Toward intelligence in photonic systems

Darko Zibar

Machine Learning in Photonic Systems (M-LiPS) group
DTU Fotonik, Department of photonics Engineering
Technical university of Denmark
Email: dazi@fotonik.dtu.dk

Acknowledgments

Technical University of Denmark, Denmark

Francesco Da Ros, Uiara C. de Moura, Metodi P. Yankov, Stenio M. Ranzini, Ognjen Jovanovic, Mehran Soltani, Ali Cem, Thyago M. Sá Pinto, Rasmus T. Jones, Simone Gaiarin , Giovanni Brajato

Politecnico di Torino, Italy

Andrea Carena, Ann M.R. Brusin

NICT, Japan

Ruben Luis, Georg Rademacher, Ben J. Puttnam

Chalmers University of Technology, Sweden

Magnus Karlsson, Victor Torres-Company

University of Santa Barbara, USA

John E. Bowers

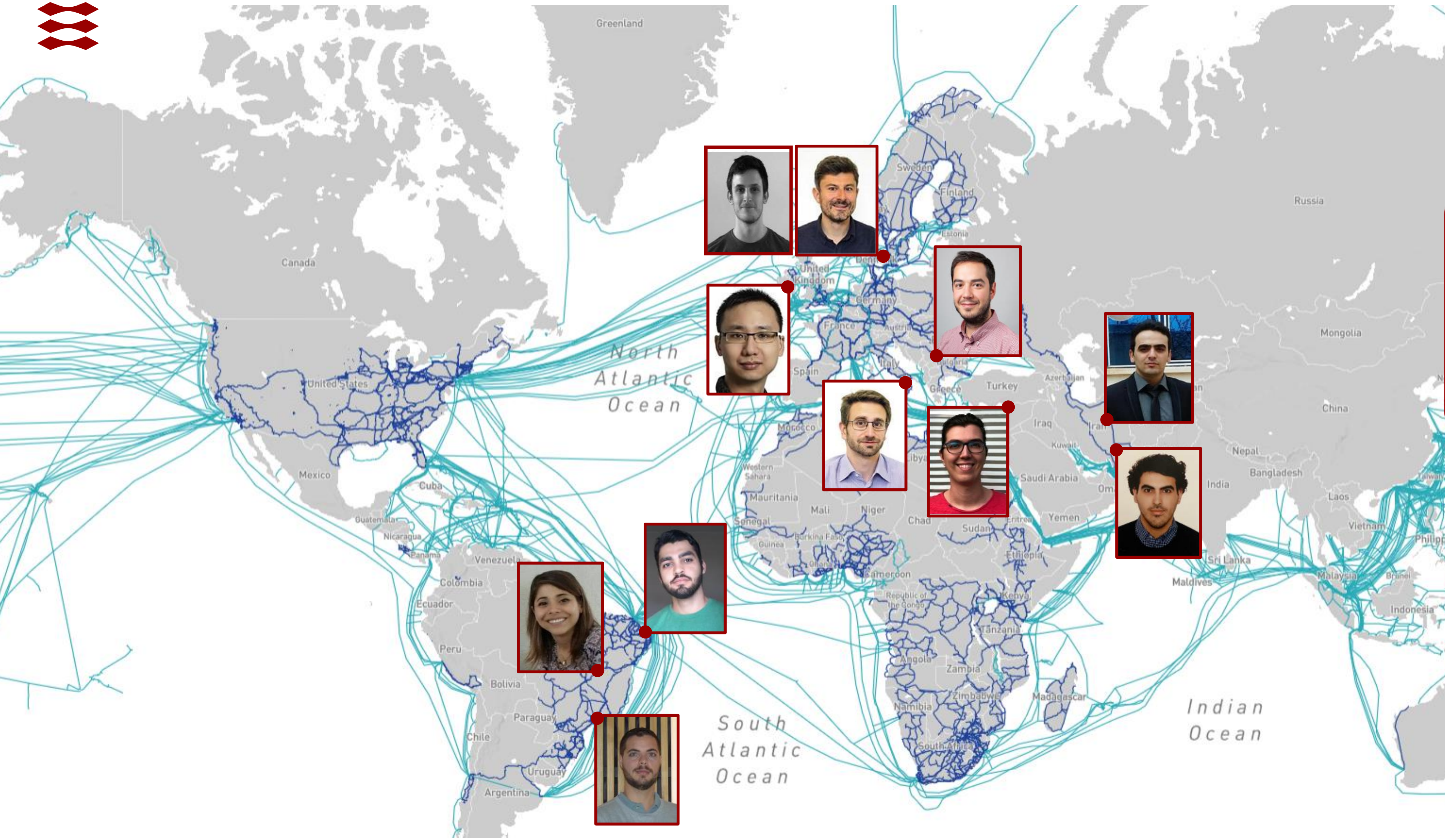
Nokia Bell Labs, Germany

Henning Bülow, Vahid Aref, Roman Dischler

Aston University, UK

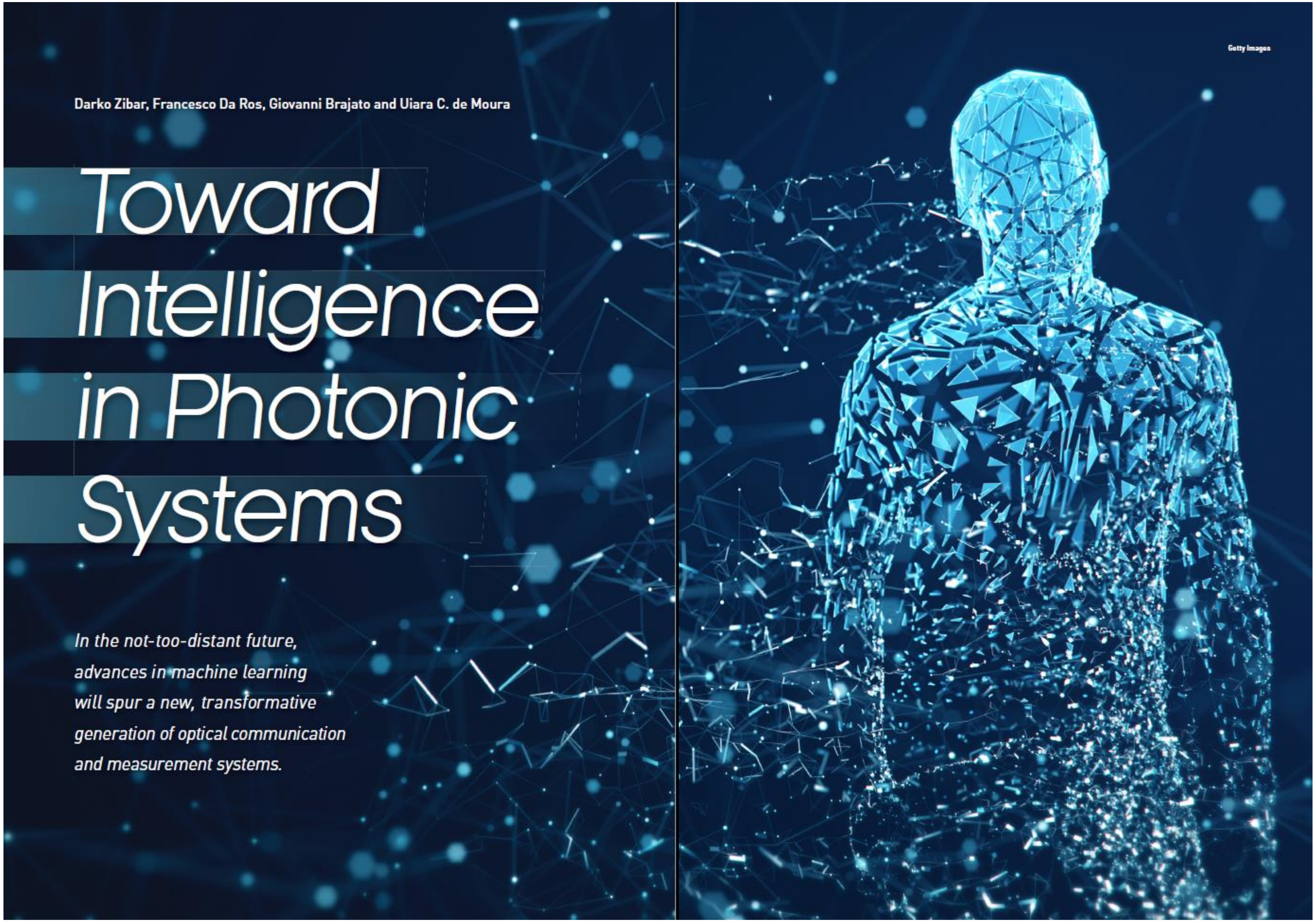
Sergei Turitsyn, Wladek Forysiak, Md Asif Iqbal, Morteza Kamalian, Lukasz Krzczanowicz

Machine Learning in Photonic Systems Group



11 Group members
8 Nationalities
1 Professor
1 Senior researcher
2 postdocs
7 PhD students

DTU


Getty Images

Darko Zibar, Francesco Da Ros, Giovanni Brajato and Uiara C. de Moura

Toward Intelligence in Photonic Systems

In the not-too-distant future, advances in machine learning will spur a new, transformative generation of optical communication and measurement systems.

Where does machine learning excel?

- Learning complex **direct** mappings:

$$\mathrm{X} \longrightarrow \boxed{f(\cdot)} \longrightarrow \mathrm{Y}$$

- Learning complex **inverse** mappings:

$$\mathrm{Y} \longrightarrow \boxed{f^{-1}(\cdot)} \longrightarrow \mathrm{X}$$

- Learning **decision rules** for complex mappings:

$$\mathrm{X} \longrightarrow \boxed{f(\cdot)} \longrightarrow P(Y = 1|X)$$

Use neural networks to learn $f(\cdot)$ and $f^{-1}(\cdot)$

Problems that could benefit from ML

- Communication over the nonlinear fiber-optic channel:
  - Channel highly complex
  - Capacity **unknown**?
  - Optimum receiver architecture **unknown**
  - Optimum modulation and pulse-shapes **unknown**

- Optical amplifiers for multiband-wavelength and SDM systems:
  - Complex relation between pumps and gain
  - Optimization of pump powers and wavelengths for target gain profiles
  - Optimization of pump powers and wavelengths for target mode dependent

- Design of optical components (inverse system design):
  - Given laser bandwidth and noise find the physical parameters
  - Given modulator BW find the physical parameters
  - Instead of running time-consuming simulation build fast ML based models

- Noise characterization of lasers and frequency combs:
  - Amplitude and phase tracking at the quantum limit
  - Extraction of noise correlation matrices, e.e amplitude, phase, amplitude-phase
  - Macroscopic comb parameters, i.e. timing jitter, amplitude jitter, carrier envelope offset

Research topics and collaborations

Machine learning enabled ultra-wideband amplifier design

Unifying framework for lasers and frequency combs noise characterization

Machine learning techniques for communicate over complex channels

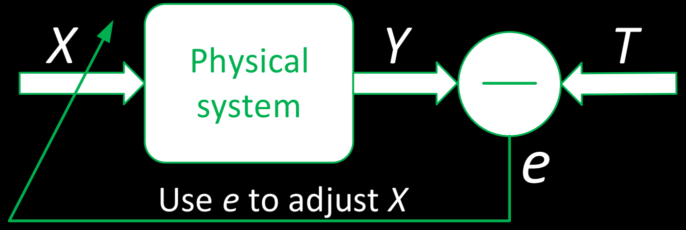
Optical technologies to accelerate AI

Quantum phase tracking and communication

Highly-sensitive fiber based sensing systems

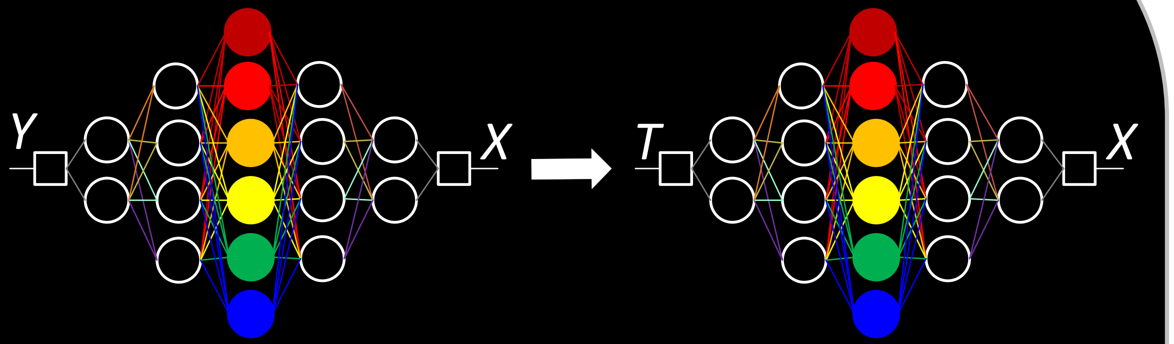
Inverse system learning

#1 Problem statement:

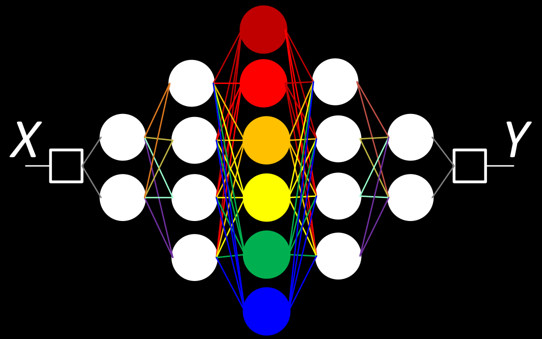


A physical system describing relation between input *X* and output *Y* is given. The objective is to determine input *X* that would result in a targeted output *T*.

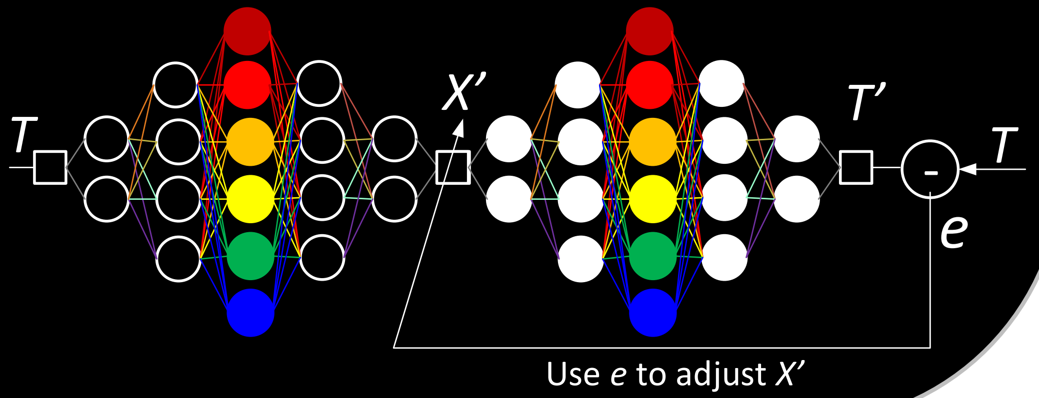
#2 Train neural network to learn *inverse* mapping (from *X* to *Y*):



#3 Train neural network to learn *forward* mapping (from *X* to *Y*):



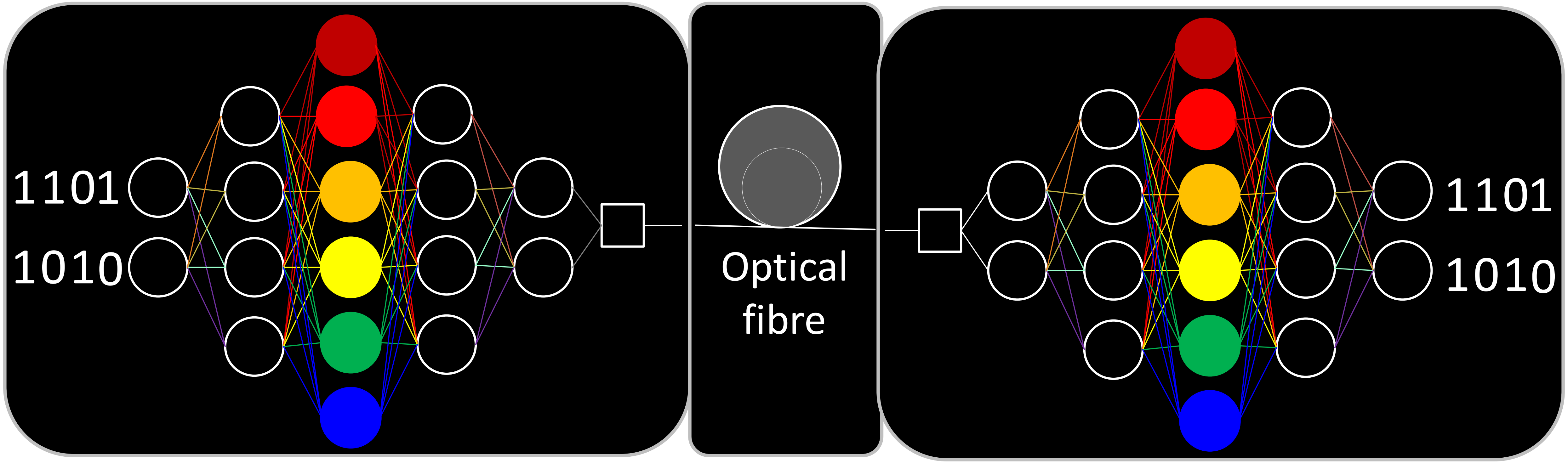
#4 Perform *final* optimization:



[1] D. Zibar et al., "Inverse system design using machine learning: the Raman amplifier case," Journal of Lightwave technology, 2019

[2] U. C. de Moura et al., "Multi-band programmable Raman amplifier," Journal of Lightwave technology, 2020

Learning to communicate over complex channels



[1] R. Jones, et al., ECOC 2018
[2] R. Jones, et al., ECOC 2019
[3] S. Gaiarin et al., JLT, 2020
[4] O. Jovanovic et al., et al, JLT 2021

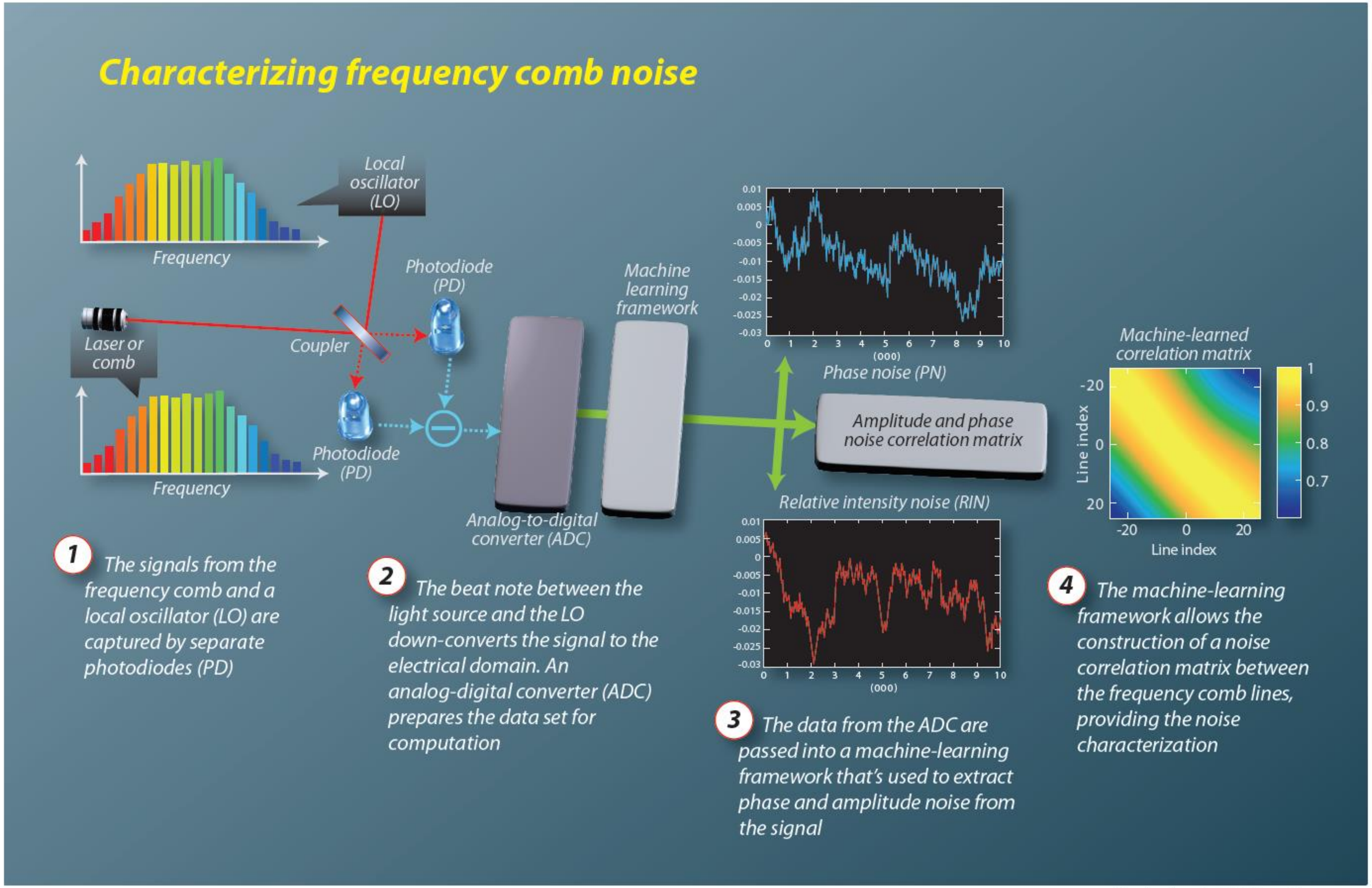
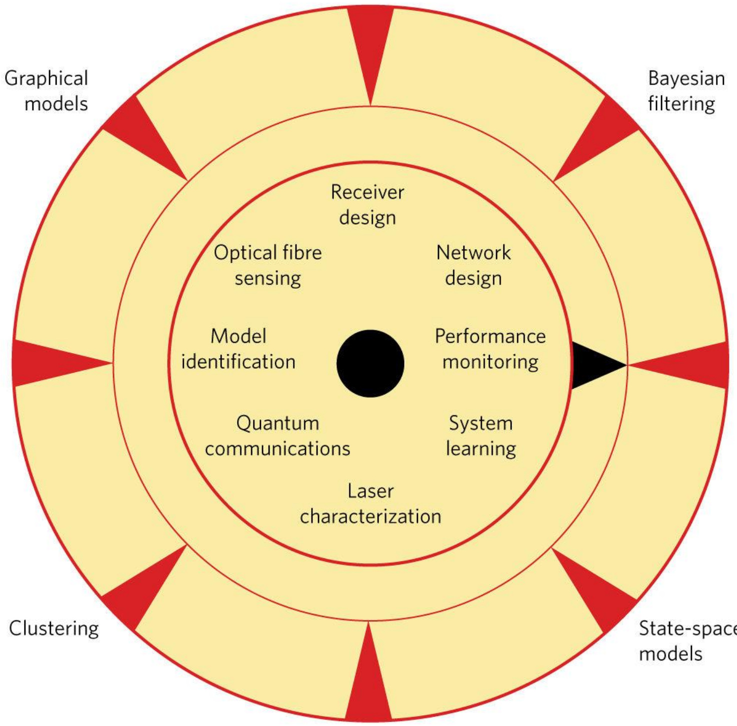


Illustration by Phil Saunders

[1] D. Zibar et al., PTL 2019

[2] G. Brajatto et al, Optics Express, 2020

State-of-the-art



D. Zibar et al., Nature Photonics, (11) 749-751, 2017

New topics anno 2020-2021:

- Photonic reservoir computing[1]
- Optical amplifier and laser design[2-3]
- End-to-end learning[4]
- Back-propagation learning[5]
- Optical network optimization[6-7]
- Frequency comb noise characetrization[8]
- Photonic component design

[1] S. Ranzini, "Tuneable optoelectronic…" JSTQE 2020

[2] D. Zibar, "Inverse system design…," JLT 2019

[3] Z. Ma, "Paremeter extraction and inverse," Optics Express 2020

[3] Karanov, "End-to-end deep learning…", JLT 2018

[4] C. Hager, "Revisiting multi-step…," ECOC 19

[5] F. Musemechi, "An overview on…," IEEE Comm. survey, 2019

[6] F. N. Khan, "An optical communication persp…," JLT 2019

[7] G. Brajato, "Bayesian filtering…," Optics Express, 2020

[8] U. C. de Moura, "Multi-band optical program. Amplifier," JLT 2020

[9] K. Kojima, "Inverse Design of Nanophotonic Devices…", OFC 2020

[10] G. Genty, "Machine learning in ultrafast photonics," Nat. Phot., 2020

Will machine learning be a *game changer*?

Research Highlights (2019-2021)

- **Record -sensitive and -accurate optical phase measurement[1,2] (quantum limited operation)**
  – Identification of fundamental laser linewidth
  – Identification of frequency comb noise sources
  – Optimum phase measurement in the presence of amplifier noise

- **Machine learning enabled ultra-wideband Raman amplifiers[3,4,5,6]**
  – Arbitrary gain profiles in S-C-L band
  – *Gain and power profile shaping in distance and frequency*
  – Noise figure prediction of Raman amplifiers

- **Learning optimum transmitter and receivers architectures[7,8,9]**
  – Channel tailored constellation
  – SNR and linewidth robust constellation
  – Equalization of IM/DD using reservoir computing

1. *D. Zibar et al., Optica, 2021*
2. *G. Brajato et al. Opt. Express, 2020*
3. *D. Zibar, J. Ligtwave Technol., 2020 (top cited JLT paper in 2020)*
4. *M. Soltani, Optics Letters, 2021*
5. *U. de Moura, J. Lightwave Technol., 2020*
6. *U. de Moura, Optics Letters, 2021*
7. *R. Jones et al, ECOC 2019*
8. *O. Jovanovic et al., sub to JLT, 2021*
9. *F. Da Ros, IEEE J. Select. Topics Quant. El. 2020*

Challenges to be addressed

- Fields focuses on the experimental demonstrations
- ML benefits on experimental data should be ideally shown
- Noise in experimental set-ups (non Gaussian, non additive)
- Experimental-set ups are prone drifts and fluctuations
- Automatizing experimental-set ups for training data acquisition (noise, drift,)
- Training of NNs using gradients computation - challenging in experimental environments
- Deep understanding of statistics, linear algebra, optimization and experimental set-up debugging necessary not to end in pitfalls

Application of multi-layer neural networks for design of Raman amplifiers

[1] D. Zibar, A. M. Rosa Brusin, U. C. de Moura, F. Da Ros, V. Curri, Andrea Carena, "Inverse system design using machine learning: The Raman amplifier case," Journal of Lightwave Technology, vol. 38, no. 4, 2020

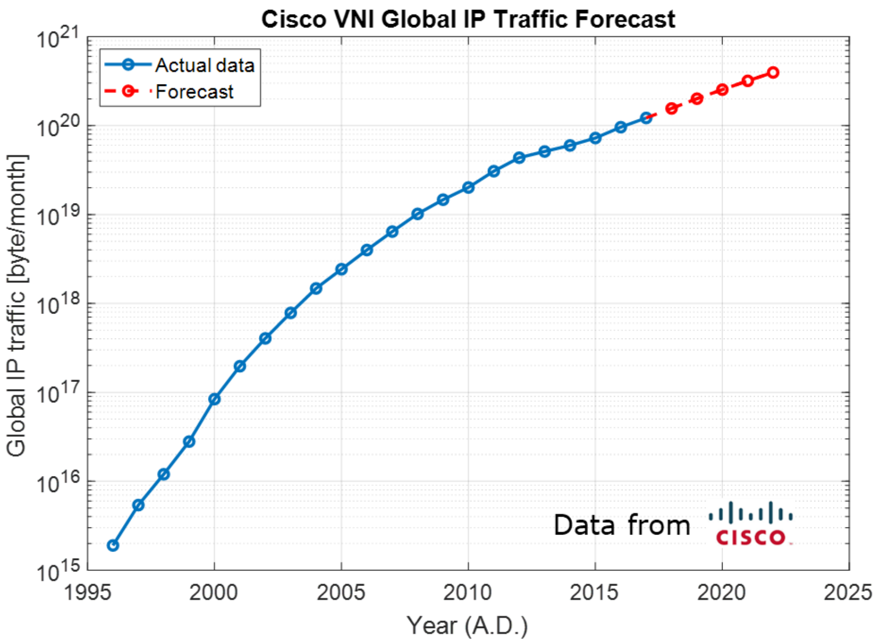
[2] M. Soltani, F. Da Ros, A. Carena, D. Zibar, "Inverse design of a Raman amplifier in frequency and distance domains using convolutional neural networks," Optics Letters, vol. 46, no. 11, 2021

[3] A. M. Rosa Brusin, V. Curri, D. Zibar, and A. Carena, "An ultrafast method for gain and noise prediction of Raman amplifiers," in proceedings of European Conference on Optical Communication, ECOC, 2019

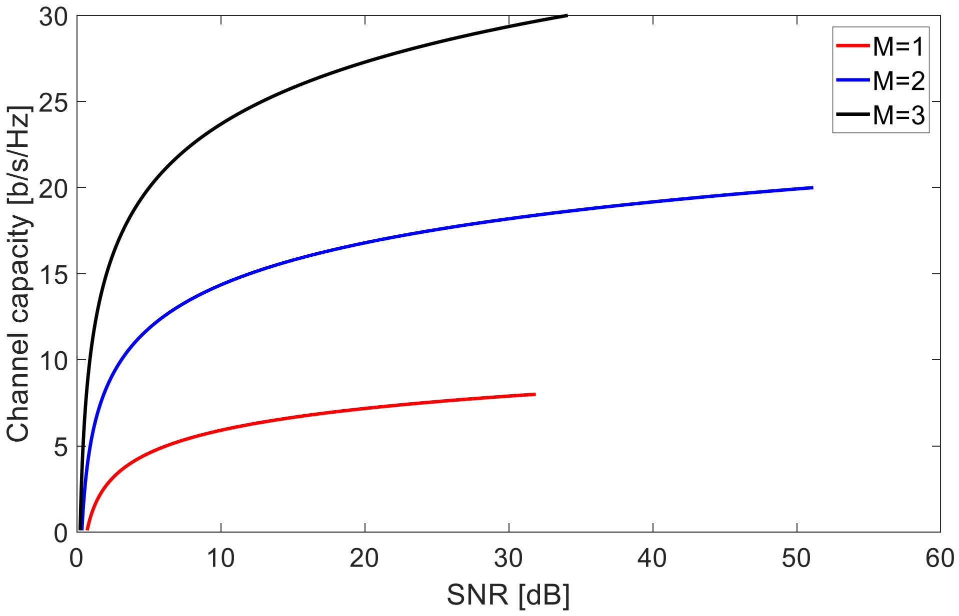
[4] U. C. de Moura, F. Da Ros, A. M. Rosa Brusin, A. Carena, and D. Zibar, "Experimental demonstration of arbitrary Raman gain–profile designs using machine learning, " in Optical Fiber Communication Conference (OFC) 2019, OSA Technical Digest (Optical Society of America), 2020

[5] U. C. de Moura, Md A. Iqbal, M. Kamalian, L. Krzczanowicz, F. Da Ros, A. M. Rosa Brusin, A. Carena, W. Forysiak, S. Turitsyn and D. Zibar, "Multi–band programmable gain Raman amplifier," Journal of Lightwave Technology, 2020

Increasing the bandwidth of optical systems



(a) Future data projection



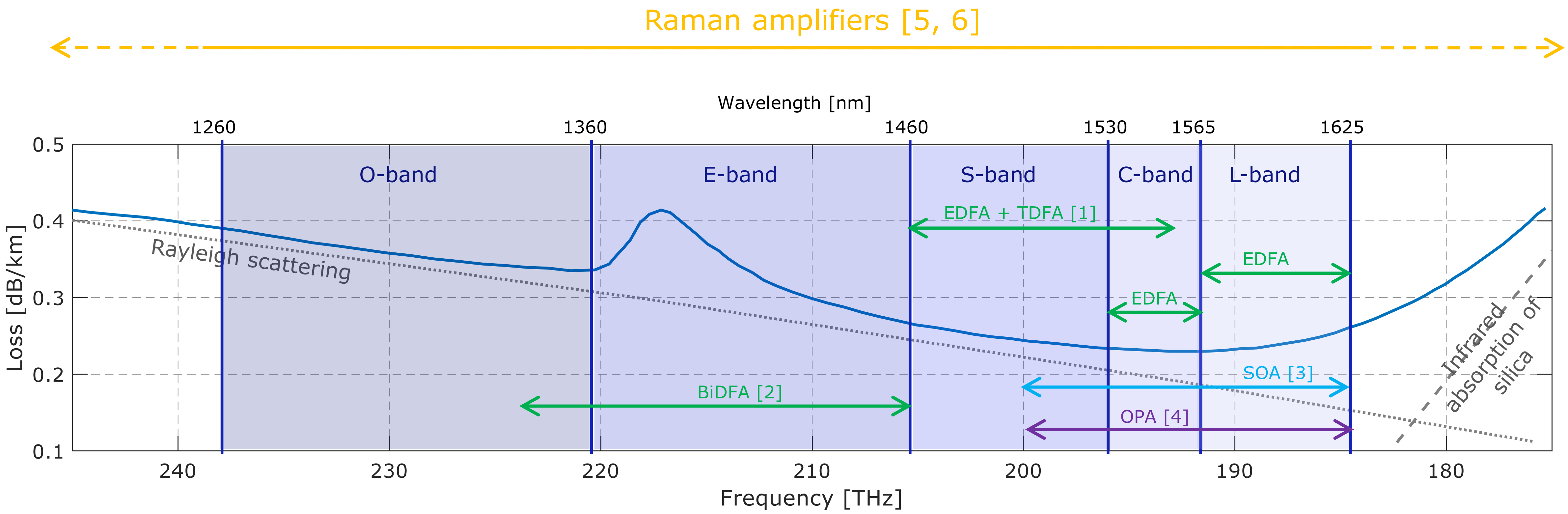
(b) Channel capacity

$$\frac{C}{B} = M \log\left(1 + \frac{E_b}{N_0}\frac{C}{B}\right)$$ [b/s/Hz]

C :capacity
M :spatial paths
B :bandwidth
$\frac{E_b}{N_0}$:signal-to-noise ratio

Significantly higher gains by increasing spatial paths than SNR

Ultra-wideband optical amplification



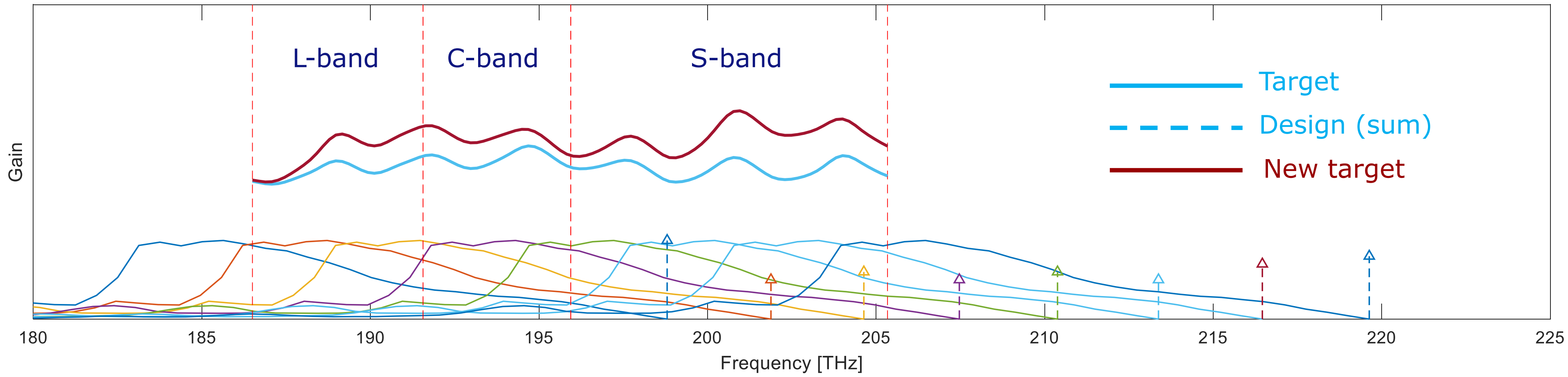
xDFA: doped fiber amplifier

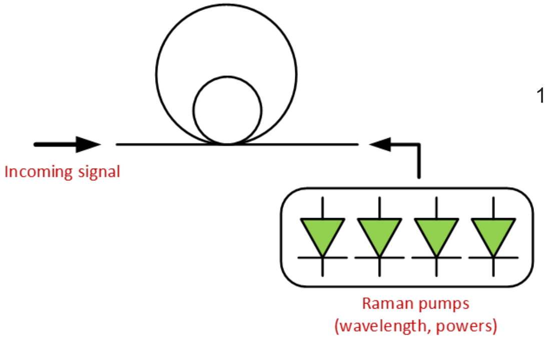
SOA: semiconductor optical amplifier

OPA: optical parametric amplifier

[1] T. Sakamoto, JLT, vol. 24, no. 6, 2006
[2] Y. Wang, OFC 2020, Th4B.1
[3] J. Renaudier, ECOC, 2018
[4] T. Kobayashi, OFC 2020, Th4C.7
[5] J. Chen, IEEE Photonics Journal, vol. 10, 2018
[6] M. A. Iqbal, OFC 2020, W3E.4

Arbitrary gain Raman amplifiers

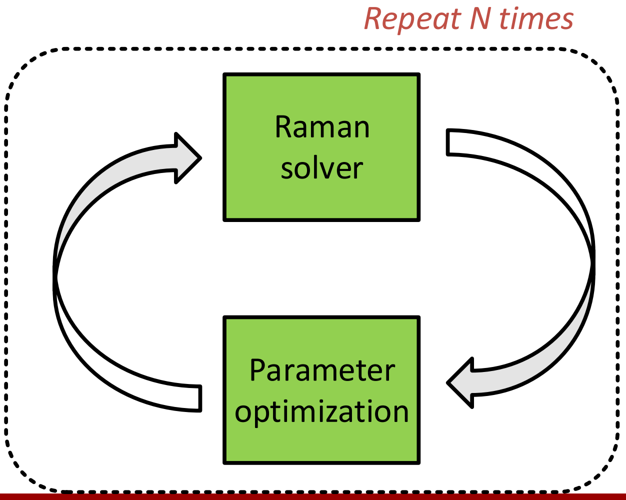




$$\frac{dP_s}{dz} = -\alpha_s P_s + C_R(\lambda_s,\lambda_p)\left[P_p^+ + P_p^-\right] P_s \qquad (1)$$

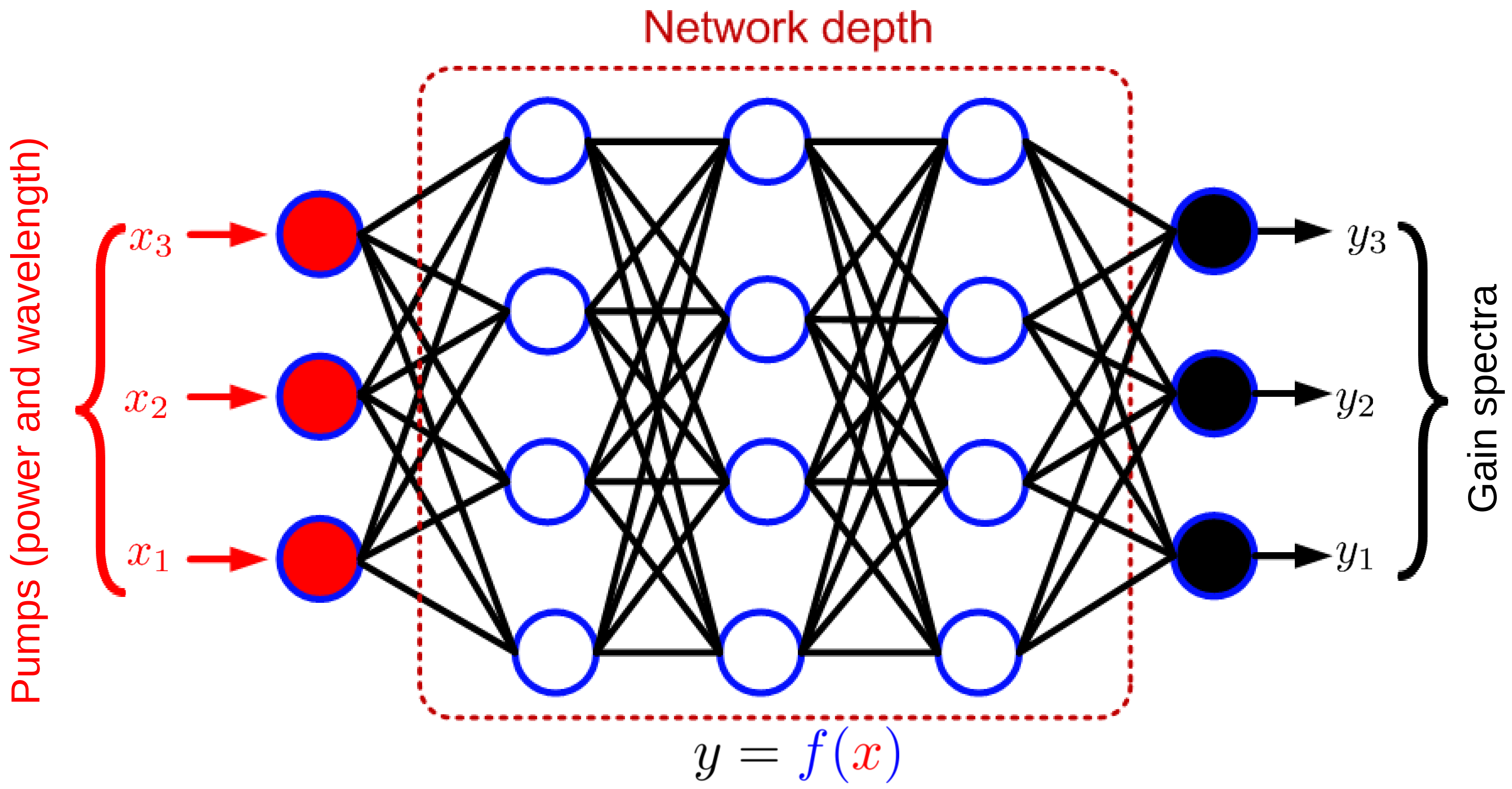
$$\pm\frac{dP_p^\pm}{dz} = -\alpha_p P_p^\pm - \left(\frac{\lambda_s}{\lambda_p}\right) C_R(\lambda_s,\lambda_p)\, P_s\, P_p^\pm \qquad (2)$$

$$\pm\frac{dP_A^\pm}{dz} = -\alpha_A P_A^\pm + C_R(\lambda_A,\lambda_p) P_p P_A^\pm + C_R(\lambda_A,\lambda_p)[1+\eta(T)] h\nu_A B_{ref} P_p \qquad (3)$$



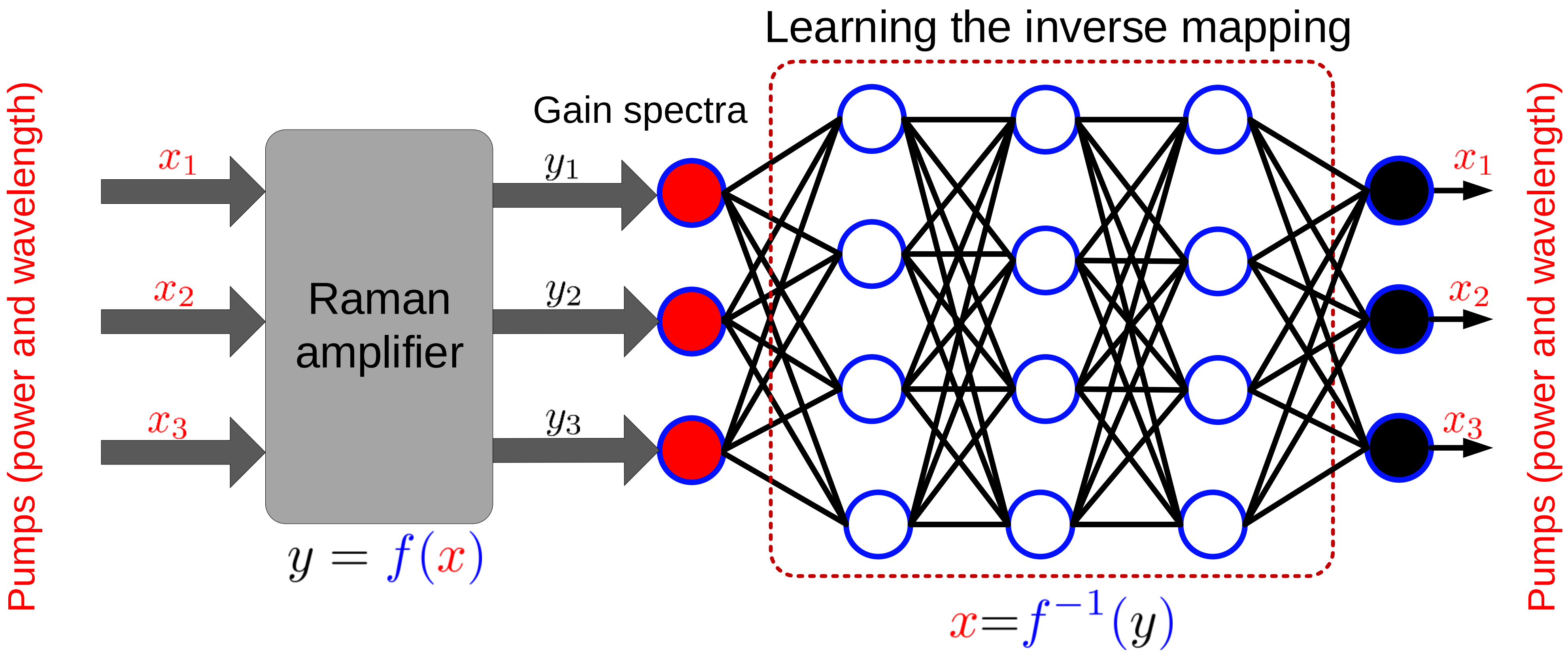
- High complexity due to Raman solver
- Long convergence time
- Restart optimization for new gain profile
- Rely on evolutional algorithms

Approximating Raman amplifier with NN



Neural network learns forward mapping, $f(\cdot)$, using training data and perform predictions for *new input* data: $y_{new} = f(x_{new})$

Learning inverse mapping



Learning the inverse mapping allows for designing arbitrary gain profile

Building the model from the data

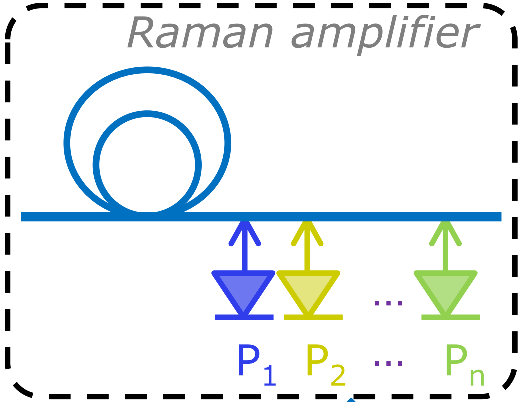
Given N pumps generate M gain profiles

$\lambda_1 \sim U[\lambda_{1,min};\ \lambda_{1,max}]$ nm
$\lambda_2 \sim U[\lambda_{2,min};\ \lambda_{2,max}]$ nm
…
$\lambda_N \sim U[\lambda_{n,min};\ \lambda_{n,max}]$ nm

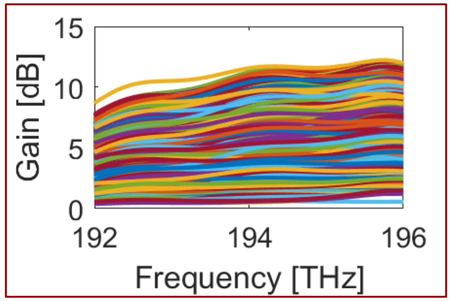
$P_1 \sim U[P_{1,min};\ P_{1,max}]$ W
$P_2 \sim U[P_{2,min};\ P_{2,max}]$ W
…
$P_N \sim U[P_{n,min};\ P_{n,max}]$ W

Numerically

Experimentally



M gain profiles each with K points

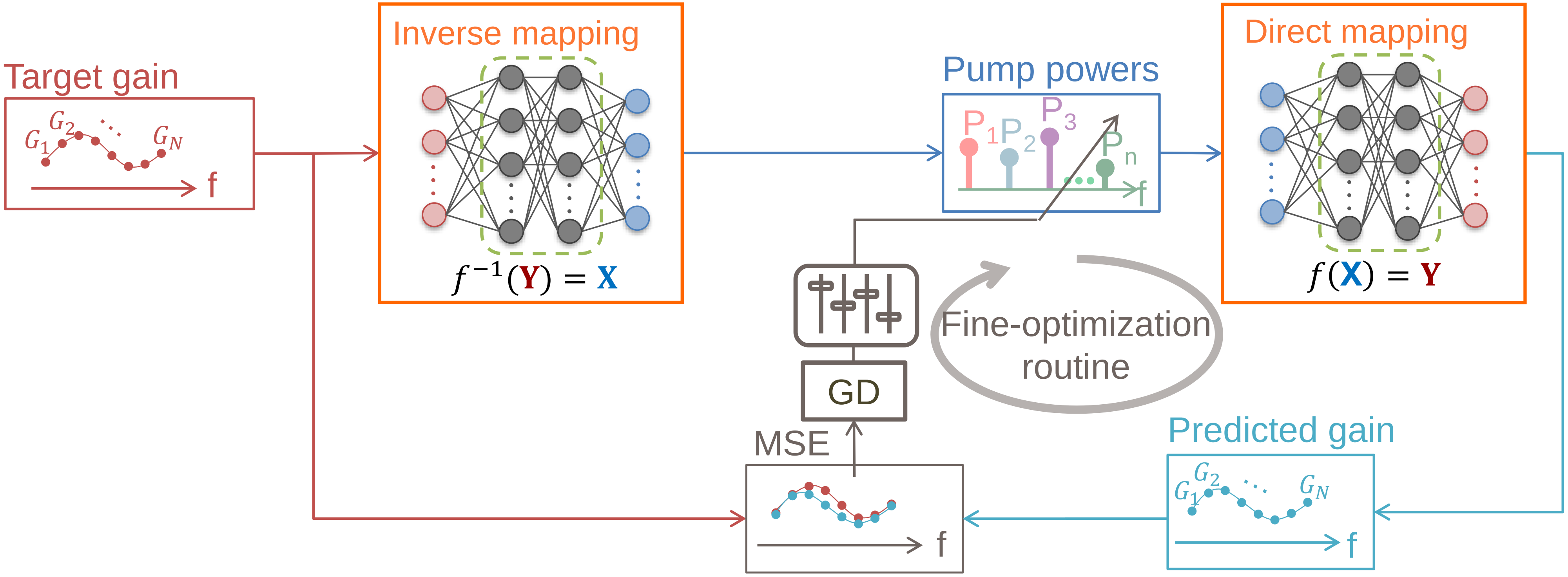


Data-set $\Rightarrow$ $\mathcal{D} = \{(\lambda_1^i, \lambda_2^i, \dots, \lambda_N^i, P_1^i, P_2^i, \dots, P_N^i, G_1^i, G_2^i, \dots, G_K^i) | i = 1, \dots, M\}$

Training

Validation

The machine learning framework



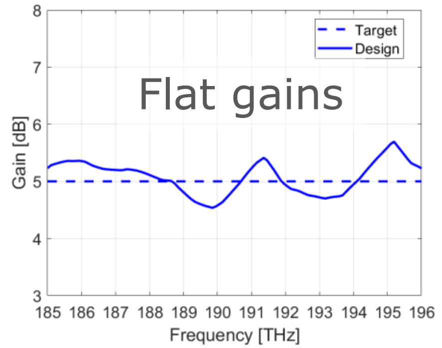
MSE: mean squared error

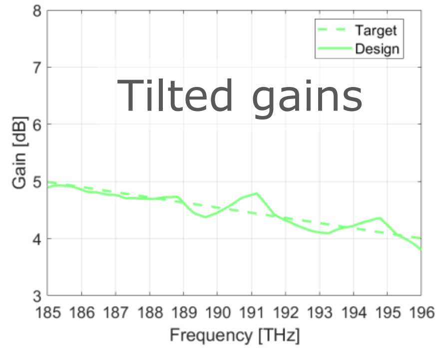
GD: gradient descent

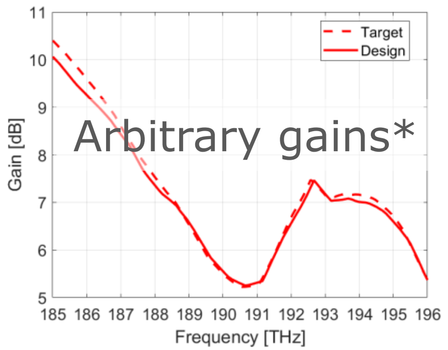
D. Zibar, J. Lightwave Technol. **38**(4), 736–753 (2019)

U. de Moura, J. Lightwave Technol. **39**(4),1162–1170 (2021)

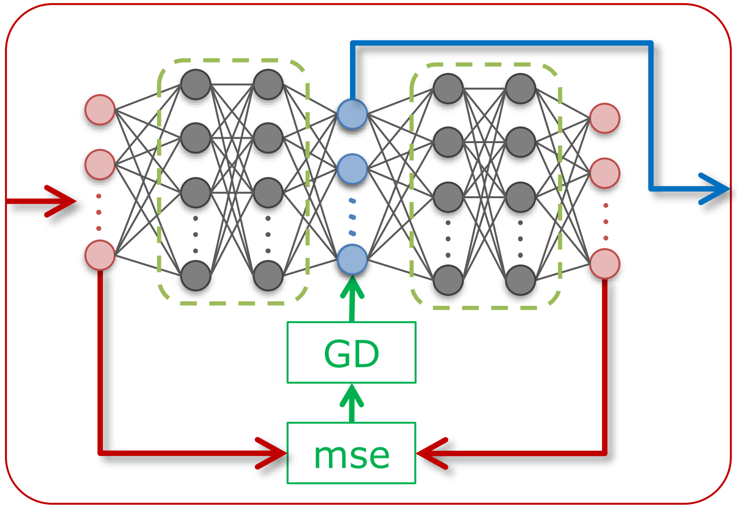
Experimental validation of the learned model







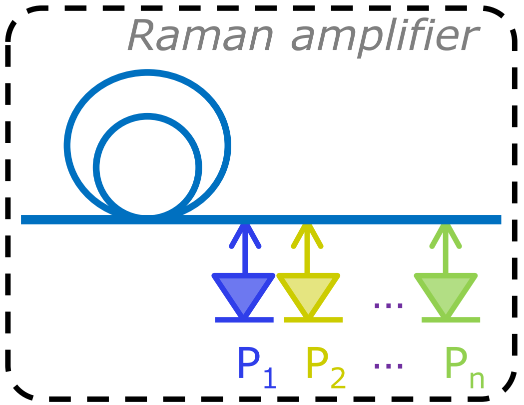
Target gains



Pump powers

P (mW)
2.87
52.57
84.94
0.00

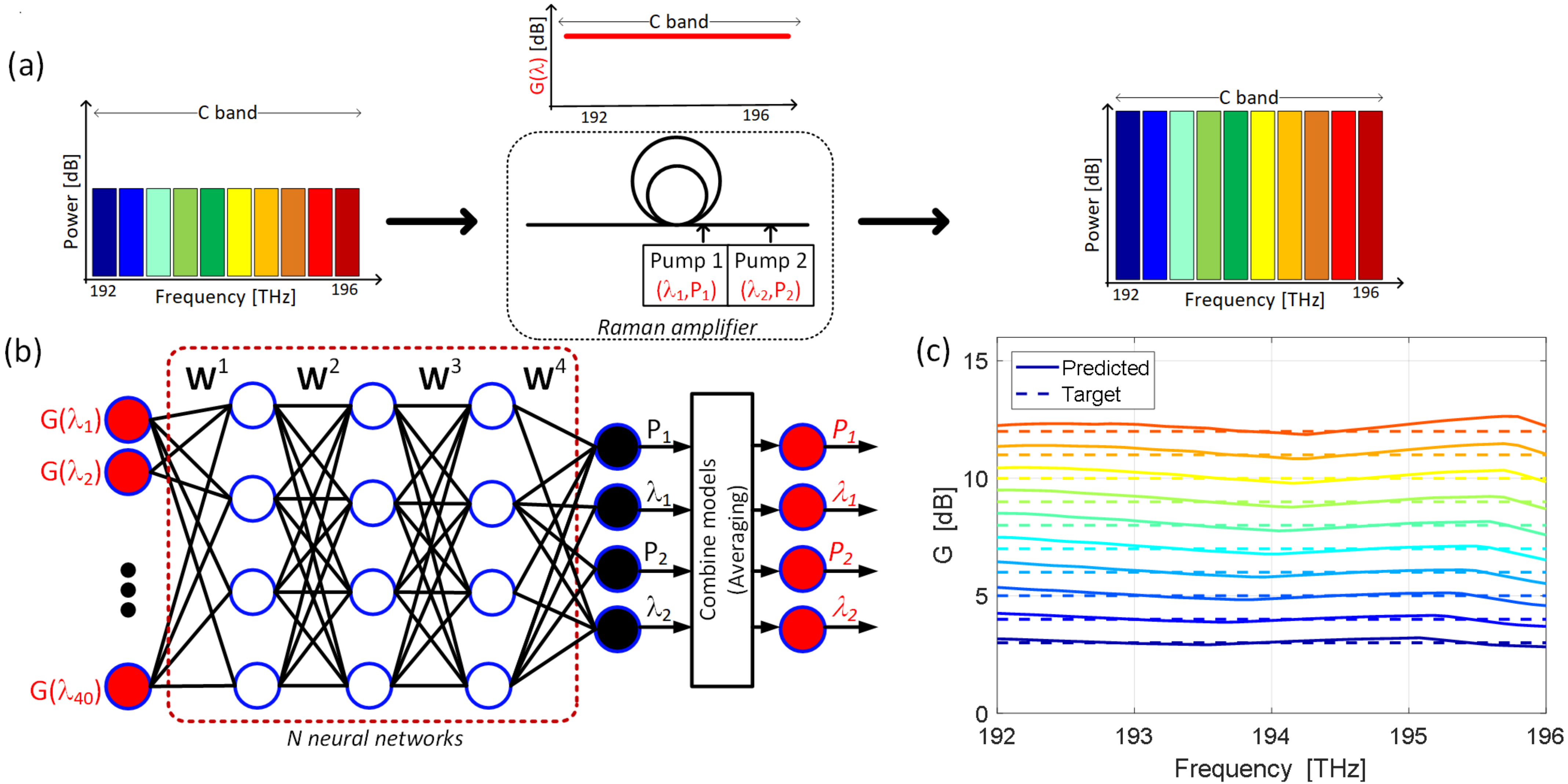
Numerical/Experiment



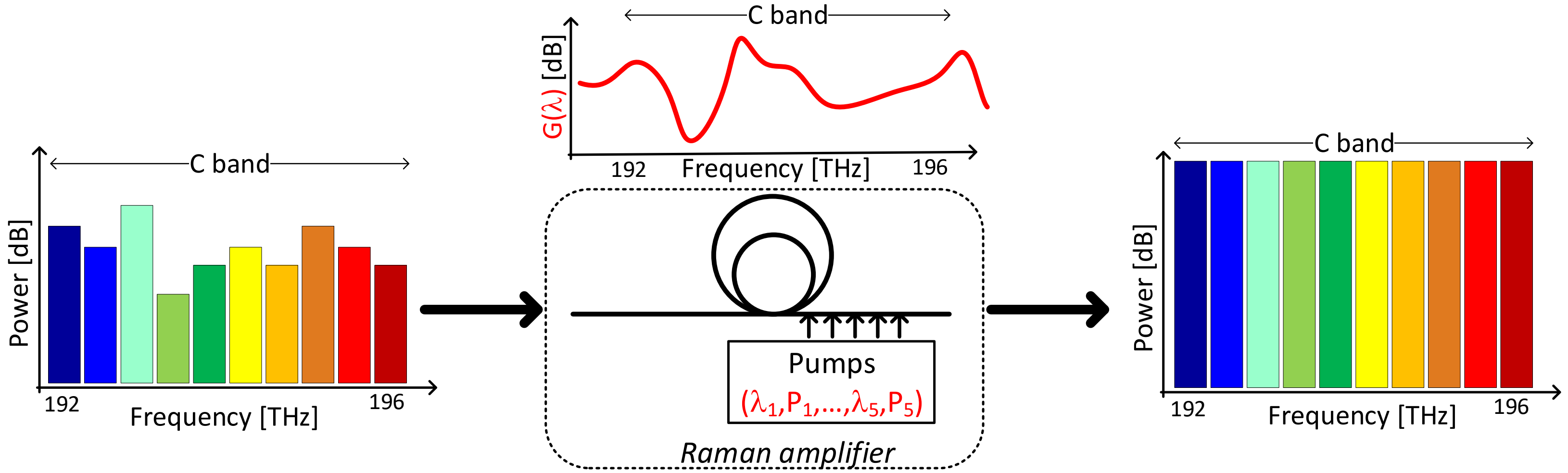
Measured gains

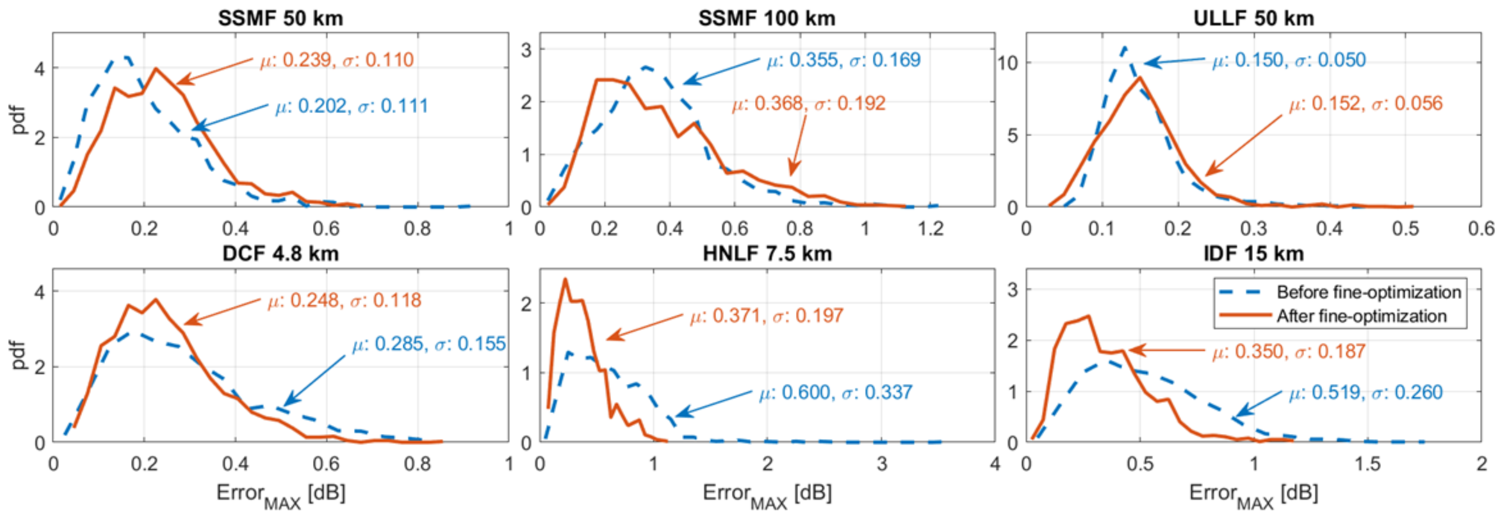
* Part from the acquired data not used on training

Flat gain profile design (C band)

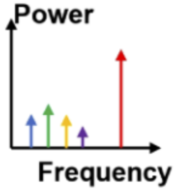


Arbitrary gain profile design (C band)





Arbitrary distance and gain profile

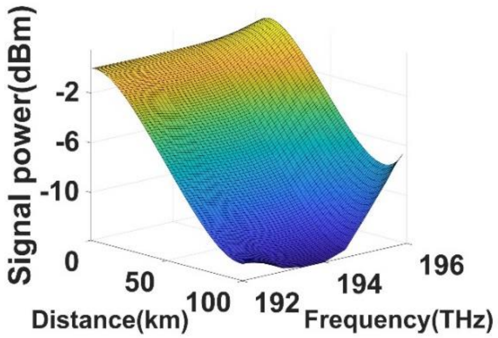


$\boldsymbol{P_{pump}} = [P_1, \ldots, P_N]$

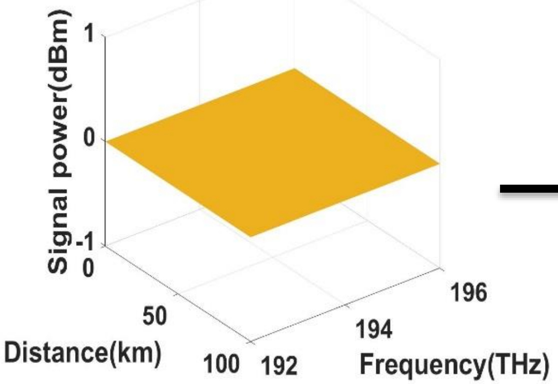
$\boldsymbol{\lambda_{pump}} = [\lambda_1, \ldots, \lambda_N]$

Raman amplifier

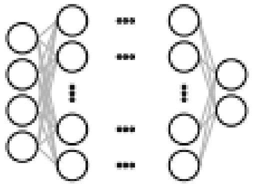
Signal power profile



Desired power profile (e.g. Flat input)



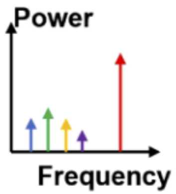
Learn the inverse model



Pump setup

$\boldsymbol{P_{pump}} = [P_1, \ldots, P_N]$

$\boldsymbol{\lambda_{pump}} = [\lambda_1, \ldots, \lambda_N]$



Power profile and gain shaping

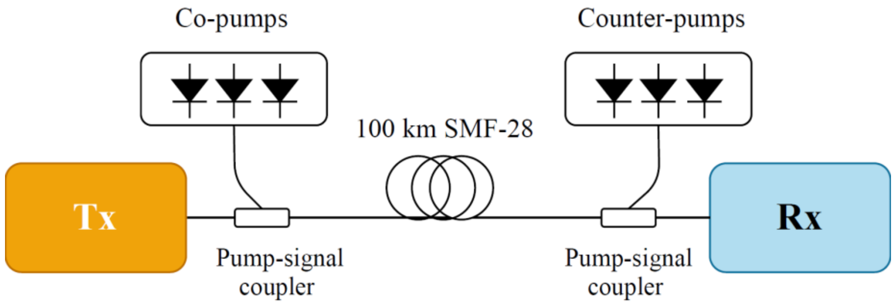
Quasi-lossless transmission with uniform distribution of power resulting in:

- Minimizing the amplified spontaneous emission (ASE) noise level
- Requirement for Nonlinear Fourier Transform (NFT) - NFT assumes lossless transmission

Symmetric power distribution:

- A requirement for nonlinearity mitigation using optical phase conjugation (OPC)

Bi-directional Raman amplifier



Direct model (Raman Solver)
Computationally demanding

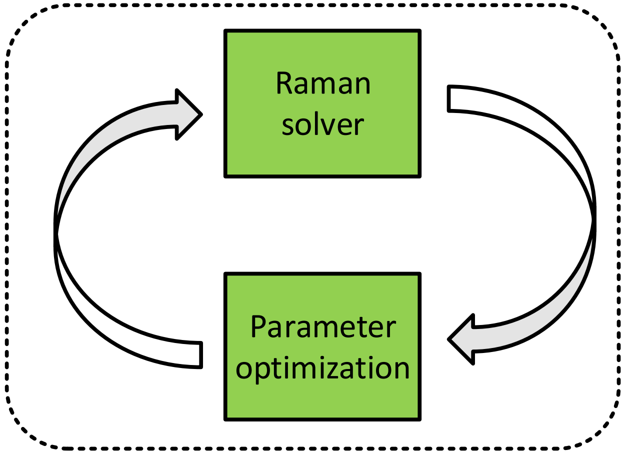
$$\frac{dP_s}{dz} = g_R P_p P_s - \alpha_s P_s,$$

$$\pm\frac{dP_p}{dz} = -\frac{\omega_p}{\omega_s} g_R P_p P_s - \alpha_p P_p,$$

Inverse model

No closed-form solution!

Repeat N times



- Long convergence time
- Restart optimization for new gain profile
- Usually based on evolutional algorithms

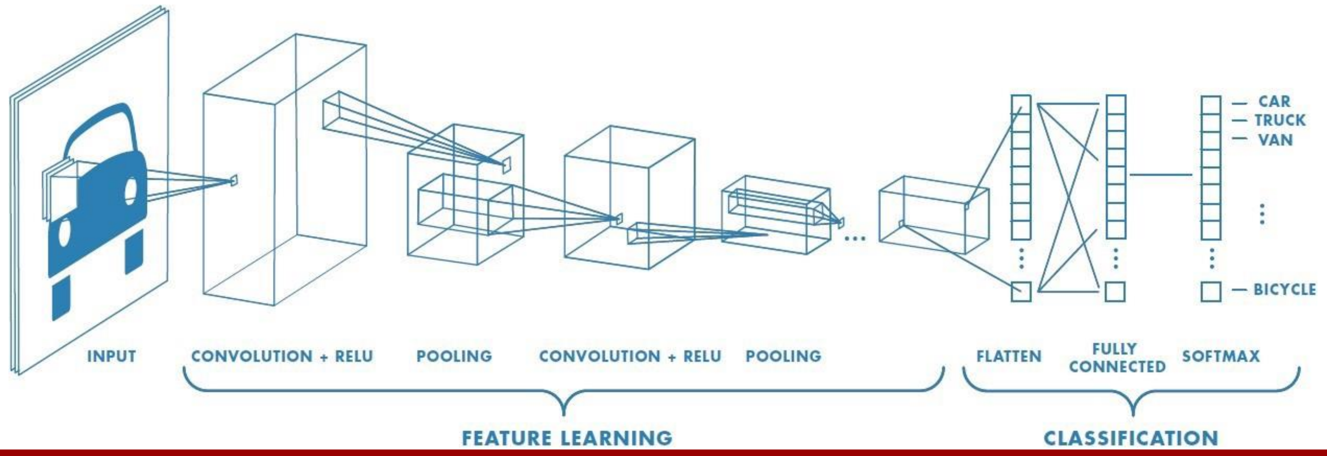
Network architecture?

Using State-of-the-art networks :

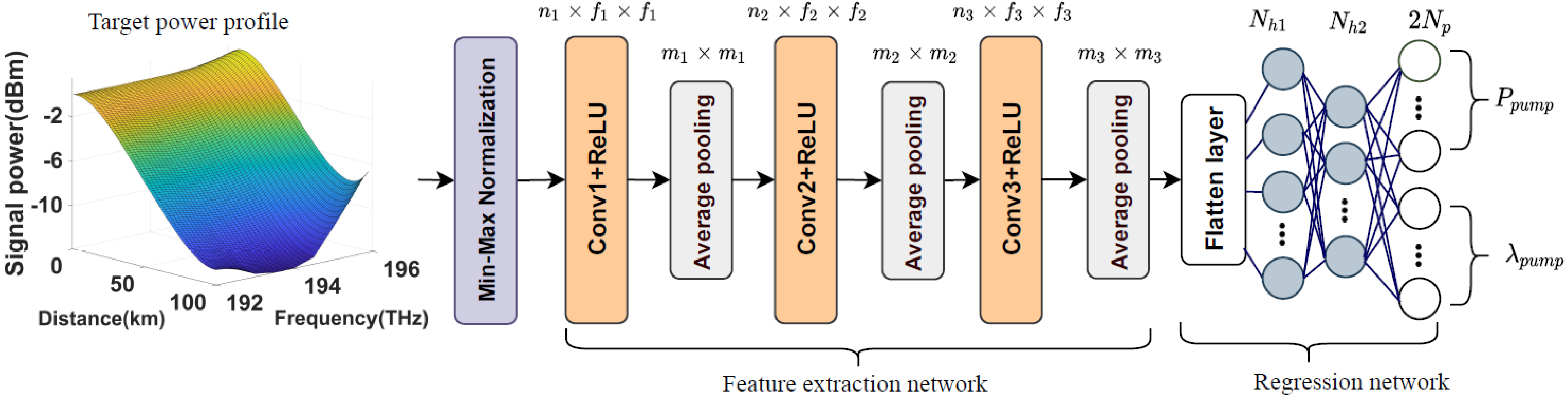
- Requires vectorising the input without removing the spatial relevancy
- Number of training parameters goes extremely high
- High training time
- Overfitting

Using Convolutional Neural Networks (CNNs)

- 2D power profile is resembled as an image
- Extracts the spatial information and decrease the redundancy
- Higher training speed and Extremely lower number of parameters

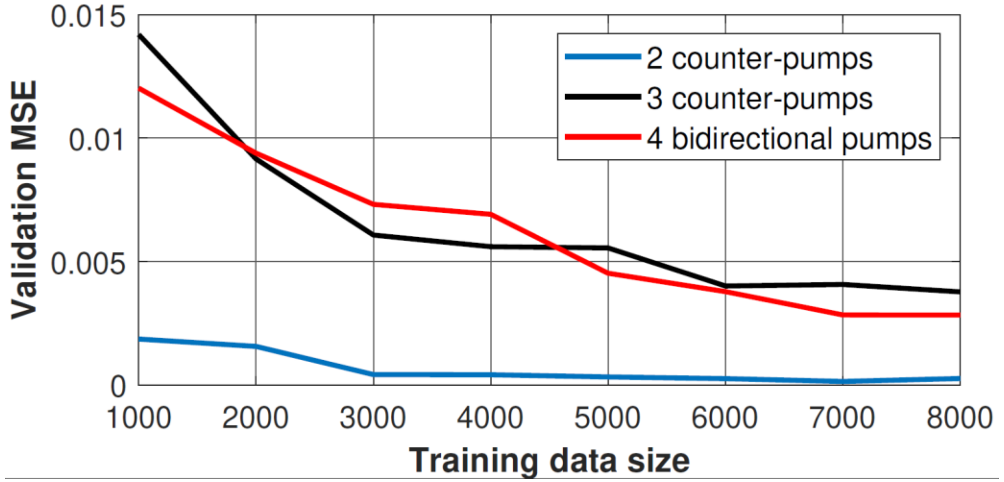


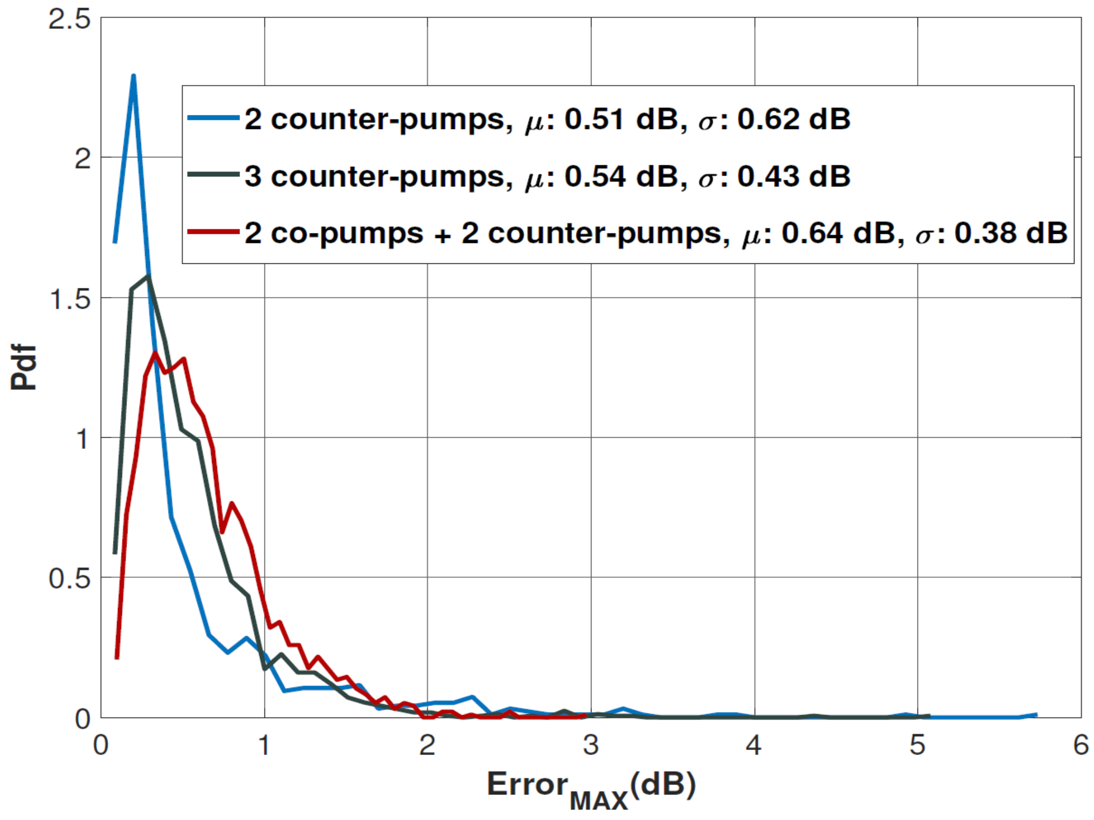
Proposed network for inverse design



M. Soltani, Opt. Lett. 46, 2650-2653 (2021).

Simulation results - 1st order pumping test results

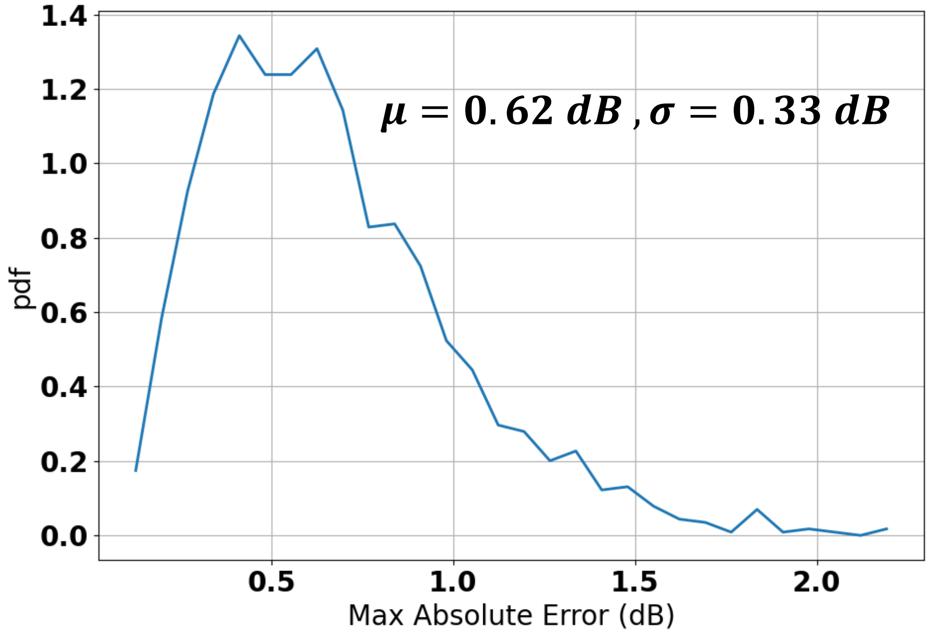




Simulation results – 2^{nd} order pumping test results

Pump parameters for 8 pumps case

Pumps	Power range	Wavelengths (fixed)
2^{nd} order co-pump	0.2 – 1.2 W	1366 nm
2^{nd} order counter-pump	0.2 – 1.2 W	1366 nm
3 1^{st} co-pumps	5 – 150 mW	[1425, 1455, 1475]
3 1^{st} counter-pumps	5 – 150 mW	[1425, 1455, 1475]



Conclusion and outlook

- Multi-layer and convolutional neural networks can learn Raman amplifier direct and inverse mappings
- Learned mappings useful for optimization of pump powers and wavelengths for:
  - *Generation of arbitrary gain profiles*
  - *Generation of arbitrary power and gain profiles*
- Maximization of information rate for ultra-wideband optical networks requires *power* and *gain* optimization
- The framework brings significant advantages for *complex* experimental optimization procedures
- Machine-learning enabled inverse system design relevant for a variety of problems in photonics

Unifying framework for noise characterization of lasers and frequency combs

[1] D. Zibar et al., "Ultra-sensitive phase and frequency noise measurement technique using Bayesian filtering," Photonics Technology Letters, 2019 (invited paper)

[2] D. Zibar et al., "Towards intelligence in photonic systems," Optics & Photonics News, 2020

[3] G. Brajato et al., "Bayesian filtering framework for noise characterization of frequency combs," Optics Express 2020

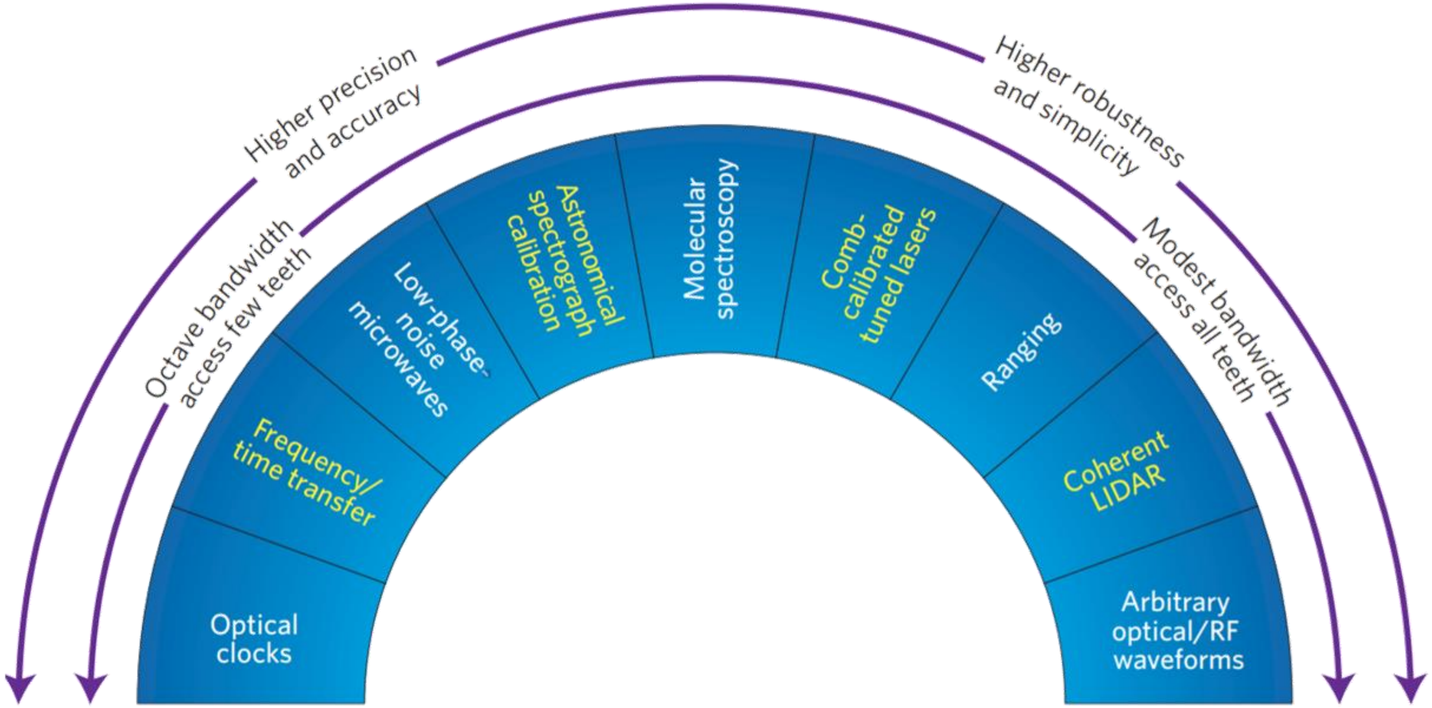
[4] H. M. Chin et al. "Machine learning aided carrier recovery in quantum key distribution," npj Qunatum Inf. 2020

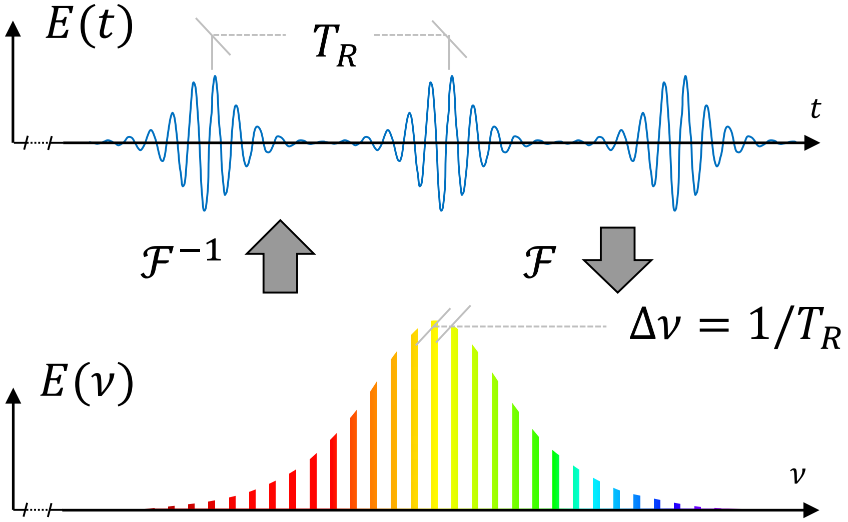
[5] N. Von Bandel et al., "Time-dependent laser linewidth: beat-note digital acquisition," Optics Express, 2016

[6] X. Xie et al., "Phase noise characterization of sub-hertz linewidth lasers via digital cross correlation," Opt. Lett. 2017

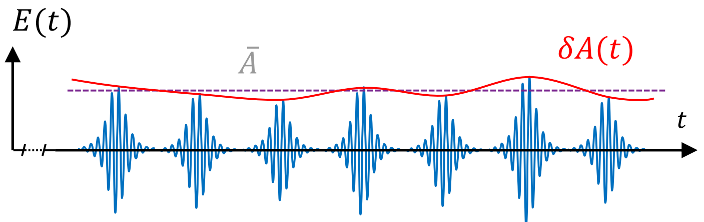
[7] D. Zibar et al., "Optimum phase measurement in the presence of amplifier noise," under review in Optica (https://arxiv.org/abs/2106.03577)

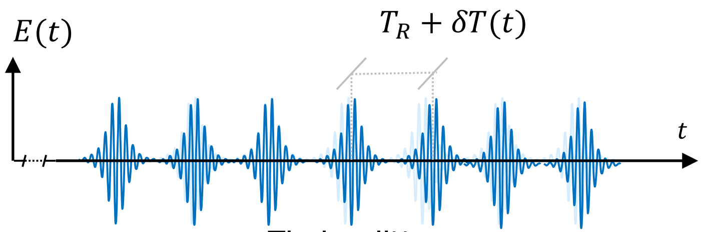
Optical frequency combs

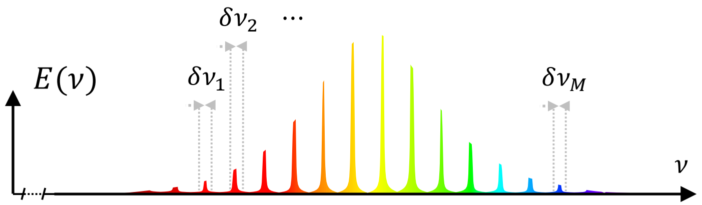




- Comb applications **performance** heavily relies on the macroscopic comb noise properties:

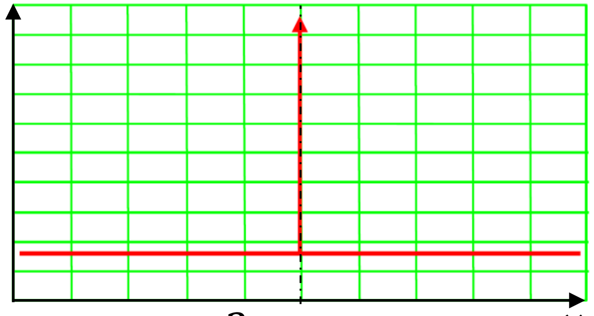






Noise definitions

Ideal oscillator

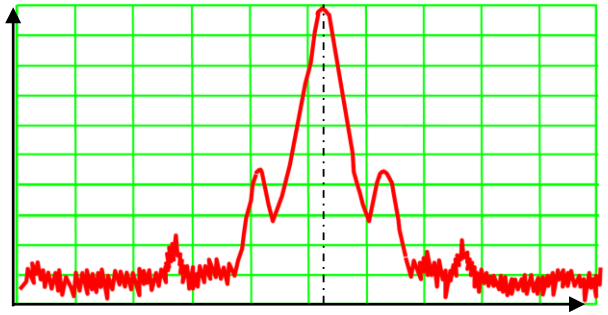


$$s(t) = \underbrace{A_0}_{\text{Mean amplitude}}\cos(\underbrace{2\pi\nu_0}_{\text{Mean frequency}} t)$$

$$\underbrace{\tilde{\phi}(t)}_{\text{Total phase}} = 2\pi\nu_0 t + \phi(t)$$

Sum of a linear growing trend and a stochastic term

Real oscillator

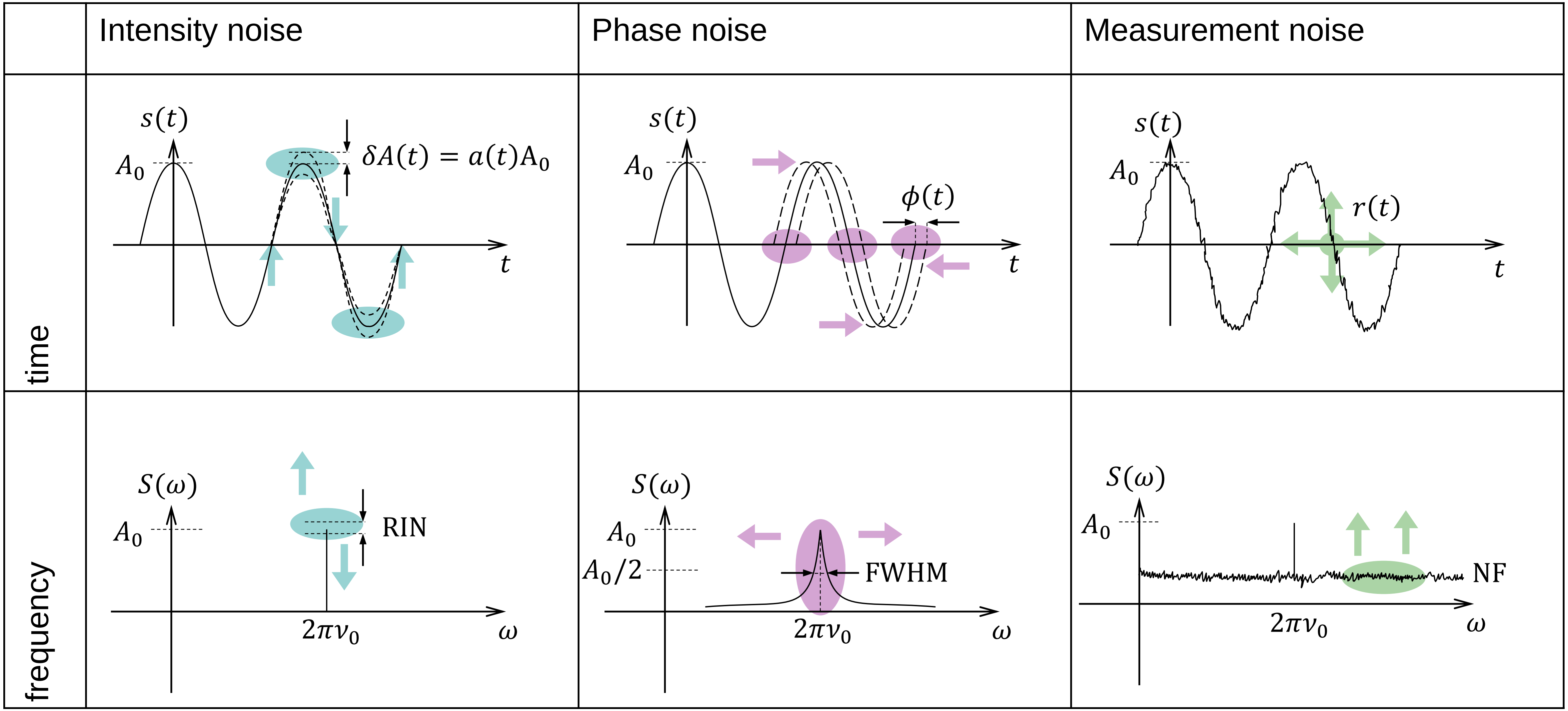


$$s(t) = A_0\big(1 + \underbrace{a(t)}_{\text{Relative intensity noise}}\big)\cos\big(2\pi\nu_0 t + \underbrace{\phi(t)}_{\text{Phase noise}}\big) + \underbrace{r(t)}_{\text{Measurement noise}}$$

$$\underbrace{\nu(t)}_{\text{Instantaneous frequency}} = \frac{1}{2\pi}\frac{d\tilde{\phi}(t)}{dt} = \nu_0 + \frac{1}{2\pi}\frac{d\phi(t)}{dt} = \nu_0 + \underbrace{\delta\nu(t)}_{\text{Frequency noise}}$$

Sum of a constant term and a stochastic term

Noise in time and frequency domain



The importance of measuring the optical phase

- Optical communication systems (DSP free data-center links)
- Noise characterization of lasers and frequency combs
- Quantum key distribution
- Classical and quantum sensing
- Gravitation wave interferometry

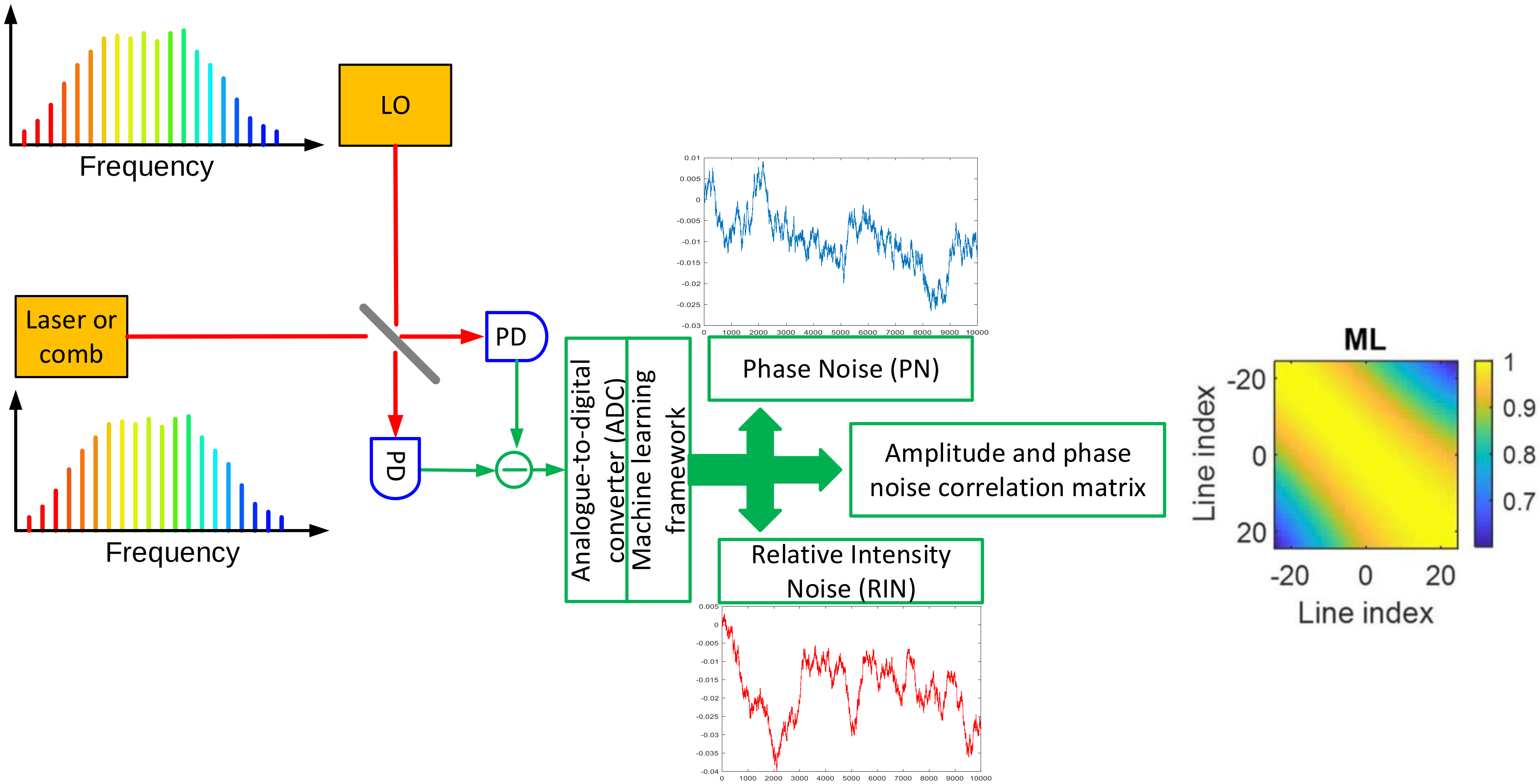
Lower bound on laser phase noise dictated by quantum noise – *how do we measure it?*

Limitation of the state-of-the-art

- Conventional methods based on delay-interferometer or cross-correlation
- Impact of optical amplifier and electronic noise
- Measurement noise floor sets a limit on frequency range (<10 MHz)
- Measurement noise floor sets a limit on range (<150 dB rad^2/Hz)
- Require relatively high input powers (>0 dBm)
- Cannot distinguish noise contribution from cavity itself and optical amplifier

State-of-the-art methods are not optimal according to statistical learning theory

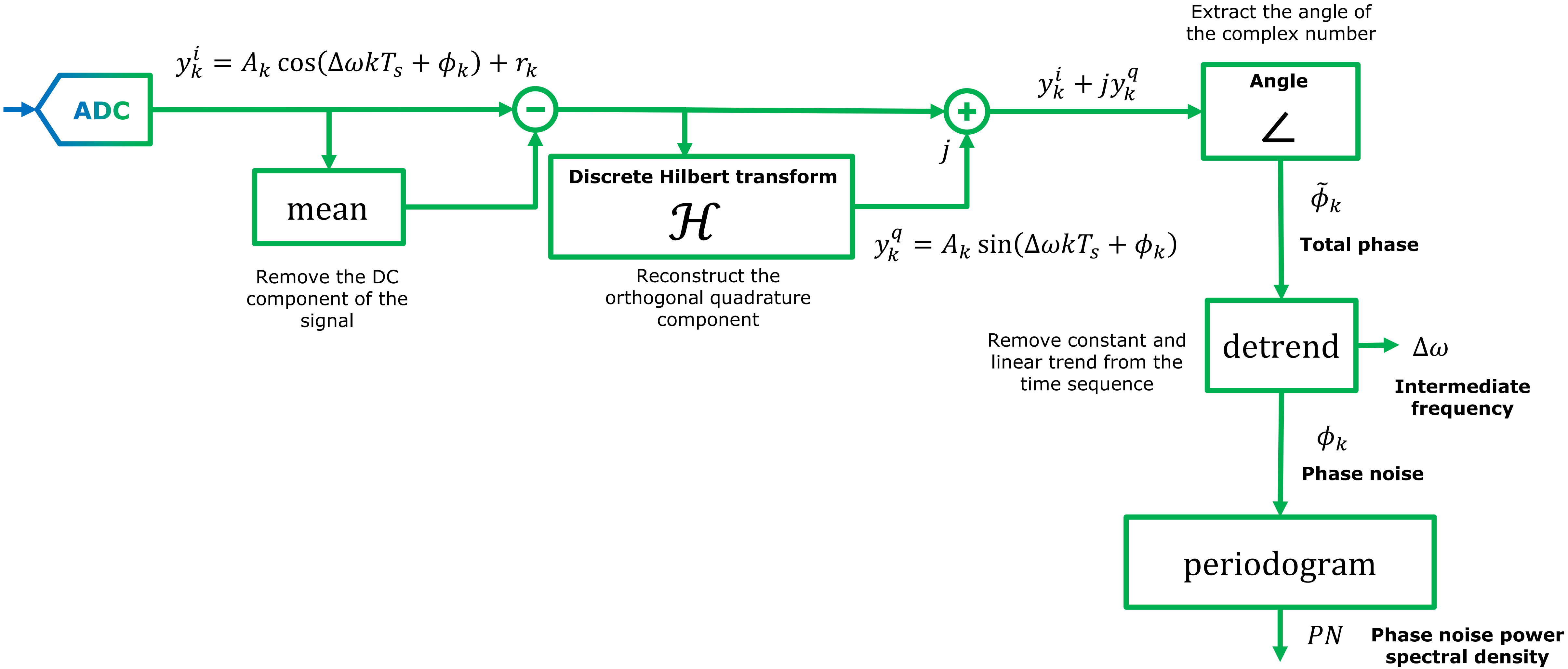
Unifying framework for noise characterization



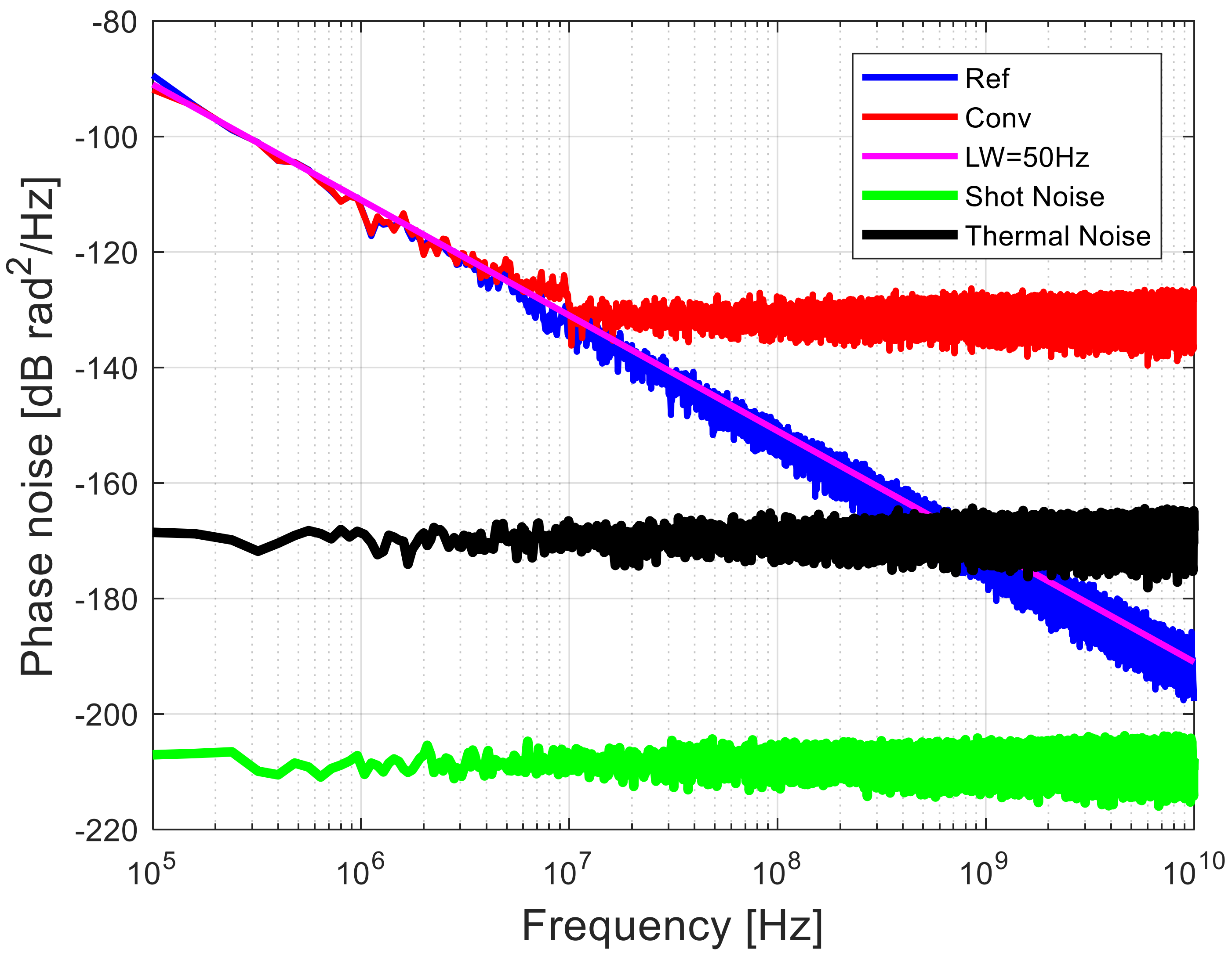
[1] D. Zibar et al, PTL 2019

Record sensitive optical phase measurement demonstrated *(-75 dBm, -200 dB rad²/Hz, 20 GHz)*

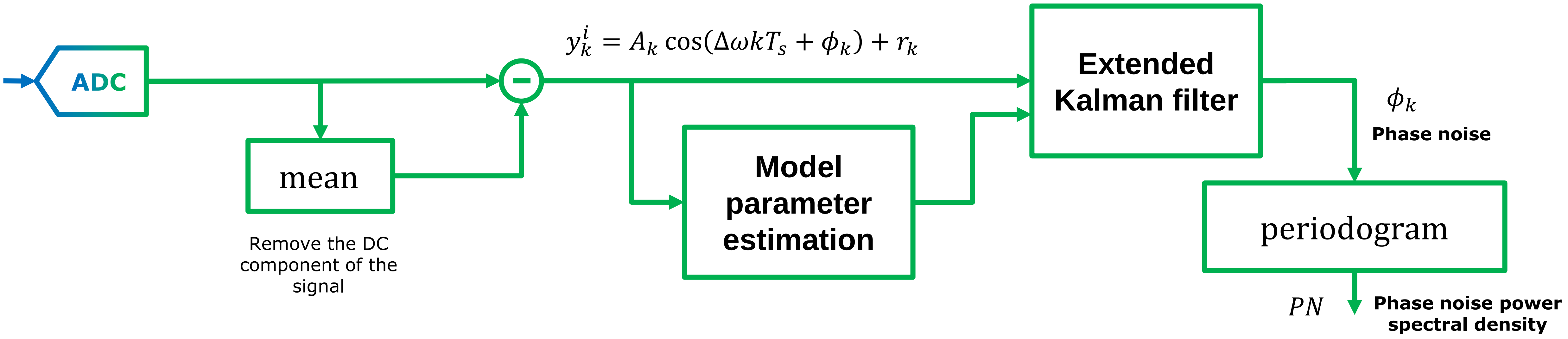
Conventional phase measurement



Numerical illustration



Bayesian filtering based phase extraction



The model we are using

$$\phi_k = \phi_{k-1} + q_k^{\phi} \qquad q_k^{\phi} \sim \mathcal{N}(0, \sigma_{\phi}^2)$$

$$y_k^i = A\cos(\Delta\omega k T_s + \phi_k) + r_k \qquad r_k \sim \mathcal{N}(0, \sigma_r^2)$$

The parameters:

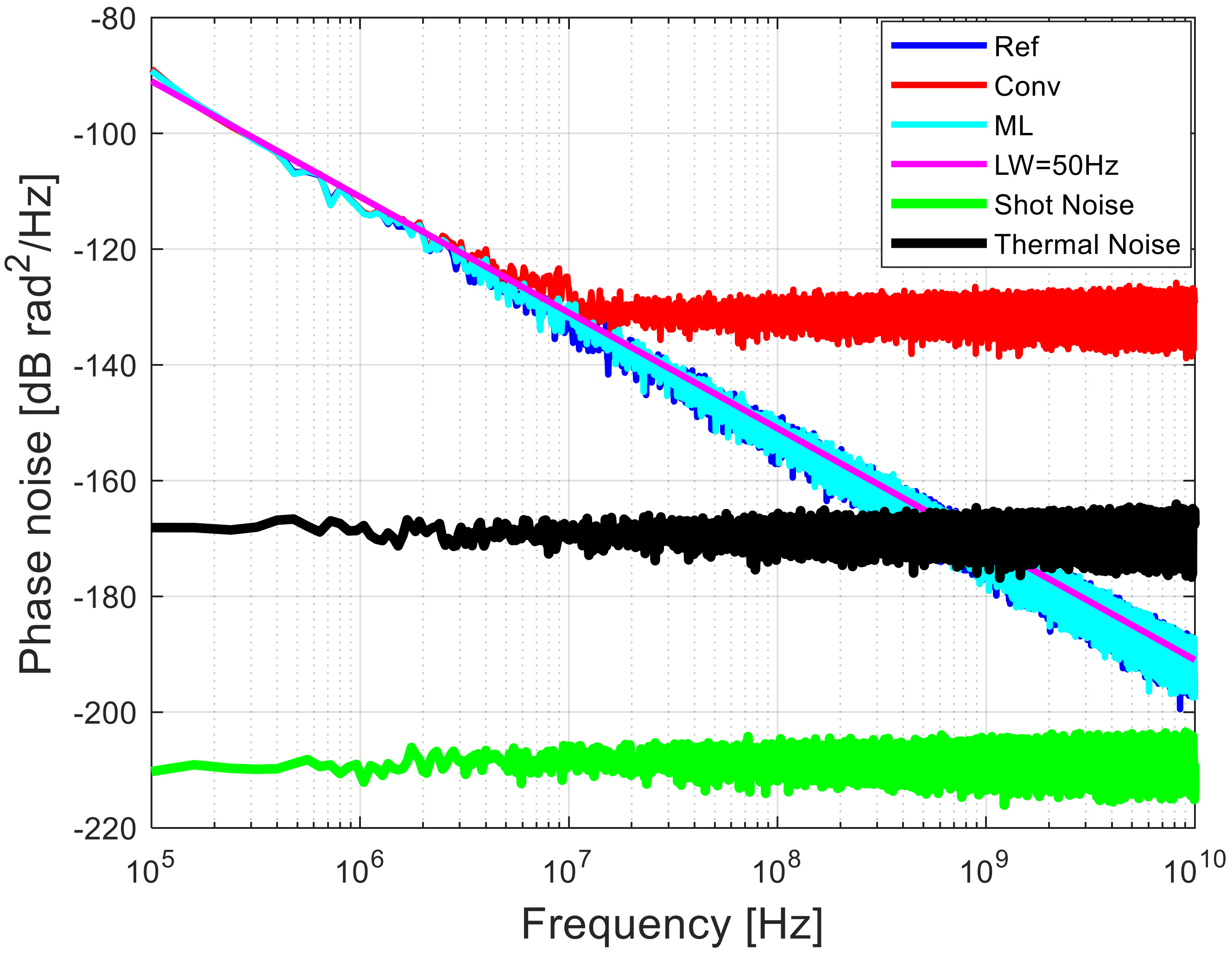
σ_{ϕ}^2 **Phase noise variance**

σ_r^2 **Measurement noise variance**

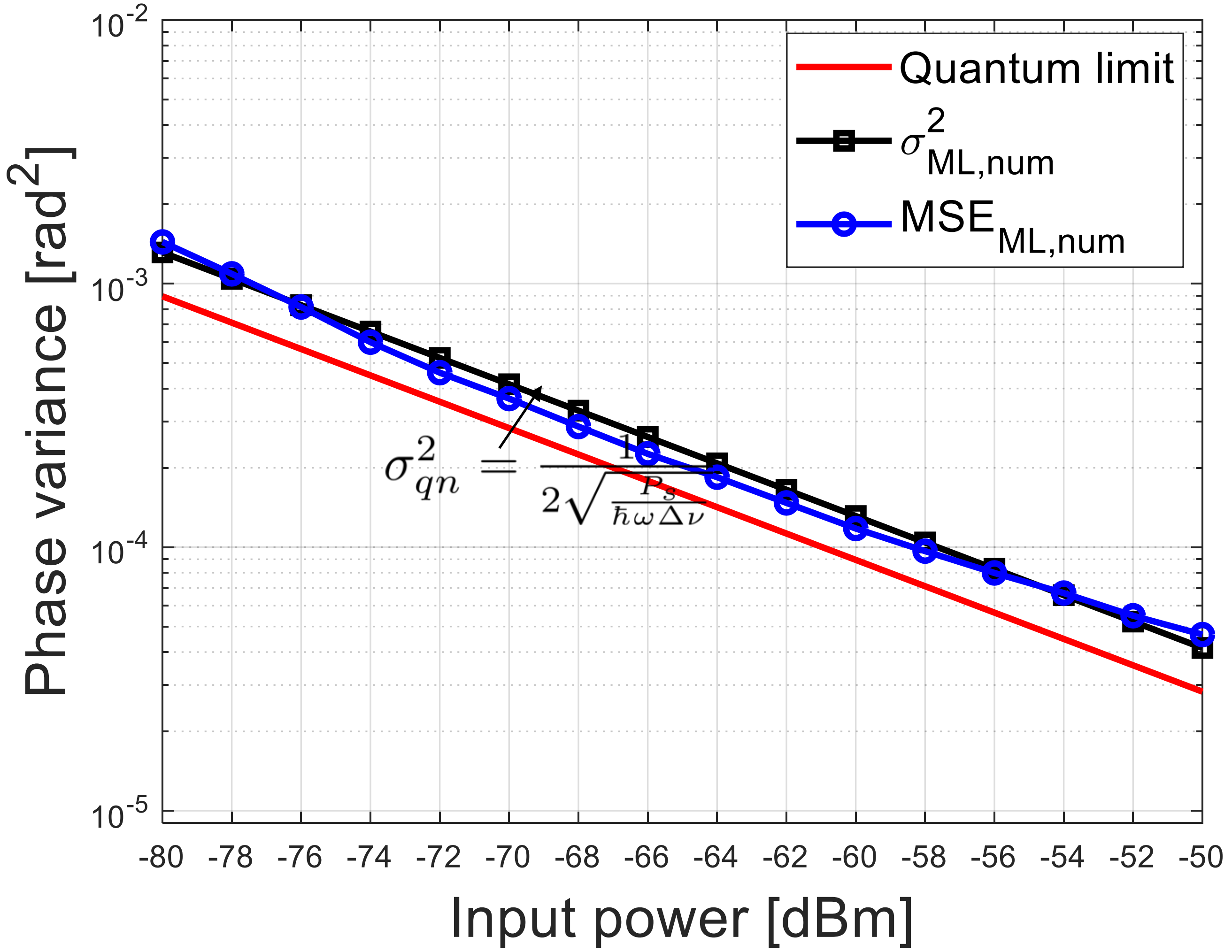
$\Delta\omega$ **Intermediate angular frequency**

A **Average signal amplitude**

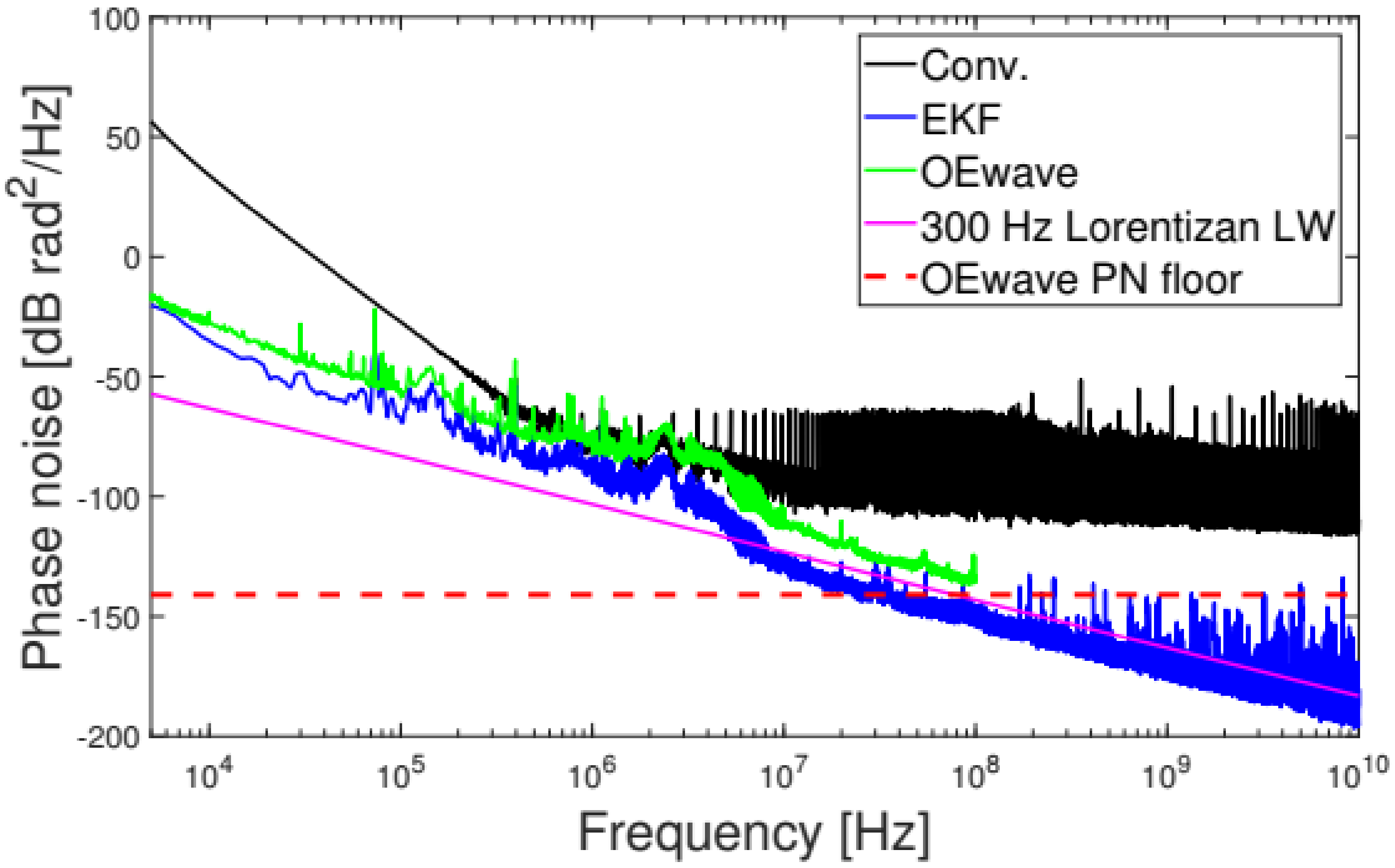
Bayesian filtering approach not limited by thermal noise



Quantum limited phase estimation

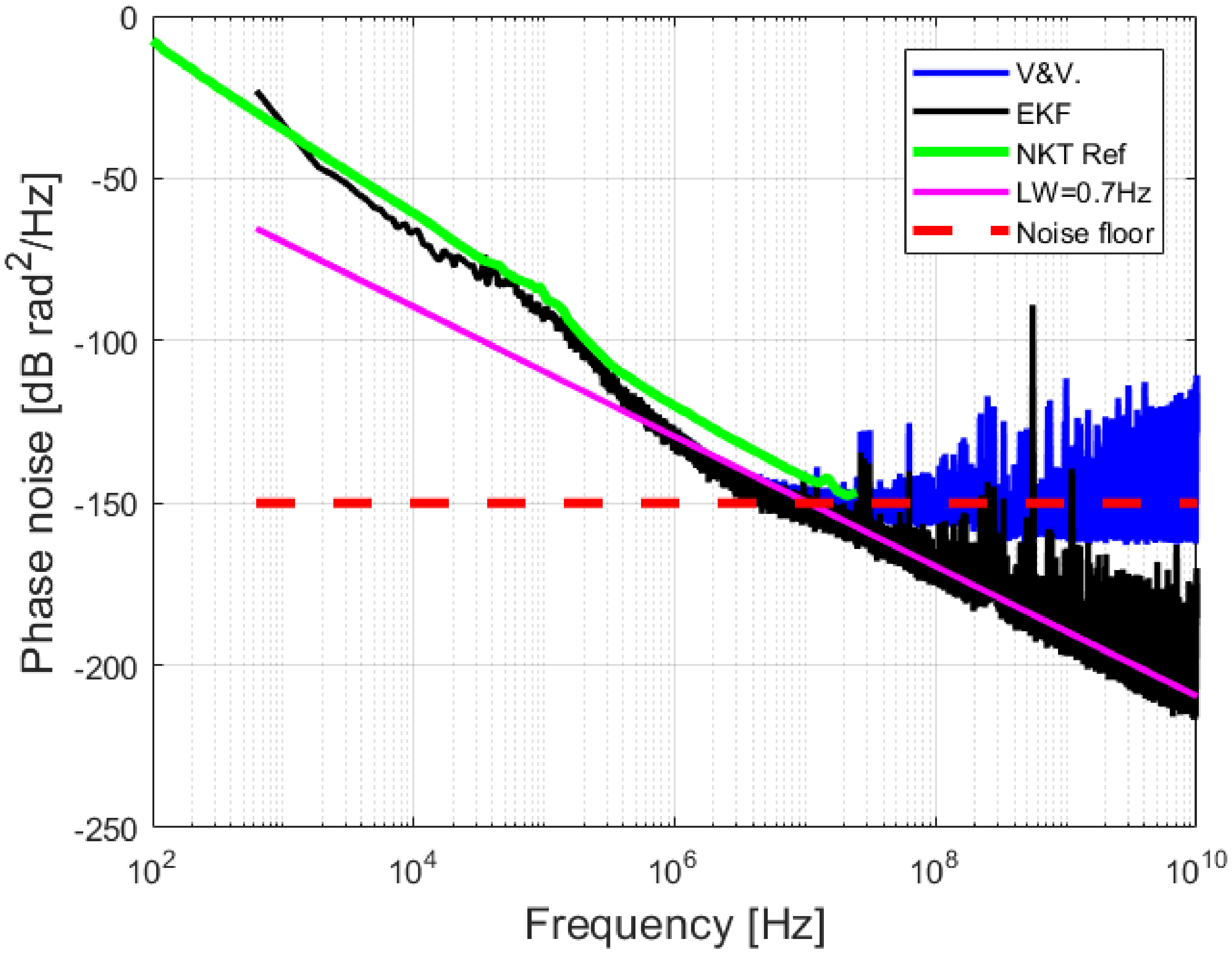


Experimental results on low-noise lasers (joint work with UCSB)

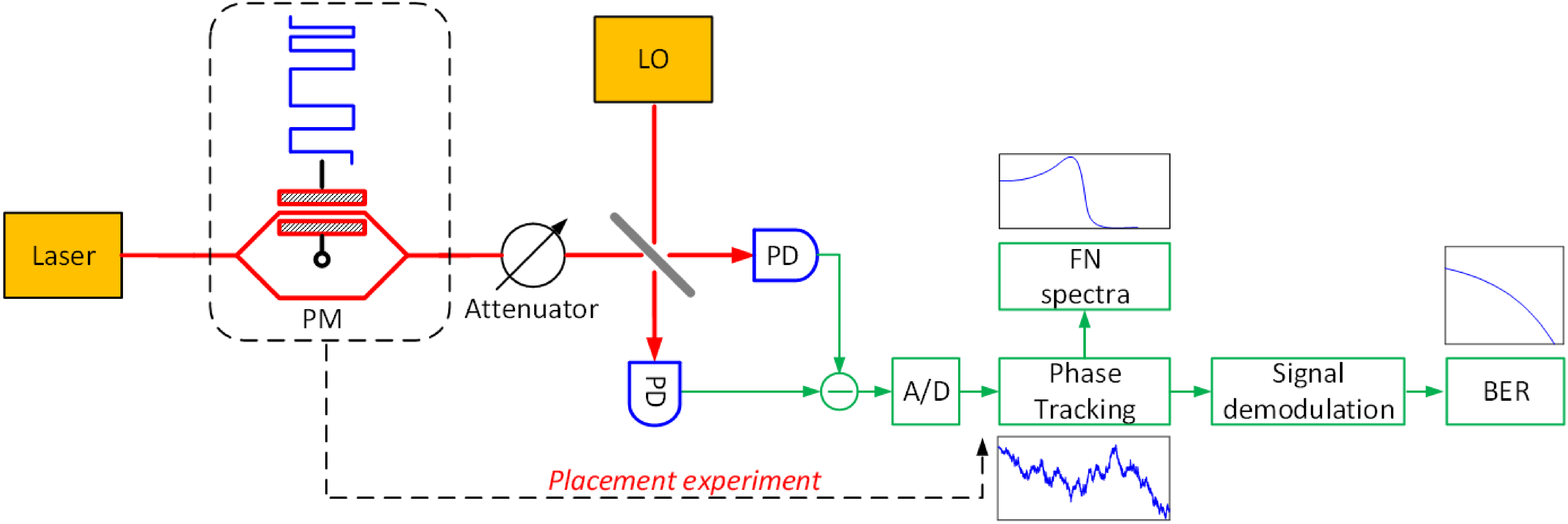


Identification of quantum noise contribution – lower bound on phase noise

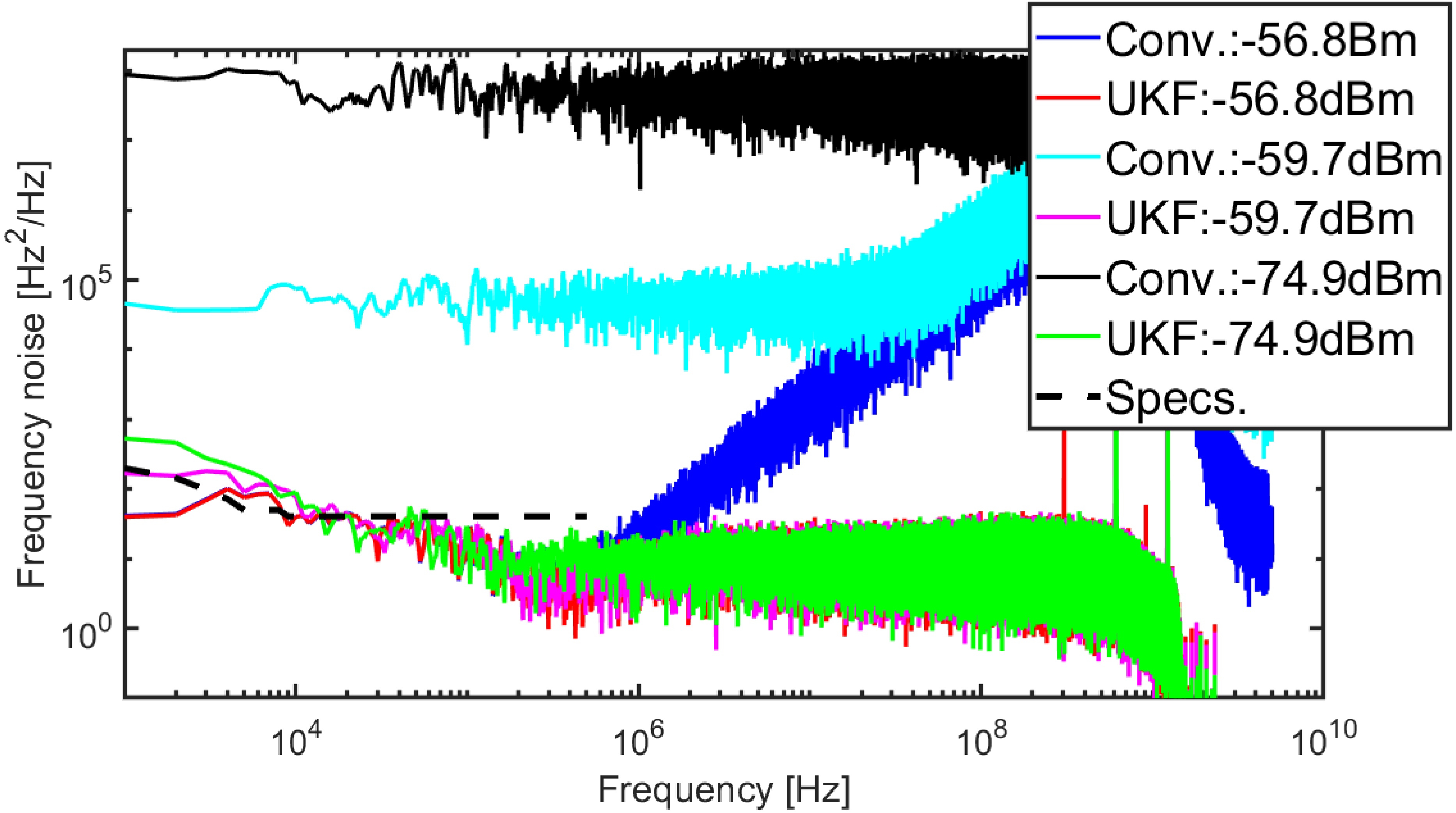
Identification of fundamental linewidth (ultra low-noise fiber laser)



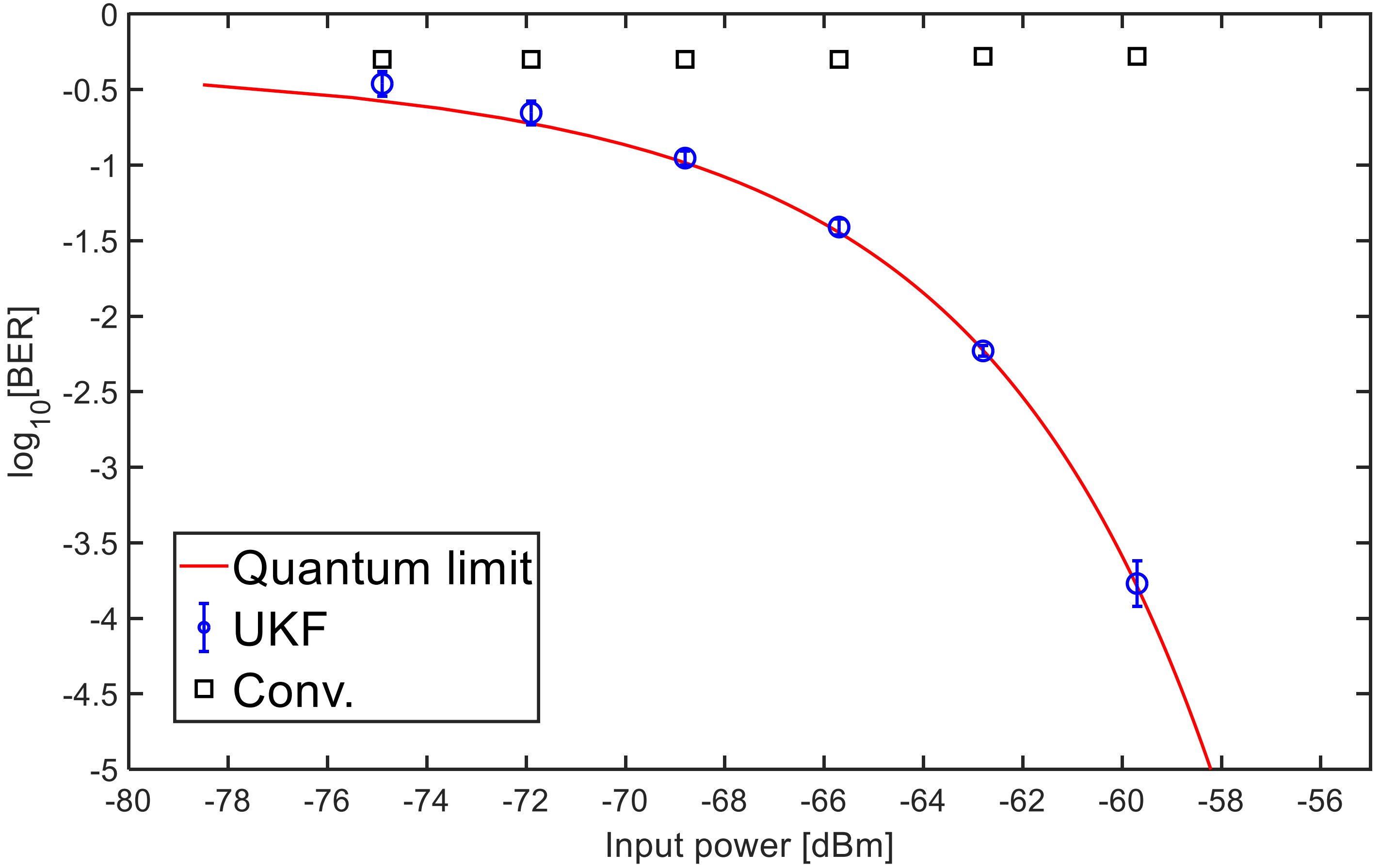
Quantum-limited optical communication



Ultra-sensitive frequency noise measurement



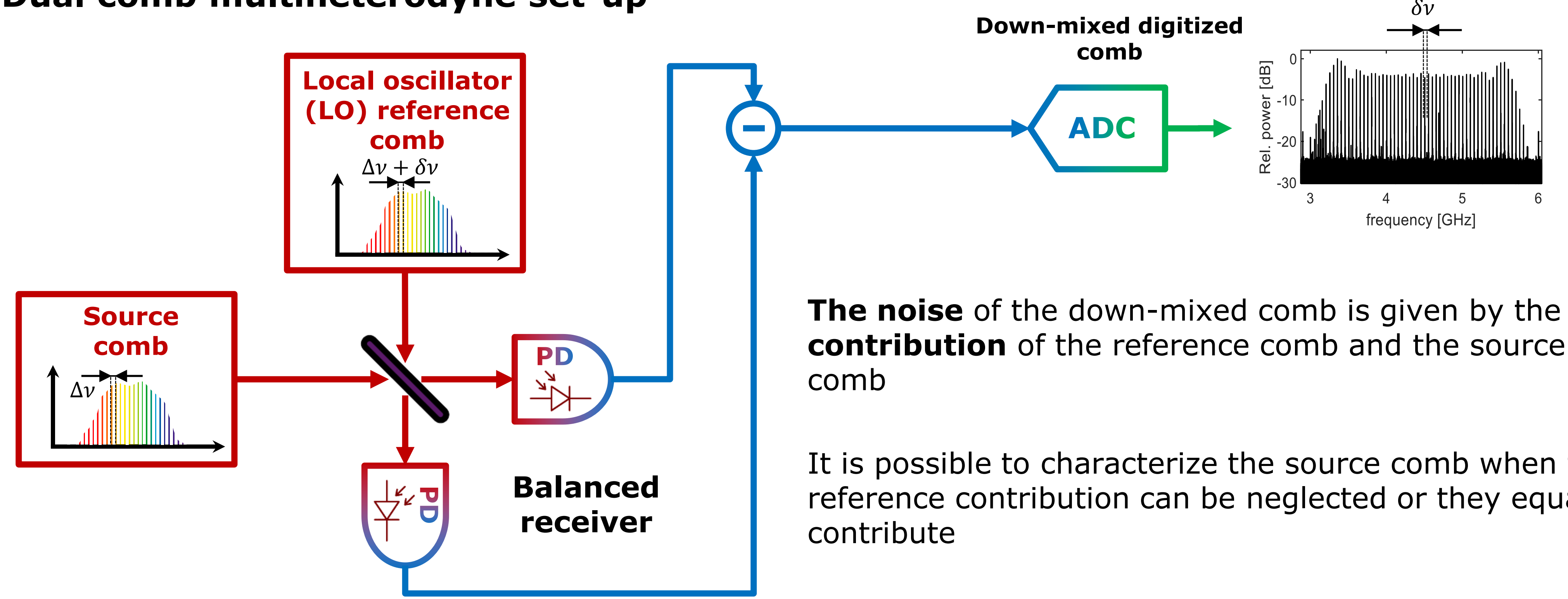
Quantum limited Bit error rate performance



Negligible penalty compared to the quantum limit bit error rate

Characterization of frequency combs

Dual comb multiheterodyne set-up



The noise of the down-mixed comb is given by the **contribution** of the reference comb and the source comb

It is possible to characterize the source comb when the reference contribution can be neglected or they equally contribute

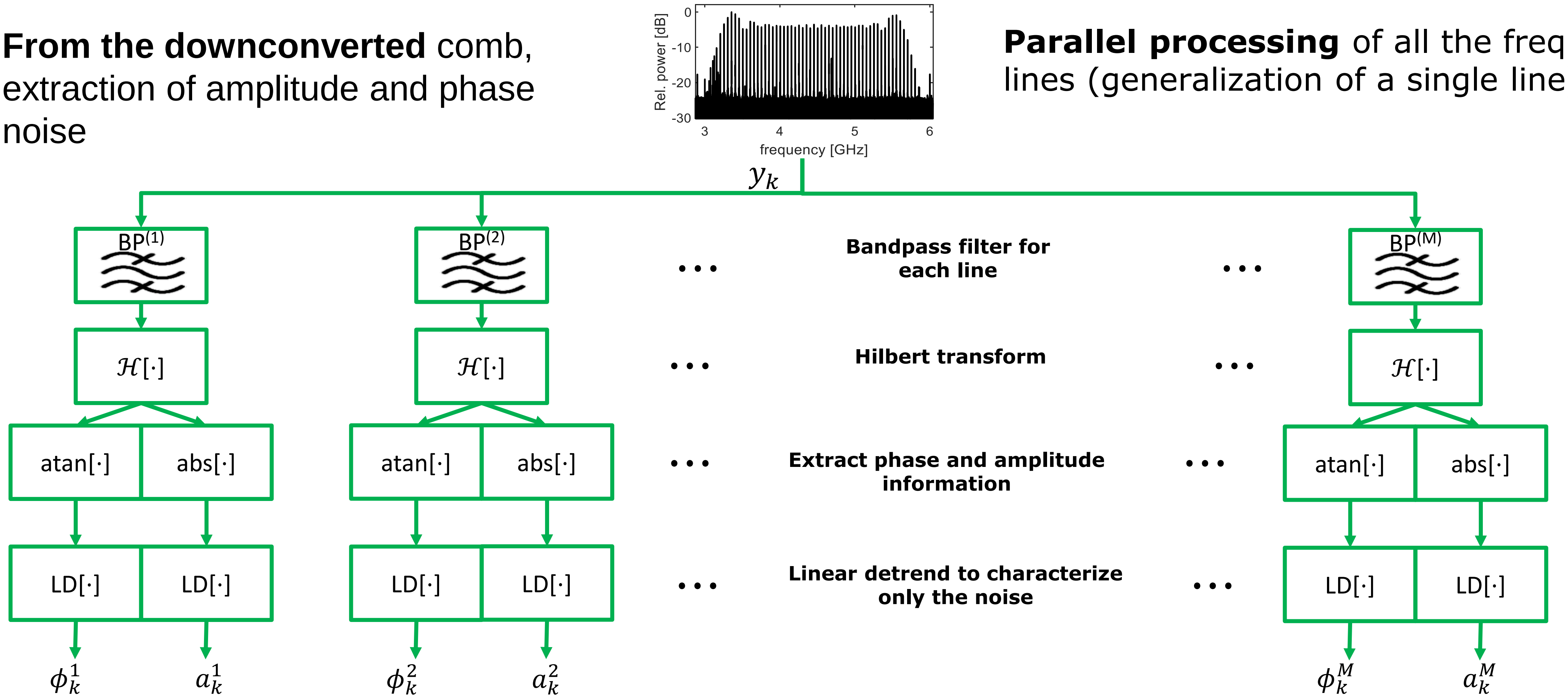
The **k-th sample of the down-digitized comb** can be described in time domain as a summation of beating tones

$$y_k = \sum_{m=1}^{M} \bar{A}^m (1 + a_k^m) \cos(\Delta\omega_m k T_S + \phi_k^m) + n_k$$

Conventional phase noise extraction

From the downconverted comb, extraction of amplitude and phase noise

Parallel processing of all the frequency lines (generalization of a single line)



Problem! Measurement noise affect the comb noise estimation

Bayesian filtering for joint amplitude and phase noise estimation

Hidden state: phase and amplitude noise of all lines

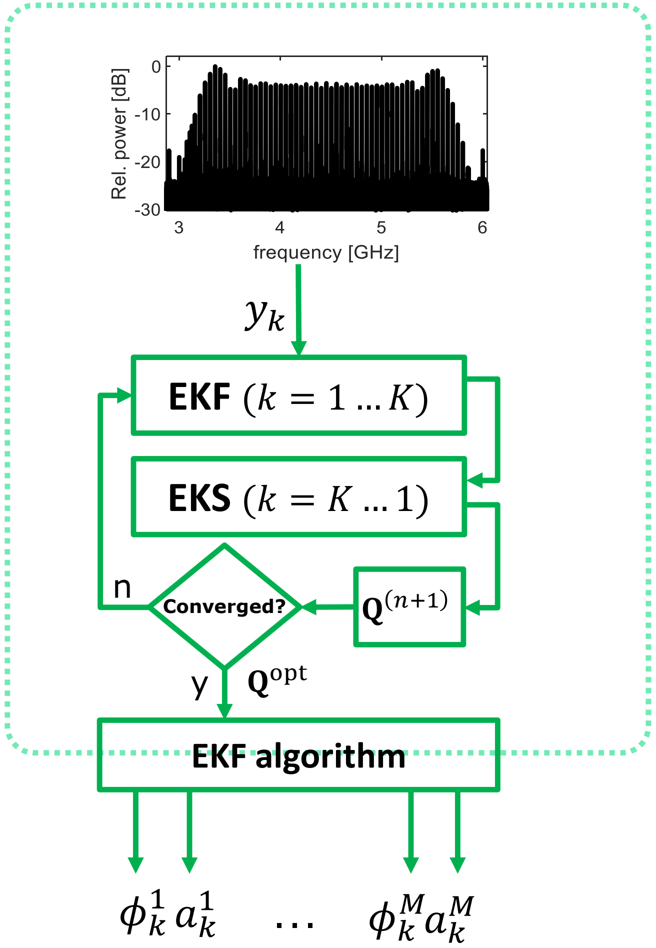
Phase and amplitude model: **Multidimensional Gaussian random walk**

$$\boldsymbol{\phi}_k = \begin{bmatrix} \phi_k^1 \\ \phi_k^2 \\ \vdots \\ \phi_k^M \end{bmatrix} \qquad \delta\boldsymbol{A}_k = \begin{bmatrix} \delta A_k^1 \\ \delta A_k^2 \\ \vdots \\ \delta A_k^M \end{bmatrix}$$

With M lines, we have M phase noise sequences and M amplitude noise sequences

$$\begin{bmatrix} \boldsymbol{\phi}_k \\ \delta\boldsymbol{A}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_{k-1} \\ \delta\boldsymbol{A}_{k-1} \end{bmatrix} + \boldsymbol{q}_{k-1}, \qquad \text{with } \boldsymbol{q}_{k-1} \sim \mathcal{N}(0, \mathbf{Q}) \text{ ,} \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_\phi & \mathbf{Q}_{\phi A} \\ \mathbf{Q}_{A\phi} & \mathbf{Q}_A \end{bmatrix}$$

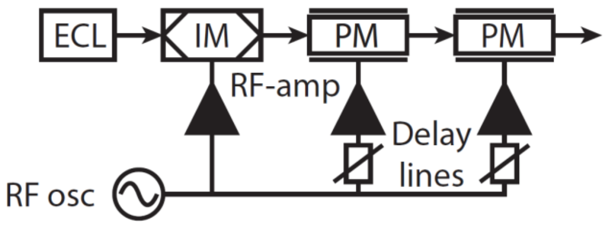
$$y_k = \sum_{m=1}^{M} \bar{A}^m (1 + \delta A_k^m) \cos(\Delta\omega_m k T_S + \phi_k^m) + n_k$$



[1] G. Brajato et al, Optics Express 2020

Data generation: EO combs

Electro optic comb generation



$$\phi_k^m = \phi_k^L + m\phi_k^{RF}$$

The phase noise of each line has two independent contributions

$$m = -\mathrm{H}, -\mathrm{H}+1, \dots, 0, \dots, \mathrm{H}-1, \mathrm{H}$$

Relative line index

$$\mathrm{H} = \left\lfloor \frac{M}{2} \right\rfloor + 1$$

Covariance of the phases

$$\mathrm{Var}[\phi_k^m] = \mathrm{Var}\left[\phi_k^L\right] + m^2 \mathrm{Var}\left[\phi_k^{RF}\right]$$

$$\mathrm{Cov}[\phi_k^m \phi_k^n] = \mathrm{Var}\left[\phi_k^L\right] + mn\mathrm{Var}\left[\phi_k^{RF}\right]$$

The covariance matrix describe how the noise variance affect different comb lines

$$\mathrm{Cov}\left[\boldsymbol{\phi}_k \boldsymbol{\phi}_k^\top\right] = \boldsymbol{\Sigma} = \boldsymbol{c}\boldsymbol{c}^\top \sigma_L^2 + \boldsymbol{h}\boldsymbol{h}^\top \sigma_{RF}^2$$

$$\boldsymbol{c} = [\underbrace{1,1,\dots 1}_{M\ lines}]^\top$$

$$\boldsymbol{h} = [\underbrace{-\mathrm{H}, -\mathrm{H}+1, \dots, 0, \dots, \mathrm{H}-1, \mathrm{H}}_{M\ lines}]^\top$$

Correlation of the phases

$$\mathrm{Corr}[\phi_k^m \phi_k^n] = \frac{\mathrm{Cov}[\phi_k^m \phi_k^n]}{\mathrm{std}[\phi_k^m]\mathrm{std}[\phi_k^m]}$$

The correlation matrix is re-scaled such that it’s maximum value is 1.

Correlation describes how similar are two lines.

1 = perfectly correlated lines

0 = uncorrelated lines

-1 = anticorrelated lines

Building phase correlation matrices

The covariance can be estimated after estimating the phases

$$\boldsymbol{\Sigma}_{N,l} = \frac{1}{N-1}\sum_{k=l}^{l+N}(\boldsymbol{\phi}_k - \overline{\boldsymbol{\phi}})(\boldsymbol{\phi}_k - \overline{\boldsymbol{\phi}})^{\mathsf{T}}$$

We are free to choose:

- **The starting point** l for estimating the covariance matrix
- **The number of contiguous samples** N we use for estimating a covariance matrix

We can **define the observation time** $\tau_{obs} = NT_s$

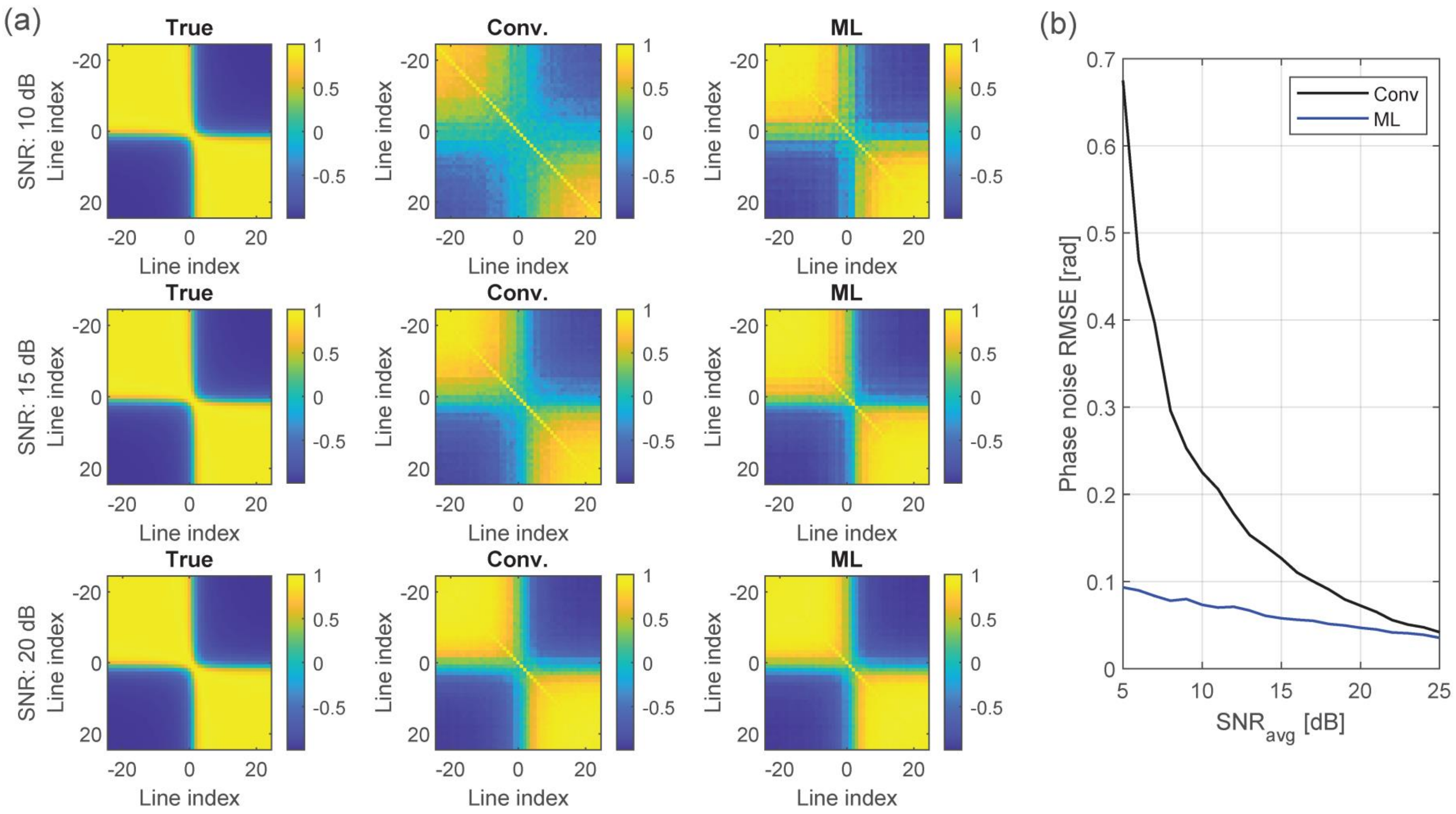
We can calculate a covariance matrix, hence do an **eigenvalue decomposition for different observation times**

This allow to identify **what are the dominant noise sources at shorter and longer timescales**

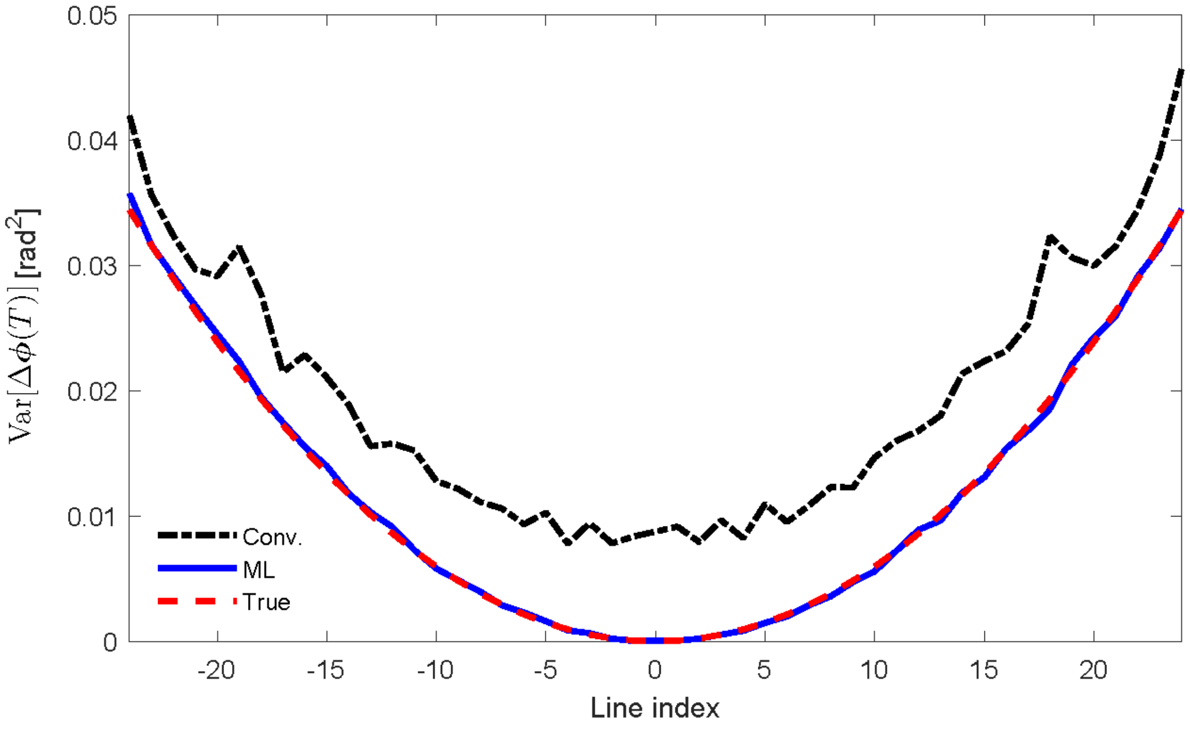
The noise sources will be identified by the corresponding eigenvector

The difference with PCA: PCA does the same but using the full signal length, this approach shows the principal component at different observation times

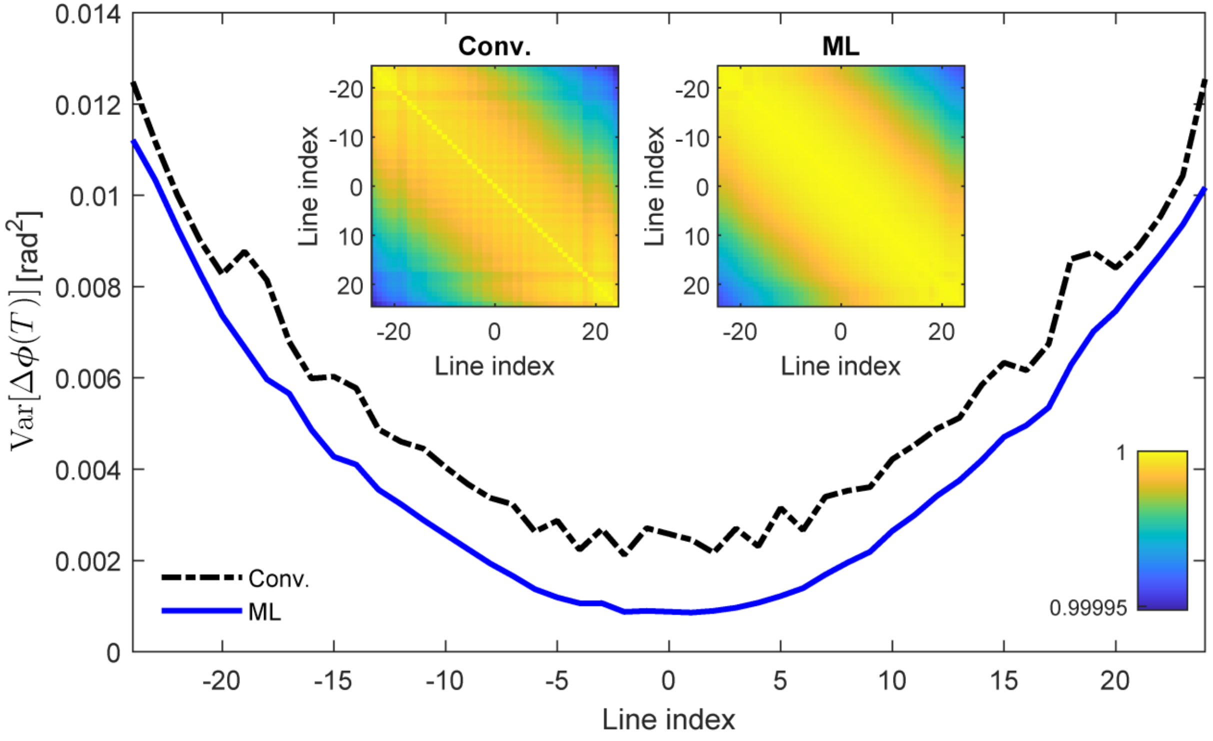
Frequency comb phase noise correlation matrix



Combs lines phase variance



(a) Simulations



(b) Experimental

Machine learning methods provides more accurate estimations

Sub-space analysis: Eigenvalue decomposition

$$\mathrm{Cov}\left[\boldsymbol{\phi}_k \boldsymbol{\phi}_k^\top\right] = \boldsymbol{\Sigma} = \boldsymbol{c}\boldsymbol{c}^\top \sigma_L^2 + \boldsymbol{h}\boldsymbol{h}^\top \sigma_{RF}^2 = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\top = \sum_{i=1}^{M} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{i\top}$$

In our case, we find that only two eigenvalues are significant (by comparing both sides and orthogonality arguments)

λ_i is the eigenvalue

$\boldsymbol{v}^i$ is the **orthonormalized** eigenvalue (or eigenmode)

By normalizing $\boldsymbol{c}$ and $\boldsymbol{k}$, and exploiting the symmetry

$$\hat{\boldsymbol{c}}\sigma_L^2\|\boldsymbol{c}\|^2 + \widehat{\boldsymbol{h}}\sigma_{RF}^2\|\boldsymbol{h}\|^2 = \boldsymbol{v}^1\lambda_1 + \boldsymbol{v}^2\lambda_2$$

$$\sigma_L^2\|\boldsymbol{c}\|^2 = \lambda_1$$

$$\sigma_{RF}^2\|\boldsymbol{h}\|^2 = \lambda_2$$

Eigenvalues are "proportional" to the strength (variance) of the noise sources

$$\hat{\boldsymbol{c}} = \boldsymbol{v}^1$$

$$\widehat{\boldsymbol{h}} = \boldsymbol{v}^2$$

Eigenvectors describe the direction of the noise source, in this case what comb lines are affected by a given noise source

$$\hat{\boldsymbol{c}}\|\boldsymbol{c}\|^2 = \boldsymbol{c}$$

$$\widehat{\boldsymbol{h}}\|\boldsymbol{h}\|^2 = \boldsymbol{h}$$

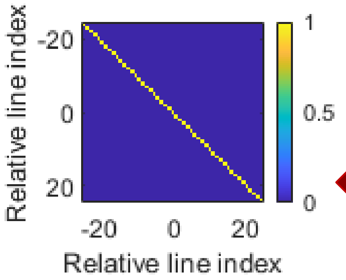
Eigenvalue decomposition on the covariance matrix helps to identify the independent and meaningful components.

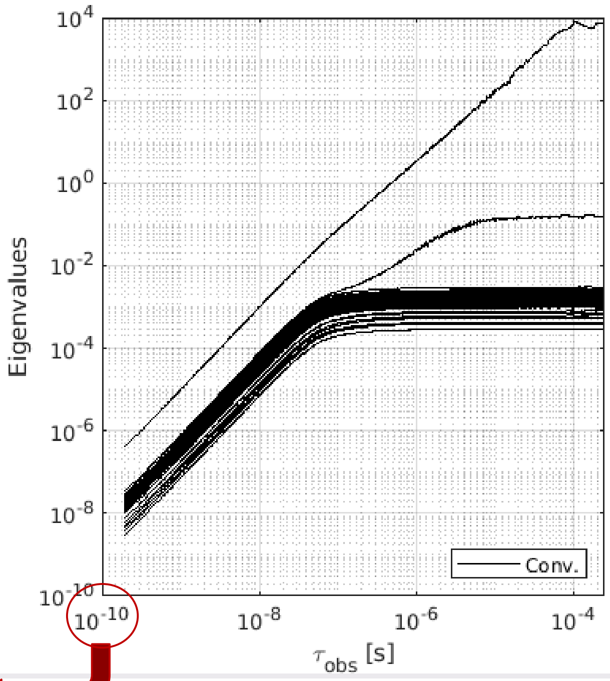
$$\boldsymbol{c} = [\underbrace{1,1,\ldots 1}_{M\ lines}]^\top$$

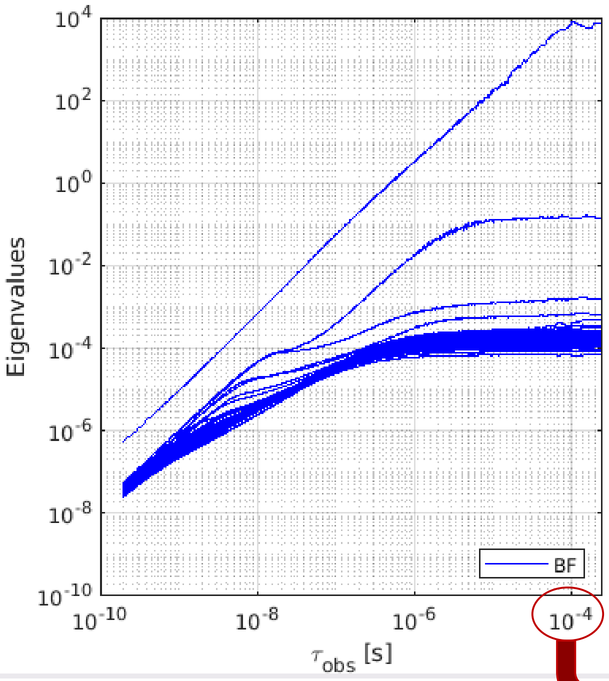
$$\boldsymbol{h} = [\underbrace{-\mathrm{H}, -\mathrm{H}+1, \ldots, 0, \ldots, \mathrm{H}-1, \mathrm{H}}_{M\ lines}]^\top$$

Covariance Eigenvalue dynamics (experimental)

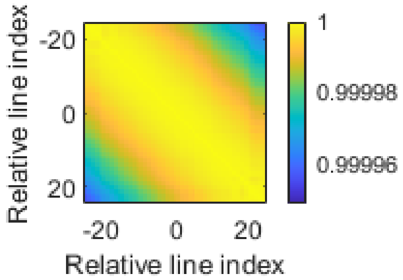
Uncorrelated phases for short observation times

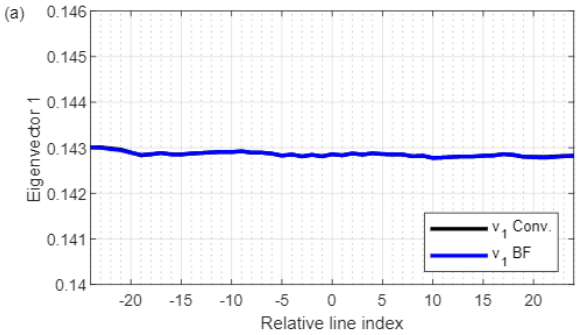


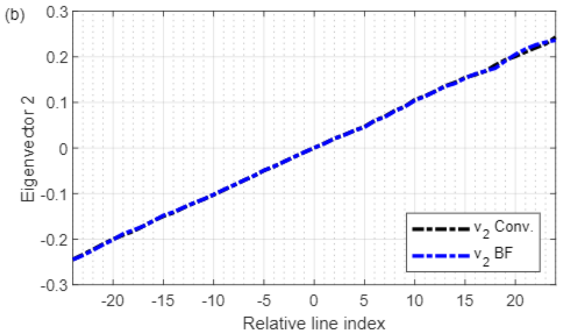




Correlated phases for long observation times



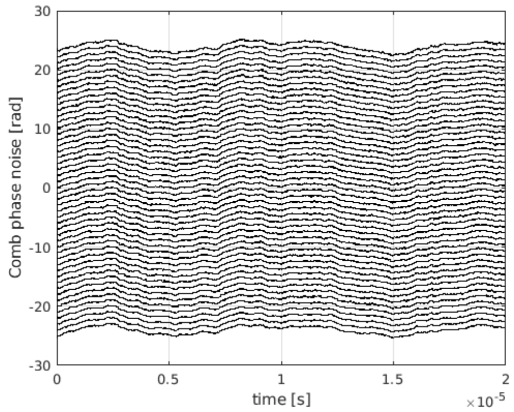


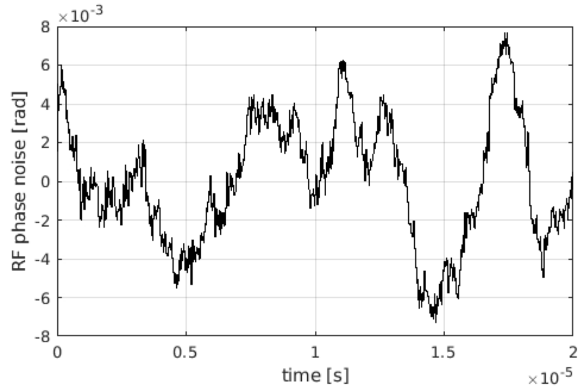


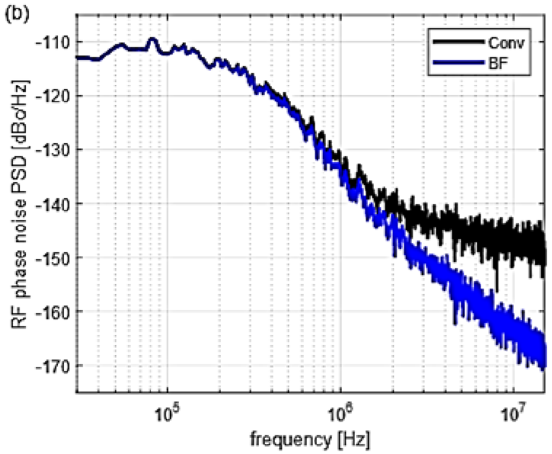
Extraction of laser and RF noise contributions

Similar to principal component projection

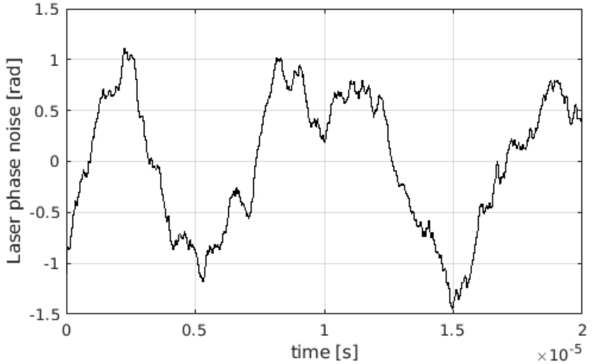
$$\mathbf{\Phi}^{\top} \cdot \frac{\boldsymbol{h}}{\|\boldsymbol{h}\|^2} = \phi^{RF}$$

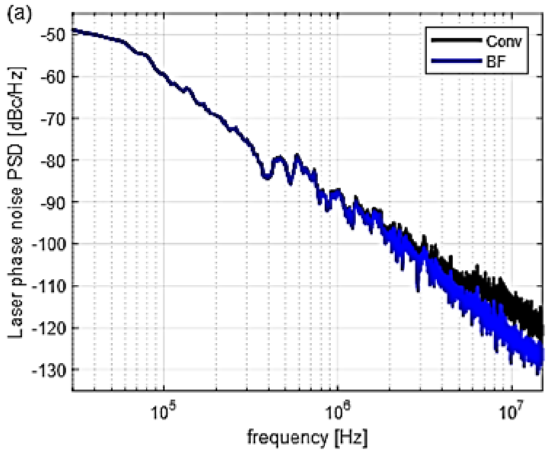






$$\mathbf{\Phi}^{\top} \cdot \frac{\boldsymbol{c}}{\|\boldsymbol{c}\|^2} = \phi^{L}$$





$$\boldsymbol{\phi}_{1:T} \coloneqq \mathbf{\Phi}_{\ M \times T}$$

$$\boldsymbol{c} = [\underbrace{1,1,\dots 1}_{M\ lines}]^{\top} \qquad M \times 1$$

$$\boldsymbol{h} = [\underbrace{-\mathrm{H}, -\mathrm{H}+1, \dots, 0, \dots, \mathrm{H}-1, \mathrm{H}}_{M\ lines}]^{\top} \qquad M \times 1$$

End-to-end learning for fiber-optic channels

[1] T. O'Shea and J. Hoydis "An introduction to deep learning for the physical layer," IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, VOL. 3, NO. 4, 2017

[2] B. Karanov et al., "End-to-end deep learning of optical fibre communication," Journal of Lightwave Technology, vol. 36, no. 20, 2018

[3] R. Jones, D. Zibar et al., "Deep learning of geometric constellation shaping including fiber nonlinearities," in Proceedings of ECOC 2018

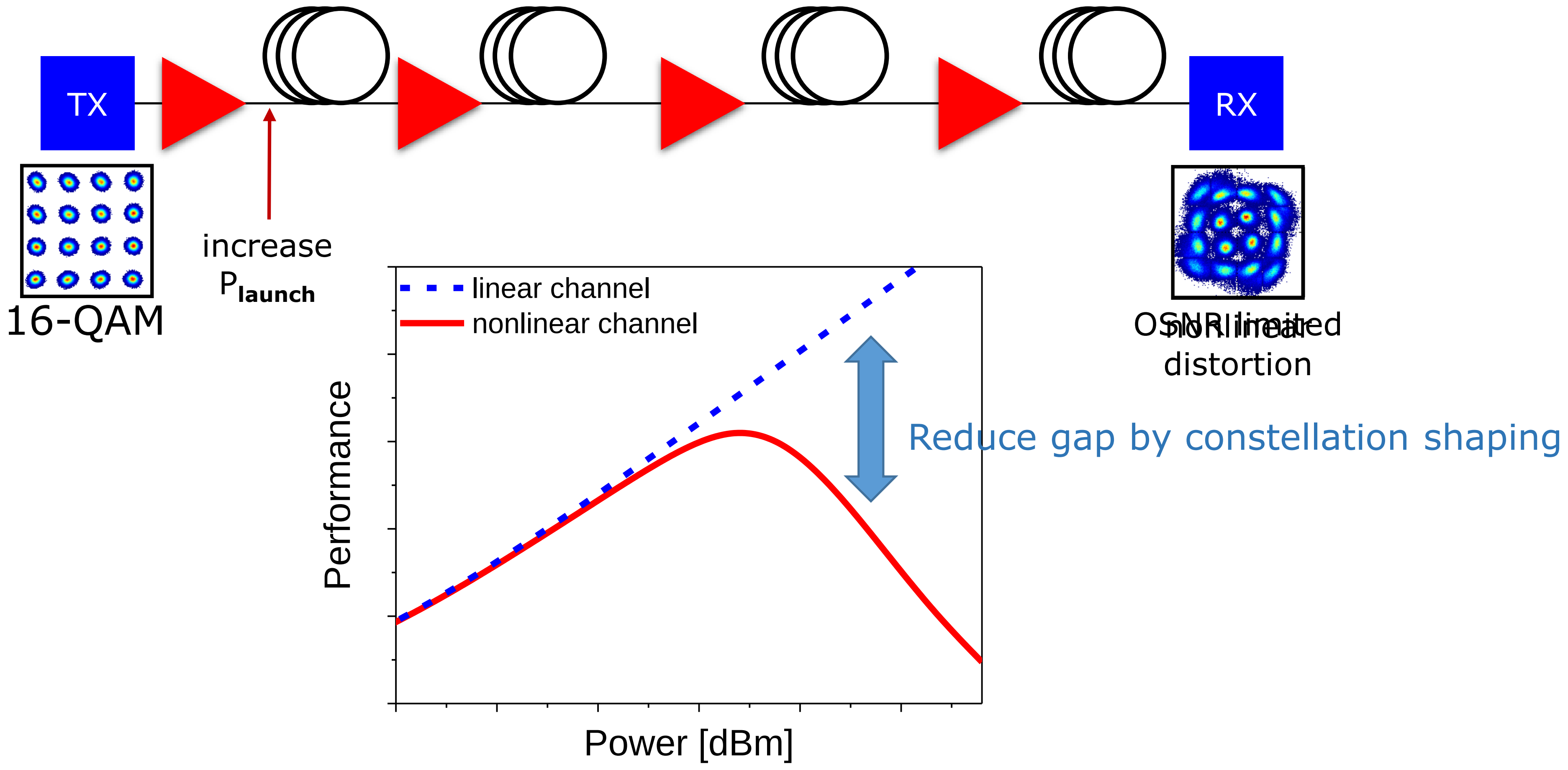
[4] R. Jones, D. Zibar et al., "End-to-end for GMI optimized geometric deep learning of geometric constellation shaping including Fiber," in Proceedings of ECOC 2019

[5] O. Jovanovic. D. Zibar et al.""Gradient-Free Training of Autoencoders for Non-Differentiable Communication Channels," Journal," vol. 31, no. 20, 2021

[6] O. Jovanovic, D. Zibar et al., "End-to-end learning of a Constellation Shape Robust to variations in SNR and laser Linewidth," In Proceedings of European Conference on Optical Communication (ECOC), 2021, (2nd place ADVA best paper award)

[7] J. Aoudia et al., "End-to-end learning of communications systems without a channel model." arXiv preprint arXiv:1804.02276 (2018)

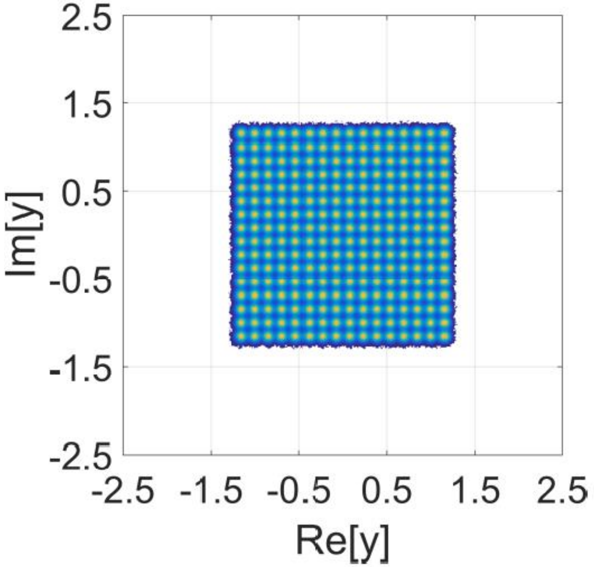
Standard coherent communications



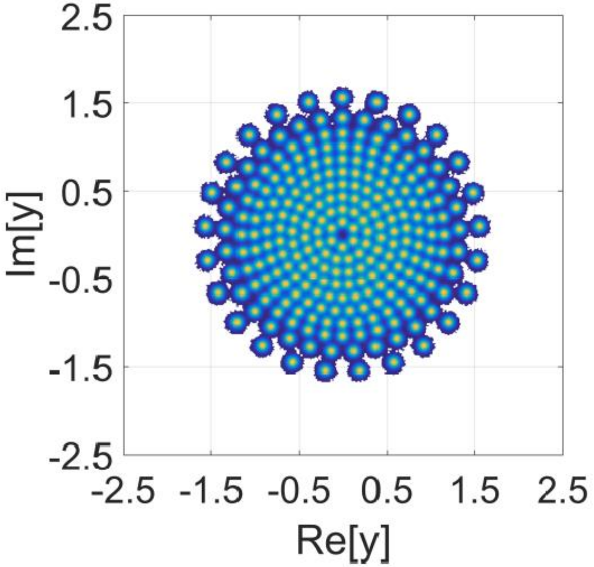
Kerr nonlinearity is one of the ultimate limits to increasing system performance

Typical optimization strategies

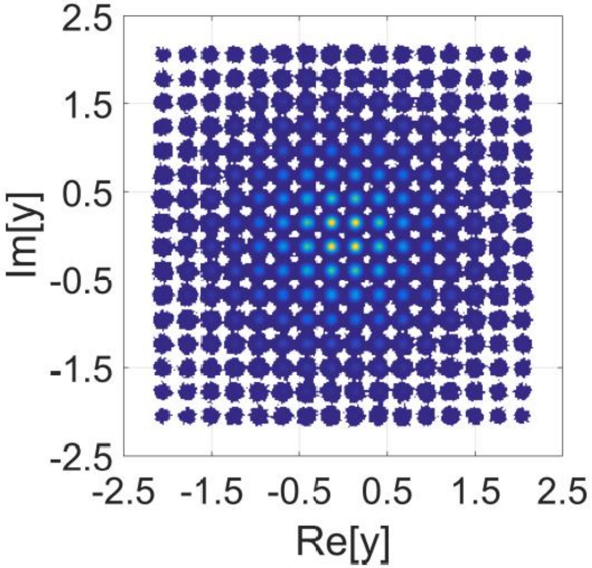
Signal $\hat{Y}$, AWGN channel with SNR=25dB



Geometric shaping – optimize the set of points $\mathcal{X}$. Typically assume $P_{\boldsymbol{X}}(\boldsymbol{X}) \sim \mathcal{U}(0, P_{av})$



Probabilistic shaping – optimize the probability mass function $P_{\boldsymbol{X}}(\boldsymbol{X})$ for a given constellation set $\mathcal{X}$



Forward channel model for the optical fiber

$$\frac{dA}{dz} = \underbrace{-\frac{\alpha}{2}A}_{\text{Fiber loss}} \overbrace{- i\frac{\beta_2}{2}\frac{d^2A}{dt^2}}^{\text{Chromatic dispersion}} \underbrace{+ i\gamma|A|^2 A}_{\text{Nonlinear Kerr effect}}$$

The nonlinear interference noise (NLIN) model:

Average power

$$y[k] = x[k] + n[k]$$

$$n[k] \sim N(0, \sigma^2_{\mathrm{ASE}} + \sigma^2_{\mathrm{NLIN/GN}}(P_{\mathrm{Tx}}, \mu_6, C))$$

Modulation
property (peak power)

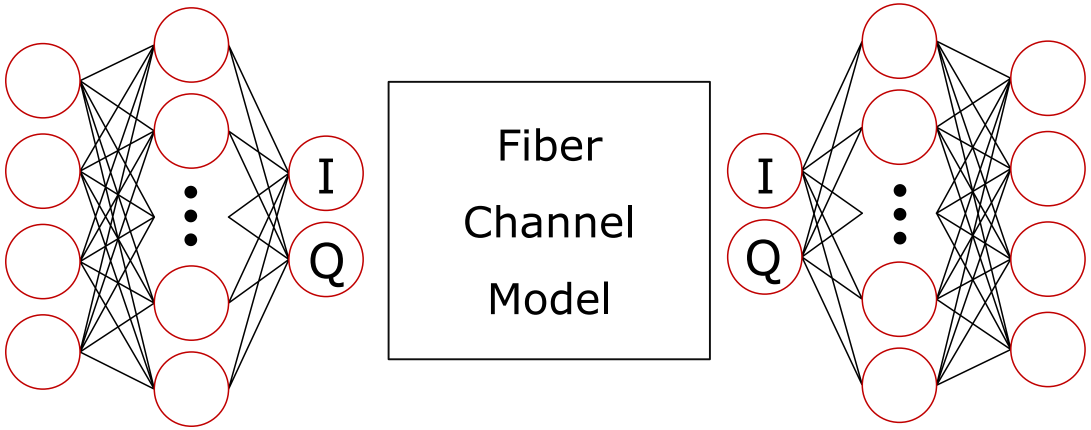
Dual power constraint – nonobvious optimal characteristics and optimization strategies

R. Dar et al., Opt. Exp. 21(22) (2013), pp. 25685-25699

Learning to mapping using auto-encoders

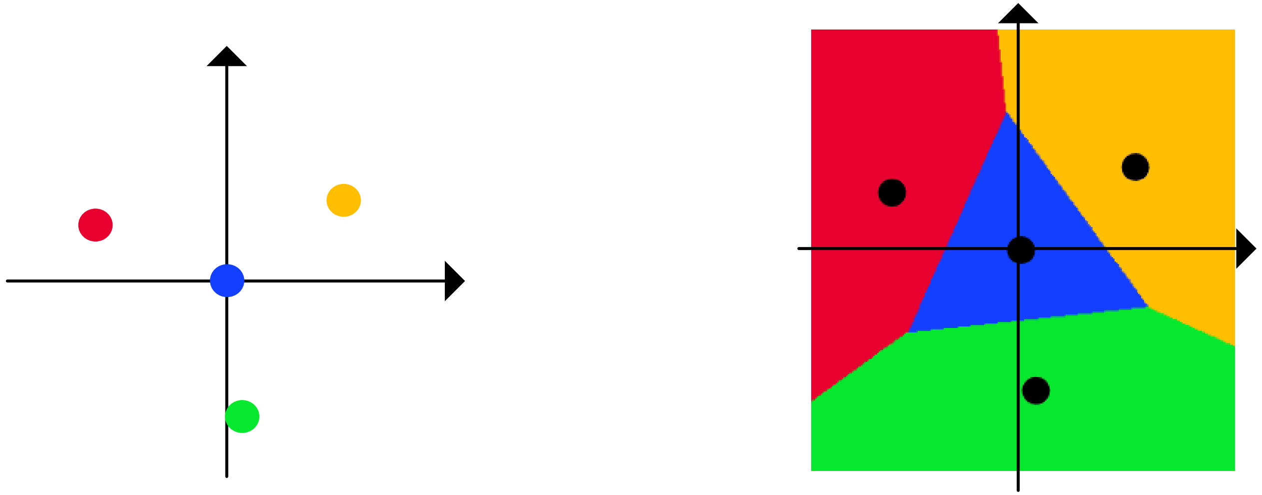
Input Space:

1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1



Output Space:

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



Auto-encoder learns constellation robust to channel impairments

Channel Models

Fiber Channel:

5 WDM Channel System

50 GHz Channel spacing

SSMF

32 GHz Bandwidth

2000 km transmission

20 Spans

- GN Model and NLIN Model[1,2] for learning
- Propagation using SSFM
- For the NLIN Model the nonlinearities depend on the moment of the constellation

For 64 QAM:

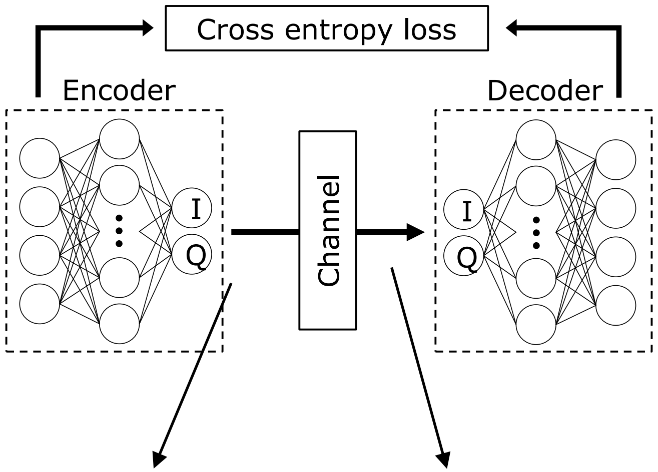
4th order moment = 1.38 $$\frac{(\mathbb{E}[b]^2)^4}{(\mathbb{E}[b]^2)^2}$$

6th order moment = 2.23 $$\frac{(\mathbb{E}[b]^2)^6}{(\mathbb{E}[b]^2)^3}$$

[1] A. Carena, et al., "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," J. Lightw. Technol., vol. 30, no. 10, pp. 1524–1539, May 15, 2012.

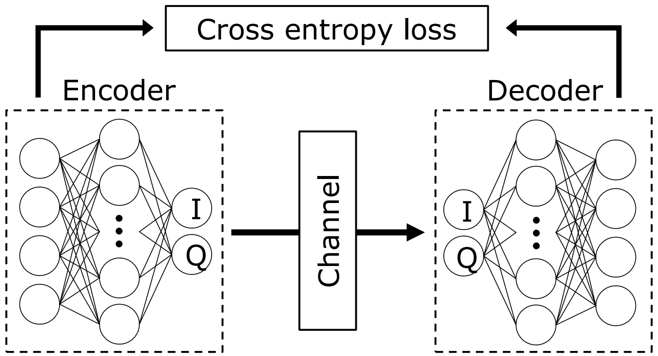
[2] R. Dar et al. ''Properties of nonlinear noise in long, dispersion-uncompensated fiber links.'' Opt. Exp. 21.22 (2013): 25685-25699.

Training process

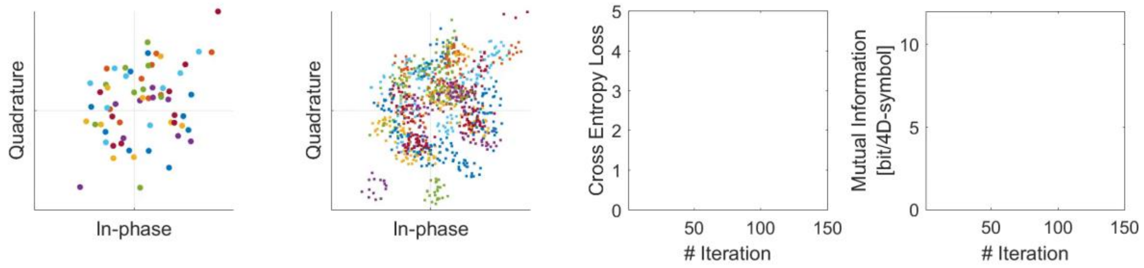


- Stochastic optimization
- Iterative training process
- Gradient based

Training process

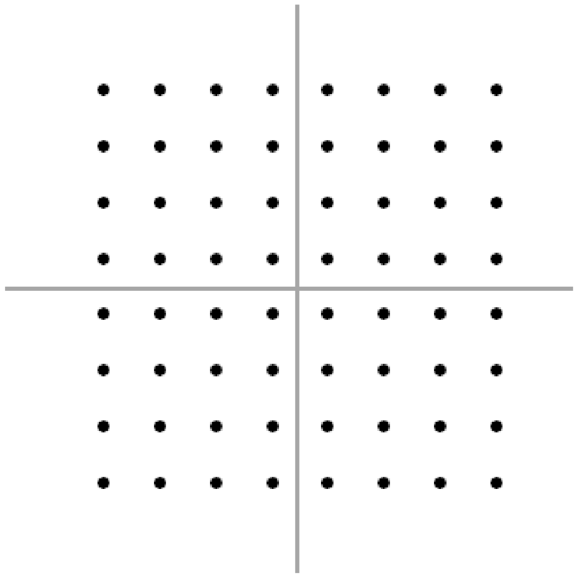


- Stochastic optimization
- Iterative training process
- Gradient based

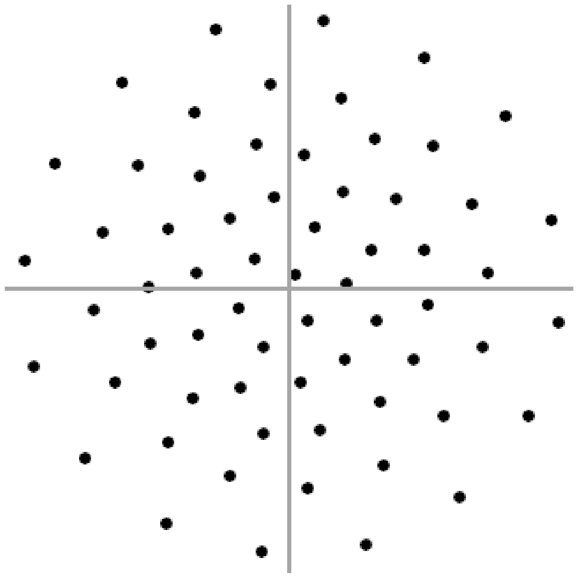


Order 64

Comparison of constellations

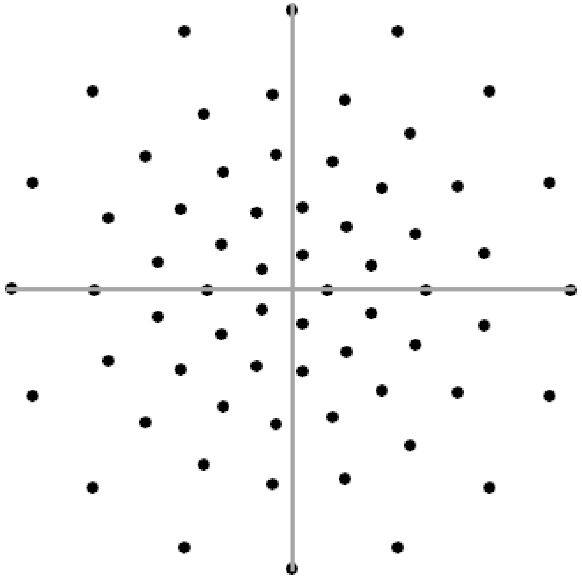


64 QAM



64 Auto-encoder

Trained via NLIN Model with optimized moments

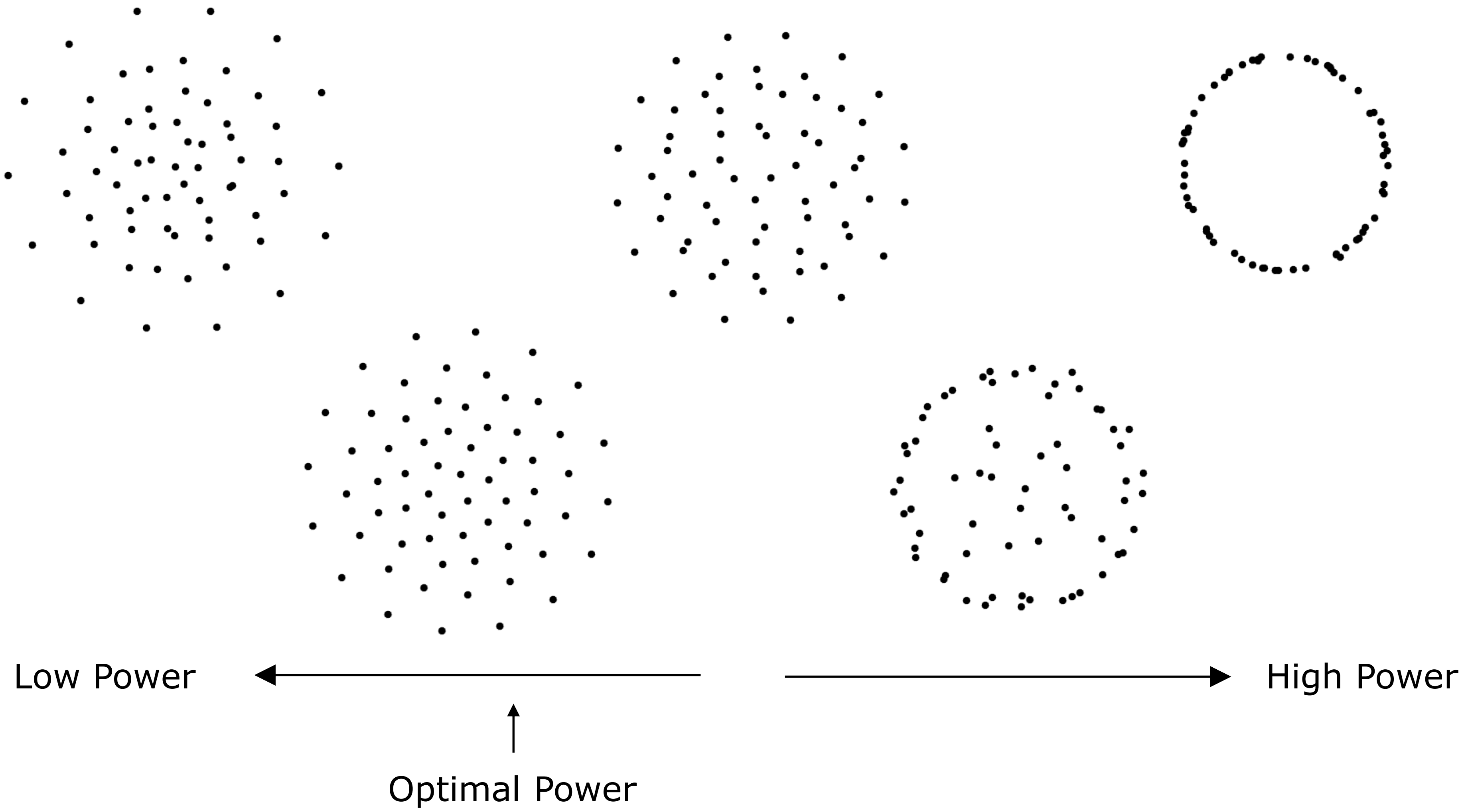


64 IPM[6]

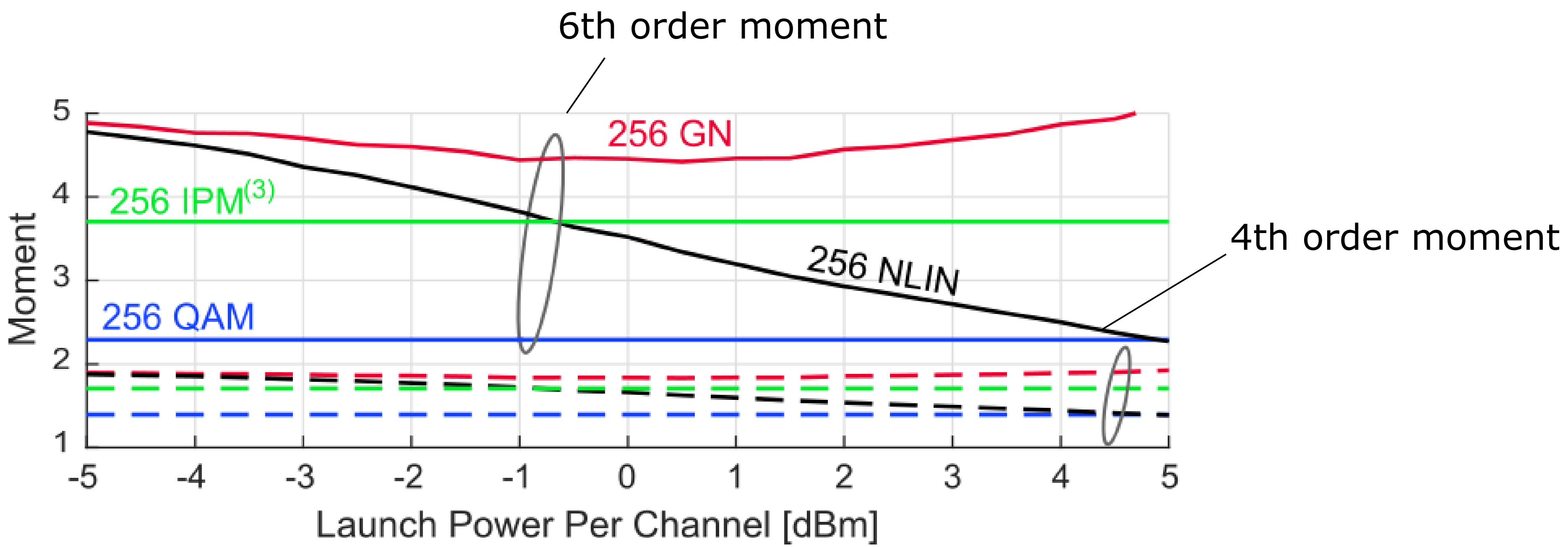
Gaussian channel assumption

[1] I. B. Djordjevic et al. ''Coded polarization-multiplexed iterative polar modulation (PM-IPM) for beyond 400 Gb/s serial optical transmission.'' OFC, paper OMK2, (2010).

Evolution of learned constellations

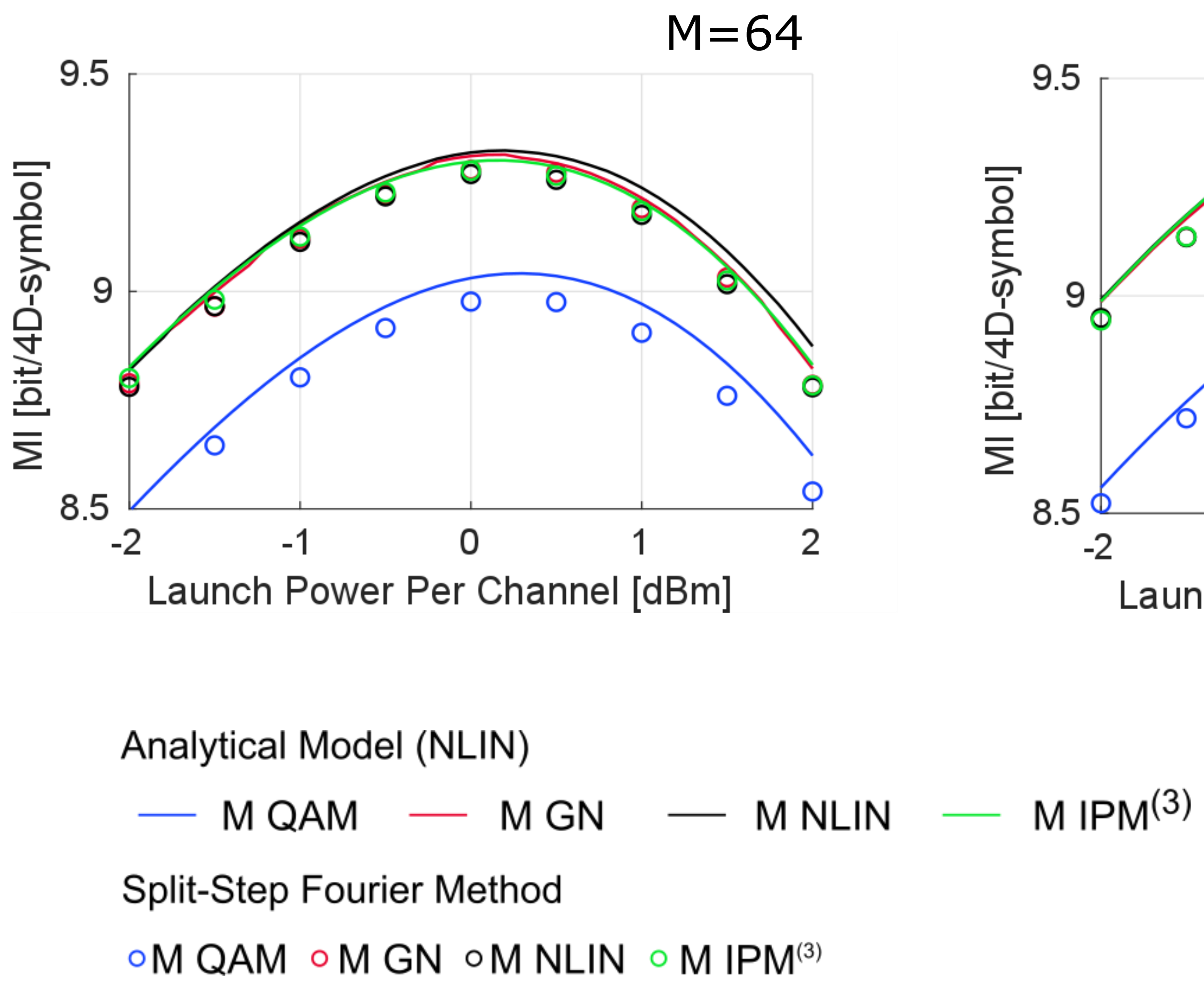


Scaling of the moments of learned constellations

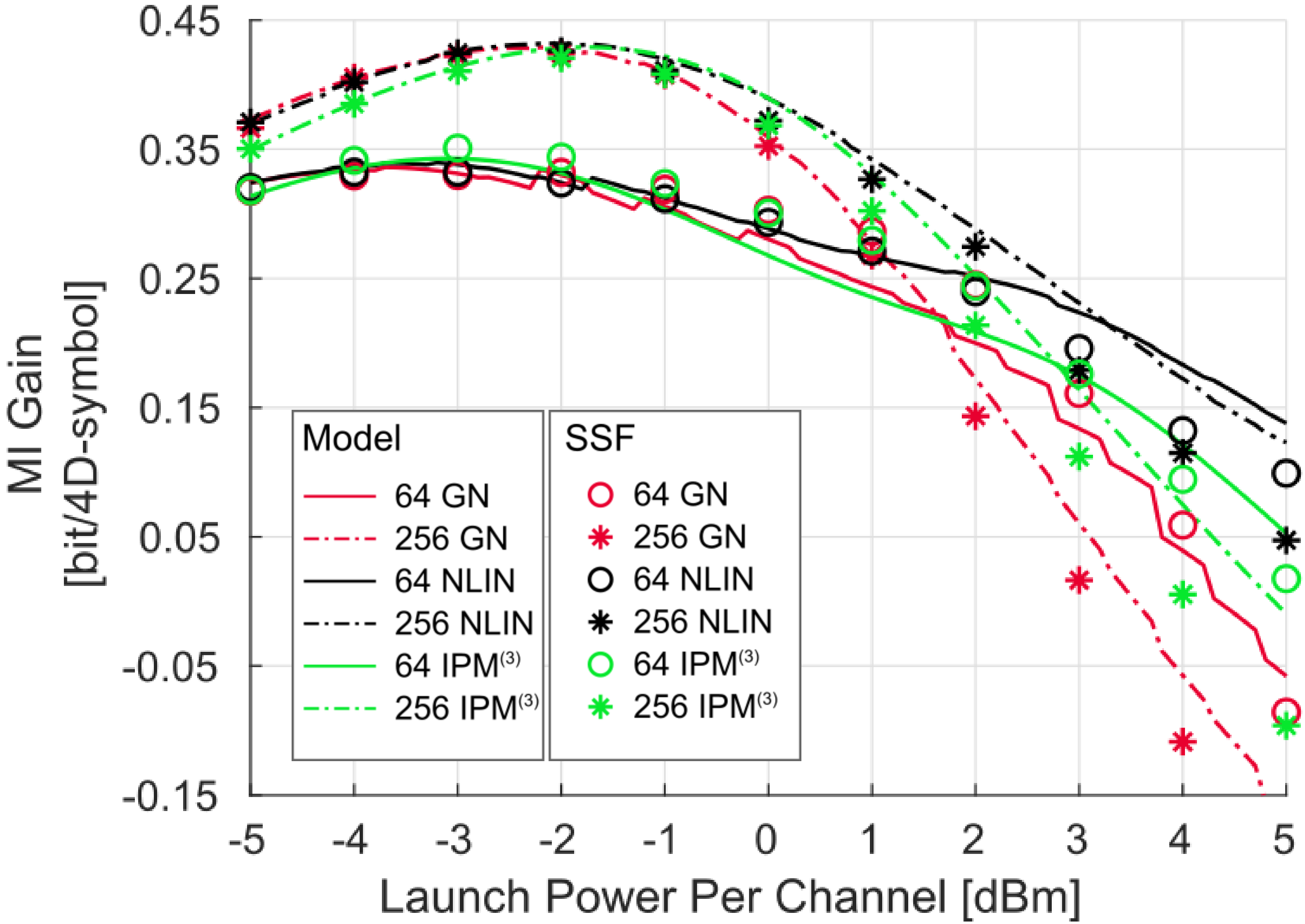


Auto-encoder learns constellation with reduced moment

Simulation results



Gains compared to the standard QAM



Conclusion and outlook

- Machine learning toolbox brings significant advantages to photonics

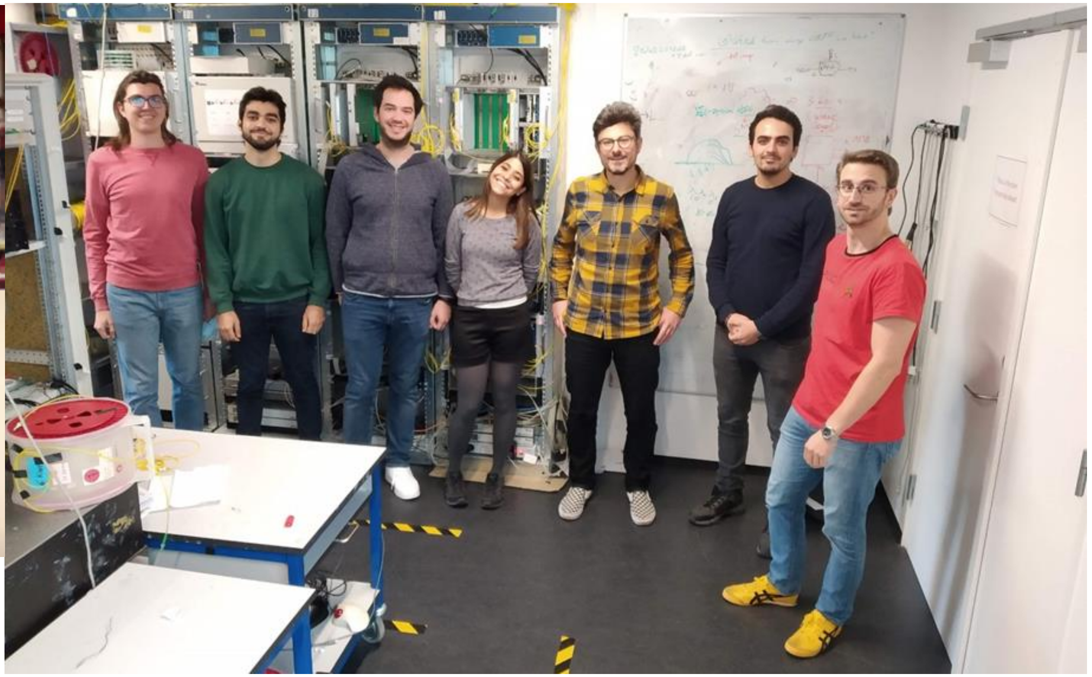
- Machine learning effective in learning complex mappings
  - Optical amplifier design
  - Communication over fiber-optic channel
  - Noise characterization of lasers and frequency combs
  - Quantum noise limited tracking

- Many other problems could benefit from ML (e.g. component design, power allocation etc)

- A lot of room for interesting research problems

- ML toolbox part of electrical and photonics engineering curriculum

- Lack of researchers that understand ML and optics to advance the field



Acknowledgements

Ministero dell'Istruzione, dell'Università e della Ricerca (PRIN 2017, Project FIRST).

CoG FRECOM (grant agreement no. 771878)

ofs Denmark
A Furukawa Company

THE VELUX FOUNDATIONS
VILLUM FONDEN VELUX FONDEN

Villum Foundations (VYI grant OPTIC-AI no.29344),