CHINESE OPTICS LETTERS

Vol. 21, No. 3 | March 2023

Speckle backpropagation for compensation of nonlinear effects in few-mode optical fibers

Pavel S. Anisimov*, Evgeny D. Tsyplakov, Viacheslav V. Zemlyakov, and Jiexing Gao (高杰星)

Central Research Institute, 2012 Labs, Huawei Technologies, Shenzhen 518129, China

*Corresponding author: anisimov.pavel1@huawei.com

Received July 12, 2022 | Accepted September 20, 2022 | Posted Online November 14, 2022

We propose an alternative approach to compensation of intermodal interactions in few-mode optical fibers by means of digital backpropagation. Instead of solving the inverse generalized multimode nonlinear Schrödinger equation, we accomplish backpropagation of the multimode signals with help of their near-field intensity distributions captured by a camera. We demonstrate that this task can successfully be handled by a deep neural network and provide a proof of concept by training an autoencoder for backpropagation of six linearly polarized (LP) modes of a step-index fiber.

Keywords: optical fibers; multimode fibers; few-mode fibers; digital signal processing; space division multiplexing; mode division multiplexing; mode decomposition.

DOI: 10.3788/COL202321.030601

1. Introduction

Multimode optical fibers (MMFs) have been actively studied over the last decade due to their high potential for overcoming the capacity crunch of optical communications with the technology referred to as mode division multiplexing (MDM)[1]. Currently, the main limiting factor of the modern systems, relying on single-mode fibers (SMFs), is the intensity-dependent Kerr nonlinearity. Coarsely, methods for mitigation of nonlinear effects in SMFs can be divided into two categories: optical[2–4] and digital[5–12]. The former implies utilization of optical devices and relevant physical effects for lowering the impact of nonlinearities, whereas the latter relies on digital signal processing (DSP) and implementation of numerical algorithms, compensating for the corresponding impairments.

Amongst DSP techniques, digital backpropagation[9–12] (DBP) stands out, as it is also capable of treating all other deterministic effects. The main idea behind DBP is solving the nonlinear Schrödinger equation[13] in the direction opposite to that of the propagating light, which turns out to be possible due to the prior knowledge on the transmitting system, i.e., fiber specifications. Typically, DBP is accomplished with the help of the split-step Fourier method[13], which has to be tuned so that it can operate within reasonable time boundaries without sacrificing performance. This trade-off between computational effort and accuracy has been recently canceled with the help of deep-neural networks (DNNs)[14,15].

In addition to the challenges posed by SMFs, MMFs propose their own, mostly originating from the nature of multimode interactions that arise in media, such as intermodal cross-phase modulation and four-wave mixing, Kerr and Raman nonlinearities[16–22]. For instance, interchannel cross talk represents a serious limitation for implementation of MMFs in telecommunication[23–26] that is known to be successfully mitigated by multimode DBP (MM-DBP)[25,26]. SSF is naturally extendable to the multimode case, but its computational complexity scales as $O(N^4)$ for N modes[27].

It is noteworthy that no standards for MDM receiver operation principles have been established yet, and not every spatial demultiplexing scenario would be appropriate for conventional MM-DBP schemes. In particular, demultiplexing may refer to utilization of bulky optics[28], physical mode separation[29], or implementation of numerical algorithms[30–33]. The latter is mostly given by the technique referred to as intensity-only mode decomposition (MD), which implies separation of modes, based on the near-field intensity distributions of multimode signals captured by a digital camera. This approach is deemed promising due to the simplicity of experimental realization of the receiver part and its reconfigurability. However, existing MM-DBP methods operate in terms of independent mode signals, i.e., assuming that spatial demultiplexing is performed prior to DBP, whereas modern cameras are not able to provide temporal exposure of multimode signals, and, as a consequence, known MM-DBP methods cannot be applied directly to the systems equipped with such a receiver.

To the best of our knowledge, DNNs have not been utilized as a tool for MM-DBP, despite the fact that it is known that they can capture nonlinear interactions in MMFs[34,35], and, analogous to the case of SMFs, can provide fast and accurate calculation. In this Letter, we turn the tables and apply MD

algorithms after DNN-based MM-DBP, thus considering the DBP problem for entangled multimode signals defined by the near-field intensity speckles. With the help of numerical simulations, we provide a proof of concept and demonstrate that such a recovery of initial modal content can take place in the case of severe nonlinearities. Additionally, we show that typical noise levels of modern cameras do not have remarkable influence on the performance of the proposed MM-DBP. We consider the case of six linearly polarized (LP) modes, and following the naming convention, we call our fiber a few-mode fiber (FMF).

2. Methods

2.1. The main idea

The flow chart illustrating our approach is displayed in Fig. 1(a). Given the output near-field intensity pattern, we want to recover the modal content of the input multimode signal. We propose utilizing two neural networks: one for MM-DBP and another one for MD. Mode configurations at the fiber output are arranged as multimode beam speckles that are fed to MM-DBP DNN, which results in those at the fiber input. The final stage is represented by MD-DNN, which provides modal content of the initial multimode signal. In order to train the MM-DBP network, we need pairs of input and output beam speckles that require numerical simulation of multimode signal propagation. Details on the MD-DNN architecture lie beyond the scope of the current work, as this part is purely technical and can be given by any other method[30,31]. In what follows, under MD-DNN, we understand an appropriately trained vision transformer[33].

2.2. Nonlinear light propagation in FMFs

Given the complex electrical field modal envelope $A_p(z,t)$ of the mode p, pulse propagation in FMFs fulfills the generalized nonlinear multimode Schrödinger equation (GMMNLSE)[27,36],

$$\frac{\partial A_p}{\partial z} = i(\beta_0^{(p)} - \Re\beta_0^{(0)})A_p - (\beta_1^{(p)} - \Re\beta_1^{(0)})\frac{\partial A_p}{\partial t} + i\sum_{n\geq 2}^{N_d}\frac{\beta_n^{(p)}}{n!}\left(i\frac{\partial}{\partial t}\right)^n A_p + i\frac{n_2\omega_0}{c}\sum_{l,m,n}[(1-f_R)S^K_{plmn}A_lA_mA_n^* + f_R S^R_{plmn}A_l\int_{-\infty}^{t}\mathrm{d}\tau h_R(\tau)A_m(z,t-\tau)A_n^*(z,t-\tau)], \quad (1)$$

where $\beta_n^{(p)} = \partial^n\beta^{(p)}/\partial\omega^n$ refers to the n^{th}-order dispersion coefficient for the p^{th} mode (limited by N_d), $\Re$ denotes the real part of a complex number, n_2 is the nonlinear index coefficient, ω_0 is the center frequency, S^K_{plmn} (S^R_{plmn}) is the modal overlap tensor for the Kerr (Raman) term, h_R is the Raman response of the fiber medium, and $f_R \simeq 0.18$ is the Raman contribution factor for fused silica. In our simulations, we set $N_d = 3$.

For obtaining the training data, we generate random amplitudes for Gaussian envelopes $A_p(t)$ with $p = 1, 2, \ldots, 6$ and simulate multimode light propagation by solving Eq. (1) with the help of the massively parallel algorithm (MPA)[27] on graphics processing unit. Then the output intensity distribution at the receiver is given by

$$I(x,y) = \left|\sum_{i=1}^{N} A_i\varphi_i(x,y)\right|^2 = \sum_{i,j=1}^{N} A_iA_j^*\varphi_i(x,y)\varphi_j(x,y), \quad (2)$$

where $\varphi_i(x,y)$ are normalized transverse mode eigenfunctions[37] and N is the number of modes. The goal of MD-DNN is inferring $\mathrm{Re}\,A_i$ and $\mathrm{Im}\,A_i$ from $I(x,y)$. Here, we mention that the complex coefficients can be recovered up to their conjugation due to the fact that intensity is conjugation-invariant. Another degree of freedom, namely, arbitrariness of the relative phase shifts, is fixed by the constraint $\mathrm{Im}\,A_0 = 0$.

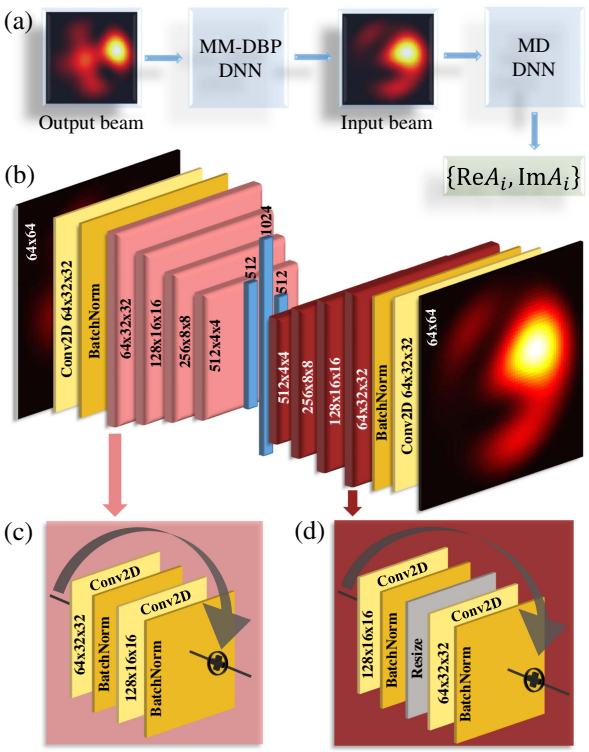

Fig. 1. (a) Flow chart of the proposed method. A near-field beam speckle, captured at the fiber output, is fed to MM-DBP DNN, resulting in the recovered input pattern. MD-DNN takes this input pattern and yields real and imaginary parts of complex mode coefficients. (b) Schematic of the proposed MM-DBP DNN. We use a residual neural network (ResNet) based autoencoder that compresses information acquired from the speckles and maps it onto a vector from the latent feature space. The decoder maps it back to the speckle space. Three middle blue blocks denote fully connected (FC) layers with 512, 1024, and 512 neurons, respectively. After each FC layer, we also place a dropout layer. BatchNorm stands for batch normalization. We use the rectified linear unit (ReLu) as the activation function. (c), (d) Detailed structure of the encoder and the decoder $64 \times 32 \times 32$ basic blocks, respectively. Those with other dimensions can be obtained analogously.

Table 1. Parameters Used in Simulations of Nonlinear Light Propagation in FMFs.

Parameter	Value	Parameter	Value
Total energy	10 nJ	NA	0.2
Fiber radius	25 μm	Wavelength	1550 nm
Fiber length	1 m	N	6
t_{FWHM}	0.1 ps	Image size	64×64

It is worth noting that even MPA[27] is computationally expensive and time-consuming, while the main goal of this work is to prove consistency of the proposed MM-DBP approach. Therefore, without loss of generality, we set simulation specifications so that the nonlinear interaction becomes colossal. Consideration of such scenarios, which are probably unsuitable for communication, allows us to observe severe interchannel cross talk on shorter distances. Detailed simulation specifications can be found in Table 1.

2.3. The network

From the versatile pool of suitable DNN architectures, we opt for the model given by a residual neural network (ResNet) based autoencoder [see Figs. 1(b)–1(d)]. Input 64×64 images are mapped by the encoder onto vectors from the latent feature space, which are processed by three dense layers and transformed to the output 64×64 pictures by the decoder.

We quantify the network's reconstruction quality by two values. The first one is the correlation,

$$C = \left| \frac{\iint \Delta I_{\text{true}}(x,y)\Delta I_{\text{rec}}(x,y)\mathrm{d}x\mathrm{d}y}{\sqrt{\iint \Delta I^2_{\text{true}}(x,y)\mathrm{d}x\mathrm{d}y \cdot \iint \Delta I^2_{\text{rec}}(x,y)\mathrm{d}x\mathrm{d}y}} \right|, \quad (3)$$

where $\Delta I_j(x,y) = I_j(x,y) - \bar{I}_j(x,y)$, and $\bar{I}_j$ is the mean intensity value, with j indexing either true or recovered patterns.

Final estimation of the network's performance is given by the mean squared error (MSE) between the genuine $\{\mathrm{Re}\, A_i, \mathrm{Im}\, A_i\}$ and those obtained by MD-DNN,

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}[(\mathrm{Re}\, A_i^t - \mathrm{Re}\, A_i^r)^2 + (\mathrm{Im}\, A_i^t - \mathrm{Im}\, A_i^r)^2], \quad (4)$$

where t and r stand for "true" and "recovered," respectively.

We train the network in two stages. At the first one, we use 150,000 image pairs and train our DNN for 100 epochs with the AdamW optimizer[38]. The learning rate is set to 5×10^{-4}, which is halved every 20 epochs. The loss function is chosen as MSE between two images. It turns out that after this stage, the network's output carries visually unnoticeable reconstruction artifacts. In other words, despite high values of the correlation between genuine input speckles and those resulting from MM-DBP, there is high discrepancy between the corresponding outputs of MD-DNN. These artifacts can be considered and treated as noise. Therefore, the second stage is aimed at rectifying this problem by transfer learning of the MM-DBP network for an extra 50 epochs with another 40,000-data set by applying the MD-DNN to ground truth and output speckles and calculating the loss function,

$$\mathcal{L} = \mathrm{MSE} + \mathrm{PSNRLoss}(\mathrm{beam}_{\text{true}}, \mathrm{beam}_{\text{rec}}), \quad (5)$$

where PSNRLoss stands for peak signal-to-noise ratio (PSNR) loss function, which is simply the negative value of PSNR calculated for the ground truth and the MM-DBP network's output. This maneuver grants substantial improvements in the MM-DBP DNN performance.

3. Results and Discussion

We validate our MM-DBP model with an unfamiliar 1000-data set of images. For each sample, we calculate the correlation and MSE. It is noteworthy that calculation of these metrics implicitly benchmarks our networks against MPA by formulation, since this algorithm has been utilized for collecting the training samples. The resulting averaged metrics achieve decent values: $\bar{C} = 0.992$ and $\mathrm{MSE} = 2.3 \times 10^{-4}$. Figure 2(a) demonstrates an MM-DBP example. The left panel shows the fiber output beam

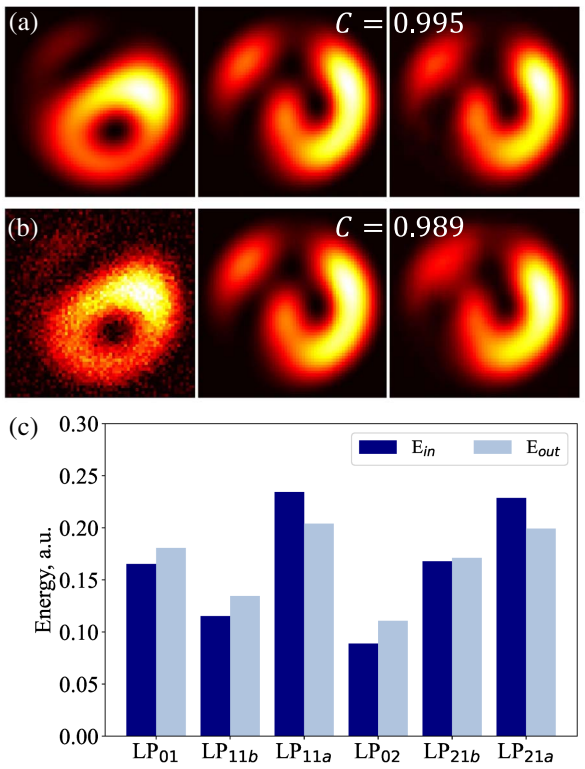

Fig. 2. (a) Example of an output, initial, and recovered speckle, respectively, in the absence of noise [see Eq. (6)]; (b) same patterns with receiver noise included (SNR = 10 dB); (c) energy redistribution for each mode after having propagated the fiber.

speckle, whereas the middle and the right panels display the true and the recovered input speckles, respectively. The resulting correlation is $C = 0.995$. The impact of nonlinear interactions is reflected in Fig. 2(c), where we demonstrate energy redistribution among the modes in the case of the sample propagation shown in Fig. 2(a).

However, random receiver noise often plays a crucial role in MD. There is always a trade-off between noise levels and the number of decomposable modes[30,31]. Therefore, investigation of the noise influence on MM-DBP performance is of paramount importance. According to the most common way, such noise can be modeled as additive white Gaussian noise[30],

$$I^{\text{noised}}(x,y) = \max[0, I^{\text{true}}(x,y) + N(0,\sigma)], \tag{6}$$

where $N(0,\sigma)$ is a normally distributed random matrix of an appropriate size. We imply that intensities are normalized so that for every speckle $\max(I^{\text{true}}) = 1$ holds.

Varying σ yields different noise levels that we quantify by the signal-to-noise ratio (SNR). We put our MM-DBP network to the test for SNR values lying between 10 and 90 dB. We follow the same pipeline using the same validation 1000-data set. The only difference is that inputs of the network are subjected to the noise procedure according to Eq. (6). We analogously calculate the quality metrics for each value of SNR. Figure 3 illustrates the behavior of the averaged MSE between the decomposed and original complex coefficients, as well as the average value of the correlation coefficient. Slightly above SNR = 20 dB, one can observe a threshold that defines the applicability of the current network architecture trained on the noise-free data. However, in some cases, the network performs well even at 10 dB [see Fig. 2(b)]. Taking into account that modern cameras provide SNR levels of 40–60 dB, our simulations reveal high potential for a realistic implementation of the proposed approach.

Herein, we proposed an alternative approach to the tedious task of compensation of intermodal interactions in FMFs relying only on the near-field intensity distribution obtained at the fiber output. As a proof of concept, we provided a model trained for DBP of six consequent LP modes propagating through a step-index fiber in the strongly nonlinear regime. This number of modes represents a typical value for the physically separable channels used for MM-DBP. We have numerically shown that this technique might be a decent alternative to the existing pipeline of nonlinearity compensation methods in the case of FMFs. We hope that this principle of MM-DBP might be useful for intensity-only receiver-based FMF communication systems.

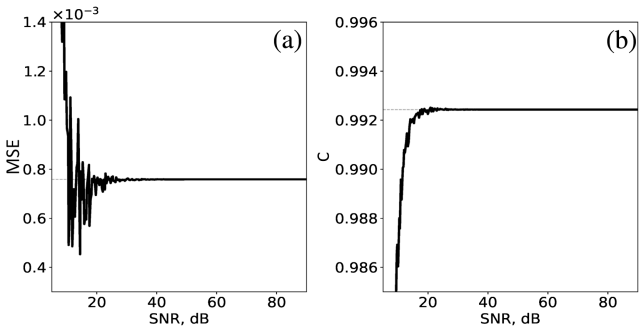

Fig. 3. (a) MSE and (b) C versus SNR. Each point represents averaging over 1000 samples from the validation data set.

References

1. J. Du, W. Shen, J. Liu, Y. Chen, X. Chen, and Z. He, "Mode division multiplexing: from photonic integration to optical fiber transmission," Chin. Opt. Lett. **19**, 091301 (2021).
2. Y. Yamamoto, Y. Kawaguchi, and M. Hirano, "Low-loss and low-nonlinearity pure-silica-core fiber for C- and L-band broadband transmission," J. Light. Technol. **34**, 321 (2016).
3. S. L. Jansen, D. van den Borne, P. M. Krummrich, S. Sälter, G.-D. Khoe, and H. de Waardt, "Long-haul DWDM transmission systems employing optical phase conjugation," IEEE J. Sel. Top. Quantum Electron. **12**, 505 (2006).
4. X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," Nat. Photonics **7**, 560 (2013).
5. R. J. Essiambre, P. J. Winzer, X. Q. Wang, W. Lee, C. A. White, and E. C. Burrows, "Electronic predistortion and fiber nonlinearity," IEEE Photon. Technol. Lett. **18**, 1804 (2006).
6. Y. Gao, J. C. Cartledge, A. S. Karar, S. S.-H. Yam, M. O'Sullivan, C. Laperle, A. Borowiec, and K. Roberts, "Reducing the complexity of perturbation based nonlinearity pre-compensation using symmetric EDC and pulse shaping," Opt. Express **22**, 1209 (2014).
7. M. Secondini and E. Forestieri, "On XPM mitigation in WDM fiber-optic systems," IEEE Photon. Technol. Lett. **26**, 2252 (2014).
8. K. Zhang, G. Gao, J. Zhang, A. Fei, and M. Cvijetic, "Mitigation of nonlinear fiber distortion using optical phase conjugation for mode-division multiplexed transmission," Opt. Fiber Technol. **43**, 169 (2018).
9. E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," J. Light. Technol. **28**, 939 (2010).
10. L. B. Du and A. J. Lowery, "Improved single channel backpropagation for intra-channel fiber nonlinearity compensation in long-haul optical communication systems," Opt. Express **18**, 17075 (2010).
11. L. Jiang, L. Yan, Z. Chen, A. Yi, Y. Pan, W. Pan, B. Luo, and G. Li, "Adaptive digital backward propagation based on variance of intensity noise," Chin. Opt. Lett. **13**, 110602 (2015).
12. R. Maher, T. Xu, L. Galdino, M. Sato, A. Alvarado, K. Shi, S. J. Savory, B. C. Thomsen, R. I. Killey, and P. Bayvel, "Spectrally shaped DP-16QAM super-channel transmission with multi-channel digital back-propagation," Sci. Rep. **5**, 8214 (2015).
13. G. P. Agrawal, *Nonlinear Fiber Optics* (Academic Press, 2019).
14. Q. Fan, G. Zhou, T. Gui, L. Chao, and A. P. T. Lau, "Advancing theoretical understanding and practical performance of signal processing for nonlinear optical communications through machine learning," Nat. Commun. **11**, 3694 (2020).
15. B. I. Bitachon, A. Ghazisaeidi, M. Eppenberger, B. Baeuerle, M. Ayata, and J. Leuthold, "Deep learning based digital backpropagation demonstrating SNR gain at low complexity in a 1200 km transmission link," Opt. Express **28**, 29318 (2020).
16. A. Mecozzi, C. Antonelli, and M. Shtaif, "Coupled Manakov equations in multimode fibers with strongly coupled groups of modes," Opt. Express **20**, 23436 (2012).
17. A. Mecozzi, C. Antonelli, and M. Shtaif, "Nonlinear propagation in multimode fibers in the strong coupling regime," Opt. Express **20**, 11673 (2012).
18. S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Nonlinear propagation in multimode and multicore fibers: generalization of the Manakov equations," J. Light. Technol. **31**, 398 (2013).
19. F. Ferreira, S. Jansen, P. Monteiro, and H. Silva, "Nonlinear semi-analytical model for simulation of few-mode fiber transmission," IEEE Photon. Technol. Lett. **24**, 240 (2012).

20. L. G. Wright, D. N. Christodoulides, and F. W. Wise, "Controllable spatiotemporal nonlinear effects in multimode fibres," Nat. Photonics **9**, 306 (2015).
21. K. Krupa, A. Tonello, B. M. Shalaby, M. Fabert, A. Barthélémy, G. Millot, S. Wabnitz, and V. Couderc, "Spatial beam self-cleaning in multimode fibres," Nat. Photonics **11**, 237 (2017).
22. M. A. Eftekhar, Z. Sanjabi-Eznaveh, H. E. Lopez-Aviles, S. Benis, J. E. Antonio-Lopez, M. Kolesik, F. Wise, R. Amezcua-Correa, and D. N. Christodoulides, "Accelerated nonlinear interactions in graded-index multimode fibers," Nat. Commun. **10**, 1638 (2019).
23. G. Rademacher, R. Ryf, N. K. Fontaine, H. Chen, R.-J. Essiambre, B. J. Puttnam, R. S. Luís, Y. Awaji, N. Wada, S. Gross, N. Riesen, M. Withford, Y. Sun, and R. Lingle, "Long-haul transmission over few-mode fibers with space-division multiplexing," J. Light. Technol. **36**, 1382 (2018).
24. G. Rademacher, R. S. Luís, B. J. Puttnam, Y. Awaji, M. Suzuki, T. Hasegawa, H. Furukawa, and N. Wada, "Wideband intermodal nonlinear signal processing with a highly nonlinear few-mode fiber," IEEE J. Sel. Top. Quantum Electron. **26**, 7702007 (2020).
25. F. M. Ferreira, C. S. Costa, S. Sygletos, and A. D. Ellis, "Nonlinear performance of few-mode fiber links with intermediate coupling," J. Light. Technol. **37**, 989 (2019).
26. F. M. Ferreira, S. Sygletos, E. Sillekens, R. Killey, A. D. Ellis, and N. J. Doran, "On the performance of digital back propagation in spatial multiplexing systems," J. Light. Technol. **38**, 2790 (2020).
27. L. G. Wright, Z. M. Ziegler, P. M. Lushnikov, Z. Zhu, M. A. Eftekhar, D. N. Christodoulides, and F. W. Wise, "Multimode nonlinear fiber optics: massively parallel numerical solver, tutorial, and outlook," IEEE J. Sel. Top. Quantum Electron. **24**, 5100516 (2018).
28. J. Li, X. Zhang, Y. Zheng, F. Li, X. Shan, Z. Han, and R. Zhu, "Fast fiber mode decomposition with a lensless fiber-point-diffraction interferometer," Opt. Lett. **46**, 2501 (2021).
29. S. G. Leon-Saval, N. K. Fontaine, J. R. Salazar-Gil, B. Ercan, R. Ryf, and J. Bland-Hawthorn, "Mode-selective photonic lanterns for space-division multiplexing," Opt. Express **22**, 1036 (2014).
30. E. Manuylovich, V. Dvoyrin, and S. Turitsyn, "Fast mode decomposition in few-mode fibers," Nat. Commun. **11**, 5507 (2020).
31. P. S. Anisimov, V. V. Zemlyakov, and J. Gao, "2D least-squares mode decomposition for mode division multiplexing," Opt. Express **30**, 8804 (2022).
32. Y. An, L. Huang, J. Li, J. Leng, L. Yang, and P. Zhou, "Learning to decompose the modes in few-mode fibers with deep convolutional neural network," Opt. Express **27**, 10127 (2019).
33. Q. Zhang, S. Rothe, N. Koukourakis, and J. Czarske, "Learning the matrix of few-mode fibers for high-fidelity spatial mode transmission," APL Photonics **7**, 066104 (2022).
34. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode optical fiber transmission with a deep learning network," Light Sci. Appl. **7**, 69 (2018).
35. U. Teğin, B. Rahmani, E. Kakkava, N. Borhani, C. Moser, and D. Psaltis, "Controlling spatiotemporal nonlinearities in multimode fibers with deep neural networks," APL Photonics **5**, 030804 (2020).
36. F. Poletti and P. Horak, "Description of ultrashort pulse propagation in multimode optical fibers," J. Opt. Soc. Am. B **25**, 1645 (2008).
37. P. S. Anisimov, V. S. Motolygin, V. V. Zemlyakov, and J. Gao, "Fast multi step-index mode solver for analysis and optimization of optical fiber performance," J. Light. Technol. **40**, 2980 (2022).
38. I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," arXiv:1711.05101 (2019).