Tailored Shaping, Improved Detection, Simpler Backpropagation: the Road to Nonlinearity Mitigation

Marco Secondini[1,2], Stella Civelli [1,2], Enrico Forestieri[1,2],

[1] Tecip Institute, Scuola Superiore Sant'anna marco.secondini@santannapisa.it
[2] PNTLab, Consorzio nazionale interuniversitario per le telecomunicazioni (CNIT)

Abstract *Several strategies for nonlinearity mitigation based on signal processing at the transmitter and/or receiver side are analyzed and their effectiveness is discussed. Improved capacity lower bounds based on their combination are presented.*

Introduction

Study and mitigation of nonlinear effects have been an important aspect of fiber-optic communication since the very beginning[1]. Aided by digital signal processing (DSP), coherent detection allows for optical systems with almost unlimited capability to modulate, demodulate, and process optical signals, improving spectral efficiency and pushing optical networks toward a seeming capacity limit due to nonlinear effects[2],[3]. This has further stimulated the research of DSP techniques to overcome fiber nonlinearity limitations[4].

The problem of mitigating nonlinear effects and improving system performance has been addressed from many different perspectives. In this work, only digital techniques to be implemented at the transmitter (TX) or receiver (RX) are considered. Therefore, we do not account for techniques that modify the fiber link or work at optical level, such as regeneration or phase conjugation.

Following[5], we consider the system described in Fig. 1. Fiber propagation is governed by the Manakov equation[6] and includes periodic in-line amplification and the simultaneous propagation of other WDM channels, all independently modulated and detected and with the same input distribution[7]. Collecting input and output symbols into the vectors $\mathbf{x}$ and $\mathbf{y}$, respectively, we try to maximize the achievable information rate (AIR)[5] by optimizing the input distribution $p(\mathbf{x})$ (blue block), the detection metric $q(\mathbf{y}|\mathbf{x})$ (pink block), and TX/RX processing (yellow blocks). Moreover, we discuss the main modulation/demodulation and coding/decoding techniques to implement such an optimized system.

Constellation shaping

Constellation shaping improves the efficiency of a digital modulation scheme by modifying the position of the symbols in the constellation diagram

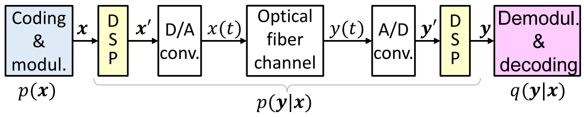

Fig. 1: Description of the system considered for AIR computation and maximization.

(geometric shaping) or the frequency with which they are used (probabilistic shaping). With reference to Fig. 1, it consists in optimizing the input distribution $p(\mathbf{x})$ (or its support) and devising a proper coded modulation scheme to encode information accordingly. On the AWGN channel, the problem is well known: the optimal distribution factorizes into the product of identical marginal distributions (i.i.d. symbols)—Gaussian in the general case[8], Maxwell–Boltzmann (MB) if the symbols are constrained on a given discrete alphabet[9]—which minimize the energy per symbol required to achieve a certain information rate. In this context, a practical coded modulation scheme that has attracted much interest in recent years is probabilistic amplitude shaping (PAS), thanks to its nearly optimal performance, simple implementation, and fine rate granularity[10],[11]. PAS uses a distribution matcher, followed by a systematic FEC encoder, to induce the desired distribution over a QAM constellation. The optimal condition of i.i.d. MB symbols is approached as the block length of the distribution matcher goes to infinity[10],[12].

Constellation shaping can be used also to mitigate nonlinear effects. In this case, often referred to as nonlinear constellation shaping, the location or probability of the constellation symbols are optimized to minimize the amount or impact of the generated nonlinear interference (NLI). There are many evidences suggesting that optimizing the marginal distribution of i.i.d. 2D symbols yields negligible gains in this case[13]. In fact, to unlock the full potentiality of nonlinear constellation shaping, the optimization should be performed in a higher dimensional space. So far, the approaches have been limited to the optimization of low-rate constellations in a low-dimensional space (e.g., geometric shaping in 4D and 8D[14],[15]), or to a highly constrained optimization of PAS in a higher-dimensional space (e.g., optimizing the block length of the distribution matcher[16],[17]). The advantages obtained in this way are moderate, and might become negligible in the presence of carrier recovery algorithms[18].

The current research challenge is the full optimization of the constellation in a high-dimensional space, possibly in combination with improved detection strategies. While this is an extremely complex and still unsolved problem, in this work we use a recently proposed sequence selection (SS) bounding technique to estimate the gain achievable by such an optimization[19].

Detection

Optimal detection requires knowledge of the conditional distribution $p(\mathbf{y}|\mathbf{x})$. The problem has been widely studied for the AWGN channel, where $p(\mathbf{y}|\mathbf{x})$ factorizes into the product of marginal Gaussian distributions, so that optimal detection can be easily implemented. For the nonlinear fiber channel, $p(\mathbf{y}|\mathbf{x})$ is unknown, so that a mismatched detection based on an approximated distribution $q(\mathbf{y}|\mathbf{x})$ is used. Often—for simplicity and in the absence of a suitable alternative—$q(\mathbf{y}|\mathbf{x})$ is still taken as the product of marginal Gaussian distributions, as in the AWGN channel, but increasing the variance to account also for NLI.

The search for more accurate and mathematically tractable mismatched channel models is the subject of current research. For instance, several models show that interchannel NLI includes relevant phase and polarization noise (PPN) components that evolve slowly in time[20]–[22]. Such components depend also on frequency and can be alternatively represented as time-varying linear ISI[23]. Their mitigation is possible[22],[24] and yields an increase of the AIR, which is more effective if combined with subcarrier multiplexing[25],[26] and an optimized per-subcarrier power allocation[27],[28]. Moreover, even the additive component of NLI has some correlation in time, which might be exploited for its mitigation[27],[28].

Another important research topic is the practical implementation of such improved detection strategies with a reasonable computational complexity. Besides particle filtering techniques[26],[29]—useful for accurate AIR estimation with complex metrics but computationally too expensive for practical implementation—various approximated implementations based, e.g., on maximum likelihood sequence detection, (extended) Kalman filtering/smoothing, recursive least square equalization, and turboequalization have been proposed[30],[31]. Relevant gains are obtained, but a further reduction of the computational complexity might be required to make the approach attractive.

Digital backpropagation

One of the most popular DSP technique for nonlinearity mitigation is digital backpropagation (DBP), a channel inversion technique that can be implemented at TX or RX to remove intrachannel NLI[32],[33]. The most classical DBP implementation is based on the split-step Fourier method (SSFM)[1], with an accuracy and complexity that increase with the number of steps. The gains achievable with DBP are well studied, both numerically and experimentally[22],[34],[35]. Unfortunately, reasonable gains require many steps (one or more per span), so that the search for less complex algorithms is still in progress. Possible approaches include perturbation methods and Volterra equalizers[36]; the simplification of the SSFM[37]; machine learning techniques[38]–[40]; and a combination of the previous approaches—e.g., where a perturbation method is used to improve the accuracy of the nonlinear step of the SSFM[41], or machine learning techniques are used to optimize and simplify a processing scheme inspired by the SSFM or by a perturbation model[42],[43].

Another interesting research topic is the possibility to include interchannel effects in DBP without increasing its complexity, either to improve the performance by jointly backpropagating several channels, or to reduce the complexity by means of subband processing in a single channel[44]–[46].

In this work, we simply consider ideal single-channel DBP as an ultimate limit for intrachannel NLI compensation, and study the AIR gain it provides when used alone, or in combination with improved shaping and detection strategies.

Achievable information rates

We investigate and compare the effectiveness of the techniques discussed above in terms of AIR. The system is depicted in Fig.1, while the scenario and link parameters are the same considered in[26]—a 1000 km standard single-mode fiber link, ideal distributed amplification, and five 50 GBd Nyquist-WDM channels with 50 GHz spacing.

As a benchmark, we consider the AIR obtained on the nonlinear channel when the system is optimized in the absence of nonlinear effects, i.e., when considering ideal electronic dispersion compensation (EDC), i.i.d. Gaussian input symbols,

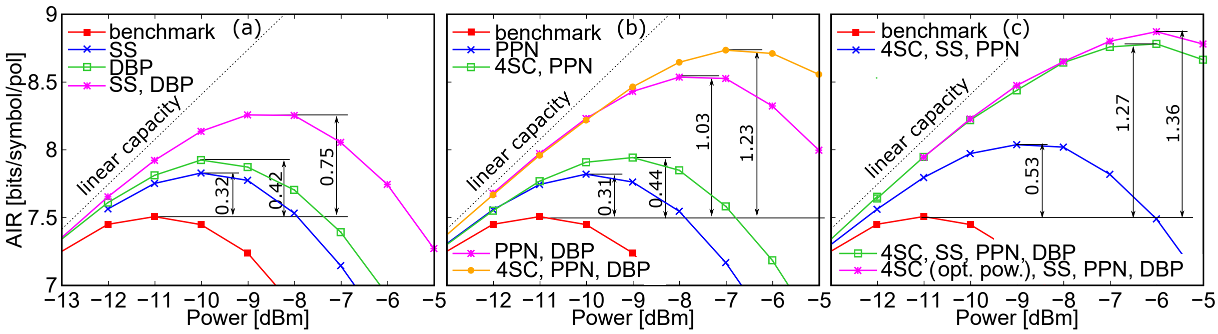

Fig. 2: AIR for different combinations of input distribution, detection metric, and processing.

and AWGN detection. This benchmark AIR, reported in Fig. 1(a)–(c) with a solid red line, reaches a peak of about 7.5 bits/sym/pol at a launch power of -11 dBm per channel per polarization, then decreases again. The other curves are obtained by modifying modulation, detection, and/or processing with respect to the benchmark, as indicated by the corresponding labels. The linear capacity $C = \log_2(1 + \mathrm{SNR})$ is also reported as a reference.

First, we investigate the gains achievable by optimizing the input distribution. The optimization employs the SS procedure mentioned above and described in[19] to minimize the average variance of intrachannel NLI, with a selection rate of 0.2% and a block length of 256 dual-polarization symbols. Besides the benchmark, Fig. 1(a) shows three different cases: SS-optimized input distribution and EDC; i.i.d. Gaussian inputs and DBP; optimized input combined with DBP. AWGN detection is considered in all the cases. The SS optimization, alone, yields an AIR gain of 0.32 bits/sym/pol with respect to the benchmark. By comparison, DBP yields a slightly higher gain of 0.42 bits/sym/pol, while the combination of the two techniques yields a higher total gain of 0.75 bits/sym/pol. This suggests that the input distribution provided by SS, though optimized to reduce intrachannel NLI, partly reduces also interchannel NLI.

Then, we investigate the gains achievable by improving the detection strategy. Besides the benchmark AIR, Fig. 2(b) shows the AIR obtained with PPN detection (with one or four subcarriers, the latter denoted as 4SC)[26], either with EDC or combined with DBP. In all the cases, i.i.d. Gaussian inputs are considered. PPN detection works better when combined with subcarrier modulation[26], providing a gain of about 0.44 bits/sym/pol, comparable with DBP. However, while PPN detection mainly addresses interchannel NLI, DBP removes only intrachannel NLI. As a result, when used alone, their effectiveness is limited by the remaining uncompensated effect. On the other hand, their combination acts synergically, mitigating both effects and yielding a much higher gain of about 1.23 bits/sym/pol.

Finally, we investigate the overall gain achievable by combining the previous techniques. Fig. 2(c) compares the benchmark AIR with that obtained by combining the SS procedure of Fig. 2(a) and the PPN detection of Fig. 2(b) (with 4SC), either with EDC or combined with DBP. The additional gain provided by SS is smaller in this case, since the optimization accounts only for intrachannel NLI and does not include PPN and DBP. Finally, including the per-subcarrier power optimization proposed in[27],[28] further improves the gain up to 1.36 bits/sym/pol.

Conclusions

Although nonlinearity mitigation appears to be an elusive target, many strategies have been devised over time, each addressing a specific aspect of the problem. By combining an optimized input distribution, a PPN-aware detection strategy, and including DBP, a gain of 1.36 bits/sym/pol in the peak AIR is achieved compared to a linearly optimized system, pushing the ultimate limit a little further and keeping alive the hope of finding a truly optimal strategy. For simplicity, the input is optimized, under some practical constraints, to minimize intrachannel NLI in the absence of any other mitigation strategy. Higher gains are expected by a full optimization that accounts also for interchannel NLI and for the actual combination of processing and detection.

References

[1] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd. San Diego, CA: Academic Press, 2001.

[2] A. D. Ellis *et al.*, “Approaching the non-linear Shannon limit,” *J. Lightw. Technol.*, vol. 28, no. 4, pp. 423–433, Feb. 2010.

[3] M. Secondini and E. Forestieri, “Scope and limitations of the nonlinear Shannon limit,” *J. Lightw. Technol.*, vol. 35, no. 4, pp. 893–902, Feb. 2017.

[4] E. Agrell *et al.*, “Roadmap of optical communications,” *J. Optics*, vol. 18, no. 6, p. 063 002, 2016.

[5] E. Agrell and M. Secondini, “Information-theoretic tools for optical communications engineers,” in *2018 IEEE Photonics Conference (IPC)*, IEEE, 2018, pp. 1–5.

[6] E. Wang and C. R. Menyuk, "Polarization evolution due to the Kerr nonlinearity and chromatic dispersion," *J. Lightw. Technol.*, vol. 17, no. 12, pp. 2520–2529, Dec. 1999.

[7] E. Agrell and M. Karlsson, "Influence of behavioral models on multiuser channel capacity," *J. Lightw. Technol.*, vol. 33, no. 17, pp. 3507–3515, Sep. 2015.

[8] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3/4, pp. 379–423/623–656, Jul. 1948.

[9] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 913–929, 1993.

[10] G. Böcherer *et al.*, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 4651–4665, 2015.

[11] F. Buchali *et al.*, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *Journal of Lightwave Technology*, vol. 34, no. 7, pp. 1599–1609, 2016.

[12] Y. C. Gültekin *et al.*, "Enumerative sphere shaping for wireless communications with short packets," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1098–1112, 2019.

[13] T. Fehenberger *et al.*, "On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel," *Journal of Lightwave Technology*, vol. 34, no. 21, pp. 5063–5073, 2016.

[14] K. Kojima *et al.*, "Nonlinearity-tolerant four-dimensional 2A8PSK family for 5–7 bits/symbol spectral efficiency," *J. Lightw. Technol.*, vol. 35, no. 8, pp. 1383–1391, 2017.

[15] B. Chen *et al.*, "Polarization-ring-switching for nonlinearity-tolerant geometrically shaped four-dimensional formats maximizing generalized mutual information," *J. Lightw. Technol.*, vol. 37, no. 14, pp. 3579–3591, 2019.

[16] O. Geller *et al.*, "A shaping algorithm for mitigating inter-channel nonlinear phase-noise in nonlinear fiber systems," *Journal of Lightwave Technology*, vol. 34, no. 16, pp. 3884–3889, 2016.

[17] T. Fehenberger *et al.*, "Mitigating fiber nonlinearities by short-length probabilistic shaping," in *Optical Fiber Communication Conference*, Optical Society of America, 2020, Th1I–2.

[18] S. Civelli *et al.*, "Interplay of probabilistic shaping and carrier phase recovery for nonlinearity mitigation," in *2020 European Conference on Optical Communication (ECOC)*, IEEE, 2020, pp. 1–3.

[19] S. Civelli *et al.*, "A sequence selection bound for the capacity of the nonlinear fiber channel," *arXiv preprint arXiv:2106.04097*, 2021.

[20] M. Secondini and E. Forestieri, "Analytical fiber-optic channel model in the presence of cross-phase modulation," *IEEE Photon. Technol. Lett.*, vol. 24, no. 22, pp. 2016–2019, Nov. 2012.

[21] R. Dar *et al.*, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Exp.*, vol. 21, no. 22, pp. 25 685–25 699, Nov. 2013.

[22] R. Dar and P. J. Winzer, "Nonlinear interference mitigation: Methods and potential gain," *J. Lightwave Technol.*, vol. 35, no. 4, pp. 903–930, 2017.

[23] M. Secondini *et al.*, "Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps," *Journal of Lightwave Technology*, vol. 31, no. 23, pp. 3839–3852, 2013.

[24] M. Secondini and E. Forestieri, "On XPM mitigation in WDM fiber-optic systems," *IEEE Photon. Technol. Lett.*, vol. 26, no. 22, pp. 2252–2255, 2014.

[25] D. Marsella *et al.*, "A simple strategy for mitigating XPM in nonlinear WDM optical systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2015.

[26] M. Secondini *et al.*, "Nonlinearity mitigation in WDM systems: Models, strategies, and achievable rates," *Journal of Lightwave Technology*, vol. 37, no. 10, pp. 2270–2283, 2019.

[27] F. J. García-Gómez and G. Kramer, "Mismatched models to lower bound the capacity of optical fiber channels," *Journal of Lightwave Technology*, vol. 38, no. 24, pp. 6779–6787, 2020.

[28] F. J. Garciagomez and G. Kramer, "Mismatched models to lower bound the capacity of dual-polarization optical fiber channels," *Journal of Lightwave Technology*, 2021.

[29] J. Dauwels and H.-A. Loeliger, "Computation of information rates by particle methods," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 406–409, Jan. 2008.

[30] E. P. da Silva *et al.*, "Perturbation-based FEC-assisted iterative nonlinearity compensation for WDM systems," *J. Lightw. Technol.*, vol. 37, no. 3, pp. 875–881, 2018.

[31] O. Golani *et al.*, "NLIN mitigation using turbo equalization and an extended Kalman smoother," *J. Lightw. Technol.*, vol. 37, no. 9, pp. 1885–1892, 2019.

[32] R. J. Essiambre and P. J. Winzer, "Fibre nonlinearities in electronically pre-distorted transmission," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, vol. 2, Sep. 2005, pp. 191–192.

[33] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, 2008.

[34] S. J. Savory *et al.*, "Impact of interchannel nonlinearities on a split-step intrachannel nonlinear equalizer," *IEEE Photon. Technol. Lett.*, vol. 22, no. 10, pp. 673–675, 2010.

[35] A. Ghazisaeidi *et al.*, "Submarine transmission systems using digital nonlinear compensation and adaptive rate forward error correction," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1886–1895, 2016.

[36] F. P. Guiomar *et al.*, "Mitigation of intra-channel nonlinearities using a frequency-domain volterra series equalizer," *Opt. Exp.*, vol. 20, no. 2, pp. 1360–1369, 2012.

[37] D. Rafique *et al.*, "Compensation of intra-channel nonlinear fibre impairments using simplified digital backpropagation algorithm," *Opt. Exp.*, vol. 19, no. 10, pp. 9453–9460, 2011.

[38] M. A. Jarajreh *et al.*, "Artificial neural network nonlinear equalizer for coherent optical OFDM," *IEEE Photon. Technol. Lett.*, vol. 27, no. 4, pp. 387–390, 2014.

[39] C. Häger and H. D. Pfister, "Nonlinear interference mitigation via deep neural networks," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, IEEE, 2018, pp. 1–3.

[40] B. Karanov *et al.*, "End-to-end deep learning of optical fiber communications," *J. Lightw. Technol.*, vol. 36, no. 20, pp. 4843–4855, 2018.

[41] M. Secondini *et al.*, "Single-step digital backpropagation for nonlinearity mitigation," *Photon. Netw. Commun.*, vol. 31, no. 3, pp. 493–502, Jun. 2016.

[42] V. Oliari *et al.*, "Revisiting efficient multi-step nonlinearity compensation with machine learning: An experimental demonstration," *J. Lightw. Technol.*, 2020.

[43] A. Redyuk *et al.*, "Compensation of nonlinear impairments using inverse perturbation theory with reduced complexity," *J. Lightw. Technol.*, vol. 38, no. 6, pp. 1250–1257, 2020.

[44] E. F. Mateo *et al.*, "Efficient compensation of inter-channel nonlinear effects via digital backward propagation in WDM optical transmission," *Opt. Exp.*, vol. 18, no. 14, pp. 15 144–15 154, 2010.

[45] C. Häger and H. D. Pfister, "Wideband time-domain digital backpropagation via subband processing and deep learning," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, IEEE, 2018, pp. 1–3.

[46] S. Civelli *et al.*, "Coupled-channel enhanced SSFM for digital backpropagation in WDM systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Optical Society of America, 2021, M3H.7.