Volterra-assisted Optical Phase Conjugation: a Hybrid Optical-Digital Scheme For Fiber Nonlinearity Compensation

Gabriel Saavedra, *Member, IEEE,* Gabriele Liga, *Member, IEEE,* and Polina Bayvel *Fellow, IEEE, Fellow, OSA*

Abstract—Digital nonlinearity compensation (NLC) schemes such as digital backpropagation (DBP) and Volterra equalization are well-known to be effective techniques in mitigating optical fiber nonlinearity, thus offering improved transmission performance. Alternatively, optical NLC, and specifically optical phase conjugation (OPC), has been proposed to relax the digital signal processing complexity.

In this work, a novel hybrid optical-digital NLC scheme combining OPC and a Volterra equalizer is proposed, termed Volterra-Assisted OPC (VAO). It has a twofold advantage: it overcomes the OPC limitation in asymmetric links and substantially enhances the performance of Volterra equalizers. When NLC is operated over the entire transmitted optical bandwidth, the proposed scheme is shown to outperform both OPC and Volterra equalization alone by up to 4.2 dB in a 5-channel, 32 GBaud PM-16QAM transmission over a 1000 km EDFA-amplified fiber link. Moreover, VAO is also demonstrated to be very robust when applied to long-transmission distances, with a 2.5 dB gain over OPC-only systems at 3000 km.

VAO combines the advantages of both optical and digital NLC offering a promising trade-off between performance and complexity for future high-speed optical communication systems.

Index Terms—Fiber nonlinear optics, optical signal processing, digital signal processing, non- linear equalization, Volterra series.

I. INTRODUCTION

FIBER nonlinearity represents a major bottleneck in increasing transmission rates of optical communication systems [1], [2]. As a result, finding a high-performance, low-complexity nonlinearity compensation (NLC) technique is a central challenge in the design of next-generation fiber systems. There are two main approaches to fiber nonlinearity mitigation, namely i) digital NLC techniques; and ii) optical NLC techniques.

Digital NLC is based on digital signal processing (DSP) (either at the transmitter or at the receiver) to mitigate nonlinear distortions accumulated during signal propagation. Digital backpropagation (DBP) is indisputably the most widespread among these digital techniques [3], [4]. DBP aims to compensate for nonlinearities by virtually propagating the received

G. Saavedra, and P. Bayvel are with the Optical Networks Group, Department of Electronic and Electrical Engineering, UCL, Torrington Place, London WC1E 7JE, UK. e-mail: (uceegsa@ucl.ac.uk).

G. Liga was previously with the Optical Networks Group. He is now with the Signal Processing Systems (SPS) Group, Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB, Eindhoven, The Netherlands (e-mail: g.liga@tue.nl).

optical field back to the transmitter in an inverted digital link. The most studied implementation of DBP to date is based on solving the Manakov equation, which describes the signal evolution in an optical fiber, through the well-known split-step Fourier method (SSFM). The SSFM algorithm is recursive and in general complex, as a minimum number of iterations (or *steps*), is required to achieve a given target performance. Such a minimum required number was shown to rapidly increase as a function of the NLC bandwidth [5], [6].

An alternative zero-forcing equalization method to DBP is based on the Volterra series expansion of the solution of the Manakov equation [7]. Volterra-based equalizers were first introduced for NLC in coherent optical communications in [8], [9], initially through a time-domain approach, and subsequently [10]–[13] in the frequency domain. A Volterra-series frequency-domain equalizer (VSFE) consists in reconstructing the transmitted optical field by using a Volterra series expansion of the Manakov equation (in the *backward direction*). This expansion, first derived in [7], is usually truncated to the third-order term for complexity reasons [10]. This leads to suboptimal NLC performance, as nonlinear distortions are only approximately compensated for at the receiver. Previous works have shown a potential advantange in complexity for VSFE over DBP [12], [14], [15], especially due to the highly parallel structure of the VSFE scheme.

Alternative to digital NLC, optical NLC techniques are based on the idea of using the physics of fiber propagation itself to compensate for, or prevent detrimental nonlinear effects on the received signal. These techniques include optical phase conjugation (OPC) [16]–[19] and phase-conjugated twin wave (PCTW) transmission [20]. Mid-link OPC consists in performing a conjugation of the optical field at the mid-point of the transmission link, such that nonlinear effects produced in the first part of the link are cancelled by the propagation effects in the second half. Mid-link OPC can jointly compensate for chromatic dispersion and nonlinear effects within the transmission link [16], [19], with the big advantage, compared to digital NLC schemes, of compensating for inter-channel effects without requiring joint detection/processing of multiple channels. As such a compensation takes place in the optical domain, NLC can be performed jointly over very large optical bandwidths, e.g., in [21] NLC was achieved over 92 channels using OPC for long-haul transmission systems . Performing digital NLC over such bandwidths would be infeasible due to the high computational complexity. In addition, OPC can notionally provide nonlinearity suppression without sacrificing

any of the available transmission degrees of freedom (e.g., orthogonal polarizations or bandwidth). A major shortcoming of OPC is the stringent requirement of a symmetric link power profile. Optical links with symmetric power profile at the mid-link point can only be designed using distributed amplification schemes. Moreover, even with Raman amplifiers, the power profile symmetry condition requires complex amplifier configurations [22], [23] that can be hard to achieve in practice. On the other hand, performing OPC in links with a certain degree of power profile asymmetry can result in a significantly reduced ability to reduce fiber nonlinear effects [18], [24]. In addition to this, OPC devices are never ideal as they typically include an optical amplifier to overcome negative conversion efficiencies and add amplified spontaneous emission (ASE) noise in the process. This leads to a performance penalty, which further reduces their nonlinearity mitigation capabilities [24].

In [25], a novel approach to nonlinearity compensation was proposed that combined OPC and Volterra equalization, termed Volterra-assisted OPC (VAO). In this paper, we extend the results through a detailed analysis of the advantages of VAO over each of its constituent techniques. The proposed approach enhances the gain of OPC in links with a low degree of power profile symmetry, such as EDFA-amplified links, through the use of VSFE. Moreover, it applies OPC as a way to boost the NLC performance of VSFE in large NLC bandwidth scenarios.

Hybrid optical NLC schemes were previously studied, e.g. in [26], with joint operation of OPC and DBP. However, this combination was designed to relax the link placement of the OPC device, rather than a way to enhance NLC performance. The technique proposed in this work is shown to markedly outperform either OPC or VSFE schemes when individually used in EDFA-amplified multi-span transmission systems, over large NLC bandwidths.

II. VOLTERRA-ASSISTED OPC

To analyze the effect of nonlinearities generated in an optical fiber transmission system, a possible approach is to use a frequency domain Volterra series expansion. Usually, for optical fiber propagation, a truncated Volterra series up to the third-order term is adopted, due to the numerical complexity in calculating the higher order terms of the series. The third-order term coincides with the first-order regular-perturbation solution to the Manakov equation (see [27]) and is a powerful analytical tool to explore the characteristics of the nonlinear interference in the optical fiber channel [7], [28]. For multi-span, EDFA-amplified, dispersion-uncompensated transmission links, the third-order Volterra series term is given by[1]:

$$
\begin{aligned}
A_{3X}(\omega) = j\gamma\frac{8}{9}\iint & [\mathcal{S}_{XXX}(\omega,\omega_1,\omega_2) \\
& +\mathcal{S}_{YYX}(\omega,\omega_1,\omega_2)]\,\Xi\,(N_s,\omega,\omega_1,\omega_2) \\
& \times F(\omega,\omega_1,\omega_2)d\omega_2 d\omega_1,
\end{aligned}
\tag{1}
$$

[1]In this work, functions are denoted by uppercase letters $G(\cdot)$, boldface uppercase letters (e.g. Ξ) denote matrices, whereas lowercase letters with either 2 or 3 sub-index (e.g. a_{ilk}) denote elements of a 2D/3D tensor.

where γ is the nonlinearity coefficient of the fiber,

$$
\begin{aligned}
\mathcal{S}_{XXX}(\omega,\omega_1,\omega_2) &\triangleq S_X^*(\omega_1)S_X(\omega_2)S_X(\omega+\omega_1-\omega_2), \\
\mathcal{S}_{YYX}(\omega,\omega_1,\omega_2) &\triangleq S_Y^*(\omega_1)S_Y(\omega_2)S_X(\omega+\omega_1-\omega_2),
\end{aligned}
\tag{2}
$$

are the *signal kernels*, and $S_X(\omega)$, $S_Y(\omega)$ are the transmitted signal spectra over two orthogonal polarizations X and Y, respectively. $\Xi\,(N_s,\omega,\omega_1,\omega_2)$ is the *phased-array* term given by

$$
\Xi\,(N_s,\omega,\omega_1,\omega_2) = \sum_{n=1}^{N_s} e^{-j\beta_2\Delta\Omega(n-1)L_s},
\tag{3}
$$

with $\Delta\Omega \triangleq (\omega-\omega_2)(\omega_1-\omega_2)$, and where β_2, L_s ,and N_s are the group velocity dispersion, the fiber span length, and the total number of spans, respectively. The factor $F(\omega,\omega_1,\omega_2)$ represents the *four-wave mixing efficiency* over one span, given by

$$
F(\omega,\omega_1,\omega_2) = \frac{1-e^{-\alpha L_s}e^{j\beta_2\Delta\Omega L_s}}{j\beta_2\Delta\Omega - \alpha},
\tag{4}
$$

where α is the fiber attenuation coefficient. The function of, ω_1 and ω_2 that multiplies the signal kernel (in this case $F(\omega,\omega_1,\omega_2)\cdot\Xi\,(N_s,\omega,\omega_1,\omega_2)$) is usually referred to *Volterra third-order kernel*.

When an OPC device is inserted in the transmission link with EDFA amplification, only part of the nonlinear distortion accumulated in the first part of the link is effectively cancelled by the OPC. Such a residual nonlinear term can be described using a third-order Volterra-series term. Similar to the case without OPC, the Volterra kernel can be obtained using a regular perturbation approach for this specific system configuration (see Appendix A). For a transmission link using mid-link OPC, the first and third-order Volterra series terms in frequency domain, are given by[2]

$$
A_{1X}(\omega) = S_X^*(-\omega),
\tag{5}
$$

$$
\begin{aligned}
A_{3X}(\omega) = j\gamma\frac{8}{9}\iint & [\mathcal{S}_{XXX}^*(-\omega,-\omega_1,-\omega_2) \\
& +\mathcal{S}_{YYX}^*(-\omega,-\omega_1,-\omega_2)]\,\Xi^*\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right) \\
& \times G(\omega,\omega_1,\omega_2)d\omega_2 d\omega_1.
\end{aligned}
\tag{6}
$$

The first-order Volterra series term simply corresponds to the conjugated version of the transmitted signal, as chromatic dispersion compensation is applied in-line by the mid-link OPC. Moreover, it is interesting to notice how the phased-array in the third-order term (Eq. (6)) only accounts for the accumulation of nonlinear effects over half of the link ($\frac{N_s}{2}$ fiber spans for a link of N_s spans). Finally, $G(\omega,\omega_1,\omega_2)$ is the characteristic nonlinear kernel for systems with OPC and is given by (see Appendix A)

$$
\begin{aligned}
G(\omega,\omega_1,\omega_2) &\triangleq \\
& \frac{(e^{-j\beta_2\Delta\Omega L_s}e^{-\alpha L_s}-1)(\alpha - j\beta_2\Delta\Omega)}{\alpha^2+\beta_2^2\Delta\Omega^2} \\
& +\frac{(e^{-j\beta_2\Delta\Omega L_s}-e^{-\alpha L_s})(\alpha + j\beta_2\Delta\Omega)}{\alpha^2+\beta_2^2\Delta\Omega^2}.
\end{aligned}
\tag{7}
$$

[2]Due to the symmetry of the Manakov equation, A_{1Y} and A_{3Y} can be obtained by simply swapping sub-indexes X and Y in (6) and (2).
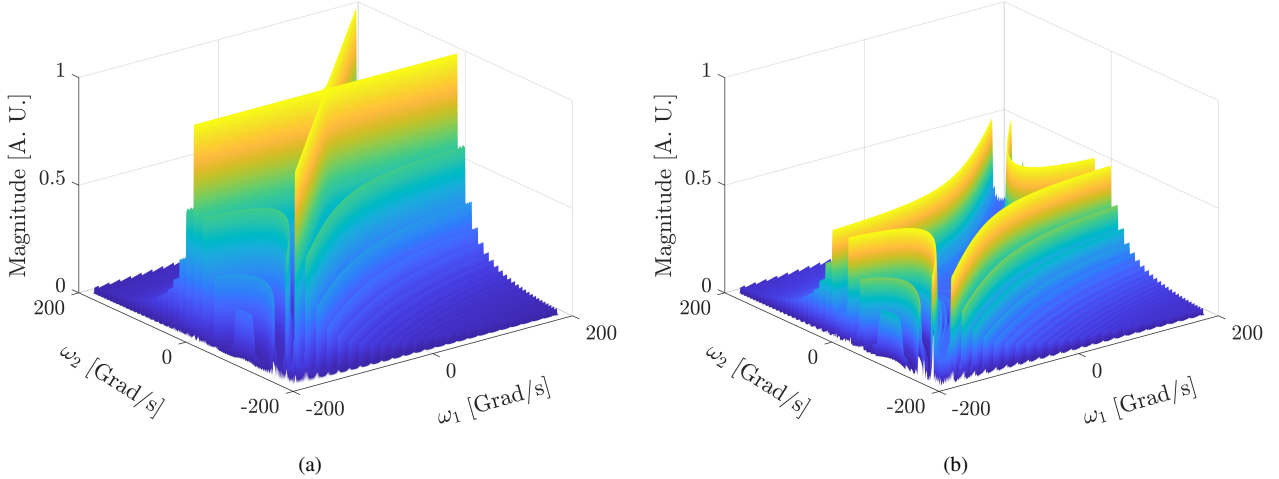
Fig. 1: Normalized absolute value of the third-order nonlinearity kernels for a 10×100 km SSMF, EDFA-amplified link (a) without OPC, and (b) with OPC.

Fig. 1 shows the normalized absolute value of the Volterra kernels in Eqs. (1) (Fig. 1a) and (6) (Fig. 1b) as a function of the 2 angular frequency variables ω_1 and ω_2 for $\omega = 0$, for an EDFA-amplified transmission link of 10x100 km spans of standard single mode fiber (SSMF). The OPC Volterra kernel in Fig. 1b is here normalised with respect to the maximum value of the kernel in (1). As expected, the kernel of the channel without OPC (Fig. 1a) shows a maximum coupling for $\omega_2 = 0$ and $\omega_1 = \omega_2$. A decrease in the coupling strength is instead observed as $|\omega_1 - \omega_2| \gg 0$ and $\omega_2 \neq 0$, due to the walk-off effect induced by fiber dispersion. The oscillations observed in the kernel arise from the coherent accumulation of nonlinearities at every span and are described by the phased-array factor.

From 1b it can be seen that OPC produces a kernel with a characteristic *dip* around the $(0,0)$ frequency and along the $\omega_1 = \omega_2$ and $\omega_2 = 0$ lines. This indicates a strong reduction of the nonlinear coupling in the low-frequency (ω_1, ω_2)–plane. On the other hand, higher frequency coupling is still present due to the power profile asymmetry[3]. As it can be observed from Eq. (6), where the phased-array is evaluated only over $N_s/2$ spans, OPC reduces the maximum magnitude of the nonlinear kernel by a factor of 2. The residual nonlinear interactions at the receiver, due to the lack of symmetry in EDFA-based links, limit the performance of OPC as an NLC method. Similar conclusions were previously derived using a different mathematical approach, such as the four-wave mixing product analysis presented in [29], [30] and follow-up publications [31]. However, no mitigation strategy specifically tailored to overcome this limitation has been proposed to date.

Eqs. (6) and (7) specify the design of a VSFE tailored to compensate for the residual nonlinearity in an EDFA-amplified link using OPC. A VSFE aims to reconstruct the transmitted field up to the third-order Volterra series term,

[3]Ideal mid-link OPC would have zero nonlinear coupling everywhere in the (ω_1, ω_2)–plane.

through a frequency-domain processing. This reconstruction can be done using: i) a fully non-recursive equalizer, that calculates the third-order Volterra kernel of the entire link in a single step (multispan VSFE) [7], [12, Sec. 2]; ii) a recursive, per-step, equalizer [10], [11]. Although the recursive method partly mitigates the so-called *energy divergence* problem, which leads to extremely inaccurate field reconstructions as the transmitted power increases [32], it also requires N_s times more calculations, for a fixed processed sequence length.

To take advantage of the action of the OPC, the VAO system proposed in this work naturally requires a non-recursive VSFE approach based on the kernel in (6) (as opposed to a per-span fiber kernel). This is due to the fact that the equalizer is supposed to benefit from the properties of the end-to-end kernel of the OPC channel, as opposed to the per-span kernel. The reason for that is twofold and it is fully justified from the results in this paper: i) OPC reduces the memory of the system as chromatic dispersion is compensated in a distributed manner along the link, thus allowing for a decreased interplay between nonlinearity and chromatic dispersion; ii) the reduction of nonlinearity induced by the OPC leads to a more accurate reconstruction of transmitted field using third-order Volterra equalizers.

III. VSFE AND VAO IMPLEMENTATION

Fig. 2 shows the difference between a conventional (non-recursive) VSFE equalizer and the Volterra equalizer in VAO. In both equalizers, N samples of the received time-domain signal (on each polarization) $\tilde{\mathbf{s}}_X \triangleq \{\tilde{s}_{X,k}\}_{k=0}^{N-1}$ over a time window T, are transformed into the frequency-domain vector $\mathbf{s}_X \triangleq \{s_{X,k}\}_{k=0}^{N-1}$, where

$$s_{X,k} = N \cdot \Delta\omega \cdot S_X(k\Delta\omega), \qquad k = -\frac{N}{2}, -\frac{N}{2}+1, ..., \frac{N}{2}-1 \tag{8}$$

and $\Delta\omega = \frac{1}{T}$ represents the sampling resolution in frequency-domain. The signal kernel matrix $\mathcal{S}_k \triangleq (\sigma_{ij} = s_i^* s_j s_{k+i-l})$

(or $\mathcal{S}_k^*$ in the case of VAO) is then computed and an Hadamard multiplication by the respective Volterra kernel matrix is performed[4]. The kernel matrices $\mathbf{F}_k \circ \boldsymbol{\Xi}_k(N_s)$, and $\boldsymbol{G}_k \circ \boldsymbol{\Xi}^*_k\left(\frac{N_s}{2}\right)$, for, respectively, the VSFE and VAO cases, are defined as

$$\mathbf{F}_k \triangleq (f_{ilk} = F'(k\Delta\omega, i\Delta\omega, l\Delta\omega)),$$
$$\boldsymbol{\Xi}_k \triangleq (\xi_{ilk} = \Xi'(k\Delta\omega, i\Delta\omega, l\Delta\omega')),$$
$$\boldsymbol{G}_k \triangleq (g_{ilk} = G'(k\Delta\omega, i\Delta\omega, l\Delta\omega)),$$
$$\forall \quad k, i, l = -\frac{N}{2}, -\frac{N}{2}+1, ..., \frac{N}{2}-1 \tag{9}$$

where F', Ξ', and G' are the backward direction counterparts of the functions F, Ξ, and G (see Appendix A).

Finally, both equalizers compute the reconstructed field (in frequency-domain) which, in the VAO case, is

$$y_{X,k} = s_{X,k} + a_{X,k},$$
$$k = -\frac{N}{2}, -\frac{N}{2}+1, ..., \frac{N}{2}-1. \tag{10}$$

where the sequence $a_{X,k}$ is the sampled version of the third-order Volterra series terms in Eqs. (1) and (6) and is given by

$$a_{X,k} = F_s A_{3X}(k\Delta\omega)$$
$$= \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{l=\max(0,k+i-N+1)}^{\min(k+i,N)} \xi_{ilk} g_{ilk} \left[s_{X,i} s_{X,l}^* s_{X,k+i-l}^* \right. \tag{11}$$
$$\left. + s_{Y,i} s_{Y,l}^* s_{X,k+i-l}^* \right]$$

where $F_s = \frac{N}{T}$ represents the sampling frequency. In the VSFE case, the chromatic dispersion accumulated in the link is then removed through the multiplication by $\mathbf{d} \triangleq \left\{ \exp\left(-j\frac{\beta_2}{2}k^2\Delta\omega^2\right) \right\}_{k=-\frac{N}{2}}^{\frac{N}{2}-1}$. Finally the sequence is transformed into the time-domain and, in the VAO case, and conjugated back to undo the effect of the OPC module (see Appendix A).

In this work, the proposed VAO method was numerically assessed in comparison with two conventional VSFE implementations: single-step and recursive VSFE. The single-step version of the VSFE was implemented using the schematic diagram in Fig. 2a. The recursive VSFE was implemented using one step per span (per-span) equalizer, where only the $\mathbf{F}_k$ kernel is used to calculate $a_{X,k}$ at each span (see [11]). After each step of the recursive VSFE, the signal was transformed back into time-domain to include the modification to the output signal presented in [32, Eq. (5)] to account for the energy divergence effect suffered by use of a truncated Volterra series for the VSFE. Modifying the Volterra output is crucial to account for the previously mentioned energy divergence effect (see Sec. II). For the single-step VSFE and VAO methods the energy divergence modification was evaluated, however, it did no present any significant performance enhancement.

The computational complexity of a VSFE, scales between $O(N)$ and $O(N^2)$ per sample, depending on the specific
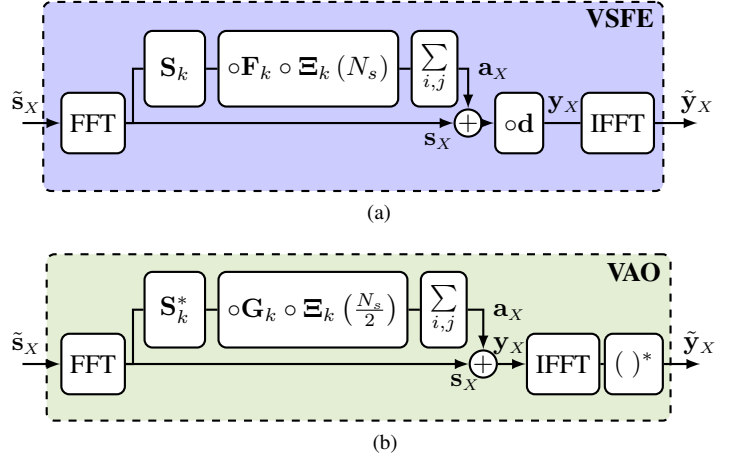
[4]The Hadamard multiplication, denoted by the $\circ$ symbol in Fig. 2, is a simple element-wise matrix multiplication.

Fig. 2: Schematic diagram of the two Volterra equalizers compared in this work: a (a) conventional VSFE, and (b) the equalizer in the proposed VAO scheme.

implementation adopted [10], [11]. For continuous data transmission, this effectively limits the window size over which the VSFE can be sequentially applied. Conversely, a minimum window size must be guaranteed to allow for an absorbing window for the fiber channel memory effects to be factored out. This is done typically through an overlap and save approach, where a certain number of symbols are discarded from each side of the window, to account for the cyclic effects induced by the discrete frequency-domain operation. In this work, VSFE was applied sequentially over subsequent sequences of samples, whose size was varied (between 128 and 1024 symbol periods) depending on the transmission distance. For each processed sequence, a certain amount of symbols were then discarded from each side of the window, to account for the above mentioned cyclic effects. Despite the increased complexity per sample (see e.g. [15]), the discarding operation considerably benefits the equalizer performance. As discussed in the following sections, the performance metrics analysed in this work (NLI suppression and SNR) are closely related to the reduction of NLI distortion. This is affected by the window size used for the VSFE. In order to guarantee the best performance, the VSFE window size was set to 4 times (up to a maximum of 1024 symbols) the estimated channel memory. A rough estimate of the channel memory can be obtained using $M \approx |\beta_2| \cdot R_s \cdot B \cdot L$ where R_s is the channel symbol rate, B is the entire transmitted bandwidth, and L is the total transmitted distance. Each sub-sequence was sequentially processed by the VSFE, and then demodulated. The boundary symbols of the demodulated subsequence were then increasingly discarded until the *performance metric* of choice converged within a desired tolerance. Multiple sub-sequences were then processed in order to accumulate enough received symbols and to guarantee the accuracy of *performance metric* calculation within 0.1 dB.

IV. SYSTEM UNDER INVESTIGATION

The performance of the proposed NLC method was evaluated through numerical simulations. All the investigated
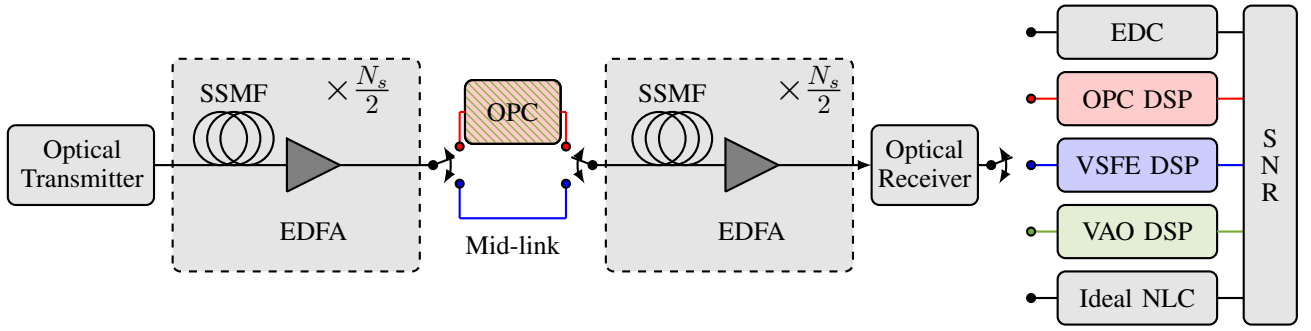
Fig. 3: Systems compared in the numerical investigation in this work.

system scenarios are shown in Fig. 3, with different colors representing different receiver schemes. An ideal optical transmitter was used to generate 5 polarization-multiplexed 16-ary quadrature amplitude modulated (PM-16QAM) channels at a symbol rate of 32 GBaud using a sequence of 2^{16} symbols. The signal pulses were shaped using a root raised cosine filter with 1% roll-off factor. Transmission over N_s 100 km-long optical fiber spans was simulated by numerically solving the Manakov equation using the split-step Fourier method. The numerical method is based on a logarithmic step size distribution as in [33], which is adjusted to achieve the required accuracy. The fiber spans were based on SSMF and EDFA amplification after each span, with all parameters shown in Table I . At the mid-link point, the signal either continues to the following half of the link (blue path), or undergoes phase conjugation (red path). OPC was performed ideally by taking the complex conjugate of the optical signal in the time domain.

After transmission, an optical coherent receiver was assumed to ideally capture the entire 5-channel optical bandwidth ($\approx$ 165 GHz) with a subsequent sampling at $\approx$ 1.2 times the Nyquist rate of the received signal ($\approx$ 192 GSa/Sym). Previously, different receiver structures were proposed and experimentally demonstrated to perform detection over such a wide bandwidth. For example, in [34] 5 *spectrally-sliced* coherent receivers were used to detect and backpropagate a bandwidth of 176 GHz. In [35] multiple channels were jointly backpropagated using a single coherent receiver of 70 GHz bandwidths. Finally, a *digitally stitched* coherent receiver, with a total bandwidth of 246 GHz was demonstrated in [36].

The sampled signal over each polarization had 5 possible DSP chain options depending on the selected system (with or without OPC), which are: (i) ideal electronic dispersion compensation (EDC) performed in the frequency-domain when no OPC module was used; (ii) OPC-only DSP; (iii) single-step/recursive VSFE; (iv) the VAO equalizer; (v) ideal DBP performed using a full-complexity (100 steps/span), *full-field* DBP implementation. The DSP chain for each of the 3 NLC schemes compared in this work, i.e. OPC-only DSP, conventional VSFE, and VAO, is shown in Figs. 4a–4c. For each DSP scheme shown in Figs. 4a–4c, the input vector (shown only for the X polarization for the sake of simplicity) is sampled at 6 samples/symbol. The OPC DSP module (Fig. 4a) consists of a downsampling stage at 2 samples/symbol to select the central channel bandwidth, followed by time-domain conjugation,

matched filter (MF) and downsampling at 1 sample/symbol. As for the VSFE DSP, the signal is first processed by a conventional VSFE (see Fig. 2a) at 6 samples/symbol corresponding to an NLC bandwidth of 192 GHz, then downsampled at 2 samples/symbol, and finally passed through a M followed by downsampling to 1 sample/symbol. Similar operation is performed in the VAO DSP chain where, however, the signal must be conjugated after the VSFE block. After each DSP scheme, the SNR was estimated using a fully data-aided method.

V. NONLINEARITY SUPPRESSION

In order to assess the NLC effectiveness of the proposed scheme, a nonlinear interference (NLI) suppression factor was characterised through numerical simulations. This NLC suppression factor was defined as

$$\zeta \triangleq \frac{\sigma_{\text{EDC}}^2}{\sigma_{\text{NLC}}^2}, \tag{12}$$

where σ_{EDC}^2 and σ_{NLC}^2 represent the variance of the NLI at the MF output when EDC and NLC are applied, respectively. The ζ factor indicates the reduction of the NLI power after NLC, and is key to understand, in the absence of any other noise sources, what is the effectiveness of that particular scheme. For an ideal DBP scheme (full compensation of any deterministic nonlinear effect), ζ tends to infinity as the NLI power is forced to 0. Conversely, this value tends to 1 if the NLI is left uncompensated (EDC-only schemes), or it is less than 1 when additional distortions are added, due for instance to NLC schemes operating with a very limited accuracy. In order to characterize the proposed VAO scheme, the ζ parameter

TABLE I: Simulation parameters

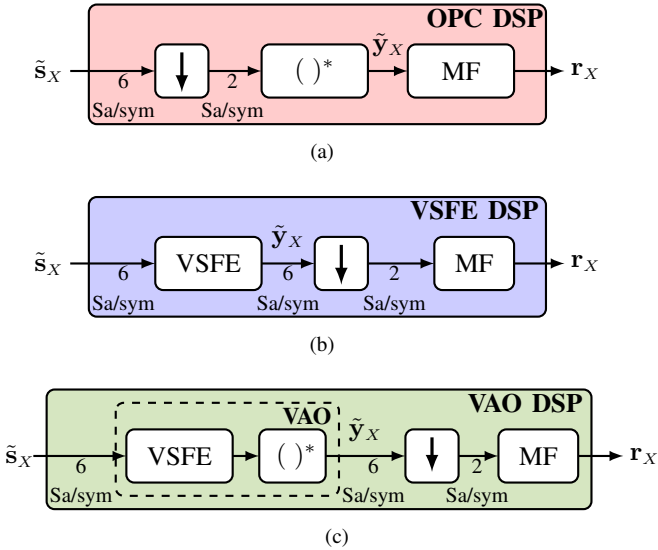| Parameter	Value
Symbol rate	32 GBd
Channel spacing	32.5 GHz
RRC roll-off	1 %
Span length (L_s)	100 km
Attenuation coefficient	0.20 dB
Dispersion parameter (D)	17 ps $\cdot$ nm^{-1} $\cdot$ km^{-1}
Nonlinear coefficient (γ)	1.2 W^{-1}km^{-1}
EDFA noise figure	5 dB

Fig. 4: Schematic diagram of the DSP chains for: (a) OPC-only system, (b) conventional VSFE and (c) VAO. MF denotes the matched filtering stage and $\mathbf{r}_X$ the output symbols at 1 sample/symbol.

Fig. 5: Nonlinearity suppression factor (ζ) as a function of transmitted power for the NLC schemes studied in this work obtained in a transmission link free of ASE noise

.

was studied as a function of the 2 optical system parameters: transmitted power, and transmission distance for a 5-channel transmission. ζ was calculated numerically through Monte-Carlo estimation of the received SNR, in the absence of ASE noise in the simulated system. An appropriate number of processed subsequences was accumulated to guarantee an SNR accuracy within 0.05 dB.

The nonlinearity suppression factor (ζ) was calculated as a function of power for the transmission of a 5×32 GBaud signal over a 1000 km link for 3 different NLC schemes illustrated in Fig. 4: mid-link OPC (red curve); VSFE (blue curves) and the proposed VAO (green curve). These results are plotted in Fig. 5. The VSFE is here implemented in two different variants: single-step per link (blue curve with circle markers), and recursive VSFE (squared markers). Both variants of VSFE and VAO were operated (repeatedly) over 512 symbol-long sequences each oversampled at 6 samples/symbol, resulting in 3072 samples per processed sequence. The results show that OPC (red curve) has a constant NL suppression of approximately 1.8 dB across the entire range of powers shown. Such a limited NLC suppression is the result of the residual NLI when OPC is used in an EDFA link (see Sec. II), and is explained by the third-order Volterra kernel illustrated in Fig. 1b. VSFE was characterized in terms of ζ in the two different implementations described in Sec. III: single-step and recursive VSFE. Single-step VSFE shows a better performance at low powers, with a maximum ζ=4 dB at -4 dBm, and monotonically decreasing ζ as the transmitted power increases. This well-known performance behaviour of the Volterra equalizer is due to the truncation of the Volterra series to the third-order term. Such a truncation becomes less accurate as the transmitted power is increased, and leads to a degradation in the equalization performance (ζ=0.5 dB at 8 dBm/channel). This degradation is partly mitigated when
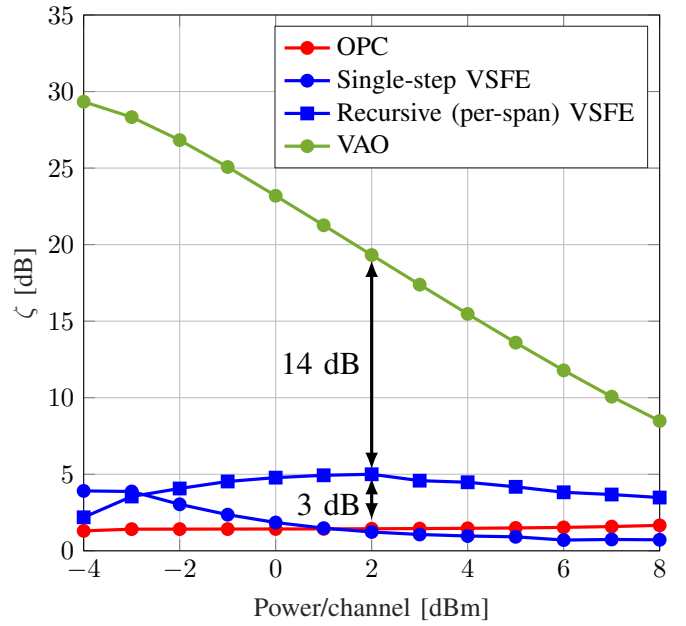
a recursive VSFE scheme is used. Although the two VSFE schemes (blue curves in Fig. 5) show similar NLI suppression ($\approx$ 3.5 dB) for around -3 dBm/channel, the recursive VSFE scheme performs better beyond -3 dBm/channel and achieves up to 5 dB NLI suppression at 2 dBm/channel. However, this improvement comes at the expense of an N_s times higher computational complexity compared to a single-step VSFE scheme, for a fixed processed sequence length [12]. Fig. 5 shows that the VAO scheme (green curve) has a superior performance compared to the all the other NLC schemes analysed, across the entire range of powers. Like in single-step VSFE, the VAO ζ factor monotonically decreases with power, and reaches a maximum value (within the range of powers shown in Fig. 5) of $\approx$ 30 dB at -4 dBm/channel. At this transmitted power, VAO exceeded the NLI suppression of all of the other NLC schemes by over 25 dB. Although the performance of VAO can be observed to degrade quickly as a function of power, even in the region of powers of interest for optimum transmission performance (0 to 4 dBm) the VAO NLI suppression is still significantly higher than other compared schemes. In particular, at 2 dBm/channel VAO achieves 17 dB higher ζ compared to using only OPC, and 14 dB higher suppression than the optimum ζ achieved by a VSFE-only scheme (ζ=5 dB at 2 dBm/channel).

The remarkable NLC performance attained by the VAO for the system studied is attributed to two reasons: i) for a fixed processed sequence length, VAO benefits from a reduced channel memory compared to VSFE; ii) the adoption of OPC leads to a more accurate reconstruction of the transmitted field using a third-order Volterra series term due to the reduced NLI at the receiver. The reduction of the channel memory is due to the compensating effect of the phase conjugation in
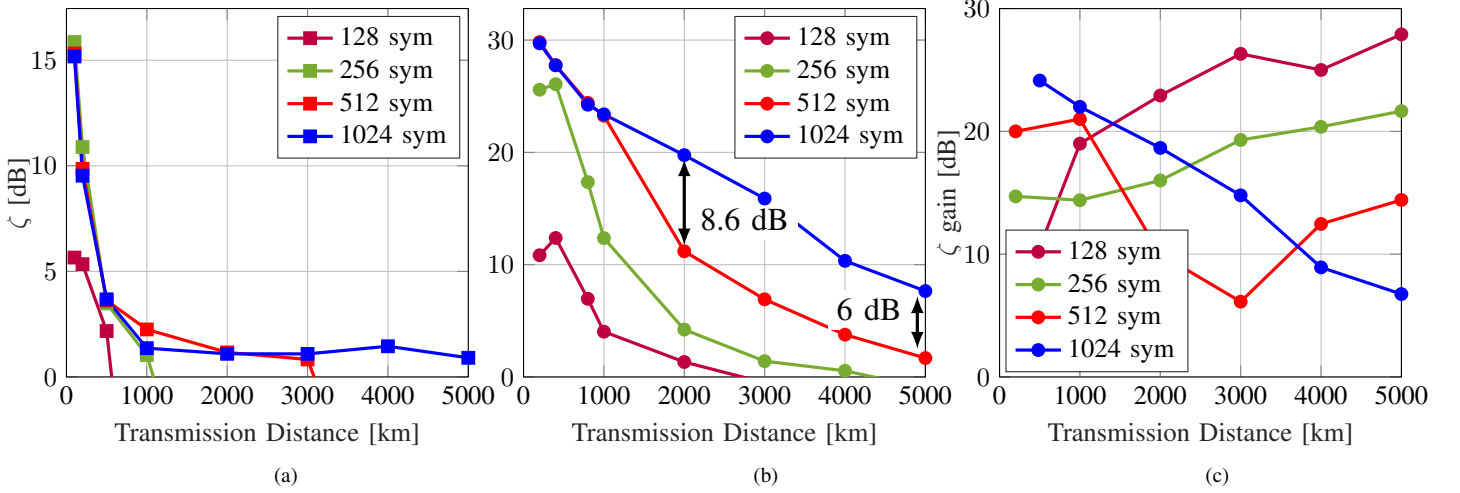
Fig. 6: Nonlinearity suppression factor (ζ) as a function of transmission distance and different equalizer window sizes for (a) VSFE, (b) VAO, and (c) VAO NLI suppression gain compared to VSFE.

the second part of the link as the OPC completely removes the fiber chromatic dispersion accumulated in the first part of the link, thus also reducing the time window over which the nonlinear intersymbol takes place. This *memory reduction* can also be inferred by the phased-array in (6) which is only evaluated over half of the link.

To assess the improved robustness of VAO against an increasing channel memory, the NLI suppression as a function of the transmitted distance is shown for the system under investigation at a transmitted power of 0 dBm/channel in Fig. 6, for single-step VSFE (Fig. 6a), and VAO (Fig. 6b). The effect of the channel memory is highlighted by showing the performance for different window sizes, spanning from 128 to 1024 symbols. As the channel memory is expected to scale linearly with increasing transmission distance, each NLC scheme shows a decreasing NLI suppression factor for a fixed window size. However, performance degradation as a function of distance is also caused by an increasing dominance of higher-order Volterra series terms. The two effects can be distinguished by observing the behaviour of ζ as the window size is increased. The results in Fig. 6a show that, up to 3000 km transmission distance, ζ saturates using time-windows between 512 and 1024 symbols. For each window size, a *threshold distance* of operation, beyond which VSFE performance breaks down (negative ζ), can be identified. This was caused by the channel memory going beyond the maximum number of discarded symbols within a given window size. Increasing from 512 to 1024 symbols the window size did not lead to any substantial increase in ζ. Thus, it can be concluded that for window sizes beyond 1024 symbols and up to 5000 km transmission distance, the VSFE performance is dominated by its compensation accuracy (NLI compensation up to the first-order).

The VAO NLI suppression as a function of distance is shown in Fig. 6b. It can be observed that for VAO the detrimental effect of the channel memory operating over finite window sizes is less pronounced. This is also confirmed

by the smoother decay of ζ as a function of transmission distance compared to VSFE, with NLC capabilities (positive ζ) observed within 3000 km for all window sizes shown. For transmission distances up to 1000 km, the value of ζ saturates already at 512 symbols. Beyond a distance of 1000 km, the increase of the window size from 512 symbols to 1024 symbols results in a substantial increase in ζ (up to 8.6 dB at 2000 km). The performance gap between using a 512-symbol window and a 1024-symbol window then gradually decreases to 6 dB at 5000 km. This can be attributed to the stronger effect of higher-order Volterra series terms as the transmission distance increases. Fig. 6c shows the NLI suppression gain achieved by VAO over VSFE as a function of distance for different window lengths. The power per channel is again fixed at 0 dBm/channel. It can be seen that the improvement introduced by the VAO scheme in terms of ζ is remarkable for all distances and window sizes used. For short distances, (up to 1000 km) the highest gains can be observed for 512- and 1024-symbol window length, up to 21 dB and 24 dB, respectively. This highlights the superior NLC efficacy of a Volterra-based equalizer when an in-line OPC device is used. For longer distances (beyond 2000 km), the highest VAO gain is achieved for the shortest window lengths (128- and 256-symbol long), due to the reduction of the channel memory provided by OPC. In contrast with VSFE, this enables the VAO equalizer to obtain a certain (albeit limited) amount of NLI suppression (≤ 4.5 dB) for distances within $2000-4000$ km, while still processing relatively short sequences (see Fig. 6b). For instance, a 19.5 dB gain is attained using VAO with a 256-symbol long sequence at 3000 km, which leads to $\zeta \approx 1.5$ dB (see Fig. 6a and 6b).

VI. SNR PERFORMANCE

In the previous section, the performance of the proposed VAO scheme was analysed in terms of NLI suppression factor. Despite this being a more direct measure of the NLC effectiveness, it does not immediately translate into more
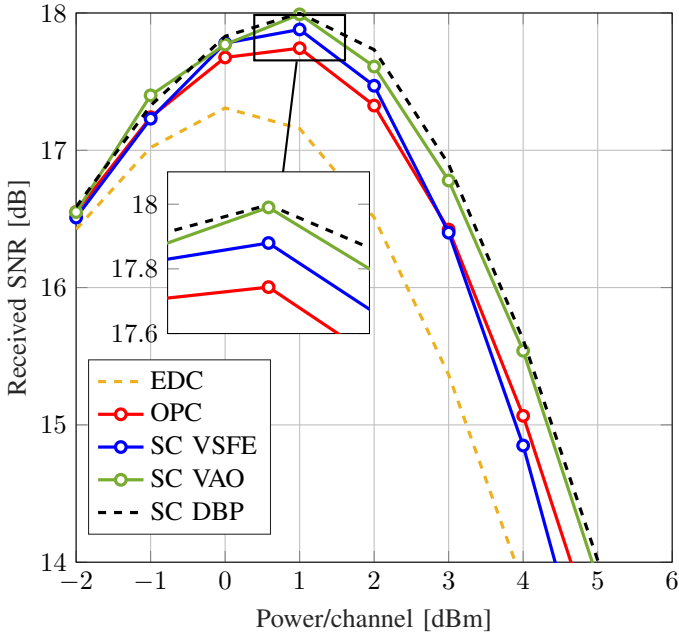
Fig. 7: Received SNR for different NLC schemes operating over a single channel (SC) as a function of signal launch power.

Fig. 8: Received SNR for different NLC schemes as a function of signal launch power.

insightful performance metrics such as signal-to-noise ratio (SNR). In this section, a study of the SNR performance in the presence of ASE noise is described for the different NLC schemes analyzed in this work.

Numerical simulations were carried out to assess the performance of different NLC methods in a realistic transmission scenario i.e. in the presence of ASE noise) and to compare their performance with the VAO method. Similar to Sec. V, the performance was first evaluated for a fixed transmission distance of 1000 km as a function of the signal launch power. The results using NLC over a single channel are shown in Fig. 7, and the results operating over the full signal bandwidth are shown in Fig. 8. Subsequently, the system performance as a function of the transmission distance at optimum power is presented in Fig. 9.

For current systems, it is of practical importance to employ a NLC method that can operate with transceivers with limited bandwidth. Hence, the performance of single channel (SC) VAO was studied and compared to SC VSFE implemented using a single step per link, and SC DBP, with the results plotted in Fig. 7. Among all the studied methods, electronic dispersion compensation (EDC) shows the worst performance of all. As EDC does not enact any compensation of nonlinear effects, this scheme serves as a baseline to evaluate the effectiveness of the NLC methods under study. EDC exhibits a maximum SNR of 17.3 dB at 0 dBm per channel. At this distance, OPC presents a gain over EDC limited to 0.4 dB, with a maximum SNR of 17.7 dB, indicating that only a small portion of the nonlinearities are compensated. As discussed in Sec. II, the spectral region where OPC offers complete cancellation of nonlinearites (see Fig. 1b) was not sufficiently large to account for the effects of all 5 transmitted channels.
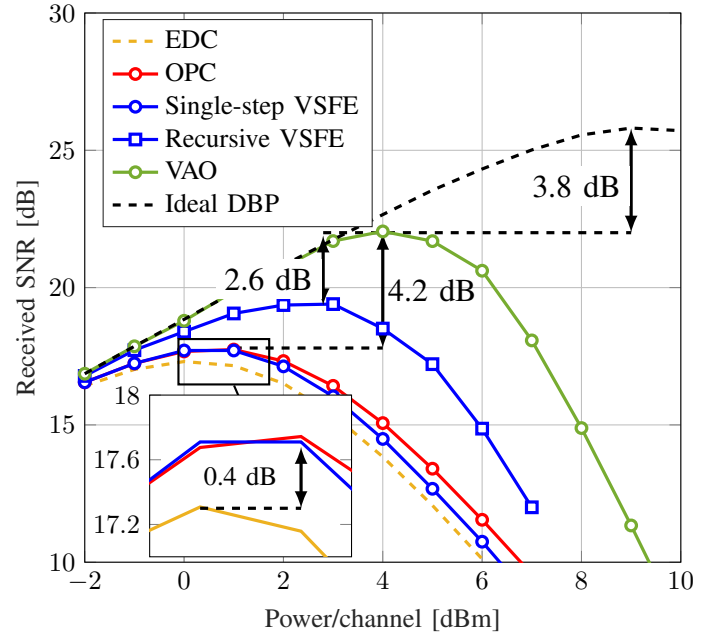
In the highly-nonlinear regime, e.g. using a power of 4 dBm, an increase of 1.2 dB in the SNR over EDC is observed with the use of OPC. This value is in agreement with the nonlinear suppression factor observed for OPC in the previous section. The use of SC VSFE presents an improvement over OPC of 0.1 dB, with a maximum SNR of 17.8 dB found at 1 dBm per channel. Additionally, both SC DBP and SC VAO exhibit the same performance at optimum power, with a SNR of 18 dB. As the compensating bandwidth is limited to a single channel, any receiver-based NLC method is limited to compensate for nonlinearities generated within that bandwidth; therefore, the efficacy of these methods is determined by the residual nonlinearity arising from outside the compensation bandwidth. For this reason, the studied single channel NLC methods show similar performance.

Significant performance enhancements are found by performing NLC over a bandwidth larger than the single channel one. As discussed in Sec. IV, NLC schemes such as DBP have been already successfully demonstrated in the previous literature over a bandwidth of up to 175 GHz, providing at least a proof of concept of wide-bandwidth NLC. Figure 8 shows the performance of the system using the different full-field NLC techniques as a function of launch power per channel after transmission of 1000 km. In this case, the performance of EDC and OPC was the same observed for the single-channel results. Similar performance, compared to OPC, is observed using single-step VSFE at the maximum SNR. In this case, one-step VSFE was implemented using a window size of 512 symbols. A maximum SNR of 17.7 dB is found using this method at a signal power of 1 dBm.

As the power is increased, the efficiency of VSFE is decreased as a result of the truncation of the Volterra series to third-order term, as described in the previous section. For
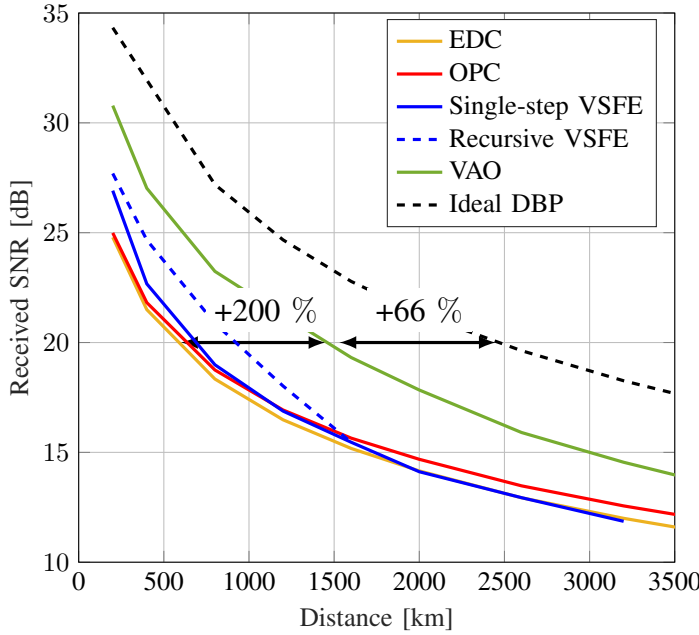
Fig. 9: Received SNR for different NLC schemes as a function of transmission distance.

the recursive VSFE a window size of 256 symbols was used. This method presents an improved performance relative to the single-step implementation, with a maximum SNR of 19.4 dB at 2 dBm/channel signal power. Using VAO results in a significant improvement relative to VSFE and OPC, with a gain of 4.6 dB compared to EDC, 4.2 dB relative to either conventional single-step VSFE or OPC, and 2.6 dB compared to recursive VSFE. The VAO was implemented using a window size of 512 symbols in the equalizer, to allow a direct comparison with the conventional VSFE methods. Whilst the VSFE truncation is one of the factors that limited the performance of conventional VSFE implementations, the reduction of nonlinear effects due to the mid-link OPC leads to a better reconstruction of the transmitted signal using the VAO scheme. Finally, for comparison, the performance of ideal DBP is shown, in a virtually reversed transmission link. This represents the maximum performance that can be obtained using any zero-forcing, receiver-based, NLC scheme [37]. Some methods have been proposed to account for signal-noise interactions, for example in [38], however with little benefit over DBP in practical scenarios. Although the VAO scheme presents improved performance relative the aforementioned NLC schemes, a 3.8 dB gap is observed compared to an ideal DBP method. The performance penalty, compared to DBP, from VAO and both VSFE methods was not observed in the single channel scenario due to the use of a smaller bandwidth. This helped to prevent energy divergence problems, and lead to a smaller penalty when VSFE is used to backpropagate the received signal.

Figure 9 shows the performance of the system as a function of transmission distance. The same trend from Fig. 8 in terms of nonlinear effects was observed, with EDC exhibiting the worst performance among all the studied methods. Interest-

ingly, OPC presents a reduced gain of 0.2 dB relative to EDC after 200 km, however greater nonlinearity compensation is obtained for longer distances and 0.6 dB gain is observed after 3200 km. The opposite behaviour is found for single-step VSFE, where a superior performance is found at short distances. As in the previous case, both standard VSFE were implemented using window sizes of 512 and 256 symbols for the single-step and recursive algorithms, respectively. Using VSFE, the gain over EDC is slowly reduced as the transmission distance increased, up to the point where no improvement is observed. This phenomenon can be understood from memory effects introduced by the fiber chromatic dispersion. When the temporal spreading of the propagated signal exceeds the VSFE window size the performance of both implementations, single-step and recursive, is dramatically reduced. For the studied system no gain from NLC is observed for transmission distances longer than 1600 km using both standard VSFE implementations.

The VAO scheme is shown to outperform both standard implementations of VSFE, and OPC at all evaluated transmission distances. For a received SNR of 20 dB, VOA is shown to attain a 200% increase in transmission distance compared to EDC, going from 500 to 1500 km. The VAO method however, did not offer full compensation of the NLI experienced during transmission and reduced performance is found compared to ideal DBP. The use of ideal DBP achieves an extra 66% increase of transmission distance for a received SNR of 20 dB compared to VAO. Additionally, the NLC gain from the VAO method presents an accelerated reduction as a function of the transmission distance, compared to ideal DBP. We attribute this behaviour to the nonlinear equalizer in the VAO scheme, which has qualitatively the same limitations as a standard VSFE in terms of channel memory and truncation of the Volterra series.

VII. COMPLEXITY

In the previous sections, the VAO scheme was shown to significantly outperform OPC and VSFE in practical transmission scenarios, such as EDFA-amplified links, for a NLC bandwidth of 165 GHz (5×32 GBaud). Although VAO offers an improved performance compared to VSFE, a substantial performance penalty (3.8 dB at 1000 km transmission distance) was still observed with respect to ideal nonlinearity compensation (see Fig. 8). Such an ideal DBP performance could be theoretically achieved through a *full-complexity* implementation of the DBP algorithm over the entire transmitted bandwidth. The question that naturally arises is then, which, between VAO and DBP, represents a better trade-off between performance and complexity.

The computational complexity of the VAO scheme proposed in this work is determined by the complexity of its modified Volterra equalizer. Comprehensive studies on the complexity of VSFE, and simplified variants thereof, can be found in literature [10], [12], [14], [15], [39], [40]. These works showed that, despite the complexity (per processed symbol) of a standard VSFE implementations scaling asymptotically as $\mathcal{O}(N^2)$, simplified implementations are possible with limited

performance penalties. In particular, in [12] and [40] the complexity of VSFE was reduced to $\mathcal{O}(N)$ and $\mathcal{O}(N \log N)$ per symbol, respectively.

The computational complexity of the DBP algorithm, for a large number of processed samples N, scales as $\mathcal{O}(N_{\text{steps}} \log N)$ where N_{steps} is the number of DBP steps performed. Alike the VSFE/VAO case discussed in Sec. V, the minimum value of N for DBP is dictated by the targeted NLC performance and the channel memory (see e.g. [41]). Due the logarithmic scaling of simplified VSFE, matching an increasing channel memory by increasing N only results in a minor complexity increase, which is also the case for DBP. The main issue in extending DBP to larger NLC bandwidths is the minimum N_{steps} required to preserve a fixed NLC suppression. The NLC performance for DBP is determined by the relationship between the accuracy of the SSFM and the signal bandwidth at a fixed power spectral density, which depends on the specific SSFM method used and, overall, it is not well known. For instance, for a logarithmic step-size SSFM [33], N_{steps} scales quadratically with the NLC bandwidth (at fixed power spectral density). Although for other SSFM variants (see, e.g., [41]) this scaling could be more benign, the advantage of (non recursive) VSFE methods, and thus VAO, is that their complexity scales on the NLC bandwidth only through N, as a result of an increasing the channel memory.

An accurate study on the performance/complexity trade-off between VAO and DBP requires a detailed description of the implementation of both algorithms, and is beyond the scope of this paper.

VIII. Conclusion

In this work, a novel nonlinearity compensation scheme combining mid-link optical phase conjugation (OPC) and frequency-domain Volterra equalization, was proposed. The new scheme, referred to as Volterra-assisted OPC (VAO) was designed to overcome the limitations of both OPC and Volterra equalization in practical transmission scenarios, whilst preserving the complexity of a conventional VSFE. VAO was demonstrated to effectively recover the NLC capabilities of OPC over a large optical bandwidth, in EDFA-amplified links. The addition of mid-link OPC was shown to significantly enhance the NLC performance of a conventional VSFE. Up to 10 dB higher NLC suppression was demonstrated for VAO and 4.2 dB enhancement in the received SNR in a 1000 km link compared to either OPC or VSFE when VAO was applied over the entire transmitted optical field. As VAO dramatically outperforms Volterra-only equalisation for large NLC bandwidths, it offers in such a scenario a attractive trade-off between NLC effectiveness and computational complexity.

Appendix
Derivation of the third-order Volterra series kernel in the presence of mid-link OPC

We begin the derivation of the third-order Volterra kernel from the Manakov equation in the frequency domain for a dual polarization signal. Only the X polarization solution will be derived here, as the solution for the Y polarization is simply obtained interchanging the subscripts X and Y. For the signal on the X polarization, the Manakov equation in frequency domain reads:

$$\frac{\partial A_X}{\partial z} = j\frac{\beta_2}{2}\omega^2 A_X - \frac{\alpha(z)}{2}A_X + j\frac{8}{9}\gamma\mathcal{F}\{(\tilde{A}_X\tilde{A}_X^* + \tilde{A}_Y\tilde{A}_Y^*)\tilde{A}_X\}, \tag{13}$$

where $A_X(\omega, z)$ is the Fourier transform of the optical field over polarization X at distance z, whereas $\tilde{A}_X(t, z) \triangleq \mathcal{F}^{-1}\{A_X\}$ is its time-domain counterpart. The equivalence between the Volterra series and the regular perturbation expansion terms shown in [27] allows us to use the latter to derive the third-order Volterra kernel. Following the regular perturbation approach for the solution of (13) [27], [42], the zeroth-order perturbation solution, corresponding to the first term of the Volterra series, is given by

$$A_{1X}(\omega, N_s L_s) = A_X(\omega, 0)e^{\left(-\frac{\alpha(z)}{2} + i\frac{\beta_2}{2}\omega^2\right)N_s L_s}, \tag{14}$$

where $A_X(\omega, 0)$ is the Fourier transform of the transmitted optical field, N_s is the number of spans and L_s is the span length. The zeroth order solution is recursively used to derive the first-order perturbation term of the solution of Eq. (13), which corresponds to its third-order Volterra series expansion, and it is given by

$$A_{3X}(\omega, N_s L_s) = j\frac{8}{9}\gamma e^{\left(i\frac{\beta_2}{2}\omega^2 - \frac{\alpha(z)}{2}\right)N_s L_s}$$
$$\times \int_0^{N_s L_s} e^{\int_0^{z'}\left(-i\frac{\beta_2}{2}\omega^2 + \frac{\alpha}{2}\right)dz''} D(\omega, z')dz', \tag{15}$$

where the term $D(\omega, z)$ comes from substituting Eq. (14) into the nonlinear term in the left-hand side of Eq. (13), and it is given by

$$D(\omega, z) = -\iint \left[A_{1X}^*(\omega_1, z)A_{1X}(\omega_2, z)A_{1X}(\omega - \omega_2 + \omega_1, z) + A_{1Y}^*(\omega_1, z)A_{1Y}(\omega_2, z)A_{1X}(\omega - \omega_2 + \omega_1, z)\right]d\omega_2 d\omega_1, \tag{16}$$

Substituting Eq. (16) into Eq. (15) we obtain

$$A_{3X}(\omega, N_s L_s) = j\frac{8}{9}\gamma e^{\left(i\frac{\beta_2}{2}\omega^2 - \frac{\alpha(z)}{2}\right)N_s L_s}$$
$$\times \iint \left[\mathcal{S}_{XXX}(\omega, \omega_1, \omega_2) + \mathcal{S}_{YYX}(\omega, \omega_1, \omega_2)\right]$$
$$\times F(\omega, \omega_1, \omega_2)\,\Xi(N_s, \omega, \omega_1, \omega_2)\,d\omega_2 d\omega_1. \tag{17}$$

where the signal kernels $\mathcal{S}_{XXX}(\omega, \omega_1, \omega_2)$ and $\mathcal{S}_{YYX}(\omega, \omega_1, \omega_2)$, the phased-array $\Xi(N_s, \omega, \omega_1, \omega_2)$, and $\Delta\Omega$ were previously defined in Sec. II. Without in-line OPC, the four-wave mixing efficiency $F(\omega, \omega_1, \omega_2)$ and the *phased array* factor $\Xi(N_s, \omega, \omega_1, \omega_2)$ are defined in (4) and (3), respectively.

To study the effect of an OPC module placed at the mid-link point, i.e. after $\frac{N_s}{2}$ fiber spans, the solution of (15) is conveniently split into 2 contributions: one before and one after the OPC module. Rewriting Eq. (15) we have:

$$A_{3X}(\omega, N_s L_s) = i\frac{8}{9}\gamma e^{\left(i\frac{\beta_2}{2}\omega^2\right)N_s L_s} P(N_s L_s)$$
$$\times \left[\int_0^{\frac{N_s}{2}L_s} P(-z')e^{\int_0^{z'} -i\frac{\beta_2}{2}\omega^2 dz''} D(z')dz' \right.$$
$$\left. + \int_{\frac{N_s}{2}L_s}^{N_s L_s} P(-z')e^{\int_0^{z'} -i\frac{\beta_2}{2}\omega^2 dz''} D(z')dz' \right]. \tag{18}$$

where $P(z)$ is the power profile at distance z given by

$$P(z) = e^{-\int_0^z \alpha(z')z'}. \tag{19}$$

Both integrals in Eq. (18) represent propagation up to specific part of the transmission link. The first one corresponds to propagation from the transmitter up the mid-link point, and the second one corresponds to propagation from the mid-link point to the end the last fiber span. Using the first-order solution for the first half of the link Eq. (18) can be rewritten as follows:

$$A_{3X}(\omega, N_s L_s) = A_{3X}\left(\omega, \frac{N_s L_s}{2}\right) P\left(\frac{N_s L_s}{2}\right) e^{i\frac{\beta_2}{2}\omega^2 \frac{N_s L_s}{2}}$$
$$+ i\frac{8}{9}\gamma P(N_s L_s)e^{i\frac{\beta_2}{2}\omega^2 N_s L_s}$$
$$\times \int_{\frac{N_s}{2}L_s}^{N_s L_s} e^{\int_0^{z'} \left(-i\frac{\beta_2}{2}\omega^2 + \frac{\alpha}{2}\right)dz''} D(z')dz'. \tag{20}$$

When an OPC module is used, both zeroth and first-order perturbation terms undergo conjugation in time-domain at $z = \frac{N_s L_s}{2}$, leading to:

$$A_{3X}(\omega, N_s L_s) = \hat{A}_{3X}\left(\omega, \frac{N_s L_s}{2}\right) P\left(\frac{N_s L_s}{2}\right) e^{i\frac{\beta_2}{2}\omega^2 \frac{N_s L_s}{2}}$$
$$+ i\frac{8}{9}\gamma P(N_s L_s)e^{i\frac{\beta_2}{2}\omega^2 N_s L_s}$$
$$\times \int_{\frac{N_s}{2}L_s}^{N_s L_s} e^{\int_0^{z'} \left(-i\frac{\beta_2}{2}\omega^2 + \frac{\alpha}{2}\right)dz''} D(z')dz'. \tag{21}$$

where the conjugation of the first-order perturbation term is given by:

$$\hat{A}_{3X}\left(\omega, \frac{N_s L_s}{2}\right) \triangleq \mathcal{F}\left\{ \tilde{A}_{3X}^*\left(t, \frac{N_s L_s}{2}\right) \right\}$$
$$= A_{3X}^*\left(-\omega, \frac{N_s L_s}{2}\right). \tag{22}$$

The term corresponding to the first half of the transmission link then becomes:

$$\hat{A}_{3X}\left(\omega, \frac{N_s L_s}{2}\right) P\left(\frac{N_s L_s}{2}\right) e^{i\frac{\beta_2}{2}\omega^2 \frac{N_s L_s}{2}}$$
$$= -i\gamma\frac{8}{9}P(N_s L_s)\iint K(\omega,\omega_1,\omega_2)\Lambda^*\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right)d\omega_2 d\omega_1, \tag{23}$$

where

$$\Lambda\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right) \triangleq \sum_{n=1}^{\frac{N_s}{2}} \left[\frac{P(nL_s)e^{i\beta_2\Delta\Omega nL_s}}{\alpha(nL_s) + i\beta_2\Delta\Omega} \right.$$
$$\left. - \frac{P((n-1)L_s)e^{i\beta_2\Delta\Omega(n-1)L_s}}{\alpha((n-1)L_s) + i\beta_2\Delta\Omega} \right], \tag{24}$$

and

$$K(\omega,\omega_1,\omega_2) = \mathcal{S}_{XXX}^*(-\omega,-\omega_1,-\omega_2) + \mathcal{S}_{YYX}^*(-\omega,-\omega_1,-\omega_2). \tag{25}$$

Beyond the conjugation point $z \geq \frac{N_s L_s}{2}$, the zeroth order perturbative solution is

$$A_{1X}(\omega, z) = A_X^*(-\omega, 0)P(z)e^{i\frac{\beta_2}{2}\omega^2(z - N_s L_s)}. \tag{26}$$

Substituting (26) into $D(z)$ we obtain:

$$D(z) = -P^3(z)e^{i\frac{\beta_2}{2}\omega^2(z - N_s L_s)}$$
$$\iint A_X^*(-\omega_2, 0)A_X(-\omega_1, 0)A_X^*(\omega_2 - \omega_1 - \omega, 0)$$
$$e^{i\beta_2\Delta\Omega(z - N_s L_s)}d\omega_2 d\omega_1. \tag{27}$$

Hence, the second term in Eq. (20) becomes

$$P(N_s L_s)e^{i\frac{\beta_2}{2}\omega^2 N_s L_s} \int_{\frac{N_s}{2}L_s}^{N_s L_s} P^{-1}(z')e^{-i\frac{\beta_2}{2}\omega^2 z'} i\frac{8}{9}\gamma D(z')dz'$$
$$= i\gamma\frac{8}{9}P(N_s L_s)\iint K(\omega,\omega_1,\omega_2)\psi\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right)d\omega_2 d\omega_1 \tag{28}$$

where

$$\Psi\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right) \triangleq e^{-i\beta_2\Delta\Omega N_s L_s}$$
$$\times \sum_{n=\frac{N_s}{2}+1}^{N_s} \left[\frac{P(nL_s)e^{i\beta_2\Delta\Omega nL_s}}{\alpha(nL_s) + i\beta_2\Delta\Omega} - \frac{P((n-1)L_s)e^{i\beta_2\Delta\Omega(n-1)L_s}}{\alpha((n-1)L_s) + i\beta_2\Delta\Omega} \right]. \tag{29}$$

Finally, we obtain:

$$A_{3X}(\omega, N_s L_s) = i\gamma\frac{8}{9}P(N_s L_s)\iint K(\omega,\omega_1,\omega_2)$$
$$\times \left[\Psi\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right) - \Lambda^*\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right) \right]d\omega_2 d\omega_1 \tag{30}$$

where the nonlinear kernel in the presence of OPC is given by:

$$\Gamma(N_s,\omega,\omega_1,\omega_2) \triangleq \Psi\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right) - \Psi^*\left(\frac{N_s}{2},\omega,\omega_1,\omega_2\right). \tag{31}$$

In order to suppress the first-order solution for all ω one condition is having $\Gamma(N_s,\omega,\omega_1,\omega_2) = 0 \quad \forall \ \omega,\omega_1,\omega_2$. This is verified if the following conditions are jointly verified

$$P((N_s - n)L_s^-) = P((n-1)L_s^+), \quad (32)$$

$$P((N_s - n - 1)L_s^+) = P(nL_s^-), \quad (33)$$

$$\alpha((N_s - n)L_s^-) = -\alpha((n-1)L_s^+), \quad (34)$$

$$\alpha((N_s - n - 1)L_s^+) = -\alpha(nL_s^-). \quad (35)$$

In the case where $\alpha = 0$ (lossless fiber), all the terms in the sum in (31) are the same and, thus, $\Gamma(N_s, \omega, \omega_1, \omega_2) = 0$. If $\alpha(z) = \alpha$ everywhere except in $z = nL_s$ for $n = 0, 1, 2, ...$ where $\alpha(z) = \delta(z - nL_s)$ (EDFA amplification) we have

$$P((N_s - n)L_s^-) = P(nL_s^-), \quad (36)$$

$$P((N_s - n - 1)L_s^+) = P((n-1)L_s^+), \quad (37)$$

$$\alpha((N_s - n)L_s^-) = \alpha((n-1)L_s^+)$$

$$= \alpha((N_s - n - 1)L_s^+) = \alpha(nL_s^-), \quad \text{for} \ \ n = 1, 2, ..., \frac{N_s}{2} \quad (38)$$

and the OPC kernel becomes:

$$\Gamma(N_s, \omega, \omega_1, \omega_2) = \Xi^* \left(\frac{N_s}{2}, \omega, \omega_1, \omega_2 \right) G(\omega, \omega_1, \omega_2) \quad (39)$$

where

$$G(\omega, \omega_1, \omega_2) \triangleq \frac{(e^{-i\beta_2 \Delta\Omega L_s} e^{-\alpha L_s} - 1)(\alpha - i\beta_2 \Delta\Omega)}{\alpha^2 + \beta_2^2 \Delta\Omega^2}$$
$$+ \frac{(e^{-i\beta_2 \Delta\Omega L_s} - e^{-\alpha L_s})(\alpha + i\beta_2 \Delta\Omega)}{\alpha^2 + \beta_2^2 \Delta\Omega^2} \quad (40)$$

is the characteristic kernel of the optical channel in the presence of mid-link OPC. The Volterra series terms discussed above corresponds to the case of signal propagation in the optical fiber. A Volterra equalizer instead computes such terms for the *reverse propagation* case, corresponding to integrating (13) in the *backward* propagation direction. From standard calculus operations one can obtain the Volterra equalizer terms (for the case with OPC) as

$$A_{1X}(\omega, 0) = A_X(\omega, N_s L_s) e^{-i\frac{\beta_2}{2}\omega^2 N_s L_s}$$

$$A_{3X}(\omega, 0) = -j\gamma \frac{8}{9} \iint K'(\omega, \omega_1, \omega_2)$$
$$\times \Xi^* \left(-\frac{N_s}{2}, \omega, \omega_1, \omega_2 \right) G'(\omega, \omega_1, \omega_2) d\omega_2 d\omega_1 \quad (41)$$

where

$$F'(\omega, \omega_1, \omega_2) = \frac{1 - e^{\alpha L_s} e^{-j\beta_2 \Delta\Omega L_s}}{j\beta_2 \Delta\Omega - \alpha},$$

$$G'(\omega, \omega_1, \omega_2) = \frac{(e^{j\beta_2 \Delta\Omega L_s} e^{\alpha L_s} - 1)(\alpha - j\beta_2 \Delta\Omega)}{\alpha^2 + \beta_2^2 \Delta\Omega^2} \quad (42)$$
$$+ \frac{(e^{j\beta_2 \Delta\Omega L_s} - e^{\alpha L_s})(\alpha + i\beta_2 \Delta\Omega)}{\alpha^2 + \beta_2^2 \Delta\Omega^2}.$$

are the *inverse* Volterra kernels for backward propagation and $K'(\omega, \omega_1, \omega_2)$ is defined as in (2) but where the transmitted signal spectrum $S(\omega)$ is replaced by the received signal spectrum.

REFERENCES

[1] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, Feb 2010.

[2] E. Agrell, M. Karlsson, A. R. Chraplyvy, D. J. Richardson, P. M. Krummrich, P. Winzer, K. Roberts, J. K. Fischer, S. J. Savory, B. J. Eggleton, M. Secondini, F. R. Kschischang, A. Lord, J. Prat, I. Tomkos, J. E. Bowers, S. Srinivasan, M. Brandt-Pearce, and N. Gisin, "Roadmap of optical communications," *Journal of Optics*, vol. 18, no. 6, p. 063002, 2016.

[3] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Opt. Express*, vol. 16, no. 2, pp. 880–888, Jan 2008.

[4] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *Journal of Lightwave Technology*, vol. 26, no. 20, pp. 3416–3425, Oct 2008.

[5] E. F. Mateo, F. Yaman, and G. Li, "Efficient compensation of interchannel nonlinear effects via digital backward propagation in WDM optical transmission," *Opt. Express*, vol. 18, no. 14, pp. 15 144–15 154, Jul 2010.

[6] G. Liga, T. Xu, A. Alvarado, R. I. Killey, and P. Bayvel, "On the performance of multichannel digital backpropagation in high-capacity long-haul optical transmission," *Opt. Express*, vol. 22, no. 24, pp. 30 053–30 062, Dec 2014.

[7] K. V. Peddanarappagari and M. Brandt-Pearce, "Volterra series transfer function of single-mode fibers," *Journal of Lightwave Technology*, vol. 15, no. 12, pp. 2232–2241, Dec 1997.

[8] Y. Gao, F. Zhang, L. Dou, Z. Chen, and A. Xu, "Intra-channel nonlinearities mitigation in pseudo-linear coherent QPSK transmission systems via nonlinear electrical equalizer," *Optics Communications*, vol. 282, no. 12, pp. 2421 – 2425, 2009.

[9] Z. Pan, B. Chatelain, M. Chagnon, and D. V. Plant, "Volterra filtering for nonlinearity impairment mitigation in DP-16QAM and DP-QPSK fiber optic communication systems," in *2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference*, no. 2, 2011, p. JThA40.

[10] F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto, "Digital post-compensation using Volterra series transfer function," *IEEE Photonics Technology Letters*, vol. 23, no. 19, pp. 1412–1414, Oct 2011.

[11] ——, "Mitigation of intra-channel nonlinearities using a frequency-domain Volterra series equalizer," *Opt. Express*, vol. 20, no. 2, pp. 1360–1369, Jan 2012.

[12] F. P. Guiomar and A. N. Pinto, "Simplified volterra series nonlinear equalizer for polarization-multiplexed coherent optical systems," *Journal of Lightwave Technology*, vol. 31, no. 23, pp. 3879–3891, Dec 2013.

[13] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, "Digital signal processing for fiber nonlinearities," *Opt. Express*, vol. 25, no. 3, pp. 1916–1936, Feb 2017.

[14] L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, and Y. Cai, "Intrachannel nonlinearity compensation by inverse Volterra series transfer function," *Journal of Lightwave Technology*, vol. 30, no. 3, pp. 310–316, 2012.

[15] A. Bakhshali, W. Y. Chan, J. C. Cartledge, M. OSullivan, C. Laperle, A. Borowiec, and K. Roberts, "Frequency-domain Volterra-based equalization structures for efficient mitigation of intrachannel Kerr nonlinearities," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1770–1777, Apr 2016.

[16] A. Yariv, D. Fekete, and D. M. Pepper, "Compensation for channel dispersion by nonlinear optical phase conjugation," *Opt. Lett.*, vol. 4, no. 2, pp. 52–54, Feb 1979.

[17] R. A. Fisher, B. R. Suydam, and D. Yevick, "Optical phase conjugation for time-domain undoing of dispersive self-phase-modulation effects," *Opt. Lett.*, vol. 8, no. 12, pp. 611–613, Dec 1983.

[18] S. Watanabe and M. Shirasaki, "Exact compensation for both chromatic dispersion and kerr effect in a transmission fiber using optical phase conjugation," *Journal of Lightwave Technology*, vol. 14, no. 3, pp. 243–248, Mar 1996.

[19] C. Lorattanasane and K. Kikuchi, "Design theory of long-distance optical transmission systems using midway optical phase conjugation," *Journal of Lightwave Technology*, vol. 15, no. 6, pp. 948–955, Jun 1997.

[20] X. Liu, A. Chraplyvy, P. Winzer, R. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nature Photonics*, vol. 7, no. 7, p. 560, 2013.

[21] T. Umeki, T. Kazama, A. Sano, K. Shibahara, K. Suzuki, M. Abe, H. Takenouchi, and Y. Miyamoto, "Simultaneous nonlinearity mitigation in 92 × 180-Gbit/s PDM-16QAM transmission over 3840 km using PPLN-based guard-band-less optical phase conjugation," *Opt. Express*, vol. 24, no. 15, pp. 16 945–16 951, Jul 2016.

[22] P. Rosa, S. T. Le, G. Rizzelli, M. Tan, and J. D. Ania-Castañón, "Signal power asymmetry optimisation for optical phase conjugation using Raman amplification," *Opt. Express*, vol. 23, no. 25, pp. 31 772–31 778, Dec 2015.

[23] A. D. Ellis, M. Tan, M. A. Iqbal, M. A. Z. Al-Khateeb, V. Gordienko, G. S. Mondaca, S. Fabbri, M. F. C. Stephens, M. E. McCarthy, A. Perentos, I. D. Phillips, D. Lavery, G. Liga, R. Maher, P. Harper, N. Doran, S. K. Turitsyn, S. Sygletos, and P. Bayvel, "4 Tb/s transmission reach enhancement using 10 x 400 Gb/s super-channels and polarization insensitive dual band optical phase conjugation," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1717–1723, Apr 2016.

[24] G. Saavedra, Y. Sun, K. R. H. Bottrill, L. Galdino, F. Parmigiani, Z. Liu, D. J. Richardson, P. Petropoulos, R. I. Killey, and P. Bayvel, "Optical phase conjugation in installed optical networks," in *Optical Fiber Communication Conference*. Optical Society of America, 2018, p. W3E.2.

[25] G. Liga, G. Saavedra, and P. Bayvel, "Combining optical phase conjugation and Volterra equalisation: a novel nonlinearity compensation scheme," in *ECOC 2018; 44th European Conference on Optical Communication, Rome, Italy*, 2018, p. Tu1F.4.

[26] J. C. Cartledge, A. Kashi, A. D. Ellis, M. Mccarthy, A. Shiner, M. Reimer, and A. Borowiec, "Signal processing techniques for reducing the impact of fiber nonlinearities on system performance," no. 1, pp. 16–18.

[27] A. Vannucci, P. Serena, and A. Bononi, "The RP method: a new tool for the iterative solution of the nonlinear Schrodinger equation," *Journal of Lightwave Technology*, vol. 20, no. 7, pp. 1102–1112, Jul 2002.

[28] M. Taghavi, G. Papen, and P. Siegel, "On the multiuser capacity of WDM in a nonlinear optical fiber: coherent communication," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5008–5022, Nov 2006.

[29] V. Pechenkin and I. J. Fair, "Analysis of four-wave mixing suppression in fiber-optic OFDM transmission systems with an optical phase conjugation module," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 2, no. 9, pp. 701–710, Sep 2010.

[30] M. A. Z. Al-Khateeb, M. A. Iqbal, M. Tan, A. Ali, M. McCarthy, P. Harper, and A. D. Ellis, "Analysis of the nonlinear Kerr effects in optical transmission systems that deploy optical phase conjugation," *Opt. Express*, vol. 26, no. 3, pp. 3145–3160, Feb 2018.

[31] M. A. Z. Al-Khateeb, M. E. McCarthy, C. Sánchez, and A. D. Ellis, "Nonlinearity compensation using optical phase conjugation deployed in discretely amplified transmission systems," *Opt. Express*, vol. 26, no. 18, pp. 23 945–23 959, Sep 2018.

[32] B. Xu and M. Brandt-Pearce, "Modified Volterra series transfer function method," *Photonics Technology Letters, IEEE*, vol. 14, no. 1, pp. 47–49, 2002.

[33] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "Suppression of spurious tones induced by the split-step method in fiber systems simulation," *IEEE Photonics Technology Letters*, vol. 12, no. 5, pp. 489–491, 2000.

[34] N. K. Fontaine, X. Liu, S. Chandrasekhar, R. Ryf, S. Randel, P. Winzer, R. Delbue, P. Pupalaikis, and A. Sureka, "Fiber nonlinearity compensation by digital backpropagation of an entire 1.2-Tb/s superchannel using a full-field spectrally-sliced receiver," in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*, Sep 2013, pp. 1–3.

[35] R. Maher, T. Xu, L. Galdino, M. Sato, A. Alvarado, K. Shi, S. J. Savory, B. C. Thomsen, R. I. Killey, and P. Bayvel, "Spectrally shaped DP-16QAM super-channel transmission with multi-channel digital back-propagation," *Scientific reports*, vol. 5, p. 8214, 2015.

[36] K. Shi, E. Sillekens, and B. C. Thomsen, "246 GHz digitally stitched coherent receiver," in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, Mar 2017, pp. 1–3.

[37] A. Ghazisaeidi, "Fundamental limits of zero-forcing nonlinear equalization in WDM coherent fiber-optic systems," in *2018 European Conference on Optical Communication (ECOC)*, no. 1. IEEE, pp. 1–3.

[38] N. V. Irukulapati, H. Wymeersch, P. Johannisson, and E. Agrell, "Stochastic digital backpropagation," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 3956–3968, Nov 2014.

[39] R. Weidenfeld, M. Nazarathy, R. Noe, and I. Shpantzer, "Volterra nonlinear compensation of 100G coherent OFDM with baud-rate ADC, tolerable complexity and low intra-channel FWM/XPM error propagation," in *Optical Fiber Communication Conference*, 2010, p. OTuE3.

[40] G. Shulkind and M. Nazarathy, "Nonlinear digital back propagation compensator for coherent optical OFDM based on factorizing the Volterra series transfer function," *Optics Express*, vol. 21, no. 11, p. 13145, 2013.

[41] M. Secondini, S. Rommel, G. Meloni, F. Fresi, E. Forestieri, and L. Potì, "Single-step digital backpropagation for nonlinearity mitigation," *Photonic Network Communications*, vol. 31, no. 3, pp. 1–10, 2015.

[42] P. Johannisson and M. Karlsson, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," *Journal of Lightwave Technology*, vol. 31, no. 8, pp. 1273–1282, Apr 2013.