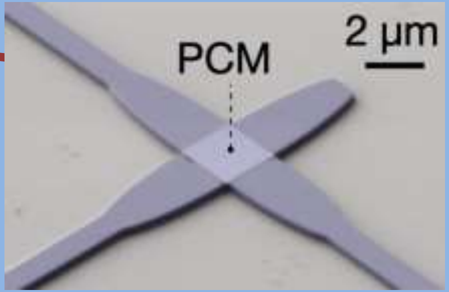
CMOS crossbar array
Source: Nature Materials 18, 309(2019)

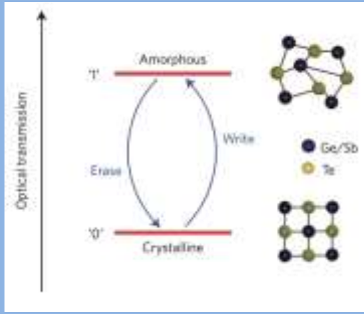


Feldmann, J., et al. "Parallel convolutional processing using an integrated photonic tensor core. »
Nature 589.7840 (2021): 52-58.

Phase-Change Materials (PCM)

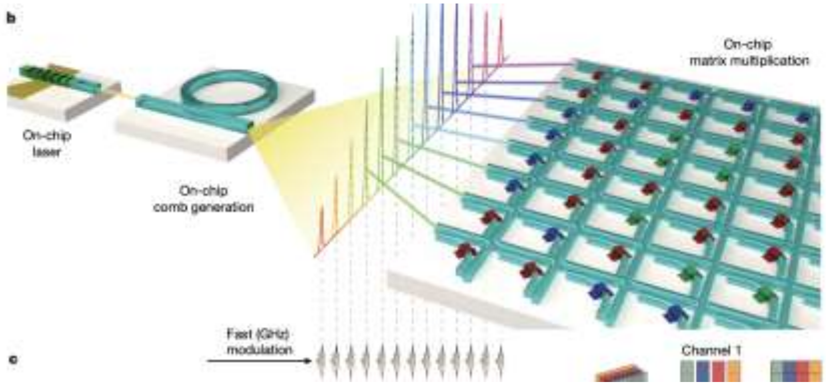
- Optically or electronically adressable
- Multilevel (6-7 bits)
- Stable – no power consumption
- Lossy



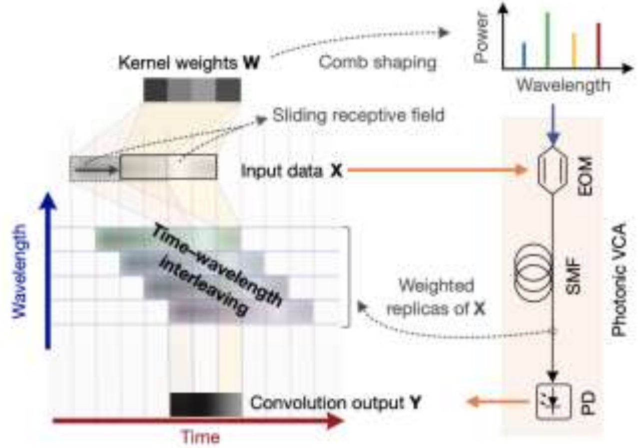


Ríos, C.,. et al.Integrated all-photonic non-volatile multi-level memory. Nature Photon 9, 725–732 (2015)

The frequency multiplexing advantage



Feldmann, J., et al. "Parallel convolutional processing using an integrated photonic tensor core." *Nature* 589.7840 (2021): 52-58.



Xu, X., Tan, M., Corcoran, B. *et al.* 11 TOPS photonic convolutional accelerator for optical neural networks. *Nature* **589,** 44–51 (2021)

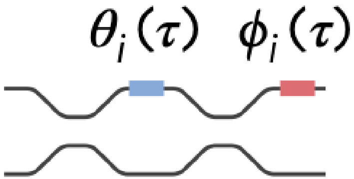
Frequency comb = N coherent lasers

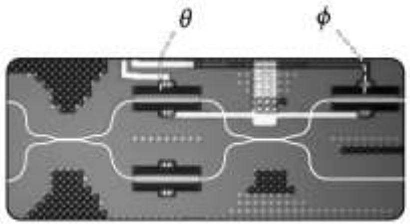
- Same matrix multiplied N times in parallel
- Different inputs can be encoded on each comb line
- Need mux-demux at the input and output

Integrated mach-Zehnder interferometers

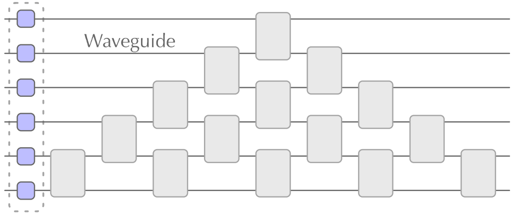
Arbitrary Lossless 2x2 coupler

Slow (thermal) or fast (EOM) tuning

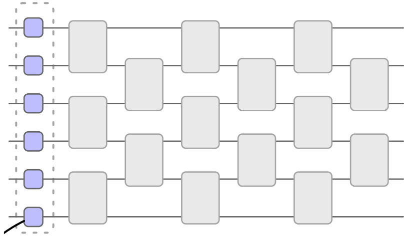


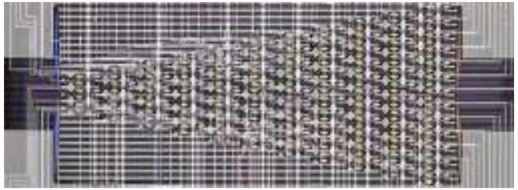


Nature Photon **11,** 447–452 (2017)



Unitary circuits

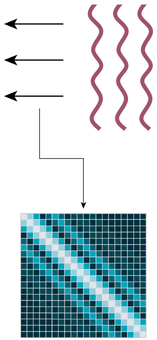




Linear programmable nanophotonic processors (review) Optica 5, 1623-1631 (2018)

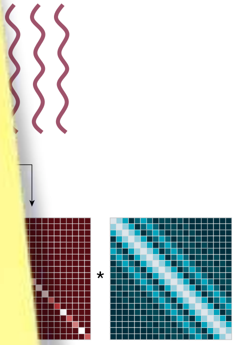
Basic building blocks:

Free space

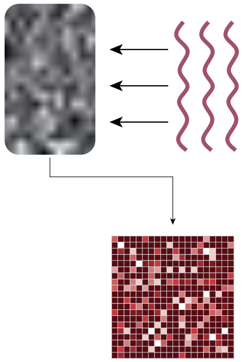


Thin mask

layered scatterer



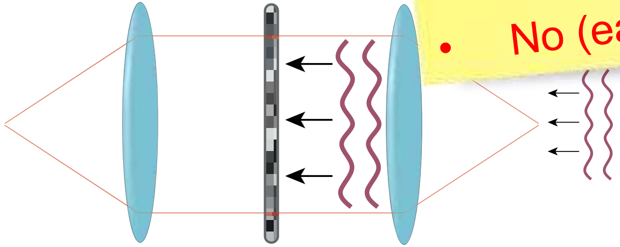
Volumetric scatterer



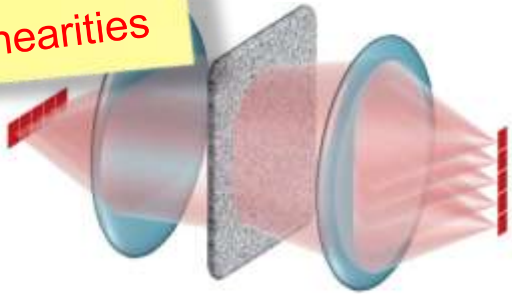
Free space optics

- Large dimension
- Energy efficient
- Slow (kHz) but high throughput
- Bulky
- No (easy) non-linearities

Advanced functions



2D convolutions



1D vector-matrix multiplier

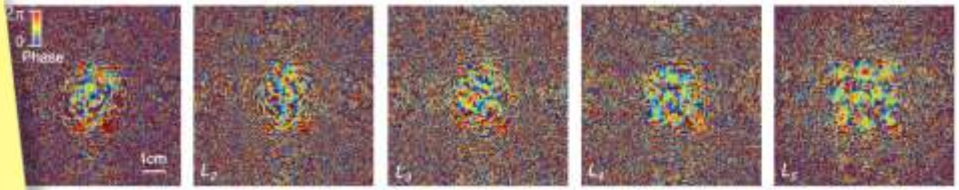
Adapted from *Nature* 588, 39 (2020)

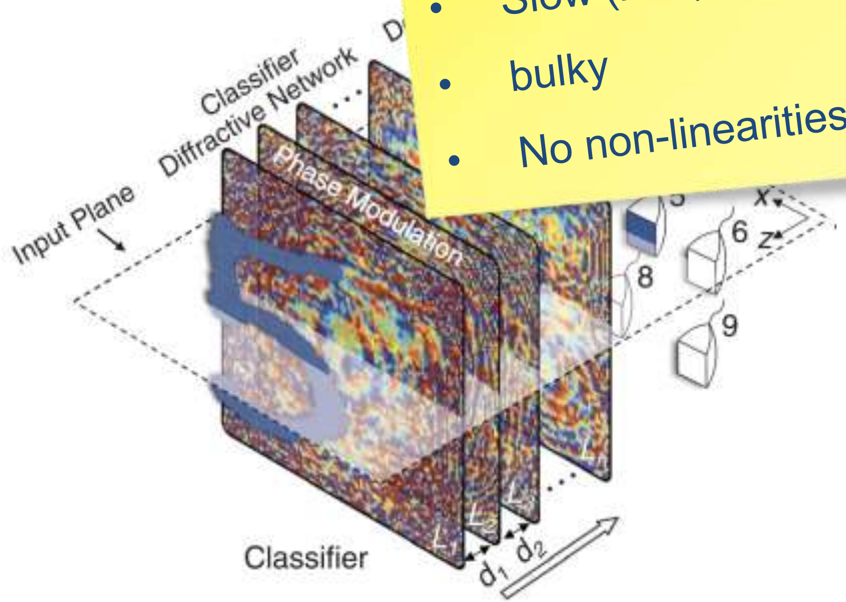
Beyond convolutions : « Deep diffractive » networks

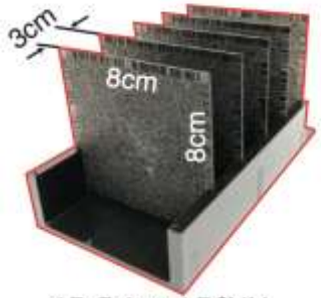
IDEA : fabricat

Free space optics

- Large size
- Slow (kHz)
- bulky
- No non-linearities







3D Printed D²NN (Classifier)

THz realization

Multilayer : « **deep** »
no non-linearity between layers:
It's a just a dense single NN layer!
« **diffractive** »

Lin, Xing, et al. "All-optical machine learning using diffractive deep neural networks." Science 361.6406 (2018): 1004-1008

The energy advantage

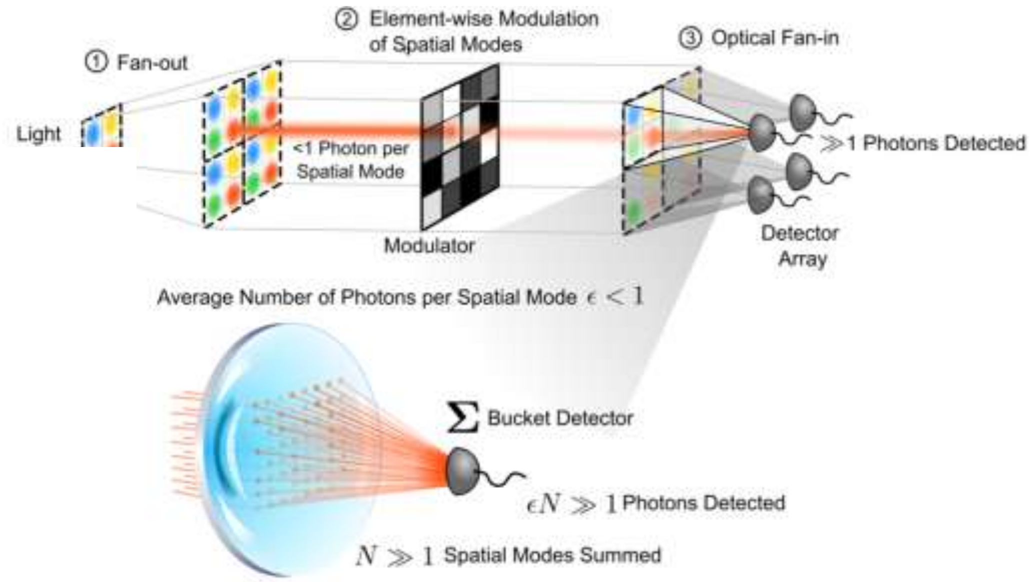
Article | Open Access | Published: 10 January 2022

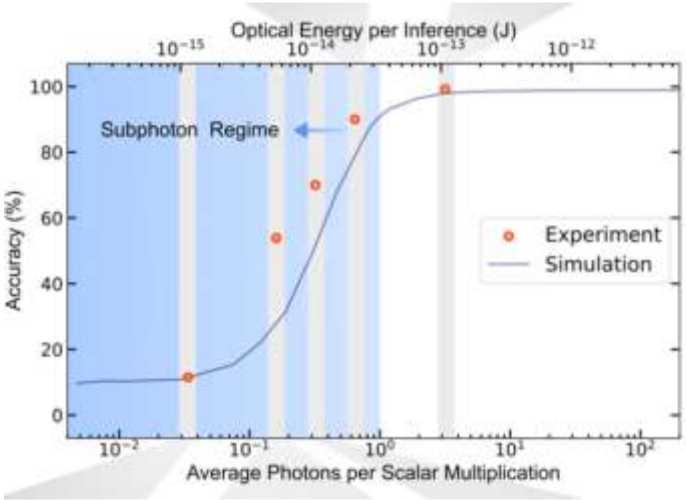
An optical neural network using less than 1 photon per multiplication

Tianyu Wang, Shi-Yuan Ma, Logan G. Wright, Tatsuhiro Onodera, Brian C. Richard & Peter L. McMahon

Nature Communications **13**, Article number: 123 (2022) | Cite this article

Accuracy on MNIST:
90% with <1 (detected) photon/MAC



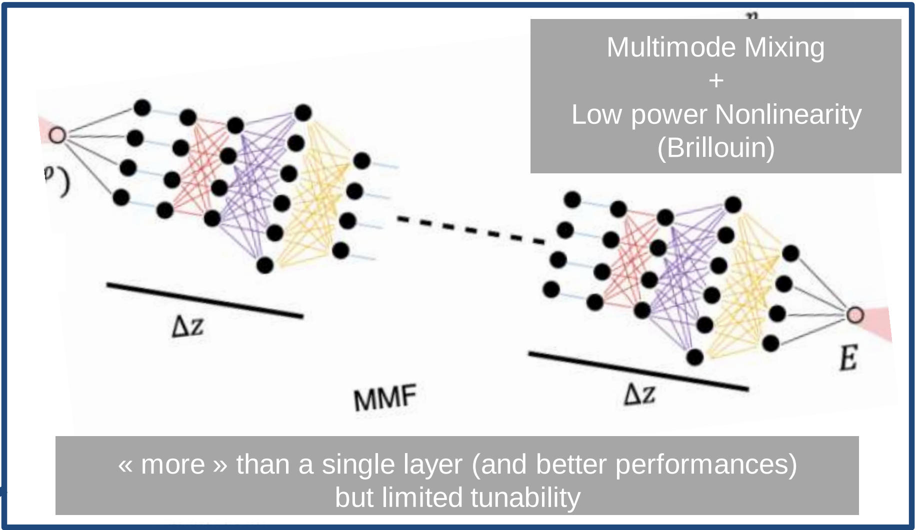


What about optical non-linearities ?

No practical implementation of optical non-linearities yet (Non-linear effects, photorefractive,…)

solutions to date:

- Intensity detection $|.|^2$
- Electronic non-linearity
- Not compatible with multi-layers

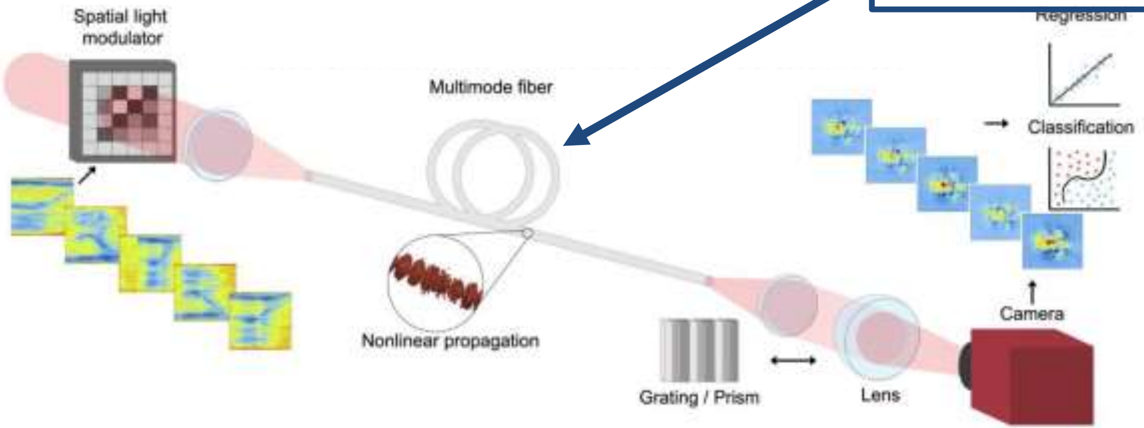


Article | Published: 20 August 2021

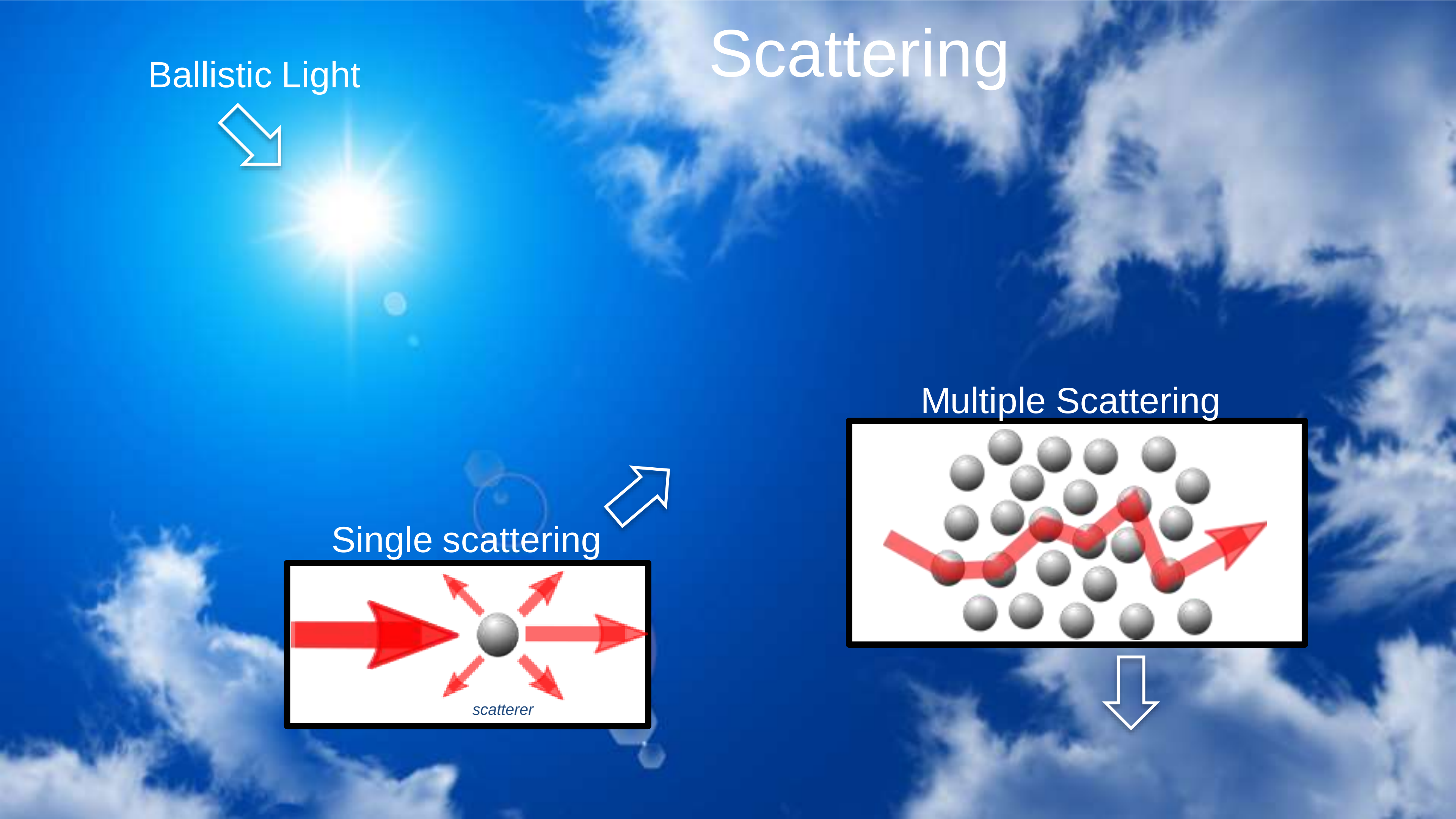
Scalable optical learning operator

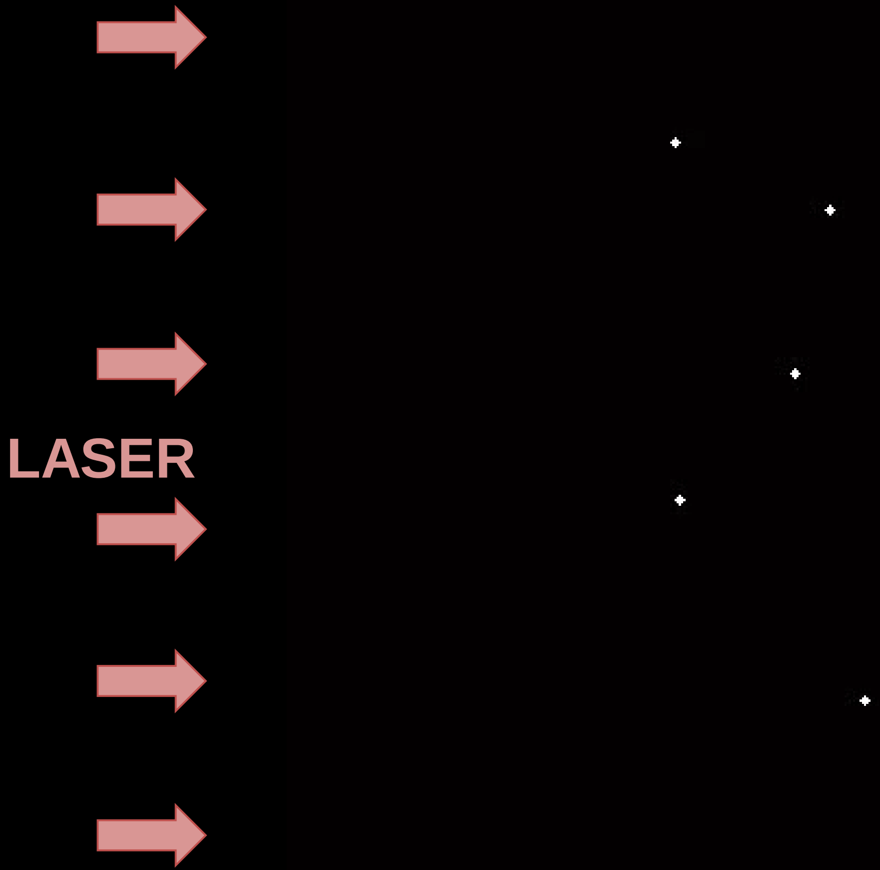
Uğur Teğin, Mustafa Yıldırım, İlker Oğuz, Christophe Moser & Demetri Psaltis

Nature Computational Science 1, 542–549 (2021) | Cite this article



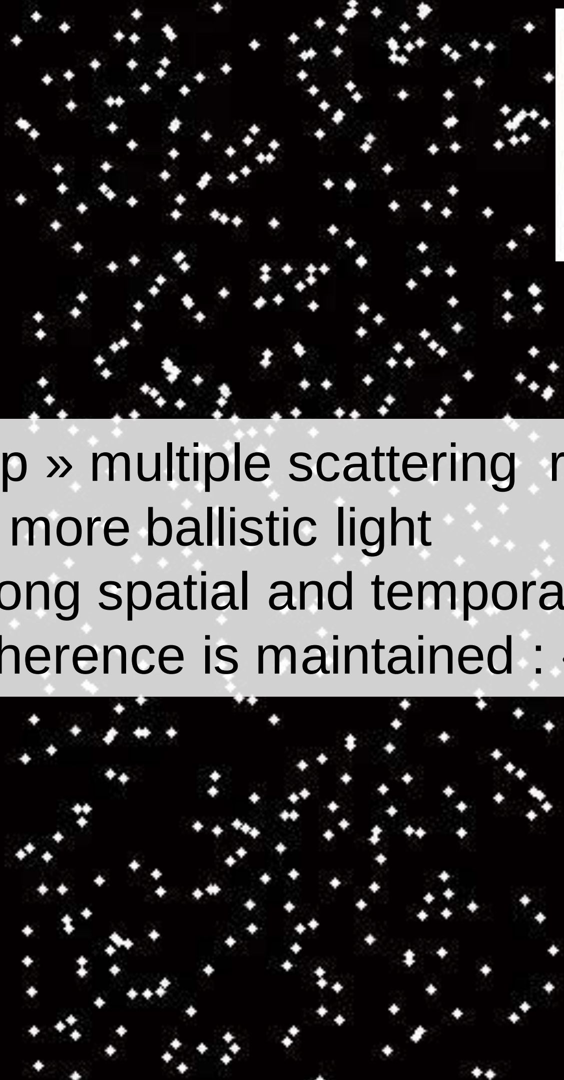
Nature Computational Science **volume 1**, pages 542–549 (2021)

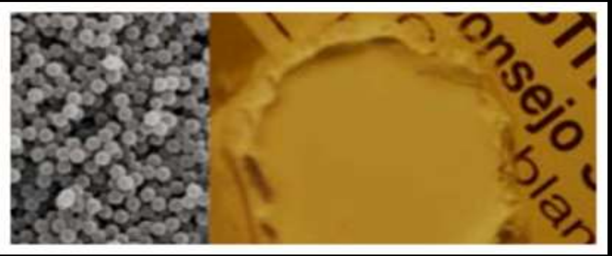

Scattering
Ballistic Light
Single scattering
scatterer
Multiple Scattering



Film courtesy of Emmanuel Bossy (Univ. Grenoble) - SIMSONIC software

LASER





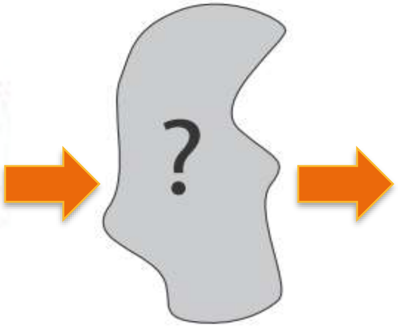
3D random Sample
« white paint »

« Deep » multiple scattering regime :

- ✗ No more ballistic light
- ✗ Strong spatial and temporal mixing
- ✓ Coherence is maintained : **« speckle »**

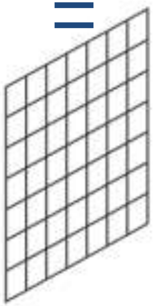
Film courtesy of Emmanuel Bossy (Univ. Grenoble) - SIMSONIC software

A general approach : the transmission matrix

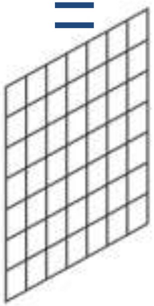


SLM : array of pixels — Linear system — CCD camera : array of pixels

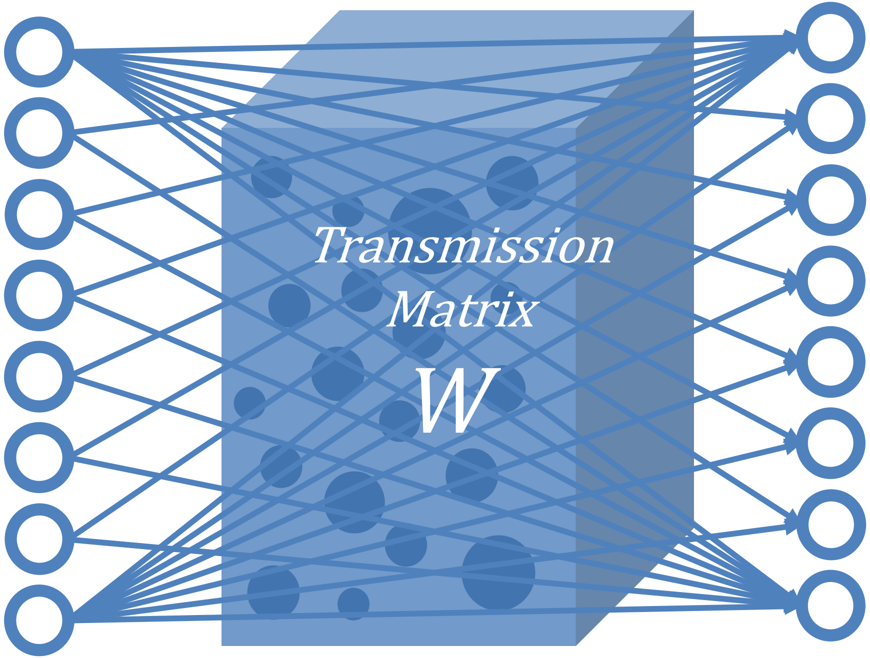
= — = — =

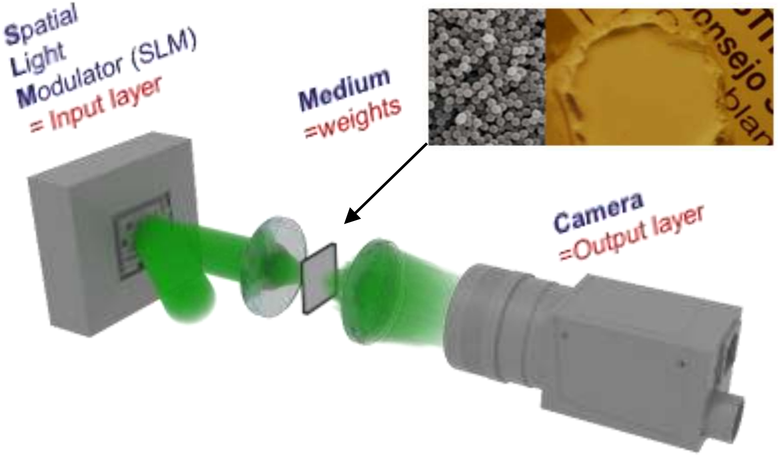


$$H = \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,N} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,N} \\ \vdots & & \ddots & \vdots \\ h_{M,1} & h_{M,2} & \cdots & h_{M,N} \end{pmatrix}$$



measurement of the TM : see Popoff et al. Phys. Rev. Lett. 104,100601 (2010)



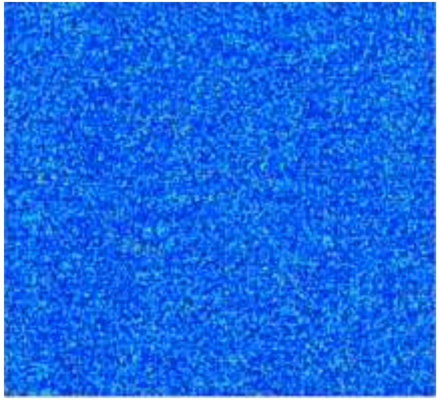


= 1-Layer neural network
(fixed weights)

$$E^{out} = W E^{in}$$

Experimentally measured Transmission Matrix (TM)

Out

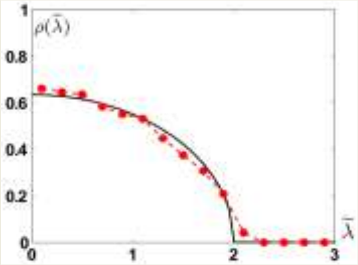


In

Popoff et al. Phys. Rev. Lett. 104,100601 (2010)

Random

Mesoscopic physics



« Quarter-circle law »

Stable for weeks

Large-dimensional

Area $A \sim 1\,mm^2$

Wavelength $\lambda \sim 1\,\mu m$

$$N \sim A/\lambda^2$$

~ many million in/out modes

as in *Yu, Lee, & Park (2017)*

Propagation of light through a disordered medium

=

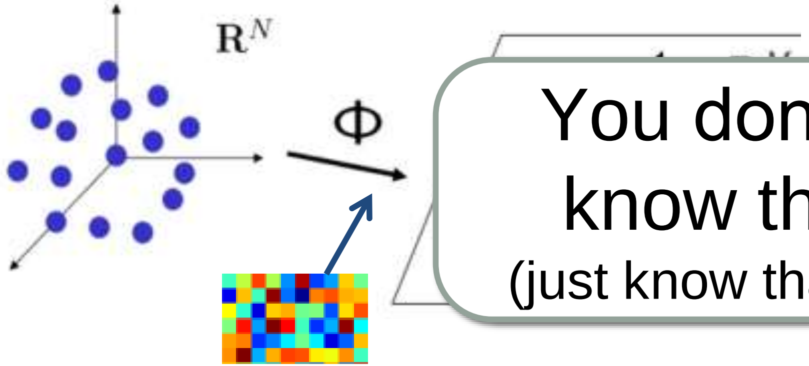
multiplication by a complex <u>i.i.d. random</u> matrix

a.k.a. in signal processing : **« random projections »**

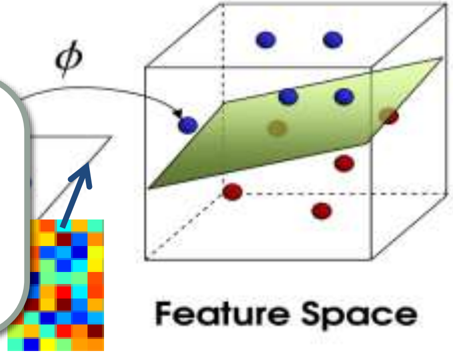
A **universal** operation

What are « random projections » good for?

Dimensional Reduction
« Johnson & Lindenstrauss »



Dimensional expansion:
The "kernel Trick" « Rahimi & Recht »



You don't need to know the matrix!
(just know that it is random)

- **Random matrix :** rotation and reduction in dimension
- **conserve distances** even for M<<N

Reference : Johnson, W. B., & Lindenstrauss, J. Extensions of Lipschitz mappings into a Hilbert space. *Contemporary mathematics*, 26,189(1984)

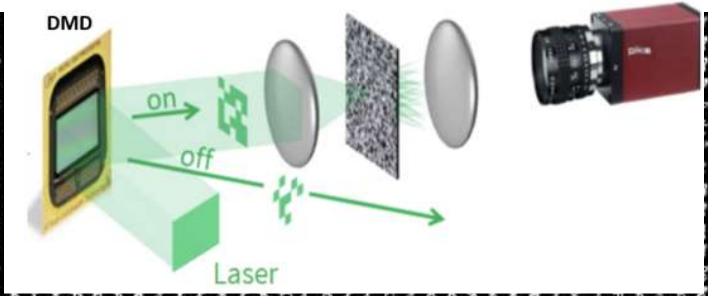
[Cited 3700 times]

- Make a **non-linear** regression problem **linear**
- Random projections are **efficient** and **universal**

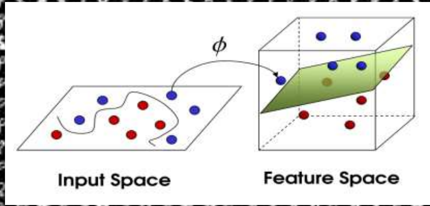
Reference : Rahimi, A., & Recht, B. (2007). Random features for large-scale kernel machines. In *Advances in neural information processing systems* (pp. 1177-1184).

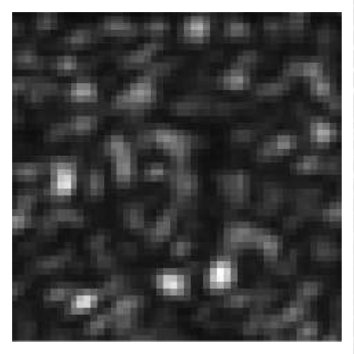
[Cited 3800 times]

Experimental Proof of Principle



MNIST database
(US postal letters)

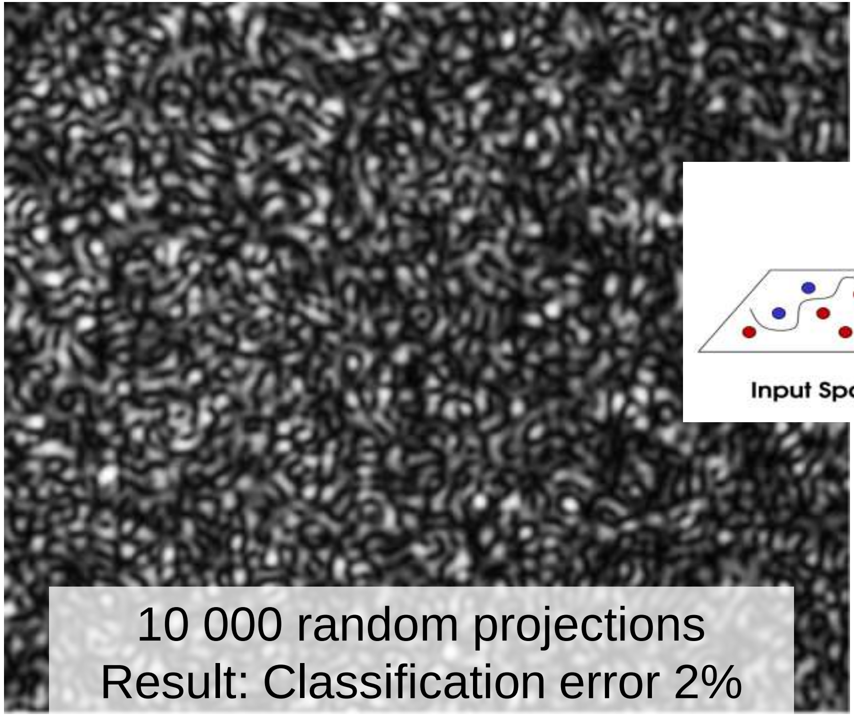


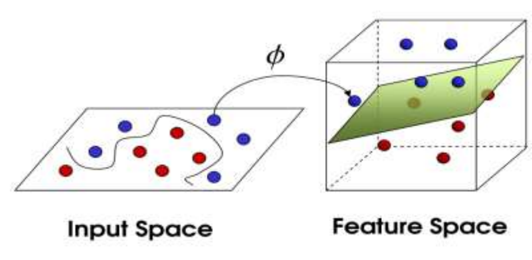


Random Optical projections

A. Saade, F. Caltagirone, I. Carron, L. Daudet, A. Drémeau, S. Gigan, F. Krzakala,
Random projections through multiple optical scattering: approximating kernels at the speed of light, ICCASP (2016)

Classification results





10 000 random projections
Result: Classification error 2%
(ridge regression)

optical computing revisited?

Why is it interesting ?

EXTRA-LARGE **&** **SUPER-FAST**

W of size higher than 10^6 x 10^6 (TBs of memory)

kHz operation →10^3 such multiplies / s

Equivalent 10^{15} operations / s : You would need a *Peta-scale* computer to do the same !

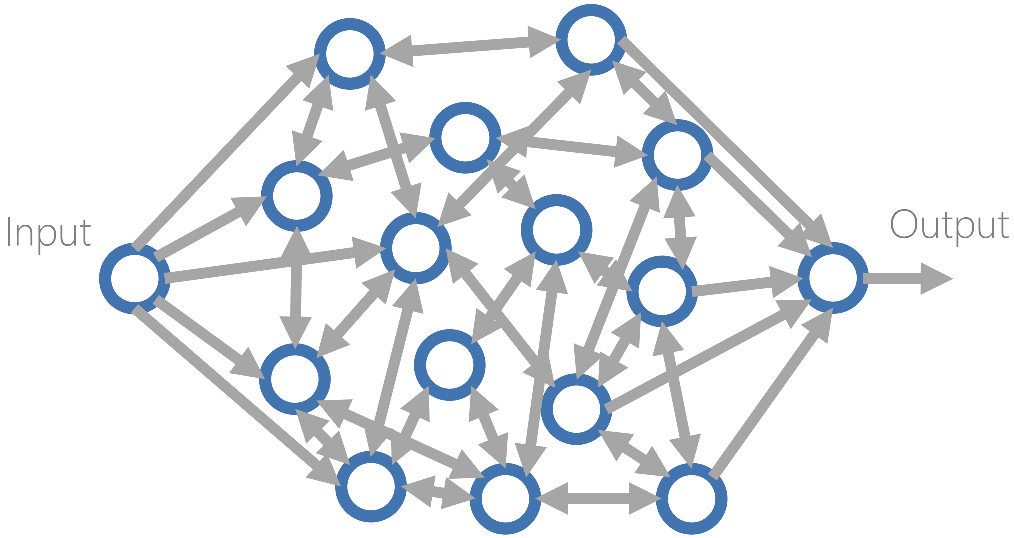
LightOn

We bring Light to AI

(CoI disclosure: S.G. acknowledges financial interest in LightOn)

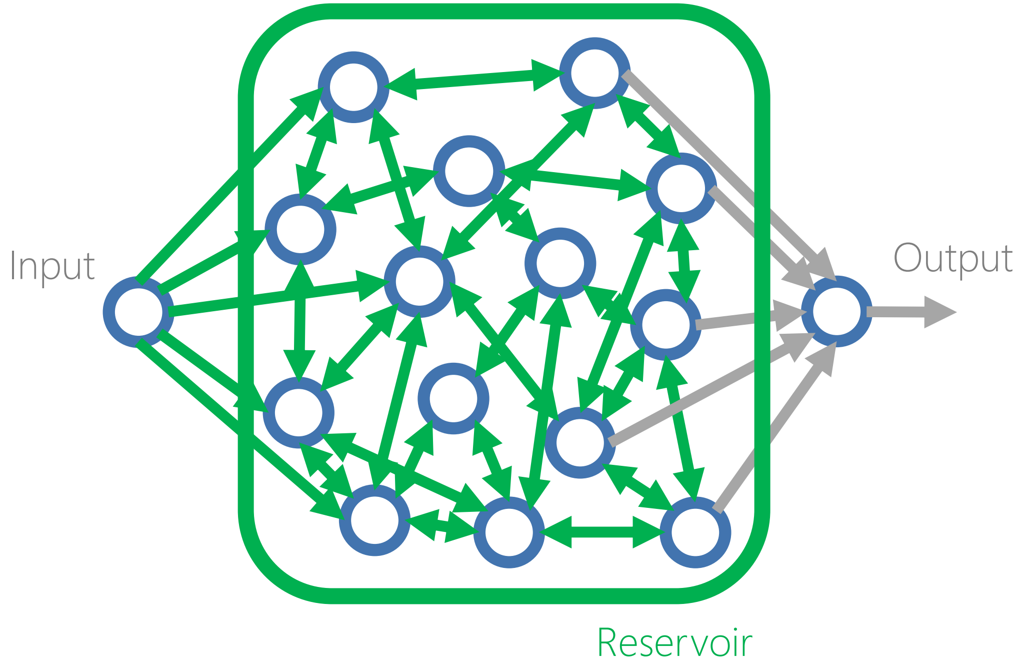
- **many, many use cases (inference, training, linear algebra…)**
- **already at scale for modern machine learning**
- **you can buy it already (1st commercial optical processor)**

www.lighton.ai



Recurrent Neural Networks are notoriously hard to train

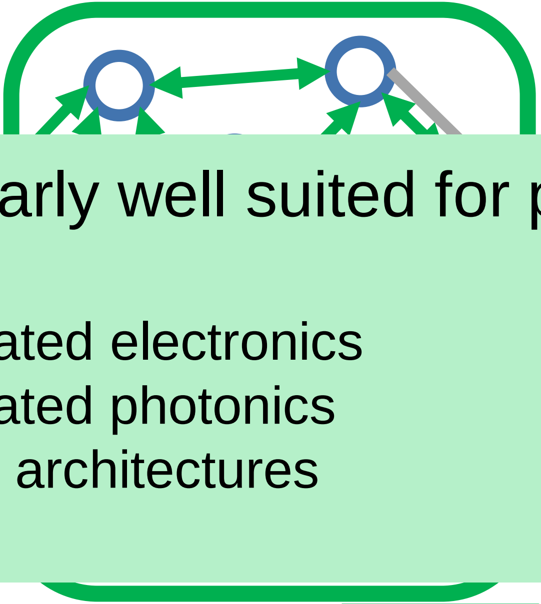
Jaeger & Haas (2004). *Science*



Recurrent Neural Networks are notoriously hard to train

Reservoir Computing fixes all internal weights randomly

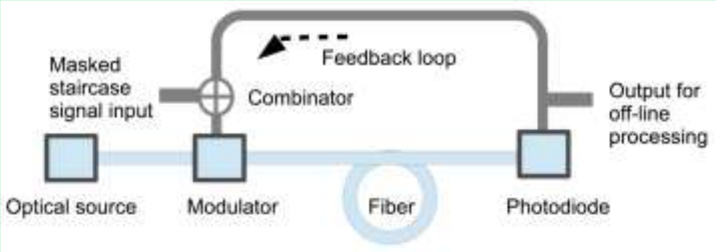
Jaeger & Haas (2004). *Science*



Recurrent Neural Networks are

Particularly well suited for physical implementations

- Dedicated electronics
- Integrated photonics
- Exotic architectures



Tanaka, Gouhei, et al. "Recent advances in physical reservoir computing: A review." *Neural Networks* 115 (2019): 100-123

Reservoir

next reservoir

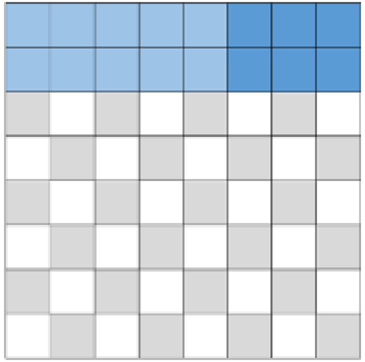
current reservoir

current input

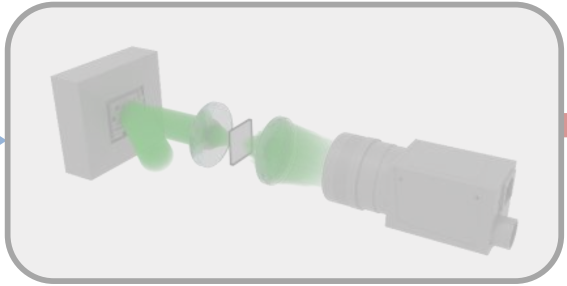
Jaeger & Haas (2004). *Science*

Reservoir computing with a complex medium ?

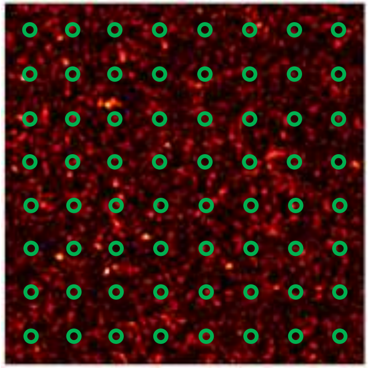
Input



Reservoir state



$$x^{(t+1)} = f\left(W_r\, x^{(t)} + W_i\, i^{(t)}\right)$$



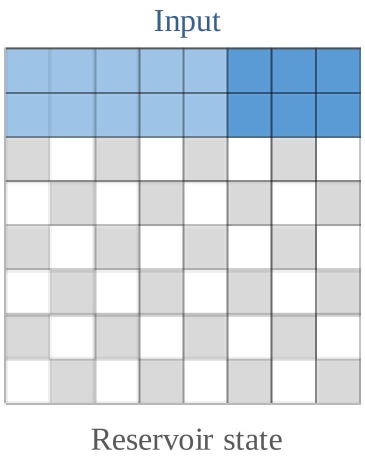
Output speckle

Dong, Gigan, Krzakala, Wainrib (2018). *IEEE SSP Workshop*

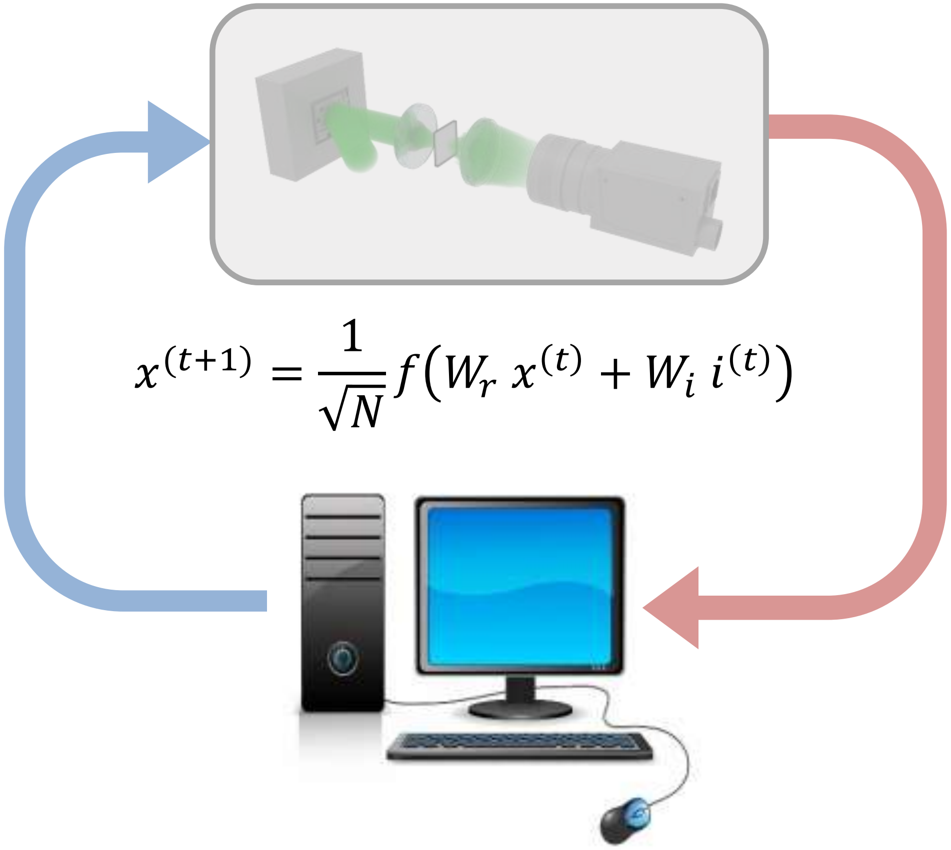
SLM encoding

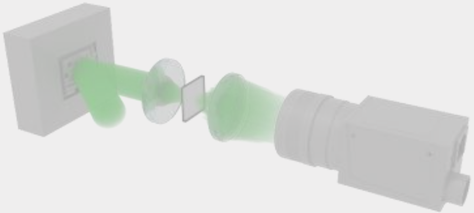
Input $i^{(t)}$

and reservoir $x^{(t)}$

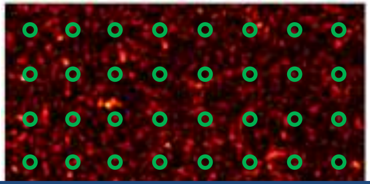


$$x^{(t+1)} = \frac{1}{\sqrt{N}} f\left(W_r\, x^{(t)} + W_i\, i^{(t)}\right)$$





$x^{(t+1)}$ 1 $f(W$ $^{(t)}$ $+$ W $i^{(t)})$



Camera readout

To get $x^{(t+1)}$

us

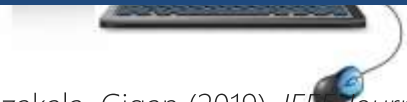
rity

Last stage / a posteriori

Predict output with a linear model

$$o^{(t)} = W_o x^{(t)}$$

(done on a CPU or GPU - Typically not the bottleneck)



Dong, Rafayelyan, Krzakala, Gigan (2019). *IEEE Journal of Selected Topics in Quantum Electronics*

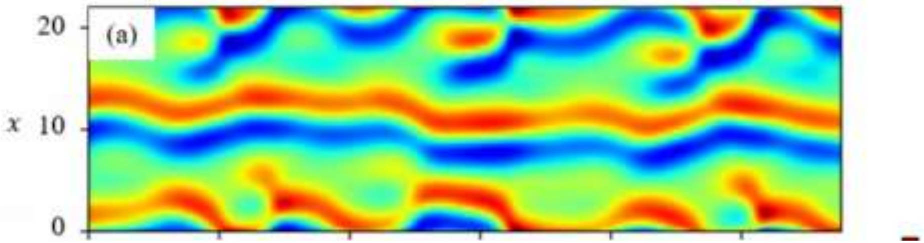
The Mackey-Glass equation (1D):

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

The Kuramoto-Sivashinsky equation (2D):

$$\frac{\partial u}{\partial t} + \nabla^4 u + \nabla^2 u + \frac{1}{2}|\nabla u|^2 = 0$$

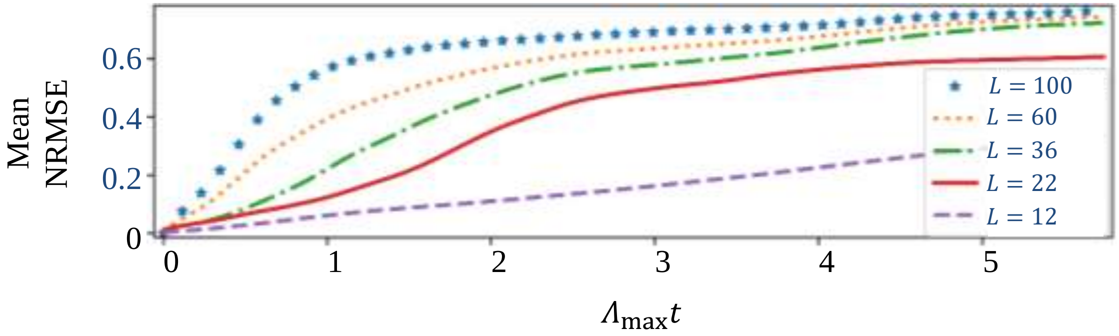
Ground truth



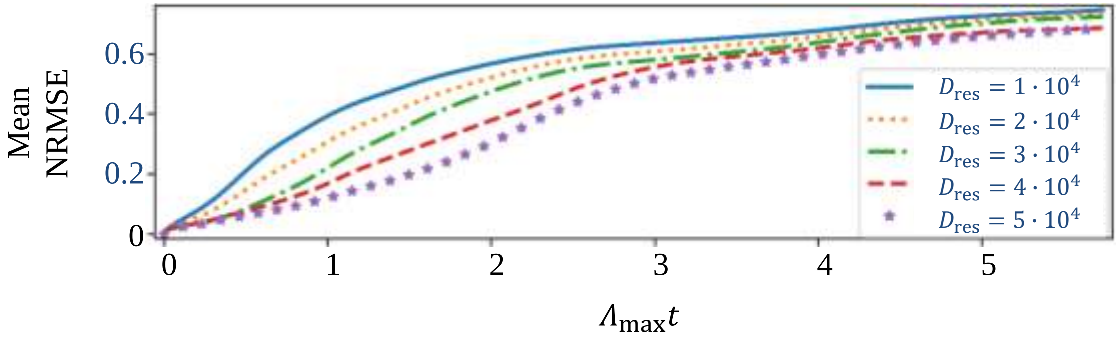
Rafayelyan, Dong, Tan, Krzakala, Gigan (2020). *Physical Review X*

Scaling behavior

Reservoir size is fixed, $D_{\text{res}} = 10000$



Spatial domain size is fixed, $L = 60$



Larger networks can predict better larger chaotic systems

Rafayelyan, Dong, Tan, Krzakala, Gigan (2020). *Physical Review X*

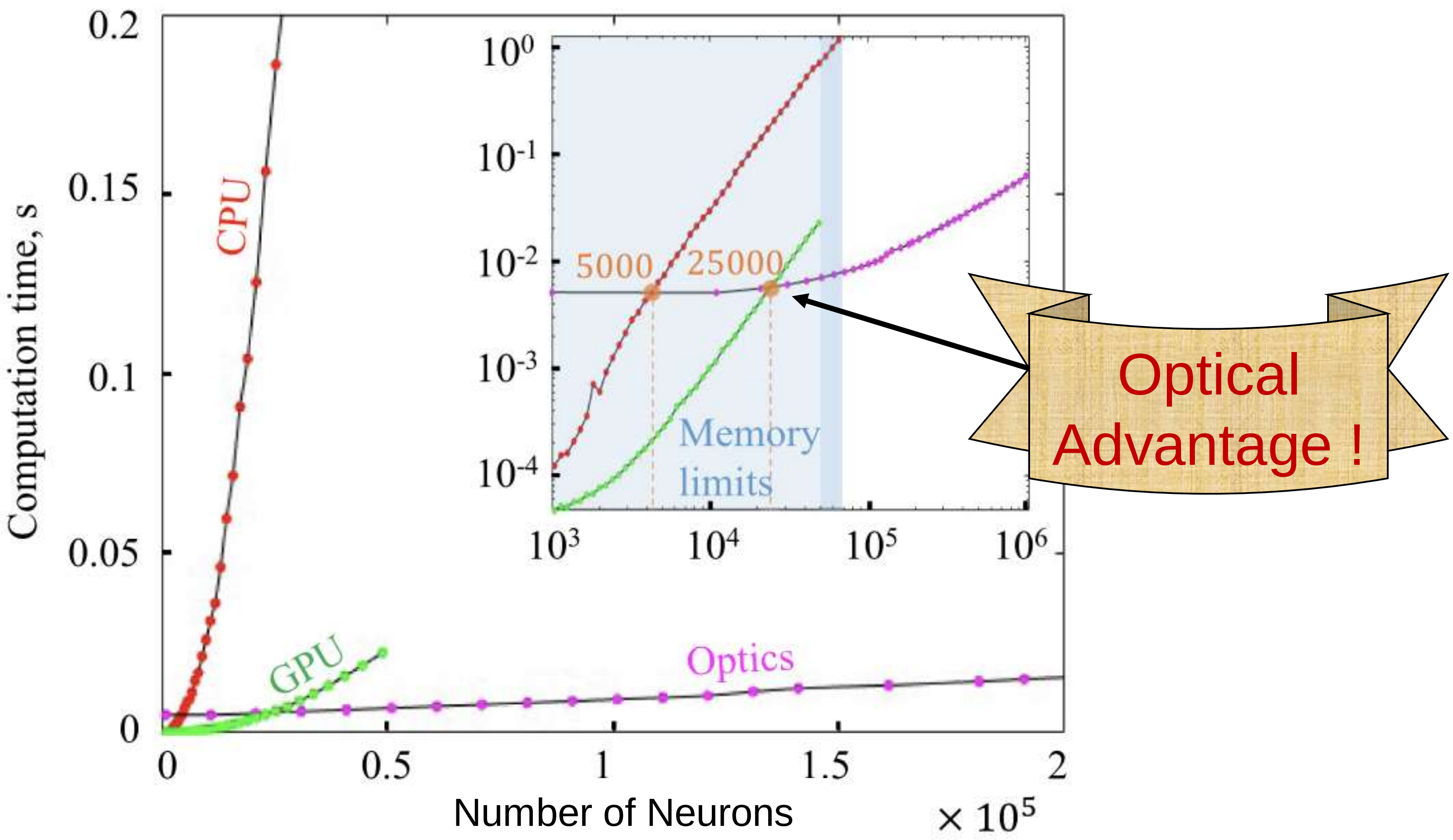
Speed

	Electronics	Optics
Speed	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$
Energy efficiency	~150 W	~30 W
Dimensionality	Memory limit (~ GB)	Resolution limit (~ TB)

Energy efficiency

Dimensionality

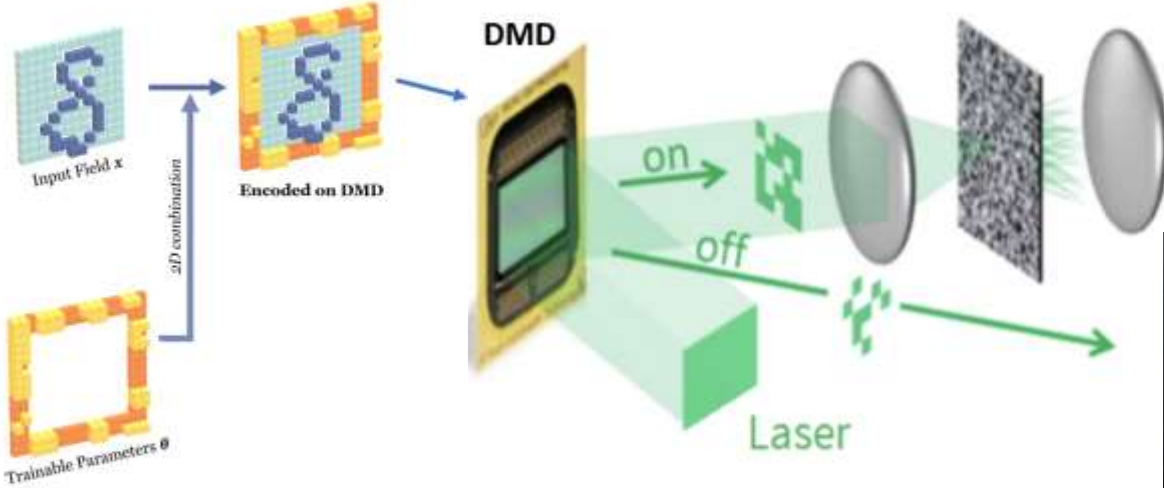
Computation time versus reservoir size (one iteration)



Rafayelyan, Dong, Tan, Krzakala, Gigan (2020). *Physical Review X*

Can we add a bit of tunability ?

- Tunable in weight by encoding on DMD
- Nonlinearity in camera detection



Laboratoire Kastler Brossel
Physique quantique et applications

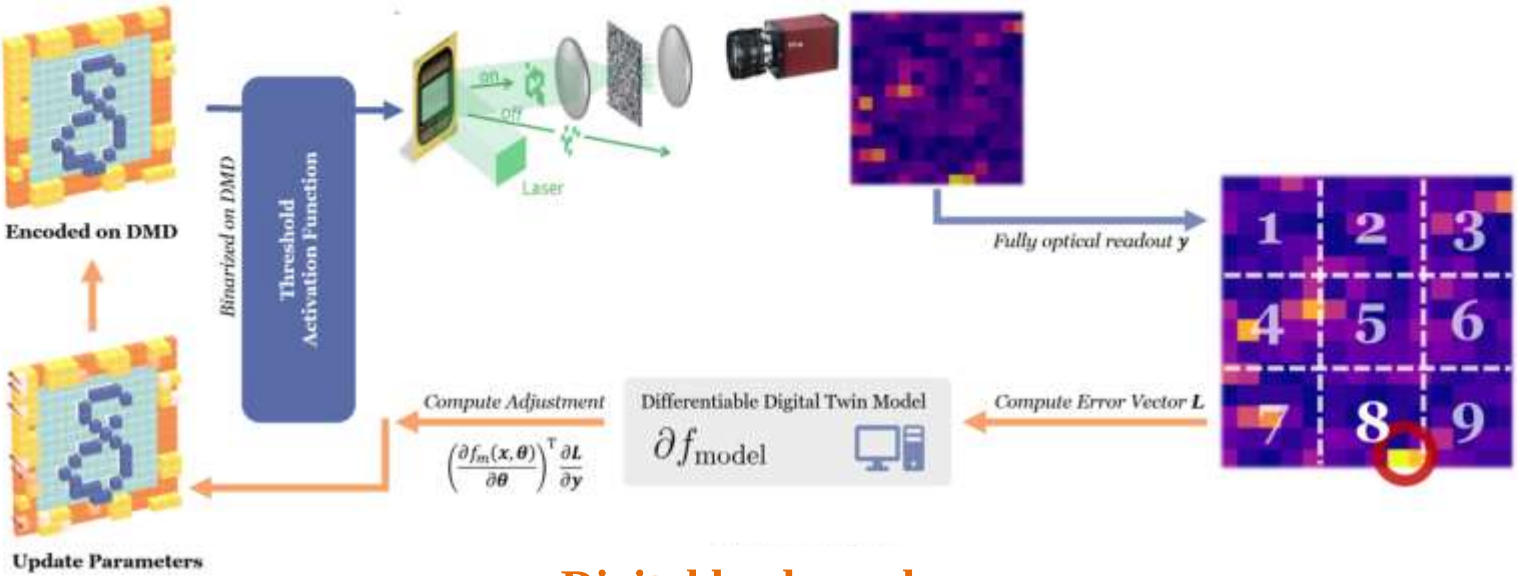
Fei Xia, Wang Ziao, Jianqi Hu

Cornell University

Collaboration with the Peter McMahon group
(Logan Wright, Tatsuhiro Onodera, Martin Stein)

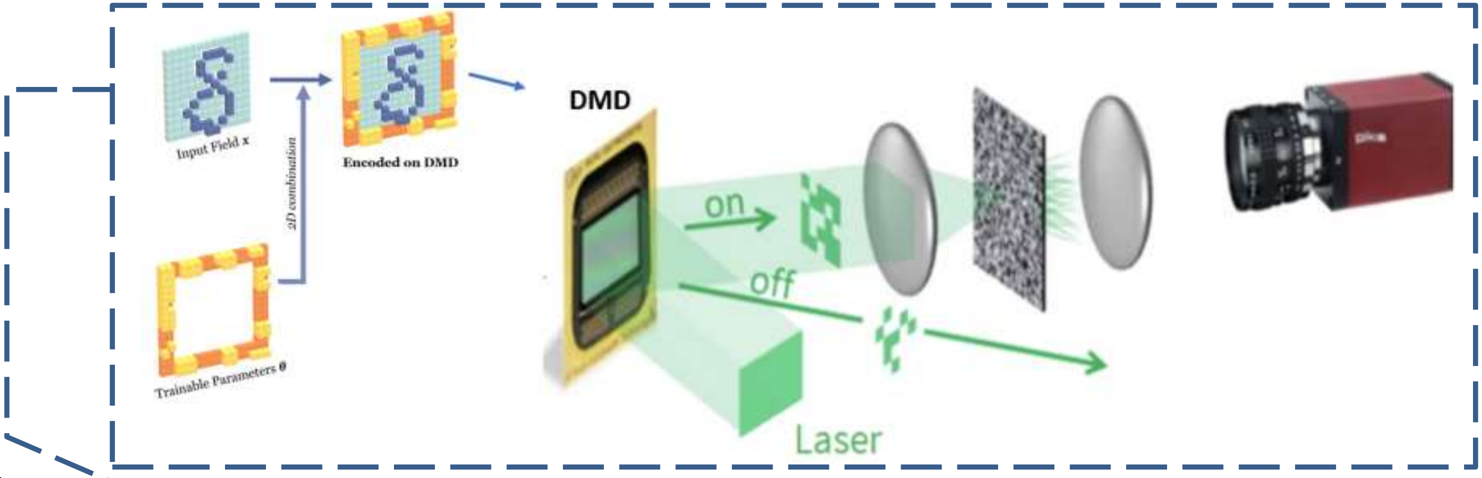
Training the optical neural network

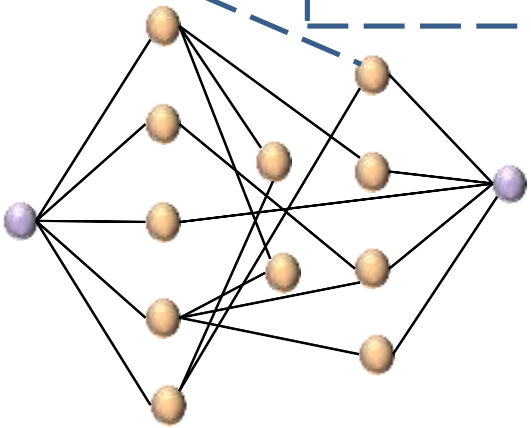
Optical forward pass



Digital backward pass

Based. On : Wright, L. G., Onodera, et al. (2022). Nature, 601(7894), 549-555.

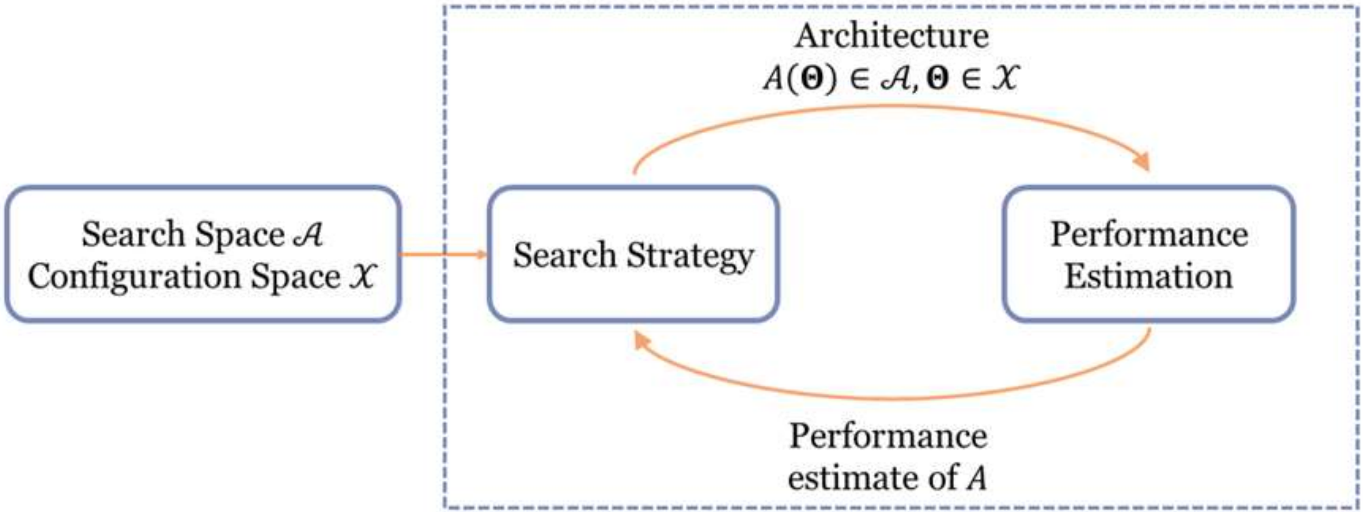


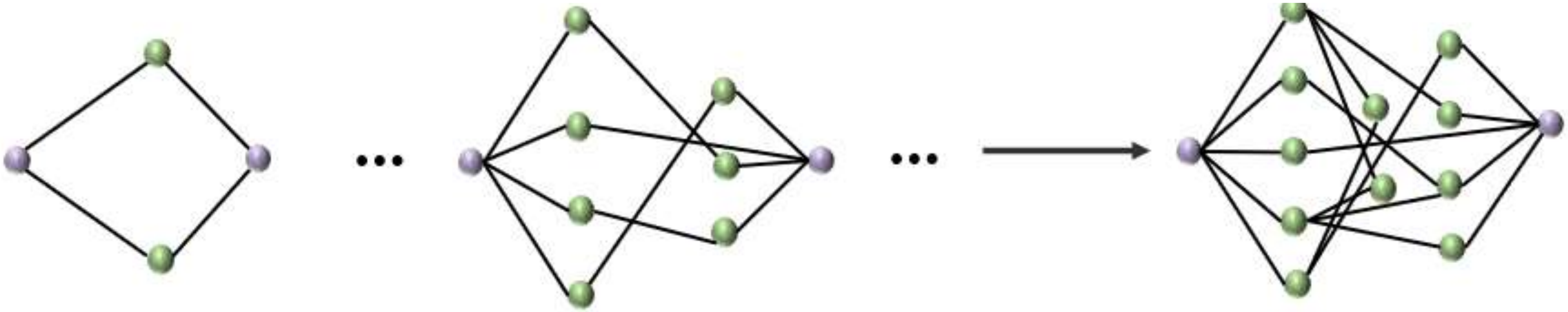


Time-multiplex the data streaming in the optical neural network enables:

- Reconfigurability in neural architecture
- Deep neural network
- Only optical layers (no linear regression)
- Not fully optical (lots of ADC and DAC)

Architecture reconfigurable optical neural network



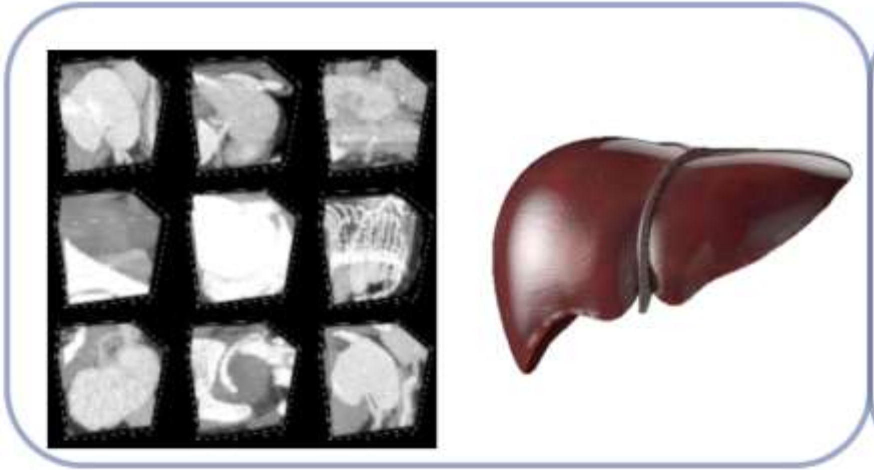


Computing task: 3D medical image classification

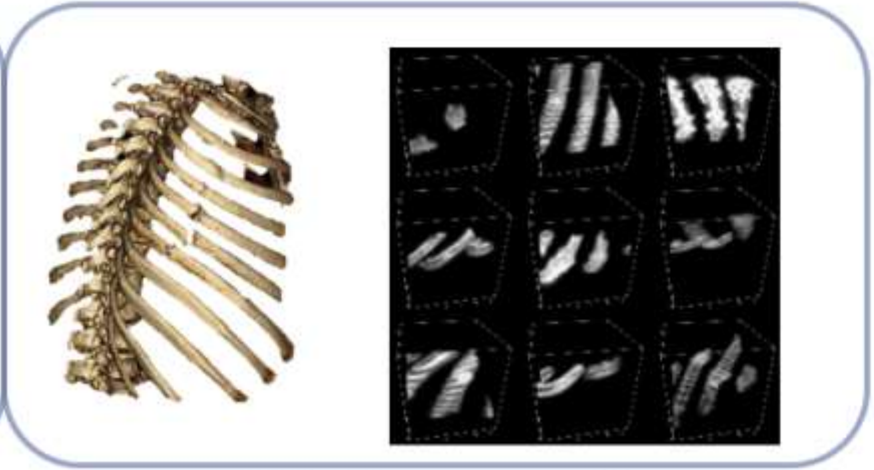
Challenging Tasks Even for Advanced Deep Learning Models

Organ 3D and Fracture 3D from M

3D CT scans for body organs or rib fractures (More than 2D images)
Large scale (Dimension is 2.20×10^4)
Not toy task, but application-level database in medicine.



Organ 3D



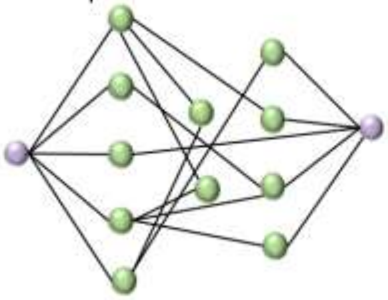
Rib Fracture 3D

arXiv preprint arXiv:1901.04056, 2019.

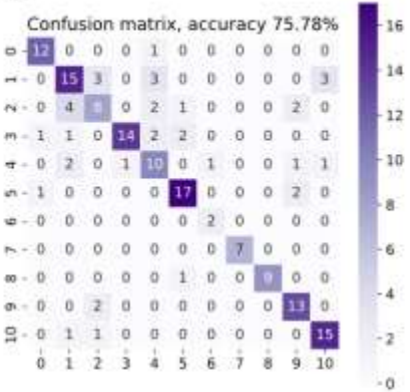
Optimal architecture is needed for best performance

Dataset 1: 3D organ classification result

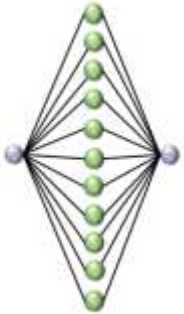
Optimal architecture

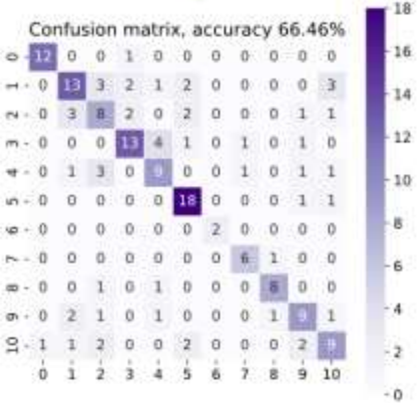


Test result

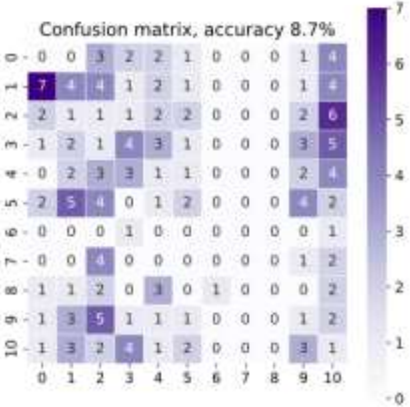


Simple architecture I: in parallel





Simple architecture II: in series

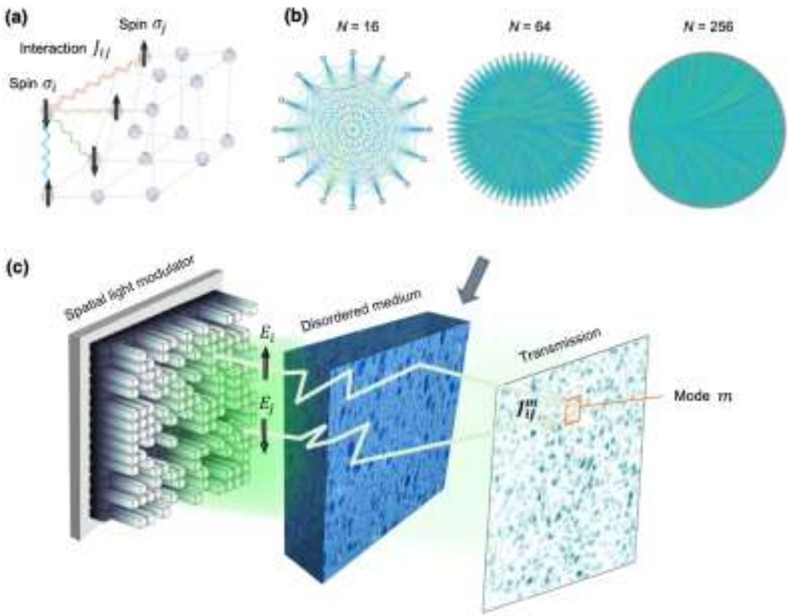


Large-scale reconfigurable ONN: performance

Methods	3D Organ	3D Fracture	
	accuracy	accuracy	# parameters
LS-ONN (ours)	**0.7578**	**0.5167**	**~2×10^5 (total)**
1-hidden layer MLP (activation: relu)	0.11	0.45	4.39×10^4
ResNet-50 + 2.5D	**0.769**	0.397	4.58×10^7
ResNet-18 + 2.5D	0.788	0.451	3.17×10^7
AutoKeras	0.804	0.458	
auto-sklearn	0.814	0.453	
ResNet-50 + 3D	0.883	0.494	4.64×10^7
ResNet-18 + 3D	0.907	**0.508**	3.33×10^7

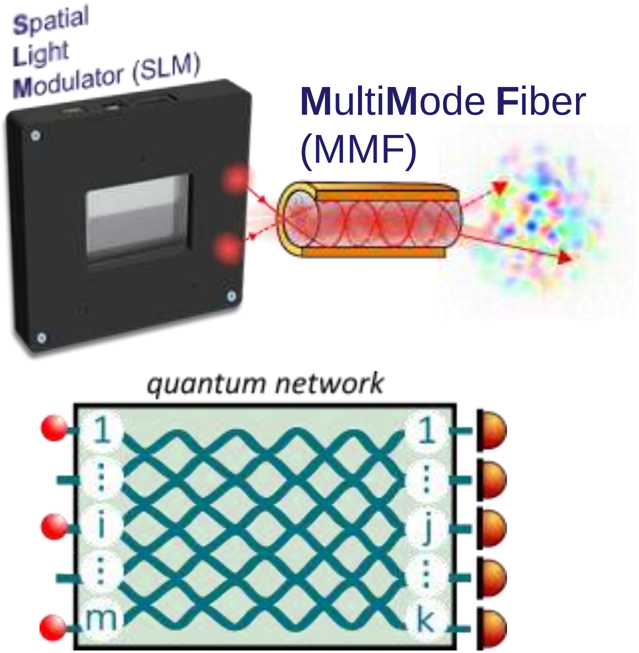
Other recent works

Scalable Spin-Glass Optical Simulator



Pierangeli et al. Phys. Rev. Applied 15, 034087 (2021)
Iaccuci et al. Phys. Rev. A (2022)
Collaboration : Claudio Conti (Roma)

Programmable linear circuits in a multimode fiber



Leedumrongwatthanakun, S.. et al.,
Nature Photonics 14, 139–142 (2020).

Take-home message

Optics & Photonics have unique advantages for current & future artificial Neural networks

Fast

Large-dimensional

Low-power consumption

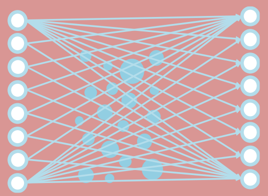
Weight tuning and training

Non-linearities

Size

- NO fully optical deep neural network
- HYBRID electro-optics implementations
- mostly <u>very</u> simple proof-of-concepts
- Some specialized architectures are successful

Optical Random Projections

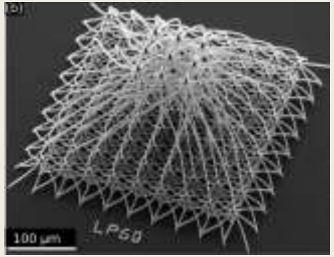


- Easy fabrication
- All-to-all connectivity
- Fixed weights
- Already at scale
- Low power consumption

➢ **Proof of principle:** classification, reservoir computing, Ising models …

Convergence between free space and integrated optics?

"Best of both worlds"



Optica 7, 640-646 (2020) (D. Brunner team – FEMTO-ST)

Convergence between algorithms & optics?

(Not just deep learning)

"the hardware lottery"

S. Hooker arXiv:2009.06489

Convergence between classical & quantum optical machine learning?

Science 370.6523 (2020)

Thanks to my team and collaborators

Thank you for your attention !

Mail : sylvain.gigan@lkb.ens.fr
Webpage: www.lkb.ens.fr/gigan

If you are interested in the field :

REVIEWS OF MODERN PHYSICS

Light fields in complex media: Mesoscopic scattering meets wave control

Stefan Rotter and Sylvain Gigan
Rev. Mod. Phys. 89, 015005 – Published 2 March 2017

Perspective | Published: 02 December 2020

Inference in artificial intelligence with deep optics and photonics

Gordon Wetzstein, Aydogan Ozcan, Sylvain Gigan, Shanhui Fan, Dirk Englund, Marin Soljačić, Cornelia Denz, David A. B. Miller & Demetri Psaltis

Nature **588**, 39–47(2020) | Cite this article

Perspective | Published: 08 September 2022

Imaging and computing with disorder

Sylvain Gigan

Nature Physics **18**, 980–985 (2022) | Cite this article