



Optics Letters

Limits of nonlinear and dispersive fiber propagation for an optical fiber-based extreme learning machine

ANDREI V. ERMOLAEV,¹  MATHILDE HARY,^{1,2}  LEV LEYBOV,² PIOTR RYCZKOWSKI,² 
ANAS SKALLI,¹ DANIEL BRUNNER,¹  GOËRY GENTY,²  AND JOHN M. DUDLEY^{1,3,*} 

¹ Université Marie et Louis Pasteur, Institut FEMTO-ST, CNRS UMR 6174 Besançon 25000, France

² Photonics Laboratory, Tampere University, Tampere FI-33104, Finland

³ Institut Universitaire de France, Paris, France

*john.dudley@univ-fcomte.fr

Received 17 March 2025; revised 27 May 2025; accepted 29 May 2025; posted 30 May 2025; published 17 June 2025

We report a generalized nonlinear Schrödinger equation simulation model of an extreme learning machine (ELM) based on optical fiber propagation. Using the MNIST handwritten digit dataset as a benchmark, we study how accuracy depends on propagation dynamics, as well as parameters governing spectral encoding, readout, and noise. For this dataset and with quantum noise limited input, test accuracies of over 91% and 93% are found for propagation in the anomalous and normal dispersion regimes, respectively. Our results also suggest that quantum noise on the input pulses introduces an intrinsic penalty to ELM performance. © 2025 Optica Publishing Group. All rights, including for text and data mining (TDM), Artificial Intelligence (AI) training, and similar technologies, are reserved.

<https://doi.org/10.1364/OL.562186>

There is currently intense interest in developing photonic-based artificial intelligence hardware. Often described as optical neuromorphic computing [1,2], experiments have reported photonic neural networks [3–8], reservoir computers [9–11], and extreme learning machines (ELMs) [7,12,13]. One specific focus has been applying nonlinear wave propagation directly as a computational resource [14], and recent experiments have reported ELM classification based on nonlinear propagation and supercontinuum generation in optical fiber [15–18]. However, while fiber-based ELM performance has been studied experimentally over broad parameter ranges [17,18], guidance from a robust simulation model is clearly needed in order to gain a more complete picture of such nonlinear wave-based computing.

Here, we report such an end-to-end numerical model of nonlinear fiber propagation in an ELM architecture, where we analyze handwritten digit classification using the MNIST Digit dataset of 60,000 training and 10,000 test images. (Supplement 1 shows results for the more complex MNIST Fashion dataset.) We specifically study how the ELM-based classification accuracy depends on data encoding, propagation dynamics, and readout, as well as input noise. Although some parameter dependencies have been studied experimentally [17–19], simulations provide significant new insights in evaluating the effect of quantum and

technical noise sources and in exploring different dynamical regimes.

We begin by describing the overall system in Fig. 1. An ELM can be considered as a neural network model computing output \mathbf{Y} from input \mathbf{X} . An ELM's ability to compute depends on the projection of input \mathbf{X} into a higher-dimensional space via a nonlinear transformation $\mathbf{H} = f_{\text{NL}}(\mathbf{X})$, with \mathbf{H} considered as a hidden layer. ELMs are computationally extremely efficient because it is only the output weights that are trained and, in contrast to deep neural networks or reservoir computers, they do not involve backpropagation or recurrence. Specifically, based on a training dataset \mathbf{H} (computed from \mathbf{X}) and corresponding target data \mathbf{Y}^T , the ELM learns a model $\mathbf{Y}^T \approx \mathbf{H}\mathbf{W}^{\text{out}}$. Importantly, \mathbf{W}^{out} can be computed in a single step according to $\mathbf{W}^{\text{out}} \approx \mathbf{H}^\dagger \mathbf{Y}^T$, where \mathbf{H}^\dagger is the Moore–Penrose pseudoinverse of \mathbf{H} . We use pseudoinverse computation here because it solves the linear regression problem directly (minimizing mean squared error between predicted and target outputs) and does not require case-by-case optimization across the wide range of parameters that we study. We discuss the effect of regularization on this regression step further in Supplement 1.

We now consider how this system is applied to handwritten digit classification. For each of the 60,000 MNIST training images, we perform (i) bicubic downsampling and flattening from a 28×28 image to a length M vector, (ii) encoding of this vector on the spectral phase (or amplitude) of a femtosecond input pulse, (iii) propagation of the encoded pulse in optical fiber, and (iv) readout of output spectra into K spectral bins after convolution with a Gaussian spectral response function and addition of a detection noise floor. The downsampling step accounts for the limited spectral bandwidth available for encoding on typical input pulses, and consistent with what would be a necessary approach in any practical system, we assume that spectral measurements are made with a single-shot basis real-time technique [20]. The underlying idea here is that input pulses encoded with different images will produce distinguishable output spectra after nonlinear propagation. Since the spectra for images corresponding to the same digit (0–9) in the training set will be expected to exhibit high-dimensional similarity in their structure, a readout step can be trained to identify these similarities and classify the corresponding digits.

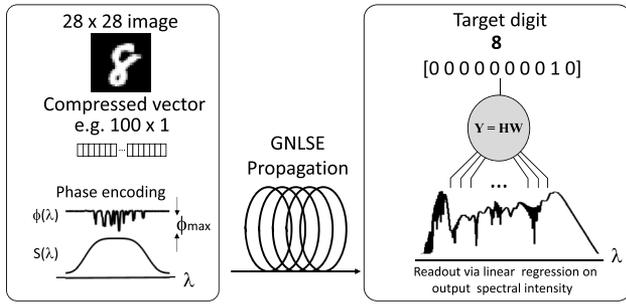


Fig. 1. Schematic of the supercontinuum ELM model showing the three steps of encoding, generalized nonlinear Schrödinger equation (GNLSSE) propagation, and spectral readout.

Specifically, after propagation and readout, the training dataset consists of a $60,000 \times K$ array of spectra, forming the hidden layer \mathbf{H} , where the propagation dynamics emulate the ELM transformation f_{NL} . The target dataset \mathbf{Y}^T is a $60,000 \times 10$ matrix of known digits (using one-hot encoding), and the $K \times 10$ weight matrix \mathbf{W}^{out} is obtained from solving $\mathbf{W}^{\text{out}} = \mathbf{H}^\dagger \mathbf{Y}^T$ using the Moore–Penrose pseudoinverse algorithm. The model accuracy is determined by applying \mathbf{W}^{out} to \mathbf{H} obtained for 10,000 test images not used in training and computing the accuracy comparing the ELM predictions and the known test digits. Accuracy associated with the training data is also typically calculated.

The propagation model is the generalized nonlinear Schrödinger equation (GNLSSE), written in dimensional form as $iA_z - 1/2 \beta_2 A_{TT} - i/6 \beta_3 A_{TTT} + 1/24 \beta_4 A_{TTTT} + \gamma(1 + i\omega_0 \partial_T)(A [R * |A|^2]) = 0$ [21]. Here, $A(z, T)$ is the complex field envelope (distance z , comoving time T), β_k are the dispersion coefficients, γ is the nonlinearity coefficient, and ω_0 is the carrier frequency. The nonlinear response function in the convolution term ($*$) is $R(t) = (1 - f_R)\delta(t) + f_R h_R(t)$, with Raman fraction $f_R = 0.18$ and h_R the experimental Raman response of fused silica [22]. Input pulse quantum noise is included via a semiclassical model [22,23] that has been found to yield quantitative agreement with experiment in reproducing supercontinuum noise characteristics [24]. We also included a Raman noise source, but this was found to have negligible influence [22,24]. For anomalous dispersion regime propagation, we consider a dispersion-shifted fiber with 1546.2 nm zero-dispersion wavelength. At a pump wavelength of 1550 nm, the parameters are $\beta_2 = -0.12 \text{ ps}^2 \text{ km}^{-1}$; $\beta_3 = 0.040 \text{ ps}^3 \text{ km}^{-1}$; $\beta_4 = 0 \text{ ps}^4 \text{ km}^{-1}$; and $\gamma = 10.7 \text{ W}^{-1} \text{ km}^{-1}$. For normal dispersion regime propagation, we consider a dispersion-flattened fiber with parameters at 1550 nm: $\beta_2 = 0.987 \text{ ps}^2 \text{ km}^{-1}$; $\beta_3 = 7.31 \times 10^{-3} \text{ ps}^3 \text{ km}^{-1}$; $\beta_4 = 4.10 \times 10^{-4} \text{ ps}^4 \text{ km}^{-1}$; and $\gamma = 7.5 \text{ W}^{-1} \text{ km}^{-1}$. These parameters correspond to a commercially available fiber. It is straightforward to also include higher-order dispersion, but these truncations yield known characteristic spectral broadening features in both dispersion regimes [22]. At the fiber lengths studied, attenuation at $\sim 0.5\%$ level is neglected. Simulations use a 2^{11} computational grid, but we can analyze the output spectra during readout using different numbers of sampling points.

Image information is encoded on the input pulses in the frequency domain on a 30 nm full width at half maximum (FWHM) second-order supergaussian spectrum centered on 1550 nm. The FWHM of the corresponding temporal intensity profile is $\Delta\tau \sim 182 \text{ fs}$ FWHM. For a particular image, the downsampled image vector of length M is scaled to a desired phase (or

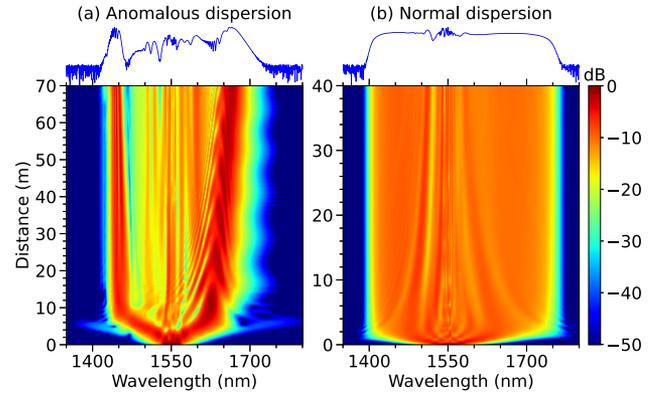


Fig. 2. Simulated spectral evolution of phase-encoded pulses for (a) anomalous and (b) normal dispersion regime dynamics.

amplitude) modulation depth ϕ_{max} . A 30 nm bandwidth allows encoding with 0.3 nm resolution, consistent with commercial Fourier-domain pulse shaping devices. After encoding, we also scale the input pulse to a particular energy at which we wish to study the dynamics.

We first show typical anomalous and normal dispersion regime spectral evolution in Figs. 2(a) and 2(b). These results use downsampling to 10×10 ($M = 100$) of one particular image, followed by phase encoding with modulation depth $\phi_{\text{max}} = 0.25\pi$. The encoding adds a low-amplitude temporal pedestal at the $\sim -50 \text{ dB}$ level, and for small ϕ_{max} , the corresponding temporal FWHM $\Delta\tau$ is unchanged from that of the unencoded pulse. As a result, it is convenient to scale the input energy so that the pulse injected in the fiber corresponds to a specified parameter $N = (\gamma P_0 T_0^2 / |\beta_2|)^{1/2}$, where time scale $T_0 \approx \Delta\tau / 1.76$. N is a characteristic measure of nonlinear strength, which corresponds to soliton number for anomalous dispersion. For anomalous dispersion regime propagation as in Fig. 2(a), input $N = 10$ corresponds to 20.7 pJ energy and $P_0 = 103 \text{ W}$ peak power, and we propagate over 70 m. For normal dispersion regime propagation as in Fig. 2(b), $N = 10$ yields $P_0 = 1215 \text{ W}$ peak power and 243 pJ energy, and we propagate over 40 m. To simulate realistic detection, the output spectrum is convolved with a 1 nm Gaussian spectral response (see Supplement 1) followed by the addition of a -30 dB random noise background to model the finite dynamic range of real-time spectral measurements [20]. The -30 dB noise background was computed relative to the peak of the mean spectral intensity of the ensemble. To generate training and testing spectra under different conditions, we define a particular parameter set and then implement these encoding and propagation steps on each of the 70,000 images.

Figures 3 and 4 explore parameter dependencies for anomalous and normal dispersion regime propagation, respectively. Specifically, Fig. 3 shows results for 10×10 downsampling, phase encoding with $\phi_{\text{max}} = 0.25\pi$, and 70 m fiber. Figure 3(a) shows how training (red) and testing (blue) accuracies vary with soliton number N over the range of 0.5–13. Figure 3(b) shows results for fixed $N = 10$, but for varying fiber lengths. Test and training results are compared with benchmark linear regression (green dashed line) based on pseudoinverse computation without fiber propagation. See Supplement 1 for associated confusion charts. These plots show the expected decrease in accuracy between training and testing and also the increase in

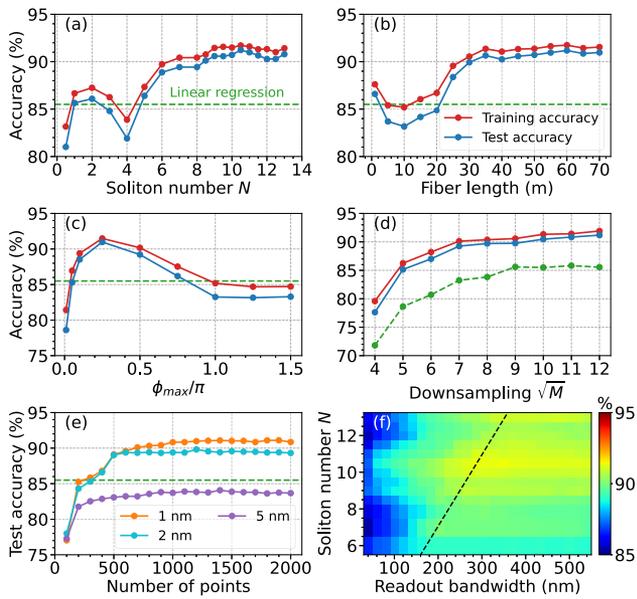


Fig. 3. Anomalous dispersion results. For 10×10 downsampling, $\phi_{\max} = 0.25\pi$, and 70 m fiber, training (red) and test (blue) accuracies are shown versus (a) soliton number N and (b) fiber length. Green dashed line: linear benchmark. (c)–(d) use the same color code. For $N = 10$ and 70 m fiber, accuracy is shown: (c) for different ϕ_{\max} and (d) for input downsampling to $\sqrt{M} \times \sqrt{M}$. (e) Test accuracy varying readout bins and convolution bandwidths as indicated. (f) False color plot of test accuracy dependence on N and readout bandwidth (around 1550 nm). Dashed line: -20 dB output bandwidth.

accuracy exceeding 90% with increasing propagation complexity (higher N and longer distance.) The decrease in accuracy seen in Figs. 3(a) and 3(b) around $N = 4$ and 10 m, respectively, is associated with the fact that we are in the onset phase of soliton fission dynamics (see Supplement 1).

Figures 3(c) and 3(d) study the aspects of input encoding for $N = 10$ and 70 m fiber. For example, for 10×10 downsampling, Fig. 3(c) shows how training (red) and testing (blue) accuracies depend on modulation depth ϕ_{\max} . There is a clear optimum around $\phi_{\max} \sim 0.25\pi$, and performance away from this point is degraded. Decreasing accuracy for lower modulation depth is expected since image information is weakly encoded and will have a limited effect on propagation. The decrease at higher modulation depth arises because greater ϕ_{\max} increasingly modifies the temporal input pulse, reducing peak power at the expense of a low-amplitude pedestal. This results in reduced spectral broadening. In Fig. 3(d), we apply optimal modulation $\phi_{\max} = 0.25\pi$, and we study how training (red) and testing (blue) accuracies depend on image downsampling, i.e., the length M vector describing a $\sqrt{M} \times \sqrt{M}$ image. We see that increasing resolution yields improved results, but test accuracy approaching 90% can be attained even with only 7×7 downsampling, representing only $\sim 6\%$ of the pixels in the original image.

The results in Figs. 3(a)–3(d) use readout over the full output spectrum (2048 points over 1317–1882 nm), convolved with a 1 nm spectral response, and with a -30 dB spectral noise floor. It is important to consider how readout parameters influence performance, and these results are shown in Figs. 3(e)–3(f). For downsampling to 10×10 , $\phi_{\max} = 0.25\pi$, and $N = 10$, Fig. 3(e)

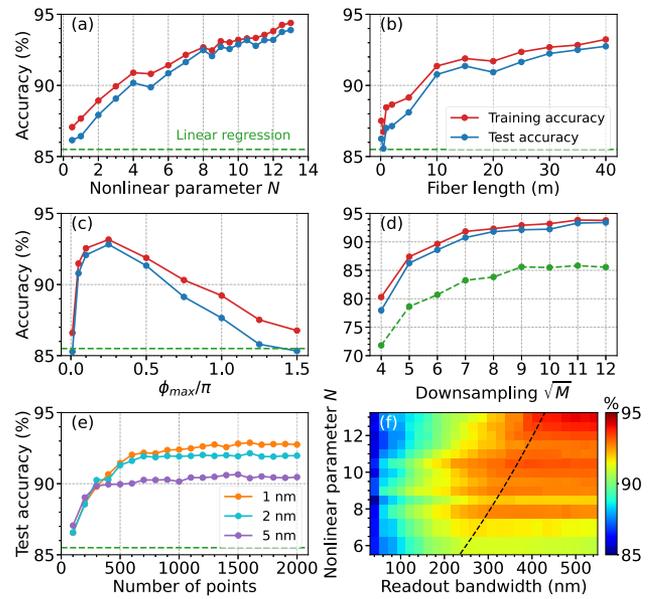


Fig. 4. Normal dispersion results. For 10×10 downsampling, $\phi_{\max} = 0.25\pi$, and 40 m fiber, training (red) and test (blue) accuracies are shown versus (a) nonlinear parameter N and (b) fiber length. Green dashed line: linear benchmark. (c)–(d) use the same color code. For $N = 10$ and 40 m fiber,

plots test accuracy reading out the spectrum over the full wavelength span, but changing the readout sampling density using different numbers (100–2000) of equispaced bins. We also compare results convolving with three different spectral responses of FWHM: 1 nm (orange), 2 nm (blue), and 5 nm (purple). Clearly, a higher sampling density yields improved accuracy, but with 1 nm resolution, $\sim 90\%$ accuracy can still be approached with only 700 bins. Figure 3(f) examines how the accuracy depends on the value of the readout bandwidth with respect to the overall output spectral bandwidth, using the same 10×10 downsampling, $\phi_{\max} = 0.25\pi$, and 1 nm convolution. The idea here is to study whether we need to read out spectral information over the full bandwidth of the output spectrum, or whether a reduced readout bandwidth is sufficient. The false color plot shows how test accuracy varies with readout bandwidth (centered on 1550 nm) while varying N over 6–13. Of course, as N increases, the bandwidth of the output spectrum will also increase, and this bandwidth (-20 dB level) is shown as the dashed line in the figure. It is clear that readout over only a portion of the output spectra (i.e., left of the dashed line) leads to reduced accuracy, but once we capture the full output spectral bandwidth (i.e., right of the dashed line), then accuracy improves (then saturates). This is consistent with the interpretation that nonlinear propagation transforms the encoded image information (initially restricted only to 30 nm around the pump) into the high-dimensional space associated with the broadened spectrum.

Figures 4(a)–4(f) show results for normal dispersion regime propagation. The overall trends and dependence on parameters such as N and fiber length are qualitatively alike, although normal dispersion regime propagation consistently shows $\sim 3\%$ improvement in training and testing accuracies. This can be attributed to the well-known observation of improved noise characteristics in the normal dispersion regime [21]. However, for both anomalous and normal dispersion, we anticipate that accuracy would ultimately degrade for increased power and/or

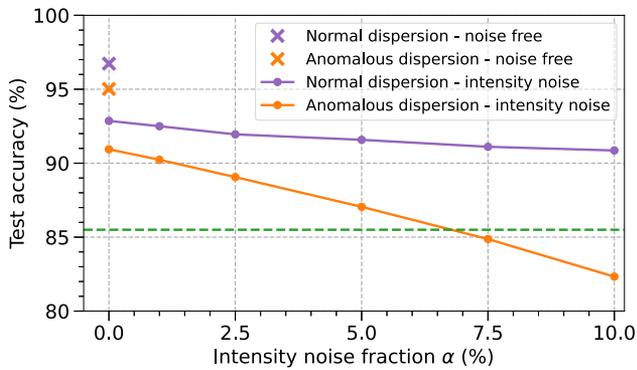


Fig. 5. Lines: Dependence of ELM test accuracy on multiplicative noise for normal (purple) and anomalous (orange) dispersion regime propagation. These results include input quantum noise and a -30 dB readout noise floor. Crosses: ideal accuracies without any input or readout noise. Green dashed line: linear regression benchmark.

fiber length as a result of deleterious effects such as incoherent supercontinuum dynamics and polarization instabilities [21].

The studies in Figs. 3 and 4 were repeated using spectral amplitude encoding on the input pulses. This involved multiplying the supergaussian input spectrum by an amplitude encoding mask with variable modulation depth. The general trends and results (see Supplement 1) were extremely similar to those obtained with phase encoding in terms of parameter dependence and classification accuracy.

In Fig. 5, we study the impact of noise on ELM performance using phase encoding for $\phi_{\max} = 0.25\pi$ and $N = 10$. Firstly, we recall that the results in Figs. 2–4 were obtained with quantum noise on the input pulses and a -30 dB readout noise background. It is straightforward with simulations to model the ELM without input quantum noise or a readout noise floor, and this allows us to study the ideal mathematical properties of the GNLSE to act as a nonlinear ELM transfer function. These results are shown as the crosses in Fig. 5 and indicate what can be considered upper limit ideal test accuracies of 96.7% and 95.0% for normal and anomalous dispersion regime propagation, respectively. Significantly, the addition of only input quantum noise (i.e., again with no imposed readout noise floor) reduces these ideal test accuracies to 94.8% and 91.9%, respectively, an important result that highlights an intrinsic quantum noise penalty for this class of fiber-based ELM. Of course, the imposed -30 dB instrumental noise floor on the output spectra reduces accuracy further, with the corresponding results in this case $\sim 93\%$ and $\sim 91\%$, respectively, for anomalous and normal dispersion regime propagation.

As might be expected, any additional input noise degrades performance. To show this, we apply a simple multiplicative intensity noise model of the form $A_{\text{in}}(T) = A_0(T)[1 + \alpha/2 \eta(T)]$, where A_0 is the temporal input field after addition of quantum noise and before spectral encoding. Parameter α corresponds to an intensity noise fraction, and $\eta(T)$ is a random variable. Figure 5 shows results applying this noise on input pulses with 10×10 downsampling, phase encoding, $\phi_{\max} = 0.25\pi$, and $N = 10$. It is clear that testing accuracy decreases with increasing

noise fraction in both anomalous (orange) and normal (purple) dispersion regimes. Results are clearly more degraded in the anomalous dispersion regime, further highlighting the noise sensitivity of anomalous dispersion regime propagation [21].

The GNLSE model framework is a powerful and flexible tool to explore nonlinear fiber propagation as applied to ELM systems and provides important insights for experimental design. Consistent with recent experiments, our results suggest that normal dispersion propagation yields improved accuracy and reduced noise sensitivity, but satisfactory results can still be obtained with anomalous dispersion regime propagation for suitable parameters [15–19]. A further key conclusion is to identify that input pulse quantum noise will likely impose an intrinsic penalty for all nonlinear fiber propagation-based ELMs.

Funding. Institut Universitaire de France; European Research Council (101044777); Research Council of Finland (368650); Agence Nationale de la Recherche (ANR-15-IDEX-0003, ANR-17-EURE-0002, ANR-20-CE30-0004).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. D. Psaltis, D. Brady, X.-G. Gu, *et al.*, *Nature* **343**, 325 (1990).
2. G. Wetzstein, A. Ozcan, S. Gigan, *et al.*, *Nature* **588**, 39 (2020).
3. M. Miscuglio, A. Mehrabian, Z. Hu, *et al.*, *Opt. Mater. Express* **8**, 3851 (2018).
4. Y. Zuo, B. Li, Y. Zhao, *et al.*, *Optica* **6**, 1132 (2019).
5. A. Jha, C. Huang, and P. R. Prucnal, *Opt. Lett.* **45**, 4819 (2020).
6. J. Liu, Q. Wu, X. Sui, *et al.*, *Photonix* **2**, 5 (2021).
7. I. Oguz, J.-L. Hsieh, N. U. Dinc, *et al.*, *Adv. Photonics* **6**, 016002 (2024).
8. N. U. Dinc, I. Oguz, M. Yildirim, *et al.*, *Opt. Mater. Express* **14**, 2413 (2024).
9. L. Larger, M. C. Soriano, D. Brunner, *et al.*, *Opt. Express* **20**, 3241 (2012).
10. F. Duport, B. Schneider, A. Smerieri, *et al.*, *Opt. Express* **20**, 22783 (2012).
11. M. Yildirim, I. Oguz, F. Kaufmann, *et al.*, *APL Photonics* **8**, 106104 (2023).
12. S. Ortín, M. C. Soriano, L. Pesquera, *et al.*, *Sci. Rep.* **5**, 14945 (2015).
13. U. Teğin, M. Yildirim, I. Oğuz, *et al.*, *Nat. Comput. Sci.* **1**, 542 (2021).
14. G. Marcucci, D. Pierangeli, and C. Conti, *Phys. Rev. Lett.* **125**, 093901 (2020).
15. B. Fischer, M. Chemnitz, Y. Zhu, *et al.*, *Adv. Sci.* **10**, 2303835 (2023).
16. K. F. Lee and M. E. Fermann, *Phys. Rev. A* **109**, 033521 (2024).
17. M. Hary, D. Brunner, L. Leybov, *et al.*, "Principles and metrics of extreme learning machines using a highly nonlinear fiber," *arXiv* (2025).
18. S. Saeed, M. Müftüoğlu, G. R. Cheeran, *et al.*, "Nonlinear inference capacity of fiber-optical extreme learning machines," *arXiv* (2025).
19. I. Muda, A. Maula, and U. Teğin, *Opt. Express* **33**, 7852 (2025).
20. M. Närhi, L. Salmela, J. Toivonen, *et al.*, *Nat. Commun.* **9**, 4923 (2018).
21. G. P. Agrawal, *Nonlinear Fiber Optics* (Elsevier, 2019).
22. J. M. Dudley, G. Genty, and S. Coen, *Rev. Mod. Phys.* **78**, 1135 (2006).
23. E. Brainis, D. Amans, and S. Massar, *Phys. Rev. A* **71**, 023808 (2005).
24. K. L. Corwin, N. R. Newbury, J. M. Dudley, *et al.*, *Phys. Rev. Lett.* **90**, 113904 (2003).