

Article

Adaptive DBP System with Long-Term Memory for Low-Complexity and High-Robustness Fiber Nonlinearity Mitigation [†]

Mingqing Zuo ^{1,‡}, Huitong Yang ^{1,2,‡}, Yi Liu ², Zhengyang Xie ^{2,*}, Dong Wang ^{1,*}, Shan Cao ¹, Zheng Zheng ²  and Han Li ¹

¹ Department of Fundamental Network Technology, China Mobile Research Institute, Beijing 100053, China; zuomingqing@chinamobile.com (M.Z.); sxx_2230@buaa.edu.cn (H.Y.); caoshan@chinamobile.com (S.C.); lihan@chinamobile.com (H.L.)

² Department of Electronic Engineering, Beihang University, Beijing 100191, China; eeleahliu@buaa.edu.cn (Y.L.); zhengzheng@buaa.edu.cn (Z.Z.)

* Correspondence: xiezy@buaa.edu.cn (Z.X.); wangdongyjjy@chinamobile.com (D.W.)

[†] This paper is an extended version of our paper published in the 2023 Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM), Wuhan, China, 4–7 November 2023, pp. 1–4. <https://doi.org/10.1109/ACP/POEM59049.2023.10369011>.

[‡] These authors contributed equally to this work.

Abstract

Adaptive digital back-propagation (A-DBP) is a promising candidate for mitigating Kerr nonlinearity due to its ability to estimate the optimal nonlinear scaling factor adaptively. However, the adaptive process relying on the gradient-dependent algorithm is prone to fluctuation, leading to extra iterations or even divergence and resulting in huge computational efforts in A-DBP. In this paper, an improved A-DBP algorithm with long-term memory (LTM) is proposed, employing root mean square propagation (RMSProp) to achieve low-complexity and high-robustness compensation performances. The A-DBP-LTM algorithm based on RMSProp was numerically validated through the simulated transmission of 69 Gbaud DP-16QAM over 2000 km and further verified through an experiment involving 26-λ 63 Gbaud DP-16QAM transmission over 1200 km. Compared with conventional digital back-propagation and A-DBP based on a gradient-descent algorithm, our proposed method allows substantial complexity reductions of 31.35% and 58.47%, respectively. Furthermore, high robustness in only a few iterations and a 0.33 dB improvement in the optical signal–noise ratio penalty were also experimentally demonstrated.

Keywords: fiber-optic communication; nonlinear compensation; digital back-propagation (DBP)



Received: 11 June 2025

Revised: 1 July 2025

Accepted: 4 July 2025

Published: 11 July 2025

Citation: Zuo, M.; Yang, H.; Liu, Y.; Xie, Z.; Wang, D.; Cao, S.; Zheng, Z.; Li, H. Adaptive DBP System with Long-Term Memory for Low-Complexity and High-Robustness Fiber Nonlinearity Mitigation. *Photonics* **2025**, *12*, 704. <https://doi.org/10.3390/photonics12070704>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the information explosion caused by high-definition video streaming, cloud computing, artificial intelligence, and so forth, global communication traffic has grown exponentially [1]. In order to sustain such dramatic traffic growth, coherent detection techniques in association with digital signal processing (DSP) techniques have been developed for optical fiber communication systems [2–4]. Today, to further increase the capacity of long-haul optical fiber communication, several effective nonlinear compensation algorithms can be used. In the optical domain, spectral inversion in the middle of the link using optical-phase conjugation (OPC) can effectively compensate for linear and nonlinear effects,

but additional optical devices and the design of symmetric optical communication links are required [5]. In the time domain, digital back-propagation (DBP) [6–17], Volterra series nonlinear equalizers (VSNEs) [18], perturbation theory [19], and neural network (NN) technology [20–25] have proved to be effective. However, a VSNE requires Fourier transform modules, and the complexity increases with the accumulation of chromatic dispersion. As for nonlinear compensation based on perturbation theory, to achieve desirable quantization accuracy, higher computational complexity is required [23]. With the rapid development of machine learning, the powerful learning ability of NNs has attracted attention. Artificial neural networks (ANNs), convolutional neural networks (CNNs), recurrent neural networks (RNNs), etc. have been introduced into fiber nonlinearity compensation strategies to further improve the performance of optical communication systems [22]. However, most of the nonlinear compensation methods based on NNs are “black box” approaches that rely only on received data to imitate the transmission model and are difficult to explain [23].

Of all the solutions presented above, DBP is the most widely used technique for the compensation of linear and nonlinear effects. It is highly effective and has a simple structure. The implementation of DBP entails solving the inverse nonlinear Schrödinger equation (NLSE) [6] using the split-step Fourier method (SSFM). In order to address the interpretability of “black box” NNs in regard to nonlinear compensation, researchers have combined theoretical models with NNs and then proposed interpretable learned DBP (LDBP) [23–25]. However, LDBP relies on received data, making the ability to adapt to different system models a key obstacle in the field of machine learning. Whether LDBP or conventional DBP, the compensation performance and processing speed of both are heavily dependent on the steps per span used in the SSFM [26]. Meanwhile, when we focus on the conventional DBP approach, we find that an inaccurate nonlinear scaling factor $k = \xi\gamma$ leads to poor accuracy in the nonlinear compensation, in which ξ represents the nonlinear compensation factor and γ represents the nonlinearity coefficient of the transmission link. This issue poses another significant challenge in regard to conventional DBP. For γ , it is difficult to obtain an accurate value in a dynamic meshed optical network and its value may vary in reconfigurable systems [13]. Also, there is no analytical method for acquiring the optimal value of ξ , which controls the optimum nonlinear compensation [6].

Hence, blind estimation of the nonlinear parameters in conventional DBP has been proposed to overcome these drawbacks, including via novel methods based on an extra filter [8,13], perturbation theory [19], the adjoint-based optimization (ABO) algorithm [17], and gradient-dependent algorithms [9,10,12,13,16,27–29]. Adaptive digital back-propagation (A-DBP) based on gradient-dependent algorithms is a desirable solution. This approach involves estimating the optimal nonlinear scaling factor k to enhance compensation performance and reducing computational complexity by employing fewer steps from the SSFM while maintaining the same effectiveness. A-DBP, which was first introduced in [9], employs the phase-noise variance as the cost function (CF) to search for the optimal γ by using the gradient-descent algorithm (GDA). After the development of this approach, several improved A-DBP methods were developed based on adjusting the CF [28], the nonlinear scaling factor k [10,12,13], and an additional filter [29], but they were all implemented using a GDA. Gradient-dependent algorithms in A-DBP control the convergence process and are strongly associated with the iterations, which show positive correlations with computational efforts. When facing changes in compensation parameters, the GDA converges slowly or is prone to fluctuation. A GDA with momentum (GDAM) was applied in A-DBP in [16] to improve convergence speed by considering the previous variation trend. However, the momentum term can intensify the fluctuations. The adaptation of the learning rate μ via a GDA was added to the GDAM to achieve faster convergence in [27]; however, fluctuations still existed and they may be worse because of

the inserted GDA. By introducing a predictive term, the fluctuation during iterations can be suppressed [30], but this method has no advantages in computational complexity due to the introduction of additional feedback.

In this paper, we extend our previous work [31] and propose an A-DBP algorithm with long-term memory (LTM) to simultaneously achieve low-complexity and high-robustness nonlinearity compensation, in which the learning rate μ is adaptively adjusted based on all of the gradients of previous iterations. The performances of classical gradient-dependent algorithms with such properties, including root mean square propagation (RMSProp), AdaGrad, Adam, and AdaDelta, are numerically compared, showing that RMSProp has the best performance. Afterwards, a proof-of-concept numerical simulation and experiment are carried out based on RMSProp to demonstrate the feasibility of the proposed A-DBP-LTM method. In the simulated transmission of 69 Gbaud DP-16QAM over 2000 km, our proposed method demonstrates strong robustness to hyperparameters by adapting μ , and it performs better with same the iterations compared to A-DBP based on a GDA and A-DBP based on a GDAM. Moreover, it outperforms the conventional DBP in the scenarios of biased γ , with the same steps per span and reconfigurable systems. It should be noted that the reductions of 31.35%, 58.47%, and 69.62% of complex multiplications (CMs) under the same performance are achieved in A-DBP-LTM based on RMSProp compared to conventional DBP, A-DBP based on a GDA, and A-DBP based on a GDAM, respectively. In order to further verify the effectiveness of this scheme through experiments, we demonstrate a 26- λ , 63 Gbaud, DP-16QAM coherent transmission system over 1200 km. Compared with A-DBP based on a GDA and A-DBP based on a GDAM, high robustness in a few iterations in the scenarios of different launching powers, different channels, or different spans and a 0.33 dB improvement of the optical signal–noise ratio (OSNR) penalty are demonstrated.

2. Principle

The concept of DBP is that the signal at the receiver is back-propagated through virtual fibers with inverse characteristics compared with the transmission fibers, thereby compensating the distortions caused by chromatic dispersion and nonlinearity. Theoretically, polarization-multiplexed optical signals in a single-mode fiber can be analytically described by the inverse Manakov equation [7] and then the equation can be numerically solved by using the SSFM [32]. The Wiener–Hammerstein model of the SSFM is accomplished by applying the dispersion operator in two equal parts, before and after the nonlinear operator, which means that the nonlinear operator is executed in the middle of each step [33], as

$$A_{x,y}(z+h,t) = \exp\left(\frac{h}{2}\hat{D}\right) \exp(h_{\text{eff}}\hat{N}) \exp\left(\frac{h}{2}\hat{D}\right) A_{x,y}(z,t) \quad (1)$$

where h_{eff} represents the effective step size involved with step size h [32], and the linear operator and the nonlinear operator are given as:

$$\hat{D} = \frac{\alpha}{2} A_{x,y} + j \frac{\beta_2}{2} \frac{\partial^2 A_{x,y}}{\partial t^2} \quad (2)$$

$$\hat{N} = -j \frac{8}{9} \zeta \gamma \left(|A_x|^2 + |A_y|^2 \right) \quad (3)$$

where the nonlinear compensation factor ζ is a free parameter that controls the optimum compensation of the nonlinear effect [14]. The nonlinear scaling factor k in the nonlinear operator is defined as:

$$k = \zeta \gamma \quad (4)$$

However, the inaccurate estimation of k in the conventional DBP degrades nonlinear compensation performance and leads to the extra complexity caused by repeated steps of the SSFM. Hence, A-DBP is required to search for the optimal k using gradient-dependent algorithms to achieve the best nonlinear compensation performance and reduce the complexity with fewer SSFM steps. As the widely used algorithm in A-DBP, the GDA suffers from slow convergence or excessive fluctuations when compensation parameters vary. A momentum term is added in the GDAM to accelerate convergence, but it brings more intensive fluctuations. Neither of these two algorithms can always maintain a small number of iterations in practical systems, which results in extra computational efforts to apply them practically. Therefore, for gradient-dependent algorithms employed in A-DBP, fast convergence and fluctuation resistance are essential requirements.

In order to achieve low-complexity and high-robustness nonlinear compensation based on A-DBP, we propose the A-DBP-LTM algorithm, where the learning rate μ is adaptively adjusted. The main purpose is that, at the beginning of the iteration approaching the convergence goal fast and jumping out of the local optimum and at the end of the iteration avoiding crossing the convergence goal, fluctuations are avoided. There are some improved gradient-dependent algorithms with the learning rate μ adaptively adjusted, which can be used for implementing the A-DBP-LTM scheme, such as RMSProp, Adam, AdaDelta, and AdaGrad. Here, taking RMSProp as an example, we theoretically demonstrate its advantages over the GDA and GDAM. The iterative formulas of RMSProp, the GDA, and the GDAM can be expressed as [16,34]:

$$k(i+1) = k(i) - \frac{\mu}{\sqrt{F[\nabla\text{CF}(k(i))^2] + \varepsilon}} \cdot \nabla\text{CF}(k(i)) \quad (5)$$

$$k(i+1) = k(i) - \mu \cdot \nabla\text{CF}(k(i)) \quad (6)$$

$$k(i+1) = k(i) - mk(i-1) - \mu \cdot \nabla\text{CF}(k(i)) \quad (7)$$

where $k(i)$ represents the nonlinear scaling factor at the iteration of i , the learning rate μ controls the adaptation speed, ε is set as a tiny value to prevent the root from being 0, CF is error vector magnitude (EVM), m is a parameter to remember the last gradient, $\nabla\text{CF}(k(i))$ is the gradient at the i -th iteration defined as:

$$\nabla\text{CF}(k(i)) = [\text{CF}(k(i)) - \text{CF}(k(i-1))]/[k(i) - k(i-1)] \quad (8)$$

The content inside the root is the exponentially weighted moving average (EWMA) of all of the previous gradients' square and is defined as:

$$F[\nabla\text{CF}(k(i))^2] = p \times F[\nabla\text{CF}(k(i-1))^2] + (1-p) \times \nabla\text{CF}(k(i))^2 \quad (9)$$

where the weighting coefficient p controls the degree of acquiring the past information of the CF's gradient. Because of LTM of the previous CF's gradient information, the A-DBP algorithm with LTM will update learning rate μ adaptively to promote compensation.

According to Equations (6) and (7), we can see that the GDA keeps the same learning rate μ to iterate; and the GDAM accelerates the convergence of the GDA by introducing the momentum term m , but when initial m is large and the update term $k(i+1)$ exceeds the optimal point, it will make it difficult to change the gradient in the next step, and then the parameter k will be updated from the current moment to the optimal point. The process will fluctuate greatly. However, in Equation (5), A-DBP-LTM based on RMSProp adaptively adjusts the learning rate μ , employing the root of the EWMA for all of the square of the previous gradients as shown in Equation (9), which is in the denominator. Hence,

μ diminishes to avoid fluctuations caused by an excessively fast descent of CF when the value of the gradient is large. In turn, μ increases to accelerate the descent of CF when the value of the gradient is small. The adaptation of μ in RMSProp-based A-DBP-LTM ensures a fast and steady convergence and makes it insensitive to changes of the compensation parameters. Therefore, our proposed A-DBP-LTM based on RMSProp is able to achieve low-complexity and high-robustness nonlinear compensation.

Figure 1 depicts the implementation flow chart of our proposed A-DBP. The dispersion compensation (DC) and the nonlinear compensation (NLC) of the SSFM are usually performed in the frequency domain (FD) and the time domain (TD). Therefore, FFT and IFFT are employed to switch signals between the FD and TD. The fiber is divided into N_{step} steps with step size h , where DC and NLC are iteratively applied during back-propagation. In addition, the overlap–save method is typically employed as the FFT strategy to divide long samples into blocks, enabling continuous-time implementation [35]. The compensation loop begins with the initial nonlinear scaling factor $k(i)$, where i is the iteration index and is set to 1 at the start. Then, the input signal is processed by the SSFM and EVM is calculated as CF after compensation. The update of $k(i + 1)$ is repeated adaptively according to Equation (5) for the new iteration loop, aiming for the descent of CF until it is minimized (i.e., the optimal k is found), which means the best compensation performance is achieved.

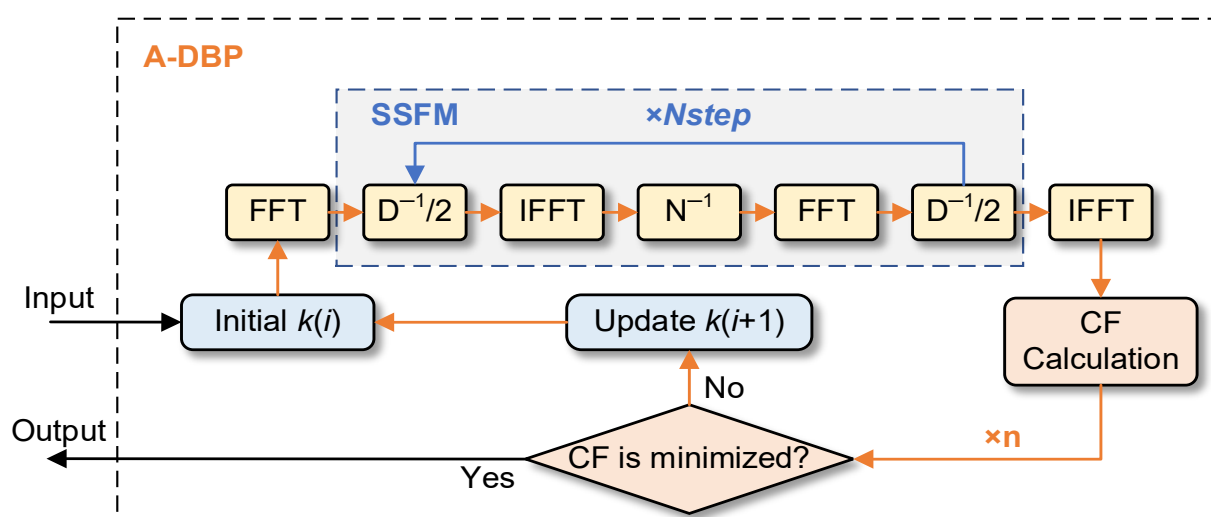


Figure 1. The implementation flow chart of proposed A-DBP.

3. Results

3.1. Simulation System and Result Analysis

3.1.1. Description of the Simulation Setup

Figure 2 illustrates the DP-16QAM simulation setup constructed using VPIphotonics 10.1 software. At the transmitter, the generated pseudo-random binary sequence (PRBS) undergoes the quadrature modulation and then digital-to-analog conversion (DAC) to obtain 69 Gbaud DP-16QAM electrical signals. The laser at 1550 nm with a 200 kHz line-width is subsequently modulated by an I/Q modulator (one per polarization component), which consists of two nested Mach–Zehnder modulators to accomplish electro-optic conversion. The optical signals are then transmitted over a link comprising 20 spans, each consisting of a 100 km standard single-mode fiber (SSMF) followed by an erbium-doped fiber amplifier (EDFA) for amplification, with noise at 5 dB. Some parameters related to the fiber are as follows: the chromatic dispersion coefficient D is 16.75 ps/nm/km, the fiber attenuation α is 0.2 dB/km, and the nonlinearity coefficient γ is 1.31 /W/km. Then, 16,384 symbols of each polarization are transmitted, with each symbol simulated using 8 samples.

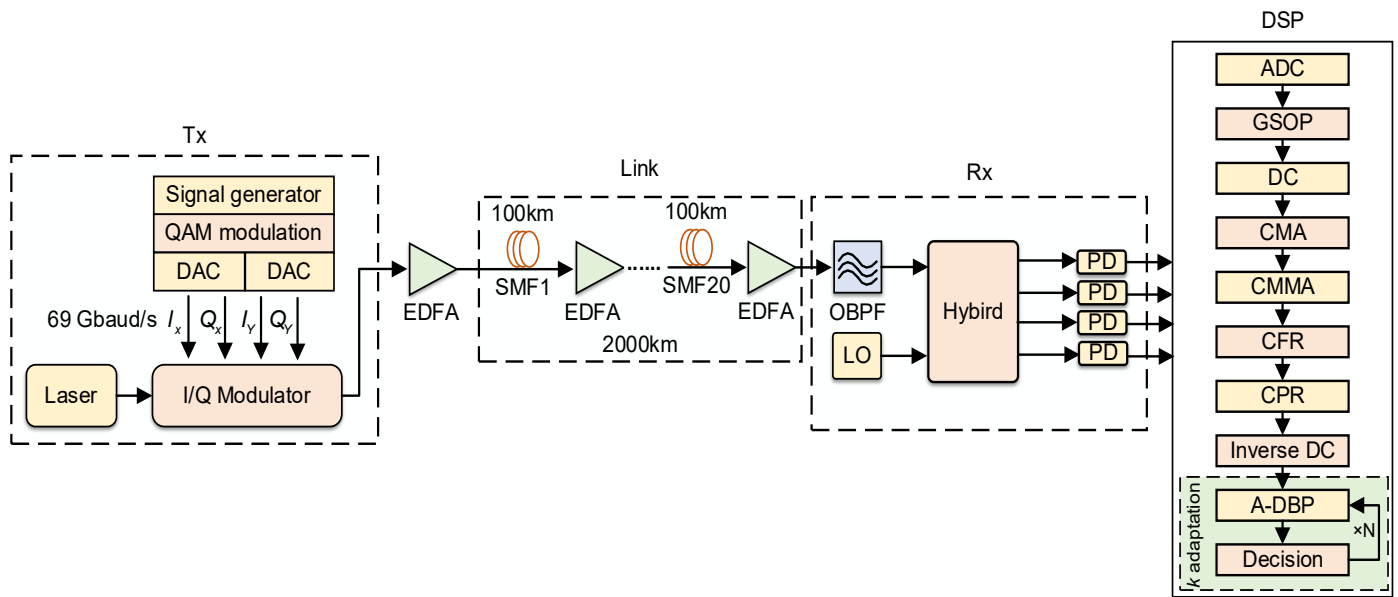


Figure 2. The simulation system setup. DAC: digital-to-analog converter. EDFA: erbium-doped fiber amplifier. SMF: single-mode fiber. OBPF: optical band-pass filter. LO: local oscillator. PD: photodetector. ADC: analog-to-digital converter. GSOP: Gram–Schmidt orthogonalization procedure. DC: dispersion compensation. CMA: constant module algorithm. CMMA: cascaded multimodule algorithm. CFR: carrier frequency recovery. CPR: carrier phase recovery. Inverse DC: inverse dispersion compensation. A-DBP: adaptive digital back-propagation.

After transmission, optical signals traverse an optical band-pass filter (OBPF) and are combined with a local oscillator (LO) at the polarization diversity hybrid to acquire I and Q optical signals of two polarizations. Subsequently, photodetectors (PDs) effectuate photoelectric conversion, and resultant electrical signals are processed in the DSP module implemented using MATLAB 2023b software. The sequence of the subsequent DSP process is shown in Figure 2. Specifically, the algorithm used for carrier frequency recovery is quadratic frequency bias estimation and the algorithm used for carrier phase recovery is a partitioned Viterbi–Viterbi algorithm. It is noted that only two samples per symbol are taken for the DSP module. Due to the different clock sources of AWG and oscilloscope, two samples per symbol are needed for achieving the time synchronization. Meanwhile, the CMA is employed for pre-convergence of the CMMA. During CMA processing in the simulation, we use 11 taps. Additionally, different from the DSP process in [30], we choose to use EVM as CF because there is no predictive term introduced in gradient-dependent algorithms.

In order to verify the applicability of the proposed algorithm in the links with different types of fibers, two kinds of heterogeneous links are set up for simulation. In Figure 3a, the first hybrid links consist of 10 spans of SSFMs with 100 km per span and 10 spans of SSFMs with 80 km per span. The EDFA gain in the 100 km part and the 80 km part are 20 dB and 16 dB, respectively. In Figure 3b, the second hybrid links are composed of 10 spans of SSFMs and 10 spans of large effective area fibers (LEAFs) with the same length of 100 km per span and the same EDFA gain of 20 dB. The γ of the SSFM and LEAF are $1.31/W/\text{km}$ and $0.75/W/\text{km}$, respectively.

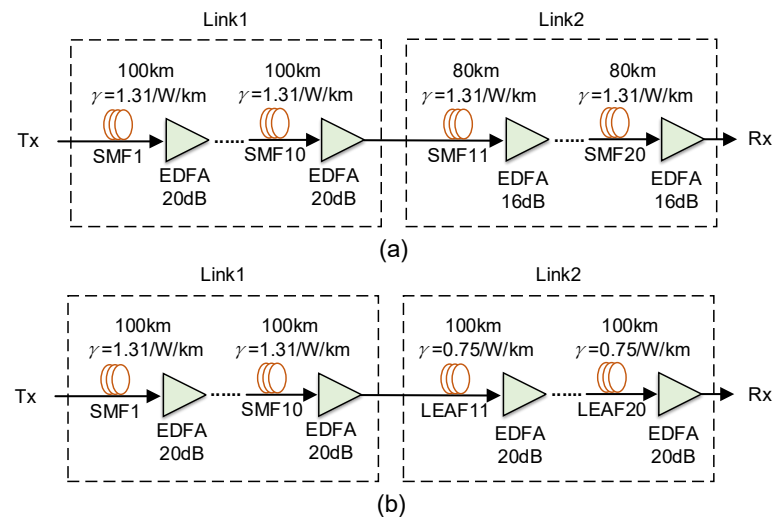


Figure 3. The heterogeneous links of length and nonlinear coefficient in the simulation: (a) the first hybrid link A; (b) the second hybrid link B.

3.1.2. Analysis and Discussion of Simulation Results

Figure 4 compares EVM performances of different algorithms versus the number of iterations in A-DBP with a five-step SSFM at 5 dBm launching power. We initialize $\gamma = 3 \text{ /W/km}$ for all algorithms and set $\mu = 1.1 \times 10^{-5}$, 8×10^{-6} , 1×10^{-5} , 2.2×10^{-4} for the GDA, GDAM, ANAGDA, A-DBP-LTM, respectively. A-DBP-LTM is implemented by RMSProp, AdaGrad, Adam, or AdaDelta. We can see that the CF of the GDA has fallen into a local minimum, while that of the GDAM fluctuates violently because of overdependence on the previous gradient. For ANAGDA, the introduction of the predictive term doubles the number of iterations, which prevents the low-complexity nonlinearity compensation. Among these algorithms, the A-DBP-LTM implemented by RMSProp simultaneously indicates the best compensation performance and the fastest and the most stable convergence process even compared with AdaGrad, Adam, and AdaDelta. Thus, we choose RMSProp to carry out the later numerical simulation and experiments to further validate the effectiveness of our proposed A-DBP-LTM method.

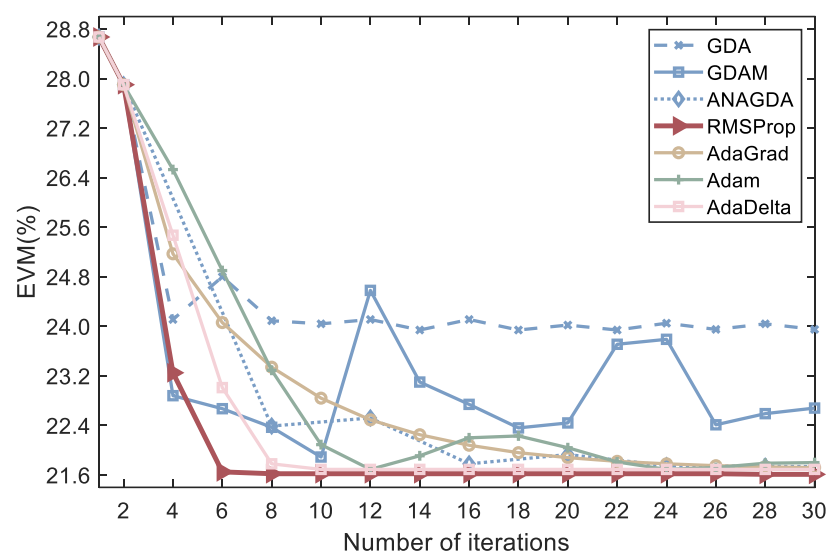


Figure 4. EVM versus number of iterations for different A-DBP with 5-step SSFM at 5 dBm launching power.

Figure 5 depicts the number of iterations for different A-DBP methods under variations of $\Delta\mu$ and k at the optimal launching power with five SSFM steps. To prevent fluctuations and ensure good convergence, the initial k is set within $[2 \times 10^{-3}, 4 \times 10^{-3}]$, exceeding the transmission link's default value. Based on the optimal μ and m (for the GDAM) at the middle value of the range of k , μ is swept across $[-1 \times 10^{-5}, 1 \times 10^{-5}]$, with $\Delta\mu$ denoting the change of μ . It is noted that the iteration process converges when $\Delta\mu < 0$, and a number of iterations greater than 20, shown as the same color as 20 in Figure 5, is caused by the low convergence speed. For $\Delta\mu > 0$, fluctuations cause more iterations or divergence, also shaded as 20 in Figure 5. Figure 5b shows the number of iterations of A-DBP based on the GDAM is reduced compared to A-DBP based on the GDA shown in Figure 5a because of the additional momentum term when $\Delta\mu < 0$. However, when $\Delta\mu > 0$, the momentum term m intensifies fluctuations, leading to more iterations or divergence compared to A-DBP based on the GDA. Note that the performance of A-DBP based on the GDAM relies on m and fewer iterations are realizable by adjusting m . But it cannot show the superiority owing to the hyperparameter m . Conversely, our proposed method demonstrates strong robustness on hyperparameters by adapting μ and consistently maintains a low number of iterations compared to A-DBP based on the GDA and A-DBP based on the GDAM.

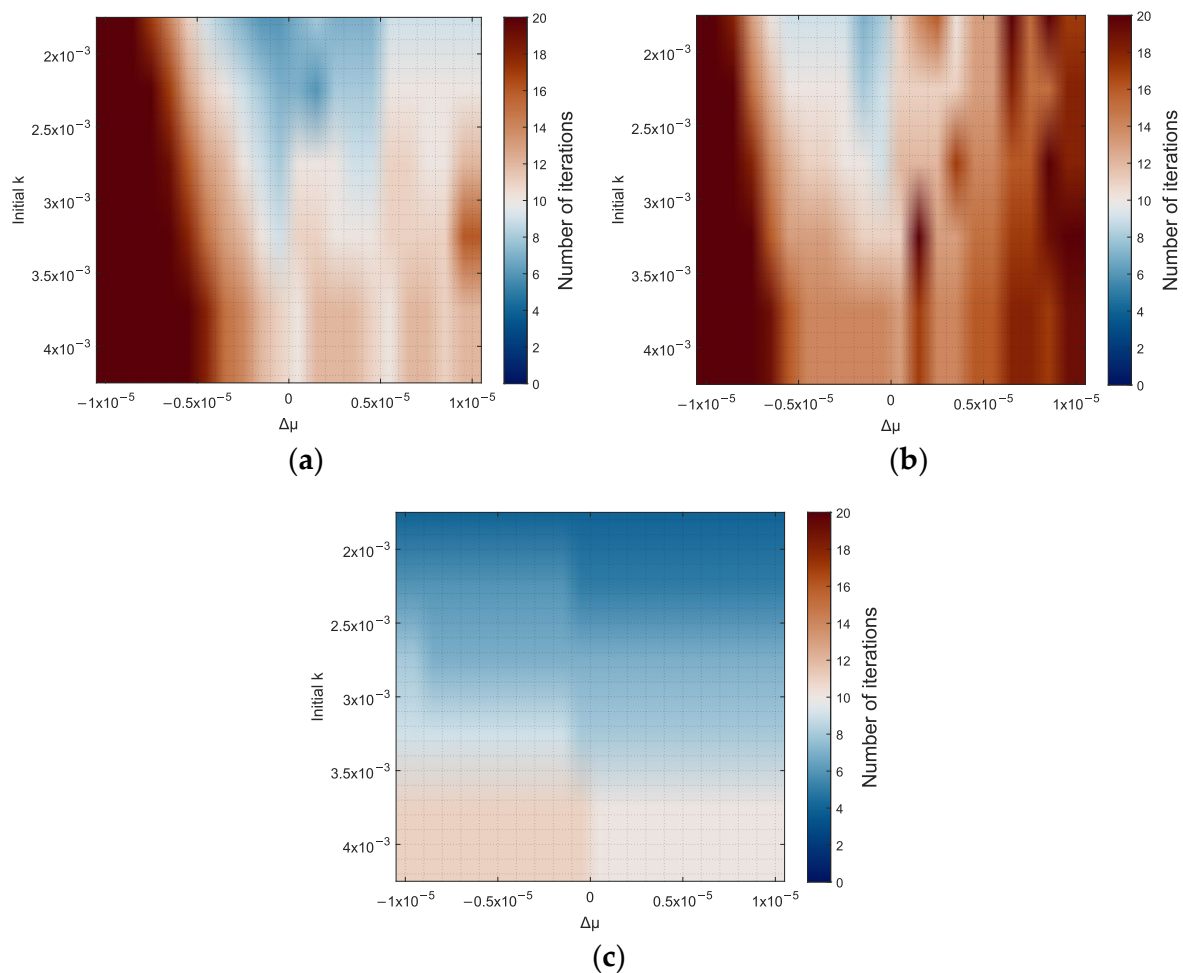


Figure 5. The number of iterations under different changes of $\Delta\mu$ and initial k using (a) A-DBP based on GDA; (b) A-DBP based on GDAM; (c) the proposed method at the best launching.

In order to verify the compensation performance in the situations of deviated γ , the conventional DBP and the proposed method are implemented by employing different γ (0.6, 1.31, 2/W/km). It should be noted that, because the nonlinear coefficient is fixed

as $1.31/\text{W}/\text{km}$ for each fiber span in the simulation system, the performance of the conventional DBP algorithm has been optimized with γ of $1.31/\text{W}/\text{km}$, which ensures a fair comparison with the proposed A-DBP-LTM implemented by RMSProp. The FEC threshold is 3×10^{-2} . Figure 6 compares the power-dependent $\log_{10}(\text{BER})$ performance of linear equalizer (LE), the conventional DBP, and the proposed method, all with five SSFM steps. The conventional DBP deteriorates significantly with inaccurate γ values and may underperform LE at certain launching powers. In contrast, our proposed method delivers reliable performance under sufficient convergence across different γ values. It outperforms the conventional DBP with accurate γ , as it optimizes k dynamically for varying launch powers, unlike the fixed k in the conventional DBP. It is difficult to compensate the Gordon–Mollenauer noise arising from the interaction between the nonlinearity effect and the amplified spontaneous emission (ASE) noise [17] employing an abnormal nonlinear scaling factor. The compensation performance gap narrows in higher launching powers because the Gordon–Mollenauer noise does not dominate in this situation, as the nonlinearity effect is not dominant.

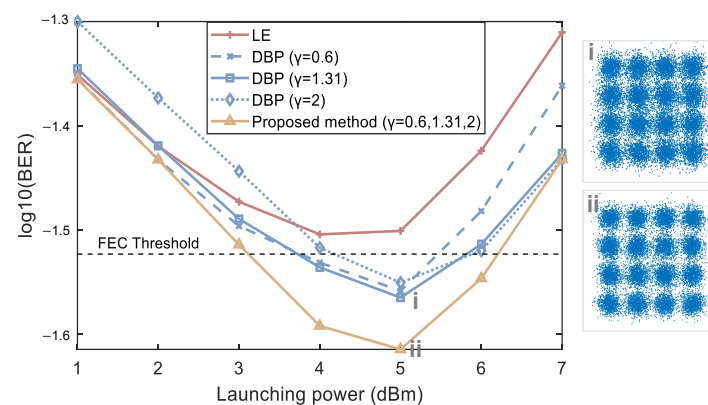


Figure 6. $\log_{10}(\text{BER})$ vs. the launching power for LE, the conventional DBP, and the proposed method with different values of γ with SSFM steps of 5.

In Figure 7, $\log_{10}(\text{BER})$ results as a function of the launching power for different numbers of steps per span. It is obvious that the proposed method takes fewer SSFM steps than the conventional DBP. There is only one step for our proposed method to reach the FEC threshold, while the conventional DBP takes six steps. It is worth noting that the increase in the compensation performance is weakened with the augmentation of SSFM steps, which signifies that the same complexity effort yields lower effectiveness. It is hence necessary to balance the complexity resulting from the increase in SSFM steps with the improvement of the compensation performance. As shown in Figure 8, our proposed method achieves better effectiveness than A-DBP based on the GDA and A-DBP based on the GDAM at four iterations with five SSFM steps, even when using their best hyperparameters. This is attributed to the fast convergence of our proposed method. It should be noted that, due to the slow convergence speeds of the GDA and GDAM, the updated γ is still far away from optimum values when RMSProp has completed convergence. Consequently, the additional nonlinear impairment to the signal would be introduced during the processing of the SSFM because of the nonideal γ , resulting in performances of the GDA and GDAM being even worse than that of the case with linear equalization only. It can be seen that different initializations of γ , μ , m affect the compensation performance of the algorithm, which consequently leads to a different result for compensation with four iterations in Figure 4 compared to Figure 8.

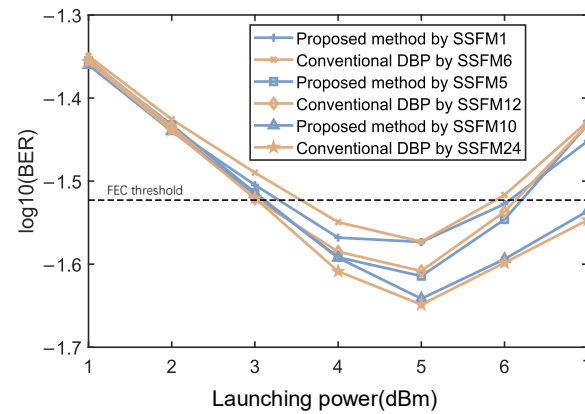


Figure 7. $\log_{10}(\text{BER})$ as the function of launching power for different steps of SSFM per span employing the conventional DBP and the proposed method.

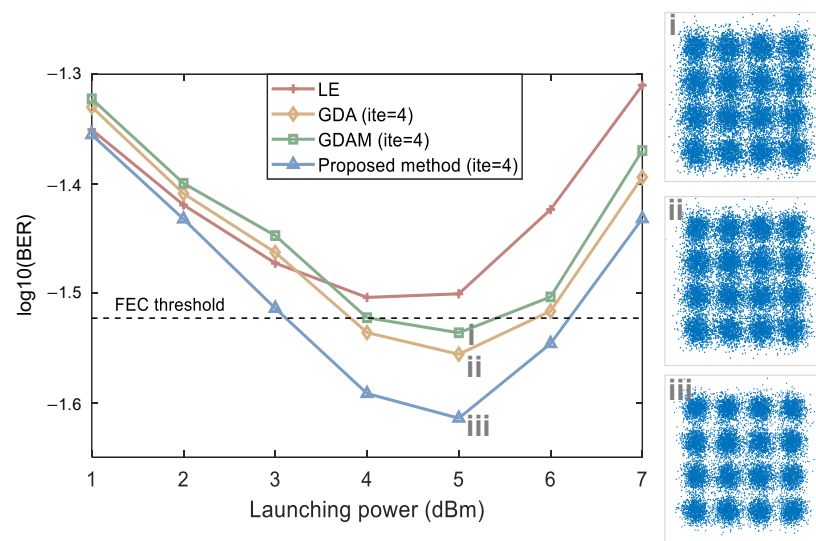


Figure 8. $\log_{10}(\text{BER})$ as the function of launching power for LE, A-DBP based on GDA, A-DBP based on GDAM and the proposed method at the same iteration of 4.

For further investigating the feasibility of our proposed method, LE, the conventional DBP, and the proposed method are implemented on the heterogeneous links which are described in Section 3.1.1 and are shown in Figure 3. Figures 9a and 9b show $\log_{10}(\text{BER})$ as functions of sweeping launching powers in the link of heterogeneous fiber lengths and types, respectively, with five SSFM steps. Compared with Link B, Link A has a shorter span length of 80 km from the 11th to 20th span, which results in a loss reduction of 4 dB. Thus, the performance of Link A is better than that of Link B. In the link of heterogeneous fiber lengths, ζ should be adjusted according to the length and ζ is smaller when being applied in shorter spans (ζ_2) with lower nonlinearity effect than in longer spans (ζ_1). The conventional DBP employing ζ_1 for all spans results in overcompensation and is worse than employing ζ_2 which is undercompensation in some launching powers. The ideal situation of the conventional DBP with ζ_1 and ζ_2 for different spans also performs worse than our proposed method for the reason that the Gordon–Mollenauer noise cannot be compensated well by the inaccurate k . In the link of heterogeneous fiber types, the conventional DBP employing $\gamma = 1.31/\text{W}/\text{km}$ for all spans leads to overcompensation and is worse than employing $\gamma = 0.6/\text{W}/\text{km}$ which is undercompensation. Our proposed method has better effectiveness than the ideal situation of the conventional DBP with $\gamma = 1.31/\text{W}/\text{km}$ and $\gamma = 0.6/\text{W}/\text{km}$ for different spans for the same reason.

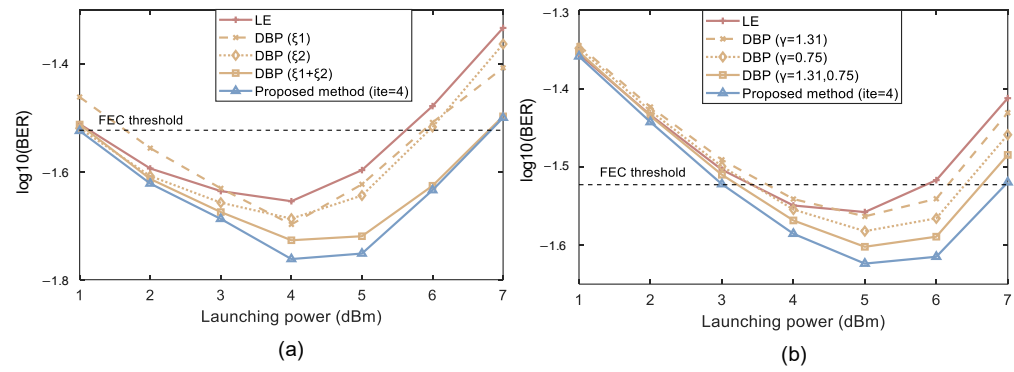


Figure 9. $\log_{10}(\text{BER})$ as the function of launching power for LE, the conventional DBP, and the proposed method in heterogeneous link of (a): fiber length, (b): fiber type.

3.1.3. Computational Effort Analysis

In this section, computational efforts of the conventional DBP and A-DBP are analyzed by complex multiplication (CM) per transmitted bit. Both of these two schemes employ a symmetric SSFM to compensate by linear and nonlinear steps. The linear step is implemented in the FD, using the overlap–save method as the FFT strategy. The nonlinear step is implemented in the TD, which needs IFFT to achieve the conversion from the FD to the TD. The CMs of the SSFM per overlap block can be obtained as [36,37]:

$$N_{\text{step}}(2M_{\text{FFT}} + 2M_D + M_N) \quad (10)$$

where N_{step} is the number of SSFM steps, M_{FFT} is the number of CMs for each FFT/IFFT and is calculated as $M_{\text{FFT}} = N_{\text{FFT}}/2\log_2(N_{\text{FFT}})$ CMs considering a standard Cooley–Tukey radix-2 implementation, in which N_{FFT} is the block size of FFT employing the overlap–save method. M_D and M_N represent the CMs for evaluating every linear and nonlinear step of each SSFM step, respectively. M_D equals N_{FFT} CMs per linear step. As for the part inside the exponential operator of the nonlinear step, $2N_{\text{FFT}}$ real multiplications (RMs) are needed for the power operator and N_{FFT} RMs are required for the multiplication of the real nonlinear constant. There are $6N_{\text{FFT}}$ RMs ($3/2 N_{\text{FFT}}$ CMs) as the total cost for solving the exponential operator by a 4th order Taylor expansion [38]. N_{FFT} CMs are needed for the multiplication of $A_{x,y}(z,t)$ outside the exponential operator which is described by the inverse Manakov equation. Hence, M_N equals $5/2 N_{\text{FFT}}$ CMs per nonlinear step. In addition, the factor of 2 associated with M_{FFT} is for the FFT/IFFT pair in one SSFM step and the factor of 2 connected with M_D is for two linear steps in each SSFM step.

Considering the whole iterations of n , extra FFT and IFFT located before and after SSFM should be considered as shown in Figure 1. The CMs of DBP/A-DBP per overlap block can be extended from Equation (10) as:

$$n \times [N_{\text{step}}(2M_{\text{FFT}} + 2M_D + M_N) + 2M_{\text{FFT}}] \quad (11)$$

When Equation (11) is applied to DBP, $n = 1$. In A-DBP, n represents the times of iterations by gradient-descent algorithms. Due to the memory length N_M of DC in the overlap–save method which is related to the channel impulse response [36], only $N_{\text{FFT}} - N_M + 1$ samples are useful in each overlap block. The useful bits per overlap block are given as [36]:

$$\frac{N_{\text{FFT}} - N_M + 1}{n_{\text{sc}}} \times \log_2(M) \quad (12)$$

where M is the order of the modulation format, n_{sc} is the oversampling ratio. It is noted that there is only one calculation of the iterative formula for plentiful samples, so it is

tiny enough to be ignored since few real multiplications generated from that formula are averaged into every sample. Therefore, the complexity N of conventional DBP and A-DBP equals the following expression:

$$N = \frac{n \times n_{sc} \times [N_{step}(2M_{FFT} + 2M_D + M_N) + 2M_{FFT}]}{(N_{FFT} - N_M + 1) \times \log_2(M)} \quad (13)$$

$$= \frac{n \times n_{sc} \times N_{FFT} \times [N_{step}(\log_2(N_{FFT}) + \frac{9}{2}) + \log_2(N_{FFT})]}{(N_{FFT} - N_M + 1) \times \log_2(M)}$$

We assess complexity efforts of the conventional DBP and our proposed method with six and one SSFM steps per span, respectively, based on the situation shown in Figure 7, showing that both methods reach the FEC threshold. Table 1 depicts the number of complex multiplications per transmitted bit calculated by (13). The reduction gain G_R which is defined as $G_R = (1 - O/O_{ref})$, in percentage, is applied for comparing computational efforts with different methods, where O_{ref} and O are the number of CMs required by the proposed method and other methods, respectively. It is shown in Table 2 that our proposed method provides a reduction gain of 31.5% in comparison with conventional DBP with the same effectiveness. In order to compare our proposed method with GDA-based A-DBP and GDAM-based A-DBP, the relative step s is introduced to deal with the situations of divergence. Considering the fastest convergence by initial $k = 2 \times 10^{-3}$ in Figure 5, s is defined as $s = 1/(1/ite_1 + ite_2 + \dots + ite_n)$, in which ite_n represents different numbers of iterations employing different μ . Table 3 shows relative steps of three methods. The reduction gain is calculated based on G_R due to the positive correlation between the number of iterations and complexity. Our proposed method achieves reduction gains of 58.7% and 69.8% compared to A-DBP based on the GDA and A-DBP based on the GDAM, respectively.

Table 1. Number of complex multiplications per transmitted bit per polarization for conventional DBP and the proposed method.

	Conventional DBP	Proposed Method
N_{FFT}	512	512
N_M	128	128
N_{step}	6×20	1×20
n_{sc}	2	2
n	1	4
N	1083	742

Table 2. Computational effort of conventional DBP versus the proposed method.

	Conventional DBP	Proposed Method
N	1083	742
G_R	31.5%	—

Table 3. Computational effort of GDA-based and GDAM-based A-DBP versus the proposed method.

	GDA-Based A-DBP	GDAM-Based A-DBP	Proposed Method
s	0.46	0.63	0.19
G_R	58.7%	69.8%	—

3.2. Experimental System and Result Analysis

3.2.1. Description of the Experimental System

Figure 10 shows the setup used in experiments. We set up a system for 26-WDM, 63 Gbaud, 16QAM signal transmission. The 63 Gbaud DP-16QAM electrical signal is generated by an arbitrary waveform generator (AWG, M8194A, Keysight, Colorado Springs, CO,

USA) operating at 120 GSa/s. To match with the sampling rate, 40-times up-sampling and 21-times down-sampling are adopted. The sampling rate of the oscilloscope is 128 GSa/s. We use an external cavity laser (ECL, IQS-2800, EXFO, Quebec City, QC, Canada) that has a central wavelength of 1550.33 nm and a linewidth of 100 kHz. The 26-WDM signals ranged from 1560.1638 nm to 1530.9051 nm with 75 GHz spacing and launched into a loop consisting of 15 spans of 80 km SSMF with a loss of 0.22 dB/km after multiplexing in a wavelength selective switch (WSS, Waveshaper 4000s, Finisar, CA, USA). And the chromatic dispersion coefficient D is 16.4 ps/nm/km, the nonlinearity coefficient γ is 1.31/W/km. We use a WSS as a gain equalizer (GEQ) in the loop and use EDFA (AEDFA-23, Amonics, Hong Kong, China) with a gain of 18.5 dB in each 80 km span to compensate for the transmission loss. Finally, the signal is detected by four balanced photodetectors (BPDs) to achieve photoelectric conversion. As shown in Figure 10, components of the off-line DSP are the same as in Figure 2.

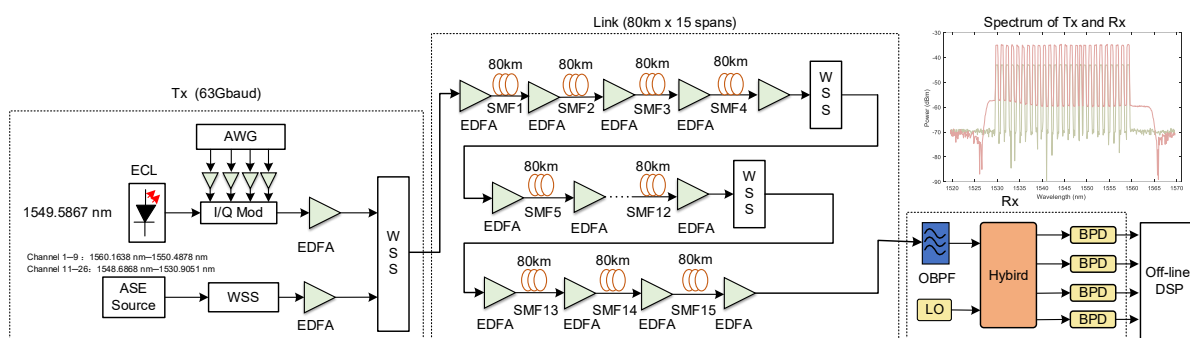


Figure 10. Experimental setup for 26-WDM, 63 Gbaud, 16QAM transmission over 1200km.

3.2.2. Analysis of Experimental Results

Figure 11a illustrates the comparison of the bit error rate (BER) for three algorithms under different input fiber powers, given the same number of SSFM steps (i.e., five steps) and iterations (i.e., four times). The complexity of these algorithms is nearly proportional to the number of iterations, so we compare their compensation performance under similar computational complexity. To ensure a fair comparison, initial values for the nonlinear scaling factor k and the hyperparameter learning rate μ are chosen such that none of the algorithms experience oscillations at the optimal input power. The experimental results demonstrate that, with a limited number of iterations, our proposed method outperforms the GDA and GDAM in terms of compensation performance. The A-DBP-LTM based on RMSProp benefits from an adaptive adjustment of the learning rate μ , requiring only four iterations to converge to the optimal k . This ensures effective compensation performance across different input powers, enhancing the algorithm's robustness. Compared with the GDA and GDAM, achieving convergence without oscillation in four iterations does not lead to the optimal k . Figure 11a shows that, at the optimal input power of 3.5 dBm, the $\log_{10}(\text{BER})$ for both the GDA and GDAM is worse than that of our proposed algorithm. Additionally, it should be noted that the GDA and GDAM have a fixed μ . This leads to oscillations during the iterative optimization process under varying input powers, resulting in fluctuations in compensation effectiveness. Moreover, their performance at certain power levels is even worse than merely compensating for dispersion alone.

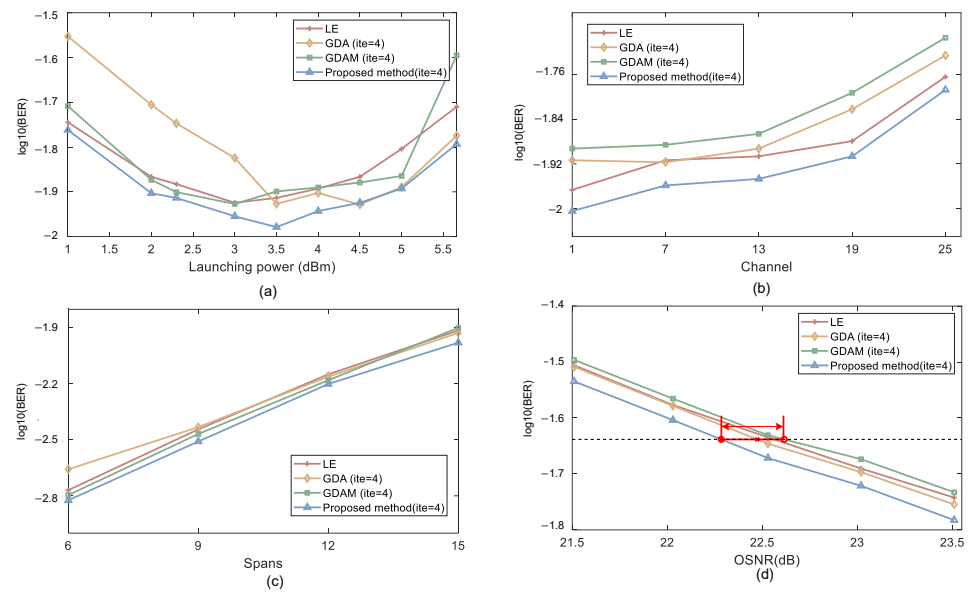


Figure 11. (a) $\log_{10}(\text{BER})$ vs. the launching power for LE, conventional A-DBP, and proposed method with different values of γ with SSFM steps of 5. (b) $\log_{10}(\text{BER})$ vs. the different WDM channels at 3.5 dBm for LE, conventional A-DBP, and proposed method. (c) $\log_{10}(\text{BER})$ vs. the different spanning cases at 3.5 dBm for LE, conventional A-DBP, and proposed method. (d) $\log_{10}(\text{BER})$ vs. the different OSNRs at 3.5 dBm for LE, conventional A-DBP, and proposed method with 4 iterations. The red arrow represents OSNR penalty when switching the algorithm from A-DBP-LTM to GDA or GDAM.

Figure 11b depicts the compensation performance of three algorithms across different WDM channels at the optimal input power of 3.5 dBm. The OSNR for channels 1, 7, 13, 19, and 25 are 25.12 dB, 25.07 dB, 25.01 dB, 24.4 dB, and 22.76 dB, respectively. As shown in Figure 11b, as OSNR decreases, the $\log_{10}(\text{BER})$ of all three algorithms increases, but our proposed method maintains superior compensation performance across all channels. The GDA and GDAM require more iterations to find the optimal k . Under the condition of a fair comparison (i.e., four iterations), their compensation performance is inferior. The GDAM exhibits the worst $\log_{10}(\text{BER})$ because the introduction of the momentum term m reduces the robustness of the algorithm during iterations, leading to oscillations and affecting compensation performance. Moreover, due to inter-channel interference, compensation performance varies across different channels.

Figure 11c illustrates the compensation performance of the three algorithms in the middle channel under different spans at the optimal input power of 3.5 dBm. Our proposed method maintains excellent compensation performance even with only four iterations. It shows that, as the number of spans increases, the accumulation of nonlinear impairments also increases. A-DBP-LTM based on RMSProp achieves a more significant BER reduction compared to the compensation of dispersion only. In contrast, for the GDA and GDAM, as nonlinear impairments accumulate, it becomes challenging to converge to the optimal k , resulting in poorer compensation performance for both algorithms. To compare the compensation performances of the three algorithms at four iterations, the BER curves versus OSNR are evaluated as shown in Figure 11d. The dashed line represents the FEC threshold value of 2.3×10^{-2} . The OSNR difference at the FEC threshold for different algorithms is defined as the OSNR penalty. The OSNR is adjusted by combining the received optical signal with ASE noise whose power is controlled by a variable optical attenuator. As shown in Figure 11d, when switching the algorithm from A-DBP-LTM to the GDA or GDAM, there will emerge a 0.189 dB or 0.3315 dB OSNR penalty, respectively.

4. Discussion

In order to figure out the optimal nonlinear scaling factor to enhance the compensation performance and reduce the computational complexity, multiple improved A-DBP methods have been proposed [16,27]. However, their performances are sensitive to hyperparameter settings such as the momentum term and learning rate. Here, we propose a low-complexity and high-robustness nonlinear compensation method termed A-DBP-LTM. In the simulation results, our proposed method, when compared to conventional DBP and A-DBP based on the GDA, achieves significant complexity reductions of 31.35% and 58.47%, respectively. Furthermore, it shows high robustness with few iterations and a 0.33 dB improvement in OSNR tolerance through experimental demonstration.

Although promising results can be seen in the simulation and experiment, the current approach still has several limitations and challenges. First, a relatively low symbol rate of 63 Gbaud was used in our experiment. It is necessary to further verify the feasibility of the proposed algorithm in a coherent system operating at a higher symbol rate such as 130 Gbaud and beyond. Second, in addition to the nonlinear scaling factor k and the learning rate μ , there are still multiple parameters which influence the compensation performance but rely on artificial selection, such as the step size h of SSFM. In the future, research on blind optimization for them should be considered.

5. Conclusions

In conclusion, a novel A-DBP-LTM implemented by RMSProp is proposed in order to achieve low-complexity and high-robustness nonlinear compensation. By adaptively adjusting the learning rate μ based on all of the gradients of the previous iterations, the A-DBP-LTM based on RMSProp delivers the best compensation performance. Not only in a 69 Gbaud DP-16QAM simulation system over 2000 km but also in a 26- λ , 63 Gbaud, DP-16QAM experimental system over 1200 km, the proposed method demonstrates low complexity and high robustness. Compared to conventional DBP, A-DBP based on the GDA, and A-DBP based on the GDAM, it reduces the complexity with the same effectiveness and improves the OSNR tolerance. It is a highly effective approach to obtain the optimal value of the nonlinear scaling factor k and adaptively tune the learning rate μ , enabling effective nonlinear compensation in large-capacity systems for long-haul transmission.

Author Contributions: Conceptualization, M.Z. and H.Y.; methodology, Y.L.; software, H.Y. and Y.L.; validation, M.Z. and H.Y.; writing—original draft preparation, M.Z. and H.Y.; writing—review and editing, Y.L., Z.X., D.W., S.C., Z.Z. and H.L.; funding acquisition, D.W. and H.L. All authors have read and agreed to the published version of the manuscript.

Funding: National Key Research and Development Program of China (No. 2022YFB2903301).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

Acknowledgments: This article is a revised and expanded version of a paper entitled Fast and High-Robustness Adaptive Digital Back-Propagation for Fiber Nonlinearity Compensation, which was presented at ACP2023.

Conflicts of Interest: The authors declare that there are no conflicts of interest related to this article.

References

- Warf, B. International Competition Between Satellite and Fiber Optic Carriers: A Geographic Perspective*. *Prof. Geogr.* **2006**, *58*, 1–11.
- Lau, A.P.T.; Gao, Y.; Sui, Q.; Wang, D.; Zhuge, Q.; Morsy-Osman, M.H.; Chagnon, M.; Xu, X.; Lu, C.; Plant, D.V. Advanced DSP Techniques Enabling High Spectral Efficiency and Flexible Transmissions: Toward elastic optical networks. *IEEE Signal Process.* **2014**, *31*, 82–92. [[CrossRef](#)]
- Winzer, P.J.; Neilson, D.T.; Chraplyvy, A.R. Fiber-optic transmission and networking: The previous 20 and the next 20 years. *Opt. Express* **2018**, *26*, 24190–24239. [[CrossRef](#)] [[PubMed](#)]
- Amari, A.; Dobre, O.A.; Venkatesan, R.; Kumar, O.S.S.; Ciblat, P.; Jaouen, Y. A Survey on Fiber Nonlinearity Compensation for 400 Gb/s and Beyond Optical Communication Systems. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 3097–3113. [[CrossRef](#)]
- Kuzucu, O.; Okawachi, Y.; Salem, R.; Foster, M.A.; Turner-Foster, A.C.; Lipson, M.; Gaeta, A.L. Spectral phase conjugation via temporal imaging. *Opt. Express* **2009**, *17*, 20605–20614.
- Ip, E.; Kahn, J.M. Compensation of Dispersion and Nonlinear Impairments Using Digital Backpropagation. *J. Light. Technol.* **2008**, *26*, 3416–3425.
- Yaman, F.; Li, G. Nonlinear Impairment Compensation for Polarization-Division Multiplexed WDM Transmission Using Digital Backward Propagation. *IEEE Photon J.* **2010**, *2*, 816–832. [[CrossRef](#)]
- Du, L.B.; Lowery, A.J. Improved single channel backpropagation for intra-channel fiber nonlinearity compensation in long-haul optical communication systems. *Opt. Express* **2010**, *18*, 17075–17088.
- Lin, C.-Y.; Napoli, A.; Spinnler, B.; Sleiffer, V.A.; Rafique, D.; Kuschnerov, M.; Bohn, M.; Schmauss, B. Adaptive Digital Back-Propagation for Optical Communication Systems. In Proceedings of the Optical Fiber Communication Conference, OSA Technical Digest (Online), San Francisco, CA, USA, 9–13 March 2014; Optica Publishing Group: Washington, DC, USA, 2014; p. M3C.4.
- Jiang, L.; Yan, L.; Chen, Z.; Yi, A.; Pan, Y.; Pan, W.; Luo, B.; Li, G. Adaptive digital backward propagation based on variance of intensity noise. *Chin. Opt. Lett.* **2015**, *13*, 110602. [[CrossRef](#)]
- Gonçalves, J.; Martins, C.S.; Guiomar, F.P.; Cunha, T.R.; Pedro, J.C.; Pinto, A.N.; Lavrador, P.M. Nonlinear compensation with DBP aided by a memory polynomial. *Opt. Express* **2016**, *24*, 30309–30316.
- Jiang, L.; Yan, L.; Yi, A.; Pan, Y.; Ge, J.; Dai, L.; Pan, W.; Luo, B. Toward Blind Nonlinearity Estimation in Back-Propagation Algorithm for Coherent Optical Transmission Systems. In Proceedings of the Optical Fiber Communication Conference, OSA Technical Digest (Online), Los Angeles, CA, USA, 19–23 March 2017; Optica Publishing Group: Washington, DC, USA, 2017; p. W1G.3.
- Zhang, F.; Zhuge, Q.; Qiu, M.; Zhou, X.-Y.; Sowailam, M.Y.; Hoang, T.M.; Xiang, M.; Plant, D.V. Blind Adaptive Digital Backpropagation for Fiber Nonlinearity Compensation. *J. Light. Technol.* **2017**, *36*, 1746–1756. [[CrossRef](#)]
- Martins, C.S.; Bertignono, L.; Nespola, A.; Carena, A.; Guiomar, F.P.; Pinto, A.N. Low-Complexity Time-Domain DBP Based on Random Step-Size and Partitioned Quantization. *J. Light. Technol.* **2018**, *36*, 2888–2895. [[CrossRef](#)]
- Wang, C.; Tian, F.; Wang, Y.; Zhang, Q.; Tian, Q.; Gao, R.; Wang, X.; Pan, X.; Yang, L.; Xin, X. Digital back propagation based on the feedback-convergence for fiber nonlinearity compensation with low computational complexity and high performance. *Opt. Commun.* **2020**, *471*, 126050. [[CrossRef](#)]
- Júnior, J.H.C.; Sutili, T.; Rossi, S.M.; Figueiredo, R.C.; Mello, D.A.A. Fast Adaptive Digital Back-Propagation Algorithm for Unrepeated Optical Systems. In Proceedings of the Optical Fiber Communication Conference, OSA Technical Digest, San Diego, CA, USA, 8–12 March 2020; Optica Publishing Group: Washington, DC, USA, 2020; p. T4I.2.
- Maghrabi, M.M.T.; Bakr, M.H.; Kumar, S. Adaptive digital back propagation exploiting adjoint-based optimization for fiber-optic communications. *Opt. Express* **2022**, *30*, 16264–16288.
- Guiomar, F.P.; Amado, S.B.; Martins, C.S.; Pinto, A.N. Time-Domain Volterra-Based Digital Backpropagation for Coherent Optical Systems. *J. Light. Technol.* **2015**, *33*, 3170–3181. [[CrossRef](#)]
- Zhou, J.; Wang, Y.; Zhang, Y. Blind back-propagation method for fiber nonlinearity compensation with low computational complexity and high performance. *Opt. Express* **2020**, *28*, 11424–11438. [[CrossRef](#)]
- Zhang, S.; Yaman, F.; Nakamura, K.; Inoue, T.; Kamalov, V.; Jovanovski, L.; Vusirikala, V.; Mateo, E.; Inada, Y.; Wang, T. Field and lab experimental demonstration of nonlinear impairment compensation using neural networks. *Nat. Commun.* **2019**, *10*, 3033.
- Deligiannidis, S.; Mesaritakis, C.; Bogris, A. Performance and Complexity Analysis of Bi-Directional Recurrent Neural Network Models Versus Volterra Nonlinear Equalizers in Digital Coherent Systems. *J. Light. Technol.* **2021**, *39*, 5791–5798. [[CrossRef](#)]
- Khan, F.N.; Fan, Q.; Lu, C.; Lau, A.P.T. An Optical Communication's Perspective on Machine Learning and Its Applications. *J. Light. Technol.* **2019**, *37*, 493–516. [[CrossRef](#)]
- Chi, X.; Bai, C.; Yang, F.; Qi, Q.; Zhang, R.; Xu, H.; Yang, L.; Bi, W.; Chen, T.; Bai, S. Joint intra and inter-channel nonlinear compensation scheme based on improved learned digital back propagation for WDM systems. *Opt. Express* **2024**, *32*, 5095.

24. Oliari, V.; Goossens, S.; Hager, C.; Liga, G.; Butler, R.M.; Hout, M.v.D.; van der Heides, S.; Pfister, H.D.; Okonkwo, C.; Alvarado, A. Revisiting Efficient Multi-Step Nonlinearity Compensation With Machine Learning: An Experimental Demonstration. *J. Light. Technol.* **2020**, *38*, 3114–3124.
25. Häger, C.; Pfister, H.D. Nonlinear Interference Mitigation via Deep Neural Networks. In Proceedings of the 2018 Optical Fiber Communication Conference, San Diego, CA, USA, 11–15 March 2018; p. W3A.4.
26. Huang, J.; Yi, A.; Yan, L.; He, X.; Jiang, L.; Yang, H.; Luo, B.; Pan, W. Fast fiber nonlinearity compensation method for PDM coherent optical transmission systems based on the Fourier neural operator. *Opt. Express* **2024**, *32*, 2245. [[CrossRef](#)] [[PubMed](#)]
27. Chen, X.; Zhang, Q.; Gao, R.; Xin, X.; Wang, X.; Tian, Q.; Tian, F.; Wang, Y.; Hou, Y.; Yang, L. Fast adaptive digital back-propagation algorithm for fiber nonlinear compensation. In Proceedings of the Asia Communications and Photonics Conference/International Conference on Information Photonics and Optical Communications, OSA Technical Digest, Beijing, China, 24–27 October 2020; Optica Publishing Group: Washington, DC, USA, 2020; p. M4A.241.
28. Yao, H.; Li, W.; Han, J.; He, Z.; Yuan, Z.; Hu, Q.; Yang, Q.; Yu, S. A Modified Adaptive DBP for DP 16-QAM Coherent Optical System. *IEEE Photon-Technol. Lett.* **2016**, *28*, 2511–2514.
29. Chen, Z.; Yan, L.; Yi, A.; Jiang, L.; Pan, Y.P.W.; Luo, B. Low complexity and adaptive nonlinearity compensation. In Proceedings of the Conference on Lasers and Electro-Optics, OSA Technical Digest (Online), San Jose, CA, USA, 10–15 May 2015; Optica Publishing Group: Washington, DC, USA, 2015; p. SM2M.6.
30. Liu, Y.; Xie, Z.; Zuo, M.; Wang, D.; Cao, S.; Li, Y.; Zhang, D. Improved Fast Anti-fluctuation Adaptive Digital Back-Propagation for Kerr-nonlinearity Compensation. In Proceedings of the 2023 Opto-Electronics and Communications Conference (OECC), Shanghai, China, 2–6 July 2023; pp. 1–3.
31. Liu, Y.; Zuo, M.; Wang, D.; Xie, Z.; Zhao, X.; Zheng, Z.; Cao, S.; Li, Y.; Zhang, D.; Li, H. Fast and High-Robustness Adaptive Digital Back-Propagation for Fiber Nonlinearity Compensation. In Proceedings of the 2023 Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM), Wuhan, China, 4–7 November 2023; pp. 1–4.
32. Agrawal, G.P. *Nonlinear Fiber Optics*, 5th ed.; Academic Press: Oxford, UK, 2012.
33. Millar, D.S.; Makovejs, S.; Behrens, C.; Hellerbrand, S.; Killey, R.I.; Bayvel, P.; Savory, S.J. Mitigation of Fiber Nonlinearity Using a Digital Coherent Receiver. *IEEE J. Sel. Top. Quantum Electron.* **2010**, *16*, 1217–1226.
34. Kudo, R.; Kobayashi, T.; Ishihara, K.; Takatori, Y.; Sano, A.; Miyamoto, Y. Coherent Optical Single Carrier Transmission Using Overlap Frequency Domain Equalization for Long-Haul Optical Systems. *J. Light. Technol.* **2009**, *27*, 3721–3728. [[CrossRef](#)]
35. Huk, M. Stochastic optimization of contextual neural networks with RMSprop. In *Intelligent Information and Database Systems*; Nguyen, N.T., Jearanaitanakij, K., Selamat, A., Trawiński, B., Chittayasothorn, S., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 343–352.
36. Spinnler, B. Equalizer Design and Complexity for Digital Coherent Receivers. *IEEE J. Sel. Top. Quantum Electron.* **2010**, *16*, 1180–1192. [[CrossRef](#)]
37. Guiomar, F.P.; Pinto, A.N. Simplified Volterra Series Nonlinear Equalizer for Polarization-Multiplexed Coherent Optical Systems. *J. Light. Technol.* **2013**, *31*, 3879–3891. [[CrossRef](#)]
38. Mateo, E.F.; Yaman, F.; Li, G. Efficient compensation of inter-channel nonlinear effects via digital backward propagation in WDM optical transmission. *Opt. Express* **2010**, *18*, 15144–15154. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.