





# Edge-Carrier-Assisted Phase Retrieval at Low CSPR and Low Dispersion Diversity with Deep Learning

Daniele Orsuti , Marco Santagiustina  , *Member, IEEE*, Andrea Galtarossa  , *Fellow, IEEE, Fellow, OPTICA*, and Luca Palmieri  , *Senior Member, IEEE, Member, OPTICA*

**Abstract**—We investigate the performance of deep learning in recovering the complex-valued field of a weak-carrier-assisted single side-band signal from two intensity measurements that are decorrelated by dispersion. The proposed scheme relies on a supervised learning-based convolutional neural network to map the intensity measurements to the full-field. Unlike conventional iterative phase retrieval schemes, the proposed scheme does not require any iterations, digital upsampling, or pilot symbols, and can operate both at low carrier-to-signal-power ratio (CSPR) and at low applied dispersion value. Through numerical simulations in relevant system settings, we compare the performance of the proposed scheme with two recently proposed carrier-assisted iterative phase retrieval schemes: one based on the solution of a nonlinear optimization problem, and the other based on a modified Gerchberg–Saxton algorithm. The results show that the proposed scheme complies with the 7% hard-decision forward error correction threshold after 24 GBaud 32-QAM transmission over 100 km of standard single-mode fiber at a CSPR of 0 dB, with 3.6 times lower applied dispersion value, 30% to 90% lower complexity, and with less than 2 dB sensitivity penalty compared to conventional iterative phase retrieval schemes. These results support the potential of deep learning to realize phase retrieval-based coherent receivers that are compatible with the low complexity requirements of short-reach optical networks.

**Index Terms**—Phase retrieval, coherent receiver, neural network, Gerchberg–Saxton algorithm, carrier-to-signal-power ratio, direct-detection, short-reach communications.

## I. INTRODUCTION

THE demand for higher data transmission rates in short-reach optical networks, such as intra- and inter-data center interconnects, has set increasingly stringent requirements on the deployed transceivers in terms of spectral efficiency, complexity, power consumption, and form factor [1]. While direct detection has emerged as the dominant solution for short-reach transmission [2], [3], it has inherent limitations compared to coherent detection, which can measure the full complex field of the optical signal. Coherent detection enables the use of higher-order modulation formats and the compensation of both linear

and nonlinear propagation impairments; however, coherent receivers are considered costly for short-reach applications, mainly due to the need of a stable local oscillator laser.

To simplify the coherent receiver structure, recent studies have focused on developing phase retrieval receivers [4]–[6], which can retain the simplicity of direct detection while recovering the complex-valued field of the optical signal through digital signal processing (DSP). In this context, self-coherent systems have been proposed as a solution to eliminate the need for a local oscillator laser [7], as they rely on a continuous-wave (CW) tone generated at the transmitter side along with the information-bearing signal. This configuration gives rise to various self-coherent detection schemes, including the Kramers-Kronig (KK) receiver, the Stokes-vector receiver [8], and carrier-assisted differential detection [9]. The Stokes-vector receiver and carrier-assisted differential detection require at least three intensity measurements and an optical hybrid, yielding a complex receiver structure. Instead, the KK receiver enables full-field recovery after single-photodiode direct detection, exploiting the KK relations to cancel (or suppress) signal-to-signal-beat interference (SSBI). The KK receiver has been shown to offer improved SSBI cancellation performance compared to other iterative phase retrieval schemes [10], and has been widely investigated since its introduction by Mecozzi *et al.* [11]. Yet, the KK receiver requires a carrier-to-signal-power ratio (CSPR) higher than 6 dB to achieve the minimum-phase condition, which enhances the impact of nonlinear fiber propagation effects, and increases the requirements of the digital-to-analog converter (DAC) at the transmitter side [12]. Therefore, there is significant interest in developing novel phase retrieval schemes that can achieve phase retrieval with low steep complexity and at low CSPR.

Recently, carrier-less phase retrieval has been achieved by measuring two (or more) intensity waveforms related by a known amount of experienced chromatic dispersion [13], [14]. The phase retrieval task is accomplished using a modified version of the Gerchberg–Saxton (GS) algorithm [15], which seeks the optimum phase satisfying the constraints set by the measured intensity waveforms. However, in the simplest configuration, i.e., where two intensity waveforms are measured before and after a dispersive element, GS-phase retrieval requires hundreds of iterations, 5% to 20% pilot symbols, and high applied dispersion values (ADV) to prevent stagnation in local minima and to achieve satisfactory bit error rates (BERs) [13], [16], [17].

Manuscript received 18 June 2023; revised 25 October 2023; revised XXXX. This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on “Telecommunications of the Future” (PE0000001 - program “RESTART”) and by MIUR (PRIN 2017, project FIRST).

Daniele Orsuti, Marco Santagiustina, Andrea Galtarossa, and Luca Palmieri are with the Department of Information Engineering, University of Padova, 35131 Padova, Italy (corresponding author e-mail: daniele.orsuti@phd.unipd.it).

Luca Palmieri and Marco Santagiustina are also with the National Laboratory of Advanced Optical Fibers for Photonics (FIBERS), National Inter-University Consortium for Telecommunications (CNIT), L’Aquila, Italy.

To relax the CSPR requirements of self-coherent systems and address the complexity and convergence issues of the GS algorithm, a promising solution is to combine self-coherent transmission with the use of dispersive elements [18]–[20]. The underlying idea is to utilize a weak CW tone to obtain a rough estimate of the true phase; although this estimate is corrupted by SSBI, it can still be used to initialize the phase retrieval algorithm to facilitate its convergence. Basically, the CW tone plays a similar role as pilot symbols, but it helps convergence without reducing the net capacity of the system. This CW tone-based initialization was first exploited in Ref. [18], where the output of a KK receiver operating at low CSPR was used to initialize a nonlinear optimization algorithm to solve for the phase satisfying two intensity measurements decorrelated by dispersion; hereinafter we refer to this scheme as the enhanced KK (EKK) receiver. EKK achieves similar performance to the KK receiver with a lower CSPR (5 dB to 6 dB lower); yet the nonlinear optimization process significantly increases the computational complexity. Another approach for phase retrieval that exploits a CW tone-based initialization is the edge-carrier-assisted (ECA)-GS algorithm proposed in Ref. [19]. The ECA-GS scheme achieves successful phase retrieval with a few tens of iterations and with reduced computational requirements compared to the EKK scheme. It is worth mentioning that, while EKK and ECA-GS schemes rely on a CW tone generated at the transmitter at the edge of the information-bearing signal spectrum, recent works have investigated the performance of central-carrier-assisted (CCA)-phase retrieval [6], [19], [21], where the CW tone lies at the center of the information-bearing signal spectrum. CCA phase retrieval enables relaxed electrical bandwidth requirements compared to ECA phase retrieval; however, it requires higher CSPR values, along with the insertion of a guard-band around 0 Hz to mitigate the impact of chromatic dispersion-induced power fading [21]. For these reasons, in this work, we focus on ECA phase retrieval rather than on CCA phase retrieval.

Both the EKK and ECA-GS algorithms face a challenging trade-off between increasing the CSPR to obtain more accurate initial phase estimates and decreasing the CSPR to reduce the impact of nonlinear impairments caused by the carrier component. Operating at a low CSPR is the preferred choice, but it can lead to suboptimal solutions due to the poor initialization of the phase retrieval algorithm, especially when the intensity measurements are not sufficiently decorrelated (i.e., at low ADVs). Consequently, at low CSPR, both EKK and GSA require high ADVs, typically higher than 1000 ps/nm to achieve the 7% hard-decision forward error correction (HD-FEC) threshold with low sensitivity penalties [19]. In order to achieve high ADVs, one approach is to realize the dispersive element using several kilometers of optical fiber, but this leads to bulky fiber modules that do not meet the small form factor requirements. A more viable option is to realize the dispersive element on an integrated chip; however, achieving high dispersion over a wide bandwidth with low losses on-chip presents a significant challenge [22]. For these reasons, developing phase retrieval schemes that operate both at low CSPR and at low ADVs is of great interest. In this context,

deep learning has recently emerged as a promising method to perform the phase retrieval task in optical fiber communication systems, showing promising results when operating at low CSPR [21], [23], [24].

In this work, we propose a deep learning-based phase retrieval scheme to recover the phase of a weak-carrier assisted single side-band (SSB) signal from two intensity measurements that are decorrelated by a dispersive element, expanding preliminary results reported in [25]. We present a comparative analysis between the proposed scheme, the EKK scheme, and the ECA-GS scheme. We show that the NN-based scheme successfully recovers 32-QAM waveforms at a CSPR of 0 dB after 5-channel WDM transmission (24 GBaud channels) over 100 km of standard single-mode fiber (SSMF). Notably, these results are achieved with a dispersion value that is 3.6 times lower, a complexity reduced by 30% to 90%, and with less than 2 dB OSNR penalty compared to conventional iterative phase retrieval schemes.

The paper is organized as follows. Section II presents the numerical transmission system developed to assess the performance of edge-carrier-assisted phase retrieval. Section III begins by reviewing the working principles of EKK and ECA-GS schemes; it then introduces the proposed NN scheme, including details about the NN model architecture as well as the training and test procedures for the NN model. Section IV reports and compares the performance of EKK, ECA-GS, and NN-based phase retrieval. Finally, Section V evaluates and compares the computational complexity of the considered phase retrieval schemes.

## II. SYSTEM DESCRIPTION

To evaluate the performance of edge-carrier-assisted phase retrieval when using either EKK, ECA-GS or the proposed NN-based scheme, we simulated the 5-channel WDM transmission system shown in Fig. 1(a). For each of the five transmitters, 32-QAM Gray-coded symbols are generated from random bits at a symbol rate of 24 GBaud, which are then upsampled and shaped with a raised-cosine (RC) fundamental waveform with a roll-off factor 0.05. The CW tone is added virtually at the transmitter, exactly at the edge of the information-bearing signal spectrum [27]. The resulting signal is sent to an ideal DAC (i.e., without quantization and without bandwidth limitation) where electrical to optical conversion is performed by an IQ modulator biased at the null point. The IQ modulator is driven by a laser source operating at 1550 nm with 1 MHz linewidth and a relative intensity noise of  $-139$  dBc/Hz, which are typical values for low-cost distributed-feedback (DFB) laser diodes [28]. The WDM channels are multiplexed with a channel spacing of 40 GHz by a WDM-MUX that has the same channel spectral response of the optical filter employed at the receiver. The fiber link consists of 100 km of G.652 SSMF (single span) with an attenuation coefficient of 0.2 dB/km, a chromatic dispersion coefficient of 17 ps/nm/km, and a nonlinear parameter of  $1.3 \text{ W}^{-1}\text{km}^{-1}$ . The output of the fiber link is amplified by an erbium-doped fiber amplifier (EDFA) with a noise figure of 5 dB operating in transparency conditions. The optical filter at the receiver has 12 th-order

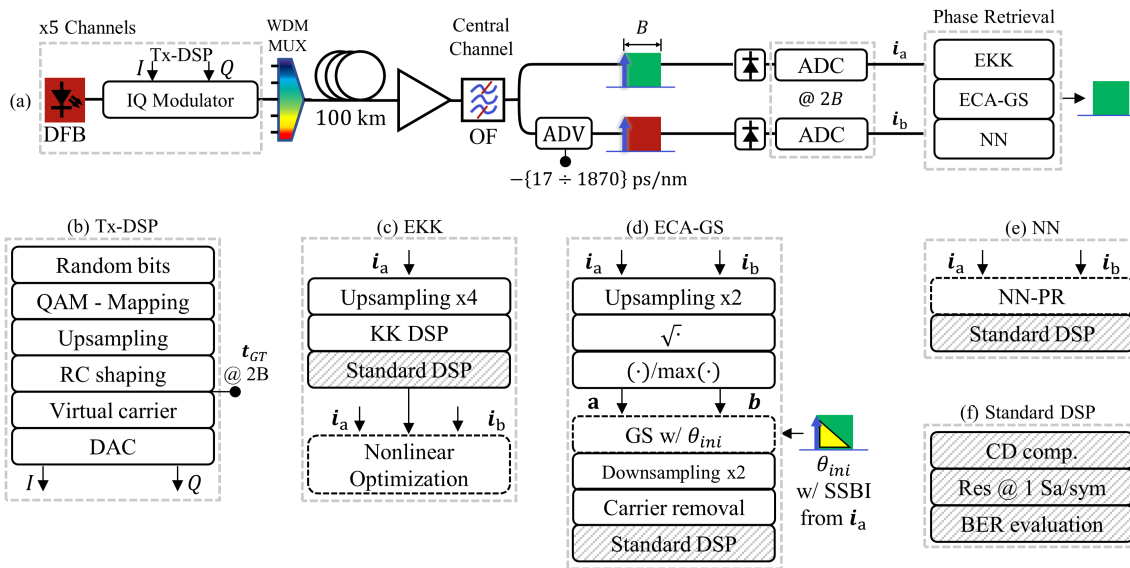


Fig. 1. (a) Simulation setup for 5-channel WDM transmission over 100 km of SSMF. DSP chain for (a) A transmitter section, (b) The Enhanced Kramer Kronig Receiver (EKK) [18], i.e., the KK receiver followed by nonlinear optimization [26]. (d) The Edge-carrier-assisted Gerchberg–Saxton (ECA-GS) algorithm [19], (e) The proposed neural network (NN) scheme. (f) Standard DSP block. The optical filter (OF) selects the central channel.  $t_{GT}$  is the transmitted ground-truth signal,  $i_a$  ( $i_b$ ) is the photocurrent waveform without (with) applied dispersion at the receiver.

super-Gaussian shape with 3 dB bandwidth of 36 GHz [29], and selects the central channel. The filtered signal enters into a two-branch receiver that measure two intensity waveforms:  $i_a$ , without passing through a dispersive element, and  $i_b$ , after passing through a dispersive element. The PIN photodiodes have 1 A/W responsivity and a bandwidth of 29 GHz; thermal and shot noise due to photodetection are included in the system. The analog-to-digital (ADC) converters have a vertical resolution of 8 bits and their sampling frequency is set to  $2B$ , where  $B$  is the information-bearing signal bandwidth after RC shaping. The output of the ADCs is fed to the phase retrieval schemes shown in Fig. 1, namely, EKK [Fig. 1(c)], ECA-GS [Fig. 1(d)], and NN [Fig. 1(e)]. For each phase retrieval scheme, after full-field recovery, standard DSP is applied, consisting in chromatic dispersion compensation of the fiber link with frequency domain-based processing, downsampling to one sample per symbol, and BER evaluation [Fig. 1(f)].

### III. PHASE RETRIEVAL RECEIVER SCHEMES

In this section, we first briefly recall the EKK and the ECA-GS schemes. Next, we describe the proposed NN-based phase retrieval scheme.

#### A. Enhanced KK Receiver

In the EKK receiver proposed in Ref. [18], the photocurrent waveform  $i_a$  (with a weak carrier) is digitally upsampled by a factor of  $R_{KK} = 4$  and is fed to the conventional KK receiver; the upsampling operation allows to accommodate the spectral broadening introduced by the nonlinear operations entailed in the KK receiver DSP (i.e., square root, logarithm and exponential functions) [11]. The upsampling/downsampling operation and the Hilbert transform in the KK-DSP are implemented using frequency domain processing. The output

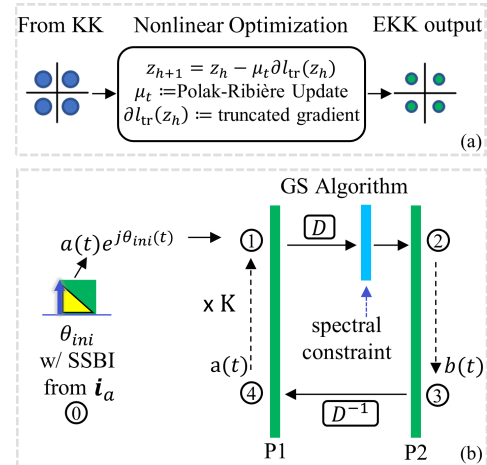


Fig. 2. Iterative phase retrieval schemes that combine the use of a CW tone (used for initialization) with the use of a dispersive element. (a) Nonlinear optimization block entailed in the EKK receiver scheme [18];  $z_h$  denotes the estimated symbols at step  $h$ . (b) ECA-GS algorithm [19];  $D$  denotes the dispersion applied at each iteration of the GS algorithm to transition between projection P1 (on the undispersed plane) and projection P2 (on the dispersed plane).

of the KK receiver is downsampled at symbol period, and the resulting symbols initialize the nonlinear optimization problem defined by the loss function given by (3) in [18]. Next, the generalized gradient expression derived in [26], is used in the Polak-Ribière version of the conjugate gradient method, which iteratively refines the initial guess provided by the KK receiver. Figure 2(a) shows a schematic representation of the steps involved in the nonlinear optimization block.

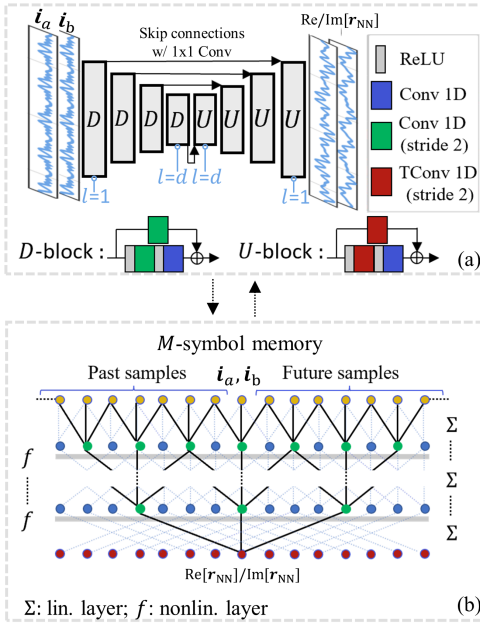


Fig. 3. (a) Encoder-decoder temporal CNN for phase retrieval task.  $r_{NN}$  denotes the complex-valued signal predicted by the NN, D/U-block: downsampling/upsampling block, Conv 1D: 1D convolutional layer, TConv 1D: 1D transposed convolutional layer,  $l$  is the D-block/U-block index, and  $d = 4$  is the model depth. (b) Equivalent representation of the input/output relation of the NN model using strided convolutions. The number of filters for each convolutional layer is  $n_f$  (see main text). The kernel size is 3 for all the convolutional layers except for the long skip connections outside the D/U-blocks where the kernel size is 1. The output convolutional layer has 2 filters, i.e.,  $\text{Re}[r_{NN}]$  and  $\text{Im}[r_{NN}]$ .

### B. Edge-Carrier-Assisted GS

In the ECA-GS algorithm proposed in Ref. [19], a SSB filtering operation is applied to  $i_a$  to obtain an estimate of the sought phase, which is required for initializing the GS algorithm. To accommodate the spectral broadening generated by the square root operation entailed in GS processing, the photocurrent waveforms  $i_a$  and  $i_b$  are upsampled by a factor  $R_{\text{ECA-GS}} = 2$ . The convergence of the GS algorithm is aided by a low-pass filter of bandwidth  $B$  applied during each iteration, which removes the spectral components outside the information-bearing signal bandwidth. Figure 2(b) shows a schematic representation of the steps involved in the GS algorithm.

### C. NN-based Phase Retrieval

Inspired by the successful results achieved by neural networks (NNs) for phase retrieval tasks in coherent imaging techniques and holographic image reconstruction [30], [31], we propose a supervised learning NN model to recover the in-phase (I) and quadrature components (Q) components of a weak-carrier assisted SSB signal from the intensity measurements  $i_a$  and  $i_b$ . Addressing the phase retrieval task with a NN is of interest for the following reasons.

- After the training process, the NN does not require an initial estimate of the I/Q components; this is in contrast to EKK and ECA-GS methods, where the initial estimate can be strongly compromised at low CSPR.

- The NN can operate at  $2B$ , i.e., without the need to upsample the photocurrent signals, as there is no spectral broadening introduced by nonlinear operations, such as the ones entailed in the KK or GS algorithms.
- During the training procedure the NN learns to extract relevant features related to the employed modulation format, which can enhance the robustness of the phase retrieval task to impairments.
- After training, the NN can achieve phase retrieval in a single forward-pass of its layers, which typically results in shorter computation times compared to iterative algorithms. Additionally, by carefully designing the NN architecture, it is possible to reduce the computational complexity of the NN below that of conventional iterative algorithms.

1) *NN model*: To realize low-complexity phase retrieval with a NN, we rely on the encoder-decoder temporal convolutional neural network (CNN) [32] shown in Fig. 3(a), which consists of stacked non-causal 1D convolutional layers alternated with rectified linear unit (ReLU) activation functions. The NN receives as input the two intensity waveforms  $i_a$  and  $i_b$  and predicts the sought I/Q components; the intensity waveforms  $i_a$  and  $i_b$  are fed to the NN as two channels of a 1D convolutional layer<sup>1</sup>. The encoder-decoder structure allows to increase the memory size of the NN model [33]; the encoder path is implemented using strided convolutions, whereas the decoder path is implemented using fractionally strided convolutions (also known as transposed convolutions). In the NN architecture, skip connections allow the features extracted from the downsampling (D)-blocks to be concatenated with the features extracted from the upsampling (U)-blocks [34]; this approach has been observed to improve the I/Q reconstruction performance. Figure 3(b) depicts an equivalent representation of the NN model, which shows that multiple input samples are used to predict a single output sample when strided convolutions are used. The memory size of the NN model needs to be properly selected to avoid introducing performance penalties in the phase retrieval task, and can be tuned by varying the number of stacked strided convolutional layers. We define the model depth,  $d$ , as the number of times the input signal is reduced in spatial dimensions (i.e., passes through a 1D strided convolutional layer) before being expanded back in the decoder; this corresponds to the number of D-blocks in the architecture. The model depth of the NN shown in Fig. 3 is  $d = 4$ , which yields a memory size  $M = 77$  symbols, computed as detailed in Ref. [23]. The chosen value of  $M$ , will allow the NN to handle the symbol mixing caused by all the investigated ADVs at the receiver.

2) *Training and Test procedure*: To ensure that the NN focuses solely on learning the phase retrieval task, we eliminate any influence of propagation-related impairments during the NN training process; to achieve this, we train the NN in back-to-back (B2B) settings. Following the training, we evaluate the phase retrieval performance of the neural network in both B2B and in the presence of a fiber link. In the latter case, the

<sup>1</sup>In temporal CNNs, the measured waveforms are typically arranged into different channels of a 1D convolutional layer, analogously to how the color planes (such as RGB) of an image are arranged in a 2D convolutional layer when dealing with image processing tasks [32].

TABLE I  
TRAINING SET PARAMETERS AND TEST SETS PARAMETERS.

	Training set	Test set B2B	Test set WDM
Settings	B2B <sup>1</sup>	B2B <sup>1</sup>	WDM <sup>2</sup>
CSPR [dB]	-1 or {-2, -1, 0, 1}	-1	-3 to 3 step: 1
OSNR [dB]	26	26	12 to 37 step: 2
ADV [ps/nm]		-17 to -1870	
Number of signals	27, 200	50	50
Signal length [symbols]	2 <sup>9</sup>	2 <sup>11</sup>	2 <sup>11</sup>

<sup>1</sup> B2B: back-to-back transmission, i.e., as depicted in Fig. 1(a), but without the fiber link.

<sup>2</sup> WDM: 5-channel WDM transmission as depicted in Fig. 1(a).

chromatic dispersion of the fiber link is compensated at the output of the NN with frequency domain processing, as for EKK and ECA-GS. Note that this training configuration differs from the typical one used for NN-based equalizers [35]–[37], where the training data must be affected by the propagation-related impairments of the channel being considered to learn the channel equalization task.

To train the NN in B2B settings, we remove the fiber link in the setup of Fig. 1, then we collect  $N$  pairs of signals denoted by  $\{\hat{i}_h, \mathbf{y}_h\}_{h=1}^N$ , where,  $\hat{i} := [\hat{i}_a, \hat{i}_b]$  contains the digitized intensity waveforms and  $\mathbf{y} := [\mathbf{I}, \mathbf{Q}]$  contains the corresponding ground truth I/Q components generated at the transmitted side. During the training phase we seek for the NN model parameters that minimize the normalized root mean squared error (NRMSE) [38], between the I/Q components predicted by the NN model ( $\hat{\mathbf{I}}_h/\hat{\mathbf{Q}}_h$ ) and the ground truth I/Q components ( $\mathbf{I}_h/\mathbf{Q}_h$ ), over all the  $N$  signal pairs ( $h = 1, \dots, N$ ). A different NN model is trained for each of the considered ADVs. For ADVs in the range 17 – 510 ps/nm the number of filters in the convolutional layer,  $n_f$ , is set to  $n_f = 28$ ; for higher ADVs we noticed that a higher number of filters was required to avoid performance penalty. Therefore, we set  $n_f = 64$  and  $n_f = 74$  for ADVs 1360 ps/nm and 1870 ps/nm, respectively.

For each ADV, the training set consists of  $N = 27, 200$  signal pairs with length 2<sup>9</sup> symbols. We consider two training sets that differ based on the CSPR values of the intensity waveforms that they include. The first training set includes intensity waveforms at a CSPR of -1 dB only, whereas the second training set includes intensity waveform with CSPR values in the set {-2, -1, 0, 1} dB. In the latter case, the number of intensity waveform for each CSPR value is  $N/4 = 6800$ . For both training sets, the intensity waveforms have been collected at an optical signal-to-noise ratio (OSNR) of 26 dB, where the OSNR includes both the optical power of the signal and the CW tone power and refers to a noise bandwidth of 0.1 nm [29]. We found that training the NN at an OSNR of 26 dB, rather than at a higher OSNR, allows the amplified spontaneous emission (ASE) noise in the data to act as a regularizer, thereby enhancing the extrapolation capabilities of the NN. Adam-based optimization with learning rate 10<sup>-3</sup> tunes the NN model parameters for 300 epochs using a batch

size of 256. The trainings take ~ 15 minutes on an NVIDIA Quadro RTX 5000 GPU.

To test the NN, the test data are generated in WDM transmission with a random number generator independent from the one used for the training phase [36], [39]. Specifically, the training set is generated using the PCG64 generator [40], whereas the test set using the Mersenne Twister generator [41], in both cases using different random seed for each generated signal pair in the datasets. We transmit 50 sequences of 2<sup>11</sup> symbols for each estimated BER point, where the choice of transmitting multiple 2<sup>11</sup>-symbol long sequences rather than a single long sequence is merely for numerical convenience. Observe that the training set the sequence length is 2<sup>9</sup> symbols, which has been selected to limit the computational requirements during the training phase, whereas in the test set the sequence length has been extended to 2<sup>11</sup>; the proposed NN is fully-convolutional (i.e., without dense layers), so the trained layers can be applied in a sliding window manner to the longer sequences in the test set without the need to re-train the NN model. The NN is tested over the OSNR range from 12 dB to 37 dB, and CSPR range from -3 dB to 3 dB. Notice that most of the test set parameters are not included in the training set data. This allows to evaluate the extrapolation capabilities of the NN model for parameters outside the training set. The training set parameters and the test set parameters are summarized in Table I.

Although in this work we focus on single-polarization phase retrieval, the EKK, ECA-GS, and NN schemes can in principle also be extended to polarization multiplexed transmission to compete with standard coherent detection. Yet, this extension is not straightforward as the phase retrieval problem requires additional constraints to reduce phase ambiguity. For example, in a recent study [42], it was proposed to generate two CW tones propagating alongside with the information-bearing signal: one at the left edge and one at the right edge. This approach allows to mitigate carrier fading induced by the random polarization state rotation along the fiber link. Therefore, to extend the NN-based phase retrieval to polarization multiplexed transmission, it is necessary to determine a new training procedure and suitable NN model parameters.

#### IV. RESULTS AND DISCUSSION

In this section, we compare the phase retrieval performance of EKK, ECA-GS, and the proposed NN scheme. We first evaluate the performance in B2B settings, then we present the performance after 5-channel WDM transmission over 100 km of SSMF.

##### A. B2B Performance

Figures 4(a)-(c) show the mean absolute phase error between the transmitted and recovered signals,  $\langle |\Delta\theta| \rangle = \langle |\arg\{t_{GT} \cdot \mathbf{r}^*\}| \rangle$ , for different ADVs in B2B settings. The OSNR and the CSPR of the test set are 26 dB and -1 dB, respectively, which correspond to the parameters used to train the NN when the training set includes intensity waveforms at the CSPR value of -1 dB only. Note that for the EKK [Fig. 4(a)] and the ECA-GSA [Fig. 4(b)], the x-axis represents the iteration

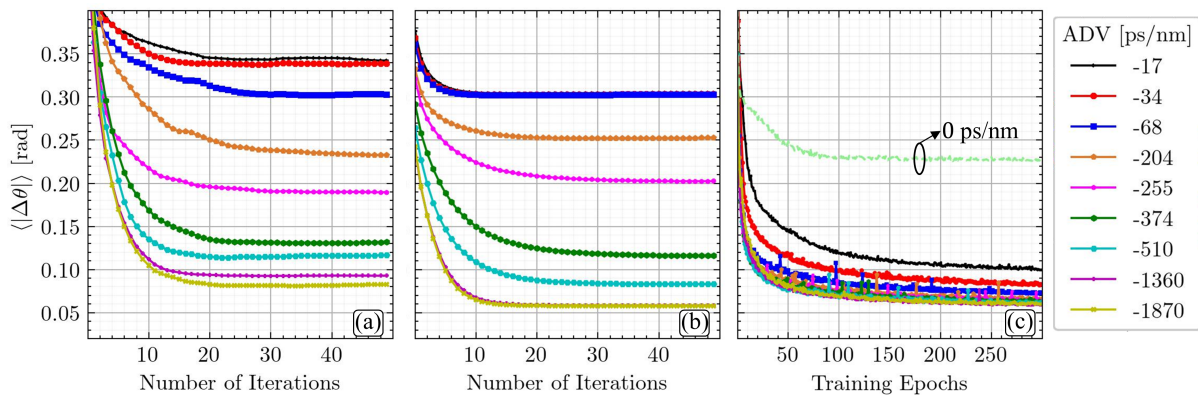


Fig. 4. Mean absolute phase error,  $\langle |\Delta\theta| \rangle$ , between the transmitted and recovered signals versus iteration number for (a) EKK scheme and (b) ECA-GS scheme. (c)  $\langle |\Delta\theta| \rangle$  versus training epoch for the NN-based scheme. The curves are obtained in B2B settings at an OSNR of 26 dB and at a CSNR of  $-1$  dB and for the different ADVs in the legend.

number, whereas for the NN the x-axis represents the training epoch number [Fig. 4(c)]. Both the EKK and the ECA-GS schemes achieve better performance as the ADV increases; this can be explained by the fact that interfering more symbols together reduces the chances for the iterative phase retrieval algorithms getting trapped in suboptimal solutions. Compared to ECA-GS, the EKK scheme has a slower convergence rate and saturates to higher  $\langle |\Delta\theta| \rangle$  values as the ADV increases. This can be attributed to the different phase initializations (KK output for EKK versus SSB filtering for ECA-GS) and to the adverse impact of noise on the nonlinear optimization algorithm entailed in EKK. Remarkably, the proposed NN-based scheme [Fig. 4(c)] enables accurate phase retrieval at low ADVs, for which the EKK and ECA-GS methods cannot converge to satisfactory  $\langle |\Delta\theta| \rangle$  values. Indeed, for the ADV in the range 17-374 ps/nm, the NN achieves a  $\langle |\Delta\theta| \rangle$  floor that is significantly lower than the other schemes. For the high ADVs 1360 ps/nm and 1870 ps/nm, the ECA-GS and the NN achieve similar performance, which are limited only by the noise floor.

As a sanity check, in Fig. 4(c) we plot the performance of the NN model when employing single-photodiode direct detection (which corresponds to 0 ps/nm of applied dispersion); it can be seen that the proposed scheme fails to converge to satisfactory  $\langle |\Delta\theta| \rangle$  values due to the ill-posed nature of the phase retrieval problem, i.e., for an ADV of 0 ps/nm and at the low CSNR of  $-1$  dB there exist many complex signals carrying the same intensity waveform [11].

The relaxed ADVs requirements of the NN-phase retrieval can also be seen in Fig. 5, where we plot the constellation diagrams for selected ADVs. The NN successfully reconstructs the constellation diagram for ADV as low as 34 ps/nm (for each constellation diagram, we report the NRMSE on the top left corner). Observe that, for ADV 1360 ps/nm, the NN achieves a higher NRMSE value compared to ECA-GS; this is explained by the higher number of filters in the NN model (from  $n_f = 28$  for ADVs up to 510 ps/nm, to  $n_f = 64$  for 1360 ps/nm), which leads to an higher number of parameters, and, therefore, to a more complex NN model training.

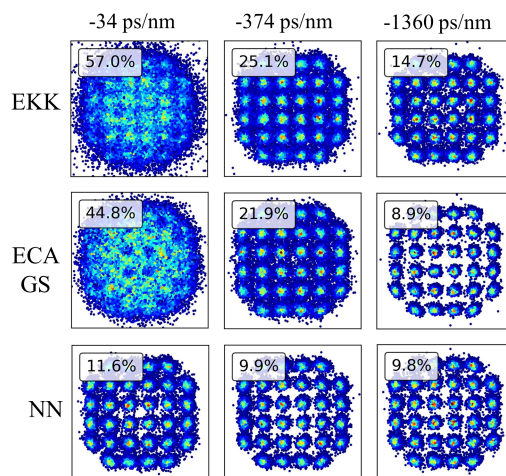


Fig. 5. Constellation diagrams reconstructed by EKK (first row), ECA-GS (second row) and NN (third row) at different ADVs shown at the top of each column. The diagrams are obtained in B2B settings at an OSNR of 26 dB and at a CSNR of  $-1$  dB. The phase retrieval scheme varies across rows, whereas the ADV varies across columns. The NRMSE, which was used as the loss function of the NN model, is displayed on the top left corner of each diagram.

## B. Transmission Performance

The performance after 5-channel WDM transmission over 100 km are shown in Figs. 6 and 7. When introducing the chromatic dispersion of the fiber link, the NN (trained in B2B settings) recovers the full-field up to a constant amplitude scaling and constant phase offset, which are dependent only on the total chromatic dispersion introduced by the fiber link and can be easily compensated for [23], [43]. The number of iterations for the EKK and the ECA-GS is set to  $K_{EKA-GS} = 20$  and  $K_{EKK} = 40$ , respectively, which were found to be the optimal values for all the considered OSNRs and ADVs in our simulations. The reason for the higher number of iterations needed for EKK compared to ECA-GS is the slower convergence rate of the nonlinear optimization algorithm entailed in EKK. Figures 6(a)-(d) show the BER performance as a function of CSNR at a fixed OSNR of 25.3 dB. As the

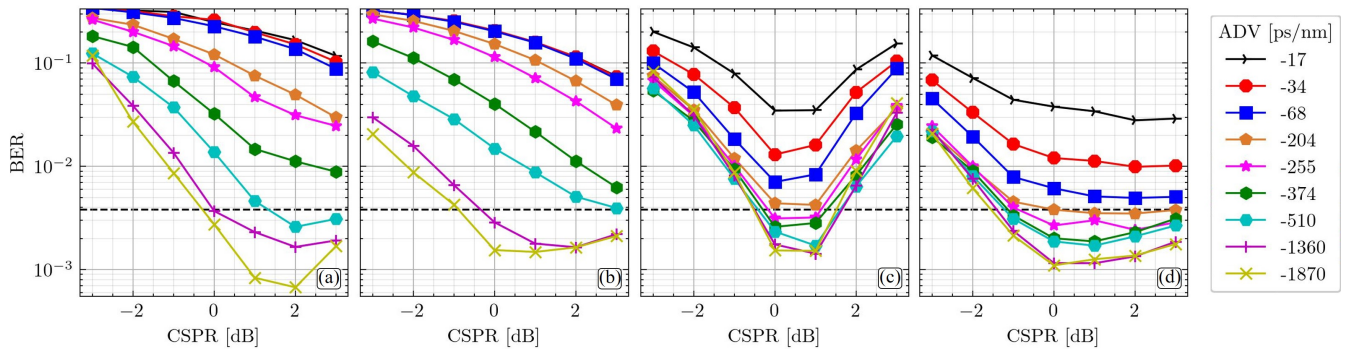


Fig. 6. BER versus CSPPR for the ADVs in the legend after 5-channel WDM transmission over 100 km of SSMF for a 24 GBaud 32-QAM modulated signal (central channel performance). (a) EKK performance, (b) ECA-GS performance, (c) NN performance for the NN model trained at a CSPPR of  $-1$  dB, and (d) NN performance for the NN model trained at CSPPRs in the set  $\{-2, -1, 0, 1\}$  dB. The test set CSPPR corresponds to the x-axis in the figures, namely  $-3$  dB to  $3$  dB. The OSNR is set to  $25.3$  dB. The horizontal black dashed line shows the 7% HD-FEC threshold.

CSPPR increases, both the EKK scheme [Fig. 6(a)] and the ECA-GS scheme [Fig. 6(b)] tend to require lower ADVs to achieve a target BER; this is because the initial phase estimate is closer to the true phase for higher CSPPRs. It can be seen that the ECA-GS scheme outperforms the EKK scheme since it requires lower CSPPRs to achieve the 7% HD-FEC threshold. For high ADVs, namely, 1360 ps/nm and 1870 ps/nm, the BER initially decreases with increasing CSPPR until an optimal value is reached, after which it starts increasing again; this trend is observed for both the EKK and the ECA-GS schemes. This optimal CSPPR value exists because increasing the CSPPR while keeping the OSNR fixed reduces the power of the information-bearing signal, leading to a higher impact of carrier-to-ASE noise beating.

Figures 6(c) and (d) show the BER versus CSPPR performance for the NN trained at a single CSPPR of  $-1$  dB and at multiple CSPPR values in the set  $\{-2, -1, 0, 1\}$  dB, respectively. It is evident that, by including intensity waveform at different CSPPRs in the training set, the NN generally offers better performance over the test set CSPPR range of  $-3$  dB to  $3$  dB. However, an interesting observation arises when the test set CSPPR is  $0$  dB, which corresponds to the minimum CSPPR at which optimum BER performance is achieved for most of the considered ADVs. In this case, training the NN across multiple CSPPRs [Fig. 6(d)] does not provide significant performance improvement compared to training it at CSPPR of  $-1$  dB only [Fig. 6(c)]. Consequently, one might choose to sacrifice the ability to operate over a broader CSPPR range to simplify the training set generation procedure, including intensity waveforms at a CSPPR of  $-1$  dB only. Notably, in Figs. 6(c) and (d), at low ADVs, the NN requires lower CSPPRs to achieve the 7% HD-FEC threshold compared to the EKK and the ECA-GS schemes. For instance, at a CSPPR of  $0$  dB, the NN achieves the 7% HD-FEC for an ADV of 255 ps/nm, while the EKK and ECA-GS schemes require an ADV of  $\sim 1360$  ps/nm to achieve the same threshold. For CSPPR values higher than  $1$  dB, the BER performance of EKK and ECA-GS benefits from the improved initial phase estimate, whereas the NN performance depends on the training configuration. When the NN is trained at CSPPR of  $-1$  dB only, it offers worse BER performance than the other phase retrieval schemes due to the limited

extrapolation capabilities at CSPPR lower than  $-1$  dB and higher than  $1$  dB. Instead, remarkably, when the NN is trained over multiple CSPPRs, it outperforms both EKK and ECA-GS across the entire test set CSPPR range and for all the considered ADV values.

Figure 7(a)-(d) shows the BER vs OSNR curves for different ADVs at a CSPPR of  $0$  dB; clearly, the NN outperforms both EKK and ECA-GS, achieving a target BER value with a significantly lower ADV. It can also be observed that all the BER curves exhibit an optimum OSNR operation point at which the minimum BER is achieved. Specifically, at low OSNR value, ASE noise limits the BER performance, whereas at high OSNR values, nonlinear impairments dominate the performance degradation due to the higher signal power (i.e., CW tone plus information signal powers). The NN outperforms both ECA-GS and EKK for all the considered ADVs, except for 1870 ps/nm, where ECA-GS and the NN achieves similar performance. For ADVs higher than 1870 ps/nm (not shown in the figure), all the schemes offer negligible sensitivity improvements at the 7% HD-FEC threshold. Notice that, at a CSPPR of  $0$  dB, for ADVs in the range  $17 - 510$  ps/nm, the NN trained at a CSPPR of  $-1$  dB offers similar BER versus OSNR performance to the NN trained over multiple CSPPR values. Therefore, in the following investigation, we consider the simpler training set configuration with a single CSPPR value.

To determine the effect of the NN model memory size on performance, in Fig. 8 we show the results for varying model depth,  $d$ , in the set  $\{2, 3, 4, 5\}$ ; these values corresponds to memory sizes of  $\{17, 37, 77, 157\}$  symbols. The ADV is 347 ps/nm. It can be seen that the BER performance improves as the model depth increases from 2 to 4. However, there is no advantage in increasing the model depth to 5, as the BER performance does not improve.

## V. COMPUTATIONAL COMPLEXITY

In this section, we determine the computational complexity,  $C$ , of the considered phase retrieval schemes defined as the number of required real multiplications per recovered output sample. We also investigate the trade-off between performance and complexity for the proposed NN-based phase retrieval scheme.

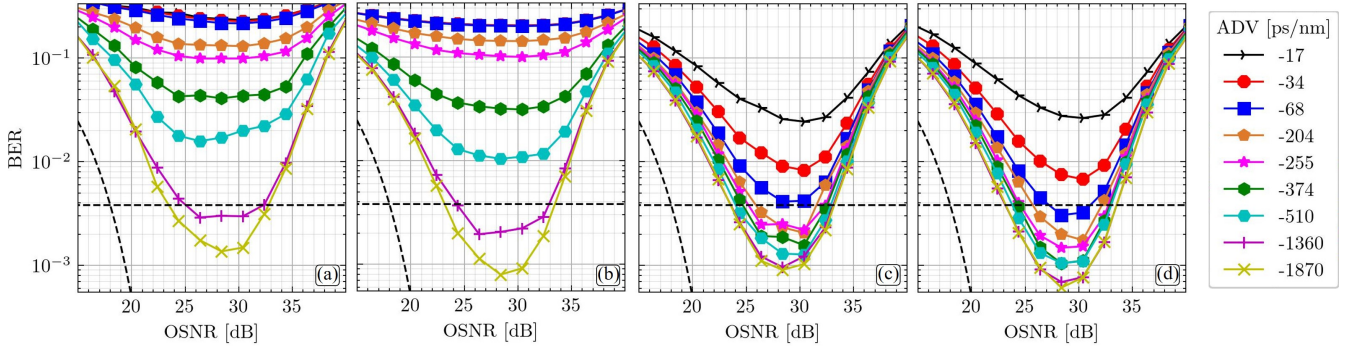


Fig. 7. BER versus OSNR for the ADVs in the legend after 5-channel WDM transmission over 100 km of SSMF for a 24 GBaud 32-QAM modulated signal (central channel performance). (a) EKK performance, (b) ECA-GS performance, (c) NN performance for the NN model trained at a CSPP of  $-1$  dB, and (d) NN performance for the NN model trained at CSPPs in the set  $\{-2, -1, 0, 1\}$  dB. The test set CSPP is set to 0 dB. The black dashed line shows the analytic expression for the BER of a 32-QAM modulated system impaired by AWGN, whereas the horizontal black dashed line shows the 7% HD-FEC threshold.

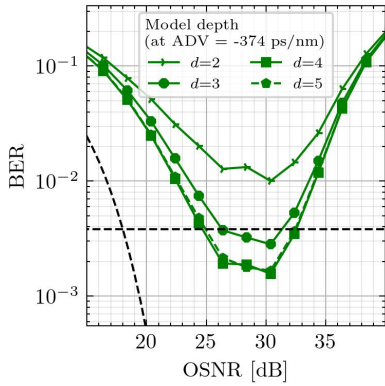


Fig. 8. BER versus OSNR performance at varying NN model memory sizes. The NN model memory is varied by tuning the NN model depth,  $d$ , in the set  $\{2, 3, 4, 5\}$ , which corresponds to the memory sizes  $\{17, 37, 77, 157\}$  symbols. The ADV is 374 ps/nm and the CSPP is 0 dB. The BER curves are obtained after 5-channel WDM transmission over 100 km of SSMF for a 24 GBaud 32-QAM modulated signal (central channel performance). The black dashed line shows the analytic expression for the BER of a 32-QAM modulated system impaired by AWGN, whereas the horizontal black dashed line shows the 7% HD-FEC threshold.

### A. Complexity of the EKK scheme

The complexity of the EKK receiver can be written as  $C_{\text{EKK}} = C_{\text{KK}} + C_{\text{NL}}$ , where  $C_{\text{KK}}$  and  $C_{\text{NL}}$  denotes the complexity of the KK receiver and the complexity of the non-linear optimization algorithm, respectively. For the complexity expression of the KK receiver we rely on the low-complexity time-domain implementation of the KK-DSP analyzed in Ref. [44]. The time-domain implementation of the KK receiver can achieve performance close to the frequency domain-based implementation provided that the number of taps of the employed FIR filters is sufficiently high [45]. The number of real multiplications per sample required by the KK algorithm is

$$C_{\text{KK}} = (3N_s + 2 + N_h/2)R_{\text{KK}}, \quad (1)$$

where  $N_s$  is the number of taps of the upsampling/downsampling filter and  $N_h$  is the number of taps of the Hilbert transform filter; we set  $N_s = N_h = 128$  [45], and the upsampling factor to  $R_{\text{KK}} = 4$ .

The primary contribution to the complexity of the nonlinear optimization algorithm comes from the gradient evaluations within the iterations of the nonlinear conjugate gradient method. Specifically, for each iteration of the conjugate gradient method, two gradient evaluations are made; the first gradient evaluation determines the steepest descent direction, whereas the second gradient evaluation is performed after the Polak-Ribière estimate of the conjugate direction. Each gradient evaluation involves two convolution operations: one between the estimated symbols and the fundamental RC pulse waveform, and the other between the estimated symbols with the same waveform but including the extra dispersion introduced by the dispersive element. Therefore, the number of real multiplications per sample required by the nonlinear optimization algorithm can be approximated as

$$C_{\text{NL}} = 4 \cdot K_{\text{EKK}} \cdot 2 \cdot (2 \cdot N_{\text{RC}}), \quad (2)$$

where the factor 4 converts from complex multiplications to real multiplications,  $N_{\text{RC}}$  denotes the number of taps of the RC fundamental waveform, which is set to  $N_{\text{RC}} = 127$ , and the number of iterations is  $K_{\text{EKK}} = 40$ . The value of  $N_{\text{RC}}$  needs to be chosen based on the roll-off factor of the RC shaping filter to achieve satisfactory pass-band and stop-band performance, as described in [46].

### B. Complexity of the ECA-GS scheme

To determine the complexity of the ECA-GS scheme,  $C_{\text{ECA-GS}}$ , we refer to implementation described in Ref. [47], which uses overlap-save processing with a 50% save ratio to enable block-wise phase retrieval with a block-size of 1024. The GS iterations are the most computationally intensive part of the ECA-GS scheme, and, therefore, we neglect the complexity associated with SSB filtering, upsampling and downsampling to determine the complexity expression. Each GS iteration involves three FFT/IFFT pairs, and the application of two intensity constraints, so that the number of real multiplications per sample reads:

$$C_{\text{ECA-GS}} = 4R_{\text{ECA-GS}}K_{\text{ECA-GS}} \left[ \frac{3(N \log_2 N + N)}{N - N_{\text{ovlp}} + 1} + 2 \right], \quad (3)$$



where the factor 4 is used to convert from complex multiplications to real multiplications,  $R_{\text{ECA-GS}} = 2$ , i.e., the photocurrent signals are upsampled by a factor of 2 as described in Section III-B, the number of iterations is set to  $K_{\text{ECA-GS}} = 20$ ,  $N = 1024$  is the FFT size, and the term  $N - N_{\text{ovlp}} + 1$ , with  $N_{\text{ovlp}} = N/2$ , denotes the number of output samples produced by each iteration of the overlap-save algorithm.

### C. Complexity of the NN scheme

Consistently with previous works in the literature, e.g., Ref. [48], we evaluate the complexity of the NN scheme by considering only the prediction phase and not accounting for the training phase. The complexity evaluation involves two steps. First, we evaluate the complexity of each D-block and U-block that are shown at the bottom of Fig. 3(a). Then, we sum the contribution of each D-block and U-block to obtain the total complexity. The complexity of the D-blocks reads [23], [49]

$$C_D^l = 3 (n_f^2 k / s^l), \quad l = 1, \dots, d \quad (4)$$

where the factor 3 accounts for the number of convolutional layers inside a D-block,  $n_f$  denotes the number of filters in each convolutional layer,  $k = 3$  is the kernel size,  $s = 2$  is the stride of the convolutional layers,  $l$  is the D-block index [see Fig. 3(a) for the values assumed by  $l$  across the NN model], and  $d$  is the total number of D-blocks in the NN model or, equivalently, the model depth (see Section III-C1). Equation (4) does not account for the complexity of ReLU activations, which have a negligible contribution compared to the convolutional layers, and holds for all the D-blocks, except for the one at the input, which has a lower complexity since the input channels are two, i.e.,  $i_a$  and  $i_b$ , instead of  $n_f$ . Analogously, the complexity of the U-blocks is

$$\begin{aligned} C_U^l &= 2 (n_f^2 k / s^l) + (n_f^2 k / s^{l-1}) \\ &= (2 + s) (n_f^2 k / s^l), \quad l = 1, \dots, d, \end{aligned} \quad (5)$$

where, the first term accounts for the complexity of the transposed convolutional layers, whereas the second term accounts for the complexity of the convolutional layer applied after upsampling. The total complexity for a NN model prediction,  $C_{\text{NN}}$ , can be obtained by summing (4) and (5).

It is worth mentioning that the number of filters in each convolutional layer,  $n_f$ , has the greatest impact on the NN complexity; therefore, it is important to properly tune the parameter  $n_f$  in order to achieve the target performance with the lowest possible complexity.

### D. Complexity Comparison

Figure 9 compares the complexity of the phase retrieval schemes considered in this work. As stated, the NN complexity strongly depends on the ADV, whereas that of EKK and ECA-GS is much less dependent on ADV (see Fig. 4). For fairness, the three approaches are compared when providing similar performances in terms of BER. Therefore, we consider EKK and ECA-GS with ADV of 1360 ps/nm, and the NN with ADV of just 374 ps/nm. The NN complexity is varied by varying the number of filters in the convolutional layers,  $n_f$ ,

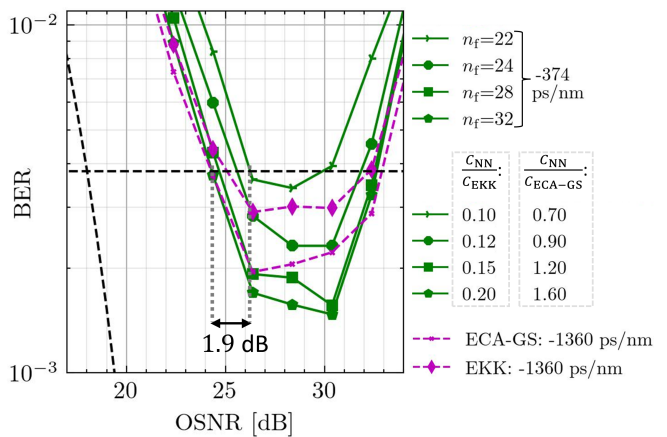


Fig. 9. Trade-off between performance and complexity for the NN-based phase retrieval scheme. The NN complexity is varied by varying,  $n_f$ , i.e., the number of filters in the convolutional layers. The NN operates with an ADV of 374 ps/nm, whereas both the ECA-KK scheme and the EKK scheme operate with an ADV of 1360 ps/nm.  $C_{\text{NN}}/C_{\text{ECA-GS}}$  and  $C_{\text{NN}}/C_{\text{EKK}}$  denote the relative complexity between the NN scheme and the ECA-GS scheme and between the NN scheme and the EKK scheme, respectively. The BER curves are obtained in 5-channel WDM transmission (central channel performance) with a CSFR of 0 dB. The black dashed line shows the analytic expression for the BER of a 32-QAM modulated system impaired by AWGN, whereas the horizontal black dashed line shows the 7% HD-FEC threshold.

in the set  $\{22, 24, 28, 32\}$ . For the EKK and ECA-GS we considered 40 and 20 iterations, respectively, as suggested by Fig. 4. The results in Fig. 9 show that the performance of the different schemes are comparable. Nevertheless, the complexity of the NN is substantially smaller than that of the EKK (see table  $C_{\text{NN}}/C_{\text{EKK}}$  in Fig. 9), and smaller to comparable with that of ECA-GS (see table  $C_{\text{NN}}/C_{\text{ECA-GS}}$ ). It can be also seen that the NN with ADV 374 ps/nm can achieve similar sensitivity performance (at the 7% HD-FEC threshold) to that of EKK and ECA-GS with ADV 1360 ps/nm, i.e., with 3.6 times lower ADV. Additionally, the NN reduces the computational complexity by up to 30% compared to ECA-GS and up to 90% compared to EKK, while incurring in less than 2 dB OSNR penalty.

## VI. CONCLUSION

We investigated the performance of deep learning in recovering the complex-valued field of a weak-carrier-assisted SSB signal from two intensity measurements that are decorrelated by a dispersive element. We presented a comparative analysis between the proposed NN-based scheme and two iterative schemes: a nonlinear-optimization-based phase retrieval scheme, and a Gerchberg-Saxton-based phase retrieval scheme. These two iterative schemes require an initial phase estimate to enable convergence, which can be strongly corrupted at low CSFR; for this reason, they require ADVs higher than 1000 ps/nm to achieve the 7% HD-FEC threshold at low CSFR. In contrast, the NN-based scheme proposed here does not require an initial phase estimate; instead, it relies on a supervised training process to find the optimal map between the intensity waveforms and the ground-truth I/Q components. Through numerical simulations in relevant transmission settings, we show that the NN offers a remarkable

improvement in performance at low ADVs, which can be attributed to the training procedure being carried out on a specific class of signals, i.e., 32-QAM RC shaped waveforms. We considered two training set configurations for the NN model. One with intensity waveforms at CSPR  $-1$  dB only, and the other with intensity waveforms at multiple CSPRs in the set  $\{-2, -1, 0, 1\}$  dB. In the multi-CSPR training configuration, the NN outperforms conventional iterative phase retrieval schemes over a broader CSPR range compared to the single-CSPR training configuration. The proposed scheme complies with the 7% HD-FEC threshold after 5-channel WDM transmission over 100 km of SSMF (per-channel symbol rate of 24 GBaud) with a CSPR of 0 dB, while requiring 3.6 times lower ADV, 30% to 90% lower complexity, and incurring in less than 2 dB OSNR penalty compared to the iterative phase retrieval schemes. We believe that this work paves the way to design dispersive elements-based phase retrieval schemes with both low hardware complexity and low computational complexity.

## REFERENCES

- [1] K. Zhong, X. Zhou, J. Huo, C. Yu, C. Lu, and A. P. T. Lau, "Digital Signal Processing for Short-Reach Optical Communications: A Review of Current Technologies and Future Trends," *J. Lightwave Technol.*, vol. 36, no. 2, pp. 377–400, Jan. 2018.
- [2] A. S. Karar, A. R. E. Falou, J. M. H. Barakat, Z. N. Gürkan, and K. Zhong, "Recent Advances in Coherent Optical Communications for Short-Reach: Phase Retrieval Methods," *Photonics*, vol. 10, no. 3, p. 308, Mar. 2023.
- [3] A. S. Karar, "Gerchberg-Saxton Based FIR Filter for Electronic Dispersion Compensation in IM/DD Transmission Part I: Theory and Simulation," *J. Lightwave Technol.*, vol. 41, no. 5, pp. 1335–1345, Mar. 2023.
- [4] Y. Yoshida, T. Umezawa, A. Kanno, and N. Yamamoto, "Phase-Retrieving Coherent Reception and its Sample Complexity," in *2020 European Conference on Optical Communications (ECOC)*, 2020, pp. 1–4.
- [5] E. S. Chou, H. Srinivas, and J. M. Kahn, "Phase Retrieval-Based Coherent Receivers: Signal Design and Degrees of Freedom," *Journal of Lightwave Technology*, vol. 40, no. 5, pp. 1296–1307, 2022.
- [6] M. Matsumoto, "Optical signal phase reconstruction based on temporal transport-of-intensity equation," *Journal of Lightwave Technology*, vol. 38, no. 17, pp. 4722–4729, 2020.
- [7] X. Chen, S. Chandrasekhar, and P. Winzer, "Self-Coherent Systems for Short Reach Transmission," in *2018 European Conference on Optical Communication (ECOC)*, 2018, pp. 1–3.
- [8] D. Che, A. Li, X. Chen, Q. Hu, Y. Wang, and W. Shieh, "Stokes Vector Direct Detection for Linear Complex Optical Channels," *J. Lightwave Technol.*, vol. 33, no. 3, pp. 678–684, Feb. 2015.
- [9] W. Shieh, C. Sun, and H. Ji, "Carrier-assisted differential detection," *Light Sci Appl*, vol. 9, no. 1, p. 18, Feb. 2020.
- [10] Z. Li, M. S. Erkılınc, K. Shi, E. Sillekens, L. Galdino, B. C. Thomsen, P. Bayvel, and R. I. Killely, "SSBI Mitigation and the Kramers–Kronig Scheme in Single-Sideband Direct-Detection Transmission With Receiver-Based Electronic Dispersion Compensation," *Journal of Lightwave Technology*, vol. 35, no. 10, pp. 1887–1893, 2017.
- [11] A. Mecozzi, C. Antonelli, and M. Shtaf, "Kramers-kronig receivers," *Adv. Opt. Photon.*, vol. 11, no. 3, pp. 480–517, Sep 2019.
- [12] M. v. d. Hout, S. van der Heide, and C. Okonkwo, "Kramers-Kronig Receiver With Digitally Added Carrier Combined With Digital Resolution Enhancer," *Journal of Lightwave Technology*, vol. 40, no. 5, pp. 1400–1406, 2022.
- [13] H. Chen, N. K. Fontaine, J. M. Gene, R. Ryf, D. T. Neilson, and G. Raybon, "Dual Polarization Full-Field Signal Waveform Reconstruction Using Intensity Only Measurements for Coherent Communications," *J. Lightwave Technol.*, vol. 38, no. 9, pp. 2587–2597, May 2020.
- [14] H. Chen, N. K. Fontaine, H. Huang, R.-J. Essiambre, M. Mazur, R. Ryf, and D. T. Neilson, "Phase Retrieval Receivers Based on Alternative Projections for Coherent Optical Communications," in *2021 European Conference on Optical Communication (ECOC)*, 2021, pp. 1–4.
- [15] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.
- [16] H. Huang, H. Chen, Y. Huang, N. K. Fontaine, R. Ryf, and Y. Song, "Carrier-less phase retrieval receiver leveraging digital upsampling," *Opt. Lett.*, vol. 45, no. 21, p. 6070, Nov. 2020.
- [17] P. Zhou, M. Xiang, W. Cheng, Y. Ma, R. Wu, G. Zhou, J. Li, C. Zhang, J. Li, S. Fu, and Y. Qin, "Weighted Decision Enabled Phase Retrieval Receiver With Adaptive Intensity Transformation," *J. Lightwave Technol.*, vol. 41, no. 19, pp. 6208–6217, Oct. 2023.
- [18] L. Blech, C. Antonelli, A. Mecozzi, Y. C. Eldar, and M. Shtaf, "Enhancing the Kramers–Kronig receiver via dispersion-based spatial diversity," *Opt. Lett.*, vol. 45, no. 13, p. 3494, Jul. 2020.
- [19] Q. Wu, Y. Zhu, and W. Hu, "Carrier-Assisted Phase Retrieval," *J. Lightwave Technol.*, vol. 40, no. 16, pp. 5583–5596, Aug. 2022.
- [20] D. Orsuti, M. Cappelletti, M. Santagiustina, A. Galtarossa, and L. Palmieri, "Edge-carrier-assisted phase retrieval receivers based on alternative projections: Performance and complexity analysis," in *press at 2023 European Conference on Optical Communication (ECOC)*, 2023, pp. 1–4.
- [21] X. Li, S. An, H. Ji, J. Li, W. Shieh, and Y. Su, "Deep-learning-enabled high-performance full-field direct detection with dispersion diversity," *Opt. Express*, vol. 30, no. 7, p. 11767, Mar. 2022.
- [22] B. Stern, H. Chen, K. Kim, and N. K. Fontaine, "Large Dispersion Silicon Bragg Grating for Full-Field 40-GBd QPSK Phase Retrieval Receiver," *Journal of Lightwave Technology*, vol. 40, no. 22, pp. 7358–7363, 2022.
- [23] D. Orsuti, C. Antonelli, A. Chiuso, M. Santagiustina, A. Mecozzi, A. Galtarossa, and L. Palmieri, "Deep Learning-Based Phase Retrieval Scheme for Minimum-Phase Signal Recovery," *Journal of Lightwave Technology*, vol. 41, no. 2, pp. 578–592, 2023.
- [24] Q. Wu, Z. Xu, Y. Zhu, Y. Zhang, H. Ji, Y. Yang, G. Qiao, L. Liu, S. Wang, J. Liang, J. Wei, J. Li, Z. He, Q. Zhuge, and W. Hu, "Machine Learning for Self-Coherent Detection Short-Reach Optical Communications," *Photonics*, vol. 10, no. 9, p. 1001, Aug. 2023.
- [25] D. Orsuti, M. Cappelletti, M. Santagiustina, A. Galtarossa, and L. Palmieri, "Edge-carrier-assisted Phase-Retrieval Based on Deep Learning Enabling low CSPR and low Applied Dispersion Values," in *2023 Optical Fiber Communications Conference and Exposition (OFC)*, San Diego, CA, USA, 2023, pp. 1–3.
- [26] G. Wang, G. B. Giannakis, and Y. C. Eldar, "Solving Systems of Random Quadratic Equations via Truncated Amplitude Flow," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 773–794, Feb. 2018, conference Name: IEEE Transactions on Information Theory.
- [27] S. T. Le, K. Schuh, M. Chagnon, F. Buchali, and H. Buelow, "1.6 Tb/s Virtual-Carrier Assisted WDM Direct Detection Transmission Over 1200 km," *J. Lightwave Technol.*, vol. 37, no. 2, pp. 418–424, Jan. 2019.
- [28] W. Yi, Z. Li, M. S. Erkılınc, D. Lavery, E. Sillekens, D. Semrau, Z. Liu, P. Bayvel, and R. I. Killely, "Performance of Kramers–Kronig Receivers in the Presence of Local Oscillator Relative Intensity Noise," *Journal of Lightwave Technology*, vol. 37, no. 13, pp. 3035–3043, 2019.
- [29] A. Mecozzi, C. Antonelli, and M. Shtaf, "Kramers-Kronig coherent receiver," *Optica*, vol. 3, no. 11, pp. 1220–1227, Nov 2016.
- [30] Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci Appl*, vol. 7, no. 2, pp. 17 141–17 141, Feb. 2018.
- [31] Y. Nishizaki, R. Horisaki, K. Kitaguchi, M. Saito, and J. Tanida, "Analysis of non-iterative phase retrieval based on machine learning," *Opt Rev*, vol. 27, no. 1, pp. 136–141, Feb. 2020.
- [32] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal Convolutional Networks for Action Segmentation and Detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, Jul. 2017, pp. 1003–1012.
- [33] A. Vedaldi and K. Lenc, "MatConvNet: Convolutional Neural Networks for MATLAB, Supplementary material: "matconvnet-manual.pdf", Chapter 5," in *Proceedings of the 23rd ACM international conference on Multimedia*. Brisbane Australia: ACM, Oct. 2015, pp. 689–692.
- [34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778.
- [35] B. Karanov, M. Chagnon, F. Thouin, T. A. Eriksson, H. Bulow, D. Lavery, P. Bayvel, and L. Schmalen, "End-to-End Deep Learning of Optical Fiber Communications," *J. Lightwave Technol.*, vol. 36, no. 20, pp. 4843–4855, Oct. 2018.
- [36] P. J. Freire, A. Napoli, B. Spinnler, N. Costa, S. K. Turitsyn, and J. E. Prilepsky, "Neural Networks-Based Equalizers for Coherent Optical

- Transmission: Caveats and Pitfalls,” *IEEE J. Select. Topics Quantum Electron.*, vol. 28, no. 4, pp. 1–23, Jul. 2022.
- [37] T. T. Nguyen, T. Zhang, E. Giacoumidis, A. A. Ali, M. Tan, P. Harper, L. P. Barry, and A. D. Ellis, “Coupled Transceiver-Fiber Nonlinearity Compensation Based on Machine Learning for Probabilistic Shaping System,” *J. Lightwave Technol.*, vol. 39, no. 2, pp. 388–399, Jan. 2021.
- [38] T. Chai and R. R. Draxler, “Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature,” *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014.
- [39] T. A. Eriksson, H. Bulow, and A. Leven, “Applying Neural Networks in Optical Communication Systems: Possible Pitfalls,” *IEEE Photon. Technol. Lett.*, vol. 29, no. 23, pp. 2091–2094, Dec. 2017.
- [40] M. O’Neill, “PCG: A family of simple fast space-efficient statistically good algorithms for random number generation,” *Harvey Mudd College Computer Science Department Technical Report*, 2014.
- [41] H. Haramoto, M. Matsumoto, T. Nishimura, F. Panneton, and P. L’Ecuyer, “Efficient jump ahead for F2-Linear random number generators,” *INFORMS Journal on Computing*, vol. 20, no. 3, pp. 385–390, 2008.
- [42] Q. Wu, Y. Zhu, Q. Zhuge, and W. Hu, “Dual-Carrier-Assisted Phase Retrieval for Polarization Division Multiplexing,” *J. Lightwave Technol.*, vol. 40, no. 22, pp. 7297–7306, Nov. 2022.
- [43] D. Orsuti, C. Antonelli, A. Chiuso, M. Santagiustina, A. Mecozzi, A. Galtarossa, and L. Palmieri, “Phase retrieval receiver based on deep learning for minimum-phase signal recovery,” in *2022 European Conference on Optical Communication (ECOC)*, 2022, pp. 1–4.
- [44] T. Bo and H. Kim, “Toward practical kramers-kronig receiver: Resampling, performance, and implementation,” *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 461–469, 2019.
- [45] C. Füllner, M. M. H. Adib, S. Wolf, J. N. Kemal, W. Freude, C. Koos, and S. Randel, “Complexity analysis of the kramers–kronig receiver,” *Journal of Lightwave Technology*, vol. 37, no. 17, pp. 4295–4307, 2019.
- [46] K. Gentile, “Digital Pulse-Shaping Filter Basics,” *Analog Devices Application Note*, no. 922, pp. 1–12, 2007.
- [47] H. Chen, N. K. Fontaine, M. Mazur, L. Dallachiesa, Y. Zhang, H. Huang, D. van Veen, V. Houtsma, A. Blanco-Redondo, R. Ryf, and D. T. Neilson, “140G/70G Direct Detection PON with >37 dB Power Budget and 40-km Reach Enabled by Colorless Phase Retrieval Full Field Recovery,” in *2021 European Conference on Optical Communication (ECOC)*. Bordeaux, France: IEEE, 2021, pp. 1–4.
- [48] P. J. Freire, Y. Osadchuk, B. Spinnler, A. Napoli, W. Schairer, N. Costa, J. E. Prilepsky, and S. K. Turitsyn, “Performance versus complexity study of neural network equalizers in coherent optical systems,” *Journal of Lightwave Technology*, vol. 39, no. 19, pp. 6085–6096, 2021.
- [49] —, “Performance Versus Complexity Study of Neural Network Equalizers in Coherent Optical Systems,” *Journal of Lightwave Technology*, vol. 39, no. 19, pp. 6085–6096, 2021.