Probabilistic Amplitude Shaping and Nonlinearity Tolerance: Analysis and Sequence Selection Method

Mohammad Taha Askari, Lutz Lampe, *Senior Member, IEEE*, Jeebak Mitra

Abstract—Probabilistic amplitude shaping (PAS) is a practical means to achieve a shaping gain in optical fiber communication. However, PAS and shaping in general also affect the signal-dependent generation of nonlinear interference. This provides an opportunity for nonlinearity mitigation through PAS, which is also referred to as a nonlinear shaping gain. In this paper, we introduce a linear lowpass filter model that relates transmitted symbol-energy sequences and nonlinear distortion experienced in an optical fiber channel. Based on this model, we conduct a nonlinearity analysis of PAS with respect to shaping block-length and mapping strategy. Our model explains results and relationships found in literature and can be used as a design tool for PAS with improved nonlinearity tolerance. We use the model to introduce a new metric for PAS with sequence selection. We perform simulations of selection-based PAS with various amplitude shapers and mapping strategies to demonstrate the effectiveness of the new metric in different optical fiber system scenarios.

Index Terms—Optical fiber communications, probabilistic amplitude shaping, nonlinear interference, nonlinearity mitigation, nonlinear shaping gain, sequence selection, first-order perturbation analysis.

I. INTRODUCTION

PROBABILISTIC constellation shaping is a well-established approach to improve over uniform signaling providing an ultimate shaping gain of 1.53 dB for the additive white Gaussian noise (AWGN) channel [1]. Among the various probabilistic shaping methods [2], PAS has emerged as an effective design for integrating shaping and forward error correction (FEC) coding [3]. The PAS architecture consists of an inner FEC encoder concatenated with an outer amplitude shaper (AS). The AS maps uniformly distributed input bits to amplitudes and generates a non-uniform amplitude distribution. It operates on finite-length blocks of amplitudes at a time, which results in a rate loss of the achievable information rate (AIR) compared to the infinite blocklength limit.

In addition to the conventional *linear* shaping gain, shaping also has the potential to exhibit a *nonlinear* gain compared to uniform signaling for the optical fiber channel [4]. Several previous works, e.g., [5]–[8], attributed the nonlinear shaping gain to a change in the moments and in particular the kurtosis of the amplitudes distribution. For PAS, it has furthermore been demonstrated in [9]–[11] that the effective signal-to-noise ratio (SNR) depends on the blocklength of the AS. As the AS

M.T. Askari and L. Lampe are with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada, e-mail: (mohammadtaha@ece.ubc.ca, lampe@ece.ubc.ca).
J. Mitra was with Huawei Technologies Canada, Ottawa, He is now with Dell Technologies Canada, e-mail: jeebak.mitra@dell.com.

blocklength increases, the nonlinear shaping gain decreases, while the linear shaping gain increases. Since this trend is not observed if a symbol interleaver is applied to shaped sequences, it can be concluded that the temporal properties of the shaped sequences play a role in the severity of nonlinear interference (NLI).

Besides the shaping blocklength, the choice of mapping of amplitudes from one shaping block to the components of the transmitted signal affects the nonlinearity tolerance of PAS transmission. In particular, it has been reported that mapping shaping blocks across in-phase and quadrature-phase components in single-polarization transmission and also across x- and y-polarization in dual-polarization transmission can increase the effective SNR [12]. This effect can be explained by noting that the optical channel acts as a lowpass filter on symbol-energy sequences causing NLI, and that finite blocklength PAS produces a spectral dip at frequency zero for such sequences [13]. The filter point of view also facilitates understanding the effect of AS together with other system parameters such as baud rate on NLI [14]. The significance of the properties of symbol-energy sequences for the NLI in PAS transmission is further highlighted by the so-called energy dispersion index (EDI) introduced in [15]. The EDI is a measure for the variance of the windowed energy of the shaped signal, where the window length is infinite in the case of the exponentially weighted EDI (EEDI) [16]. For PAS implemented with constant composition distribution matching (CCDM) [17], both EEDI and EDI have been demonstrated to be good predictors for effective SNR.

The observations of the interplay between PAS and NLI suggest a modification of the traditional AS design that includes the nonlinear shaping gain as an objective. Two types of such modified designs have emerged in the literature. The first pairs a conventional AS with a selection module, in which the latter chooses from a candidate set of shaped amplitude sequences generated by the AS. The selection is aided by a criterion that accounts for the nonlinearity of the optical channel, such as the EDI in the list-encoding CCDM (L-CCDM) of [18] or the effective SNR obtained from channel emulation using the split-step Fourier method (SSFM) in [19]. The generation of candidate sets can be integrated into the operation of a conventional AS by inserting additional flipping bits into each information block that is input to the AS. The second approach for nonlinearity tolerant shaping directly modifies the sequence generation in the conventional AS. For example, the trellis of the enumerative sphere shaping (ESS) algorithm [20] can be adjusted to obtain sequences with improved nonlinearity tolerance. This has been done based
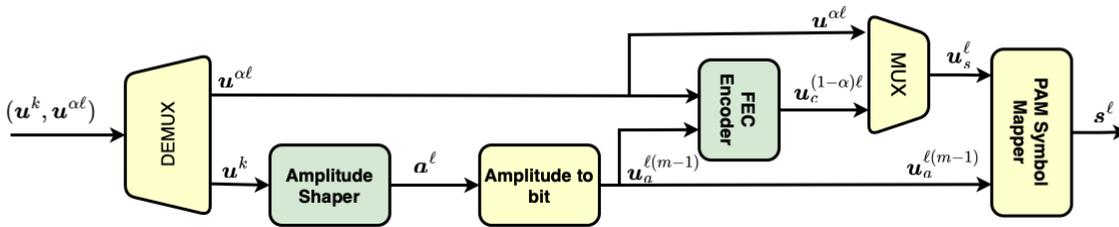
Fig. 1. Block diagram for conventional PAS. Superscripts denote the sequence length. $\boldsymbol{u}$, $\boldsymbol{a}$, $\boldsymbol{s}$ denote a bit, PAM amplitude, and PAM symbol sequences, respectively.

on the per-sequence kurtosis in K-ESS [21] and considering sequence-energy variations in band-trellis ESS (B-ESS) [22]. The bisection-based implementation of CCDM (BS-CCDM) presented in [23] exploits the redundancy in CCDM to re-move unfavorable CCDM sequences and thus mitigate fiber nonlinearity.

In this paper, we revisit the interplay of PAS and fiber nonlinearity using the first-order perturbation based model for NLI. First, we develop an effective *linear channel model* for the conversion of symbol-energy sequences into NLI. Our model extends the lowpass filter model from [13] and accounts for self-phase modulation (SPM) and cross-phase modulation (XPM) as well as intra-polarization and inter-polarization effects. Using the frequency domain representation of the effective linear channel, we provide insights into the nonlin-earity tolerance of finite-length PAS as a function of system parameters such as link length and baud rate. This analysis permits us to explain several phenomenological observations reported in previous works, such as the optimization of AS blocklength and mapping strategy to maximize the effective SNR. As a second contribution, we employ our channel model to introduce the lowpass-filtered symbol-amplitude sequence (LSAS) metric for nonlinearity tolerant PAS using sequence selection.[1] We identify its relation to the EDI, which is interesting in itself as it validates the EDI as a metric for nonlinearity tolerance of shaped sequences. It furthermore highlights that different from the EDI, the derivation of the LSAS metric does not rely on simplifying assumptions for quantifying NLI caused by symbol-energy sequences, and it generalizes systematically to XPM and inter-polarization NLI. The effectiveness of the new metric for PAS with sequence selection is demonstrated through simulations of single-span and multi-span fiber links for single-polarization and dual-polarization setups, which show that the LSAS metric outper-forms the EDI metric in terms of effective SNR, AIR, and error rate. Finally, we integrate the different strategies for improved nonlinearity tolerance with PAS. That is, for the first time, we combine (i) multidimensional mapping and (ii) nonlinearity tolerant K-ESS with LSAS-based sequence selection. Our nu-merical results demonstrate (i) that multidimensional mapping is effective in reducing the search space and thus complexity for PAS with sequence selection and (ii) that the combination of K-ESS and sequence selection yields notable benefits in terms of, e.g., effective SNR for the nonlinear fiber channel,

compared to applying them individually.

The remainder of this paper is organized as follows. In Section II, we briefly explain the system model for PAS, symbol mapping, and sequence selection. In Section III, we derive the linear channel model between symbol-energy se-quences and NLI using perturbation theory, and we discuss the effect of fiber link and system parameters on the channel properties. We then introduce the LSAS metric in Section IV, we show how the EDI can be derived as an approximation of LSAS, and we discuss the computational complexity associated with sequence selection using EDI and LSAS. In Section V, we present quantitative comparisons of different nonlinearity tolerant PAS methods. We highlight the benefits of the LSAS-metric and its combination with multidimensional mapping and different AS methods by simulating both short-haul and long-haul optical fiber transmission scenarios. In Section VI, we reflect on aspects of our work that merit further consideration and are thus suggested as future research. Section VII concludes the paper.

II. PROBABILISTIC AMPLITUDE SHAPING, MAPPING, AND SEQUENCE SELECTION

In this section, we briefly review conventional PAS and PAS with sequence selection. We also discuss the options for mapping symbols to the signal components in dual-polarized optical fiber transmission.

A. Probabilistic Amplitude Shaping (PAS)

Figure 1 shows the operation of PAS following the de-scription in [3]. The AS maps a block $\boldsymbol{u}^k$ of k independent and uniformly distributed information bits to a block $\boldsymbol{a}^\ell$ of ℓ amplitudes of a 2^m-ary pulse-amplitude modulation (PAM) constellation. The amplitudes of the PAM symbols have a marginal distribution P_a. The elements of $\boldsymbol{a}^\ell$ are amplitude-to-bit converted into the sequence $\boldsymbol{u}_\text{a}^{\ell(m-1)}$ of $\ell(m-1)$ amplitude bits. A sequence $\boldsymbol{u}_\text{s}^\ell$ of ℓ sign bits is obtained from $\alpha\ell$ information bits $\boldsymbol{u}^{\alpha\ell}$, $0 \le \alpha \le 1$, and $(1-\alpha)\ell$ bits $\boldsymbol{u}_\text{c}^{(1-\alpha)\ell}$ generated using a systematic FEC encoder. Each m bits consisting of $m-1$ amplitude bits from $\boldsymbol{u}_\text{a}^{\ell(m-1)}$ and one sign bit from $\boldsymbol{u}_\text{s}^\ell$ are mapped to a PAM symbol resulting in ℓ symbols $\boldsymbol{s}^\ell$.

---

[1]We note that a special case of the LSAS metric was originally introduced in our conference paper [24].
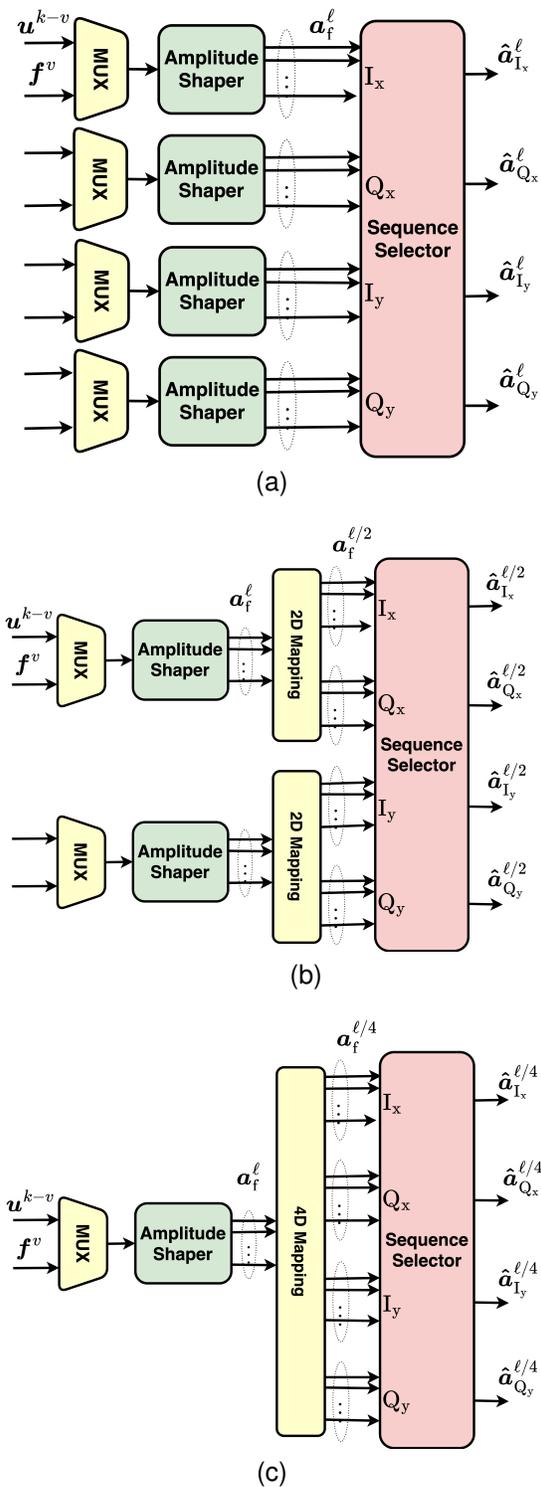
(a)

(b)

(c)

Fig. 2. Block diagram for different mapping strategies: (a) 1D, (b) 2D, and (c) 4D mapping, followed by sequence selection.

The shaping rate of the system is $R_\mathrm{s} = \frac{k}{\ell}$, and the overall data rate of this PAS scheme is $R = R_\mathrm{s} + \alpha$. The rate loss is defined as

$$R_\mathrm{loss} = H(P_\mathrm{a}) - R_\mathrm{s}, \qquad (1)$$

where $H(P_\mathrm{a})$ is the entropy associated with the distribution P_a.

B. Mapping Strategy

We consider dual-polarized optical fiber transmission using quadrature amplitude modulation (QAM). This means that PAM blocks s^ℓ are mapped to four-dimensional (4D) symbol sequences, where each symbol has I_x, Q_x, I_y, and Q_y components for in-phase and quadrature and x- and y-polarization dimensions, respectively. We differentiate three strategies for this mapping, which we refer to as 1D, 2D, and 4D mapping. For the case of 1D mapping, each shaping block s^ℓ is mapped to a separate quadrature and polarization dimension. The 2D mapping strategy maps each shaping block across the two quadrature components but to a single polarization. Finally, each shaping block is mapped to all four 4D-symbol components in 4D mapping. The different mapping strategies are illustrated in Figure 2 in terms of amplitude blocks a^ℓ, i.e., we map the shaped amplitude blocks to the signal components and apply the sign bits afterward, as this facilitates the discussion of sequence selection further below.

The choice of mapping has an effect on nonlinearity tolerance with PAS for two reasons. First, the mapping strategy changes the time duration during which the total energy of one shaping block is transmitted. This can be seen in Figure 2 comparing the lengths of the amplitude sequences at the output of the mappers. Thus, the mapping affects the effective shaping blocklength. Second, 2D and 4D mapping introduce dependencies between the different components of the 4D symbols. We will discuss the consequences of these interactions between the mapping and finite-length PAS in Section III-D.

C. Sequence Selection

When the AS of the PAS scheme produces several shaped amplitude blocks for representing the same block of information bits, the mapping step discussed in the previous section is followed by a sequence selector module (see Figure 2). The generation of candidate shaping blocks is enabled by allocating the first v out of k bits as redundant bits f^v, which are referred to as flipping bits [18]. Flipping the first bits provides more variation to the set of generated candidates for common shaping methods such as CCDM and ESS. The remaining $k-v$ bits are considered as information bits u^{k-v}. The candidate generation, which is adopted from [18], is independent of the selected amplitude shaper, i.e., the same process is applied to e.g. CCDM or ESS. For each combination of flipping bits, a shaping block of ℓ PAM amplitude symbols a_f^ℓ is generated. The amplitudes are mapped to a block of ℓ/d 4D-amplitude symbols according to the $(d \in \{1, 2, 4\})$-dimensional mapping strategy as shown in Figure 2. The sequence selector computes a metric for each 4D candidate and selects the best candidate $\hat{a}^{\ell/d} = [\hat{a}_{I_\mathrm{x}}^{\ell/d}, \dots, \hat{a}_{Q_\mathrm{y}}^{\ell/d}]$. We note that the size of the candidate set depends on the mapping strategy for a fixed shaping blocklength ℓ. Since $4/d$ ASs contribute to the generation of 4D-amplitude symbols, the selection module computes $2^{\frac{4}{d}v} \times d$ metrics for producing ℓ 4D symbols. Hence, computational complexity decreases with the dimension d of the mapping. After selection, amplitude-to-bit conversion and FEC opera-

tions are performed in the same way as for the conventional PAS scheme shown in Figure 1.

When introducing flipping bits, the rate loss expression (1) generalizes to

$$R_{\text{loss},v} = H(P_{\text{a}}) - R_{\text{s},v} = H(P_{\text{a}}) - \frac{k-v}{\ell}. \tag{2}$$

In order to have a fixed effective rate $R = \frac{k-v}{\ell} + \alpha$, $k - v$ should be remained fixed. Hence, as v increases, k also increases, resulting in lower linear shaping gain for given ℓ. Therefore, sequence selection is a mechanism to trade off linear with nonlinear shaping gain.

III. LINEAR CHANNEL APPROXIMATION FOR NONLINEAR DISTORTION

In this section, we derive the approximation of the nonlinear distortion as the output of a linear channel that is excited by the symbol-energy sequences as its input. We discuss the properties of the linear filter representing the channel as a function of fiber-link parameters. We then use this channel model to explain the effects of finite-length PAS and multidimensional mapping on the nonlinear tolerance of transmission.

A. Linear Channel Model Formulation

We consider the optical fiber channel after compensation of linear distortions at the receiver and focus on nonlinear distortion from signal-to-signal interactions. We assume dual-polarization transmission in a wavelength division multiplexing (WDM) system with a set of channels $\mathcal{C}$. The transmit symbol in polarization $p \in \mathcal{P} = \{x, y\}$ and channel $c \in \mathcal{C}$ at discrete time $n \in \mathbb{Z}$ is denoted by $s_p^{(c)}(n)$, and the symbol energy is

$$E_p^{(c)}(n) \triangleq |s_p^{(c)}(n)|^2. \tag{3}$$

Then, using the first-order perturbation based distortion model, as briefly reviewed in the Appendix, and only retaining symbol-energy terms, we can approximate the corresponding received sample as

$$r_p^{(c)}(n) \approx s_p^{(c)}(n) \Bigg[1 \\ + \mathrm{j}\gamma \underbrace{\left(\sum_{p' \in \mathcal{P}} \sum_{c' \in \mathcal{C}} \sum_{m \in \mathbb{Z}} E_{p'}^{(c')}(n+m) h_{p,p'}^{(c,c')}(m) \right)}_{\triangleq D_p^{(c)}(n)} \Bigg], \quad (4)$$

where γ is the fiber nonlinearity parameter, and $h_{p,p'}^{(c,c')}(m) \in \mathbb{R}_{\geq 0}$ are coefficients determined by the fiber channel. We observe that the distortion term can be written as ($*$ denotes convolution)

$$D_p^{(c)}(n) = \sum_{p' \in \mathcal{P}} \sum_{c' \in \mathcal{C}} \left(E_{p'}^{(c')} * h_{p,p'}^{(c,c')} \right)(n), \tag{5}$$

i.e., the summation of the linear convolution between the filters $h_{p,p'}^{(c,c')}$ and the symbol-energy sequences $E_{p'}^{(c')}$ for all WDM channels and both polarizations. The filters $h_{p,p'}^{(c,c')}$ represent intra- ($p = p'$) and inter-polarization ($p \neq p'$) interferences of SPM ($c = c'$) and XPM ($c \neq c'$), respectively. We note that

the linear channel model (5) is not meant to be sufficiently accurate for perturbation-based nonlinearity compensation. However, it is suitable for explaining the interplay between shaping and multidimensional mapping and NLI, and for guiding the design of nonlinearity tolerant PAS.

It is insightful to develop (4) and (5) further, which also allows us to connect to other related work. For this, we introduce the deterministic mean distortion term

$$\bar{D}_p^{(c)} = \sum_{p' \in \mathcal{P}} \sum_{c' \in \mathcal{C}} \bar{E}_{p'}^{(c')} \sum_{m \in \mathbb{Z}} h_{p,p'}^{(c,c')}(m) \tag{6}$$

and the distortion variation term

$$\Delta D_{p,p'}^{(c,c')}(n) = \left(\left[E_{p'}^{(c')} - \bar{E}_{p'}^{(c')} \right] * h_{p,p'}^{(c,c')} \right)(n), \tag{7}$$

where ($\mathbb{E}$ denotes statistical expectation)

$$\bar{E}_p^{(c)} = \mathbb{E}\left[E_p^{(c)}(n) \right] \tag{8}$$

is the mean symbol energy. Then, the total distortion (5) decomposes as

$$D_p^{(c)}(n) = \sum_{p' \in \mathcal{P}} \sum_{c' \in \mathcal{C}} \Delta D_{p,p'}^{(c,c')}(n) + \bar{D}_p^{(c)}. \tag{9}$$

Applying (9) and the first order Taylor series approximation of the exponential function to (4), we obtain the approximation

$$r_p^{(c)}(n) \approx s_p^{(c)}(n) \exp\left[\mathrm{j}\gamma \left(\sum_{p' \in \mathcal{P}} \sum_{c' \in \mathcal{C}} \Delta D_{p,p'}^{(c,c')}(n) + \bar{D}_p^{(c)} \right) \right]. \tag{10}$$

The NLI is now approximated as phase noise, which suggests that a carrier phase recovery (CPR) may mitigate the distortion. Since the effectiveness of a CPR depends on the dynamic of the distortion term, the temporal structure of the symbol-energy and thus the shaped amplitude sequence are important. Simulation results in [25], [26] suggest that a powerful CPR can be similarly effective in mitigating NLI as finite-length PAS, at least in the high SNR regime [25]. But even a mean-phase recovery would compensate for the deterministic term $\bar{D}_p^{(c)}$, and therefore we will exclude it in the design of a selection metric in Section IV.

Furthermore, the relationship for the filter coefficients $h_{p,p}^{(c,c')}(n)$ provided in (30) in the Appendix motivates the approximation

$$h_{p,p'\neq p}^{(c,c')}(n) \approx \frac{1}{2} h_{p,p}^{(c,c')}(n) \triangleq h_p^{(c,c')}(n), \quad \forall c, c' \in \mathcal{C}. \tag{11}$$

This leads to the simplified approximation

$$D_p^{(c)}(n) \approx \sum_{c' \in \mathcal{C}} \left(\left[2E_p^{(c')} + E_{p'\neq p}^{(c')} \right] * h_p^{(c,c')} \right)(n), \tag{12}$$

for (5), where the nonlinear distortion is approximated as linear filtering of the sum of symbol-energy sequences for the two signal polarizations. The simplified formulation (12) is equivalent to [13, Eqs. (3) and (4)].

This article has been accepted for publication in IEEE/OSA Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2023.3264032
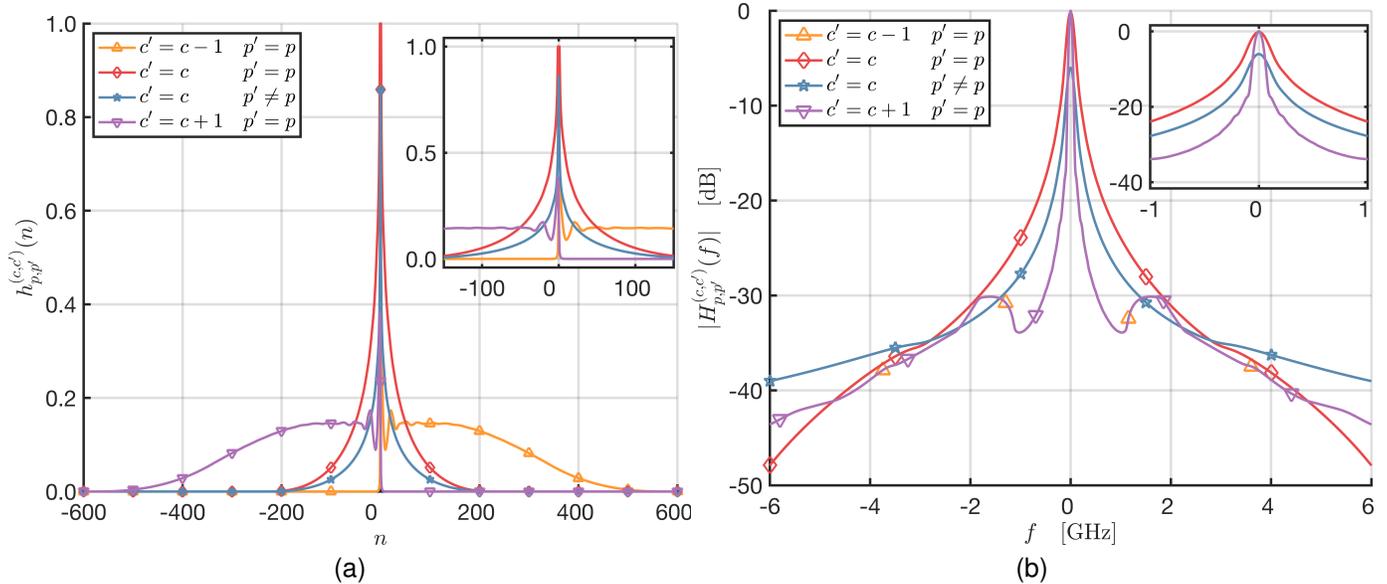
5

Fig. 3. Normalized filters for a fiber link with length 1600 km, 32 GBd, and 50 GHz channel spacing. SPM filter (intra- and inter-polarization) and XPM (intra polarization) filter from two adjacent channels. (a) Time domain, maximum coefficient normalized to 1. (b) Frequency domain (magnitude), maximum normalized to 0 dB at $f = 0$. The insets show an interval around the origin.

B. Filter Properties

We now evaluate the expression for the filters $h_{p,p'}^{(c,c')}$ for selected optical fiber links. If not specified otherwise, the link parameters are as given in Table I.

Figures 3a and 3b show the normalized filter coefficients and their normalized magnitude Fourier transforms, respectively, for SPM and for XPM from two adjacent channels for a 1600 km link, i.e., 20 spans of length 80 km. We observe that all filters have a lowpass characteristic and that the XPM filters ($c \neq c'$) are more frequency-selective than their SPM counterparts ($c = c'$). This suggests that the modulation of the temporal characteristics of amplitude sequences via shaping will be more effective in suppressing NLI from XPM than from SPM. We also observe from the two SPM filters in Figure 3a that the intra-polarization coefficients are two times higher than the inter-polarization coefficients except for $n = 0$ (see Appendix). The two XPM filters for $c' = c + 1$ and $c' = c - 1$ in Figure 3a are time-reversed functions and therefore their magnitude frequency responses in Figure 3b overlap completely.

Figure 4 shows the normalized magnitude frequency response for the intra-polarization SPM filters for fiber links with 1, 10, 20, and 40 spans of 80 km. As the link length increases, the channel memory increases, and as a result, the lowpass filters become more selective. For this, the inset in Figure 4 shows the 3 dB filter bandwidth versus the fiber link length in a log-log plot. We observe that the trend is linear with an approximate slope of -1. This is consistent with channel memory being linearly proportional to link length, e.g., [15, Eq. (2)]. Similar observations apply for XPM filters, which we have not shown for brevity. Furthermore, we note the higher magnitude of filter frequency responses at zero for longer links, which indicates a more substantial overall nonlinear distortion.
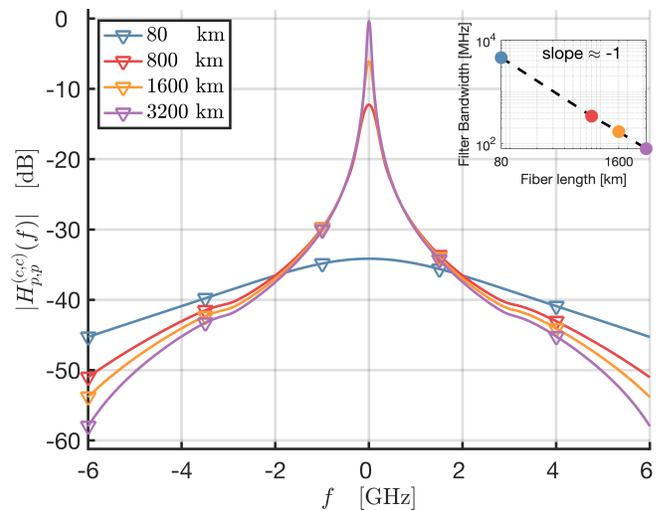

Fig. 4. Magnitude frequency response for intra-polarization SPM filter for a transmission with 32 GBd and 50 GHz channel spacing over fibers with various link lengths. Normalized to a maximum of 0 dB at $f = 0$. Inset: 3 dB filter bandwidth vs. fiber link length. The corresponding bandwidth of each filter is shown with the same color code.

Next, we consider a fiber link with a fixed length of 1600 km and different baud rates of 16 GBd, 32 GBd, and 64 GBd. Figure 5 shows the normalized magnitude frequency response for the intra-polarization SPM filter for these cases. A larger baud rate and thus signal bandwidth results in a more distinct lowpass characteristic in the frequency domain. The inset figure shows the 3 dB filter bandwidth versus transmission baud rate in a log-log plot. We observe that the bandwidth decreases as the baud rate increases with a linear slope of about -1. Again, this is consistent with channel memory, measured in time duration, being linearly proportional to baud rate, e.g., [15, Eq. (2)]. We also expect more significant nonlinear
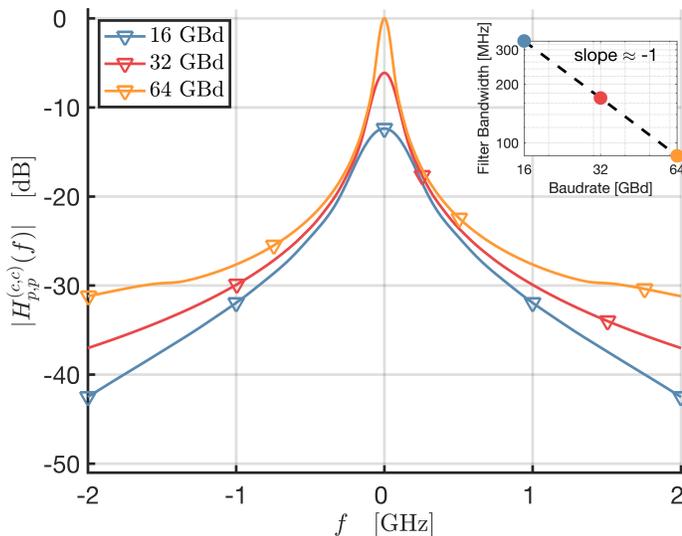
Fig. 5. Magnitude frequency response for intra-polarization SPM filter for a transmission over a 1600 km fiber link with three different baud rates. Normalized to a maximum of 0 dB at $f = 0$. Inset: 3 dB filter bandwidth vs. baud rate. The corresponding bandwidth of each filter is shown with the same color code.

Fig. 6. Left: Frequency domain representation of x-polarization symbol-energy sequences generated by CCDM with blocklength 108 and 300, and by uniform signaling and ideal AS. The launch power is -6.5 dBm per polarization. Right: Normalized magnitude frequency response of intra-polarization SPM filter. Dashed lines mark the 3 dB bandwidth for the filter.

SPM effects with increasing baud rate, which is confirmed by the larger frequency response value at zero. Based on the observations in Figures 4 and 5, we suggest the relationship BW $\propto \frac{1}{L \cdot B}$ for the 3 dB bandwidth of the SPM intra-polarization filter, fiber link length L, and baud rate B.

C. Interplay Between Filter and PAS Blocklength

The linear channel model is an effective tool to provide insights into the role of PAS in the generation of NLI. In particular, from the channel model (5) and the lowpass filter characteristics illustrated in the previous section, we conclude that symbol-energy sequences with weaker low-frequency components would result in less NLI. Such an argument has already been made in [13] using the simplified approximation (12) for comparing multidimensional mappings in conjunction with finite-length PAS.

Figure 6 shows the magnitude Fourier transform F_{x} of the symbol-energy sequence $\left(E_{\mathrm{x}}^{(c)}(n) - \bar{E}_{\mathrm{x}}^{(c)} \right)$ in the x-polarization generated by PAS and for 1D mapping. We note that we subtract the mean-energy term as suggested in (7). Link and system parameters from Table I are applied. For PAS, we consider CCDM blocklength 108 and 300, and ideal AS, where we directly draw amplitude symbols from the desired distribution. The curve for uniform signaling is also included as a reference. Overlaid to these curves is the magnitude frequency response of the intra-polarization SPM filter, with the axis on the right-hand side of the figure. The figure reveals several important observations that support previous findings considering effective SNR, e.g., [9]. First, ideal AS, which corresponds to practical AS such as CCDM with an infinite blocklength, causes more severe NLI than uniform signaling. Second, finite-length CCDM creates sequences with a high-pass characteristic, which becomes more pronounced as blocklength decreases. Hence, due to the lowpass channel
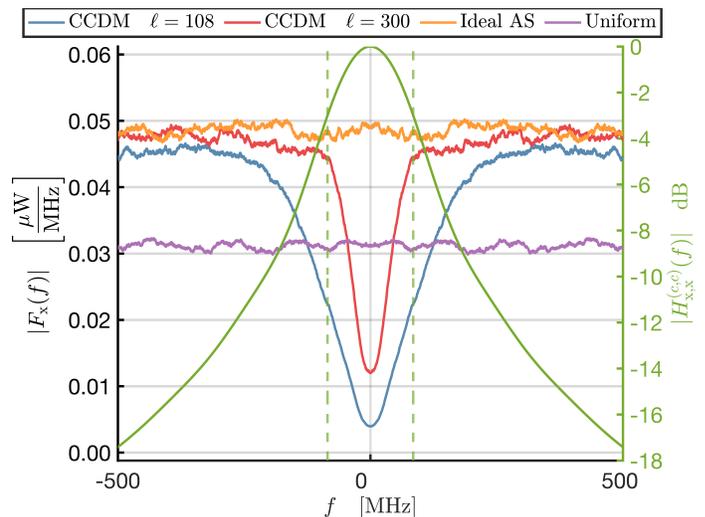
filter, nonlinear distortion is reduced, and shaped transmission can become more nonlinearity tolerant than uniform signaling. This is the reason for the nonlinear shaping gain obtained using PAS with short blocklengths. The exact cross-over point for when a gain will be observed depends on the filter bandwidth.

D. Role of Multidimensional Mapping

We can apply the same methodology as above to explain the role of multidimensional mapping of shaped sequences (see Section II-B) for the nonlinearity tolerance of PAS. However, before doing so, we consider a summary statistic called running digital sum (RDS), which was suggested in [13], to motivate the effect of multidimensional mapping on NLI. The RDS is based on the energy terms $2E_p^{(c')} + E_{p' \neq p}^{(c')}$ in approximation (12). The authors of [13] only considered the sum $E_p^{(c')} + E_{p' \neq p}^{(c')}$, which is common for both polarizations $p \in \{\mathrm{x, y}\}$, and introduced the RDS as

$$\tilde{\lambda}_{\mathrm{RDS}}(t) = \sum_{n=0}^{t-1} \sum_{c' \in \mathcal{C}} \sum_{p' \in \mathcal{P}} \left(E_{p'}^{(c')}(n) - \bar{E}_{p'}^{(c')} \right) \quad (13)$$

for $t \in \mathbb{N}$. We modify this definition by considering all energy terms in (12), i.e., we have the RDS for polarization p as

$$\lambda_{\mathrm{RDS}}(t) = \tilde{\lambda}_{\mathrm{RDS}}(t) + \sum_{n=0}^{t-1} \sum_{c' \in \mathcal{C}} \left(E_p^{(c')}(n) - \bar{E}_p^{(c')} \right), \quad (14)$$

which is a more accurate predictor for nonlinearity tolerance as we will show.

The RDS can be interpreted as a lowpass filtered version of the symbol-energy sequence. A small absolute RDS would thus suggest a weaker nonlinear distortion. For example, for CCDM with blocklength ℓ and 1D mapping (see Figure 2a), we have $\tilde{\lambda}_{\mathrm{RDS}}(t) = \lambda_{\mathrm{RDS}}(t) = 0$ for $t = k \cdot \ell$ and $k \in \mathbb{N}$, i.e., the RDS remains smaller for decreasing shaping

blocklength. A similar argument applies to sphere-shaping based AS such as ESS, as most of the sequences are selected from energies near the surface of a sphere. We can extend the discussion to 2D mapping as done in [13]. When CCDM-shaped amplitude sequences are mapped in two dimensions (see Figure 2b), the sequence energy is spread over length $\ell/2$ and thus $\tilde{\lambda}_{\mathrm{RDS}}(t) = \lambda_{\mathrm{RDS}}(t) = 0$ for all $t = k \cdot \ell/2$, which predicts a further reduction in NLI. However, for 4D mapping (see Figure 2c), the energy is spread over both polarizations, and while $\tilde{\lambda}_{\mathrm{RDS}}(t) = 0$ for all $t = k \cdot \ell/4$, this is not the case for $\lambda_{\mathrm{RDS}}(t)$. Based on the behavior of $\tilde{\lambda}_{\mathrm{RDS}}(t)$, it was suggested in [13] that PAS with 4D and 2D mapping is generally superior to PAS with 2D and 1D mapping in terms of nonlinearity tolerance, respectively. While we concur with the advantage of 2D over 1D mapping, our refined RDS metric (14) does not permit the conclusion that 4D mapping is generally advantageous.

We illustrate our point by plotting the absolute value of the Fourier transform F_p^{s} of

$$E_p^{\mathrm{s}}(n) = 2\left(E_p^{(c)}(n) - \bar{E}_p^{(c)}\right) + \left(E_{p' \neq p}^{(c)}(n) - \bar{E}_{p' \neq p}^{(c)}\right), \quad (15)$$

which is the aggregate symbol-energy sequence in (12), for $p = \mathrm{x}$ in Figure 7. We consider CCDM with blocklength 180 and 1D, 2D, and 4D mapping, and we overlay these curves with the magnitude frequency response of the inter-polarization SPM filter. Again, link and system parameters from Table I are applied. We observe that the curve for 2D mapping is always below that for 1D mapping, which means that 2D mapping always performs better than 1D mapping with respect to NLI. The comparison with 4D mapping is not that simple, as its curve intersects with those for 1D and 2D mapping. It has a shallower but wider dip around zero frequency than the 1D and 2D mapping curves. Hence, 4D mapping does not necessarily result in a reduced (or increased) NLI compared to 1D and 2D mapping, but it depends on the optical fiber link, which determines the lowpass filter curve in Figure 7. In particular, we expect 4D mapping to (only) provide benefits for shorter fiber links, which produce wider lowpass filters, and long shaping blocklengths, which lead to a narrower spectral dip. This conclusion is consistent with the results in [12] and the performance results presented in Section V.

IV. METRIC FOR PAS WITH SEQUENCE SELECTION

In this section, we utilize the linear channel model (5) to derive a metric for PAS with sequence selection as introduced in Section II-C.

A. The LSAS Metric

An effective selection of shaped sequences suppresses the distortion variations defined in (7). Hence, the proposed metric is

$$\lambda_{\mathrm{LSAS}} = \sum_{n=0}^{\ell/d-1} \sum_{p \in \mathcal{P}} \left| s_p^{(c)}(n) \sum_{p' \in \mathcal{P}} \sum_{c' \in \tilde{\mathcal{C}} \subseteq C} \Delta D_{p,p'}^{(c,c')}(n) \right|^2, \quad (16)$$
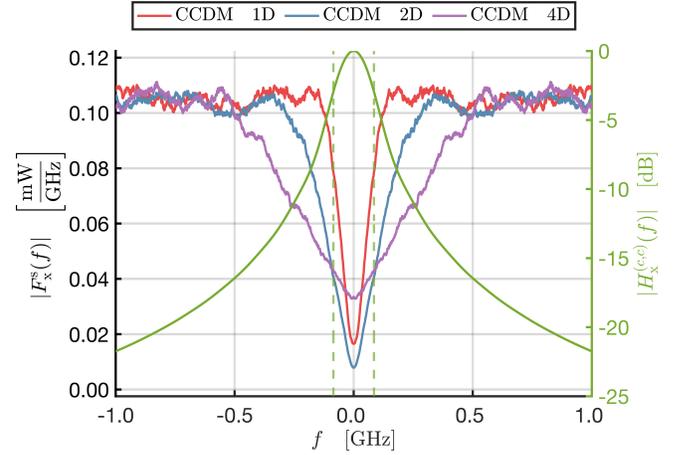

Fig. 7. Left: Frequency domain representation of aggregate symbol-energy sequence from (15) for x-polarization generated by CCDM with blocklength 180 and 1D, 2D, and 4D mapping. CCDM with $\ell = 180$. Launch power is -6.5 dBm per polarization. Right: Normalized magnitude frequency response of SPM filter $h_p^{(c,c')}$ from (11). Dashed lines mark the 3 dB bandwidth for the filter.

where, we recall, ℓ is the blocklength of the AS and d the dimensionality of the mapping. $D_{p,p'}^{(c,c')}$ is obtained from the convolution of the energy sequence of each candidate with the filters $h_{p,p'}^{(c,c')}$ in (7) and only the central ℓ/d convolution outputs are used in (16). This renders λ_{LSAS} independent for successive shaping blocks. The metric accounts for nonlinearity expressed in terms of symbol amplitudes and filtering through the fiber channel, and we established in the previous section that the filters have lowpass characteristics. Therefore, we refer to it as the lowpass filtered symbol amplitude sequence (LSAS) metric, as already used in [24] for the special case of 1D mapping. Since LSAS measures the total NLI from a multidimensional shaped amplitude sequence with length ℓ/d, the selection will be based on the minimum LSAS metric. In relation to the analysis in Section III-C, we expect that LSAS-based sequence selection results in a temporal structure of symbol-energy sequences with a deeper and wider spectral dip.

We note that we consider XPM nonlinearity from only a subset $\tilde{\mathcal{C}}$ of WDM channels in the LSAS definition (16) because the significant interference is caused by adjacent channels. Ideally, a joint selection of amplitude sequences for all channels is performed. However, since then the number of candidates would grow exponentially with the number of WDM channels, we chose to perform the selection of sequences for different WDM channels sequentially, i.e., in a greedy fashion, from outer to inner channels of the system. The details of our greedy selection approach are provided in Algorithm 1, for a WDM system with a set of channels $\mathcal{C}$, set of candidate sequences $\{s^{(c,\nu)}, c \in \mathcal{C}, \nu \in \{1, \ldots, 2^{\frac{2|\mathcal{P}|}{d} v}\}\}$, and considering the XPM effects from the N_{XPM} closest outer channels. Fig. 8 illustrates the execution of Algorithm 1 for a WDM system with $|\mathcal{C}| = 5$ channels and $N_{\mathrm{XPM}} = 1$.
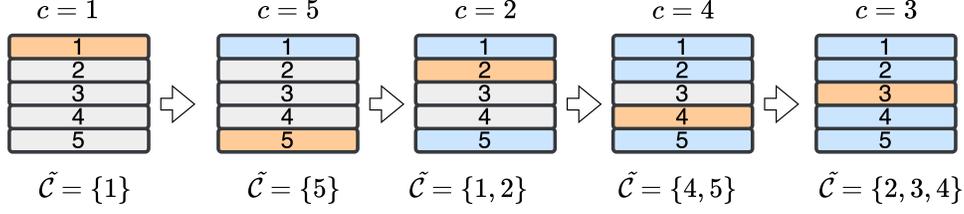
Fig. 8. Sequence selection for a WDM system with 5 channels $\mathcal{C} = \{1, 2, 3, 4, 5\}$ and $N_{\mathrm{XPM}} = 1$. c and $\tilde{\mathcal{C}}$ denote the channel under selection and the subset of channels from which nonlinear effects are accounted for, respectively. The channel under selection is highlighted in orange, while the channels with selected symbols are shown in light blue.

Algorithm 1 Greedy LSAS sequence selection algorithm for WDM systems

---

Input: set of channels $\mathcal{C} = \{1, 2, \ldots, |\mathcal{C}|\}$, set of candidates $\{s^{(c,\nu)}, c \in \mathcal{C}, \nu \in \{1, \ldots, 2^{\frac{2|\mathcal{P}|}{d}v}\}\}$, number of considered XPM channels N_{XPM}
Output: selected sequences $\{\hat{s}^{(c)}, c \in \mathcal{C}\}$
1: $\mathcal{T} \leftarrow \emptyset$ ▷ Set of channels for which selection is completed
2: **for** $\alpha \in \{1, 2, \ldots, \frac{|\mathcal{C}|+1}{2}\}$ **do** ▷ Loop over channels from outer to center
3:      **for** $\beta \in \{0, 1\}$ **do** ▷ For channels equidistant to the center channel
4:          $c \leftarrow (-1)^\beta \times \alpha + \beta \times (|\mathcal{C}| + 1)$ ▷ Channel under selection
5:          $\mathcal{K} \leftarrow \{c - N_{\mathrm{XPM}}, \ldots, c, \ldots, c + N_{\mathrm{XPM}}\}$
6:          $\tilde{\mathcal{C}} \leftarrow (\mathcal{T} \cap \mathcal{K}) \cup \{c\}$
7:          $\hat{s}^{(c)} \leftarrow$ candidate with minimum LSAS
8:          $\mathcal{T} \leftarrow \mathcal{T} \cup \{c\}$
9:      **end for**
10: **end for**

B. Relation to the EDI Metric

It is insightful to compare the LSAS metric with the EDI metric from [18]. For this purpose, we show how the latter can be derived as a simplified version of the former. We specialize in the case of 1D mapping, single polarization, and single channel transmission, as EDI has been defined for this scenario.

The EDI is computed based on the empirical average and variance of symbol-energy sequences. If we approximate the average $\bar{E}_p^{(c)}$ in (8) with the empirical mean

$$\tilde{E}_p^{(c)} = \frac{1}{\ell - w} \sum_{n=1+w/2}^{\ell-w/2} |s_p^{(c)}(n)|^2, \qquad (17)$$

where w is the EDI window length, and introduce the average-free symbol-energy sequence

$$\Delta\tilde{E}_{p,p}^{(c,c)}(n) = \sum_{m=-w/2}^{w/2-1} (E_p^{(c)}(n+m) - \tilde{E}_p^{(c)}), \qquad (18)$$

then the EDI metric can be written as

$$\lambda_{\mathrm{EDI}} = \frac{1}{(\ell - w - 1)w\tilde{E}_p^{(c)}} \sum_{n=1+w/2}^{\ell-w/2} \left(\Delta\tilde{E}_{p,p}^{(c,c)}(n)\right)^2. \quad (19)$$

The relation to the LSAS metric follows from noticing that the sequence $\Delta\tilde{E}_{p,p}^{(c,c)}(n)$ is an approximation of the distortion variation $\Delta D_{p,p}^{(c,c)}(n)$ from (7). In particular, the two quantities would be identical if $\bar{E}_p^{(c)} = \tilde{E}_p^{(c)}$ and

$$h_{p,p}^{(c,c)}(n) = \begin{cases} 1, & \text{if } n \in \{-\frac{w}{2}, \ldots, \frac{w}{2} - 1\} \\ 0, & \text{otherwise} \end{cases} \qquad (20)$$

were true. Hence, different from the LSAS metric, the EDI in (19) does not include the amplitude terms $|s_p^{(c)}(n)|$, uses an approximation of the lowpass filter of the channel, and it discards $w/2$ symbols at both ends of the length-ℓ shaped sequence to have meaningful empirical estimates. In the performance results section, we show that for these reasons, LSAS outperforms EDI for sequence selection.

C. Computational Complexity Analysis

While the main focus of our work has been on insights and improvement of NLI tolerance through shaping, we would also like to provide expressions for the associated computational cost. For this, we count the number of real-valued additions and multiplications, i.e., floating-point operations (flops), that are required for candidate generation and metric calculation.

Denoting the number of operations for sequence generation as O_{gen} and the number of operations for metric calculation as O_{met}, the total number of real-valued operations for shaping with sequence selection is given by

$$O_{\mathrm{tot}} = \underbrace{2^{\frac{2|\mathcal{P}|v}{d}}}_{\substack{\text{number of} \\ \text{candidates}}} O_{\mathrm{met}} + \underbrace{2^v}_{\substack{\text{number of} \\ \text{sequences}}} O_{\mathrm{gen}}. \qquad (21)$$

The generation of each QAM sequence with length ℓ requires $O_{\mathrm{gen}} = 2 \times 2^{m-1}$ and $O_{\mathrm{gen}} = 4 \times 2^{m-1}$ flops per QAM symbol for ESS/K-ESS and arithmetic coding implementation of CCDM (AC-CCDM), respectively [2].

The EDI metric calculation with window size w, candidate length ℓ, single polarization, and 1D mapping requires

$$O_{\mathrm{met}} = 6 + w - \frac{w^2 + 2w - 4}{\ell} \qquad (22)$$

flops per QAM symbol. We note that EDI complexity increases linearly with window size w in the long blocklength regime $w \ll \ell$, where EDI is most useful. Finally, the number of flops per QAM symbol for evaluating the LSAS metric using filters of length q, mapping dimension d, $|\mathcal{P}|$ polarizations, shaping blocklength ℓ, and $|\mathcal{C}|$ channels can be obtained as

$$O_{\mathrm{met}} = 6 + |\mathcal{P}|N_{\mathrm{avg}} - \frac{d}{|\mathcal{P}|\ell} + \frac{|\mathcal{P}|N_{\mathrm{avg}}O_{\mathrm{conv}}d}{\ell}. \qquad (23)$$

In (23), O_{conv} is the number of flops required for each convolution in (16), and N_{avg} is the average number of convolutions that are evaluated for the metric calculation of different channels as

$$N_{\mathrm{avg}} = \frac{2 + (2N_{\mathrm{XPM}} + 1) + (|\mathcal{C}| - 3)(N_{\mathrm{XPM}} + 1)}{|\mathcal{C}|}, \qquad (24)$$

TABLE I
SIMULATION PARAMETERS FOR SETUP 1 [18]

Parameter	Value
Modulation	256 QAM
Amplitude shaper	CCDM
Shaping rate	2.4 bits/amplitude
Polarization	Single/Dual
Center wavelength	1550 nm
Symbol rate	32 GBd
WDM spacing	50 GHz
# WDM channels	11
Total bandwidth	550 GHz
Pulse shape	Root-raised cosine
Pulse roll-off	0.1
Span length	80 km
# Spans	20
Fiber loss	0.2 dB/km
Dispersion parameter	17 ps/nm/km
Nonlinearity parameter (γ)	1.37 1/W/km
EDFA noise figure	6 dB
Oversampling factor	2
# QAM symbols per run	243,000
# Simulation runs	5

TABLE II
SIMULATION PARAMETERS FOR SETUP 2 [21]

Parameter	Value
Modulation	64 QAM
Amplitude shaper	ESS/K-ESS
Shaping rate	1.5 bits/amplitude
Polarization	Dual
Center wavelength	1550 nm
Symbol rate	50 GBd
WDM spacing	55 GHz
# WDM channels	1/11
Total bandwidth	55/605 GHz
Pulse shape	Root-raised cosine
Pulse roll-off	0.1
Span length	205 km
# Spans	1
Fiber loss	0.2 dB/km
Dispersion parameter	17 ps/nm/km
Nonlinearity parameter (γ)	1.30 1/W/km
EDFA noise figure	5 dB
Oversampling factor	2
# QAM symbols per run	324,000
# Simulation runs	5

where N_{XPM} is the number of closest outer channels considered in the greedy procedure of sequence selection (see Algorithm 1). For simplicity, we assumed that the same filter lengths are used in all convolutions, which makes the expression in (23) an upper bound. We note that the LSAS complexity scales with the product $N_{\mathrm{avg}}O_{\mathrm{conv}}$, where O_{conv} is a monotonically increasing function of filter length q when $\ell/d \geq \frac{q+1}{2}$.

For typical settings (see Section V), $O_{\mathrm{gen}} \lesssim O_{\mathrm{met}}$. Since furthermore $2^v \leq 2^{\frac{2|\mathcal{P}|v}{d}}$, we will present a quantitative comparison for O_{met} with EDI and LSAS in Section V-C.

V. PERFORMANCE RESULTS

Sequence selection using EDI and nonlinearity tolerant amplitude shaping using K-ESS are state-of-the-art schemes for mitigating nonlinear effects in optical fiber channel. Accordingly, we consider the two simulation setups from the respective references to showcase the validity of the energy-based model and effectiveness of LSAS for sequence selection.

A. Simulation Parameters

The parameters for the two simulation setups are adopted from [18] and [21] and shown in Tables I and II, respectively.

a) Setup 1 [18]: A multi-span WDM transmission with 11 channels, 32 GBd baud rate, and 50 GHz channel spacing is simulated using the SSFM. Two times oversampling is applied to account for fiber nonlinearity. The modulation is 256 QAM with root raised cosine pulse shaping (roll-off 0.1). Amplitudes are generated using CCDM with a Maxwell-Boltzmann distribution to achieve an effective rate of 2.4 bits/amplitude, i.e., the extra rate loss due to sequence selection is compensated for by increasing the AS shaping rate.. The PAS scheme uses a low-density parity-check (LDPC) code with rate 4/5. The overall rate is thus 5.2 bit/QAM-symbol. The symbols are transmitted through 20 spans of a standard single-mode fiber with 80 km span length, fiber loss 0.2 dB/km, chromatic

dispersion parameter 17 ps/nm/km, and nonlinearity parameter $1.37~\mathrm{W}^{-1}\mathrm{km}^{-1}$. At the end of each span, an erbium doped amplifier (EDFA) with a 6 dB noise figure is deployed. Both single and dual polarization transmissions are considered. At the receiver, chromatic dispersion is compensated, and a mean phase rotation is applied as a CPR, which is the same as used in [18] and [21]. The central channel is the channel of interest for performance results. Results are obtained as the average from transmitting 243,000 QAM symbols in five independent simulation runs.

b) Setup 2 [21]: The main differences from Setup 1 are the adoption of 64 QAM transmission, PAS with ESS and K-ESS with a shaping rate of 1.5 bits/amplitude and a rate-5/6 FEC, i.e., the overall rate is 4 bit/QAM-symbol, a baud rate of 50 GBd, a channel spacing of 55 GHz for the WDM case, and a single-span link of 205 km. We note that there are several pairs of maximum energy and kurtosis that result in K-ESS with the same shaping rate. We used the optimal pair for blocklength $\ell = 108$ as reported in [21] for a fair comparison. As in [21], we will also consider 64 QAM transmission with uniform signaling and a rate 2/3-FEC as a baseline to compare various shaped transmission systems.

The other parameters are identical or similar to Setup 1 (see Tables I and II).

B. Performance metrics

We use three common performance metrics for assessing the nonlinearity tolerant schemes. The first is the effective SNR of symbols after linear equalization:

$$\mathrm{SNR}_{\mathrm{eff}} = \frac{\mathbb{E}\left[|S|^2\right]}{\mathbb{E}\left[|S - R|^2\right]}, \tag{25}$$

where S and R denote transmitted and received symbol, respectively. The second is the Q-factor, which is defined based on uncoded bit error rate (BER) as

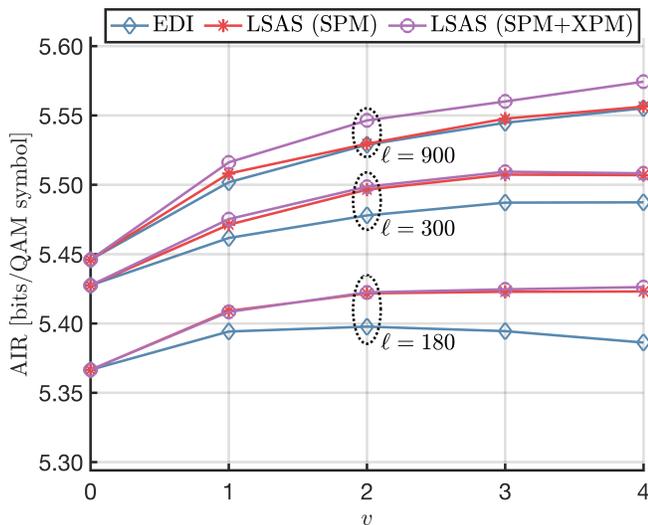$$\mathrm{Q} = \sqrt{2}\mathrm{erfc}^{-1}(2\mathrm{BER}), \tag{26}$$

Fig. 9. AIR vs. v for sequence selection using EDI, LSAS with only SPM components, and LSAS with both SPM and XPM components. Setup 1 with single-polarization transmission and 1D mapping.

where erfc denotes the complementary error function. The third metric is the AIR [27]–[29]

$$\text{AIR} = H(P_{\text{s}}) - \sum_{i=1}^{2m} H(P_{\text{b}|\text{r},i}) - 2R_{\text{loss},v}, \qquad (27)$$

where P_{s} and $P_{\text{b}|\text{r},i}$ denote the distributions of the transmitted QAM symbol and its ith bit label given the received symbol, respectively. The effective SNR is useful to highlight the strength of NLI associated with different shaping schemes. Q-factor and AIR also account for the linear shaping gain.

C. Setup 1: 256 QAM Transmission

We start with 256 QAM transmission in Setup 1 to compare the effectiveness of the LSAS metric for sequence selection with that of the EDI metric from [18].

a) Sequence Selection: We first focus on sequence selection with 1D mapping. If not stated otherwise, single-polarization transmission is used. For SNR and AIR results, a launch-power optimization has been performed for each measurement point.

Figure 9 shows the AIR as a function of the number of flipping bits v for PAS with different blocklengths and EDI and LSAS metrics for sequence selection. EDI uses a window length of $w = 100$, as suggested in [18]. For the LSAS metric, we distinguish two cases: (i) $\tilde{\mathcal{C}} = \{c\}$, i.e., only SPM is considered ("LSAS (SPM)"), and (ii) $\tilde{\mathcal{C}} = \{c-1, c, c+1\}$, i.e., XPM from two adjacent channels is considered ("LSAS (SPM+XPM)") for the center channel of interest. We observe that LSAS is superior to EDI for all blocklengths. For $\ell \in \{180, 300\}$, we suspect that discarding w symbols from each shaping block as well as the lowpass-filter approximation affect the EDI metric. Accounting for XPM does not make a notable difference for short blocklengths, which is due to the wide spectral dip for CCDM at $\ell \in \{180, 300\}$ (see Figure 6) and the narrow XPM filter (see Figure 3b). For $\ell = 900$, the LSAS is superior to the EDI metric because of the inclusion

of XPM in the selection. Since this requires knowledge of the data in neighboring channels, it could be impractical in a conventional WDM scenario. However, it would be a legitimate option for a digital subcarrier multiplexing (DSCM) system. Furthermore, we note that there is an optimum value for v, which shifts toward larger v for larger blocklengths. This is because of the trade-off between linear and nonlinear shaping gain in the sequence selection scheme. An initially higher linear shaping gain in the case of larger blocklength provides more opportunity for increasing the nonlinear shaping gain using more redundant bits. However, since the computational complexity of sequence selection grows with v, we choose $v = 2$ for the following results.

Figure 10 compares the sequence selection metrics from Figure 9 for different shaping blocklengths. Shaping without sequence selection, i.e., $v = 0$ is shown as the baseline curve. Selection using EDI converges to the baseline for short blocklengths, as it requires a blocklength ℓ notably larger than the window size w to work effectively (see Section IV-B). As blocklength increases, EDI and LSAS with only SPM converge to the same performance. As already noted in Figure 9, LSAS with SPM and XPM components outperforms the other selection metrics for larger ℓ, with a gain of about 0.3 dB in SNR and 0.1 bits/QAM-symbol in AIR over shaping without selection at $\ell = 900$. From the AIR in Figure 10b we note that the optimal shaping blocklength increases from $\ell = 900$ in the case of no selection to $\ell = 1350$ for selection using LSAS, which is due to the reduced nonlinear impairments.

To gauge the computational complexity associated with the two sequence selection metrics, we evaluate the expressions (22) and (23) in Section IV-C for two shaping blocklengths $\ell = 108$ and $\ell = 900$ used in Figure 10. The LSAS metric uses SPM and XPM filters of lengths 400 and 500, respectively, as shown in Figure 3. For these values, the evaluation of EDI and LSAS (SPM) require 11 and 220 flops per QAM symbol, respectively. At the blocklength $\ell = 900$, we have 95, 719, and 1367 flops per QAM symbol for EDI, LSAS (SPM), and LSAS (SPM+XPM), respectively, i.e., incorporating the XPM terms from the first nearest outer channels increases the computational complexity by a factor of $N_{\text{avg}} \approx 2$. We thus observe a significantly increased computational complexity for LSAS compared to EDI, which is due to the scaling of complexity with filter length and N_{avg}. However, the LSAS filter lengths can be shortened with little effect on the performance results. For example, we can simply truncate the LSAS filters using a threshold compared to the maximum coefficient. Table III shows the number of flops and the corresponding AIR for different truncation thresholds. We observe that the computational complexity of LSAS can be reduced to values almost identical to those for EDI, i.e., 16 vs. 11 flops at a threshold of -8 dB for $\ell = 108$ and 87 vs. 95 flops at a threshold of -15 dB for $\ell = 900$, with a negligible loss in AIR. We speculate that this can further be improved with more sophisticated filter approximation approaches.

In the remainder of this section, results for LSAS refer to LSAS with only SPM components, which is suitable for blocklengths $\ell \leq 300$.

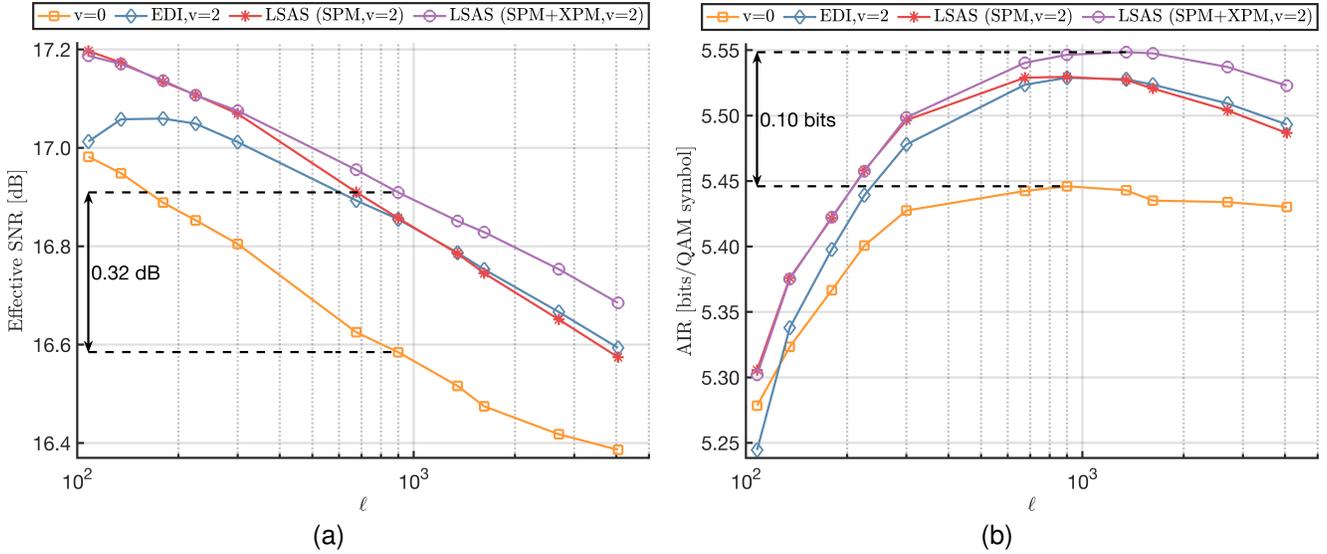Next, we present results for dual-polarization transmission.

Fig. 10. (a) Effective SNR and (b) AIR vs. shaping blocklength for different selection metrics with $v = 2$ redundant bits. Shaping without selection ($v = 0$) shown as a reference. Setup 1 with single-polarization transmission and 1D mapping.

TABLE III
NUMBER OF REAL-VALUED OPERATIONS PER QAM SYMBOL AND AIR (IN [BITS/QAM SYMBOL]) FOR LSAS (SPM) AT $\ell = 108$ AND $\ell = 900$ FOR DIFFERENT TRUNCATION THRESHOLDS (IN dB).

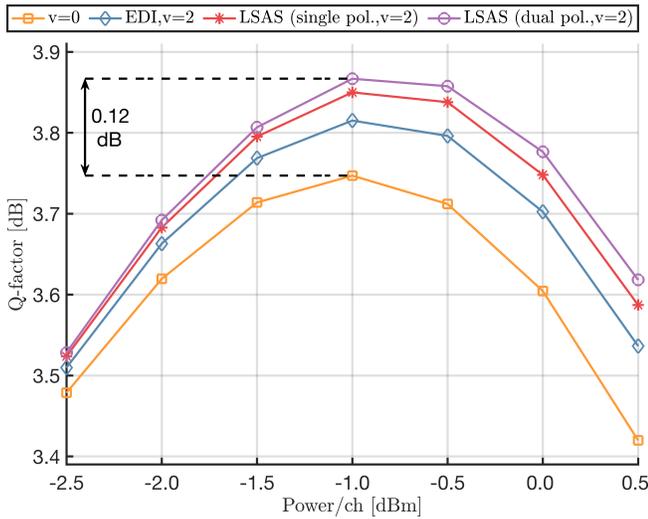| ℓ	Threshold	$-\infty$	-20	-15	-8
108	O_{LSAS}	220	145	80	16
	AIR	5.30	5.30	5.30	5.29
900	O_{LSAS}	719	176	87	
	AIR	5.53	5.53	5.52	

Fig. 11. Q-factor vs. launch power for different selection metrics with $v = 2$ redundant bits. Shaping without selection ($v = 0$) shown as a reference. Setup 1 with dual-polarization transmission, 1D mapping, and shaping blocklength $\ell = 180$.
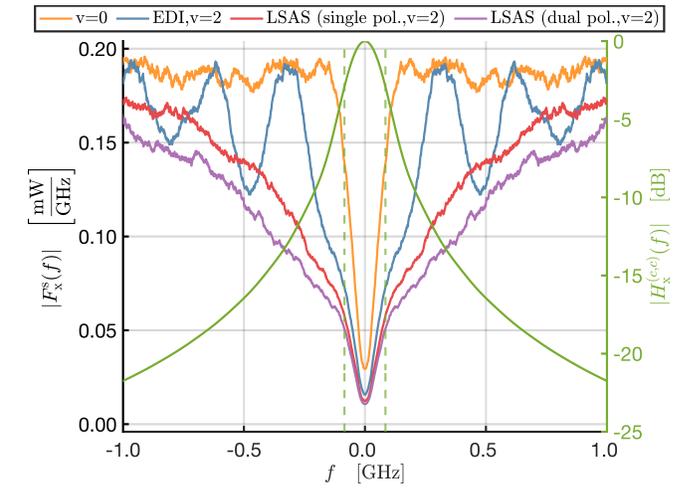

Fig. 12. Left: Frequency domain representation of aggregate symbol-energy sequence from (15) for x-polarization for shaping methods and system setup from Figure 11. Launch power is -1 dBm per channel. Right: Normalized magnitude frequency response of SPM filter $h_p^{(c,c')}$ from (11). Dashed lines mark the 3 dB bandwidth for the filter.

Figure 11 shows the Q-factor as a function of launch power for dual-polarization transmission and shaping with different selection metrics. In the case of the LSAS metric, we distinguish between ($\mathcal{P} = \{x\}, \mathcal{P} = \{y\}$) and $\mathcal{P} = \{x, y\}$ in (16),

i.e., a single-polarization and a dual-polarization version. We observe that the advantage of the LSAS metric also manifests in the dual-polarization case, with some slight additional gain due to performing joint selection over the two polarization dimensions. We highlight the effect of shaping sequence selection and the different selection metrics in Figure 12, which shows the Fourier transforms F_p^{s} of the aggregate symbol-energy sequences defined in (15) corresponding to the different shaping methods. We overlay the curves with the frequency response for the SPM filter $h_p^{(c,c')}$ defined in (11), which provides the approximation for the distortion in (12). It can be seen that sequence selection makes the spectral dip wider and deeper, which translates to a decrease in NLI considering the lowpass channel filter. Moreover, the order of
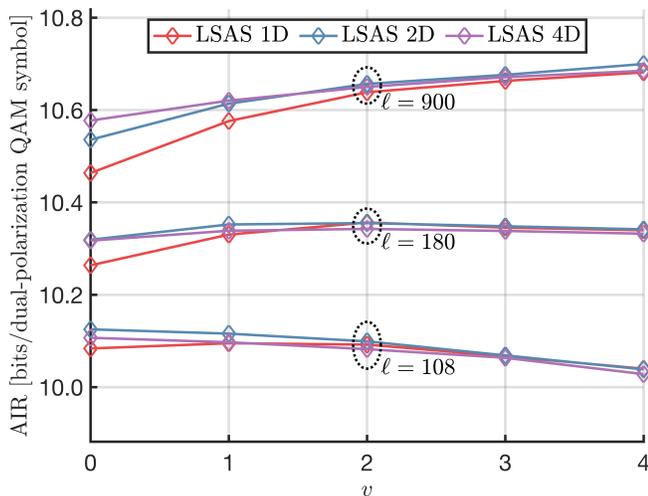
This article has been accepted for publication in IEEE/OSA Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2023.3264032

12

Fig. 13. AIR vs. v for sequence selection with LSAS and different mapping strategies. Setup 1 with dual-polarization transmission.

Fig. 14. Effective SNR vs. launch power for Setup 2. (a) Single channel transmission and (b) WDM transmission with 11 channels. Shaping block-length $\ell = 108$ and 4D mapping.

the curves in Figure 12 matches the order of the corresponding curves in Figure 11. This again demonstrates the usefulness of the linear channel model for nonlinearity analysis. As a final detailed observation, we point out the oscillation of the Fourier transform signal for the EDI metric in Figure 12. This results from rectangular-filter approximation in (20), which corresponds to a sinc-function in the frequency domain, and thus frequency components at its nulls do not contribute to the metric.

b) Multidimensional mapping: As the last set of results for Setup 1, we consider sequence selection with different amplitude mapping strategies. For this, we need to adopt the dual-polarization transmission scenario. Furthermore, since the components of the 4D signal need to be chosen jointly, only the dual-polarization version of the LSAS metric but not the EDI can be applied. Figure 13 shows the AIR as a function of v for the three mapping strategies from Section II-B. First, we observe that for shaping without selection ($v = 0$), the dimensionality of mapping has a notable effect on the system performance. 4D mapping is superior to 1D and 2D mapping strategies only for long blocklengths. This result has been predicted in Section III-D from the channel model for distortion and the frequency-domain visualization in Figures 6 and 7. Second, as the number of selection bits increases, the performances for all three mapping strategies become identical even though they select among different number of candidates. Therefore, it is preferable to use 4D mapping in conjunction with sequence selection for its lower complexity. Regarding the latter, one could also compare the different mapping strategies for the same number of candidates instead of fixed shaping blocklength ℓ. For example, denoting the number of redundant bits for mapping dimension d as v_d, then we can compare $v_1 = 1$ from the 1D mapping curve with $v_2 = 2$ and $v_4 = 4$ from the 2D and 4D mapping curves in Figure 13, respectively. From this comparison we observe that multidimensional mapping is superior to 1D mapping for relatively longer blocklengths, i.e., $\ell = 900$ in Figure 13. For shorter blocklengths, it is preferable to use a smaller number
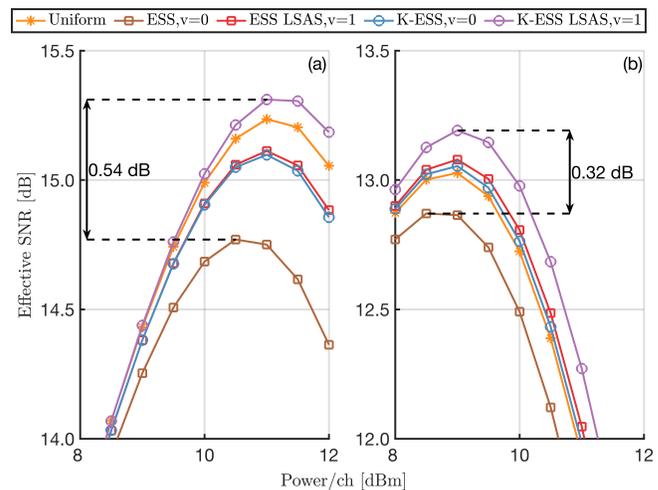
of candidates to avoid the larger rate loss with increasing v.

D. Setup 2: 64 QAM transmission

For Setup 2 from [21], we focus on the comparison of sequence selection using LSAS with K-ESS, which was proposed for improved nonlinearity tolerance in [21], as well as for the first time, the combination of the two methods.

Figure 14 shows the effective SNR for the various shaping schemes for (a) single channel and (b) WDM transmission. PAS with $\ell = 108$ and 4D mapping are applied, where the latter is beneficial because of the short link length. The effective SNR for uniform signaling is included as a reference. We observe that PAS with ESS suffers an SNR loss compared to uniform signaling, which is due to an increased NLI. A good fraction of this loss (single-channel case) or even the entire loss (WDM case) is recovered using K-ESS or adding sequence selection with LSAS. The combination of K-ESS with LSAS-based selection outperforms all methods with a gain of more than 0.5 dB and 0.3 dB over ESS-based shaping in the single-channel and WDM cases, respectively. The SNR-performance results can be explained by considering the frequency-domain representation of the corresponding symbol-energy signals and the linear channel filter in Figure 15. It can be seen that K-ESS is effective in lowering the higher-frequency components, which is a result of generating a distribution with lower kurtosis. The selection-based schemes optimize the temporal structure of symbols to minimize the NLI with the same distribution, which translates to a wider and deeper spectral dip at frequency zero. As a result, we expect that the selection-based schemes perform well for a wide variety of transmission setups, while K-ESS is only beneficial for short fiber links corresponding to a wide lowpass filter. This conclusion is consistent with the observations in [21].

Finally, to examine our hypothesis about K-ESS, we consider Setup 2 but with a single channel transmission through a long-haul fiber link with 20 spans of length 80 km. Figure 16 shows the Q-factor curves obtained with ESS and K-ESS
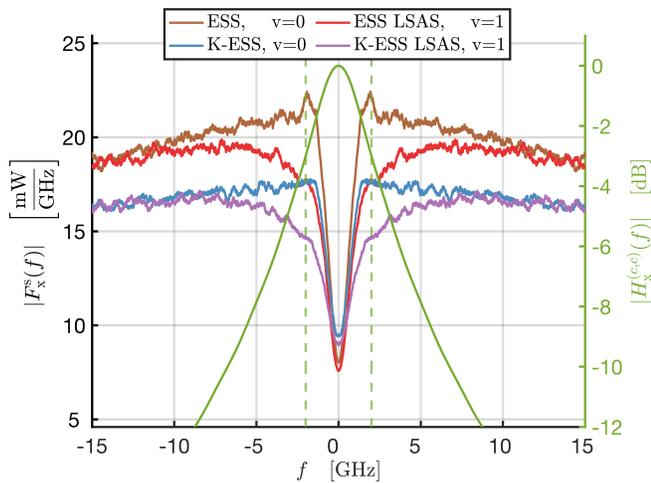
Fig. 15. Left: Frequency domain representation of aggregate symbol-energy sequence from (15) for x-polarization for shaping methods from Figure 14(b). Launch power is 9 dBm per channel. Right: Normalized magnitude frequency response of SPM filter $h_p^{(c,c')}$ from (11). Dashed lines mark the 3 dB bandwidth for the filter.
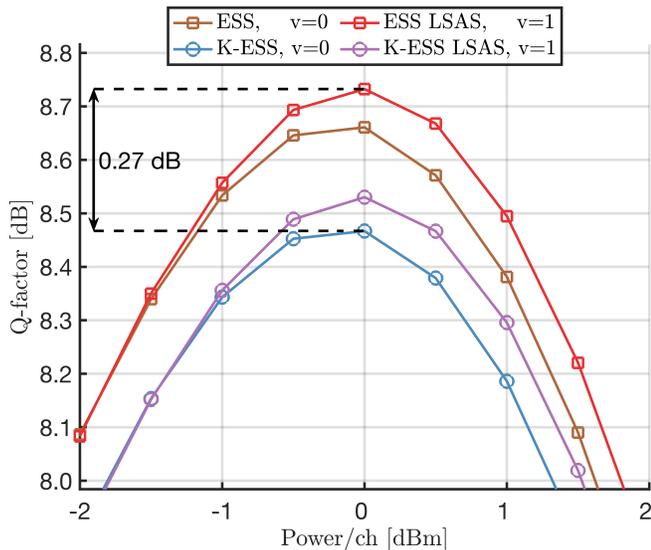

Fig. 16. Q-factor vs. launch power for Setup 2 but with a single channel transmission through 20 spans of 80 km. Shaping blocklength $\ell = 108$ and 2D mapping strategy.

with $\ell = 108$ and with and without LSAS-based sequence selection. 2D mapping is chosen as it is the best mapping strategy for this longer link. As suggested by our linear channel model analysis, ESS outperforms K-ESS in this scenario. Furthermore, sequence selection provides a gain when applied to either baseline AS.

VI. FURTHER DISCUSSION

In the derivation of the LSAS metric we applied several simplifications that permitted us to apply sequence selection for each shaping block independently. First, we ignored the terms in the perturbative model (28) that also depend on signal phases. This is because the latter are determined by the sign bits that are generated by the FEC encoder (see Section II-A). As one FEC block typically consists of several shaping blocks,

sign bits and thus signal phases are dependent on multiple shaping blocks. Second, the channel memory, i.e., the convolution of energy sequences with the filters $h_{p,p'}^{(c,c')}$ in (7), creates dependencies between NLI of adjacent shaping blocks. In the LSAS metric (16) we ignored these dependencies by applying energy sequences for each shaping block separately to the convolution operation. Despite these simplifications, our results in Section V showed that the proposed linear channel model and selection metric are sensitive enough to guide the interaction between shaping and fiber nonlinearity. A natural extension of this work is a selection mechanism that accounts for sign bits and fully includes the channel memory. In this case, a joint selection over multiple shaping blocks is in order. However, a naive approach would substantially increase the computational complexity, as the number of possible sequences grows exponentially with the number of shaping blocks. Therefore, it is interesting to explore possible alternative methods that do not suffer from the complexity growth.

Besides the sign bits, we note that the implementation of the CPR also affects the nonlinear gain of PAS [25]. The proposed LSAS metric accounts for a mean phase rotation at the receiver. Similarly, the simulation setup used in Section V applies such a CPR, which has been consistent with the setups in related previous works [18], [21]. This choice is motivated as it enables us to better delineate the effects of nonlinearity tolerant shaping and different selection metrics, which is the main focus of this paper. To further demonstrate the performance with sequence selection entirely independent of a CPR implementation, we consider the nonlinear phase noise introduced by the fiber channel based on the additive-multiplicative distortion model obtained from first-order perturbation theory [30]. We apply Setup 1 with single polarization transmission and evaluate the phase noise terms in the distortion model to account for SPM and XPM from the two adjacent channels on either side of the channel of interest. In Figure 17, we present the estimated standard deviation of nonlinear phase noise plotted versus the shaping blocklength for different selection metrics. The figure compares PAS without sequence selection ($v = 0$), selection using EDI, and selection using LSAS. The results show that sequence selection successfully suppresses nonlinear phase noise. Similar to the effective-SNR performance metric in Figure 10a, selection using the EDI metric does not produce any nonlinear gain in the short blocklength regime. However, for longer blocklengths, all selection metrics decrease the standard deviation of nonlinear phase noise. Furthermore, the order of curves in Figure 17 matches the effective-SNR result in Figure 10a. Considering the interplay of PAS and CPR for suppressing nonlinear phase noise, a desirable extension of this work is the formulation of a sequence-selection metric that accounts for receiver processing with a more potent CPR method than a mean phase rotation.

VII. CONCLUSION

In this paper, we revisited the signal-dependent generation of NLI in optical channels and presented a linear lowpass channel model to approximately describe this process. The
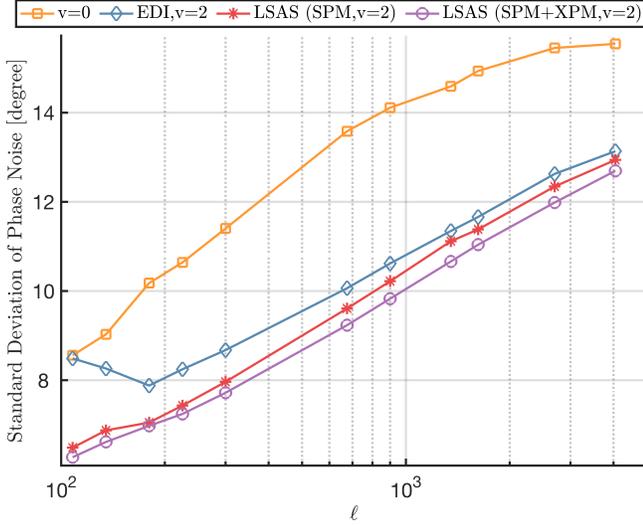
Fig. 17. Estimated standard deviation of nonlinear phase noise vs. shaping blocklength for PAS without ($v = 0$) and with sequence selection ($v = 2$). Sequence selection using EDI and LSAS metrics. Setup 1 with single-polarization transmission and 1D mapping.

model considers symbol-energy sequences as input and therefore permits the analysis of the interplay between amplitude shaping and nonlinear distortion. We showed the evolution of filter coefficients for different fiber setups and introduced a frequency domain visualization as a useful tool for the analysis and design of PAS schemes for nonlinearity mitigation. We applied the model to develop the LSAS metric for PAS with sequence selection. LSAS naturally includes inter-polarization and inter-channel nonlinear effects and extends to multidimensional mapping. The latter permits the combination of sequence selection with multidimensional mapping for lower complexity. Through extensive simulations, we showed the performance benefits resulting from LSAS, and we confirmed the accuracy of the predictions based on the linear model.

APPENDIX

In this appendix, we derive the approximation (4) for the linear channel model between symbol-energy sequences and nonlinear distortions.

We start with the expression [31, Eq. (3)] for the nonlinear distortion term from signal-signal interactions based on the first-order perturbation analysis of a multi-span dual-polarization coherent WDM system. Accordingly, the input-output relationship for the x-polarization signal in WDM

channel $c \in \mathcal{C}$ after linear equalization can be written as

$$
\begin{aligned}
r_{\mathrm{x}}^{(c)}(n) &= s_{\mathrm{x}}^{(c)}(n) \\
&+ \mathrm{j}\gamma \sum_{m\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\bigg(s_{\mathrm{x}}^{(c)}(m+n)[s_{\mathrm{x}}^{(c)}(k+m+n)]^*s_{\mathrm{x}}^{(c)}(k+n) \\
&+ s_{\mathrm{y}}^{(c)}(m+n)[s_{\mathrm{y}}^{(c)}(k+m+n)]^*s_{\mathrm{x}}^{(c)}(k+n) \bigg) h_{\mathrm{spm}}^{(c)}(m,k) \\
&+ \mathrm{j}\gamma \sum_{\substack{c'\in\\ \mathcal{C}\setminus\{c\}}}\sum_{m\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\bigg(2s_{\mathrm{x}}^{(c')}(m+n)[s_{\mathrm{x}}^{(c')}(k+m+n)]^*s_{\mathrm{x}}^{(c)}(k+n) \\
&+ s_{\mathrm{y}}^{(c')}(m+n)[s_{\mathrm{y}}^{(c')}(k+m+n)]^*s_{\mathrm{x}}^{(c)}(k+n) \\
&+ s_{\mathrm{x}}^{(c')}(m+n)[s_{\mathrm{y}}^{(c')}(k+m+n)]^*s_{\mathrm{y}}^{(c)}(k+n) \bigg) h_{\mathrm{xpm}}^{(c,c')}(m,k),
\end{aligned}
\tag{28}
$$

where $s_p^{(c)}(n)$ and $r_p^{(c)}(n)$ are the nth transmitted symbol and received sample in polarization $p \in \{\mathrm{x,y}\}$ and channel $c \in \mathcal{C}$, respectively, and γ is the fiber nonlinearity parameter. $h_{\mathrm{spm}}^{(c)}$ and $h_{\mathrm{xpm}}^{(c,c')}$ contain the perturbation coefficients for intra-channel and inter-channel NLI, respectively. An analogous expression to (28) can be written for the y-polarization signal.

To arrive at the representation used for the distortion model based on symbol-energy sequences, we approximate (28) by retaining only the symbol-energy dependent terms in the summations:

$$
\begin{aligned}
r_{\mathrm{x}}^{(c)}(n) &\approx s_{\mathrm{x}}^{(c)}(n)\Big[1+ \\
&\mathrm{j}\gamma\sum_{m\in\mathbb{Z}}\bigg(|s_{\mathrm{x}}^{(c)}(m+n)|^2+|s_{\mathrm{y}}^{(c)}(m+n)|^2\bigg)h_{\mathrm{spm}}^{(c)}(m,0) \\
&+ \mathrm{j}\gamma\sum_{k\in\mathbb{Z}\setminus\{0\}}\bigg(|s_{\mathrm{x}}^{(c)}(k+n)|^2\bigg)h_{\mathrm{spm}}^{(c)}(0,k) \\
&+ \mathrm{j}\gamma\sum_{\substack{c'\in\\ \mathcal{C}\setminus\{c\}}}\sum_{m\in\mathbb{Z}}\bigg(2|s_{\mathrm{x}}^{(c')}(m+n)|^2+|s_{\mathrm{y}}^{(c')}(m+n)|^2\bigg)h_{\mathrm{xpm}}^{(c,c')}(m,0)\Big].
\end{aligned}
\tag{29}
$$

Then, noting that $h_{\mathrm{spm}}^{(c)}(m,0) = h_{\mathrm{spm}}^{(c)}(0,m)$ [32], and defining filters $h_{p,p'}^{(c,c')}$ as

$$
\begin{aligned}
c'=c, p'=p &: h_{p,p'}^{(c,c')}(n) = \begin{cases} 2h_{\mathrm{spm}}^{(c)}(n,0), & n\neq 0 \\ h_{\mathrm{spm}}^{(c)}(n,0), & n=0 \end{cases} \\
c'=c, p'\neq p &: h_{p,p'}^{(c,c')}(n) = h_{\mathrm{spm}}^{(c)}(n,0) \\
c'\neq c, p'=p &: h_{p,p'}^{(c,c')}(n) = 2h_{\mathrm{xpm}}^{(c)}(n,0) \\
c'\neq c, p'\neq p &: h_{p,p'}^{(c,c')}(n) = h_{\mathrm{xpm}}^{(c)}(n,0)
\end{aligned}
\tag{30}
$$

to describe both intra- and inter-channel NLI effects, we obtain (4) in Section III.

Furthermore, as it can be seen in (30), the intra-polarization coefficients are two times the inter-polarization coefficients, except for $n = 0$ for SPM. This motivates the approximation of the relation between filter coefficients in (11) in Section III-A.

REFERENCES

[1] R. F. Fischer, *Precoding and signal shaping for digital transmission.* John Wiley & Sons, 2005.

[2] Y. C. Gültekin, T. Fehenberger, A. Alvarado, and F. M. Willems, "Probabilistic shaping for finite blocklengths: Distribution matching and sphere shaping," *Entropy*, vol. 22, no. 5, p. 581, 2020.

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

[4] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "On shaping gain in the nonlinear fiber-optic channel," in *IEEE International Symposium on Information Theory (ISIT)*, 2014, pp. 2794–2798.

[5] H. E. Hansen, M. P. Yankov, L. K. Oxenløwe, and S. Forchhammer, "Optimization of probabilistic shaping for nonlinear fiber channels with non-Gaussian noise," *Entropy*, vol. 22, no. 8, p. 872, 2020.

[6] J. Cho and R. Tkach, "On the kurtosis of modulation formats for characterizing the nonlinear fiber propagation," *Journal of Lightwave Technology*, 2022.

[7] C. Pan and F. R. Kschischang, "Probabilistic 16-QAM shaping in WDM systems," *Journal of Lightwave Technology*, vol. 34, no. 18, pp. 4285–4292, 2016.

[8] T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, "On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel," *Journal of Lightwave Technology*, vol. 34, no. 21, pp. 5063–5073, 2016.

[9] T. Fehenberger, D. S. Millar, T. Koike-Akino, K. Kojima, K. Parsons, and H. Griesser, "Analysis of nonlinear fiber interactions for finite-length constant-composition sequences," *Journal of Lightwave Technology*, vol. 38, no. 2, pp. 457–465, 2019.

[10] T. Fehenberger, H. Griesser, and J.-P. Elbers, "Mitigating fiber nonlinearities by short-length probabilistic shaping," in *Optical Fiber Communication Conference (OFC)*, 2020, pp. Th1I–2.

[11] T. Fehenberger, "On the impact of finite-length probabilistic shaping on fiber nonlinear interference," in *Signal Processing in Photonic Communications.* Optical Society of America, 2020, pp. SpTh1I–3.

[12] P. Skvortcov, I. Phillips, W. Forysiak, T. Koike-Akino, K. Kojima, K. Parsons, and D. S. Millar, "Huffman-coded sphere shaping for extended-reach single-span links," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 27, no. 3, pp. 1–15, 2021.

[13] W.-R. Peng, A. Li, Q. Guo, Y. Cui, and Y. Bai, "Transmission method of improved fiber nonlinearity tolerance for probabilistic amplitude shaping," *Optics Express*, vol. 28, no. 20, pp. 29430–29441, 2020.

[14] W.-R. Peng, A. Li, Q. Guo, Y. Cui, and Y. Bai, "Baud rate and shaping blocklength effects on the nonlinear performance of super-symbol transmission," *Optics Express*, vol. 29, no. 2, pp. 1977–1990, 2021.

[15] K. Wu, G. Liga, A. Sheikh, F. M. Willems, and A. Alvarado, "Temporal energy analysis of symbol sequences for fiber nonlinear interference modelling via energy dispersion index," *Journal of Lightwave Technology*, vol. 39, no. 18, pp. 5766–5782, 2021.

[16] K. Wu, G. Liga, Y. C. Gültekin, and A. Alvarado, "Exponentially-weighted energy dispersion index for the nonlinear interference analysis of finite-blocklength shaping," in *European Conference on Optical Communication (ECOC)*, 2021, pp. 1–4.

[17] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, 2015.

[18] K. Wu, G. Liga, A. Sheikh, Y. C. Gültekin, F. M. Willems, and A. Alvarado, "List-encoding CCDM: A nonlinearity-tolerant shaper aided by energy dispersion index," *Journal of Lightwave Technology*, vol. 40, no. 4, pp. 1064–1071, 2022.

[19] M. Secondini, S. Civelli, E. Forestieri, and L. Z. Khan, "New lower bounds on the capacity of optical fiber channels via optimized shaping and detection," *Journal of Lightwave Technology*, vol. 40, no. 10, pp. 3197–3209, 2022.

[20] A. Amari, S. Goossens, Y. C. Gültekin, O. Vassilieva, I. Kim, T. Ikeuchi, C. M. Okonkwo, F. M. Willems, and A. Alvarado, "Introducing enumerative sphere shaping for optical communication systems with short blocklengths," *Journal of Lightwave Technology*, vol. 37, no. 23, pp. 5926–5936, 2019.

[21] Y. C. Gültekin, A. Alvarado, O. Vassilieva, I. Kim, P. Palacharla, C. M. Okonkwo, and F. M. Willems, "Kurtosis-limited sphere shaping for nonlinear interference noise reduction in optical channels," *Journal of Lightwave Technology*, vol. 40, no. 1, pp. 101–112, 2021.

[22] Y. C. Gültekin, A. Alvarado, O. Vassilieva, I. Kim, P. Palacharla, C. M. Okonkwo, and F. M. Willems, "Mitigating nonlinear interference by limiting energy variations in sphere shaping," in *Optical Fiber Communication Conference (OFC)*, 2022.

[23] M. Fu, Q. Liu, H. Lun, H. Jiang, Y. Wu, X. Liu, Z. Yang, L. Yi, W. Hu, and Q. Zhuge, "Parallel bisection-based distribution matching for nonlinearity-tolerant probabilistic shaping in coherent optical communication systems," *Journal of Lightwave Technology*, vol. 39, no. 20, pp. 6459–6469, 2021.

[24] M. T. Askari, L. Lampe, and J. Mitra, "Nonlinearity tolerant shaping with sequence selection," in *European Conference on Optical Communications (ECOC)*, 2022, pp. 1–4.

[25] S. Civelli, E. Forestieri, and M. Secondini, "Interplay of probabilistic shaping and carrier phase recovery for nonlinearity mitigation," in *European Conference on Optical Communications (ECOC)*, 2020, pp. 1–4.

[26] R. R. Borujeny and F. Kschischang, "Why constant-composition codes reduce nonlinear interference noise," *arXiv preprint arXiv:2203.02641*, 2022.

[27] A. Alvarado, T. Fehenberger, B. Chen, and F. M. Willems, "Achievable information rates for fiber optics: Applications and computations," *Journal of Lightwave Technology*, vol. 36, no. 2, pp. 424–439, 2017.

[28] T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, "On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel," *Journal of Lightwave Technology*, vol. 34, no. 21, pp. 5063–5073, 2016.

[29] T. Fehenberger, D. S. Millar, T. Koike-Akino, K. Kojima, and K. Parsons, "Multiset-partition distribution matching," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 1885–1893, 2018.

[30] X. Liang and S. Kumar, "Multi-stage perturbation theory for compensating intra-channel nonlinear impairments in fiber-optic links," *Optics express*, vol. 22, no. 24, pp. 29733–29745, 2014.

[31] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Inter-channel nonlinear interference noise in WDM systems: Modeling and mitigation," *Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1044–1053, 2014.

[32] W. Yan, Z. Tao, L. Dou, L. Li, S. Oda, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "Low complexity digital perturbation back-propagation," in *European Conference on Optical Communication (ECOC)*, 2011, pp. 1–3.

Mohammad Taha Askari received the B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2015, and the M.A.Sc. degree in electrical and computer engineering from the University of British Columbia, Vancouver, BC, Canada, in 2022. He is currently pursuing the Ph.D. degree in electrical and computer engineering at the University of British Columbia. His field of interest includes probabilistic shaping for optical fiber communications, machine learning for communications, and information and coding theory.

Lutz Lampe received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Erlangen, Erlangen, Germany, in 1998 and 2002, respectively. Since 2003, he has been with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada, where he is a Full Professor. His research interests include theory and application of wireless, optical wireless, optical fiber, power line, and underwater acoustic communications. He has served as an associate editor and a guest editor for several IEEE journals, and as the general co-chair and the technical program committee co-chair for IEEE conferences. He has been a Distinguished Lecturer of the IEEE Communications Society and a (co-)recipient of a number of best paper awards. He is a Co-Editor of the book Power Line Communications: Principles, Standards and Applications from Multimedia to Smart Grid (2nd edition, John Wiley & Sons).

Jeebak Mitra currently works as a Member of Technical Staff in the Advanced Wireless Technologies (AWT) group at the office of the CTO, Dell Technologies, involved in design and development of next generation mobile network architectures and is responsible for various strategic initiatives related to Dell's wireless portfolio. Prior to that, Jeebak worked as a principal engineer at Huawei Technologies Research Center in Ottawa, Canada from 2013 to 2022, working as part of a world class DSP team that designed several generations of market leading optical DSP transceivers focusing primarily on algorithm design and implementation for optical links for metro and transport network. From 2010 to 2013, he was a technical leader for DSP architecture and algorithms for wireless communications with several startup companies focused on wireless backhaul and access, all of which were acquired. He received the M.A.Sc. and Ph.D. degrees in Electrical Engineering from The University of British Columbia in 2005 and 2010, respectively. His research interests lie in the area of high-performance communication systems design focusing on wireless and optical transmission. Jeebak has over 20 issued patents and over 50 publications in various conferences and journals and received the Best Student Paper Award at the IEEE Canadian Conference in Electrical and Computer Engineering 2009. He was a co-recipient of the Best Paper Award at IEEE/ACM/IFIP CNSM 2017 and 2019.