Tu.1.B.1 FEC

Neural Turbo Equalization to Mitigate Fiber Nonlinearity

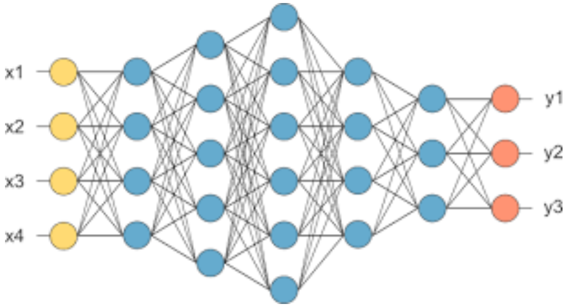
Toshiaki Koike-Akino, Ye Wang, David S. Millar,
Keisuke Kojima, Kieran Parsons

September 24, 2019

MITSUBISHI ELECTRIC RESEARCH LABORATORIES (MERL)
Cambridge, Massachusetts, USA
http://www.merl.com

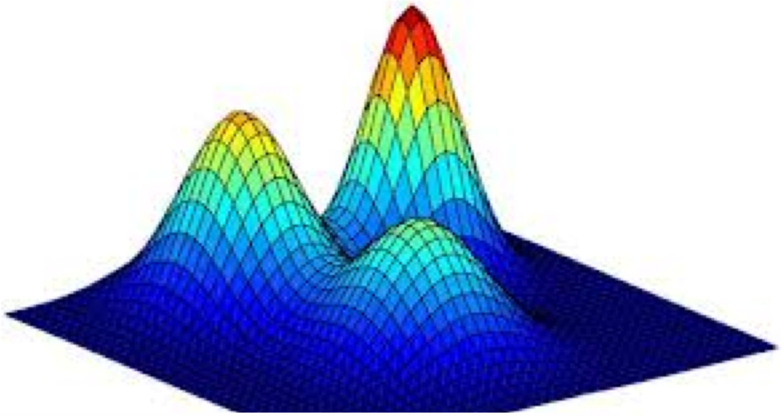
Outline

- Machine learning for optical communications
  - Research trend
  - Optics applications
  - Nonlinearity compensation
- Deep neural network (**DNN**) for DP-QAM
  - From maximum-likelihood to machine learning
  - Multi-label binary cross-entropy loss
  - Turbo equalization (**TEQ**)
- Performance analysis
  - Turbo feedback can reduce decoder iteration
  - **Up to 1.5dB gain** over conventional methods
- Summary

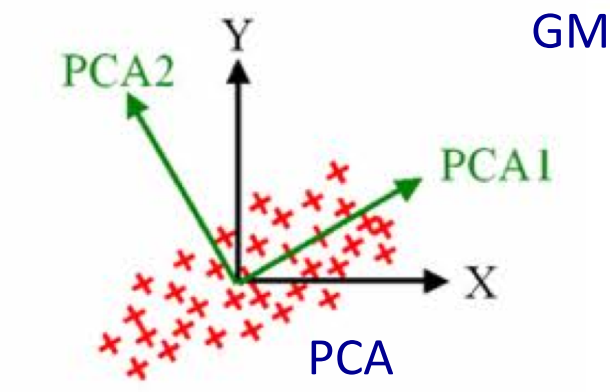


Machine Learning (ML)

- K-means
- Gaussian mixture model (GMM)
- Principal component analysis (PCA)
- Independent component analysis (ICA)
- Support vector machine (SVM)
- Self-organizing map (SOM)
- Hidden Markov model (HMM)
- Artificial neural networks (ANN)
- Deep learning (DL)
- …

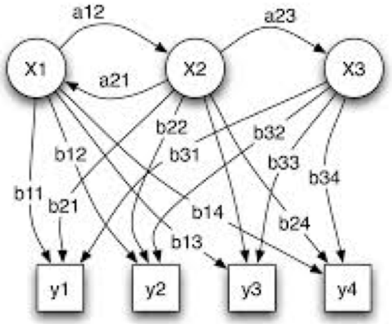


GMM

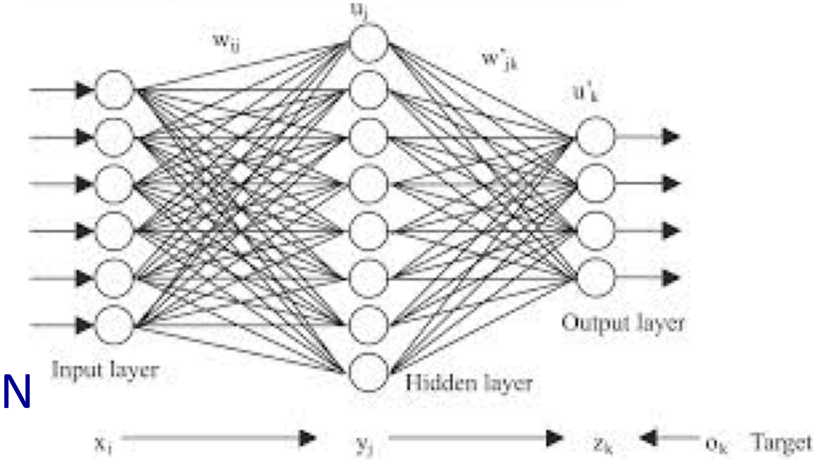


PCA

HMM

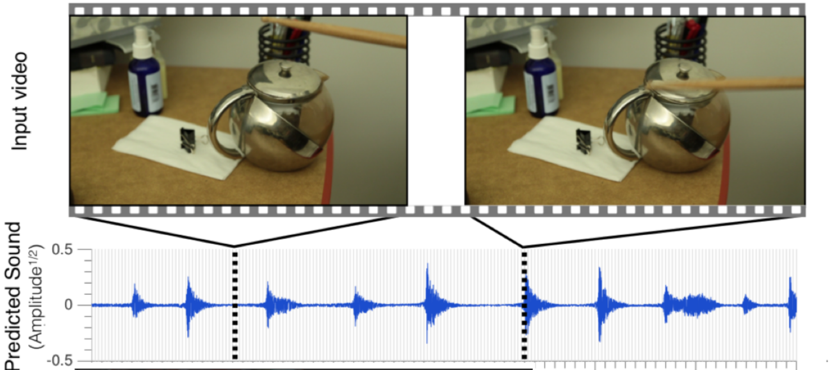


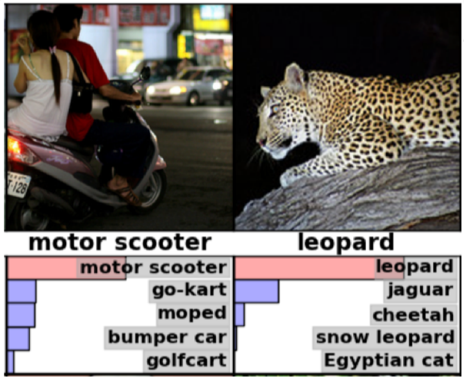
ANN



ML Success in Audio & Visual Signal Processing

- Denoising, segmentation, classification, translation, dialog, recognition, decomposition, generation, super-resolution, …





"man in black shirt is playing guitar."

ML surpassed Human-Level Performance

- For some applications, …

May 11th, 1997

Computer won world champion of chess

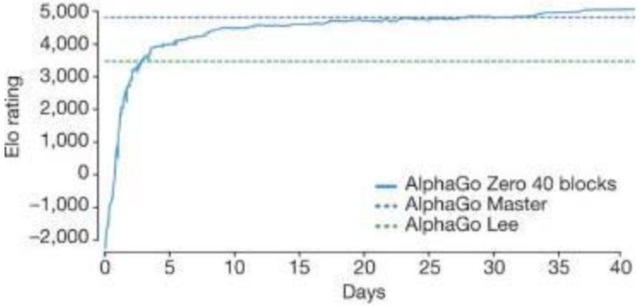
(Deep Blue) (Garry Kasparov)

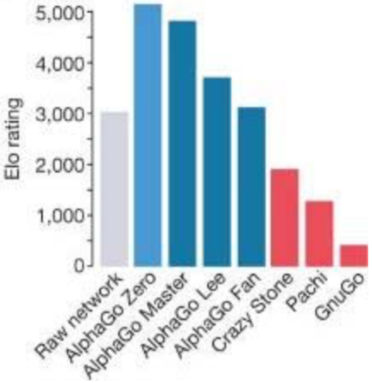
(Reuters = Kyodo News)

DARPA Grand Challenge

Autonomous Vehicle Races

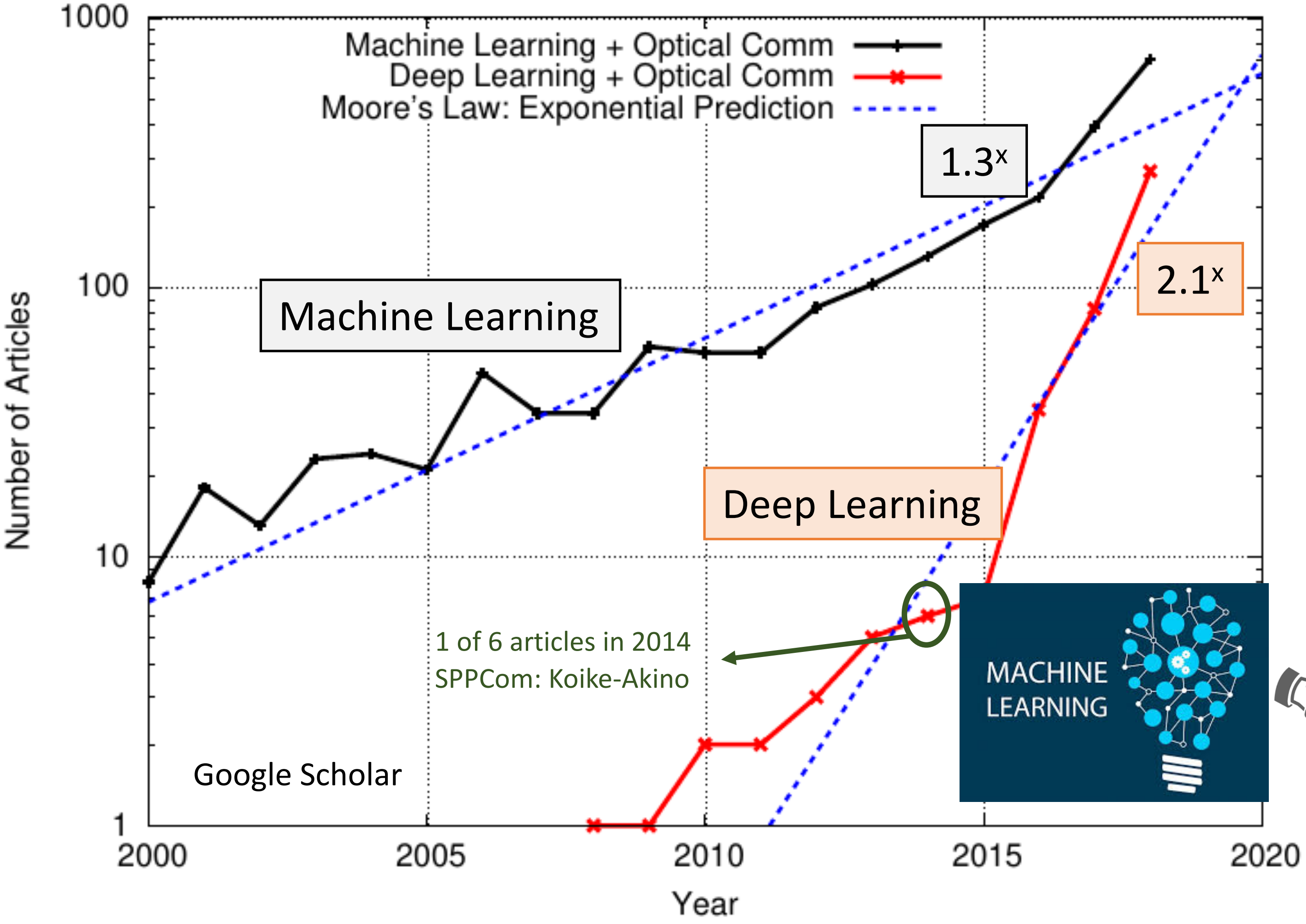
DGC I **Barstow to Primm** March 13, 2004	142 miles 10 hours $1M
DGC II **Desert Classic** October 8, 2005	132 miles 10 hours $2M
DGC III **Urban Challenge** November 3, 2007	60 miles 6 hours $3.5M





ML meets Optical Communications

- New **Moore's Law** rediscovered here:
  Number of articles grows exponentially, nearly **doubling** every year
  - Beyond 2020, thousands of publications per year will appear….



1.3^x

2.1^x

Machine Learning

Deep Learning

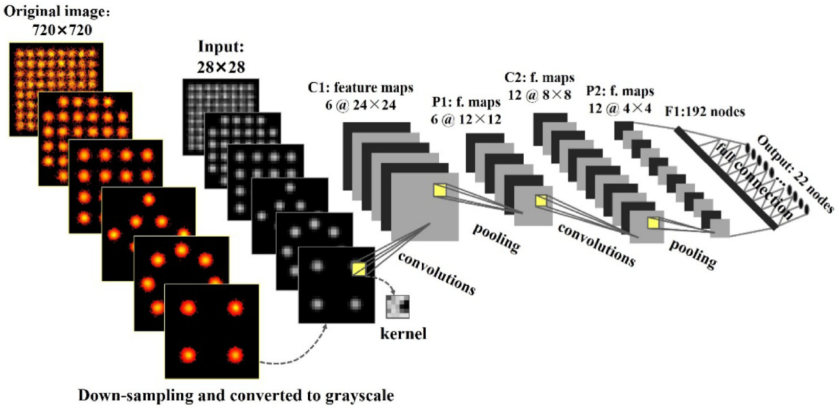
1 of 6 articles in 2014
SPPCom: Koike-Akino

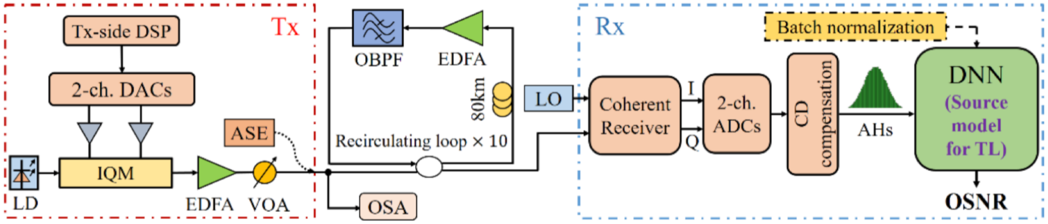
Google Scholar

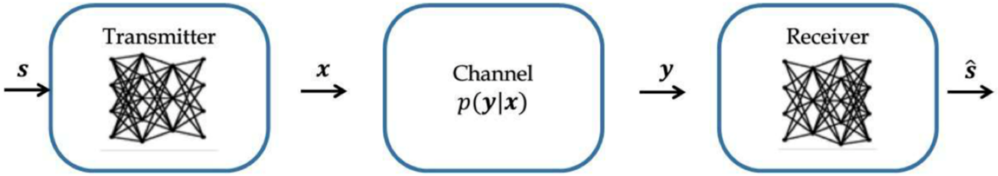
Deep Learning Applications for Optics

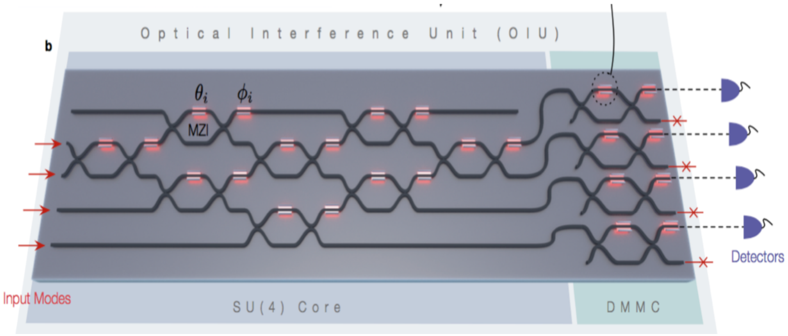
Already approx. 1000 related articles:

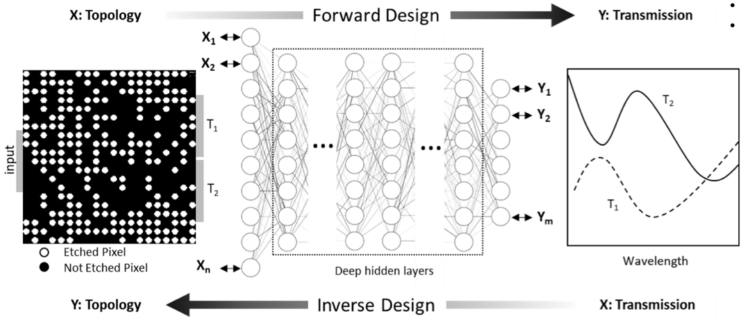
- Modulation classification
- Link quality monitoring
- Resource allocation
- End-to-end design
- Signal detection
- **Nonlinear compensation**
- Photonic circuit design
- Optical neural processor

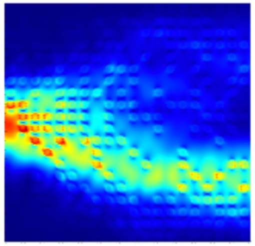












Why ML for Nonlinearity Compensation?

- Fiber channels are governed by **nonlinear physics in nature**
  - Self-phase modulation, cross-phase modulation, four-wave mixing, etc.
- Spectral efficiency can be improved by nonlinearity compensation
  - **Complicated model-based approaches** are required to capture real physics
- **Terabit-class massive data within a second** can be obtained
  - Deep learning: New **data-driven approach**. Suited for **massive parallel computing**

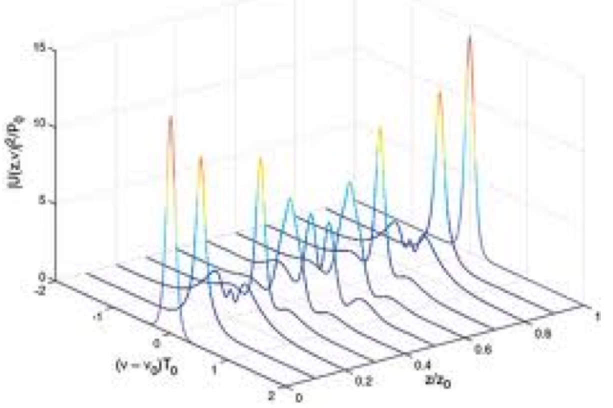
Nonlinear Schrodinger Equation:

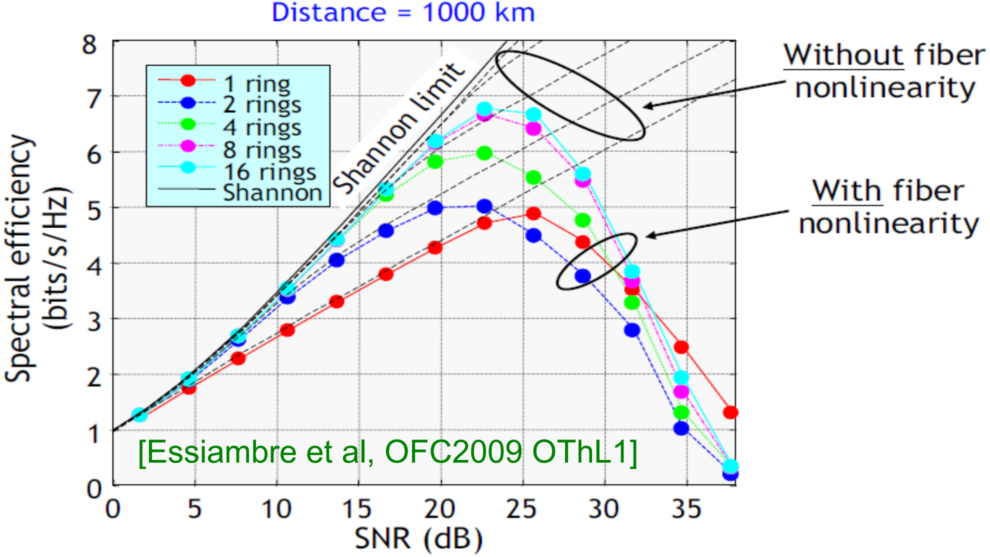
PMD CD (GVD) SPM/XPM

$$\frac{\partial \boldsymbol{E}}{\partial z} = \underbrace{\left(-\frac{1}{2}\boldsymbol{\alpha} - \boldsymbol{\beta}_1 \frac{\partial}{\partial t} - \jmath \boldsymbol{\beta}_2 \frac{1}{2!}\frac{\partial^2}{\partial t^2} + \frac{1}{3!}\boldsymbol{\beta}_3 \frac{\partial^3}{\partial t^3}\right)}_{\hat{\boldsymbol{D}}} \boldsymbol{E} + \jmath\gamma \underbrace{\left(\|\boldsymbol{E}\|^2 \boldsymbol{I} - \frac{1}{3}(\boldsymbol{E}^\dagger \boldsymbol{\sigma}_3 \boldsymbol{E})\boldsymbol{\sigma}_3\right)}_{\hat{\boldsymbol{N}}} \boldsymbol{E}$$

CD slope XPolM

Nonlinear propagation





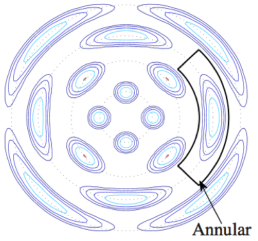
Without fiber nonlinearity

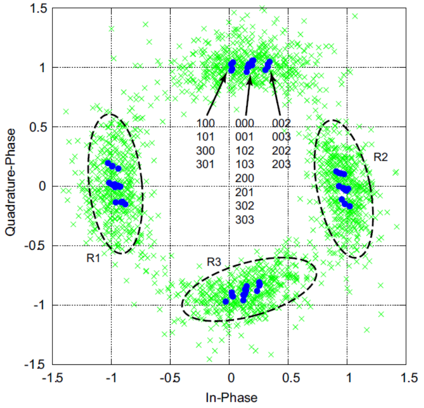
With fiber nonlinearity

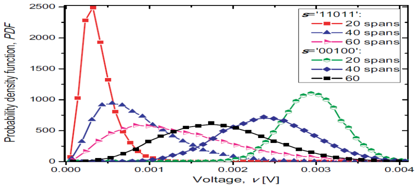
[Essiambre et al, OFC2009 OThL1]

Nonlinear Equalization

- Nonlinear impairments may be compensated by *nonlinear equalization:*
  - Decision feedback equalizer
  - Maximum-likelihood sequence equalizer
  - Volterra equalizer
  - Digital back-propagation
  - **Turbo equalizer (TEQ)**
  - **Deep neural networks (DNN)**



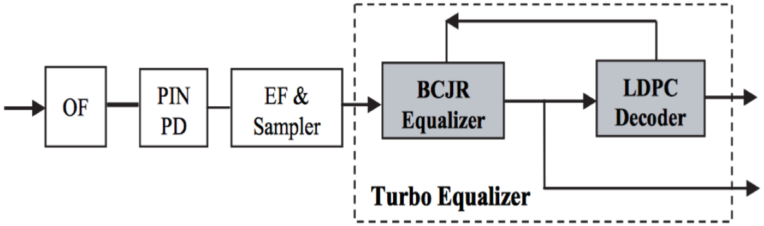




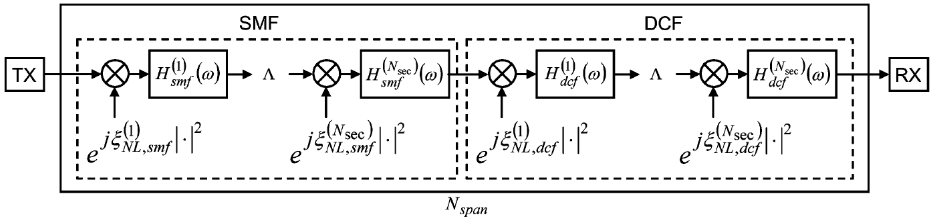
$$y(n) = \sum_{p=0}^{P} \sum_{l_1,\ldots,l_p=0}^{L_p} h(l_1,\ldots,l_p)\boxed{x(n-l_1)x^*(n-l_2)\cdots x(n-l_p)} + z(n)$$

Volterra series expansion

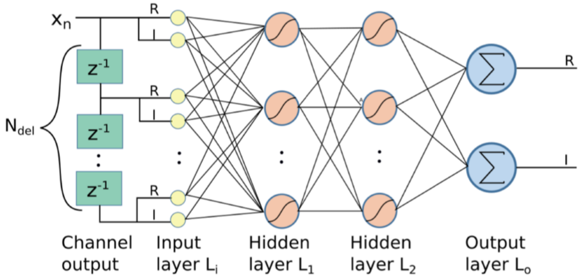
Volterra [Peddanarappagari '97]



TEQ [Haunstein '04, Djordjevic '07]



Digital back-propagation [Li et al '08, Ip-Kahn '08]



DNN [Sidelnikov '18, Koike-Akino '18, Kamalov '18]

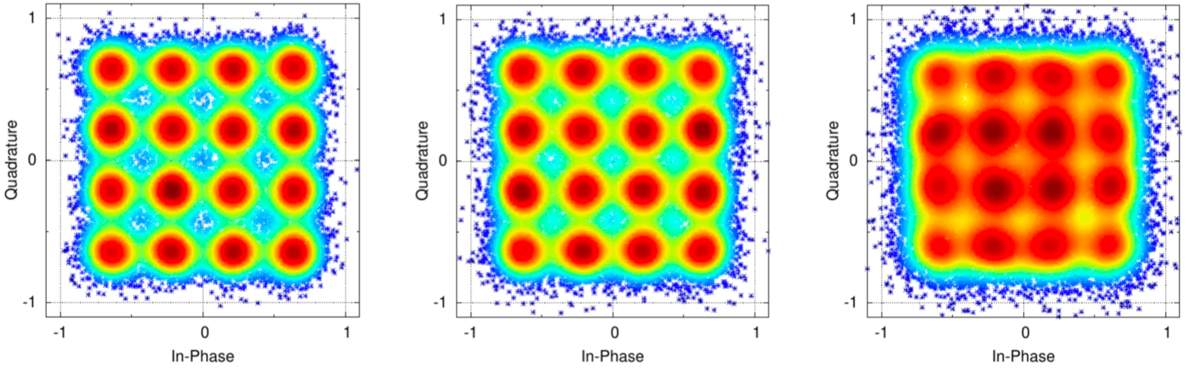
ML2ML: Maximum-Likelihood to Machine Learning

- Nonlinear equalization based on **maximum-likelihood (ML)**
  - Log-likelihood maximization, depending on nonlinear channel statistics

Maximum-Likelihood (ML)

$$\max_i \log \Pr(x_i|y)$$

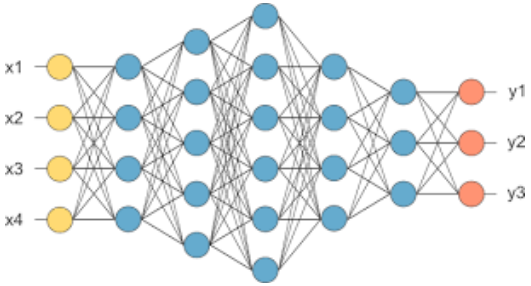
How to determine?
Model based?
Model mismatch?



(a) -5 dBm Launch (b) -3 dBm Launch (c) -1 dBm Launch

Post-Linear Equalization Distortion (16 spans)

- Cross-entropy minimization based on **machine learning (ML)**
  - Learning nonlinear channel statistics given data
  - Lower bound maximization of **GMI** (generalized mutual information)
  - Analogy to SSFM: sequence of **linear** transform and **nonlinear** operation

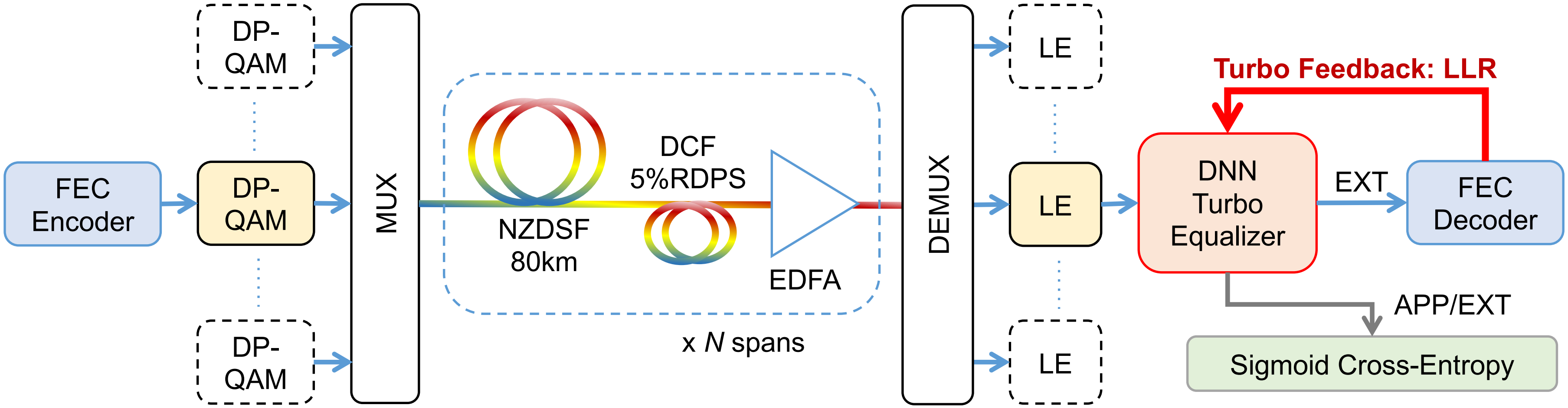


Binary cross entropy (BCE) corresponds to GMI

$$\mathbb{E}[\sum_i -\log \Pr(x_i|y)] \to 1 - \mathrm{GMI}$$

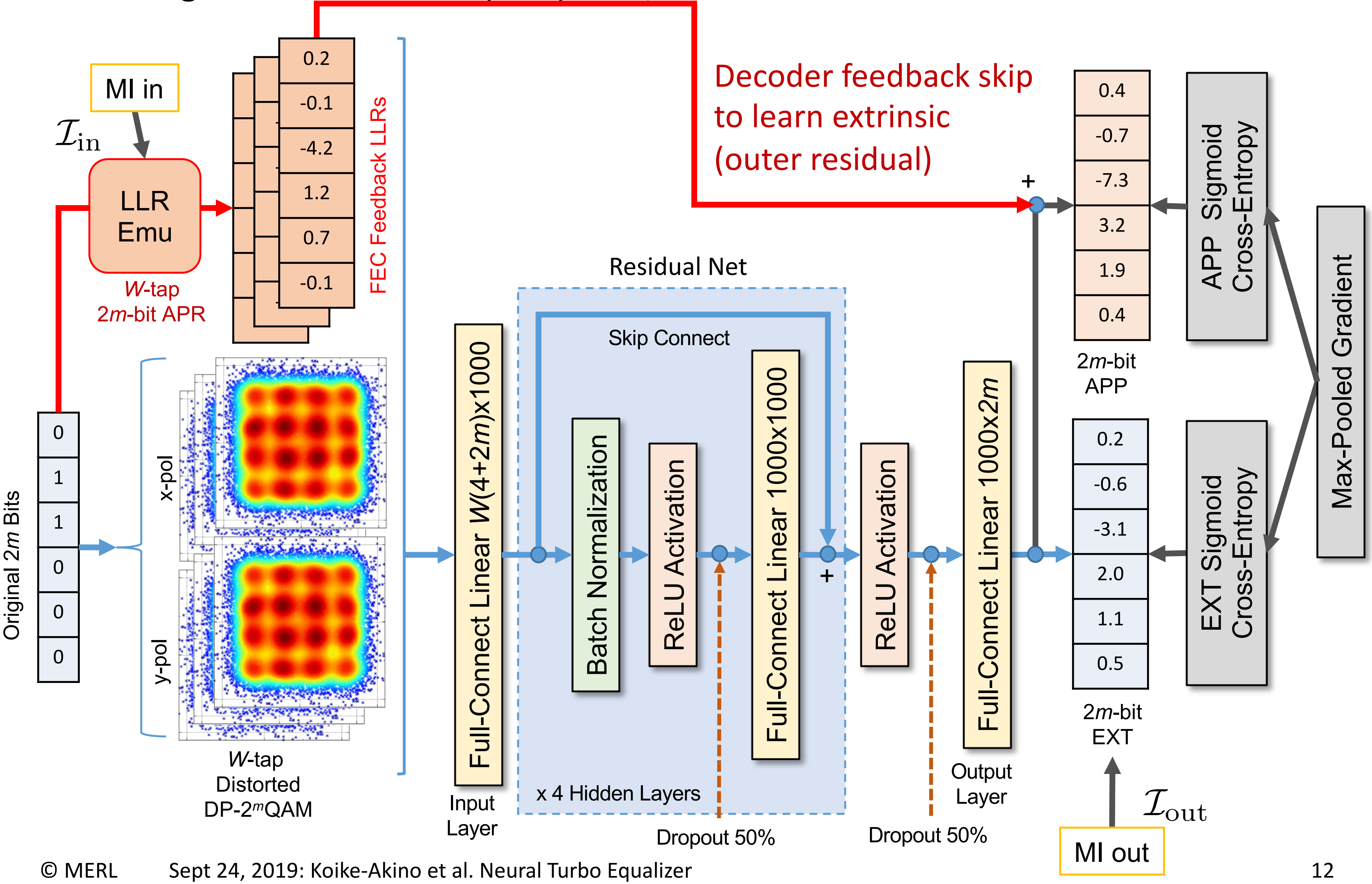
Proposed Method: DNN x TEQ

- We propose a new **TEQ based on DNN**
- We learn nonlinear statistics over 500,000 symbols on system model:
  - Non-zero dispersion-shifted fiber (NZDSF) 80km x *N* spans
    - 3.9ps/nm/km, 1.6/W/km, 0.2dB/km
  - 5% residual dispersion per span (RDPS)
  - Erbium-doped fiber amplifier (EDFA) 5dB noise figure
  - 3-channel DP-QAM at 34GBd, root-raised cosine role-off 0.1
  - DVB-S2 standard LDPC codes
  - 31-tap least-squares linear equalizer prior to DNN-TEQ



DNN-TEQ Architecture: Learning Residual Nets

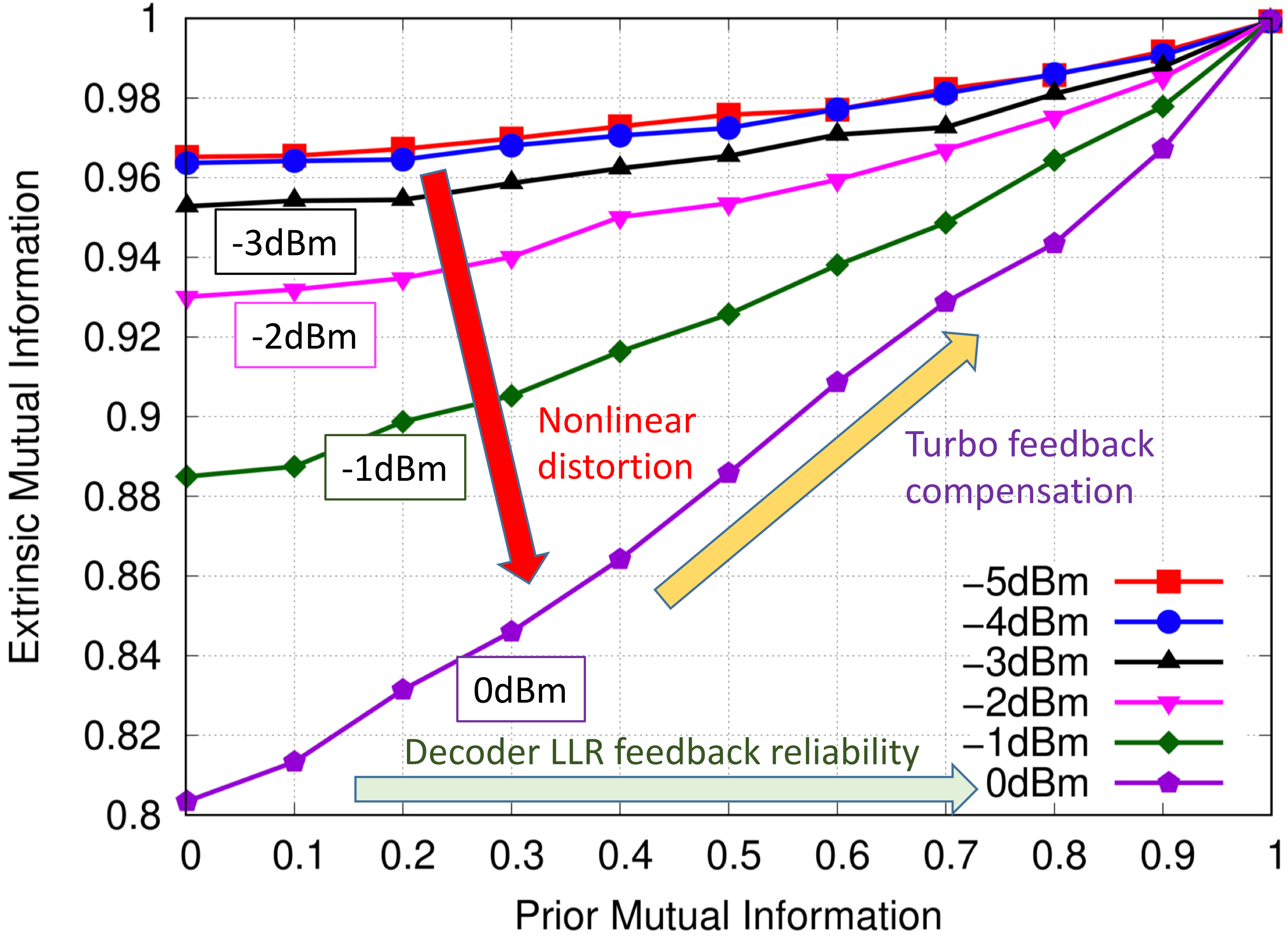
- Feeding Gaussian *A Priori* (APR) LLRs, emulated as decoder feedback



Decoder feedback skip to learn extrinsic (outer residual)

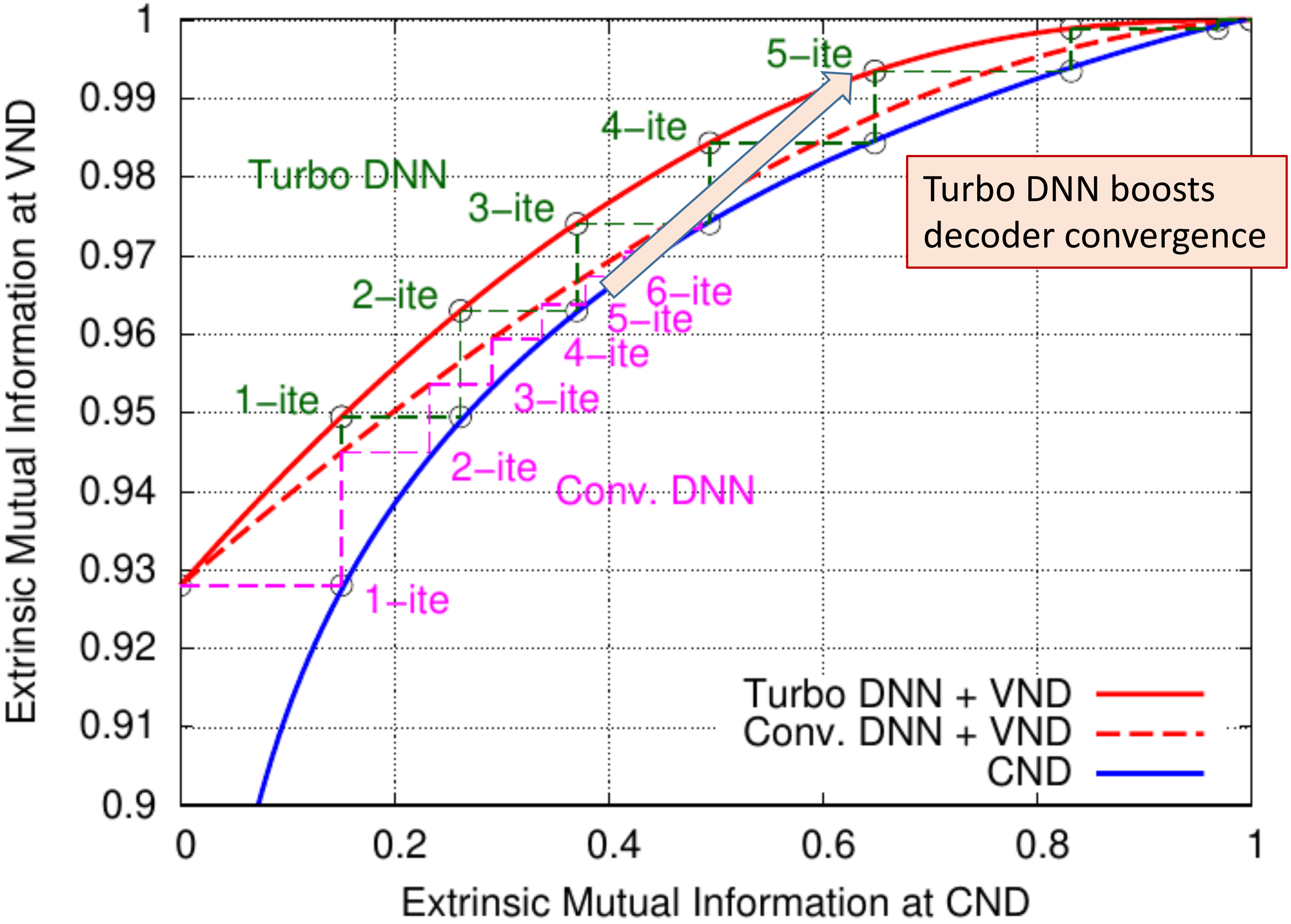
Extrinsic Information Transfer (EXIT) Analysis

- Trained DNN can produce improved LLR, given decoder feedback



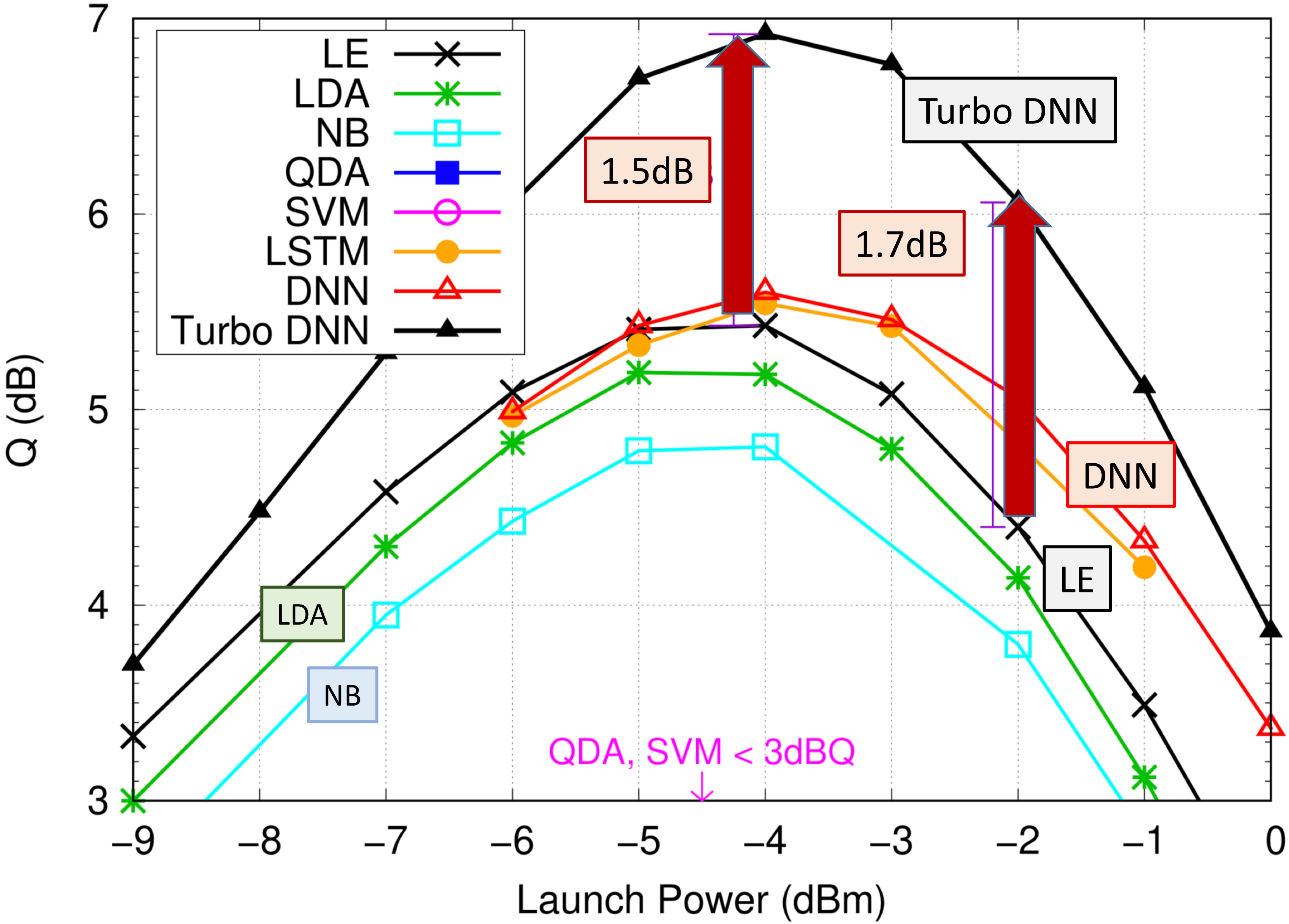
Combined EXIT Trajectory for DNN x LDPC

- EXIT of trained DNN, combined with LDPC decoder (-2dBm, 9/10 DVB-S2)



Turbo DNN boosts decoder convergence

Performance Evaluations (DP-64QAM, 4/5-LDPC)



Turbo DNN

1.5dB

1.7dB

DNN

LE

LDA

NB

QDA, SVM < 3dBQ

Summary

- We showed some perspectives of deep learning techniques for nonlinear optical fiber communications
  - Nonlinear fiber distortion may call for **nonlinear** signal processing
  - **Data-driven approach** can be a viable alternative to model-based approaches as massive data are available in high-speed optical transmission
- We proposed **DNN-based TEQ** which is scalable to high-order QAM
  - Turbo DNN can **accelerate** LDPC decoder iterations
  - Showing up to **1.3dB gain** over conventional DNN for DP-64QAM
  - DNN output can be directly used as extrinsic **LLR** for FEC decoder
- There are a great amount of open research fields to apply deep learning techniques to optical communications because of the nature of nonlinear physics

MITSUBISHI
ELECTRIC

Changes for the Better