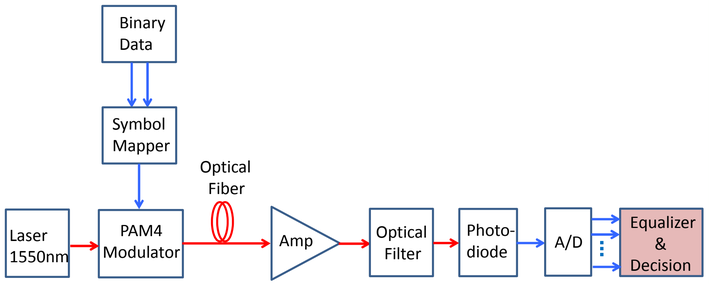
Machine Learning Equalization Techniques for High Speed PAM4 Fiber Optic Communication Systems

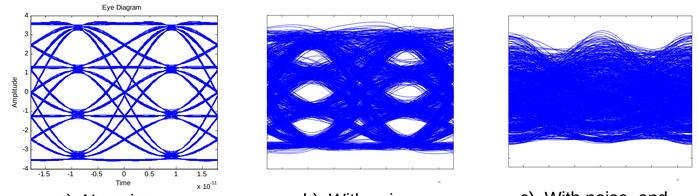
Ilya Lyubomirsky
CS229 Poster Presentation
Stanford University
8 December 2015

PAM4 Fiber Optic Communication System



We simulate a high-speed 112 Gb/s PAM4 fiber optic link, including the impact of fiber dispersion and amplifier noise. A/D samples the received waveform at 2 samples per symbol, and feeds N samples to the equalizer/decision circuit.

Simulated PAM4 Eye Diagrams at Receiver



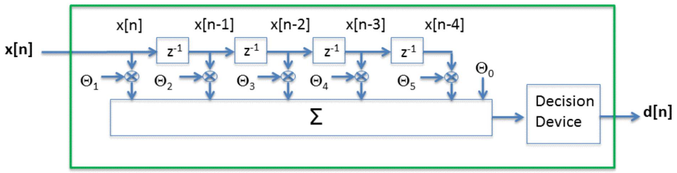
a). No noise, no dispersion

b). With noise, no dispersion

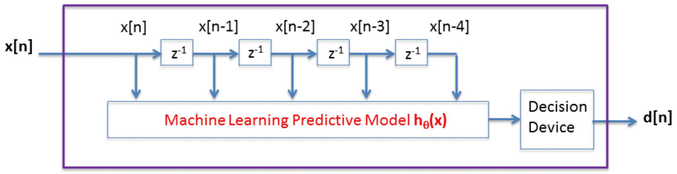
c). With noise, and dispersion ~4km SMF

An eye diagram provides a visualization on the impact of transmission impairments. Ideal PAM4 eye diagram should show "open" eyes, as in a) with no noise/dispersion. The addition of amplifier noise in b). closes the eyes, shows more noise on higher PAM4 levels due to square law detection. When both noise and dispersion are included in c)., the eye is completely closed and conventional receiver is not able to successfully distinguish the PAM4 symbols.

Machine Learning for Digital Equalization

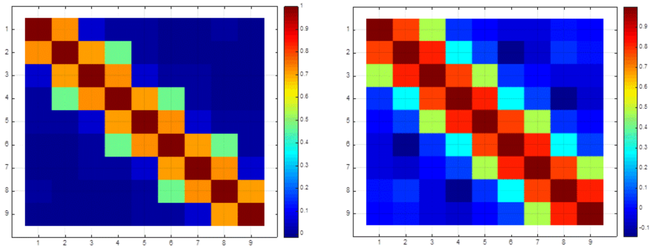


a). Conventional linear equalizer



b). Generalized "equalizer" based on machine learning

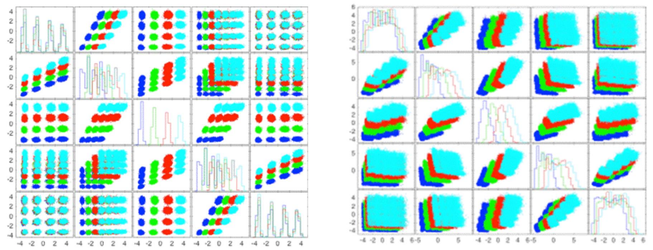
Feature Correlation Heat Map



a). Correlation plot for 9 samples, with noise but no dispersion. Note adjacent samples weakly correlated due to system bandwidth limitation.

b). Correlation plot for 9 samples, with noise and dispersion. Note adjacent samples now strongly correlated due to dispersion.

Pairwise Feature Scatter Diagrams



a). Pairwise scatter plot for 5 samples, with noise but no dispersion. The 4 PAM4 classes are labeled by color. In this case it's easy to separate the 4 classes.

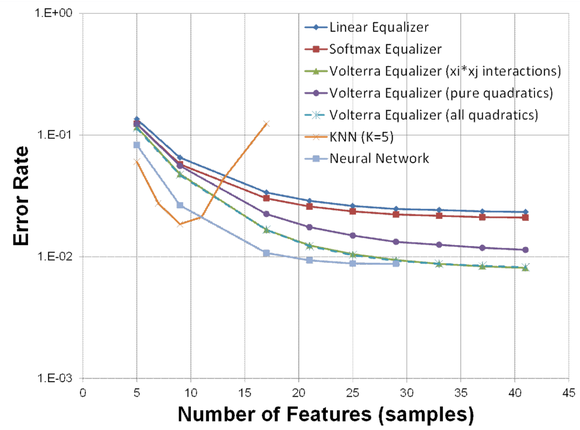
b). Pairwise scatter plot for 5 samples, with noise and dispersion included. The 4 PAM4 classes are labeled by color. In this case there is significant overlap; need nonlinear machine learning algorithms to separate the 4 classes.

Machine Learning Algorithms

All the algorithms are implemented using Matlab's Statistics and Machine Learning Toolbox:

- Linear equalization (or linear regression) with minimum MSE parameter fitting; using the **fitlm** function
- Volterra equalization (or polynomial regression) with minimum MSE parameter fitting; using the **fitlm** function
- Softmax regression with maximum likelihood estimate of parameters; using the **mnrfit** function with logit link function
- K nearest neighbors (KNN) classification; using **fitcknn** function with K optimized using a separate validation data set
- Neural Network, designed with 2 hidden layers, N neurons in each layer; using **feedforwardnet** function with training based on Levenberg-Marquardt algorithm

Monte-Carlo Simulation Results



Conclusions

- Machine learning algorithms for regression and classification may provide novel ideas for equalization and detection in digital communication systems
- We employ Monte-Carlo simulation of an 112 Gb/s PAM4 fiber optic communication system to compare several interesting machine learning algorithms in terms of test data set symbol error rate
- The simplest linear regression scheme shows relatively poor performance; polynomial regression with linear + interaction ($x_i x_j$) terms offers good performance with modest increase in complexity
- The most complex regression scheme based on Neural Networks gives excellent performance but with a high price in complexity
- Interestingly, the relatively simple KNN (K=5) gives the best performance for small number of samples ~ 9; KNN performs poorly as the feature space expands to > 13 dimensions due to the well known "curse of dimensionality"