Multi-dimensional Probabilistic Shaping for Optical Superchannels

Mrinmoy Jana[1], Lutz Lampe[1], Jeebak Mitra[2], Chuandong Li[2]
(1) ECE Dept., University of British Columbia, BC, Canada, {mjana, lampe}@ece.ubc.ca
(2) Huawei Technologies, Ottawa, ON, Canada, jeebak.mitra@huawei.com

Abstract *We present a multi-dimensional probabilistic shaping algorithm for optical superchannels, complemented with machine-learning assisted soft demapping at the receiver. The proposed system outperforms traditional shaping methods by 0.3-1.05 dB for the same block-length, to enable a 300-km superchannel transmission achieving 800 Gbps data-rate per carrier.*

Introduction

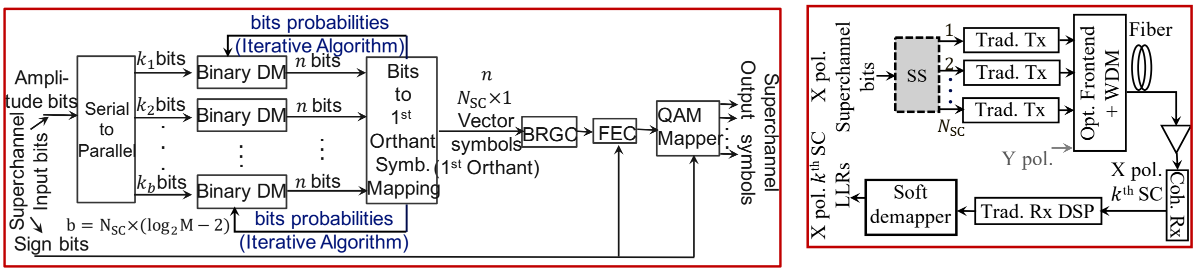
To enable high spectral efficiency (SE) in the next generation optical fiber networks, probabilistic shaping (PS) has received a considerable attention more recently[1–3]. In addition to providing an ultimate shaping gain up to 1.53 dB, PS can also offer rate adaptivity for systems employing a fixed forward-error-correction (FEC) code. Distribution matching (DM) is considered to be a key element for the implementation of PS. One of the most popular DM methods is a symbol-level constant composition DM (SL-CCDM) algorithm[2]. Although SL-CCDM has vanishing rate loss for asymptotically large block-lengths, it suffers from high rate loss as the block-length reduces[1,2]. On the other hand, implementing SL-CCDM with large block-lengths entails significant complexity and latency[1]. Therefore, to strike a good trade-off, several alternative DM methods, such as the product distribution matching (PDM)[4,5] and the two-dimensional DM (2D-DM)[6], have been proposed, that apply bit-level CCDMs. Moreover, our recent work[7] investigated bit-level shaping across both polarizations of a dual-polarized (DP) system. To further exploit the additional degrees of freedom available in a wavelength-division-multiplexing (WDM) superchannel transmission, in this work, we consider jointly shaping multiple sub-channels (SCs). Due to the practical limitations of the opto-electronics to facilitate high baud-rate single-carrier transmissions, such a superchannel system is an attractive choice to achieve high data rates. While the proposed superchannel shaping (SS) enables joint processing of SCs at the transmitter, we employ separate demappers at the receiver for each SC to parallelize the receiver signal processing. Such an individual processing together with the presence of fiber nonlinearity may cause a traditional soft demapper to produce mismatched log-likelihood ratios (LLRs)[8]. Therefore, to improve the accuracy of the LLRs, we also investigate a machine learning (ML) assisted soft demapping algorithm. Our numerical results show that the combined benefits of the SS in tandem with the ML demapper offer a 0.3-1.05 dB performance improvement over existing shaping methods when an optical superchannel is transmitted over 300 km enabling 800 Gbps data rate per SC.

978-1-6654-3868-1/21/$31.00 ©2021 IEEE

Superchannel Shaping (SS)

The block diagram of the proposed SS system is shown in Fig. 1a. Uniformly distributed bits for each polarization (pol.) stream and all N_{SC} SCs of the superchannel are processed in parallel by multiple binary DMs to output n shaped bits. Thereafter, the shaped bit-sequences are mapped to the $N_{\mathrm{SC}} \times 1$ symbols vector in the first orthant of the superchannel signal space. A binary reflected Gray code (BRGC) mapping is subsequently used to convert the first-orthant symbols to bits, followed by FEC coding and quadrature amplitude modulation (QAM) symbols generation for all SCs in the superchannel.

To obtain the optimal binary probabilities $\boldsymbol{p}$ required for the binary DMs, and the bits-to-first-orthant symbols mapping $\mathcal{M}$, we employ an iterative algorithm. We initialize the algorithm with a multi-dimensional Maxwell-Boltzmann (MB) probability mass function (pmf) f_{SS} assumed for the first orthant superchannel symbols to meet a desired SE target, and the well-known natural binary code (NBC)[6] for the bit-mapping $\mathcal{M}$. Thereafter, we perform the following sequential operations, iteratively: (a) marginalize the pmf f_{SS} to obtain the constituent bit probabilities for a given $\mathcal{M}$, (b) quantize the bit probabilities to obtain $\boldsymbol{p}$ by accounting for the finite block-length n, (c) multiplying the quantized binary bit-probabilities to recompute the pmf f_{SS}, and (d) update $\mathcal{M}$ by sort-



(a) Detailed SS block diagram. **(b)** Transmission model, shown for X pol.

Fig. 1: SS WDM system model.

ing the symbol probabilities such that the largest probabilities are assigned to the lowest amplitudes. The iterative method is performed until the root-mean square (RMS) power of the multi-dimensional constellation reaches a steady state. At the end of the final iteration, $\boldsymbol{p}$ and $\mathcal{M}$ are stored as a look-up table (LUT). The above iterative algorithm can be computed offline for a target SE without additional run-time computational complexity requirement.

By implementing the proposed SS, we aim to improve the shaping performance with short block-lengths. To accomplish this, the first key observation is that the RMS power of the QAM constellation should be minimized[6], which is dictated by the bit probabilities $\boldsymbol{p}$ and the mapping $\mathcal{M}$. To this end, our second key observation is that an NBC mapping, which is used in traditional PS schemes such as[5], is not optimal for the superchannel multidimensional signal space. Third, we note that by jointly configuring the lengths $k_1, k_2, \ldots$ of the binary DM input bits for all SCs in the superchannel, as shown in Fig. 1a, we introduce additional degrees of freedom into the shaping design in order to minimize the RMS power of the transmitted symbols, compared to signal-shaping performed separately on individual SCs.

Once the QAM symbols for all SCs of each pol. are generated from the SS algorithm as described above, the transceiver data-path involves traditional transmitter and receiver (Rx) processing of a typical WDM transmission as shown in Fig. 1b. At the coherent Rx, individual SCs are processed separately by a conventional digital signal processing (DSP) unit, followed by soft demapping to produce the LLRs. In the following section, we provide more details on the soft demapper.

ML-Assisted Soft Demapping

The jointly shaped SCs in SS are separately processed at the receiver to parallelize the Rx-DSP. A traditional soft demapping performs marginalization of the multi-dimensional SS symbols pmf to compute the a-priori probability of the QAM symbols for each SC. Moreover, a nonlinear

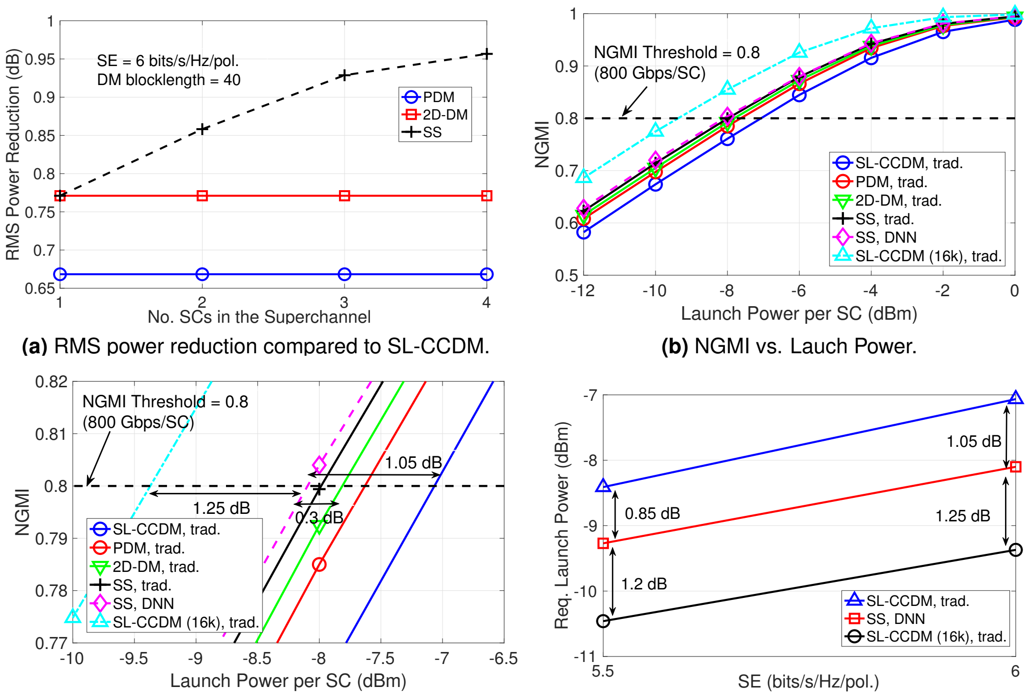
Tab. 1: Simulation parameters

Parameters	Values	
Root raised cosine roll-off	0.1	
Raw SE (bits/Hz/pol.)	6	5.5
Baud rate/SC (Gbaud)	80	90
Channel Spacing (GHz)	88	99
Data rate/SC (Gbps)	800	801
Superchannel BW (GHz)	264	297
SSMF spans	4×75 km	
CD	17 ps/nm/km	
Nonlinearity parameter	$1.4\ \mathrm{W}^{-1}\mathrm{km}^{-1}$	
Fiber attenuation	0.2 dB/km	
Amplifier noise-figure	4.5 dB	
DNN neurons ($H = 3$)	$100 \times 200 \times 100$	
DNN training symbols	4×10^5	
DNN validation symbols	1×10^5	
DNN test symbols	5×10^5	
DNN mini-batch size	7000	
DNN epochs	300	

fiber channel may exhibit characteristics different from a traditional additive white Gaussian noise (AWGN) model[9]. Due to such mismatches in the underlying model, a traditional demapping may lead to imperfect LLRs[8]. Therefore, to improve the accuracy of the LLRs for the SS transmission, we present an ML based demapping employing a deep neural network (DNN). In contrast to a previous work[10], which applies DNN for device nonlinearity mitigation, we apply DNN for the SS systems transmitted through a nonlinear fiber channel. The DNN uses the received symbols after the Rx DSP as inputs, followed by H hidden layers, and the bit-level LLRs computed as the final output. For training the DNN, we employ stochastic gradient descent algorithm with binary cross-entropy as the cost function, such that the LLR computation can be formulated as a multi-label binary classification problem.

Numerical Results

We simulate a DP 256-QAM SS Nyquist WDM superchannel transmission having 3 SCs. For our numerical results, we choose a normalized generalized mutual information (NGMI) threshold of 0.8, which corresponds to 800 Gbps data rate per SC. We consider two raw SE values of 5.5 and 6 bits/s/Hz/pol. with varying baud rates per SC, such



(a) RMS power reduction compared to SL-CCDM. **(b)** NGMI vs. Lauch Power.

(c) Zoomed-in version of Fig. 2b. **(d)** Minimum required launch power vs. raw SE.

Fig. 2: Numerical results for a DP 3×256-QAM Nyquist WDM superchannel transmission.

that the 800 Gbps SCs occupy less than 100 GHz bandwidth (BW) in the superchannel. We choose a fixed DM output block-length of 40 for our results. Such a short block-length in tandem with multiple binary DMs operating in parallel requires significantly low serialism and latency compared to long block-length traditional SL-CCDM architectures. The remaining transmission, fiber, Rx-DSP and DNN parameters are listed in Table 1.

In Fig. 2a, we show the reduction of the RMS power of the QAM constellation for different shaping algorithms compared to the SL-CCDM method to achieve a target raw SE = 6 bits/s/Hz/pol., as a function of the number of SCs in a superchannel. The plots in the figure suggest that SS offers increasing gains compared to SL-CCDM as the number of SCs in the superchannel increases, by jointly shaping more SCs together. For example, SS yields close to 1 dB reduction in the RMS power over SL-CCDM for the same block-length when 3 and 4 SCs are employed in a superchannel. In Fig. 2b, we show the NGMI achieved by different shaping methods for a raw SE of 6 bits/s/Hz/pol., applying either a traditional or the DNN based demapper at the receiver, as a function of the launch power. As a benchmark, we also added a plot for the SL-CCDM using a block-length of 16000 (labeled “SL-CCDM (16k), trad.”), which suffers negligible rate loss at the cost of significantly higher complexity and latency requirement[1,5]. For the sake of clarity, a zoomed-in version of Fig. 2b is shown in Fig. 2c around an NGMI value of 0.8. The plots show that SS with traditional demapping (labeled “SS, trad.”) offers 0.95 dB gain compared to SL-CCDM having the same block-length of 40 (labeled “SL-CCDM, trad.”) as predicted from Fig. 2a. When the DNN based demapping is employed for the SS superchannels, an additional gain of 0.1 dB can be observed, resulting in an aggregate performance improvement of 1.05 dB, 0.5 dB and 0.3 dB over the SL-CCDM, PDM, and 2D-DM, respectively, for the same block length. However, SS with a block-length of 40 using the DNN demapper is outperformed by SL-CCDM with 16000 block-length by 1.25 dB, with the benefit of a significantly lower complexity and latency due to a much smaller block-length. Finally, in Fig. 2d, we plot the minimum launch power required by SL-CCDM and SS to achieve an NMGI value of 0.8 when raw SE = 5.5 and 6 bits/s/Hz/pol. We observe a similar gain of 0.85 dB by the SS using DNN based demapper compared to SL-CCDM employing a traditional demapper for the SE value of 5.5 bits/s/Hz.

Conclusion

We presented SS to jointly shape multiple SCs in a superchannel. Moreover, to improve the accuracy of the mismatched LLRs, we investigated a DNN assisted soft demapper. For the same DM block-length, SS in conjunction with the ML demapping outperformed SL-CCDM, PDM and 2D-DM employing a traditional demapper by up to 1.05 dB, 0.5 dB and 0.3 dB, respectively, to enable a WDM superchannel transmission over 300 km achieving 800 Gbps data rate per SC.

References

[1] Y. C. Gültekin *et al.*, "Probabilistic Shaping for Finite Blocklengths: Distribution Matching and Sphere Shaping," *Entropy*, vol. 22, no. 5, 2020.

[2] G. Böcherer *et al.*, "Probabilistic Shaping and Forward Error Correction for Fiber-Optic Communication Systems," *J. Lightw. Technol*, vol. 37, no. 2, 2019.

[3] M. Jana *et al.*, "Probabilistic Shaping in Time-Frequency-Packed Terabit Superchannel Transmission," *IEEE Photon. Technol. Lett.*, vol. 32, no. 17, 2020.

[4] G. Böcherer *et al.*, "High Throughput Probabilistic Shaping with Product Distribution Matching," *arXiv e-prints*, p. arXiv:1702.07510, 2017.

[5] M. Pikus *et al.*, "Bit-Level Probabilistically Shaped Coded Modulation," *IEEE Commun. Lett.*, vol. 21, no. 9, 2017.

[6] M. Fu *et al.*, "Multi-dimensional Distribution Matching for Probabilistically Shaped High Order Modulation Format," *Opt. Fiber Commun. Conf. (OFC)*, 2020.

[7] M. Jana *et al.*, "Jointly Shaped Dual Polarization Systems," *Opt. Fiber Commun. Conf. (OFC)*, 2021.

[8] A. Alvarado *et al.*, "Improved Soft-Decision Forward Error Correction via Post-Processing of Mismatched Log-Likelihood Ratios," *European Conf. on Opt. Commun. (ECOC)*, 2016.

[9] A. Amari *et al.*, "A Machine Learning-Based Detection Technique for Optical Fiber Nonlinearity Mitigation," *IEEE Photon. Technol. Lett.*, vol. 31, no. 8, 2019.

[10] M. Schaedler *et al.*, "Neural Network-Based Soft-Demapping for Nonlinear Channels," *Opt. Fiber Commun. Conf. (OFC)*, 2020.