

Near-Equilibrium Propagation Training in Nonlinear Wave Systems

Karol Sajnok^{1,*} and Michał Matuszewski^{1,2,†}¹*Center for Theoretical Physics, Polish Academy of Sciences, Aleja Lotników 32/46, 02-668 Warsaw, Poland*²*Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, 02-668 Warsaw, Poland*

(Received 17 October 2025; revised 20 February 2026; accepted 4 August 2026; published 22 September 2026)

Backpropagation learning algorithm, the workhorse of modern artificial intelligence, is notoriously difficult to implement in physical neural networks. Equilibrium propagation (EP) is an alternative with comparable efficiency and strong potential for *in situ* training. We extend EP learning to both discrete and continuous complex-valued wave systems. In contrast to previous EP implementations, our scheme is valid in the weakly dissipative regime and readily applicable to a wide range of physical settings, even without well-defined nodes, where trainable internode connections can be replaced by trainable local potential. We test the method in driven-dissipative exciton-polariton condensates governed by generalized Gross-Pitaevskii dynamics. Numerical studies on standard benchmarks, including a simple logical task and handwritten-digit recognition, demonstrate stable convergence, establishing a practical route to *in situ* learning in physical systems in which system control is restricted to local parameters.

DOI: [10.1103/qd75-znm](https://doi.org/10.1103/qd75-znm)

I. INTRODUCTION

Neuromorphic learning in physical substrates promises orders-of-magnitude gains in throughput and energy efficiency by exploiting native dynamics for training and inference [1–6]. Optical and solid-state platforms provide reconfigurable couplings with fast readout, yet end-to-end *in situ* training remains challenging because backpropagation assumes an explicit computational graph and accurate adjoint dynamics, assumptions fragile in analog hardware with hidden couplings, latency, and dissipation [1,7]. Driven-dissipative wave systems, including photonic neural networks, are especially attractive due to their throughput and scalability [5–9]. While *in silico* training can work for small devices, model-experiment mismatch in larger systems motivates efficient *in situ* alternatives [10–15]. However, nonlinear complex-valued dynamics couple amplitude and phase to gain and loss, complicating accurate gradient estimation.

Equilibrium propagation (EP) is an energy-based two-phase algorithm that estimates gradients locally from contrasts between free and nudged steady states [16]. The system first relaxes under the input and then under weak output-dependent nudging; their steady-state difference yields a local update rule. EP is equivalent to recurrent backpropagation [17,18] and achieves high accuracy on large-scale digital benchmarks [18–20]. It has been extended to convolutional,

spiking, oscillator, thermodynamic, and quantum networks, as well as to Lagrangian and agnostic learning frameworks [19,21–32]. EP was demonstrated experimentally in mechanical [33], resistor [34,35], and memristor [36] networks. It was also partially implemented experimentally on a D-Wave quantum circuit [37] and optical coherent Ising machines [38,39].

Here, we introduce near-equilibrium propagation (NEP), a generalization of EP to driven-dissipative complex-wave systems (Fig. 1). Our approach extends previous results in several ways: (1) operates in the near-equilibrium regime of weak pumping and dissipation; (2) is applicable to oscillatory and wave systems, extending beyond conventional EP based on energy relaxation; (3) can train *local parameters* instead of internode weights, thus covering a broader class of physical systems; and (4) accommodates both discrete and continuous complex-valued fields. Some of these aspects were partially explored before. Generalization of EP to real-valued vector-field dynamics without a well-defined energy function was considered in Ref. [40]. Training oscillatory systems with EP was studied in Refs. [23–25], but those models assumed energy-gradient descent rather than Hamiltonian dynamics, yielding relaxation instead of oscillatory behavior. EP in complex networks was introduced in Ref. [41] for holomorphic functionals, which do not correspond to physical real-valued quantities like energy or Hamiltonian. In Ref. [42], the analogy between continuous wave systems and neural networks was recognized, but no method for *in situ* training was provided, relying instead on backpropagation in a discretized digital model.

To test the framework, we model NEP learning in exciton-polariton condensates, a platform combining light and matter [43–45]. Polaritonic neural networks have been explored in several architectures [46–53], and recent work highlights their energy efficiency for inference [6,54]. We demonstrate stable NEP training on XOR and MNIST tasks, derive local

*Contact author: ksajnok@cft.edu.pl†Contact author: mmatuszewski@cft.edu.pl

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

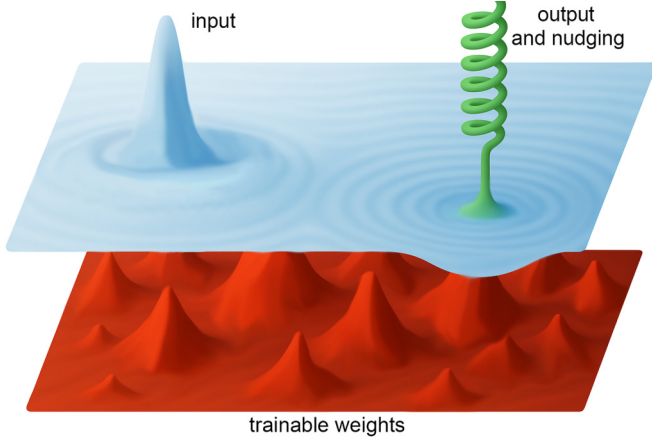


FIG. 1. Scheme of NEP implementation in a continuous wave system (light blue) with input drive, output and nudging force (green spring), and trainable potential parameters (red).

update rules from free-nudged contrasts, and model a realistic implementation using the driven-dissipative Gross-Pitaevskii equation (GPE), including pumping geometry and output nudging. Numerical results confirm convergence and indicate experimental feasibility.

II. NEAR-EQUILIBRIUM PROPAGATION IN COMPLEX-VALUED SYSTEMS

We define the NEP learning protocol for driven-dissipative systems. An oscillatory or wave system is described by a time-dependent complex field $\psi(\vec{r}, t)$ defined over a spatial domain $\vec{r} \in \mathcal{M}$, which may be continuous or discrete depending on the physical realization.

Training proceeds via a two-step procedure for each input sample $X = (x_1, x_2, \dots)$. In the first, free-evolution phase, the system follows

$$\partial_t \psi(\vec{r}, t) = \kappa_\psi(\psi(\vec{r}, t), \theta, X), \quad (1)$$

where $\theta = (\theta_1, \theta_2, \dots)$ is the vector of trainable parameters, such as local potentials, internode couplings, or input weights. Unlike in conventional EP, Eq. (1) need not derive from a variational principle with an energy functional. We assume that after sufficiently long evolution, the system reaches an oscillating steady state $\psi(\vec{r}, t) = \Psi_0(\vec{r}) e^{i\omega t}$, where ω denotes frequency, implying $\kappa_0(\Psi_0(\vec{r}), \theta, X) := \kappa_\psi e^{-i\omega t} = i\omega \Psi_0(\vec{r})$. We assume ω is locked to a resonant drive, $\omega = \omega_D = \text{const}$, independent of θ and X . Moving to the rotating frame $\Psi(\vec{r}, t) = \psi(\vec{r}, t) e^{-i\omega_D t}$ gives

$$\partial_t \Psi(\vec{r}, t) = \kappa(\Psi(\vec{r}, t), \theta, X) \stackrel{\text{s.s.}}{=} 0, \quad (2)$$

with $\kappa = \kappa_0(\Psi, \theta, X) - i\omega_D \Psi$, where the last equality in Eq. (2) is fulfilled in the steady state $\Psi(\vec{r}, t) = \Psi_0(\vec{r})$.

In the second, nudged phase, the system is perturbed with respect to the target. The goal is to tune θ to minimize the cost function $C(\Psi_0)$, quantifying output-target discrepancy. To achieve this, we add a driving term proportional to

the cost gradient

$$\partial_t \Psi = \kappa - i\beta \frac{\partial C(\Psi_0, \bar{\Psi}_0)}{\partial \bar{\Psi}_0} =: \kappa^\beta, \quad (3)$$

where β is a small real parameter and $\frac{\partial C}{\partial \bar{\Psi}}$ is a Wirtinger derivative. The perturbation includes an imaginary factor, and derivatives are taken with respect to the conjugate variable. This yields a simple *in situ* gradient-descent rule. We assume the system relaxes to a new, nudged steady state satisfying $\kappa^\beta(\Psi, \theta, X, \beta) \stackrel{\text{s.s.}}{=} 0$.

For complex-field systems, the output can depend on both amplitude and phase; here, we focus on amplitude. For instance, the mean-squared error (MSE) cost in the continuous case is

$$C_{\text{MSE}}(\Psi_0, \bar{\Psi}_0) = \frac{1}{2} \int_{\mathcal{Y}} (|\Psi_Y(\vec{r})|^2 - |\Psi_0(\vec{r})|^2)^2 d\vec{r}, \quad (4)$$

where $\mathcal{Y}$ is the output region. This choice leads, via Eq. (3), to the evolution equation

$$\kappa_{\text{MSE}}^\beta = \kappa + i\beta \Theta_{\mathcal{Y}}(|\Psi_Y|^2 - |\Psi_0|^2) \Psi_0, \quad (5)$$

with Heaviside step function $\Theta_{\mathcal{Y}}(\vec{r}) = 1$ for $\vec{r} \in \mathcal{Y}$ and 0 otherwise. Another example is the categorical cross-entropy (CCE) cost, useful for multiclass tasks, given in the discrete case by

$$C_{\text{CCE}}(\Psi_0, \bar{\Psi}_0) = - \sum_i |\Psi_Y(\vec{r}_i)|^2 \ln \sigma(|\Psi_0(\vec{r}_i)|^2), \quad (6)$$

where $\sigma(|\Psi(\vec{r}_i)|^2) = \frac{\exp(|\Psi(\vec{r}_i)|^2)}{\sum_j \exp(|\Psi(\vec{r}_j)|^2)}$ is the softmax function and the sum runs over all output nodes $\vec{r}_i \in \mathcal{Y}$ corresponding to output classes. The nudged evolution then becomes

$$\kappa_{\text{CCE}}^\beta = \kappa + i\beta \sum_i (|\Psi_Y(\vec{r}_i)|^2 - \sigma(|\Psi_0(\vec{r}_i)|^2)) \Psi_0(\vec{r}_i). \quad (7)$$

Thus, the nudged perturbation depends on the chosen cost function and is confined to $\mathcal{Y}$.

Parameter updates θ can be applied per sample, batch, or epoch. They minimize the cost along the steepest gradient

$$\Delta \theta \propto - \frac{\partial C}{\partial \theta} \Big|_{\beta=0} = i \left\langle \frac{\partial \bar{\Psi}}{\partial \beta}, \frac{\partial \bar{\kappa}}{\partial \theta} \right\rangle - i \left\langle \frac{\partial \Psi}{\partial \beta}, \frac{\partial \kappa}{\partial \theta} \right\rangle, \quad (8)$$

where $\langle f, g \rangle$ is the inner product, $\langle f, g \rangle := \int_{\mathcal{M}} d\vec{r} \bar{f}(\vec{r}) g(\vec{r})$ in the continuous case, and $\langle f, g \rangle := \sum_{i \in \mathcal{M}} \bar{f}(\vec{r}_i) g(\vec{r}_i)$ in the discrete one. The derivations for both cases are given in Appendix A, while Appendix B provides a detailed comparison with the results of Ref. [40].

NEP training generalizes [40] and holds for any complex wave equation, provided derivatives of $\kappa(\Psi, \bar{\Psi}, \theta, X)$ satisfy near-equilibrium conditions

$$\frac{\partial \kappa(\vec{r})}{\partial \bar{\Psi}(\vec{s})} = \frac{\partial \kappa(\vec{s})}{\partial \bar{\Psi}(\vec{r})}, \quad \frac{\partial \kappa(\vec{r})}{\partial \Psi(\vec{s})} \approx - \frac{\partial \bar{\kappa}(\vec{s})}{\partial \bar{\Psi}(\vec{r})}. \quad (9)$$

Here, *equilibrium* is understood more broadly than in previous EP works [16,23,24], referring to a stationary fixed point of Eq. (2). When a Hamiltonian functional $H(\Psi, \bar{\Psi})$ exists such that $\kappa(\vec{r}) \approx \{\Psi(\vec{r}), H\} = -i \frac{\partial H}{\partial \bar{\Psi}(\vec{r})}$ with $\{\cdot\}$ as the Poisson bracket, Eq. (9) is naturally satisfied as a consequence of

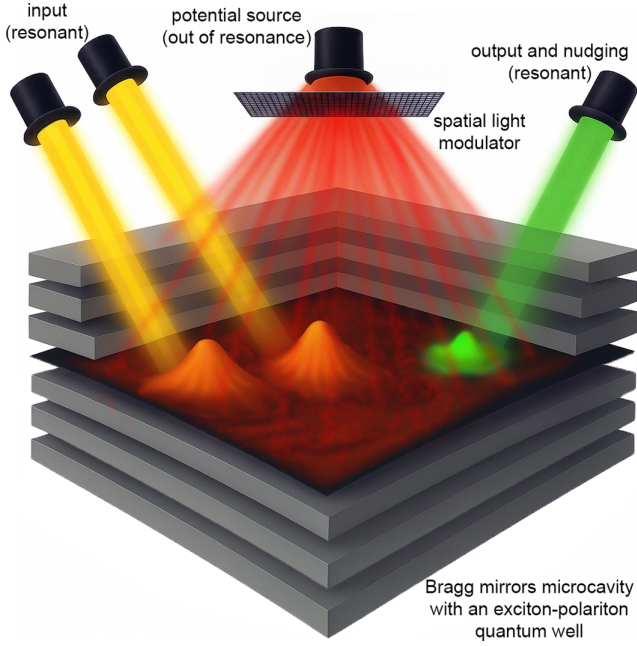


FIG. 2. Example of a polaritonic system with two inputs and one output. Two localized laser beams (yellow) encode the inputs. A spatially modulated trainable potential (red) is applied via an SLM. The selected area (green) marks the output region where optical nudging and readout occur. Input and output beams share the same frequency ω_D .

Clairaut's theorem. As shown below, even when not exactly fulfilled, training can still converge successfully.

During inference, nudging β is switched off, and outputs are read directly from the steady-state amplitudes Ψ_Y in the output region, with the system driven by input X as in the free phase.

III. IMPLEMENTATION IN EXCITON-POLARITON SYSTEMS

Exciton-polariton systems are ideal for optical neural networks owing to their strong nonlinearity, ultrafast dynamics, and high energy efficiency [6,47,48,54]. The NEP setup, shown in Fig. 2, consists of a cavity sample hosting polaritons, resonant laser beams for input pumping and output nudging, and a nonresonant laser forming the potential. Trainable parameters include both software input weights and the physical potential landscape controlled by a spatial light modulator (SLM).

For illustration, we adapt the general NEP framework to driven-dissipative polaritons described by a generalized continuous GPE with dissipation γ , complex potential $V(\vec{r})$, resonant pumping $P(\vec{r})$, and nonlinear function $f(\Psi)$ —either density response $f_d(\Psi) = g_d|\Psi|^2$ or saturation response $f_s(\Psi) = \frac{g_s}{1+|\Psi|^2}$. In the rotating frame, the polariton wave function $\Psi(\vec{r}, t)$ evolves as

$$\partial_t \Psi \equiv \kappa = -\frac{i}{\hbar} \left(-\frac{\hbar^2}{2m} \nabla^2 + V + f(\Psi) - i\gamma \right) \Psi + P, \quad (10)$$

where $P(\vec{r}) = \sum_k w_k X_k G(\vec{r} - \vec{r}_k)$, with complex input weights w_k for inputs X_k and normalized Gaussian envelopes $G(\vec{r} - \vec{r}_k) = \frac{1}{W\sqrt{2\pi}} \exp[-\frac{(\vec{r}-\vec{r}_k)^2}{2W^2}]$ of width W centered at $\vec{r}_k$. The complex potential $V(\vec{r})$ and pumping weights w_k constitute the trainable parameters θ . Although the terms P , γ , and $\text{Im } V$ break Hamiltonian dynamics, we show below that NEP remains accurate when γ and $\text{Im } V$ are small compared to the other terms (the near-equilibrium regime). The Laplacian term, being parity symmetric, satisfies both conditions in Eq. (9).

In the nudged evolution phase, the perturbation is introduced in the output region $\mathcal{Y}$, as defined in Eq. (4), depending on the chosen cost function. Experimentally, this corresponds to applying a resonant pump in $\mathcal{Y}$ with amplitude proportional to the error and a $\frac{\pi}{2}$ phase shift relative to the phase of the steady state in the output region. The steady state $\Psi_0(\vec{r})$ satisfies $\kappa^\beta(\Psi^\beta, V, w, P, \beta) = 0$.

To determine update rules for the two training parameters—the potential $V(\vec{r})$ and pumping weights $w(\vec{r}_k)$ —we employ the NEP rule [Eq. ((8)]. Using the derivative of κ with respect to the real potential $V(\vec{r})$ gives

$$\Delta V(\vec{r}) \propto -\frac{\partial \mathcal{C}}{\partial V(\vec{r})} = -\frac{1}{\hbar} \frac{\partial |\Psi_0(\vec{r})|^2}{\partial \beta}. \quad (11)$$

Analogously, for the real pump weights,

$$\Delta w_k \propto -\frac{\partial \mathcal{C}}{\partial w_k} = -2 \int d\vec{r} X_k G(\vec{r} - \vec{r}_k) \text{Im} \frac{\partial \Psi_0(\vec{r})}{\partial \beta}. \quad (12)$$

Thus, updates for both the potential and pump weights depend on local steady-state fields $\Psi_0(\vec{r}, \beta)$ obtained in the free and driven phases. The parameter β controls the nudging strength and must be small enough for accurate derivative estimation at $\beta = 0$, yet large enough to induce a measurable change in the wave function.

IV. NUMERICAL RESULTS

In simulations, we use a discretized version of the GPE and implement the two-phase NEP algorithm on a finite spatial grid. In dimensionless units, the discretized GPE reads

$$\kappa_i = \frac{i}{2}(\Psi_{i+1} - 2\Psi_i + \Psi_{i-1}) - i(V_i + f(\Psi_i) - i\gamma)\Psi_i + P_i. \quad (13)$$

Dirichlet boundary conditions are imposed, realizable experimentally via a steep out-of-resonant potential wall, though the framework is independent of boundary choice. The evolution equation is integrated using a fourth-order Runge-Kutta (RK4) scheme.

As a proof of principle, we train a one-dimensional (1D) polariton network to implement the two-input XOR function. The simulation employs a nine-node grid, with two input nodes and one output node $\mathcal{Y}$. Figure 3 shows the system geometry and optimization results. Input samples X correspond to four binary combinations (00, 01, 10, 11) and are encoded as pump amplitudes $P_i = w_i(\delta_i^1 X_1 + \delta_i^2 X_2)$, where w_i are real input weights and X_k ($k = 1, 2$) are the inputs located at sites i_k . The target pattern Ψ_Y requires high amplitude ($|\Psi_Y|^2 = 1$) for (01, 10) and low amplitude ($|\Psi_Y|^2 = 0$) otherwise. The

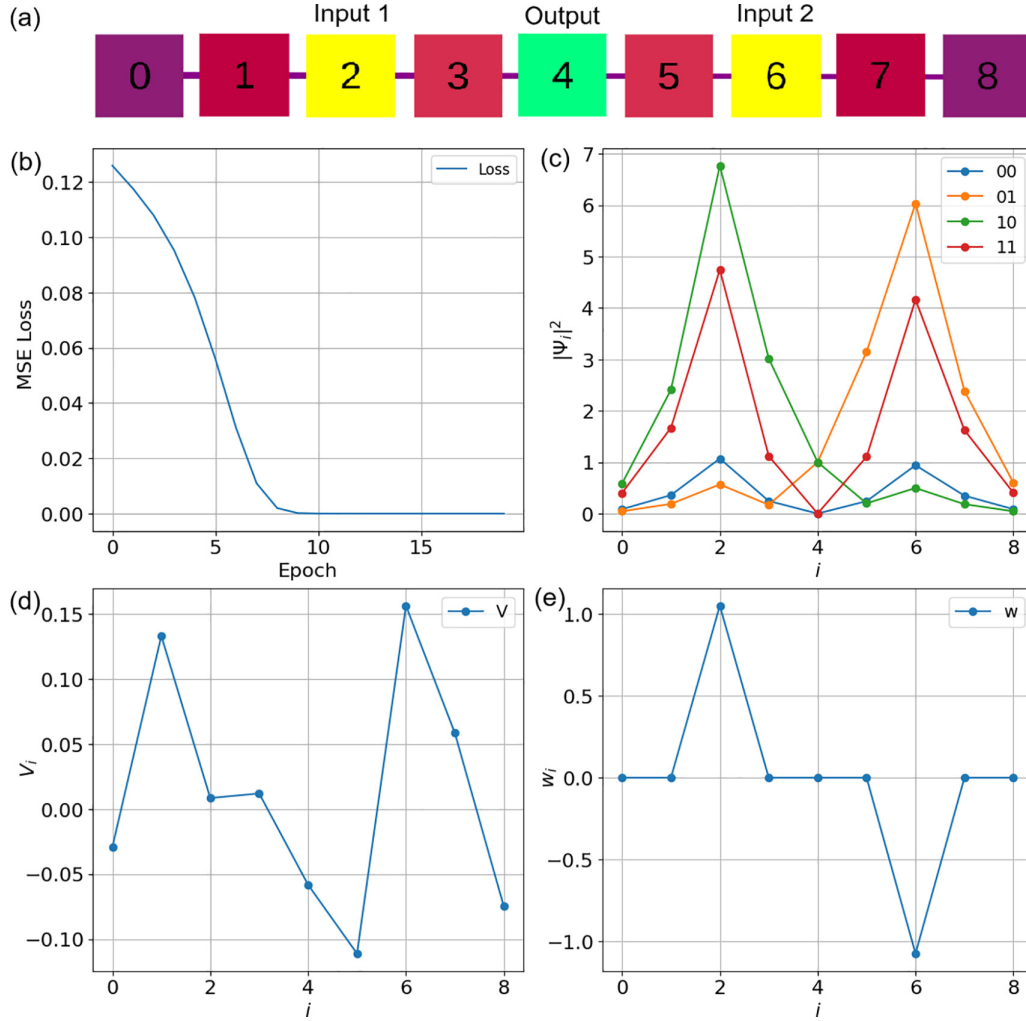


FIG. 3. Training a nine-node 1D network for the two-input XOR task. Inputs are applied symmetrically at sites 2 and 6, and the output is read at site 4. (a) Both potential V and pump weights w are trained. (b) Training loss, (c) steady-state field magnitudes $|\Psi_i|^2$ for each input, (e) trained potential V , and (d) pump weights w . Final outputs: $|\Psi_4|^2 = (0.00, 0.92, 1.06, 0.01)$ for inputs (00, 01, 10, 11).

cost function is the MSE between output intensity and target intensity for each input.

As shown in Fig. 3(b), convergence occurs after about ten epochs with nudging parameter $\beta = 0.01$. Both potential V and pump weights w are trained following Eqs. (11) and (12), using learning rates $\text{lr}_V = \text{lr}_w = 0.1$. The GPE parameters are $g = 0.1$, $\gamma = 0.1$, and nonlinear saturation response $f_s(\Psi_i) = \frac{g_s}{1+|\Psi_i|^2}$. Other panels of Fig. 3 show steady-state amplitudes, potential, and input weights in the trained system. Appendix B demonstrates that optimization remains effective regardless of input/output node positions and can rely on training both V_i and w_i , or only V_i with fixed w_i . The method is robust to random potential perturbations—e.g., structural inhomogeneities in experimental samples—verified by training with fixed small random fluctuations, which did not affect convergence.

To demonstrate the capability of the NEP algorithm to train multiclass classification tasks, we trained a two-dimensional (2D) discrete polariton network to classify the MNIST handwritten digits dataset (1000 samples per digit) (see Fig. 4). The simulation is performed on a 15×150 grid, where the

input region $\mathcal{X}$ corresponds to the entire network area, with each of ten 15×15 cells receiving a copy of the same input. Input patterns are generated by resizing and normalizing MNIST images. The output region $\mathcal{Y}$ consists of ten central nodes, each representing one class. The output pattern Ψ_Y is one-hot encoded, with each output node required to exhibit high intensity [$\sigma(|\Psi_Y|^2) = 1$] for the correct digit and low intensity [$\sigma(|\Psi_Y|^2) = 0$] for all others, with the cost function given by Eq. (6). Convergence to a test accuracy of $\approx 90.3\%$ is achieved after about 20 training epochs with nudging parameter $\beta = 0.01$, learning rates $\text{lr}_V = \text{lr}_w = 0.1$, resonant pumping amplitude normalized to unity, GPE parameters $g = 0.04$, $\gamma = 0.07$, and density-response nonlinearity $f_d(\Psi_i) = |\Psi_i|^2$. Notably, the pumping weights shown in Fig. 4 form patterns that resemble handwritten digits, recognizable to the human eye.

Varying the nonlinearity g reveals optimal performance for a moderate value $g = 0.04$, corresponding to the mean nonlinearity-to-loss ratio $\frac{\langle g|\Psi|^2 \rangle}{\hbar\gamma} \approx 0.3$, which yields the lowest final loss and highest validation accuracy of $\approx 91\%$ (see Fig. 5). Larger values of g tend to destabilize learning.

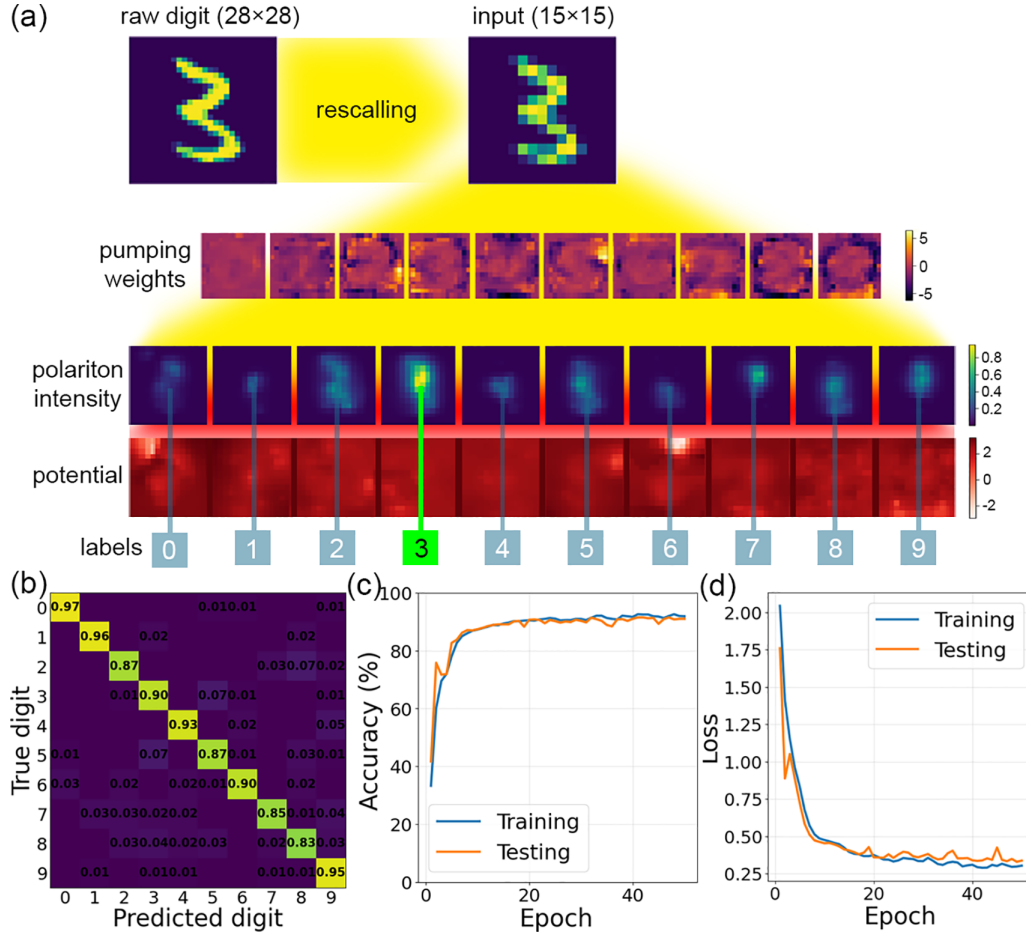


FIG. 4. Training a 15×150 2D network to classify handwritten MNIST digits. Each of the ten 15×15 cells receives a rescaled input image. The output region $\mathcal{V}$ contains ten central nodes corresponding to digits 0–9. (a) Data flow, including rescaling, replication, optical injection with pumping weights, nonresonant potential, and final polariton intensity. The cells are separated by a blocking potential (dark red). (b) Confusion matrix with final test-set accuracies. (c), (d): training and testing accuracy and loss.

Training only the potential, without optimizing the pumping weights, reduces performance but confirms that optimization can proceed via purely optical, rather than electronic, updates (Appendix C). A benchmark fully connected linear network with no hidden neurons and ten outputs, trained by

backpropagation, achieves $\approx 89.9\%$ accuracy on the same MNIST test subset. Previous studies reported that oscillatory networks with local connections perform poorly on MNIST [24]. We attribute the improved performance to nonequilibrium wave dynamics, which overcome the locality of node couplings.

Appendix C presents additional simulation results for the XOR and MNIST tasks, demonstrating the robustness of NEP to restricted parameter training and variations in network architecture.

V. EXPERIMENTAL OUTLOOK AND SUMMARY

The NEP protocol can be experimentally realized with current polariton systems. Numerical simulations for the MNIST task (1000 samples per digit) converge after about 20 epochs for $\gamma = 0.1$, requiring two steady-state measurements per sample, each with ≈ 5000 RK4 steps and $dt = 0.1$. In physical units, the full training procedure completes within $T_{\text{phys}} \approx 0.1\text{--}10$ ms for typical $\gamma^{-1} \approx 1\text{--}100$ ps [43,55], nearly 6 orders of magnitude faster than GPU-based numerical integration (NVIDIA GeForce RTX 2080 Ti, 11 GB). Thus, NEP provides both conceptual and substantial practical advantages. For GaAs-based microcavity [56], a typical linewidth

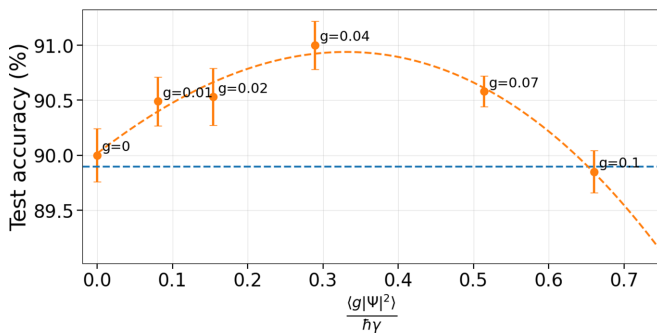


FIG. 5. Mean test set accuracy over the last 10 out of 50 epochs for the setup of Fig. 4, plotted as a function of the mean nonlinearity-to-loss ratio. The dashed orange line shows a third-order polynomial fit, whereas the dashed blue line marks the accuracy of a fully connected linear network trained by backpropagation on the same inputs.

$\hbar\gamma \approx 0.1$ meV and interaction constant $g = 1 \mu\text{eV} \mu\text{m}^2$ [56], the numerical parameters used in the simulations shown in Fig. 4 correspond to exciton density $|\Psi|^2 \approx 30 \mu\text{m}^{-2}$. The trained parameters correspond to coupling energy of ≈ 1 meV, optically induced potential of ≈ 0.4 meV, and resonant pumping amplitude of ≈ 0.8 meV, all well within experimentally accessible range.

We have presented the NEP scheme, which transforms a nonlinear wave system into a machine-learning device trained via steady-state contrasts. Unlike previous approaches, our algorithm operates near equilibrium, trains local parameters instead of internode weights, and applies to both discrete and continuous systems. The protocol uses standard optical tools (SLMs and camera-based readouts), is robust to moderate inhomogeneities, and naturally supports tiling into weakly coupled cells for large-scale tasks. Altogether, these features make NEP a realistic framework for laboratory implementation, enabling *in situ*, ultrafast, energy-efficient training on polaritonic hardware within experimental reach.

Note added. Recently, Scurria *et al.* reported an EP scheme for nonequilibrium steady states formulated in a real-valued state space [57], whereas our NEP targets driven-dissipative polariton networks with complex fields and implements nudging via Wirtinger derivatives.

ACKNOWLEDGMENTS

We thank Andrzej Opala for useful discussions. The authors acknowledge support from the European Union EIC Pathfinder project “Neuromorphic Polariton Accelerator” (PolArt, No. 101130304) and from the National Science Center, Poland (Grant No. 2021/43/B/ST3/00752).

DATA AVAILABILITY

The data that support the findings of this article are openly available [58].

APPENDIX A: DERIVATION OF THE UPDATE RULE

Continuous variable case. We derive the update rule for a trainable parameter $\theta(\vec{r})$ in an arbitrary complex wave system. Using the chain rule for functional derivatives, the derivative of the cost function $C(\Psi, \bar{\Psi})$ with respect to $\theta(\vec{r})$ reads

$$\frac{\partial C}{\partial \theta(\vec{r})} = \int d\vec{s} \left(\frac{\partial C}{\partial \Psi(\vec{s})} \frac{\partial \Psi(\vec{s})}{\partial \theta(\vec{r})} + \frac{\partial C}{\partial \bar{\Psi}(\vec{s})} \frac{\partial \bar{\Psi}(\vec{s})}{\partial \theta(\vec{r})} \right), \quad (\text{A1})$$

where integration is over the spatial domain $\mathcal{M}$, and $\Psi = \Psi(\theta, \bar{\theta})$, $\bar{\Psi} = \bar{\Psi}(\theta, \bar{\theta})$ are steady-state fields depending implicitly on θ and $\bar{\theta}$.

From Eq. (3),

$$\frac{\partial C}{\partial \Psi} = -i \frac{\partial \bar{\kappa}^\beta}{\partial \beta}, \quad (\text{A2})$$

so Eq. (A1) becomes

$$\frac{\partial C}{\partial \theta(\vec{r})} = -i \int d\vec{s} \left(\frac{\partial \bar{\kappa}^\beta}{\partial \beta} \frac{\partial \Psi(\vec{s})}{\partial \theta(\vec{r})} - \frac{\partial \kappa^\beta}{\partial \beta} \frac{\partial \bar{\Psi}(\vec{s})}{\partial \theta(\vec{r})} \right), \quad (\text{A3})$$

since $\frac{\partial C}{\partial \bar{\Psi}} = \overline{\frac{\partial C}{\partial \Psi}}$ for real-valued C .

In the nudged steady state, $\kappa^\beta(\Psi^\beta, \bar{\Psi}^\beta, \theta, \bar{\theta}, \beta) = 0$, where Ψ^β and $\bar{\Psi}^\beta$ depend explicitly on $\theta, \bar{\theta}, \beta$. Treating these as independent variables and differentiating with respect to β yields

$$\frac{\partial \kappa^\beta(\vec{s})}{\partial \beta} = - \int d\vec{u} \left(\frac{\partial \kappa^\beta(\vec{s})}{\partial \Psi^\beta(\vec{u})} \frac{\partial \Psi^\beta(\vec{u})}{\partial \beta} + \frac{\partial \kappa^\beta(\vec{s})}{\partial \bar{\Psi}^\beta(\vec{u})} \frac{\partial \bar{\Psi}^\beta(\vec{u})}{\partial \beta} \right), \quad (\text{A4})$$

evaluated at $\beta = 0$.

Similarly, differentiating the stationary condition with respect to $\theta(\vec{r})$ gives

$$\frac{\partial \kappa^\beta(\vec{u})}{\partial \theta(\vec{r})} = - \int d\vec{s} \left(\frac{\partial \kappa^\beta(\vec{u})}{\partial \Psi^\beta(\vec{s})} \frac{\partial \Psi(\vec{s})}{\partial \theta(\vec{r})} + \frac{\partial \kappa^\beta(\vec{u})}{\partial \bar{\Psi}^\beta(\vec{s})} \frac{\partial \bar{\Psi}(\vec{s})}{\partial \theta(\vec{r})} \right), \quad (\text{A5})$$

also evaluated at $\beta = 0$. These relations describe how variations in β or θ propagate through steady-state fields via the system’s linearized response near equilibrium.

Using Eq. (A4), we expand Eq. (A3), exchange integration order, and regroup terms by their β derivatives

$$\begin{aligned} \frac{\partial C}{\partial \theta(\vec{r})} = & i \int d\vec{u} \int d\vec{s} \\ & \times \left[\left(\frac{\partial \bar{\kappa}^\beta(\vec{s})}{\partial \Psi^\beta(\vec{u})} \frac{\partial \Psi(\vec{s})}{\partial \theta(\vec{r})} - \frac{\partial \kappa^\beta(\vec{s})}{\partial \Psi^\beta(\vec{u})} \frac{\partial \bar{\Psi}(\vec{s})}{\partial \theta(\vec{r})} \right) \frac{\partial \Psi^\beta(\vec{u})}{\partial \beta} \right. \\ & \left. + \left(\frac{\partial \bar{\kappa}^\beta(\vec{s})}{\partial \bar{\Psi}^\beta(\vec{u})} \frac{\partial \Psi(\vec{s})}{\partial \theta(\vec{r})} - \frac{\partial \kappa^\beta(\vec{s})}{\partial \bar{\Psi}^\beta(\vec{u})} \frac{\partial \bar{\Psi}(\vec{s})}{\partial \theta(\vec{r})} \right) \frac{\partial \bar{\Psi}^\beta(\vec{u})}{\partial \beta} \right]. \end{aligned} \quad (\text{A6})$$

Assuming that derivatives of $\bar{\kappa}^\beta$ with respect to Ψ^β are invariant under spatial transposition and that the near-equilibrium conditions hold,

$$\frac{\partial \bar{\kappa}^\beta(\vec{s})}{\partial \Psi^\beta(\vec{u})} = \frac{\partial \bar{\kappa}^\beta(\vec{u})}{\partial \Psi^\beta(\vec{s})}, \quad \frac{\partial \kappa^\beta(\vec{s})}{\partial \Psi^\beta(\vec{u})} \approx - \frac{\partial \bar{\kappa}^\beta(\vec{u})}{\partial \bar{\Psi}^\beta(\vec{s})}, \quad (\text{A7})$$

comparison of Eqs. (A5) and (A6) gives

$$\frac{\partial C}{\partial \theta(\vec{r})} \approx -i \int d\vec{u} \left(\frac{\partial \bar{\kappa}^\beta(\vec{u})}{\partial \theta(\vec{r})} \frac{\partial \Psi(\vec{u})}{\partial \beta} - \frac{\partial \kappa^\beta(\vec{u})}{\partial \theta(\vec{r})} \frac{\partial \bar{\Psi}(\vec{u})}{\partial \beta} \right), \quad (\text{A8})$$

evaluated at $\beta = 0$.

Gross-Pitaevskii equation. Returning to the polariton GPE, we now write the update rule explicitly. Condition (A7) holds when the imaginary part of the potential $V(\vec{r})$ and the dissipation parameter γ are negligible (near-equilibrium regime), since the Liouvillian is spatially symmetric and the diagonal Jacobians are

$$\frac{\partial \kappa^\beta(\vec{u})}{\partial \Psi(\vec{s})} = \left(\frac{i\hbar}{2m} \nabla^2 - i \frac{V}{\hbar} - i \frac{2g}{\hbar} |\Psi(\vec{u})|^2 - \frac{\gamma}{\hbar} \right) \delta(\vec{s} - \vec{u}), \quad (\text{A9})$$

$$\frac{\partial \bar{\kappa}^\beta(\vec{u})}{\partial \bar{\Psi}(\vec{s})} = - \left(\frac{i\hbar}{2m} \nabla^2 - i \frac{\bar{V}}{\hbar} - i \frac{2g}{\hbar} |\Psi(\vec{u})|^2 + \frac{\gamma}{\hbar} \right) \delta(\vec{s} - \vec{u}), \quad (\text{A10})$$

where $\delta(\vec{s} - \vec{u})$ is the Dirac delta function. For an arbitrary nonlinearity $f(\Psi)$, condition (A7) remains valid when

$$f(\Psi) + \Psi \frac{\partial f(\Psi)}{\partial \Psi} = \bar{f}(\Psi) + \bar{\Psi} \frac{\partial \bar{f}(\Psi)}{\partial \bar{\Psi}}. \quad (\text{A11})$$

Discrete variable case. In the discretized system, the derivation simplifies. Define the state vector $s = (\Psi_i, \bar{\Psi}_i)$ and evolution vector $K = (-i\kappa_i^\beta, i\bar{\kappa}_i^\beta)$. Using Eq. (A2) and $K(s, \theta, \beta) \equiv 0$, we find

$$\Delta\theta_i \propto -\frac{\partial C}{\partial \theta_i} = -\frac{\partial C}{\partial s_j} \frac{\partial s_j}{\partial \theta_i} = \frac{\partial \bar{K}_j}{\partial \beta} \frac{\partial s_j}{\partial \theta_i} = -\frac{\partial \bar{K}_j}{\partial s_k} \frac{\partial s_k}{\partial \beta} \frac{\partial s_j}{\partial \theta_i}.$$

Under transposition invariance and the near-equilibrium condition, which now have the compact form

$$\frac{\partial \bar{K}_j}{\partial s_k} \approx \frac{\partial \bar{K}_k}{\partial s_j}, \quad (\text{A12})$$

and using again $K(s, \theta, \beta) \equiv 0$, we obtain

$$\begin{aligned} \Delta\theta_i &\propto -\frac{\partial \bar{K}_k}{\partial s_j} \frac{\partial s_k}{\partial \beta} \frac{\partial s_j}{\partial \theta_i} = \frac{\partial \bar{K}_k}{\partial \theta_i} \frac{\partial s_k}{\partial \beta} \\ &= i \frac{\partial \bar{\kappa}_k}{\partial \theta_i} \frac{\partial \Psi_k}{\partial \beta} - i \frac{\partial \kappa_k}{\partial \theta_i} \frac{\partial \bar{\Psi}_k}{\partial \beta}. \end{aligned} \quad (\text{A13})$$

APPENDIX B: COMPARISON WITH THE VECTOR-FIELD EQUILIBRIUM PROPAGATION

In this Appendix, we clarify the difference between the NEP training principle and the vector-field equilibrium-propagation result of Scellier *et al.* [40], applied separately to the real and imaginary parts of the wave function.

A direct application of the update rule of Ref. [40] to the real and imaginary parts of the wave function does not lead to a useful update rule for complex wave systems. To demonstrate this, we provide below a step-by-step application of the vector-field equilibrium-propagation scheme in the real and imaginary bases.

We consider the dimensionless GPE as a simple example of a nonlinear wave equation,

$$i\partial_t \Psi = (-\nabla^2 + V + f(\Psi) - i\gamma)\Psi + iP. \quad (\text{B1})$$

In the real and imaginary bases, it reads

$$\begin{aligned} \partial_t \begin{pmatrix} \psi_R \\ \psi_I \end{pmatrix} &= \begin{pmatrix} \kappa_R \\ \kappa_I \end{pmatrix} \\ &= \begin{pmatrix} -\gamma\psi_R + (-\nabla^2 + V + f(\psi))\psi_I + P_R \\ -\gamma\psi_I - (-\nabla^2 + V + f(\psi))\psi_R + P_I \end{pmatrix}, \end{aligned} \quad (\text{B2})$$

where we assumed $V \in \mathbb{R}$. Notice the sign difference between the real and imaginary mixing terms. This structure is characteristic of wave equations and differs from the symmetric coupling matrix assumed in Ref. [40].

We denote $s = (\psi_R, \psi_I)$, $K = (\kappa_R, \kappa_I)$. The condition in Ref. [40],

$$\frac{\partial C}{\partial s} = -\frac{\partial K}{\partial \beta}, \quad (\text{B3})$$

leads to the nudging prescription

$$\kappa_R^\beta = \kappa_{0R} - \beta \frac{\partial C}{\partial \psi_R}, \quad (\text{B4})$$

$$\kappa_I^\beta = \kappa_{0I} - \beta \frac{\partial C}{\partial \psi_I}. \quad (\text{B5})$$

Following the corresponding derivation in Ref. [40], the update rule requires a symmetric Jacobian matrix,

$$\frac{\partial K}{\partial s} = \left(\frac{\partial K}{\partial s} \right)^T, \quad (\text{B6})$$

or, explicitly,

$$\begin{pmatrix} \frac{\partial \kappa_R}{\partial \psi_R} & \frac{\partial \kappa_R}{\partial \psi_I} \\ \frac{\partial \kappa_I}{\partial \psi_R} & \frac{\partial \kappa_I}{\partial \psi_I} \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial \kappa_R}{\partial \psi_R} \right)^T & \left(\frac{\partial \kappa_I}{\partial \psi_R} \right)^T \\ \left(\frac{\partial \kappa_R}{\partial \psi_I} \right)^T & \left(\frac{\partial \kappa_I}{\partial \psi_I} \right)^T \end{pmatrix}. \quad (\text{B7})$$

However, due to the sign difference between the real and imaginary parts of the equation,

$$\frac{\partial \kappa_R}{\partial \psi_I} \neq \left(\frac{\partial \kappa_I}{\partial \psi_R} \right)^T \quad (\text{B8})$$

and therefore

$$\frac{\partial K}{\partial s} \neq \left(\frac{\partial K}{\partial s} \right)^T, \quad (\text{B9})$$

even when the evolution equation contains only symmetric differential operators, such as the Laplacian.

In contrast, in the NEP training scheme, the nudging can be written in the real and imaginary bases as

$$\kappa_R^\beta = \kappa_{0R} + \beta \frac{\partial C}{\partial \psi_I}, \quad (\text{B10})$$

$$\kappa_I^\beta = \kappa_{0I} - \beta \frac{\partial C}{\partial \psi_R}. \quad (\text{B11})$$

This prescription is different from Eq. (B5) and results in a symmetric Jacobian matrix in the case of a conservative wave equation.

Physically, in contrast to the original vector-field equilibrium-propagation prescription of Ref. [40], the output is not nudged directly toward the target. Instead, the real part is nudged according to the difference between the imaginary part of the output and the corresponding target, whereas the imaginary part is nudged in the opposite direction according to the difference between the real part of the output and its target. In other words, each component is nudged according to the error associated with its conjugate variable.

The NEP nudging prescription therefore physically and mathematically differs from the vector-field equilibrium-propagation rule introduced in Ref. [40].

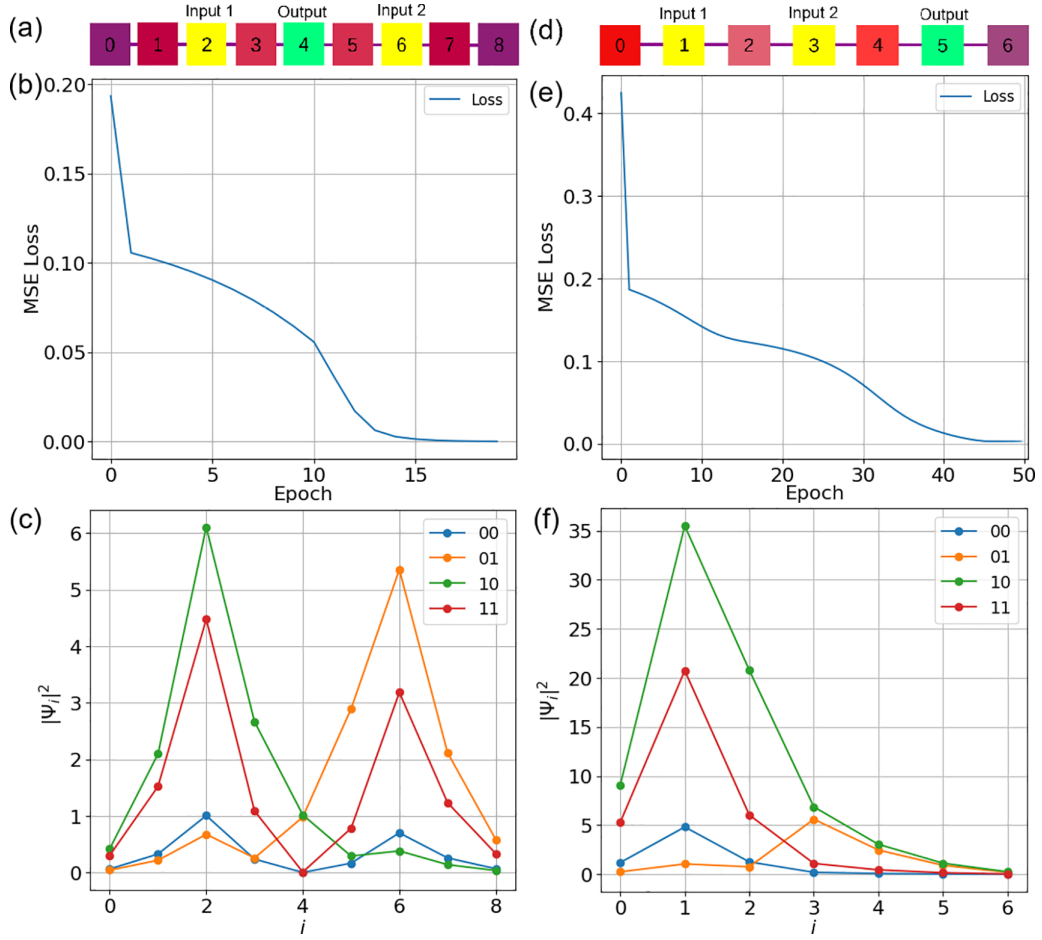


FIG. 6. XOR task in (a) a nine-node 1D network with training restricted to V , fixed near-optimal pump weights w , and (d) a seven-node network with asymmetric input placement (sites 1 and 3) and output at site 5. Top panels: Training losses [panels (b) and panel (e)] for panels (a) and (d), respectively. Bottom: Steady-state field magnitudes $|\Psi|^2$ for each input [panels (c) and (f)]. Panels (a)–(c) show that V -only training solves XOR for near-optimal fixed w ; panels (d)–(f) show successful training despite broken spatial symmetry.

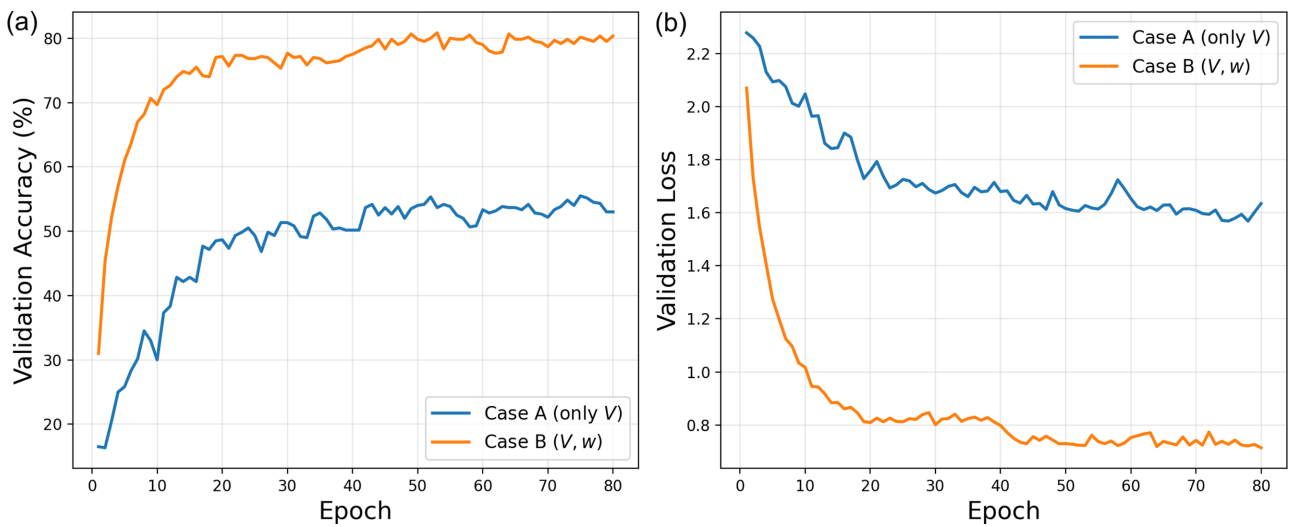


FIG. 7. Impact of the choice of trainable parameters, V and w , on (a) validation accuracies and (b) CCE losses vs epochs during training of a 5×50 -node 2D network on MNIST using NEP for 50 epochs. Classification achieves $\approx 80\%$ accuracy when both V and w are trainable, and $\approx 53\%$ when w is fixed to unities.

APPENDIX C: ADDITIONAL SIMULATION RESULTS

XOR task. Figures 6(a)–6(c) show XOR training on the same nine-node architecture as in the main text, but with optimization limited to the potential V . Pump weights w are fixed near optimal ($w_2 = 1$, $w_6 = -0.75$), serving as static input encoding. Even under this restriction, the network minimizes loss and reproduces correct outputs, confirming that proper input encoding via w enables V -only training, though convergence saturates later than with both V and w optimized. Figures 6(d)–6(f) present a reduced seven-node network with asymmetric input placement (sites 1 and 3) and

output at site 5, where both V and w are trained. Results show that spatial symmetry of inputs is unnecessary—The network still performs the XOR mapping under asymmetric geometry. This illustrates NEP’s robustness to architectural variation.

MNIST task. Figure 7 presents classification of the MNIST dataset (after principal component analysis) using a 5×50 polariton grid with ten 5×5 input cells and ten output nodes. Training both V and w yields $\approx 80\%$ test accuracy, while training only V reduces accuracy to $\approx 53\%$, still demonstrating purely optical optimization.

-
- [1] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, Physics for neuromorphic computing, *Nat. Rev. Phys.* **2**, 499 (2020).
 - [2] V. K. Sangwan and M. C. Hersam, Neuromorphic nanoelectronic materials, *Nat. Nanotechnol.* **15**, 517 (2020).
 - [3] J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, Neuromorphic spintronics, *Nat. Electron.* **3**, 360 (2020).
 - [4] D. Kudithipudi, C. Schuman, C. M. Vineyard, T. Pandit, C. Merkel, R. Kubendran, J. B. Aimone, G. Orchard, C. Mayr, R. Benosman, *et al.*, Neuromorphic computing at scale, *Nature (London)* **637**, 801 (2025).
 - [5] B. J. Shastri, A. N. Tait, T. F. de Lima, W. H. Pernice, H. Bhaskaran, C. D. Wright, and P. R. Prucnal, Photonics for artificial intelligence and neuromorphic computing, *Nat. Photon.* **15**, 102 (2021).
 - [6] M. Matuszewski, A. Prystupki, and A. Opala, Role of all-optical neural networks, *Phys. Rev. Appl.* **21**, 014028 (2024).
 - [7] A. Momeni, B. Rahmani, B. Scellier, L. G. Wright, P. L. McMahon, C. C. Wanjura, Y. Li, A. Skalli, N. G. Berloff, T. Onodera, *et al.*, Training of physical neural networks, *Nature (London)* **645**, 53 (2025).
 - [8] P. L. McMahon, The physics of optical computing, *Nat. Rev. Phys.* **5**, 717 (2023).
 - [9] G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. Miller, and D. Psaltis, Inference in artificial intelligence with deep optics and photonics, *Nature (London)* **588**, 39 (2020).
 - [10] T. W. Hughes, M. Minkov, Y. Shi, and S. Fan, Training of photonic neural networks through *in situ* backpropagation and gradient measurement, *Optica* **5**, 864 (2018).
 - [11] M. Nakajima, K. Inoue, K. Tanaka, Y. Kuniyoshi, T. Hashimoto, and K. Nakajima, Physical deep learning with biologically inspired training method: Gradient-free approach for physical hardware, *Nat. Commun.* **13**, 7847 (2022).
 - [12] L. G. Wright, T. Onodera, M. M. Stein, T. Wang, D. T. Schachter, Z. Hu, and P. L. McMahon, Deep physical neural networks trained with backpropagation, *Nature (London)* **601**, 549 (2022).
 - [13] I. Oguz, J. Ke, Q. Weng, F. Yang, M. Yildirim, N. U. Dinc, J.-L. Hsieh, C. Moser, and D. Psaltis, Forward-forward training of an optical neural network, *Opt. Lett.* **48**, 5249 (2023).
 - [14] S. Pai, Z. Sun, T. W. Hughes, T. Park, B. Bartlett, I. A. Williamson, M. Minkov, M. Milanizadeh, N. Abebe, F. Morichetti, *et al.*, Experimentally realized *in situ* backpropagation for deep learning in photonic neural networks, *Science* **380**, 398 (2023).
 - [15] A. Momeni, B. Rahmani, M. Malléjac, P. Del Hougne, and R. Fleury, Backpropagation-free training of deep physical neural networks, *Science* **382**, 1297 (2023).
 - [16] B. Scellier and Y. Bengio, Equilibrium propagation: Bridging the gap between energy-based models and backpropagation, *Front. Comput. Neurosci.* **11**, 24 (2017).
 - [17] B. Scellier and Y. Bengio, Equivalence of equilibrium propagation and recurrent backpropagation, *Neural Comput.* **31**, 312 (2019).
 - [18] M. Ernoult, J. Grollier, D. Querlioz, Y. Bengio, and B. Scellier, Updates of equilibrium prop match gradients of backprop through time in an RNN with Static Input, in *Advances in Neural Information Processing Systems (NeurIPS)* (Curran Associates, Inc., Red Hook, NY, 2019), Vol. 32.
 - [19] A. Laborieux, M. Ernoult, B. Scellier, Y. Bengio, J. Grollier, and D. Querlioz, Scaling equilibrium propagation to deep ConvNets by drastically reducing its gradient estimator bias, *Front. Neurosci.* **15**, 633674 (2021).
 - [20] T. Kerjan, R. Høier, and B. Scellier, Training a predictive coding network on ImageNet using equilibrium propagation, *arXiv:2606.03584*.
 - [21] E. Martin, M. Ernoult, J. Laydevant, S. Li, D. Querlioz, T. Petrisor, and J. Grollier, EqSpike: Spike-driven equilibrium propagation for neuromorphic implementations, *iScience* **24**, 102222 (2021).
 - [22] L. S. Ricco, E. Imos, A. Opala, and H. Sigurðsson, Training polariton Stuart-Landau oscillatory neural networks via equilibrium propagation, *Conference talk at 19th International Conference on Optics of Excitons in Confined Systems (OECS 19)*, Warsaw (2025).
 - [23] T. Rageau and J. Grollier, Training and synchronizing oscillator networks with equilibrium propagation, *Neuromorph. Comput. Eng.* **5**, 034008 (2025).
 - [24] Q. Wang, C. C. Wanjura, and F. Marquardt, Training coupled phase oscillators as a neuromorphic platform using equilibrium propagation, *Neuromorph. Comput. Eng.* **4**, 034014 (2024).
 - [25] G. Zoppo, F. Marrone, M. Bonnin, and F. Corinto, Equilibrium propagation and (memristor-based) oscillatory neural networks, in *2022 IEEE International Symposium on Circuits and Systems (ISCAS), Austin, TX, USA, 2022* (IEEE, Piscataway, NJ, 2022), pp. 639–643.

- [26] B. Scellier, Quantum equilibrium propagation: Gradient-descent training of quantum systems, *APL Quantum* **3**, 026117 (2026).
- [27] S. Massar and B. M. Mognetti, Equilibrium propagation: The quantum and the thermal cases, *Quantum Stud.: Math. Found.* **12**, 6 (2025).
- [28] C. C. Wanjura and F. Marquardt, Quantum equilibrium propagation for efficient training of quantum systems based on Onsager reciprocity, *Nat. Commun.* **16**, 6595 (2025).
- [29] B. Scellier, A deep learning theory for neural networks grounded in physics, [arXiv:2103.09985](https://arxiv.org/abs/2103.09985).
- [30] S. Massar, Equilibrium propagation for learning in Lagrangian dynamical systems, *Phys. Rev. E* **112**, 035304 (2025).
- [31] B. Scellier, S. Mishra, Y. Bengio, and Y. Ollivier, Agnostic physics-driven deep learning, [arXiv:2205.15021](https://arxiv.org/abs/2205.15021).
- [32] P. Lipka-Bartosik, G. Blasi, J. L. Puértolas, G. Haack, M. Perarnau-Llobet, and N. Brunner, Thermodynamic networks: Harnessing non-equilibrium steady states for computation, [arXiv:2605.15985](https://arxiv.org/abs/2605.15985).
- [33] L. E. Altman, M. Stern, A. J. Liu, and D. J. Durian, Experimental demonstration of coupled learning in elastic networks, *Phys. Rev. Appl.* **22**, 024053 (2024).
- [34] S. Dillavou, M. Stern, A. J. Liu, and D. J. Durian, Demonstration of decentralized physics-driven learning, *Phys. Rev. Appl.* **18**, 014040 (2022).
- [35] S. Dillavou, B. D. Beyer, M. Stern, A. J. Liu, M. Z. Miskin, and D. J. Durian, Machine learning without a processor: Emergent learning in a nonlinear analog network, *Proc. Natl. Acad. Sci. USA* **121**, e2319718121 (2024).
- [36] S.-I. Yi, J. D. Kendall, R. S. Williams, and S. Kumar, Activity-difference training of deep neural networks using memristor crossbars, *Nat. Electron.* **6**, 45 (2023).
- [37] J. Laydevant, D. Marković, and J. Grollier, Training an Ising machine with equilibrium propagation, *Nat. Commun.* **15**, 3671 (2024).
- [38] D. V. Abeele, D. Veraldi, D. Pierangeli, C. Conti, and S. Massar, Optical implementation of equilibrium propagation using spatial photonic Ising machines, [arXiv:2606.13454](https://arxiv.org/abs/2606.13454).
- [39] C.-R. Fan, B. Lu, Z.-H. Zhang, R.-Q. Zhang, J.-W. Wen, and C. Wang, Optimizing energy-based neural network training with coherent Ising machine, [arXiv:2606.09117](https://arxiv.org/abs/2606.09117).
- [40] B. Scellier, A. Goyal, J. Binas, T. Mesnard, and Y. Bengio, Generalization of equilibrium propagation to vector field dynamics, [arXiv:1808.04873](https://arxiv.org/abs/1808.04873).
- [41] A. Laborieux and F. Zenke, Holomorphic equilibrium propagation computes exact gradients through finite size oscillations, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., Red Hook, NY, 2022), Vol. 35, pp. 12950–12963.
- [42] T. W. Hughes, I. A. Williamson, M. Minkov, and S. Fan, Wave physics as an analog recurrent neural network, *Sci. Adv.* **5**, eaay6946 (2019).
- [43] T. Byrnes, N. Y. Kim, and Y. Yamamoto, Exciton–polariton condensates, *Nat. Phys.* **10**, 803 (2014).
- [44] D. Sanvitto and S. Kéna-Cohen, The road towards polaritonic devices, *Nat. Mater.* **15**, 1061 (2016).
- [45] M. D. Fraser, S. Höfling, and Y. Yamamoto, Physics and applications of exciton–polariton lasers, *Nat. Mater.* **15**, 1049 (2016).
- [46] A. Opala, S. Ghosh, T. C. Liew, and M. Matuszewski, Neuromorphic computing in Ginzburg–Landau polariton-lattice systems, *Phys. Rev. Appl.* **11**, 064029 (2019).
- [47] D. Ballarini, A. Gianfrate, R. Panico, A. Opala, S. Ghosh, L. Dominici, V. Ardizzone, M. De Giorgi, G. Lerario, G. Gigli, *et al.*, Polaritonic neuromorphic computing outperforms linear classifiers, *Nano Lett.* **20**, 3506 (2020).
- [48] R. Mirek, A. Opala, P. Comaron, M. Furman, M. Król, K. Tyszk, B. Seredyński, D. Ballarini, D. Sanvitto, T. C. H. Liew, W. Pacuski, J. Suffczynski, J. Szczytko, M. Matuszewski, and B. Piętk, Neuromorphic binarized polariton networks, *Nano Lett.* **21**, 3715 (2021).
- [49] A. Kavokin, T. C. Liew, C. Schneider, P. G. Lagoudakis, S. Klemmt, and S. Höfling, Polariton condensates for classical and quantum computing, *Nat. Rev. Phys.* **4**, 435 (2022).
- [50] A. Opala and M. Matuszewski, Harnessing exciton-polaritons for digital computing, neuromorphic computing, and optimization [Invited], *Opt. Mater. Express* **13**, 2674 (2023).
- [51] K. Tyszk, M. Furman, R. Mirek, M. Król, A. Opala, B. Seredyński, J. Szczytko, W. Pacuski, M. Matuszewski, and B. Piętk, Leaky integrate-and-fire mechanism in exciton–polariton condensates for photonic spiking neurons, *Laser Photon. Rev.* **17**, 2370001 (2023).
- [52] A. Opala, R. Panico, V. Ardizzone, B. Piętk, J. Szczytko, D. Sanvitto, M. Matuszewski, and D. Ballarini, Training a neural network with exciton-polariton optical nonlinearity, *Phys. Rev. Appl.* **18**, 024028 (2022).
- [53] A. Opala, K. Tyszk, M. Kędziora, M. Furman, A. Rahmani, S. Świerczewski, M. Ekielski, A. Szerling, M. Matuszewski, and B. Piętk, Perovskite microwires for room temperature exciton-polariton neural network, *Adv. Mater.* **37**, e07612 (2025).
- [54] M. Matuszewski, A. Opala, R. Mirek, M. Furman, M. Król, K. Tyszk, T. C. H. Liew, D. Ballarini, D. Sanvitto, J. Szczytko, and B. Piętk, Energy-efficient neural network inference with microcavity exciton polaritons, *Phys. Rev. Appl.* **16**, 024045 (2021).
- [55] H. Deng, H. Haug, and Y. Yamamoto, Exciton-polariton Bose-Einstein condensation, *Rev. Mod. Phys.* **82**, 1489 (2010).
- [56] E. Estrecho, T. Gao, N. Bobrovska, D. Comber-Todd, M. D. Fraser, M. Steger, K. West, L. N. Pfeiffer, J. Levinsen, M. Parish, *et al.*, Direct measurement of polariton-polariton interaction strength in the Thomas-Fermi regime of exciton-polariton condensation, *Phys. Rev. B* **100**, 035306 (2019).
- [57] A. E. Scurria, D. V. Abeele, B. M. Mognetti, and S. Massar, Equilibrium propagation for non-conservative systems, [arXiv:2602.03670](https://arxiv.org/abs/2602.03670).
- [58] K. Sajnok and M. Matuszewski, Simulations of the near-equilibrium propagation training in nonlinear wave systems [Computer software], Zenodo, 2025, <https://doi.org/10.5281/zenodo.17369400>.