

# Machine learning algorithms for optimized data transmission links

Francesco Da Ros

Technical University of Denmark  
[fdro@fotonik.dtu.dk](mailto:fdro@fotonik.dtu.dk)

# Outline

## Part I – Communicating over the optical fiber with ML

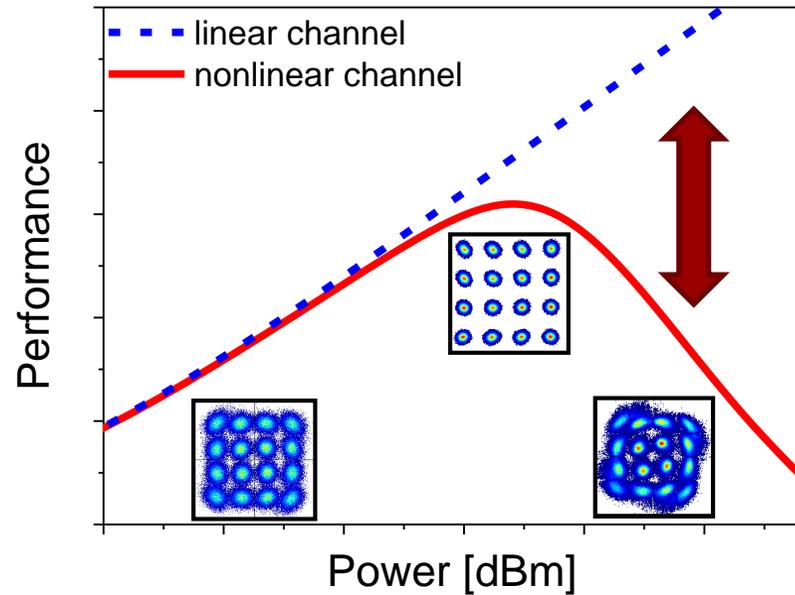
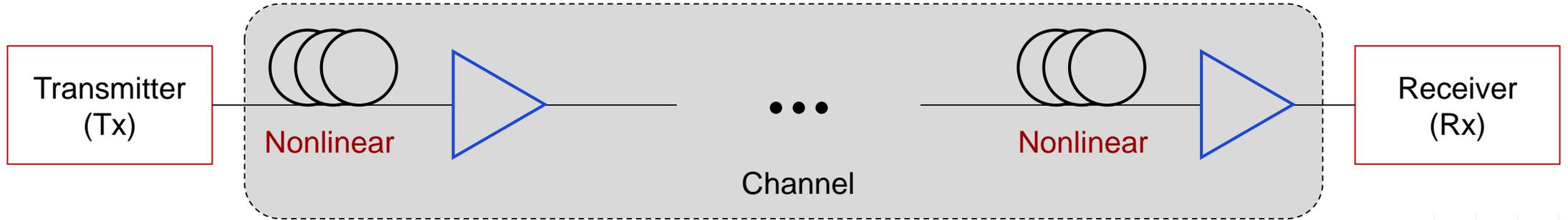
- Long-haul communication
  - Optimizing signalling schemes for performance/robustness
- Short-reach communication
  - Combating power-fading effects with equalization

## Part II – Photonic ML accelerators

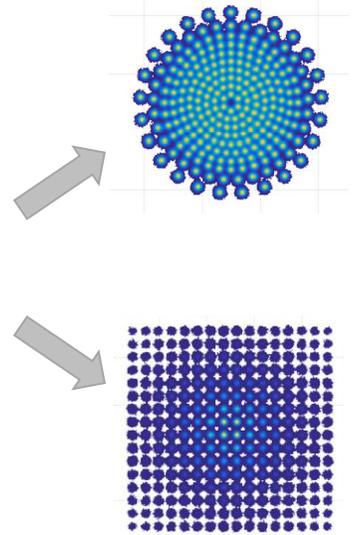
- Weight training on integrated photonics
  - Impact of thermal crosstalk
  - Offline training: chip-model accuracy vs. task performance

- Conclusions

# Long-haul optical communication



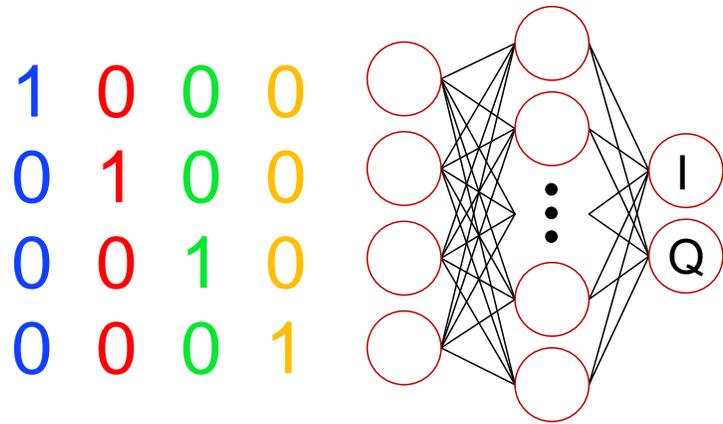
Constellation shaping



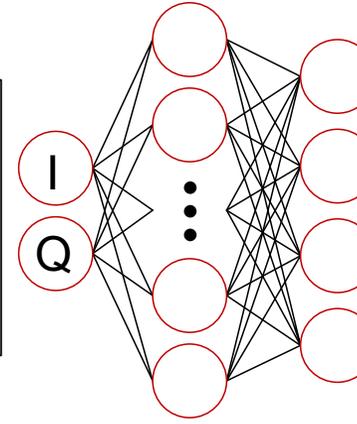
Improve the information rate by optimizing the signalling scheme for the desired channel.

# Autoencoders for end-to-end learning

Input Space:



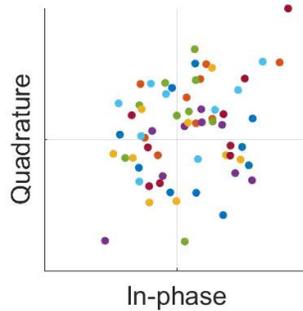
Fiber Channel Model



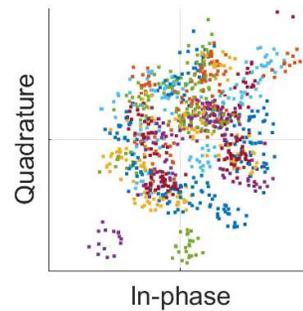
Output Space:

.8	.1	.1	.0
.1	.8	.0	.1
.1	.0	.8	.1
.0	.1	.1	.8

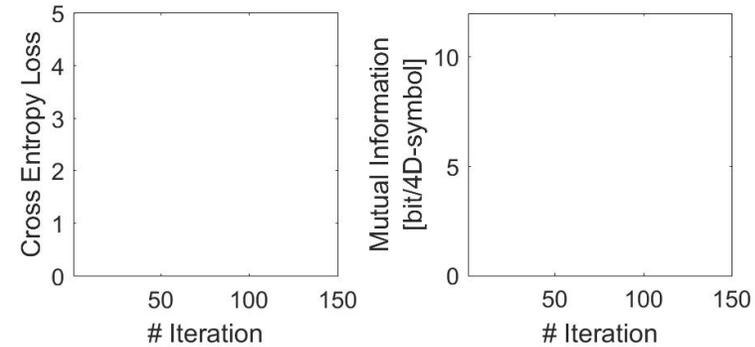
Channel input



Channel output

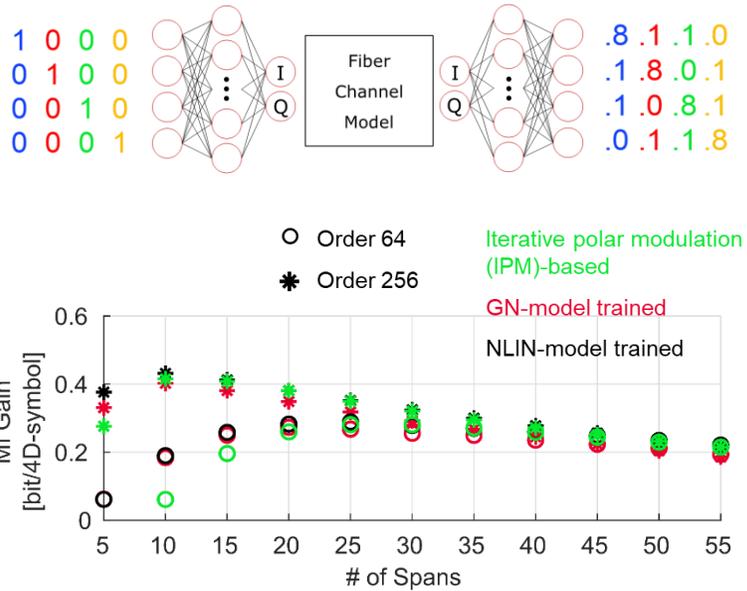


Performance

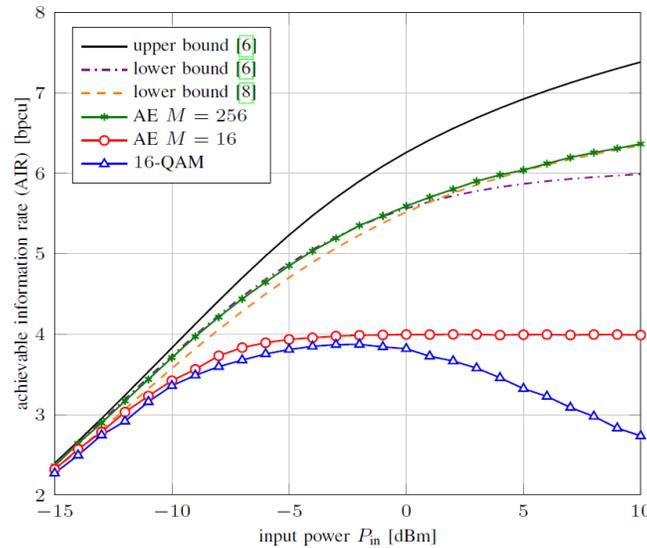


# AE-based constellation shaping

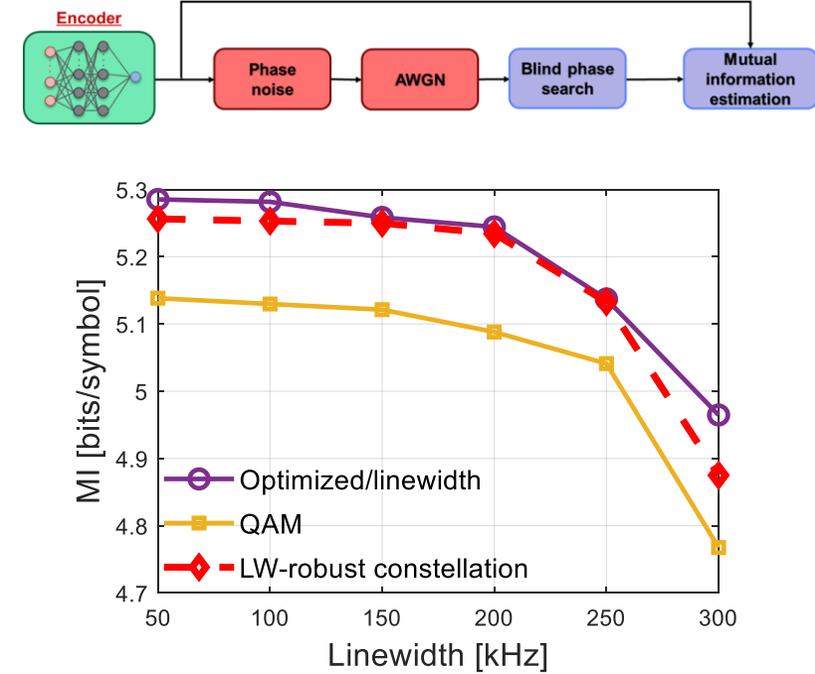
## Shaping trained over **NLIN** and **GN** fiber model



## Shaping trained over **NLPN** channel model



## Shaping **robust** to system uncertainties, enabling **interoperability**



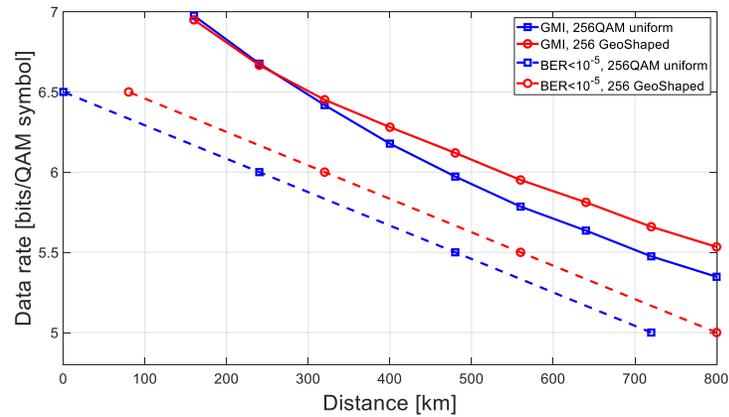
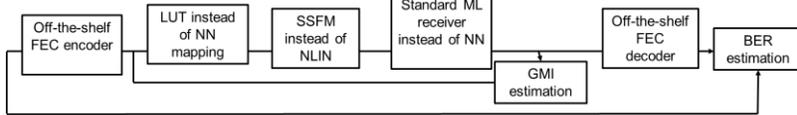
NLIN: nonlinear interference noise  
 GN: Gaussian noise  
 R. T. Jones, ECOC 2018

NLPN: nonlinear phase noise  
 S. Li, ECOC 2018

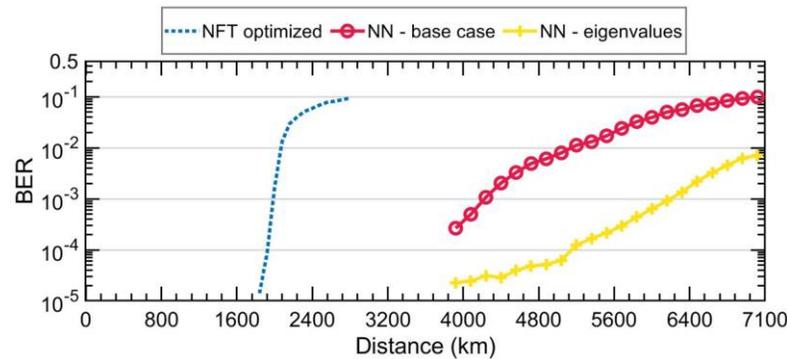
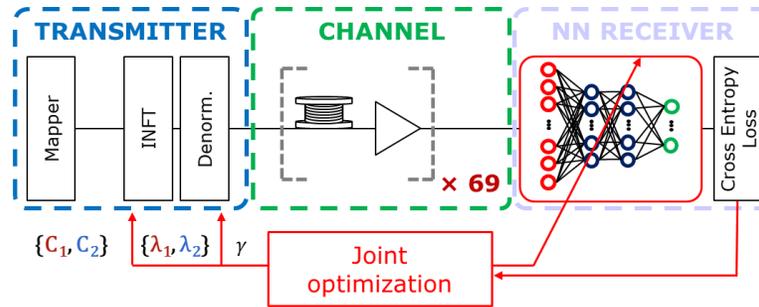
LW: linewidth  
 RPN: residual phase noise  
 O. Jovanovic, ECOC 2021

# Beyond constellation shaping

## Shaping and bit labelling trained over NLIN channel

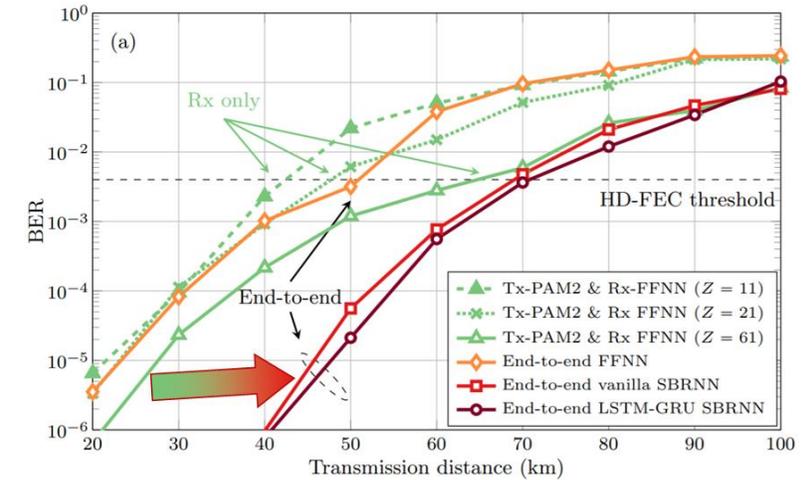
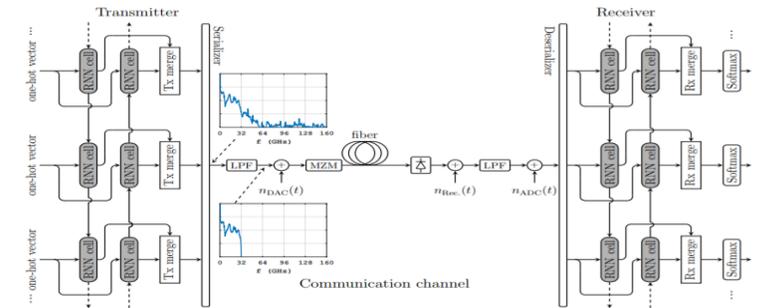


## Shaping and waveform design for transmission over SSFM channel



NFDM: nonlinear frequency division multiplexing  
 SSFM: split-step Fourier method  
 S. Gaiairin, JLT 2020

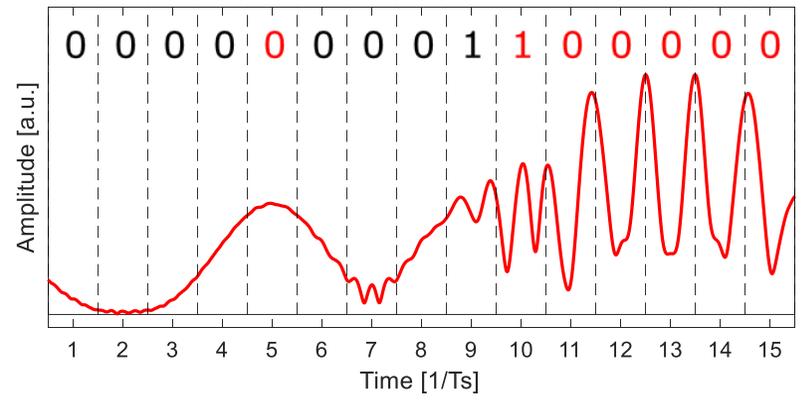
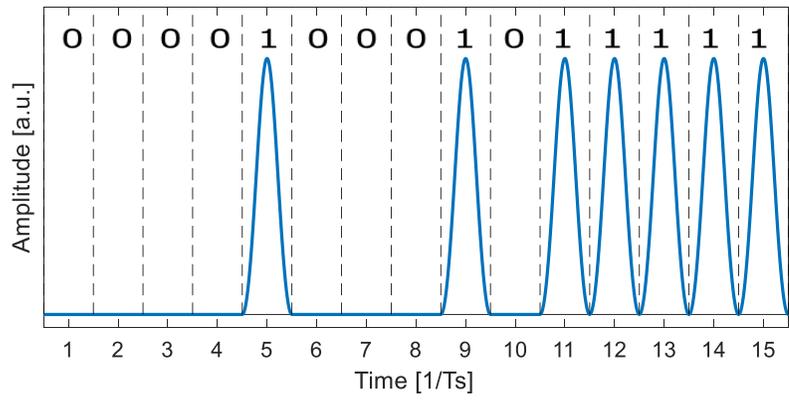
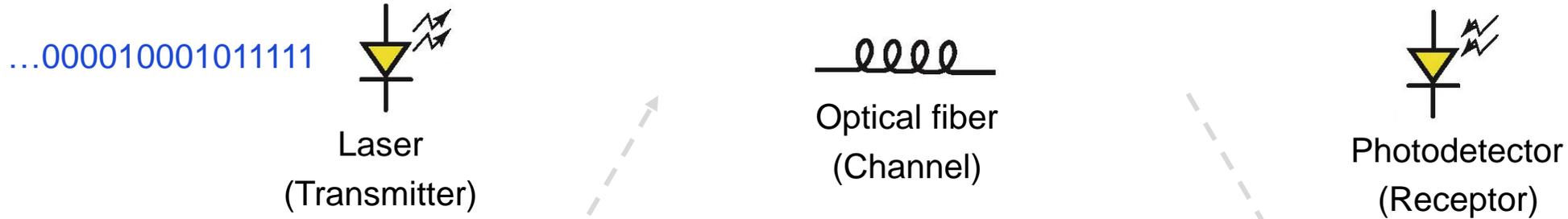
## Shaping and waveform design for transmission over dispersive channel



B. Karanov, JLT 2018

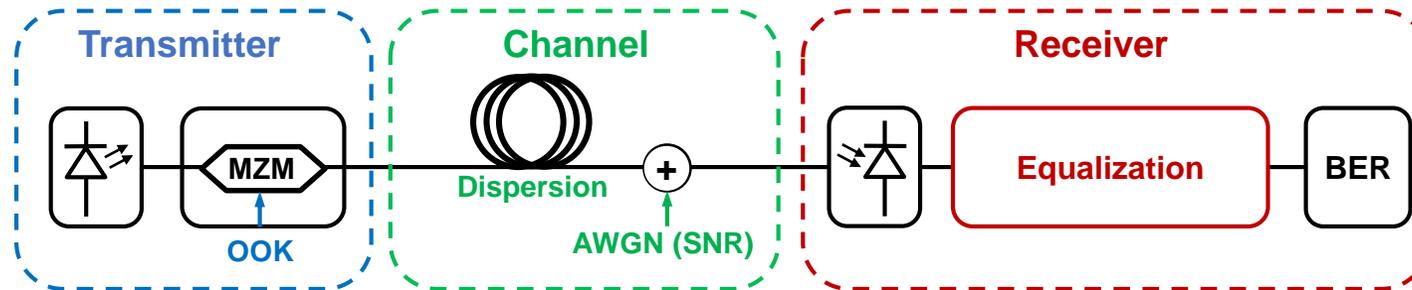
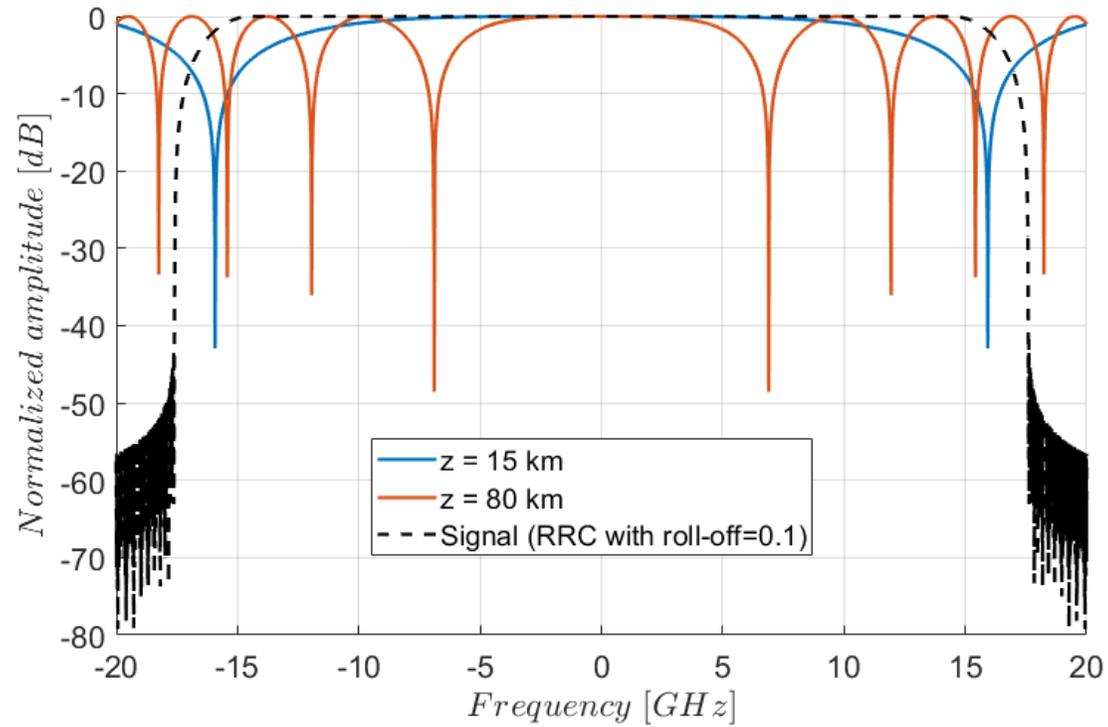
B. Karanov, Opt. Expr. 2019

# Short-reach optical communication



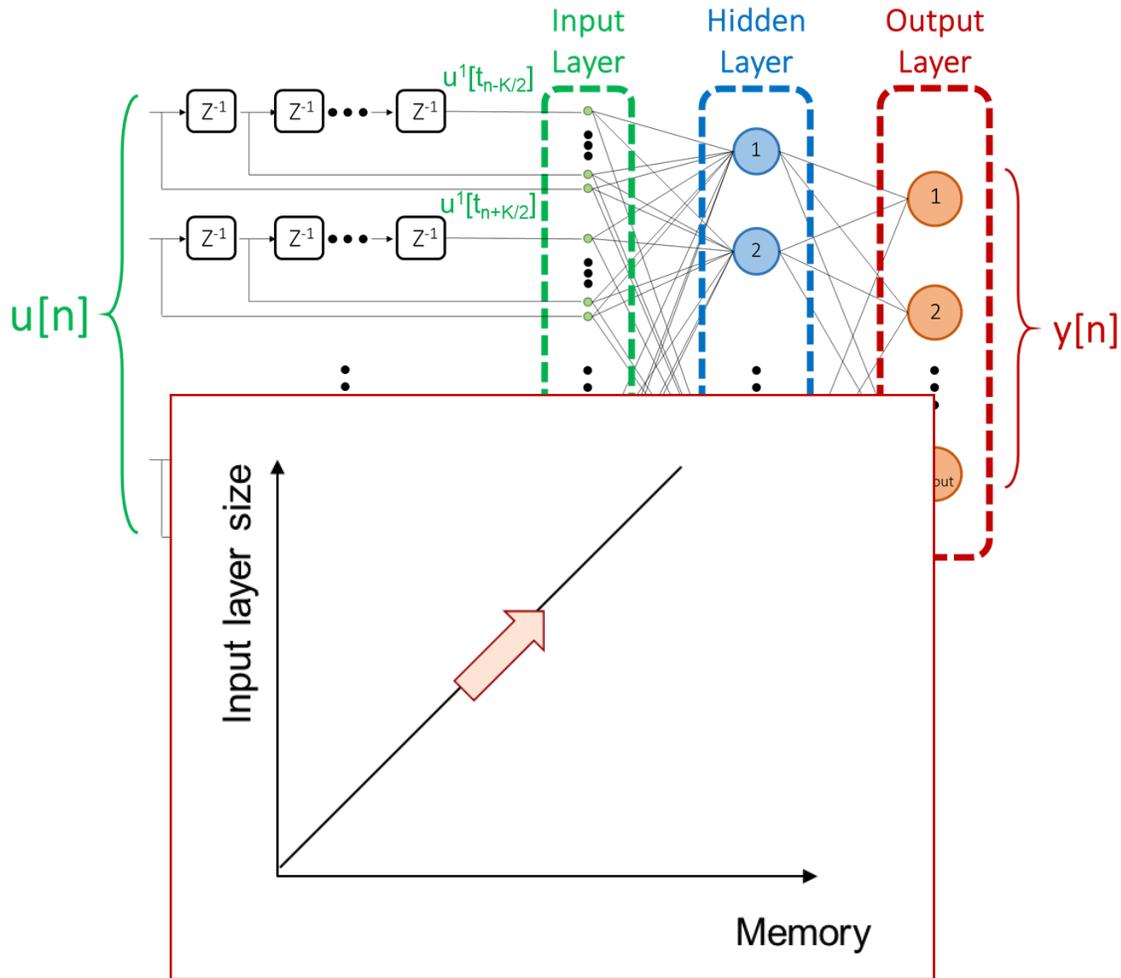
Fiber dispersion + direct-detection = inter-symbol interference

# Power-fading and equalization

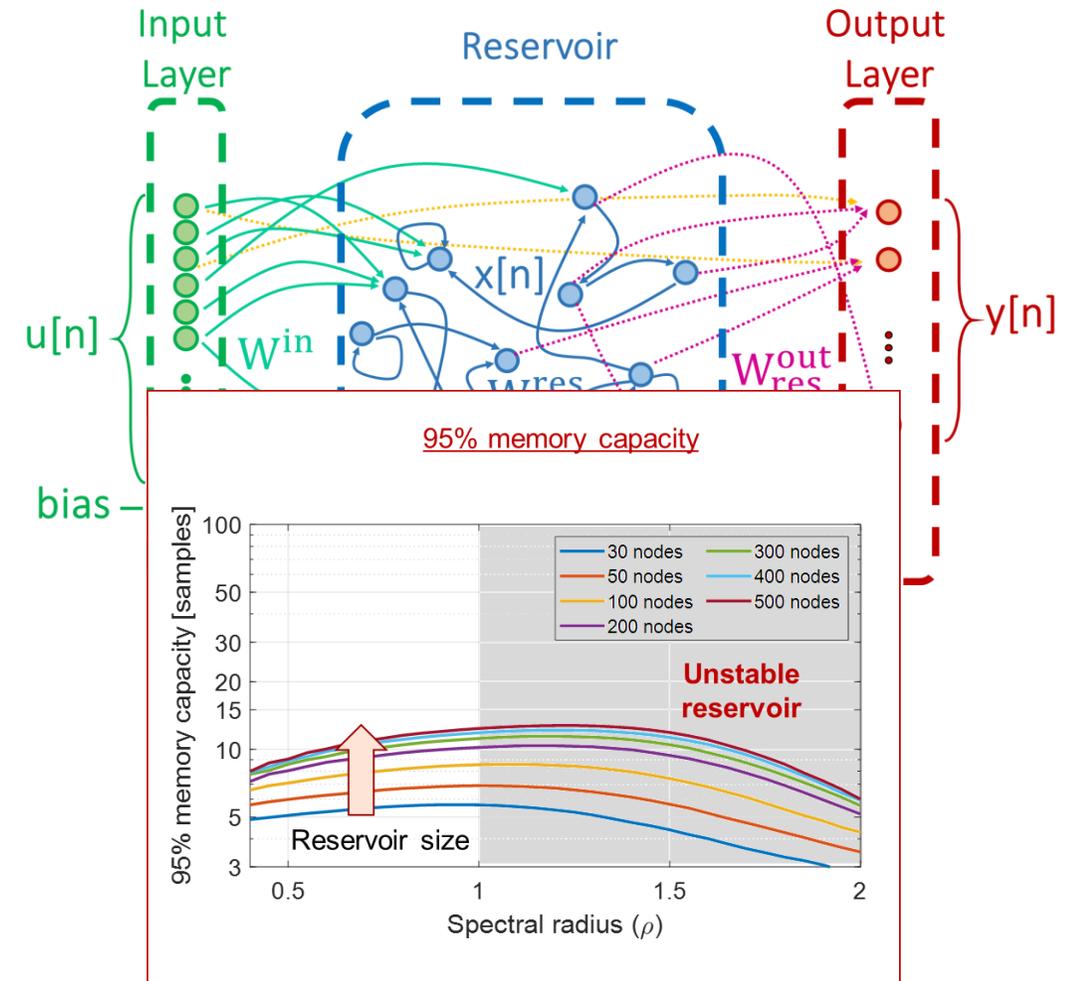


# Equalizers with memory

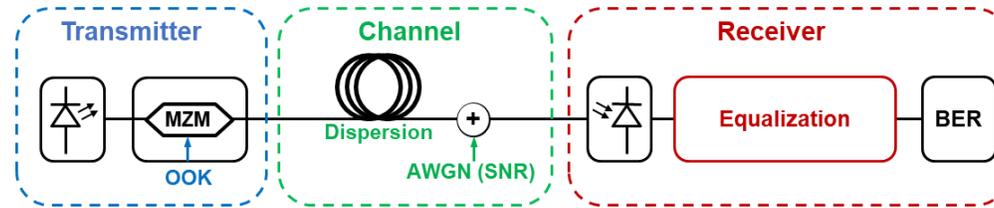
Time-delay feed-forward neural network



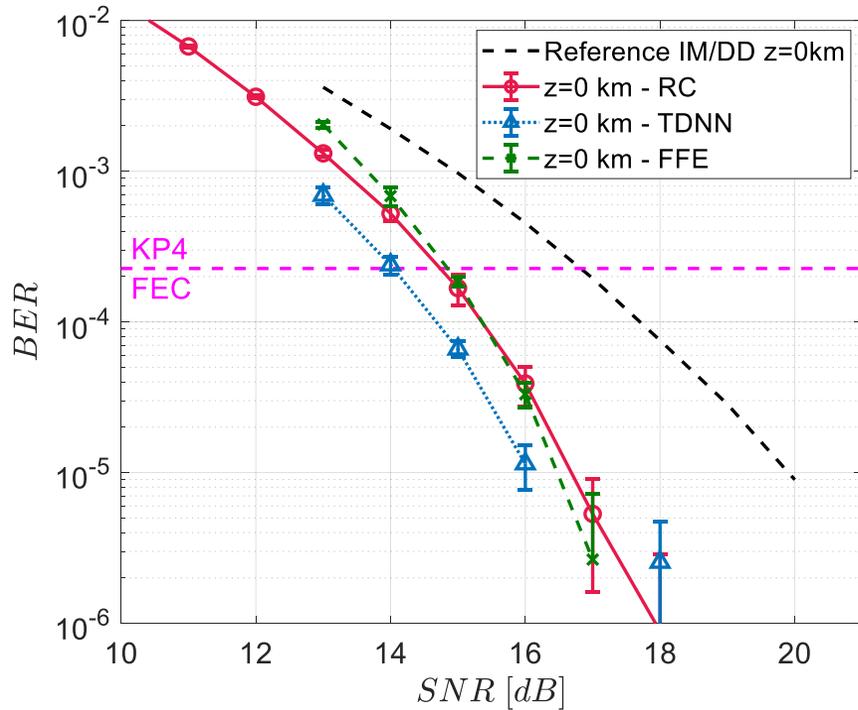
Reservoir computing



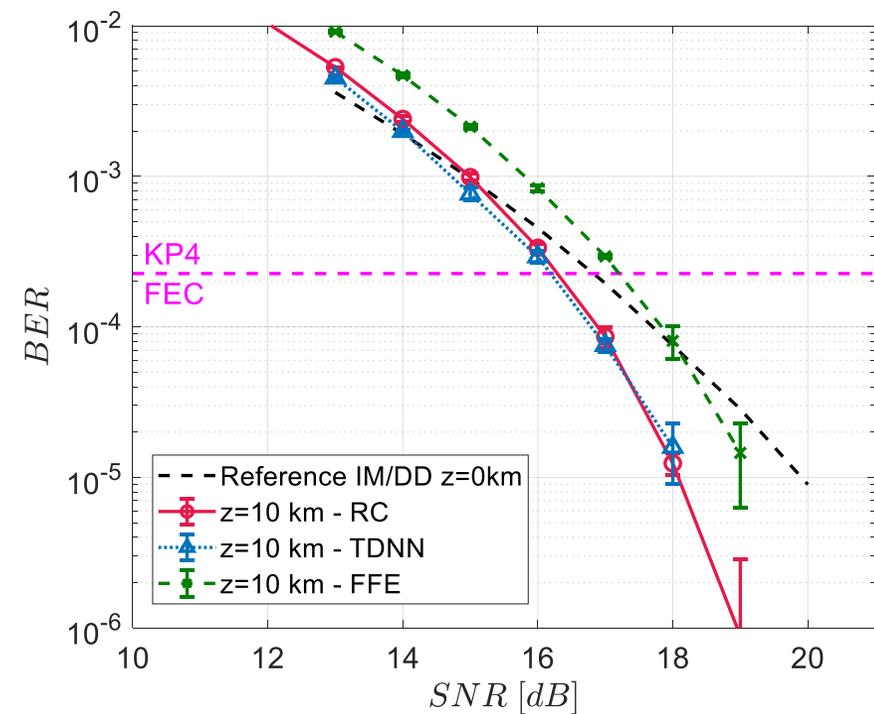
# Numerical comparison



Back-to-back



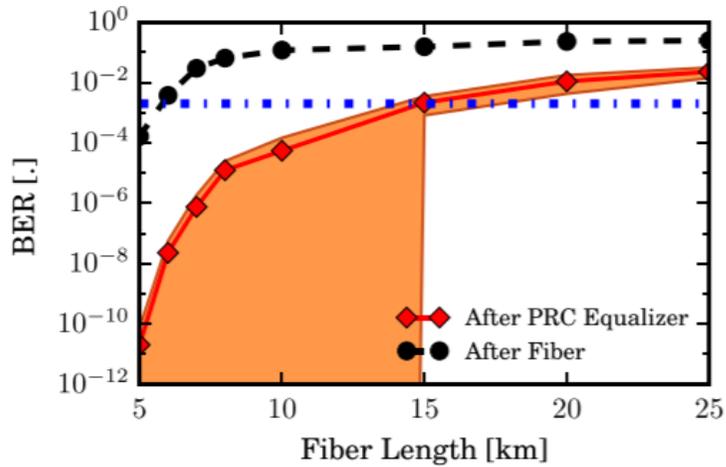
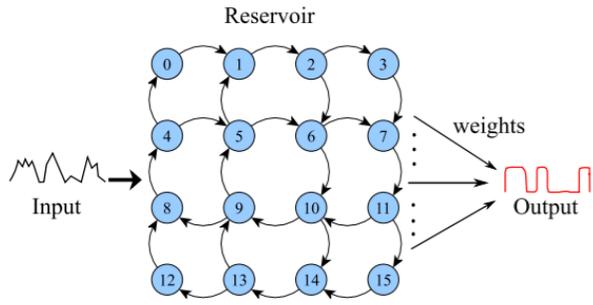
Fiber transmission



Nonlinear equalizers with memory help in compensating for power fading after photodetection.

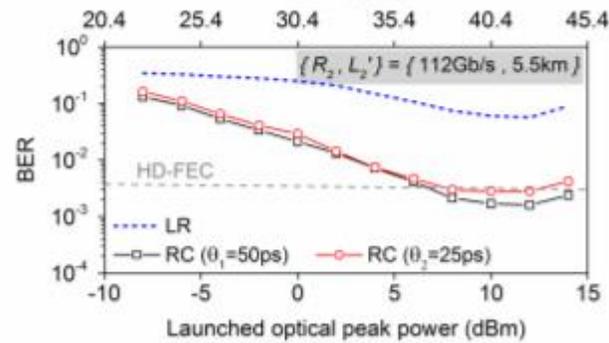
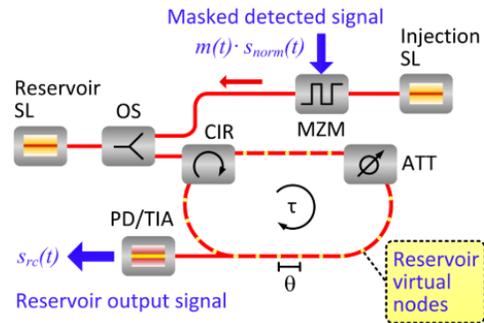
# Optoelectronic reservoir computing

## Silicon-photonic reservoir with optoelectronic output layer



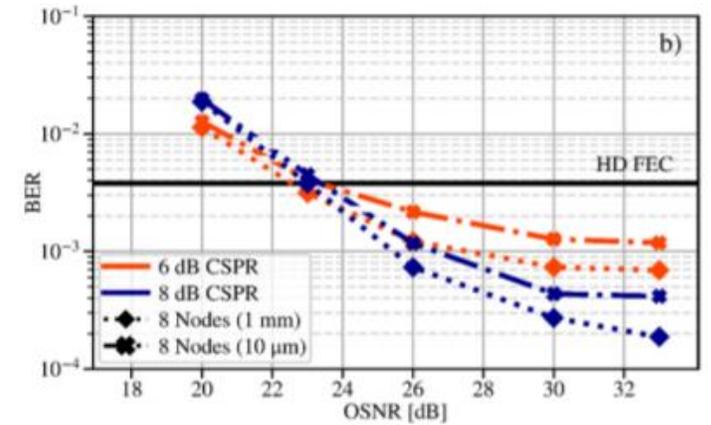
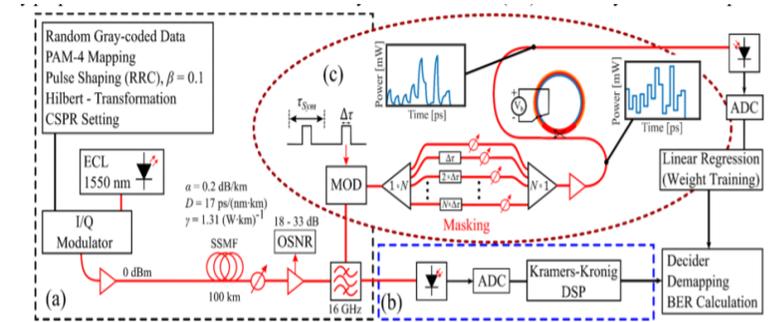
Katumba, et al., JLT 2019

## Laser-based reservoir for time-delayed RC



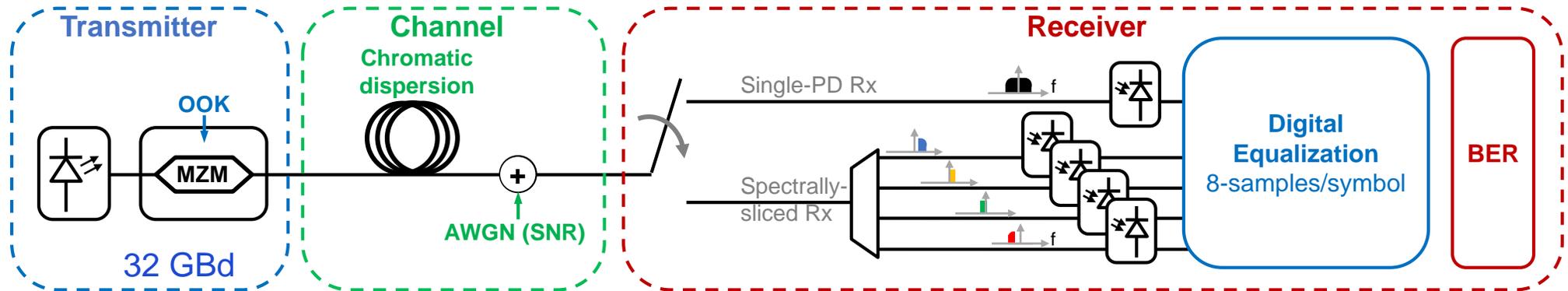
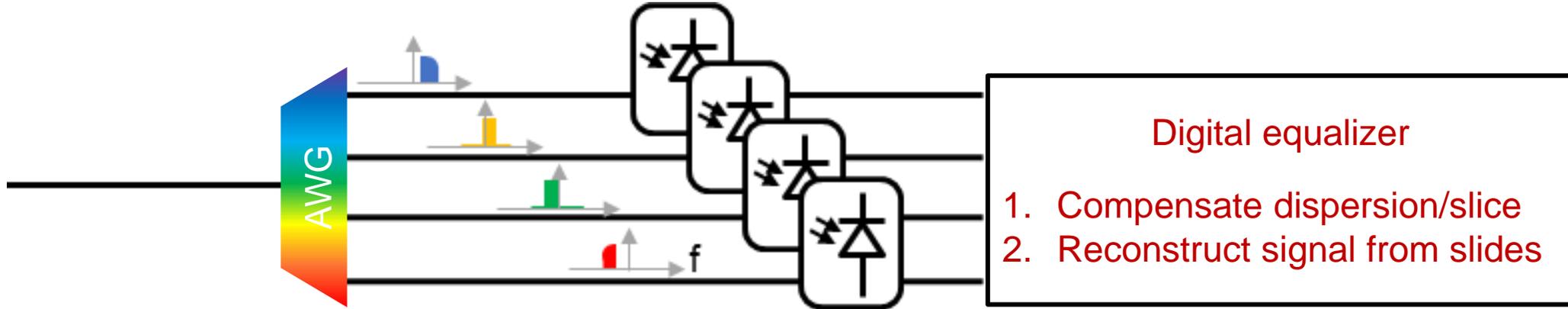
Argyris, et al, IEEE Access 2019

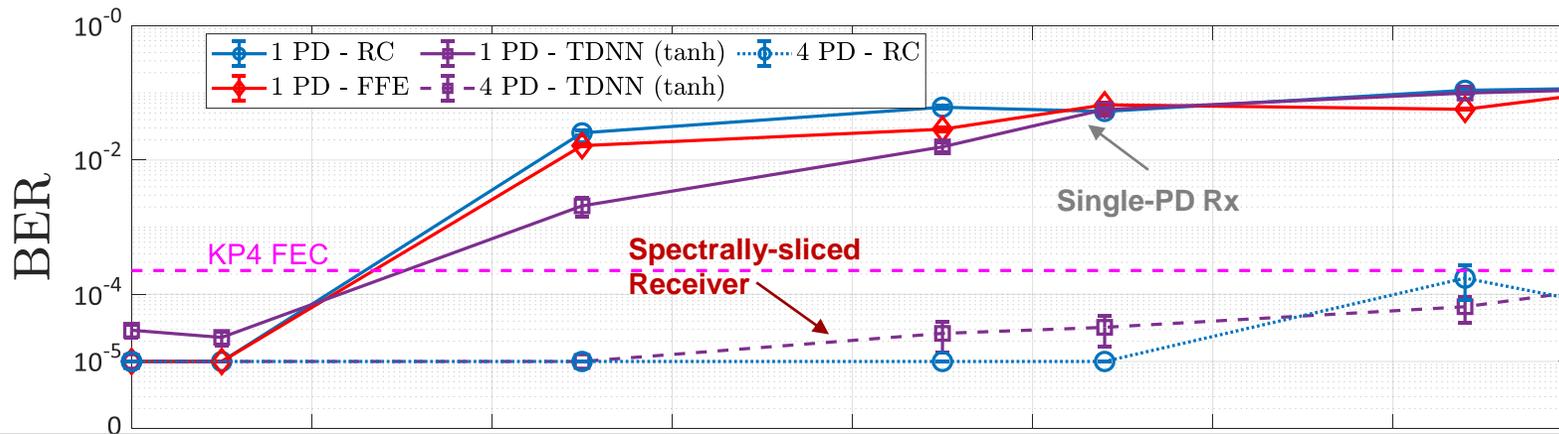
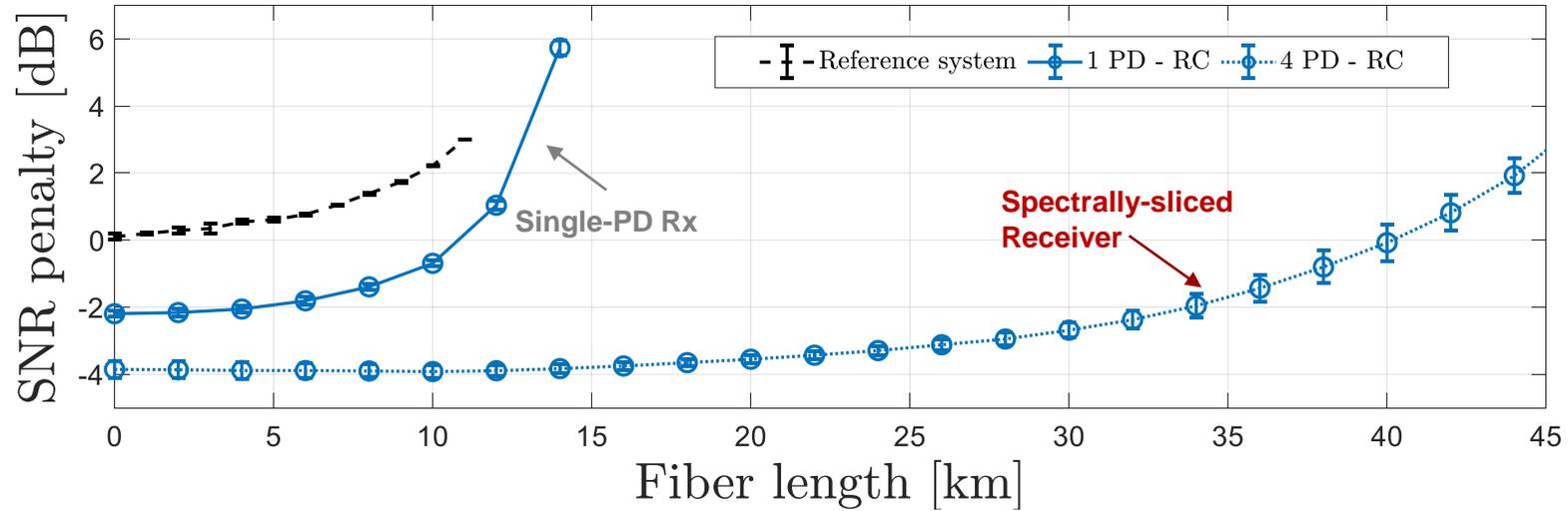
## Micro-ring resonator based reservoir



Li, et al., CLEO 2020

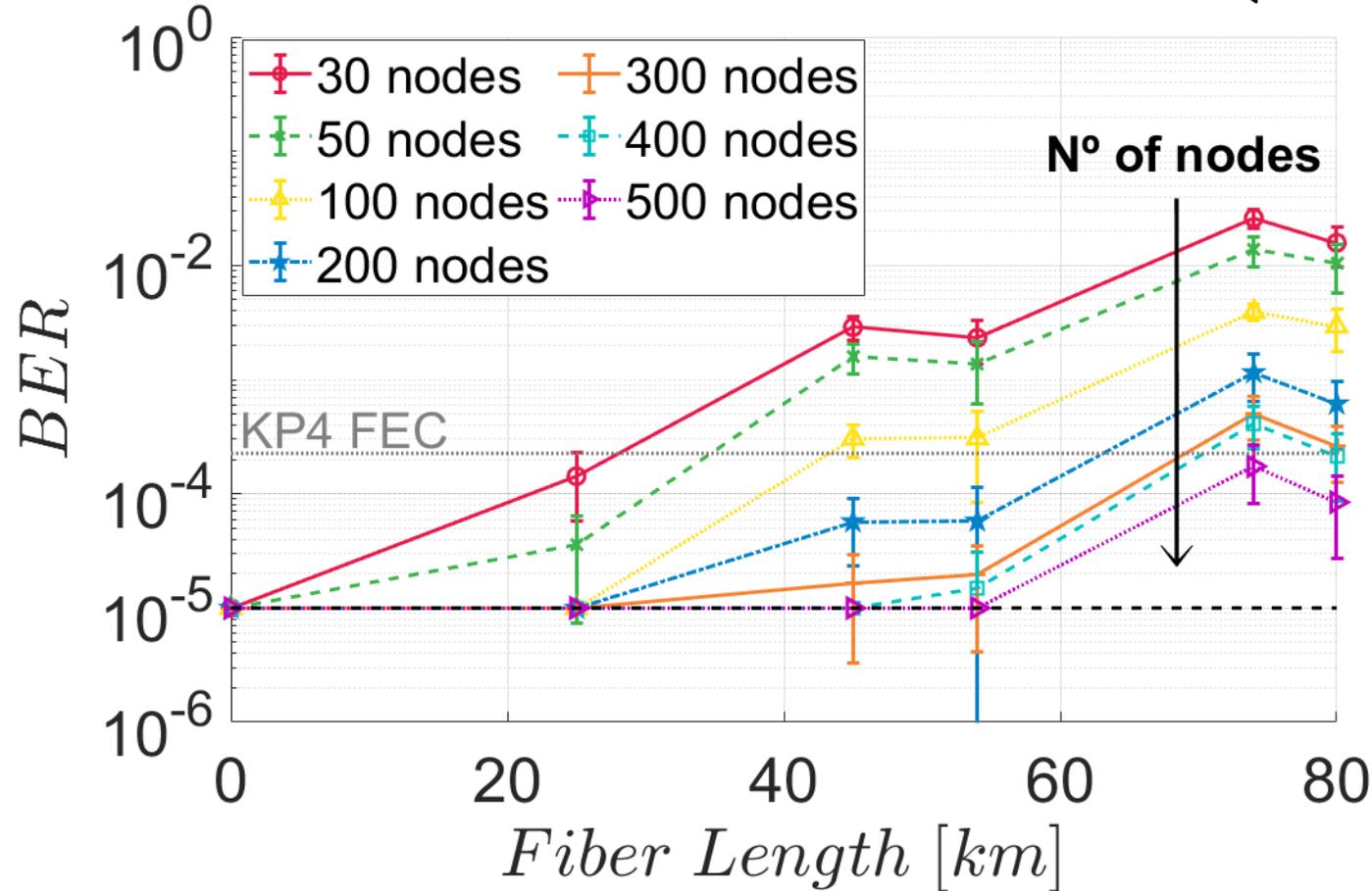
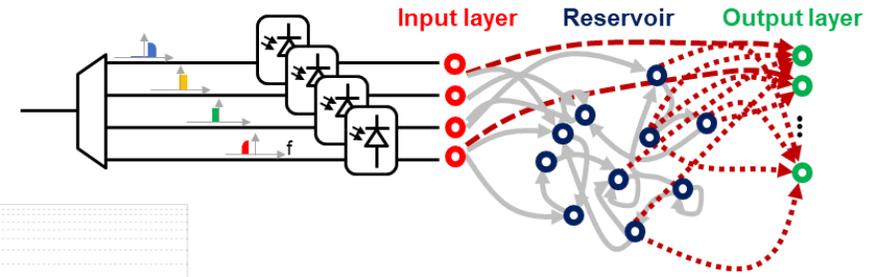
# Optoelectronic receiver with digital equalizers





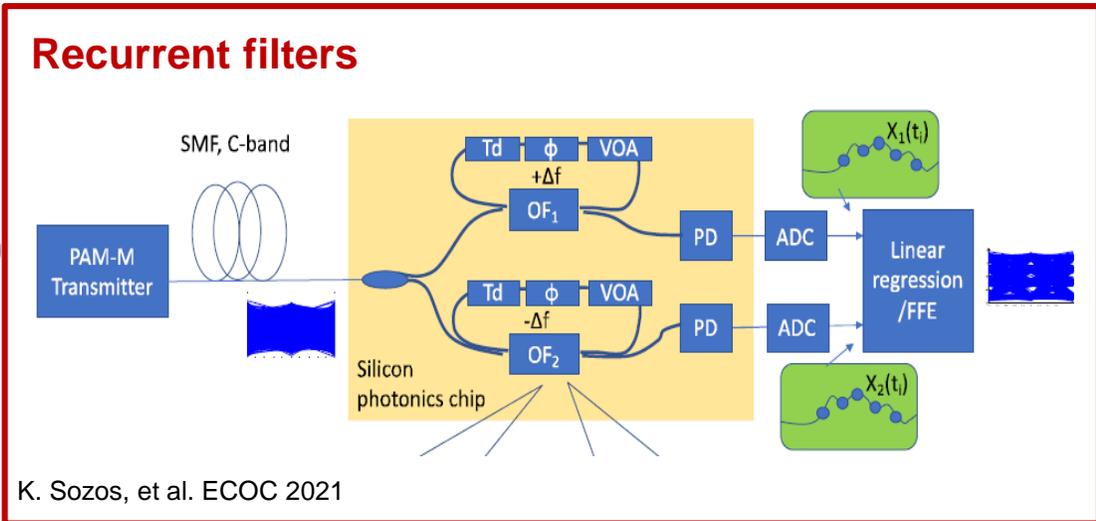
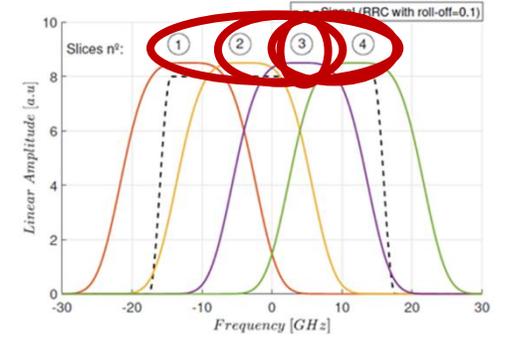
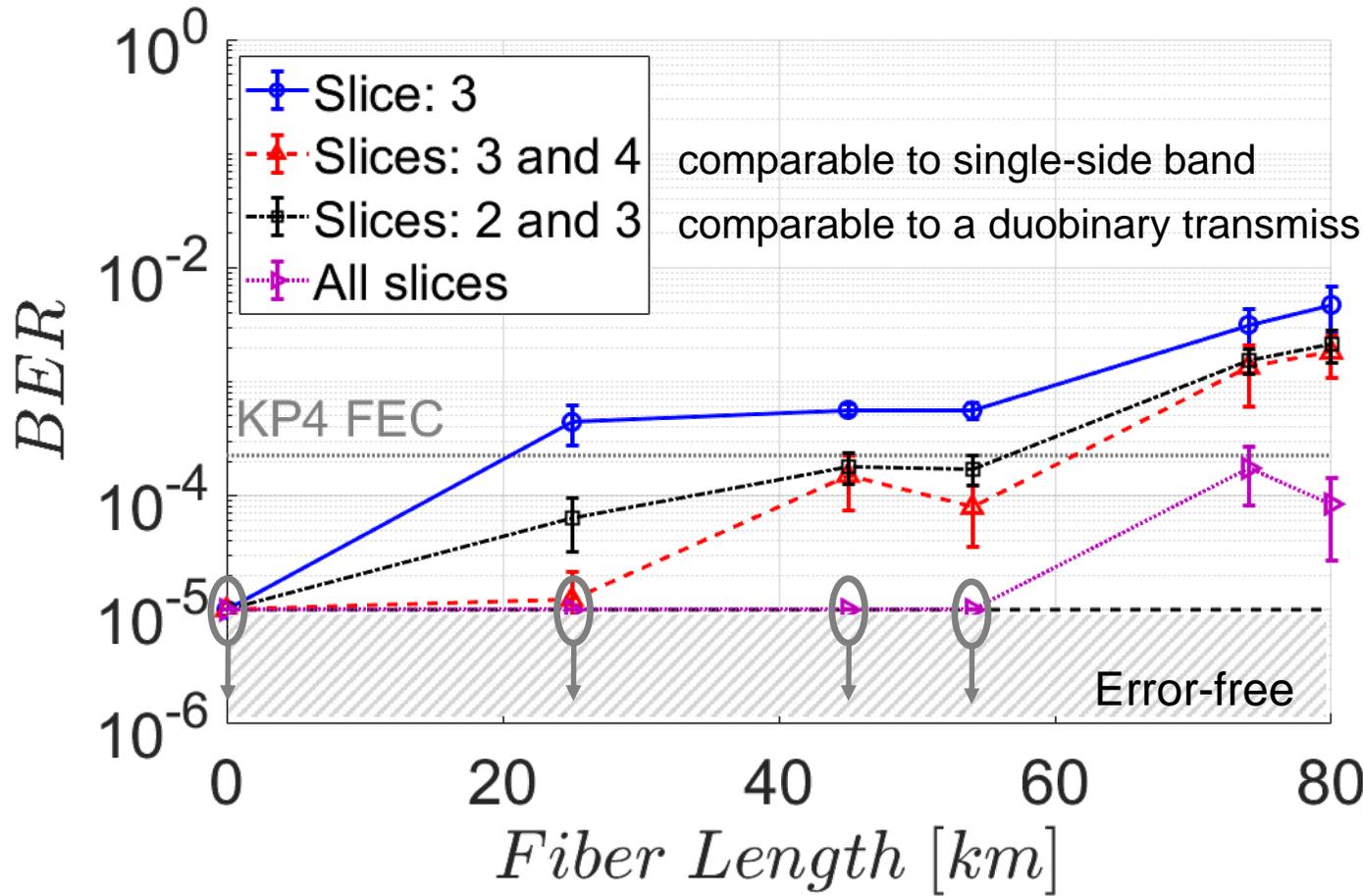
Spectral-slicing + equalization significantly enhances the transmission reach.

# Impact of reservoir memory



The equalizer memory needs to be adjusted for the specific transmission distance.

# Complexity-performance trade-off



K. Sozos, et al. ECOC 2021

# Outline

## Part I – Communicating over the optical fiber with ML

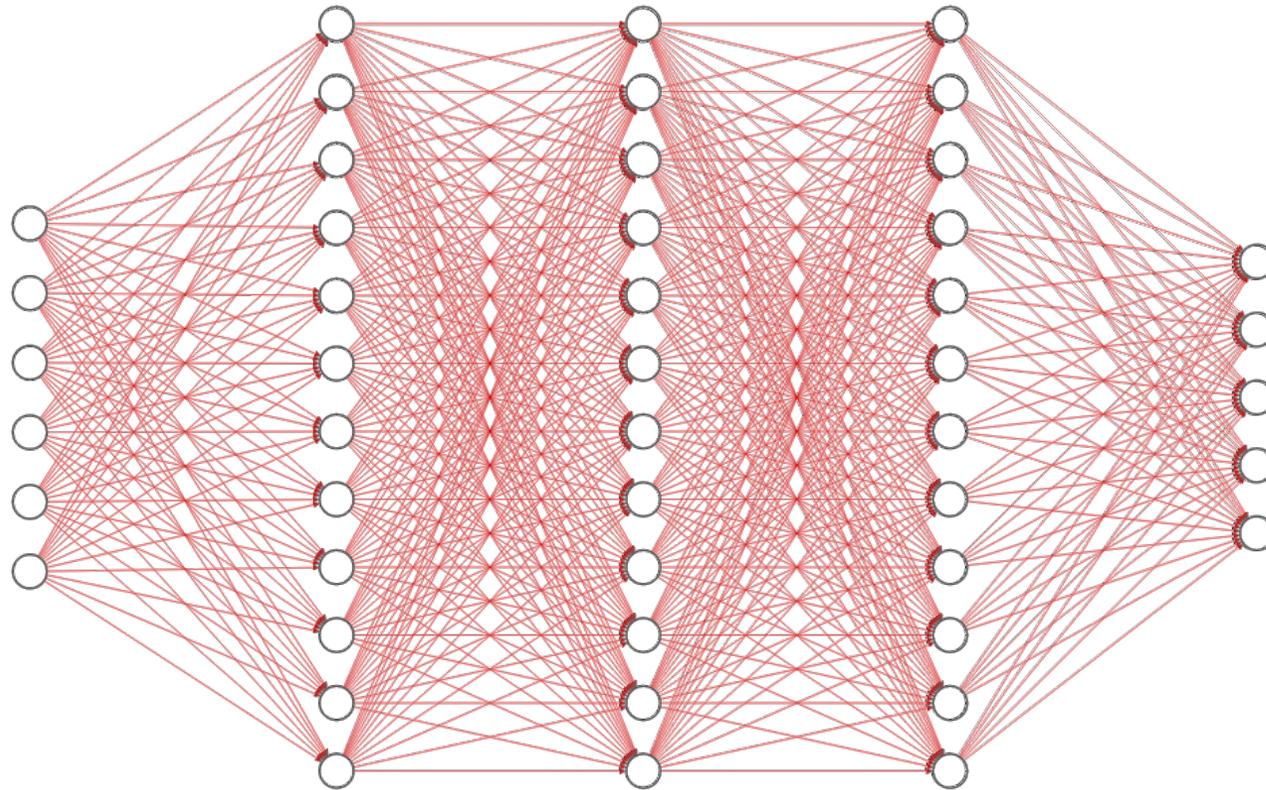
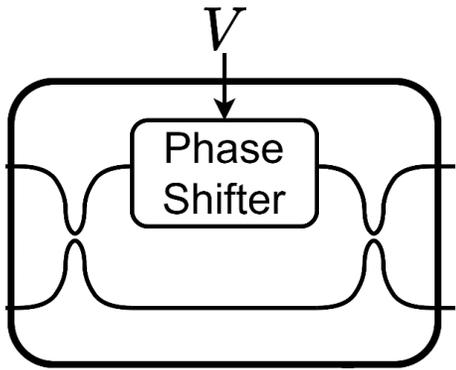
- Long-haul communication
  - Combating fiber nonlinearity
  - Signalling scheme resilient to system uncertainties
- Short-reach communication
  - Combating power-fading effects
  - Optoelectronic receivers

## Part II – Towards optical ML

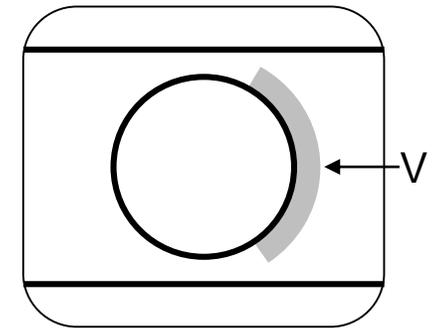
- Weight training on integrated photonics
  - Impact of thermal crosstalk
  - Offline training: chip-model accuracy vs. task performance
- Conclusions

# Photonic accelerators

Mach-Zehnder interferometers (MZI)

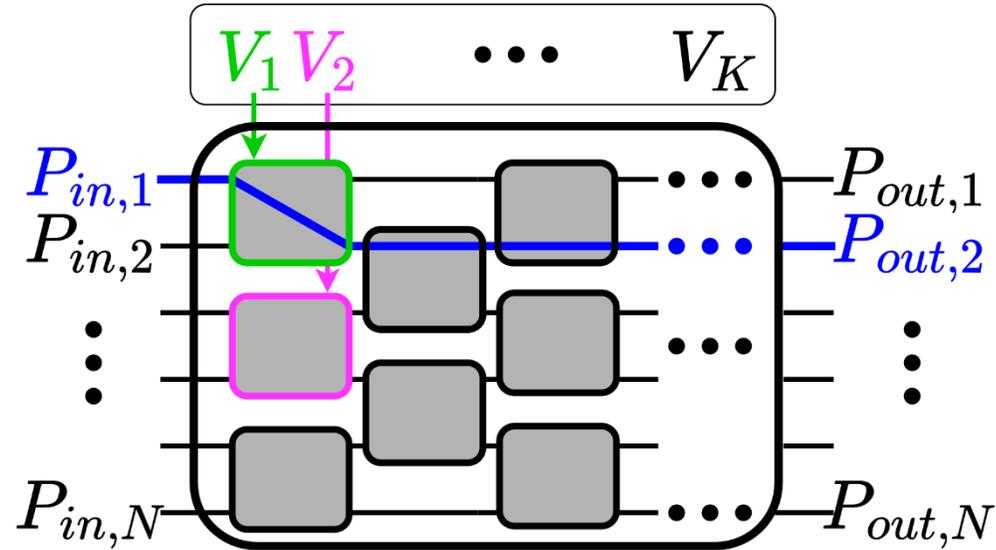


Micro-ring resonators (MRR)



Photonics is particularly effective in providing weigh-tunable interconnection → matrix multiplications.

# Training photonic networks



## Off-line training

- Relies on a PIC model
- Allows for faster re-configurability

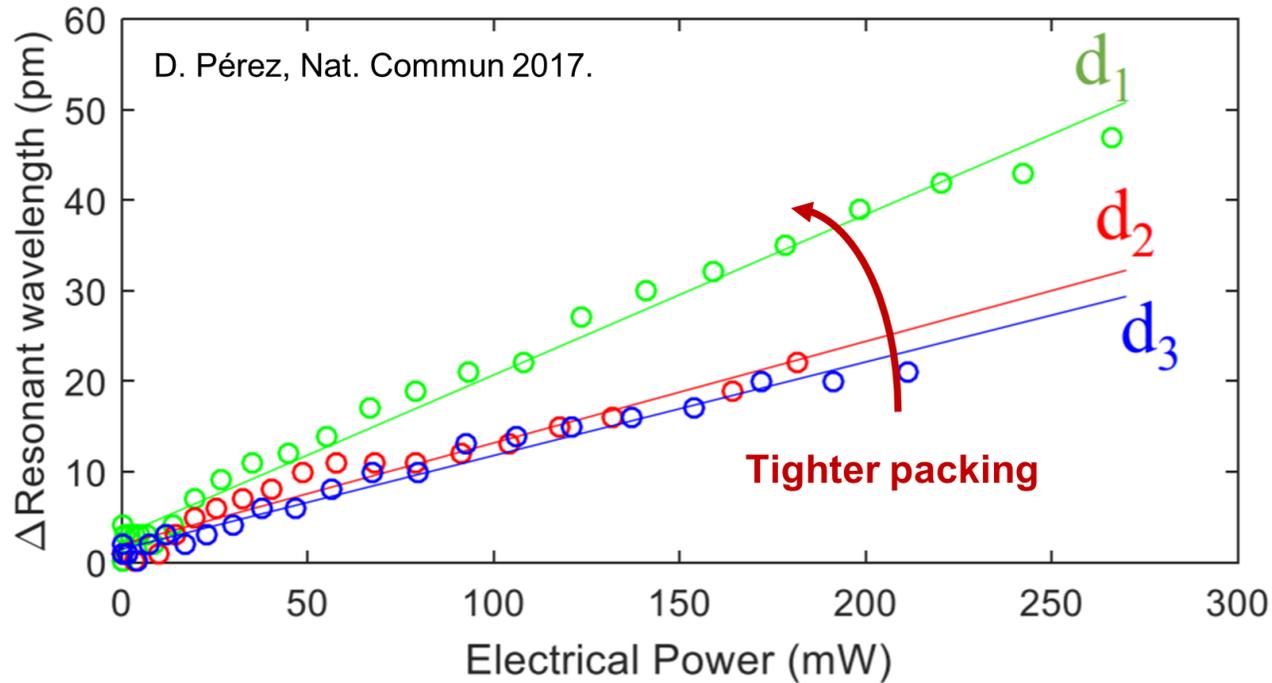
## On-line training

- Performed on the specific PIC
- Iterative procedure (re-train → re-start)

D. Pérez, Nat. Commun 2017.  
M. Fang Opt. Expr. 2019  
S. Bandyopadhyay, Optica 2021

Z. Thang Opt. Expr. 2019  
S. Pai JSTQE 2020  
M. Milanizadeh JSTQE 2020  
H. Zhang ACS Phot 2021

# Model-based off-line training



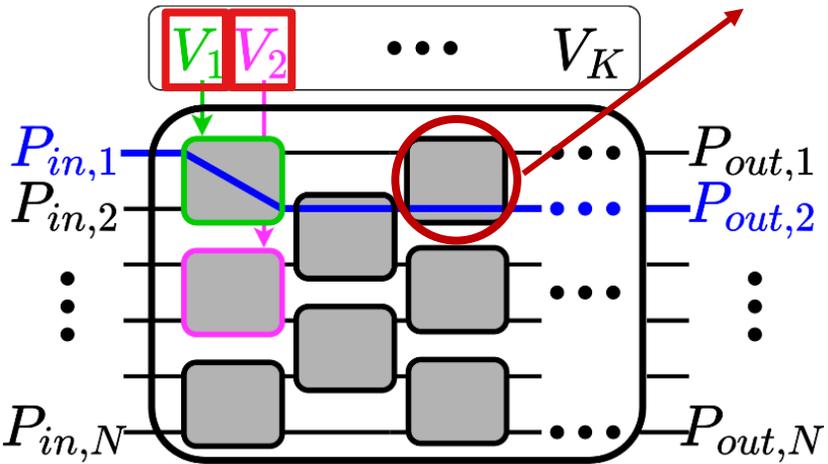
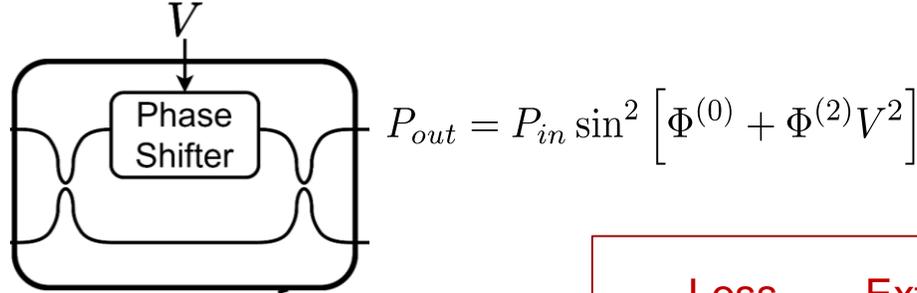
Packing MZI/MRR meshes tightly:

- Optical crosstalk – waveguide crossing
- Thermal crosstalk – thermal diffusion
- Electrical crosstalk – voltage delivery network



Are simple models accurate enough?

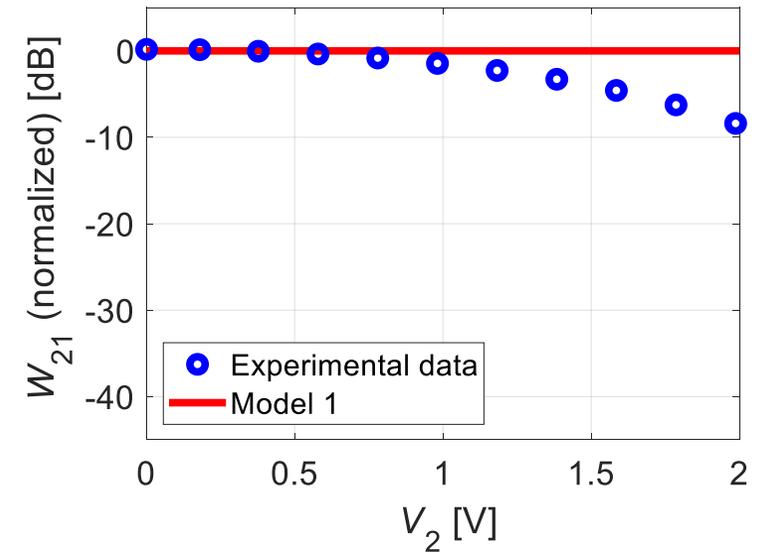
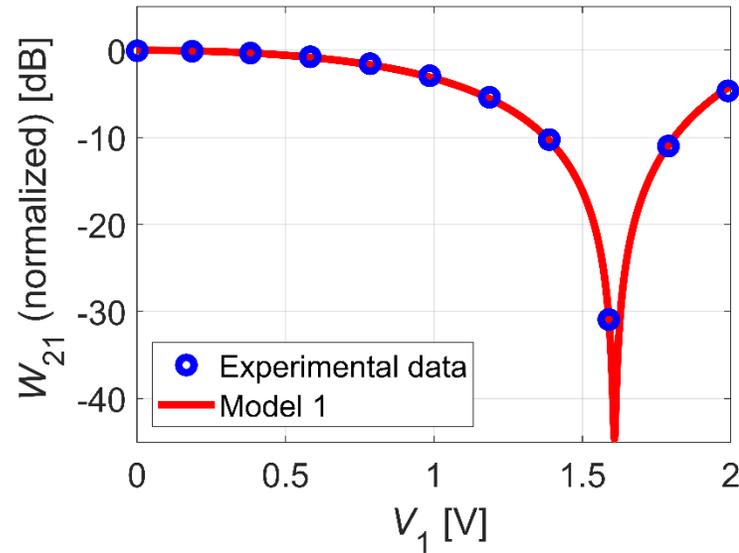
# Simple MZI mesh model



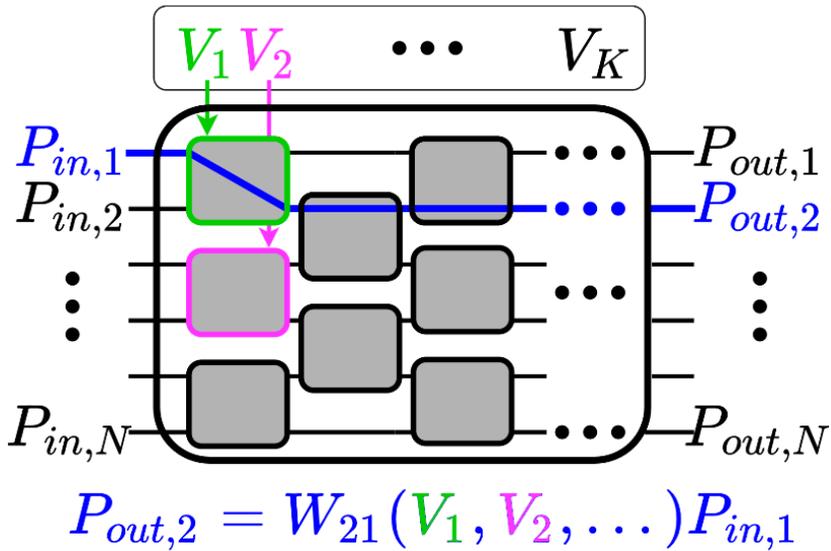
$$P_{out,2} = W_{21}(V_1, V_2, \dots) P_{in,1}$$

Loss      Extinction ratio

$$W_{ij} = L_{ij} \prod_{k \in K_{ij}} \frac{1}{4} \left| \frac{\sqrt{ER} - 1}{\sqrt{ER} + 1} - \exp\left(\sqrt{-1}(\phi_k^{(0)} + \phi_k^{(2)}V_k^2)\right) \right|^2$$

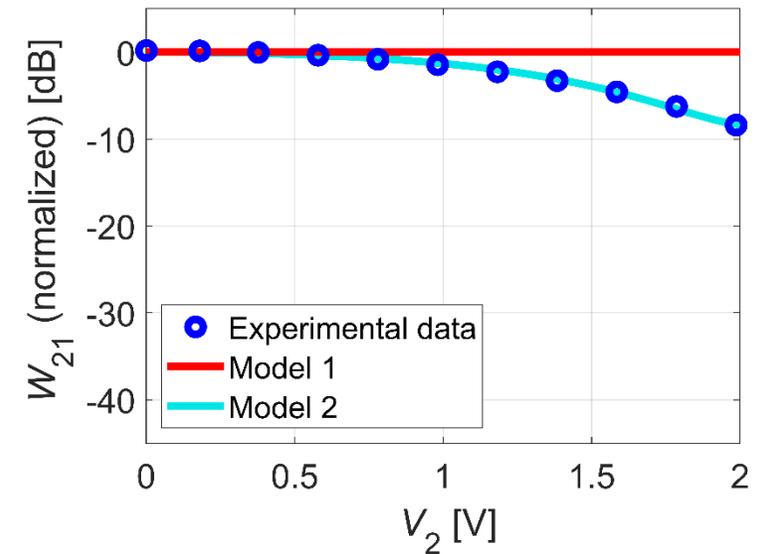
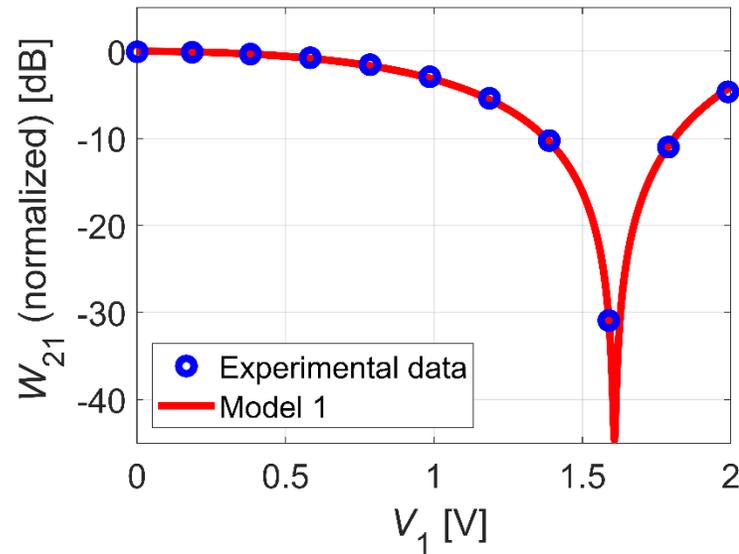


# MZI mesh model with crosstalk

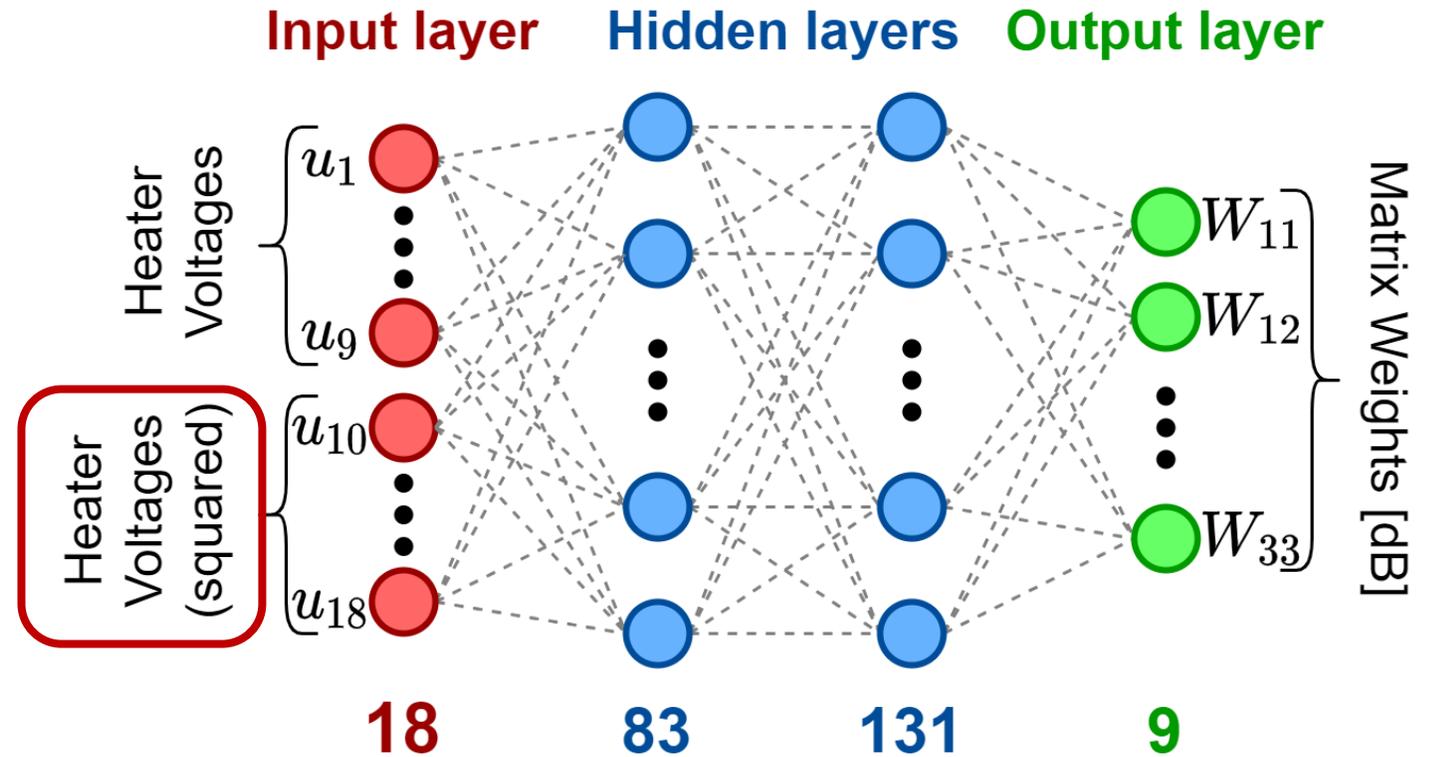
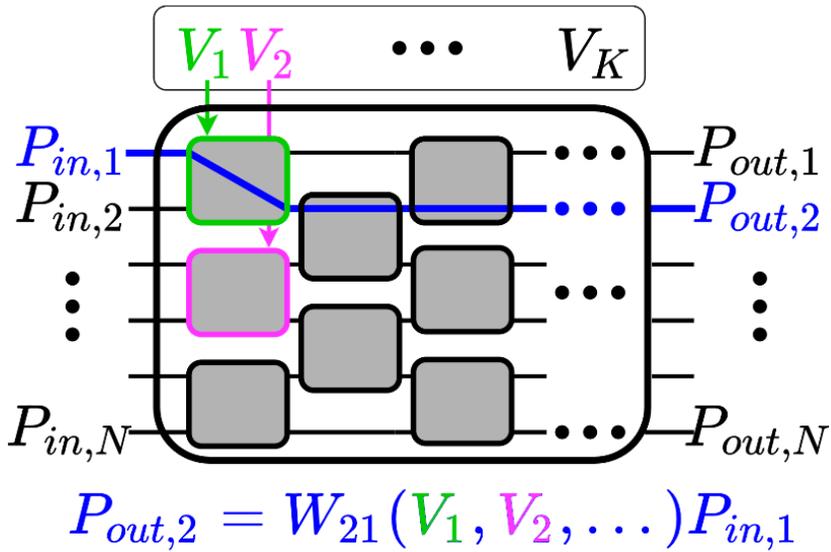


Self & crosstalk

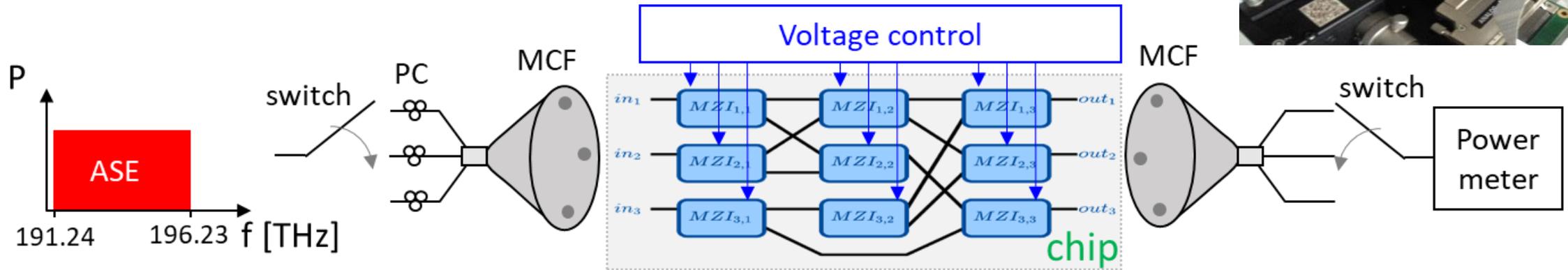
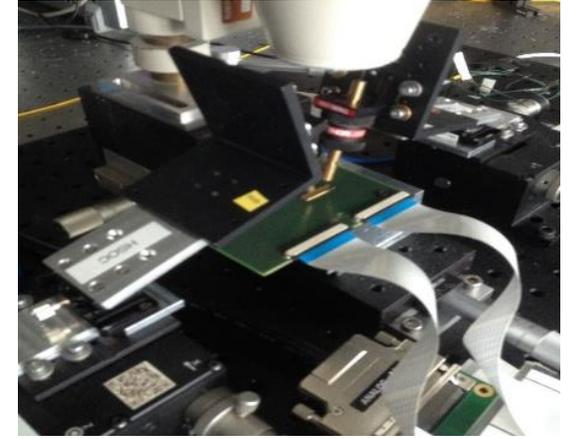
$$W_{ij} = L_{ij} \prod_{k \in K_{ij}} \frac{1}{4} \left| \frac{\sqrt{ER} - 1}{\sqrt{ER} + 1} - \exp \left( \sqrt{-1} \left( \phi_{ij,k}^{(0)} + \sum_{m=1}^M \phi_{ij,k,m}^{(2)} V_m^2 \right) \right) \right|^2$$



# Black-box NN MZI mesh model



# Experimental setup and 3x3 MZI mesh



Measurements:

1. Individually sweep one voltage  $[0, V_{\pi}]$
2. Randomly chosen voltages



Dataset = { Voltages | Transmission }

# Performance comparison

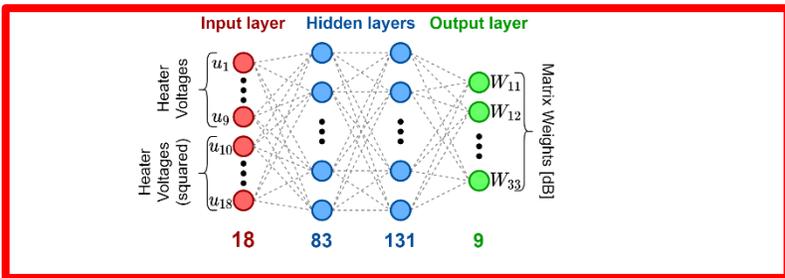
$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (w_i - \hat{w}_i)^2}$$

$$W_{ij} = L_{ij} \prod_{k \in K_{ij}} \frac{1}{4} \left| \frac{\sqrt{ER}-1}{\sqrt{ER}+1} - \exp\left(\sqrt{-1}(\phi_k^{(0)} + \phi_k^{(2)}V_k^2)\right) \right|^2$$

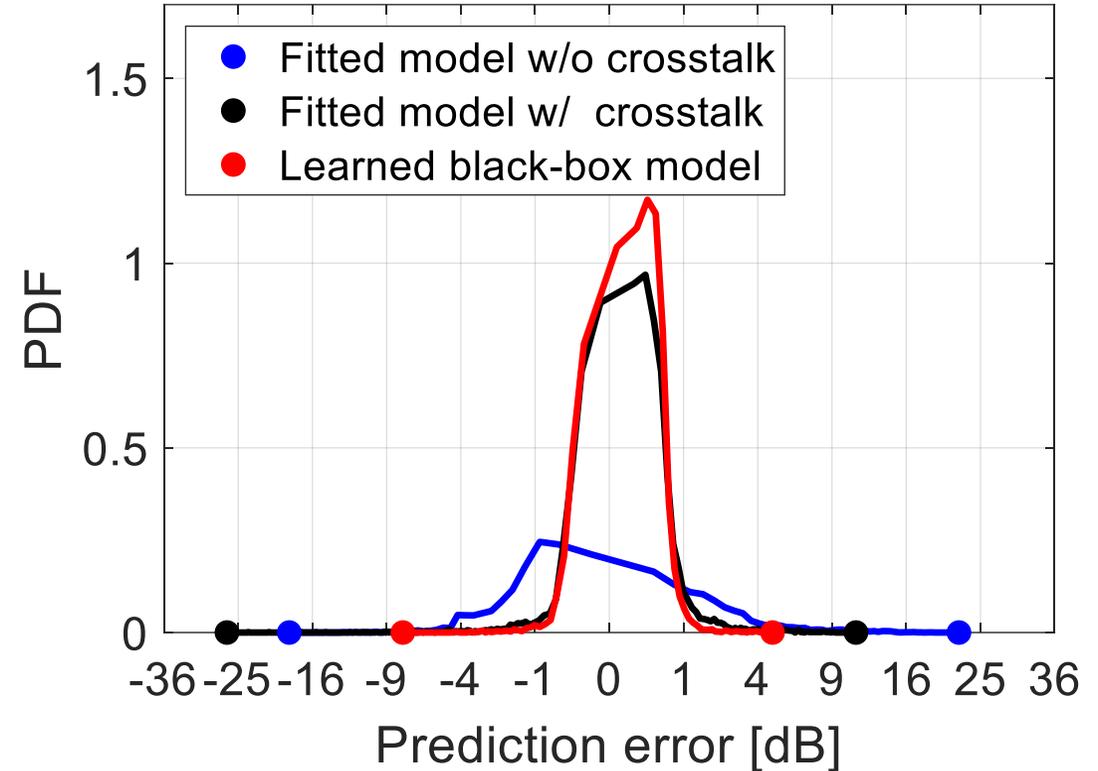
**RMSE = 3.26 dB**

$$W_{ij} = L_{ij} \prod_{k \in K_{ij}} \frac{1}{4} \left| \frac{\sqrt{ER}-1}{\sqrt{ER}+1} - \exp\left(\sqrt{-1}(\phi_{ij,k}^{(0)} + \sum_{m=1}^M \phi_{ij,k,m}^{(2)}V_m^2)\right) \right|^2$$

**RMSE = 1.44 dB**



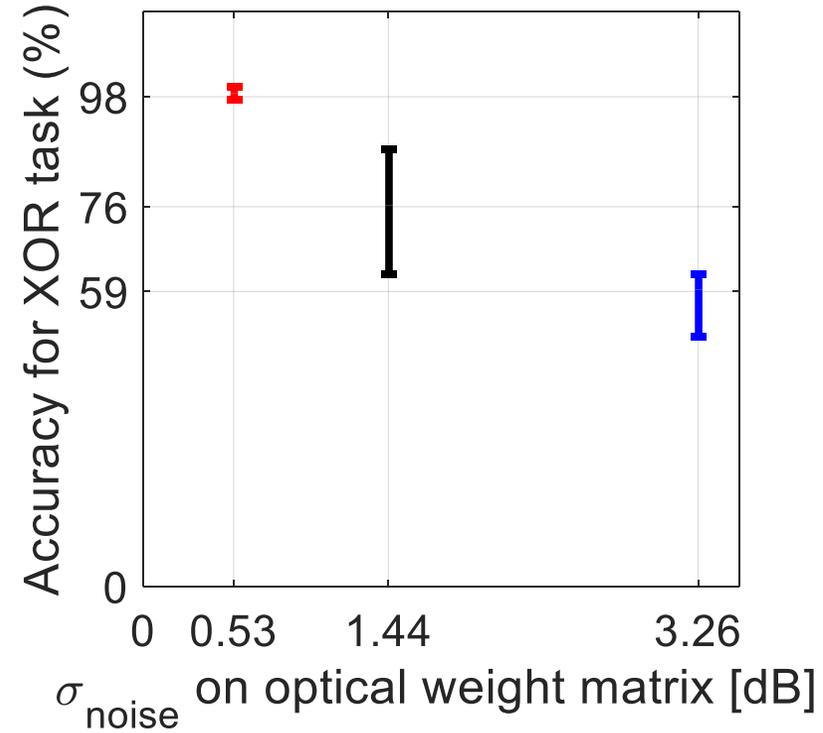
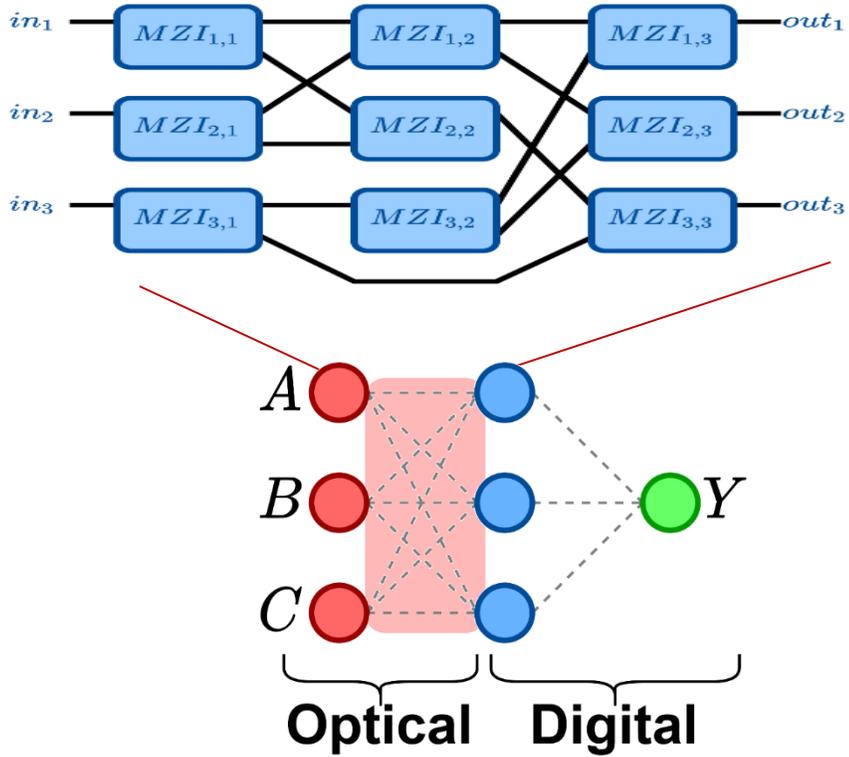
**RMSE = 0.53 dB**



Including thermal cross-talk improve performance but not as much as a black-box ML model.

# Testing on a 3-bit XOR task

A	B	C	Y
0	0	1	1
0	1	0	1
1	0	0	1
1	1	1	0
1	1	0	0
0	1	1	0
1	0	1	0
0	0	0	0



Training with inaccurate models leads to a significant performance penalty during inference.

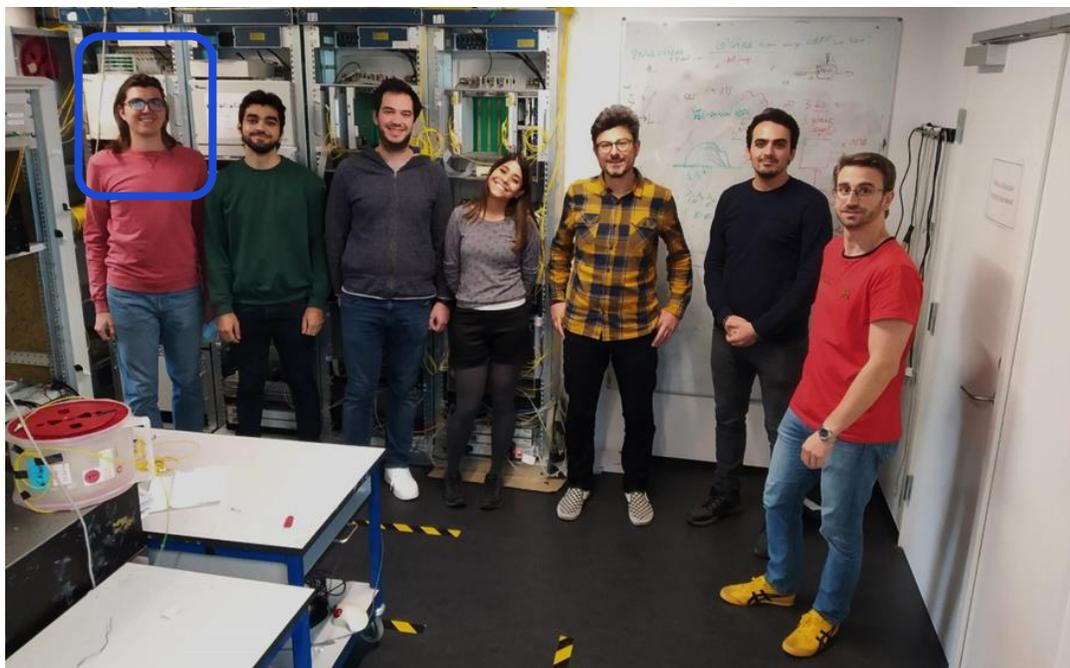
# Summary

- **ML techniques** provide significant benefit for optical communication
  - Learning **how to communicate** over nonlinear/dispersive channels
  - Enhancing performance through **equalization** (digital/optical/optoelectrical)
- Photonics provides an effective **hardware for ML**
  - Training is an **open challenge** with off-line training requiring **accurate models**
  - Model inaccuracies **directly impact the performance** of the tasks running over the photonic hardware

# Thank you for your attention



MENTOR  
{Machine LEarning in optical NeTwORks}



Open  
positions



## Acknowledgements

This work is supported by the Villum Foundations (VYI grant OPTIC-AI no.29344), the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreements No 766115 and 956713, and the European Research Council through the ERCCoG FRECOM project (grant No 771878).