Evolution from 8QAM live traffic to PS 64-QAM with Neural-Network Based Nonlinearity Compensation on 11000 km Open Subsea Cable

Valey Kamalov[1], Ljupcho Jovanovski[1], Vijay Vusirikala[1], Shaoliang Zhang[2], Fatih Yaman[2], Kohei Nakamura[3], Takanori Inoue[3], Eduardo Mateo[3], and Yoshihisa Inada[3]

[1]Google, Inc, Mountain View, CA 94043 USA
[2]NEC Laboratories America, Inc., Princeton, NJ 08540, USA
[3]Submarine Network Division, NEC Corporation, Japan
Emails: vkamalov@google.com

Abstract: We report on the evolution of the longest segment of FASTER cable at 11,017 km, with 8QAM transponders at 4b/s/Hz spectral efficiency (SE) in service. With offline testing, 6 b/s/Hz is further demonstrated using probabilistically shaped 64QAM, and a novel, low complexity nonlinearity compensation technique based on generating a black-box model of the transmission by training an artificial neural network, resulting in the largest SE-distance product 66,102 b/s/Hz-km over live-traffic carrying cable.

OCIS codes: (060.2330) Fiber optics communications; (060.1660) Coherent communications

1. Introduction

With the introduction of coherent technology, legacy submarine cables underwent massive upgrade cycles. Rapid development of coherent technology enables evolution of modern cables too, as was demonstrated for the newly built FASTER cable, a Trans-Pacific cable designed specifically for coherent transponders [1]. The longest segment of FASTER cable that connects Tanshui, Taiwan to Bandon, USA (TWN-USA) was initially designed at spectral efficiency (SE) of 2 b/s/Hz, and within few months upgraded to 3 b/s/Hz. Here we report that, less than 6 months later, a second upgrade increased the SE to 4 b/s/Hz using dual-carrier 300 Gb 8QAM modulation, carrying production traffic in Google data center network.

Considering the quick pace of upgrades, it is important to have an accurate estimate of how far the modern Trans-Pacific cables can be upgraded. Offline field trials over dark fibers in new build cable with Trans-Pacific distance showed that the SE can be pushed further to 5.68 b/s/Hz [2], while measurements over straight line test bed showed 6.06 b/s/Hz over 10285 km [3] using truncated probabilistic shaped (PS)-64QAM and nonlinearity compensation (NLC). Our offline testing over FASTER cable with PS-64QAM confirmed that 6 b/s/Hz is achievable with NLC, resulting in the highest SE-distance product over live-traffic carrying cable, 66,102 b/s/Hz-km. Considering that NLC complexity per symbol is expected to grow with increasing baud-rates, especially at transoceanic distances [4] we propose and demonstrate a novel, low-complexity, artificial-intelligence based NLC (AI-NLC) algorithm. The technique is based on building a black-box model of the transmission link using deep learning algorithms. As such, this is the first demonstration of a data-driven, working model of an open submarine cable used for the purpose of computationally efficient NLC.

2. Real-time 300G Transmission with 4 b/s/Hz SE

FASTER holds the record for carrying live traffic over the longest distance (TWN-USA) with 8QAM, at a SE of 3.0 b/s/Hz. Figure 1a shows the system setup for the 300Gb/s upgrade pushing the SE to 4.0 b/s/Hz by using live-traffic dual-carrier 300 Gb/s transponders (TPND) at a channel spacing of 75 GHz. Figure 1b shows > 0.5dB Q margin over the FEC limit, equivalent to 0.7 dB SNR margin for the four 300G channels at the nominal channel power, and the long-term Q-stability for the center 300G channels are plotted in Fig.1c with standard deviation <0.015dB of Q for both subcarriers.

3. Improved SE and Data Rate Adaptation via PS-64QAM

PS-64QAM is recognized as a Shannon capacity limit approaching modulation with a straightforward implementation of flexible data rate [5]. Performance of PS-64QAM over FASTER cable is investigated in terms of achievable Q margin in the SE range 5 b/s/Hz to 6 b/s/Hz. Figure 1a, shows the schematic of the transmitter. The 49 Gbaud center probe signal is produced by a DP-IQ-modulator driven by 92 Gsa/s AWG. A constant modulus distribution matcher [6] is applied to the pseudo-random binary bits of 2^{31}-1 to match the target data rate of PS-64QAM given by $R = 2(H_p\text{-}6*(1\text{-}R_c))$ [5]. A fixed regular binary QC-LDPC code rate R_c=0.8 is used. The entropy, H_p, at 250 Gb/s, 275 Gb/s and 300 Gb/s data rate is 3.83 b/s, 4.091 b/s and 4.354 b/s, respectively. The four

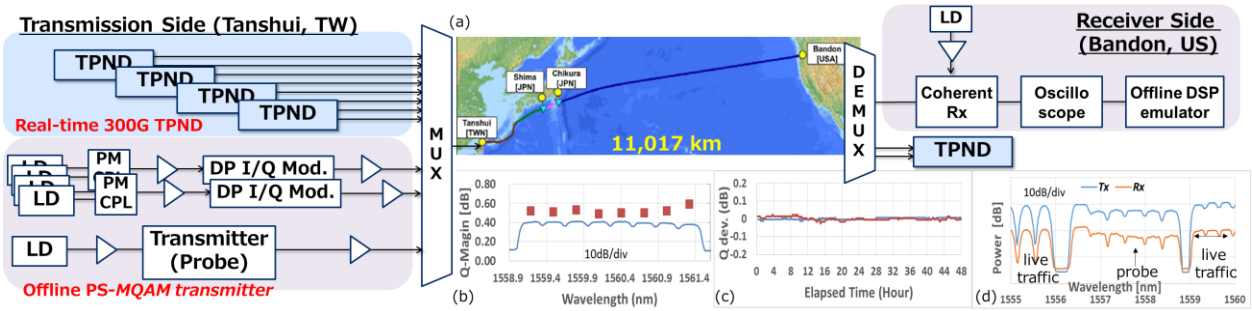

Fig.1 (a) The transmitter and receiver setup for real-time dual-carrier 8QAM/300G transponders (TPND) and offline PS-MQAM testing. (b) the Q margin and received spectra of four 300G real-time channels; (c) long-term Q-stability at the nominal channel power; (d) the spectra of offline signals.

neighboring 50GHz WDM odd/even channels are generated by two sets of independent AWG and DP-IQ-modulator.

Figure 1d shows the signal spectra at the transmitter and receiver side for the offline setup. The offline DSP uses 3% pilots to assist convergence of polarization de-multiplexing and carrier phase recovery before switching to decision-directed mode between the pilot frames. The blindly recovered data is fed to LDPC decoder and then distribution dematcher to ensure error-free binary bits stream.

Q-factor performance of PS-64QAM is plotted in Fig.2 against relative channel power with respect to the nominal channel power, together with the FEC limit at 5.0 dB [7]. At the nominal channel power, PS-64QAM achieves 5.5 b/s/Hz SE with 0.45 dB margin without NLC, and 6 b/s/Hz with 0.25 dB margin with NLC. In Fig.2, NLC is implemented using digital back propagation (DBP) method with 40 steps/span [4], to estimate the maximum possible improvement from NLC. It also emerges from the Fig.2 that, roughly a 1dB drop is expected in Q margin as SE increases by 0.5 b/s/Hz. The granularity of 25 Gb/s data rate is very appealing to maximize the cable capacity without significantly increasing cost to combine client data streams. The recovered constellation after 11017 km is plotted in Fig.3.

4. Single-Step AI-NLC and Transmission Results

Most NLC methods are based on solving or approximating the solutions of the nonlinear Schrodinger equation with a trade-off between the required complexity and degree of nonlinearity compensation [4]. An alternative approach is to set aside the deterministic model of the transmission, and instead use the abundance of transmitted data to build a simpler yet effective model of the nonlinear transmission. We built such a model by training a deep neural network (DNN) shown in Fig. 4 with only 2 hidden layers consisting of 2, and 10 nodes, where SELU() activation function is applied at both hidden layers [8]. A dropout layer with probability of 0.5 is placed after the 2nd hidden layer during training only to avoid overfitting. The network is trained by transmitting known but randomly generated patterns, and searching for the best node tensor parameters that minimize the mean square error between the transmitted and received symbols. A major differentiation in our approach compared to previous attempts that produced negligible benefits [9], or required large over-heads [10], is the inclusion of intra-channel cross-phase (IXPM) modulation and intra-channel four-wave mixing (IFWM) triplets as defined in time-domain perturbation pre/post-distortion (PPD) algorithm [11] in addition to the received soft symbols H and V in both polarizations with a symbol window L centered around the symbol of interest H_0 or V_0. The triplets provide the network with

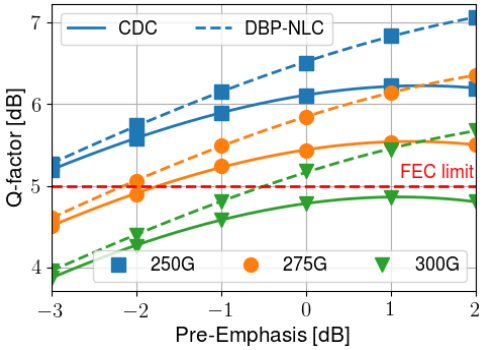

Fig.2 The Q-factor of PS-64QAM as a function of channel power with CDC only (solid), and NLC (dashed).

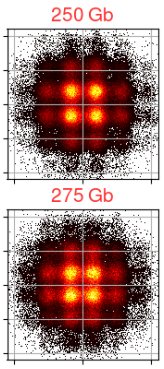

Fig. 3 The recovered constellations with CDC only.

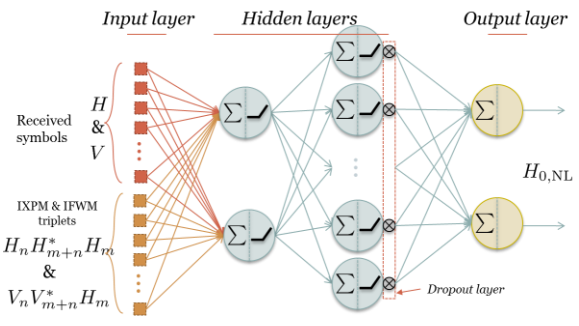

Fig.4 The block diagram of the proposed AI –NLC with 2 hidden layers. 2 output nodes represent real and imaginary part of the estimated nonlinear perturbation $H_{0,NL}$.

underlying physical processes that govern the nonlinear impairments. Note that all the triplets and received symbols are separated into real and imaginary parts before being fed into the DNN to estimate the nonlinear perturbation $H_{0,NL}$ in the received signal H_0 in H-polarization. As a result, the AI-NLC signal becomes $H_0 - \alpha H_{0,NL}$, where α is a scaling factor corresponding to operating channel power with respect to the trained channel power where α = 1. The same methodology is also applied to the V-polarization. The number of triplets N_t is selected based on the criterion $20\log_{10}|C_{mn}/C_{00}|$, where C_{mn} represents the analytical coefficients in the PPD algorithm for the triplets between symbol m and n [12].

For testing the AI-NLC, digital subcarrier modulation (DSM) 4×12.25Gbaud PS-64QAM at RRC 0.01 with 50MHz guard band is used instead of a single-carrier modulation to carry in total 300Gb/s. The lower baud rate of DSM reduces the complexity of the AI-NLC further by reducing the number of IXPM and IFWM triplets [11,12]. DDN is trained using a total of 80,000 random symbols per subcarrier transmitted over the FASTER cable at 2 dB above the nominal channel power. After convergence of the DNN nodes, the same DDN is used to compensate all four subcarriers, at all power levels. For testing, 7 independent sets, each containing 40,000 symbols per subcarrier, are transmitted. The experimental GMI averaged over all four subcarriers in the independent test sets are plotted in Fig.5 with respect to the relative channel power. At the nominal power, AI-NLC improves GMI by 0.15 b/s/2-pol at a much lower complexity compared to 0.2 b/s/2-pol by full DBP, while receiver-side PPD [13] achieves none. All the subcarriers are recovered error-free after LDPC decoding after applying AI-NLC.

Since PPD is more effective at the transmitter side, AI-NLC is compared to Tx-side PPD through simulations over 11,017 km. Figure 6a shows the simulated Q-factor vs channel power for a single channel (SC) and WDM 37×12.25 Gbaud DP-16QAM at 12.5GHz spacing. At the optimum channel power, AI-NLC outperforms CDC-only and Tx-side PPD by 0.8dB and 0.3dB at SC, and 0.4dB and 0.3 dB at WDM, respectively. Since the major complexity of both AI-NLC and PPD are determined by the number of triplets, their Q-improvement is plotted in Fig.6b against the selection criterion $20\log_{10}|C_{mn}/C_{00}|$. At SC case, -30dB selection criterion is required in Tx-side PPD to achieve ~0.5dB Q improvement at the cost of ~1800 triplets. In contrast, AI-NLC only needs ~730 triplets to achieve ~0.8dB NLC gain thanks to the self-learning algorithm in the DNN. The symbol window length L = 51 in both experimental and simulation. The large gain in the simulation results confirms that the DDN is equalizing nonlinear penalty, rather than other unforeseen experimental transmission related impairments.

Since the AI model is data driven, it is completely system agnostic. Indeed, in this experiment the model was generated without feeding any of the typical transmission link parameters such as dispersion, fiber nonlinearity, fiber length, etc., confirming the versatility, and adaptability of the method. The training is performed once at a fixed channel power. Same DNN coefficients operated close to optimum for other power levels, requiring adjustment of a single parameter α.

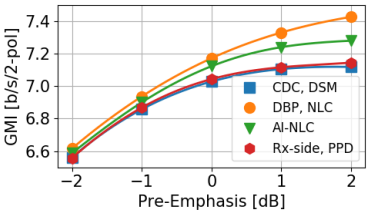

Fig.5 The experimental GMI of DSM PS-64QAM with CDC, DBP-NLC, AI-NLC and Rx-side PPD.

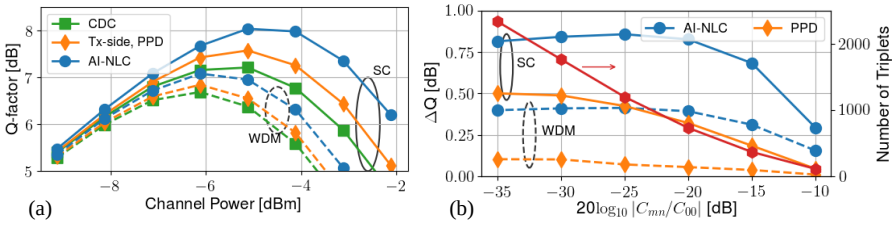

Fig. 6 (a) Q-factor vs channel power. (b) NLC gain comparison against the number of triplets in simulated single-channel and WDM 12.25Gbaud 16QAM between CDC, AI-NLC and Tx-side PPD

5. Conclusion

The longest 11,017 km segment of FASTER open cable connecting Tanshui, Taiwan and Bandon, USA, is upgraded to double the design capacity with 4b/s/Hz spectral efficiency (SE), carrying production traffic based on 8QAM. Experimental investigation based on the offline field trial shows realistic target of SE= 6 b/s/Hz for Trans-Pacific cables with probabilistically shaped 64QAM. A data-driven nonlinear model of the cable is generated and tested through a deep-neural network architecture. We achieved what we believe today is the highest SE-distance product of 66,102 b/s/Hz-km over live-traffic carrying cable.

References

[1] V. Kamalov et al., Proc. OFC'17, M2E.2.
[2] J. Cho et al., Proc. OFC'17, PDP Th5B.3
[3] I. Ruiz et al., J. Lightw.Technol. 36, 1354-1361 (2018)
[4] E. Ip and J. M. Kahn, J.Lightw. Technol. 26, 3416–3425 (2008).
[5] F. Buchali et al., J. Lightwave Technol. 34, 1599-1609 (2016)
[6] P. Schulte and G. Bocherer, IEEE Trans. Inf. Theory 62, 430 (2016)
[7] S. Zhang and F. Yaman, J. Lightwave Technol. 36, 416-423 (2018)
[8] G. Klambauer et al., arXiv preprint arXiv:1706.02515, 2017.
[9] D. Zibar et al., Opt. Express 20, B181-B196 (2012)
[10] E. Giacoumidis et al., J. Lightwave Technol. 36, 721-727 (2018)
[11] Z. Tao et al., J. Lightw. Technol. 29, 2570-2576 (2011)
[12] Y. Gao et al., Opt. Express 22, 1209-1219 (2014)
[13] T.Oyama et al., OFC'14, Tu3A.3