Forming complex neurons by four-wave mixing in a Bose-Einstein condensate

Kai Niklas Hansmann [*] and Reinhold Walser
Institut für Angewandte Physik, Technische Universität Darmstadt, Hochschulstraße 4a, 64289 Darmstadt, Germany

(Received 29 September 2023; accepted 11 December 2023; published 2 January 2024)

A physical artificial complex-valued neuron is formed by four-wave mixing in a homogeneous three-dimensional Bose-Einstein condensate. Bragg beam-splitter pulses prepare superpositions of three plane-wave states as an input and the fourth wave as an output signal. The nonlinear dynamics of the nondegenerate four-wave mixing process leads to Josephson-like oscillations within the closed four-dimensional subspace and defines the activation function of a neuron. Due to the high number of symmetries, closed-form solutions can be found by quadrature and agree with the numerical simulation. The ideal behavior of an isolated four-wave mixing setup is compared to a situation with additional population of rogue states. We observe a robust persistence of the main oscillation. As an application for neural learning of this physical system, we train it on the XOR problem. After 100 training epochs, the neuron responds to input data correctly at the 10^{-5} error level.

DOI: 10.1103/PhysRevA.109.013302

I. INTRODUCTION

Neural networks and deep learning methods have evolved dynamically into a far-reaching research field [1]. Nowadays, applications can be found in diverse areas like biochemistry [2], medicine [3], image analysis [4], computer games [5,6], gravitational-wave detection [7], and sundry more. A variety of implementations of artificial neural networks exist: not only electronic implementations using graphical processing units [8,9] but also other physical implementations can be considered [10,11]. In particular, optical implementations receive a lot of attention [12–21].

The key issue in setting up a novel physical implementation of artificial neural networks is the description of their constituents, the artificial neurons. Diverse approaches can realize artificial neurons in photonic systems [22–26]. In this paper, we consider an artificial neuron using the inherent nonlinearity of ultracold coherent bosonic matter waves.

Coherent matter waves show a wide range of nonlinear effects, which, for example, have been used to detect the phase transition towards a Bose-Einstein condensate (BEC) experimentally [27,28]. For our purposes, we investigate the process of four-wave mixing (FWM) in coherent matter waves, which is well known from nonlinear optics [29]. If phase-matching conditions are fulfilled in a nonlinear optical medium, three frequencies interact in a way such that an initially absent fourth frequency can be observed. Following the advent of the BEC, theoretical investigations [30–34], as well as experiments [35,36], demonstrated the equivalent FWM process. There, momentum components of the BEC took over the role of optical frequencies from the initial scenario. In an idealized homogeneous BEC, we can show that the FWM process of plane waves exhibits Josephson-like oscillations [37–43].

We utilize this highly nonlinear process to implement a complex-valued neuron, where we identify three-momentum components as the input and the fourth component as the output. The input-output relations of the neuron are highly nonlinear and will be investigated in detail. These two features offer up the possibility to implement deep neural networks with increased data throughput compared to real-valued neurons with linear activation functions. As an application, we train the FWM neuron to solve the benchmark XOR problem.

This paper is organized as follows. In Sec. II, we introduce the isolated FWM problem in a three-dimensional homogeneous BEC, revealing the dynamics of populations and phases of the four-momentum components. In Sec. III, we solve the FWM dynamics analytically in the form of Josephson oscillations. After the investigation of FWM under ideal conditions, we look at the influence of additional population in momentum components outside of the FWM manifold in Sec. IV. Finally, we introduce the artificial FWM neuron in Sec. V, discuss the nonlinear activation function of the neuron, and introduce the steepest descent learning method for complex-valued neurons. As an application, we train the FWM neuron on the XOR problem. In the Appendix, we discuss the preparation of FWM input states using Bragg pulses well known in atom interferometry [44–46].

II. IDEAL FOUR-WAVE MIXING

The dynamics of a weakly interacting BEC described by the order parameter $\psi(\mathbf{r}, t)$ are given by the Gross-Pitaevskii equation [47,48]

$$i\hbar\partial_t\psi = \left[-\frac{\hbar^2}{2m}\nabla^2 + U + gn\right]\psi, \tag{1}$$

where $n(\mathbf{r}) = |\psi(\mathbf{r})|^2$ is the density, $N = \int d^3r\, n$ is the total particle number, $U(\mathbf{r})$ describes an external potential, and the coupling constant $g = 4\pi\hbar^2 a_s/m$ is proportional to the atomic

[*]kai.hansmann@physik.tu-darmstadt.de

 ©2024 American Physical Society

s-wave scattering length a_s and the mass of the atoms m. The Gross-Pitaevskii Lagrangian functional for such a system reads [49,50]

$$L = \int \mathrm{d}^3 r\, (i\hbar\psi^* \partial_t \psi - \mathcal{E}), \tag{2}$$

$$\mathcal{E} = \frac{\hbar^2}{2m}|\nabla\psi|^2 + Un + \frac{g}{2}n^2. \tag{3}$$

In the following, we consider the case of a homogeneous BEC with $U = 0$ and periodic boundary conditions. Then, a wave function $|\psi\rangle = |\psi_\alpha\rangle + |\psi_\beta\rangle$ is a coherent superposition of plane waves $|\mathbf{k}_j\rangle$ with complex amplitudes α_j and β_l. It consists of a FWM state

$$|\psi_\alpha\rangle = \sum_{j=1}^{4} \sqrt{N}\alpha_j|\mathbf{k}_j\rangle \tag{4}$$

and a residual wave

$$|\psi_\beta\rangle = \sum_{l>4} \sqrt{N}\beta_l|\mathbf{k}_l\rangle, \tag{5}$$

which is orthogonal $\langle\psi_\alpha|\psi_\beta\rangle = 0$ to the FWM state.

The complex amplitudes α_j, in terms of absolute value and phase, are given by

$$\alpha_j = \sqrt{n_j}e^{-i\varphi_j}. \tag{6}$$

Thus, $n_j = |\alpha_j|^2$ is the probability of being in the momentum state $|\mathbf{k}_j\rangle$. The mode functions $\langle\mathbf{r}|\mathbf{k}_j\rangle = \mathrm{e}^{i\mathbf{k}_j\mathbf{r}}/\sqrt{V}$ are normalized in a cuboid with lengths (L_1, L_2, L_3) and a volume $V = L_1L_2L_3$. For periodic boundary conditions, the wave numbers $k_j = 2\pi\kappa_j/L_j$ are quantized with $\kappa_j \in \mathbb{Z}$, and the plane-wave states are orthonormal $\langle\mathbf{k}_i|\mathbf{k}_j\rangle = \delta_{ij}$.

The conditions for FWM are energy and momentum conservation [35],

$$\omega_1 + \omega_2 = \omega_3 + \omega_4, \quad \mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{k}_4. \tag{7}$$

We choose for the whole population [see (4)] to reside in momentum states with finite k, similar to spinor condensates [41]. Hence, the dispersion relation for free massive particles holds,

$$\omega_j = \omega(\mathbf{k}_j) = \frac{\hbar|\mathbf{k}_j|^2}{2m}. \tag{8}$$

Experimentally, momentum states fulfilling these conditions can be prepared using atomic beam splitters based on Bragg diffraction [44,51], as discussed in the Appendix.

In the ideal FWM scenario, the residual wave is absent, $\beta_l = 0$. Consequently, $\sum_{j=1}^{4} n_j = 1$. In this case, the nondimensionalization of the physical Lagrangian functional (2) is achieved by measuring time $\tau = \gamma t$ by a clock that ticks with frequency $\gamma = gN/\hbar V$ and by scaling the frequencies $\bar{\omega}_j = \omega_j/\gamma$, as well as by scaling and shifting the Lagrangian function $\mathcal{L} = 1 + VL/gN^2$. Thus, the mathematical Lagrangian functional reads

$$\mathcal{L} = \sum_{j=1}^{4} i\alpha_j^*\dot{\alpha}_j - \mathcal{E}, \tag{9}$$

$$\mathcal{E} = \sum_{j=1}^{4} \varepsilon_j + 2(\alpha_1^*\alpha_2^*\alpha_3\alpha_4 + \text{c.c.}), \tag{10}$$

where $\dot{\alpha}_j$ denotes $\partial_\tau\alpha_j$ and the mean-field shifted single-particle energies ε_j and chemical potentials μ_j are defined as

$$\varepsilon_j = \bar{\omega}_j n_j - \frac{n_j^2}{2}, \quad \mu_j = \frac{\partial\varepsilon_j}{\partial n_j} = \bar{\omega}_j - n_j. \tag{11}$$

According to the Euler-Lagrange equations [41]

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial\mathcal{L}}{\partial\dot{\alpha}_j} = \frac{\partial\mathcal{L}}{\partial\alpha_j}, \tag{12}$$

the complex amplitudes evolve as

$$\begin{aligned} i\dot{\alpha}_1 &= \mu_1\alpha_1 + 2\alpha_2^*\alpha_3\alpha_4, \\ i\dot{\alpha}_2 &= \mu_2\alpha_2 + 2\alpha_1^*\alpha_3\alpha_4, \\ i\dot{\alpha}_3 &= \mu_3\alpha_3 + 2\alpha_4^*\alpha_1\alpha_2, \\ i\dot{\alpha}_4 &= \mu_4\alpha_4 + 2\alpha_3^*\alpha_1\alpha_2. \end{aligned} \tag{13}$$

Using the polar decomposition of the complex amplitudes (6), a number of symmetries can be identified. The dynamics are symmetric with regard to the total phase

$$\Phi = \tfrac{1}{4}(\varphi_1 + \varphi_2 + \varphi_3 + \varphi_4), \tag{14}$$

as well as the relative phases between α_1 and α_2,

$$\varphi = \tfrac{1}{2}(\varphi_1 - \varphi_2), \tag{15}$$

and between α_3 and α_4,

$$\theta = \tfrac{1}{2}(\varphi_3 - \varphi_4). \tag{16}$$

According to the Noether theorem [52], this implies conservation of the conjugate momenta.

The interaction term in (10) coherently couples the subspaces $\{|\mathbf{k}_1\rangle, |\mathbf{k}_2\rangle\} \longleftrightarrow \{|\mathbf{k}_3\rangle, |\mathbf{k}_4\rangle\}$ through the relative phase difference

$$\phi = \varphi_1 + \varphi_2 - \varphi_3 - \varphi_4 \tag{17}$$

and population imbalance

$$m = n_1 + n_2 - n_3 - n_4. \tag{18}$$

We evaluate the dynamics of the system (13) numerically. In Fig. 1(a), we show periodic oscillations of the populations $n_j(\tau)$ and quasilinear evolution of the phases $\varphi_j(\tau)$. In the ϕ-m phase space, we find periodic as well as aperiodic orbits of a mathematical pendulum with a separatrix in between [see Fig. 1(b)]. The population imbalance m is constrained by the conserved quantities $m_{12} = n_1 - n_2$ and $m_{34} = n_3 - n_4$ discussed in Sec. III.

III. JOSEPHSON OSCILLATIONS OF FOUR-WAVE MIXING AMPLITUDES

A. Coordinate transformation

Due to the Lagrangian field theory, the time-independent Hamiltonian energy (2), and the FWM state ansatz (4), we obtain a discrete nonlinear set of four Hamiltonian equations with a number of symmetries. It constrains the dynamics to a two-dimensional phase space, analogous to the mathematical pendulum. Due to the phase-invariant structure of the self-energy gn^2, typical Josephson oscillations [37–43] emerge. Similar equations appear in the study of semiclassical methods in the theory of Rydberg atoms [53].

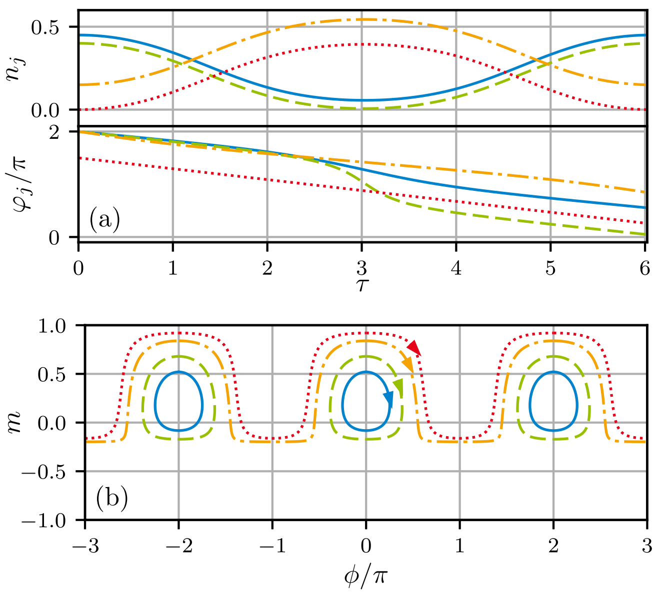

FIG. 1. (a) Populations n_j and phases φ_j of FWM states $|\mathbf{k}_1\rangle$ (blue solid line), $|\mathbf{k}_2\rangle$ (green dashed line), $|\mathbf{k}_3\rangle$ (orange dash-dotted line), and $|\mathbf{k}_4\rangle$ (red dotted line) versus dimensionless time τ for $\bar{\omega}_j = 1$. (b) FWM trajectories $(\phi(\tau), m(\tau))$ in phase space for the initial conditions $\phi(0) = 0$, $m(0) = 0.52$ (blue solid line), $m(0) = 0.68$ (green dashed line), $m(0) = 0.84$ (yellow dash-dotted line), and $m(0) = 0.92$ (red dotted line). The other constants of motion read $m_{12} = n_1 - n_2 = 0.4$ and $m_{34} = n_3 - n_4 = 0.02$.

Guided by this idea, we introduce adapted coordinates

$$\alpha_1 = \sqrt{n_1}\mathrm{e}^{-i(\Phi+\phi/4+\varphi)}, \quad \alpha_2 = \sqrt{n_2}\mathrm{e}^{-i(\Phi+\phi/4-\varphi)},$$
$$\alpha_3 = \sqrt{n_3}\mathrm{e}^{-i(\Phi-\phi/4+\theta)}, \quad \alpha_4 = \sqrt{n_4}\mathrm{e}^{-i(\Phi-\phi/4-\theta)}. \tag{19}$$

From the global phase invariance of (2) or (9), one finds that the total occupation $\sum_{j=1}^{4} n_j = \text{const}$. This can be used to construct a generating function $R(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \Phi, \phi, \varphi, \theta)$ as

$$R = \frac{i}{2}\mathrm{e}^{2i\Phi}\big(\alpha_1^2\mathrm{e}^{2i(\phi/4+\varphi)} + \alpha_2^2\mathrm{e}^{2i(\phi/4-\varphi)} + \alpha_3^2\mathrm{e}^{2i(-\phi/4+\theta)} + \alpha_4^2\mathrm{e}^{2i(-\phi/4-\theta)}\big). \tag{20}$$

According to the rules of Hamiltonian mechanics [54], this generating function relates old coordinates $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ to new coordinates $(\Phi, \phi, \varphi, \theta)$. In turn, one can obtain the old momenta

$$\pi_j = \frac{\partial R}{\partial \alpha_j} = i\alpha_j^*, \tag{21}$$

as well as the new momenta

$$P_\Phi = -\frac{\partial R}{\partial \Phi} = n_1 + n_2 + n_3 + n_4, \tag{22}$$

$$P_\phi = -\frac{\partial R}{\partial \phi} = \frac{n_1 + n_2 - n_3 - n_4}{4} \equiv \frac{m}{4}, \tag{23}$$

$$P_\varphi = -\frac{\partial R}{\partial \varphi} = n_1 - n_2 \equiv m_{12}, \tag{24}$$

$$P_\theta = -\frac{\partial R}{\partial \theta} = n_3 - n_4 \equiv m_{34}. \tag{25}$$

As conjugate momenta to Φ, φ and θ are given by the total particle number and the population differences m_{12} and m_{34}; these quantities are conserved due to the phase symmetry of the dynamics. Consequently, the equations of motion for Φ, φ, and θ can be solved by quadrature.

In terms of the new coordinates the dimensionless Lagrangian $\mathcal{L}$ reads

$$\mathcal{L} = \dot{\Phi} + \frac{m}{4}\dot{\phi} + m_{12}\dot{\varphi} + m_{34}\dot{\theta} - H(m, \phi), \tag{26}$$

with a generic Josephson Hamiltonian energy

$$H(m, \phi) = \frac{\eta}{4}\cos\phi - \frac{m^2}{8} + \mathcal{C}, \tag{27}$$

$$\eta = \sqrt{\big[(1+m)^2 - 4m_{12}^2\big]\big[(1-m)^2 - 4m_{34}^2\big]}. \tag{28}$$

Here, we have denoted an energy offset $\mathcal{C} = [m_{12}^2 + m_{34}^2 + 2(\bar{\omega}_{12}m_{12} + \bar{\omega}_{34}m_{34}) - 7/2 + \sum_{j=1}^{4}\bar{\omega}_j]/4$ and transition energies $\bar{\omega}_{12} = \bar{\omega}_1 - \bar{\omega}_2$ and $\bar{\omega}_{34} = \bar{\omega}_3 - \bar{\omega}_4$.

Clearly, H (27) is the Legendre transform of $\mathcal{L}$ (26). Accordingly, the dynamics of the system, using (23), read

$$\dot{\phi} = 4\partial_m H = \cos\phi\,\partial_m\eta - m,$$
$$\dot{m} = -4\partial_\phi H = \eta\sin\phi. \tag{29}$$

These are Josephson-like differential equations [38–40].

B. General solution

In simple classical mechanics problems of particles with position x and momentum p, Hamiltonian energies $H(x, p) = T(p) + V(x)$ separate into kinetic $T(p)$ and potential $V(x)$ energy. At the turning points $\dot{x} = \partial_p H = 0$, the Hamilton function is purely determined by potential energy $H(x, p = 0) = V(x)$. A similar investigation can be performed in the given case [53,55]. Through a canonical transformation, we can exchange the roles of position and momentum and consider m as the position and ϕ as the momentum variable. Thus, at the turning points $\dot{m} = -4\partial_\phi H = 0$, remarkably, two momenta,

$$\phi^+ = 0, \quad \phi^- = \pi, \tag{30}$$

are possible. In turn, this defines two potentials,

$$H(m, \phi^\pm) = V^\pm(m) = \pm\frac{\eta}{4} - \frac{m^2}{8} + \mathcal{C}. \tag{31}$$

Physical solutions with energies $\varepsilon = H(m, \phi)$ must be constrained by these two potentials, $V^- < \varepsilon < V^+$. This limits the value range of m and ϕ depending on the system parameters m_{12} and m_{34} (see Fig. 2). As the energy of the system is conserved, the equation of motion (29) for $m(\tau)$ can be expressed using the potentials $V^\pm$ as

$$\dot{m} = \pm 4\sqrt{[V^+(m) - \varepsilon][\varepsilon - V^-(m)]}. \tag{32}$$

Thus, the dynamical solution $\tau(m)$ can be calculated as

$$\tau(m) - \tau_0 = \int_{m_0}^{m} \frac{\pm\mathrm{d}\zeta}{4\sqrt{[V^+(\zeta) - \varepsilon][\varepsilon - V^-(\zeta)]}}. \tag{33}$$

This relation can be inverted piecewise to obtain $m(\tau)$.

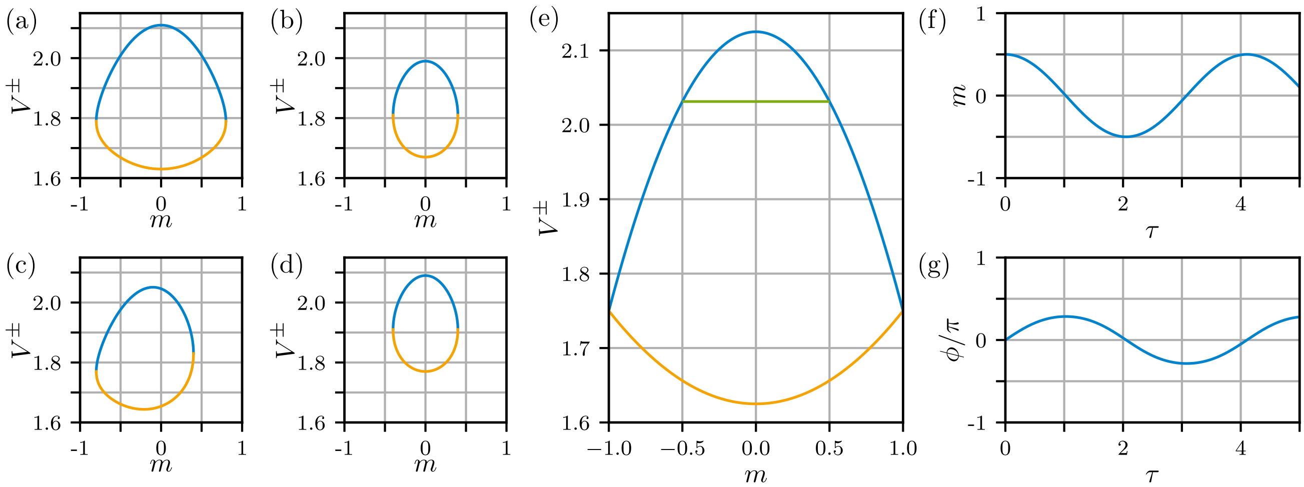

FIG. 2. (a)–(c) Potentials $V^+(m)$ (blue) and $V^-(m)$ (orange) versus population imbalance m (31) for $\bar{\omega}_j = 1$. For $m_{12} = m_{34}$ equal to 0.1 in (a) and 0.3 in (b), the potentials are symmetric around the origin. Otherwise, this symmetry is broken [$m_{12} = 0.1$, $m_{34} = 0.3$ in (c)]. (d) The potentials can be shifted in dimensionless energy by varying the value of the recoil frequencies ($\bar{\omega}_j = 1.1$). (e) For $m_{12} = m_{34} = 0$, V^+ and V^- are symmetric in m and $m \in [-1, 1]$. For initial values $m_0 = 0.5$, $\phi_0 = 0$, and $\tau_0 = 0$, the dynamics of the Josephson variables described by (34) and (35) are shown in (f) and (g). The energy is constant during the oscillation [green in (e)].

C. Analytical solution for $m_{12} = m_{34} = 0$

For the special case $m_{12} = m_{34} = 0$, implying $n_1 = n_2$ and $n_3 = n_4$, an analytical expression for the dynamical solution $m(\tau)$ can be given in terms of the elliptic cosine $\mathrm{cn}(u)$ [56] as

$$m(\tau) = \pm\sqrt{\frac{\mu + 2}{3}}\,\mathrm{cn}(\xi(\tau - \tau_0), \rho^2), \tag{34}$$

where $\mu = m_0^2 + 2(m_0^2 - 1)\cos\phi_0$, $\xi = \sqrt{6 - 3\mu}/2$, and $\rho^2 = (\mu + 2)/(6 - 3\mu)$. With that, the dynamical solution of the phase $\phi(\tau)$ can be calculated by integration of (29), yielding

$$\begin{aligned}\phi(\tau) = 2\arctan(\sqrt{3}\tanh\{&\ln(1 - \rho)\\ &- \ln[\mathrm{dn}(\xi(\tau - \tau_0), \rho^2) - \rho\,\mathrm{cn}(\xi(\tau - \tau_0), \rho^2)]\\ &+ \mathrm{arctanh}[\tan(\phi_0/2)/\sqrt{3}]\}),\end{aligned} \tag{35}$$

with the delta amplitude $\mathrm{dn}(u)$ [56]. The analytical solutions for $m(\tau)$ and $\phi(\tau)$ as well as visualizations of the potentials V^+ and V^- can be seen in Fig. 2.

The period of motion can be calculated as

$$T = \frac{4\,\mathrm{K}(\rho^2)}{\xi}. \tag{36}$$

Here, K is the complete elliptic integral of the first kind [56]. The basic frequency of the oscillation $T_0 = T(m_0 = 0)$ can be calculated as

$$T_0 = 4\pi/\sqrt{12}. \tag{37}$$

As can be seen in Fig. 3, the period of the FWM oscillation diverges when nearing the regime of aperiodic solutions.

IV. FOUR-WAVE MIXING WITH BACKGROUND POPULATION

In the ideal FWM setting, the residual wave $|\psi_\beta\rangle$ is absent. However, additional momentum states might be populated accidentally during the initialization procedure or system evolution. To investigate this scenario, we simulate the dynamics of the system, described by the Gross-Pitaevskii equation (1), using a Runge-Kutta scheme and fast Fourier transforms (FFTs) on a discrete periodic lattice. We use a two-dimensional lattice with 16×16 sites while setting $\gamma = 1/\mathrm{s}$ and discretizing dimensionless time with $\Delta\tau = 10^{-6}$. For implementation we choose the geometry of FWM states described in the Appendix, yielding $\bar{\omega}_j = 1$ for $j = 1, \ldots, 4$. The populations are set to $n_1 = n_2 = 0.375$ and $n_3 = n_4 = 0.125$, resulting in $m_{12} = m_{34} = 0$ and $m_0 = 0.5$. All phases are set to $\varphi_j = 0$, yielding $\phi_0 = 0$.

As can be seen in Fig. 4, the numerical results of the GP simulation start to deviate noticeably from the four-mode approximation (13) already after a few cycles. Looking at $m(\tau)$ and $\phi(\tau)$, the numerical results show a larger period of the oscillation. However, the general shape of the oscillations remains unchanged.

This behavior is caused by an instability of the simulation due to numerical noise of the FFTs producing population on the grid outside of the FWM states. As depicted in Fig. 5(a),

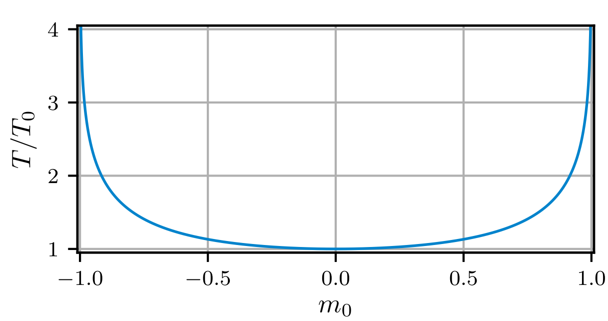

FIG. 3. Period of the FWM oscillation T versus the initial population imbalance m_0, normed to T_0. The period diverges when nearing aperiodic solutions.

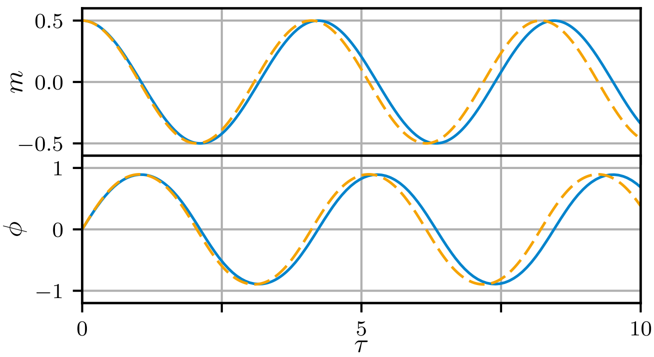

FIG. 4. Population imbalance $m(\tau)$ and relative phase $\phi(\tau)$ versus dimensionless time τ for analytical (orange dashed line) and numerical GP simulations on a discrete periodic lattice (blue solid line).

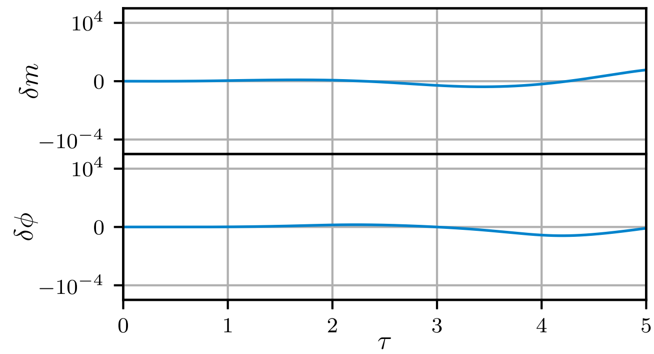

FIG. 6. Deviations $\delta m = (m_n - m_a)/\max(m_a)$ and $\delta\phi = (\phi_n - \phi_a)/\max(\phi_a)$ versus dimensionless time τ from numerical (m_n, ϕ_n) to analytical (m_a, ϕ_a) solutions. All population outside of the FWM states is eliminated after each simulation step.

the system is prepared at $\tau = 0$ with population present in only the FWM states. However, the histogram in Fig. 5(b) at $\tau = 5$ clearly shows that additional states in the vicinity of the FWM states have been populated. As this background population is located at the center of the lattice, the chosen grid is large enough that no edge effects occur during the simulation.

However, the instability caused by accidental population of additional momentum states is not destructive in nature. Looking at Fig. 5(c), the total background population

$$n_B = \sum_{l>4} |\beta_l|^2 \tag{38}$$

grows rapidly at the beginning of the oscillation. Subsequently, the dynamics of $n_B(\tau)$ stabilize and show oscillations

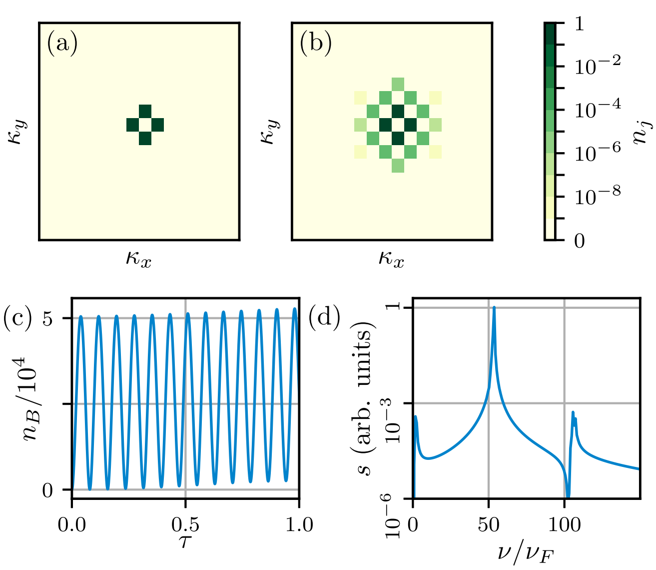

FIG. 5. Histograms of populations on a discrete 16×16 lattice in the κ_x-κ_y-plane at (a) $\tau = 0$ and (b) $\tau = 5.0$. (c) The background population n_B starts oscillating and quickly reaches the maximum value. (d) Spectral density s versus frequency ν. The oscillation frequency of n_B is about 50 times bigger than the FWM frequency ν_F.

with a maximum value of around $n_B \simeq 5 \times 10^{-4}$. As can be seen in Fig. 5(d), the frequency of the ensuing oscillation is about 50 times larger than the FWM frequency,

$$\nu_F = \frac{1}{T(m_0 = 0.5)} \simeq 0.244. \tag{39}$$

The non-negligible background population is the cause of the change in the dynamics of the FWM process. Because

$$n_F + n_B = 1, \quad n_F = \sum_{j=1}^{4} |\alpha_j|^2, \tag{40}$$

growing n_B reduces the population in the FWM states n_F in comparison to the ideal case. As the FWM process is caused by the density-density-interaction terms in the Gross-Pitaevskii equation (1), even small changes in the particle number participating in the process have profound effects on the dynamics.

The analytical solution can be recovered by eliminating all numerical noise produced by FFTs after each simulation step. Using such masks in k space, the numerical simulation and analytical solution agree within about 10^{-5} (see Fig. 6). However, this procedure yields a loss in total particle number of about $\Delta N/N = 10^{-6}$, far surpassing typical numerical noise.

For the implementation of the FWM neuron, we are interested in rather short timescales and more qualitative behavior of the system. Therefore, we accept the change in frequency of the FWM oscillations and perform the simulation on a discrete periodic lattice in the investigations without additionally applying a filter mask in k space. This is beneficial due to the high flexibility of the simulation regarding the initial conditions of the FWM states. However, the deviation between the ideal case and the case with the present background population should be kept in mind, especially when looking at increasing simulation times.

V. FOUR-WAVE MIXING NEURON

Artificial neurons are the basic computation units in neural networks [57]. In addition to the processing of real numbers, such systems are also able to operate with complex-valued inputs and outputs [58]. As the FWM process is described in terms of complex amplitudes α_j, the presented implementation of the FWM neuron constitutes a complex-valued neuron. Due to the experimental accessibility of particle numbers and phases, we choose to describe the nonlinear activation function and the learning process in terms of absolute values and phases, rather than using real and imaginary parts of the complex amplitudes α_j [58].

In general, complex-valued artificial neurons process an n-dimensional input $x_j = |x_j|e^{i\kappa_j}$, $j = 1, \dots, n$, by multiplying individually by weights $w_j = |w_j|e^{i\vartheta_j}$, summing up the weighted inputs $v_j = w_j x_j$ and yielding an output y via a nonlinear activation function Ω,

$$y = \Omega(u), \quad u = \sum_{j=1}^{n} v_j. \tag{41}$$

We implement such a computational unit with the FWM process on coherent matter waves. The phase flow

$$\tilde{\boldsymbol{\alpha}} = \Phi(\boldsymbol{\alpha}; \tau_F) \tag{42}$$

maps the initial state $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_4)$ to the evolved state $\tilde{\boldsymbol{\alpha}} = (\tilde{\alpha}_1, \dots, \tilde{\alpha}_4)$ after the duration τ_F of the FWM process (13). Identifying the three amplitudes α_1, α_2, and α_3 as weighted inputs v_j and $\tilde{\alpha}_4$ as output y, a rule similar, although not identical, to (41) can be established:

$$\tilde{\alpha}_4 = \Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_F). \tag{43}$$

The fourth component of the phase-flow map constitutes a nonlinear activation function of a complex-valued FWM neuron with three input channels. In an experiment, we use externally stored weights w_j for the neuron and the classical input data x_j and prepare the weighted input amplitude

$$\alpha_j = w_j x_j \tag{44}$$

by a sequence of Bragg pulses (see the Appendix).

A. Nonlinear activation function

In order to quantify the nonlinear activation function, τ_F has to be determined. To do so, we choose $n_1 = n_2 = 0.45$ and $n_3 = 0.1$ while setting $\varphi_j = 0$. The resulting FWM oscillation can be seen in Fig. 7. To maximize the output in terms of $\tilde{n}_4$ for this scenario, we set

$$\tau_F = T/2, \tag{45}$$

where T is the oscillation period as in (36).

The FWM neuron response is calculated for varying weighted inputs numerically (see Sec. IV). We tune n_j from 0 to 1 subject to the constraint $\sum_{j=1}^{4} n_j = 1$. Due to probability (number) conservation, all admissible combinations of n_j form a plane in n_1-n_2-n_3 space. The input phases φ_j are varied from 0 to 2π. The results can be seen in Fig. 8. The output particle number

$$\tilde{n}_4 = |\Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_F)|^2 \tag{46}$$

is independent of the input phases φ_j. Hence, only the input particle numbers n_j determine this part of the output. While there is no analytical expression for the relation, it can be determined from Fig. 8 that there has to be an exchange symmetry regarding n_1 and n_2.

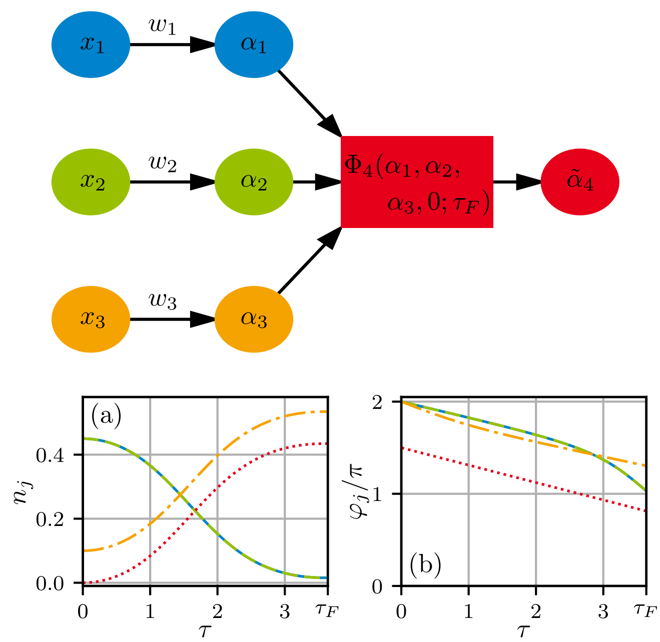

FIG. 7. Initialization sequence for a FWM neuron. Classical inputs x_j are weighted with w_j, yielding amplitudes α_j. The nonlinear relation $\Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_F)$ yields output $\tilde{\alpha}_4$. The duration $\tau_F = T/2$ is determined for $n_1 = n_2 = 0.45$ (blue solid and green dashed lines) and $n_3 = 0.1$ (orange dash-dotted line; n_4: red dotted line), while $\varphi_j = 0$, as a half-oscillation period leading to maximal response.

The output phase

$$\tilde{\varphi}_4 = \arg[\Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_F)] \tag{47}$$

exhibits a remarkably simple behavior. By analyzing Fig. 8(b), we find

$$n_\varphi = 3n_1 + 3n_2 + 5n_3, \quad \varphi_\varphi = \varphi_1 + \varphi_2 - \varphi_3. \tag{48}$$

Accordingly, the input-output relation reads

$$\tilde{\varphi}_4 = s n_\varphi + \varphi_\varphi + d, \tag{49}$$

where the slope and offset of the phase were determined from a fit as $s = (-1.77 \pm 0.01)$ and $d = (2.67 \pm 0.04)$.

The numerical results in Fig. 8 can be used to determine the partial derivatives $\partial\tilde{n}_4/\partial n_j$, $\partial\tilde{\varphi}_4/\partial n_j$, and $\partial\tilde{\varphi}_4/\partial\varphi_j$. These are needed to train the neuron according to the steepest descent method.

B. Steepest descent learning for complex-valued neurons

Steepest descent methods are common procedures in optimization, as well as in supervised learning in neural networks [59]. We consider the case of a single output neuron. In so-called error-correction learning, this neuron is stimulated by an input vector $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, x_3^{(i)})$, where i denotes an instant in time at which the excitation is applied to the system. The training data set is described by

$$\mathcal{T} : \{\mathbf{x}^{(i)}, \hat{\alpha}_4^{(i)}; i = 1, \dots, \mathcal{M}\}, \tag{50}$$

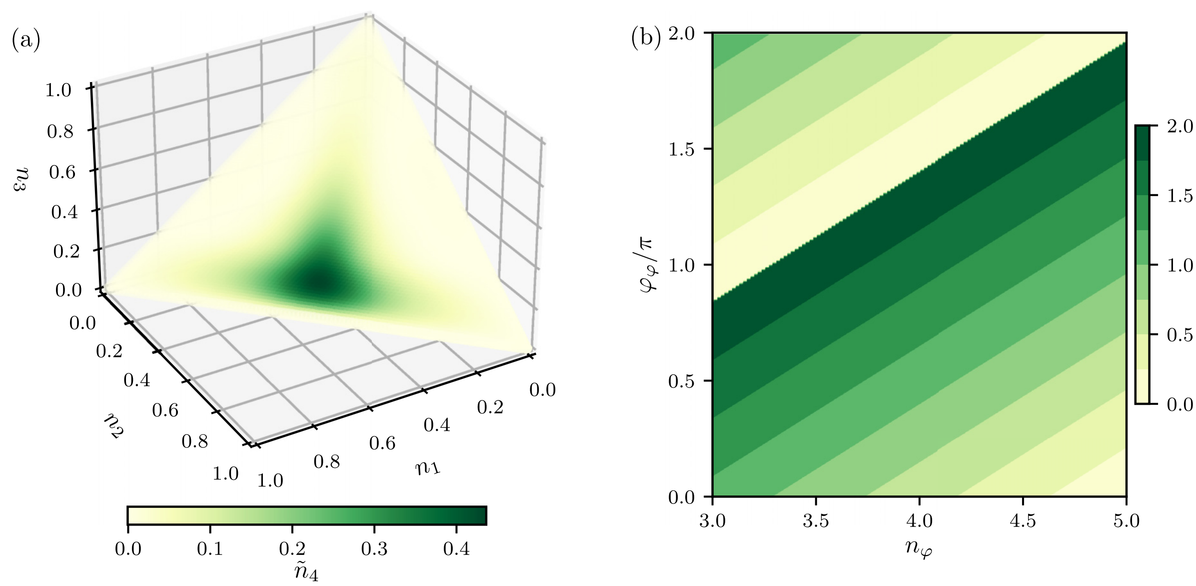

FIG. 8. Nonlinear activation function $\Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_F)$ of the FWM neuron in terms of (a) $\tilde{n}_4$ (46) versus n_1, n_2, and n_3 and (b) $\tilde{\varphi}_4$ (47) versus n_φ and φ_φ (48).

where $\hat{\alpha}_4^{(i)}$ is the desired response associated with $\mathbf{x}^{(i)}$ and $\mathcal{M}$ is the size of the data set. In response to this stimulus, the neuron produces an output $\tilde{\alpha}_4^{(i)}$.

Starting from initial weights $\mathbf{w} = (w_1, \dots, w_n)$, the goal of the learning procedure is to adjust the weights to minimize the difference between the desired and actual outputs, described by means of a cost function $\mathcal{F}$. A typical cost function is the square error averaged over the training sample set [60]

$$\mathcal{F} = \frac{1}{\mathcal{M}} \sum_{i=1}^{\mathcal{M}} \mathcal{F}^{(i)}, \quad \mathcal{F}^{(i)} = \frac{1}{2}\left|\tilde{\alpha}_4^{(i)} - \hat{\alpha}_4^{(i)}\right|^2. \tag{51}$$

This yields an unconstrained optimization problem with the necessary condition for optimality $\nabla\mathcal{F} = 0$, where ∇ denotes the gradient operator in weight space. In a steepest descent method, adjustments applied to the weight vector are performed in the direction of the negative gradient

$$\Delta\mathbf{w}(n) = \mathbf{w}(n+1) - \mathbf{w}(n) = -\lambda\nabla\mathcal{F}, \tag{52}$$

where n symbolizes one iteration and λ is a positive learning rate.

In the online learning approximation [61], adjustments to the weights are performed on an example-by-example basis. The cost function to minimize is therefore the instantaneous error energy $\mathcal{F}^{(i)}$. An epoch consists of $\mathcal{M}$ training samples. At an instant i, a pair $\{\mathbf{x}^{(i)}, \hat{\alpha}_4^{(i)}\}$ is presented to the neuron, and weight adjustments are performed. Subsequently, the next sample is presented to the network until all $\mathcal{M}$ samples have been evaluated.

The absolute values and phases of the weights can be updated independently [58],

$$\Delta\left|w_j^{(i)}\right| = -\lambda_a \partial_{|w_j|}\mathcal{F}^{(i)}, \quad \Delta\vartheta_j^{(i)} = -\lambda_p \partial_{\vartheta_j}\mathcal{F}^{(i)}, \tag{53}$$

where λ_a and λ_p are the learning rates for the absolute value and phase, respectively. The required gradients for the update rules (53), keeping in mind the variable dependences of the nonlinear activation function, are calculated using the chain rule as

$$\begin{aligned} \frac{\partial\tilde{n}_4}{\partial|w_j|} &= \frac{\partial\tilde{n}_4}{\partial n_j}\frac{\partial n_j}{\partial|w_j|} = |x_j|\frac{\partial\tilde{n}_4}{\partial n_j}, \\ \frac{\partial\tilde{\varphi}_4}{\partial|w_j|} &= \frac{\partial\tilde{\varphi}_4}{\partial n_j}\frac{\partial n_j}{\partial|w_j|} = |x_j|\frac{\partial\tilde{\varphi}_4}{\partial n_j}, \\ \frac{\partial\tilde{\varphi}_4}{\partial\vartheta_j} &= \frac{\partial\tilde{\varphi}_4}{\partial\varphi_j}\frac{\partial\varphi_j}{\partial\vartheta_j} = \frac{\partial\tilde{\varphi}_4}{\partial\varphi_j}. \end{aligned} \tag{54}$$

Hence, the update rules for $|w_j|$ and ϑ_j are

$$\begin{aligned} \Delta\left|w_j^{(i)}\right| = &-\lambda_a\left\{\left[\tilde{n}_4^{(i)} - \hat{n}_4^{(i)}\cos\left(\tilde{\varphi}_4^{(i)} - \hat{\varphi}_4^{(i)}\right)\right]\partial_{n_j}\tilde{n}_4^{(i)}\right. \\ &\left.+\,\tilde{n}_4^{(i)}\sin\left(\tilde{\varphi}_4^{(i)} - \hat{\varphi}_4^{(i)}\right)\partial_{n_j}\tilde{\varphi}_4^{(i)}\right\}\left|x_j^{(i)}\right|, \end{aligned} \tag{55}$$

$$\Delta\vartheta_j^{(i)} = -\lambda_p\tilde{n}_4^{(i)}\hat{n}_4^{(i)}\sin\left(\tilde{\varphi}_4^{(i)} - \hat{\varphi}_4^{(i)}\right)\partial_{\varphi_j}\tilde{\varphi}_4^{(i)}. \tag{56}$$

C. Application: XOR problem

To investigate the learning abilities of the FWM neuron, we use it to solve the XOR problem. This represents a basic benchmark problem which all implementations of machine learning should be able to master. Its input-output mapping is shown in Table I. The XOR problem consists of two real-valued binary inputs. Its output is 0 if the two inputs are identical and 1 if they are different. This problem is not solvable for a single real-valued neuron; that is, hidden layers are required [62].

TABLE I. Input-output mapping for the XOR problem.

Input 1	Input 2	Output
0	0	0
0	1	1
1	0	1
1	1	0

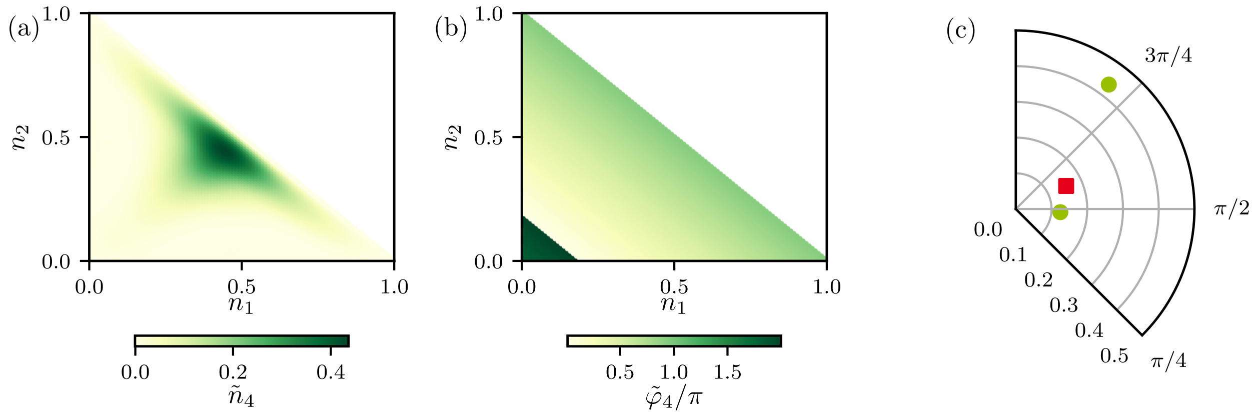

FIG. 9. Input-output relations (a) $\tilde{n}_4(n_1, n_2)$ and (b) $\tilde{\varphi}_4(n_1, n_2, 0, 0)$ of the FWM neuron to solve the XOR problem. (c) By choosing the outputs of the individual cases according to Table II (green, output 0; red, output 1), the XOR problem is solvable using a single FWM neuron.

However, a single complex-valued neuron is able to solve this problem [63].

1. Input and output encoding

To use the full value range of the nonlinear activation function of the FWM neuron to solve the XOR problem, an encoding scheme for the inputs and the output has to be developed. The inputs $x_{1,2}$ are chosen to lie on the positive real axis ($\kappa_j = 0$). While an input 0 is identified by $|x_j| = 0.3$, an input 1 is given by $|x_j| = 0.45$.

The weights w_j of the neuron are still allowed to possess nonvanishing phases ϑ_j. Therefore, the weighted inputs presented to the FWM neuron will be given by

$$\sqrt{n_j} = |w_j||x_j|, \quad \varphi_j = \vartheta_j. \tag{57}$$

As two input particle numbers, chosen to be n_1 and n_2, of the FWM neuron are set using this encoding, the third, in this case n_3, is automatically determined to ensure $\sum_j n_j = 1$. Consequently, the combinations of inputs n_1 and n_2 are constrained by $0 \leqslant n_1 + n_2 \leqslant 1$.

The particle number response of the FWM neuron to the inputs is completely determined by the input particle numbers $\tilde{n}_4(n_1, n_2)$. The neuron response in terms of the phase follows

$$\tilde{\varphi}_4(n_1, n_2, \varphi_1, \varphi_2) = \tilde{\varphi}_4(n_1, n_2, 0, 0) + \varphi_1 + \varphi_2. \tag{58}$$

These input-output relations can be seen in Fig. 9.

The possible outputs of the XOR problem are encoded in a similar fashion. An output 0 is encoded via $\tilde{n}_4 = 0.125$ and $\tilde{\varphi}_4 = 1.5$ or $\tilde{n}_4 = 0.435$ and $\tilde{\varphi}_4 = 2.5$ for the input cases [0,0] and [1,1] respectively. The output 1 is always encoded as $\tilde{n}_4 = 0.155$ and $\tilde{\varphi}_4 = 2$. The complete encoding of the XOR problem for the FWM neuron can be seen in Table II. The presented encoding is completely equivalent to the original XOR problem. Hence, it can be used to solve the problem by means of the FWM neuron.

TABLE II. Encoded input-output mapping for the XOR problem using the FWM neuron.

Input 1	Input 2	$\|x_1\|$	$\|x_2\|$	Output	$\tilde{n}_4$	$\tilde{\varphi}_4$
0	0	0.3	0.3	0	0.125	1.5
0	1	0.3	0.45	1	0.155	2.0
1	0	0.45	0.3	1	0.155	2.0
1	1	0.45	0.45	0	0.435	2.5

2. Training results

Starting from random initial weights, the update rules (55) and (56) are used to train the FWM neuron to solve the XOR problem. Training epochs are performed with $\mathcal{M} = 1000$ random samples. The learning rate of the phase $\lambda_p = 10^{-8}$ is kept constant for all epochs, while the absolute value learning rate λ_a is gradually reduced from 10^{-3} to 10^{-4} during the training. After each epoch, the performance of the neuron is evaluated by calculating the averaged square error $\mathcal{F}$ according to (51) using all four possible input-output pairs of the XOR problem.

As can be seen in Fig. 10, the FWM neuron is able to learn to solve the XOR problem. After 100 training epochs, the initial error is reduced to $\mathcal{F} = 7.8 \times 10^{-6}$. A sample is categorized as being identified correctly if the neuron output is within ± 0.005 in terms of the particle number and within ± 0.05 in terms of the phase of the desired value. At the end of the training procedure, every test sample is identified correctly.

VI. CONCLUSION AND PERSPECTIVES

We investigated the ideal FWM process in a three-dimensional homogeneous BEC. By introducing appropriate

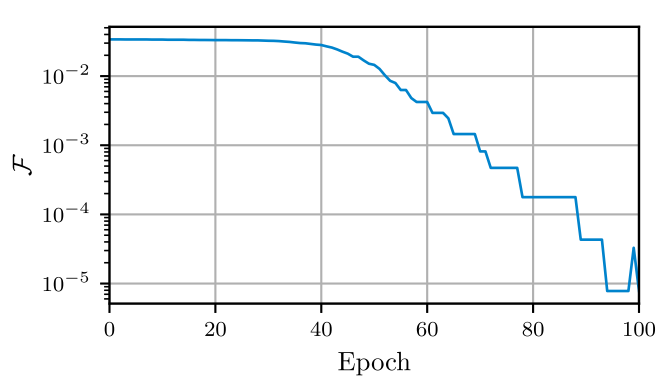

FIG. 10. Averaged square error $\mathcal{F}$ (51) over all four possible input-output pairs of the XOR problem versus the number of training epochs.

coordinates, we showed that the dynamics of the system exhibit Josephson-like oscillations, which can be described analytically by means of elliptic functions. These analytical expressions agree with numerical simulations of the Gross-Pitaevskii equation on a discrete periodic lattice. We investigated the influence of additional population outside of the FWM states on these dynamics. While the frequency of the oscillations changes, the main characteristics of the dynamics persist.

Identifying three complex amplitudes of the FWM setup as the input and the fourth amplitude as the output, we introduced an implementation for a complex-valued artificial neuron. We investigated the nonlinear activation function of the FWM neuron and showed its learning capabilities using steepest descent learning for complex-valued neurons. With the increased data throughput due to its complex-valued nature, the XOR problem can be solved using a single FWM neuron. After completing 100 learning epochs, it was able to identify every test sample presented to it correctly.

As the activation function of the FWM neuron is nonlinear, an implementation in an appropriate structure yields a deep neural network. To realize this, two key aspects have to be investigated: parallelization ability and communication between layers of the network. Theoretical investigations, as well as the study of experimental feasibility, have to be taken into account to realize a four-wave mixing neural network.

ACKNOWLEDGMENTS

This work is supported by the DLR German Aerospace Center with funds provided by the Federal Ministry for Economic Affairs and Climate Action (BMWK) due to an enactment of the German Bundestag under Grant No. 50MW2250E (QUANTUS+).

APPENDIX: FOUR-WAVE MIXING STATE PREPARATION

The desired state after initialization for FWM is a superposition of plane waves with wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$, and $\mathbf{k}_4$, fulfilling the conditions (7). There, all possible combinations for populations n_1, n_2, n_3, and n_4 with $\sum_{j=1}^{4} n_j = 1$ should be realizable.

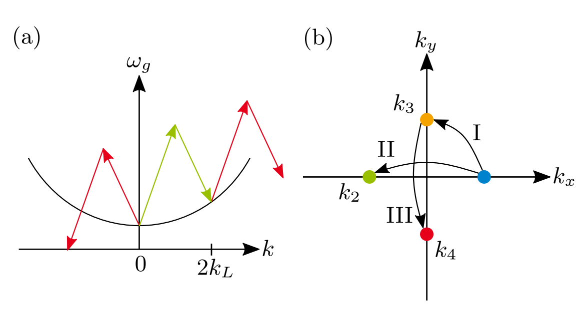

FIG. 11. (a) Energy diagram for Bragg diffraction versus wave number k. A ground-state BEC initially at rest experiences population transfer to a state with $2\mathbf{k}_L$ (green). Population transfer to other momentum states does not appear (red), as the conditions (A1) are not fulfilled. (b) Proposed initialization sequence for the FWM setup fulfilling (7). Three Bragg pulses are used to set up any combination of populations between the FWM states while ensuring that no population is transferred outside of the FWM states.

TABLE III. Particle numbers in the FWM states after each Bragg pulse of the sequence described in Fig. 11.

Pulse	n_0	n_1	n_2	n_3	n_4
0	0	1	0	0	0
I	0	$(1-p_2)$	0	p_2	0
II	0	$(1-p_3)(1-p_2)$	$p_3(1-p_2)$	p_2	0
III	0	$(1-p_3)(1-p_2)$	$p_3(1-p_2)$	$(1-p_4)p_2$	p_4p_2

We suggest using atomic beam splitters based on Bragg diffraction to populate the momentum states. This method is based on the interaction between the BEC in its internal ground state and two counterpropagating laser beams. In this scenario, energy and momentum conservation have to hold [44],

$$\hbar\omega_1 + \frac{\hbar^2 k_i^2}{2m} = \hbar\omega_2 + \frac{\hbar^2 k_f^2}{2m}, \quad \mathbf{k}_i + \mathbf{k}_1 = \mathbf{k}_f + \mathbf{k}_2, \tag{A1}$$

with the initial wave vector $\mathbf{k}_i$ and the final wave vector $\mathbf{k}_f$ of the BEC and the frequencies ω_1 and ω_2 and wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ of the laser beams, respectively. If the two laser beams are perfectly anticollinear, the momentum transfer in the BEC can be characterized as

$$\mathbf{k}_f - \mathbf{k}_i = \mathbf{k}_1 - \mathbf{k}_2 = 2\mathbf{k}_L, \tag{A2}$$

where $\mathbf{k}_L = (\mathbf{k}_1 + \mathbf{k}_2)/2$.

For a shallow lattice $[U(\mathbf{r}) = 0]$, the ground-state energy $\hbar\omega_g$ of the BEC scales quadratically with the wave number, $\omega_g \propto k^2$. Hence, the laser frequencies have to be chosen carefully, such that population transfer between momentum states is energetically permitted (see Fig. 11).

A controlled initialization of momentum states can be performed, as initial states can be targeted individually and final states are given by the momentum and energy conditions (A1).

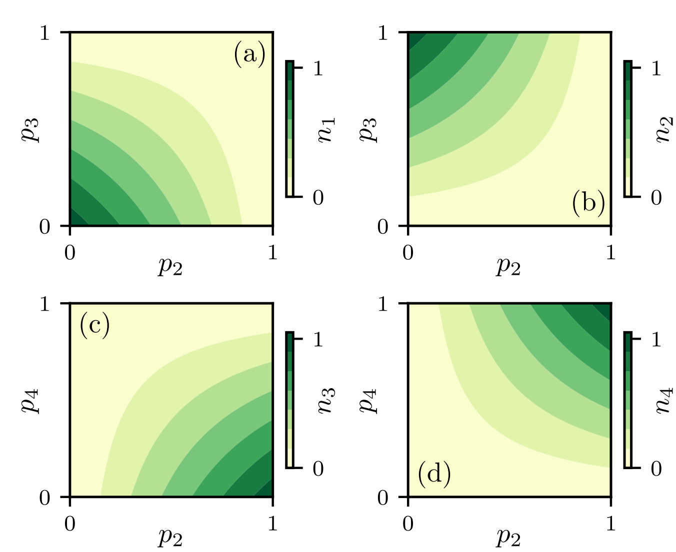

FIG. 12. Resulting particle numbers (a) n_1, (b) n_2, (c) n_3, and (d) n_4 for the initialization sequence described by Fig. 11 and Table III. All probabilities range from 0 to 1, conserving the total probability (particle number) (A3).

In Bragg diffraction, the portion of the population $0 \leqslant p_j \leqslant 1$ transferred between the momentum states can be controlled via the interaction duration between the BEC and the laser beams [44]. To avoid unwanted transitions outside of the FWM states, the preparation sequence shown in Fig. 11 and Table III is developed. For visualization, we choose $\mathbf{k}_1 = \hat{\mathbf{k}}_x$, $\mathbf{k}_2 = -\hat{\mathbf{k}}_x$, $\mathbf{k}_3 = \hat{\mathbf{k}}_y$, and $\mathbf{k}_4 = -\hat{\mathbf{k}}_y$. However, all combinations fulfilling (7) can be prepared by the described procedure. After the pulse sequence, the total particle number is transferred into the FWM states,

$$(1 - p_3)(1 - p_2) + p_3(1 - p_2) + (1 - p_4)p_2 + p_4 p_2 = 1, \tag{A3}$$

and all combinations can be realized (see Fig. 12).

[1] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature (London) **521**, 436 (2015).

[2] A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A. W. R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D. T. Jones, D. Silver, K. Kavukcuoglu, and D. Hassabis, Improved protein structure prediction using potentials from deep learning, Nature (London) **577**, 706 (2020).

[3] H. A. Elmarakeby, J. Hwang, R. Arafeh, J. Crowdis, S. Gang, D. Liu, S. H. AlDubayan, K. Salari, K. Kregel, C. Richter, T. E. Arnoff, J. Park, W. C. Hahn, and E. M. Van Allen, Biologically informed deep neural network for prostate cancer discovery, Nature (London) **598**, 348 (2021).

[4] S. Ravindran, Five ways deep learning has transformed image analysis, Nature (London) **609**, 864 (2022).

[5] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the game of Go with deep neural networks and tree search, Nature (London) **529**, 484 (2016).

[6] P. R. Wurman, S. Barrett, K. Kawamoto, J. MacGlashan, K. Subramanian, T. J. Walsh, R. Capobianco, A. Devlic, F. Eckert, F. Fuchs, L. Gilpin, P. Khandelwal, V. Kompella, H. C. Lin, P. MacAlpine, D. Oller, T. Seno, C. Sherstan, M. D. Thomure, H. Aghabozorgi *et al.*, Outracing champion Gran Turismo drivers with deep reinforcement learning, Nature (London) **602**, 223 (2022).

[7] N. Heimann, J. Petermann, D. Hartwig, R. Schnabel, and L. Mathey, Predicting the motion of a high-Q pendulum subject to seismic perturbations using machine learning, Appl. Phys. Lett. **122**, 254101 (2023).

[8] S. Mittal and S. Vaishay, A survey of techniques for optimizing deep learning on GPUs, J. Syst. Archit. **99**, 101635 (2019).

[9] S. Mittal, A Survey on optimized implementation of deep learning models on the NVIDIA Jetson platform, J. Syst. Archit. **97**, 428 (2019).

[10] C. Kaspar, B. J. Ravoo, W. G. van der Wiel, S. V. Wegner, and W. H. P. Pernice, The rise of intelligent matter, Nature (London) **594**, 345 (2021).

[11] H. P. Nautrup, T. Metger, R. Iten, S. Jerbi, L. M. Trenkwalder, H. Wilming, H. J. Briegel, and R. Renner, Operationally meaningful representations of physical systems in neural networks, Mach. Learn.: Sci. Technol. **3**, 045025 (2022).

[12] X. Sui, Q. Wu, J. Liu, Q. Chen, and G. Gu, A review of optical neural networks, IEEE Access **8**, 70773 (2020).

[13] Y. Kuratomi, A. Takimoto, K. Akiyama, and H. Ogawa, Optical neural network using vector-feature extraction, Appl. Opt. **32**, 5750 (1993).

[14] S. Gao, J. Yang, Z. Feng, and Y. Zhang, Implementation of a large-scale optical neural network by use of a coaxial lenslet array for interconnection, Appl. Opt. **36**, 4779 (1997).

[15] D. Psaltis, D. Brady, X.-G. Gu, and S. Lin, Holography in artificial neural networks, Nature (London) **343**, 325 (1990).

[16] T.-Y. Cheng, D.-Y. Chou, C.-C. Liu, Y.-J. Chang, and C.-C. Chen, Optical neural networks based on optical fiber-communication system, Neurocomputing **364**, 239 (2019).

[17] A. N. Tait, T. F. de Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, Neuromorphic photonic networks using silicon photonic weight banks, Sci. Rep. **7**, 7430 (2017).

[18] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, Deep learning with coherent nanophotonic circuits, Nat. Photonics **11**, 441 (2017).

[19] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, All-optical spiking neurosynaptic networks with self-learning capabilities, Nature (London) **569**, 208 (2019).

[20] H. Zhang, M. Gu, X. D. Jiang, J. Thompson, H. Cai, S. Paesani, R. Santagati, A. Laing, Y. Zhang, M. H. Yung, Y. Z. Shi, F. K. Muhammad, G. Q. Lo, X. S. Luo, B. Dong, D. L. Kwong, L. C. Kwek, and A. Q. Liu, An optical neural chip for implementing complex-valued neural network, Nat. Commun. **12**, 457 (2021).

[21] F. Ashtiani, A. J. Geers, and F. Aflatouni, An on-chip photonic deep neural network for image classification, Nature (London) **606**, 501 (2022).

[22] Y. Zuo, B. Li, Y. Zhao, Y. Jiang, Y.-C. Chen, P. Chen, G.-B. Jo, J. Liu, and S. Du, All-optical neural network with nonlinear activation functions, Optica **6**, 1132 (2019).

[23] A. Ryou, J. Whitehead, M. Zhelyeznyakov, P. Anderson, C. Keskin, M. Bajcsy, and A. Majumdar, Free-space optical neural network based on thermal atomic nonlinearity, Photonics Res. **9**, B128 (2021).

[24] M. Miscuglio, A. Mehrabian, Z. Hu, S. I. Azzam, J. George, A. V. Kildishev, M. Pelton, and V. J. Sorger, All-optical nonlinear activation function for photonic neural networks, Opt. Mater. Express **8**, 3851 (2018).

[25] S. Skinner, J. Steck, and E. Behrman, Optical neural network using Kerr-type nonlinear materials, in *Proceedings of the Fourth International Conference on Microelectronics for Neural Networks and Fuzzy Systems* (IEEE, Turin, 1994), pp. 12–15.

[26] A. Dejonckheere, F. Duport, A. Smerieri, L. Fang, J.-L. Oudar, M. Haelterman, and S. Massar, All-optical reservoir computer based on saturation of absorption, Opt. Express **22**, 10868 (2014).

[27] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, Observation of Bose-Einstein condensation in a dilute atomic vapor, Science **269**, 198 (1995).

[28] M. R. Andrews, C. G. Townsend, H.-J. Miesner, D. S. Durfee, D. M. Kurn, and W. Ketterle, Observation of interference between two Bose condensates, Science **275**, 637 (1997).

[29] P. D. Maker and R. W. Terhune, Study of optical effects due to an induced polarization third order in the electric field strength, Phys. Rev. **137**, A801 (1965).

[30] M. Trippenbach, Y. Band, and P. Julienne, Four wave mixing in the scattering of Bose-Einstein condensates, Opt. Express **3**, 530 (1998).

[31] M. Trippenbach, Y. B. Band, and P. S. Julienne, Theory of four-wave mixing of matter waves from a Bose-Einstein condensate, Phys. Rev. A **62**, 023608 (2000).

[32] Y. Wu, X. Yang, C. P. Sun, X. J. Zhou, and Y. Q. Wang, Theory of four-wave mixing with matter waves without the undepleted pump approximation, Phys. Rev. A **61**, 043604 (2000).

[33] C. Sun, C. Hang, G. Huang, and B. Hu, Investigation of four-wave mixing of matter waves in Bose-Einstein condensates, Mod. Phys. Lett. B **18**, 375 (2004).

[34] E. Rowen, R. Ozeri, N. Katz, R. Pugatch, and N. Davidson, Dressed-state approach to matter-wave mixing of bosons, Phys. Rev. A **72**, 053633 (2005).

[35] L. Deng, E. W. Hagley, J. Wen, M. Trippenbach, Y. Band, P. S. Julienne, J. E. Simsarian, K. Helmerson, S. L. Rolston, and W. D. Phillips, Four-wave mixing with matter waves, Nature (London) **398**, 218 (1999).

[36] J. M. Vogels, K. Xu, and W. Ketterle, Generation of macroscopic pair-correlated atomic beams by four-wave mixing in Bose-Einstein condensates, Phys. Rev. Lett. **89**, 020401 (2002).

[37] A. J. Leggett and F. Sols, On the concept of spontaneously broken gauge symmetry in condensed matter physics, Found. Phys. **21**, 353 (1991).

[38] A. J. Leggett, Bose-Einstein condensation in the alkali gases: Some fundamental concepts, Rev. Mod. Phys. **73**, 307 (2001).

[39] A. Smerzi, S. Fantoni, S. Giovanazzi, and S. R. Shenoy, Quantum coherent atomic tunneling between two trapped Bose-Einstein condensates, Phys. Rev. Lett. **79**, 4950 (1997).

[40] J. Williams, R. Walser, J. Cooper, E. Cornell, and M. Holland, Nonlinear Josephson-type oscillations of a driven, two-component Bose-Einstein condensate, Phys. Rev. A **59**, R31 (1999).

[41] J. Kronjäger, C. Becker, M. Brinkmann, R. Walser, P. Navez, K. Bongs, and K. Sengstock, Evolution of a spinor condensate: Coherent dynamics, dephasing, and revivals, Phys. Rev. A **72**, 063619 (2005).

[42] R. Walser, E. Goldobin, O. Crasser, D. Koelle, R. Kleiner, and W. Schleich, Semifluxons in superconductivity and cold atomic gases, New J. Phys. **10**, 045020 (2008).

[43] M. Grupp, W. P. Schleich, E. Goldobin, D. Koelle, R. Kleiner, and R. Walser, Emergence of atomic semifluxons in optical Josephson junctions, Phys. Rev. A **87**, 021602(R) (2013).

[44] A. Neumann, M. Gebbe, and R. Walser, Aberrations in (3+1)-dimensional Bragg diffraction using pulsed Laguerre-Gaussian laser beams, Phys. Rev. A **103**, 043306 (2021).

[45] W. Ertmer, C. Schubert, T. Wendrich, M. Gilowski, M. Zaiser, T. V. Zoest, E. Rasel, C. J. Bordé, A. Clairon, F. Landragin, P. Laurent, P. Lemonde, G. Santarelli, W. Schleich, F. S. Cataliotti, M. Inguscio, N. Poli, F. Sorrentino, C. Modugno, G. M. Tino *et al.*, Matter wave explorer of gravity (MWXG), Exp. Astron. **23**, 611 (2009).

[46] S. Marksteiner, R. Walser, P. Marte, and P. Zoller, Localization of atoms in light fields: Optical molasses, adiabatic compression and squeezing, Appl. Phys. B **60**, 145 (1995).

[47] E. P. Gross, Structure of a quantized vortex in boson systems, Nuovo Cimento **20**, 454 (1961).

[48] L. P. Pitaevskii, Vortex lines in an imperfect Bose gas, Sov. Phys. JETP **13**, 451 (1961)

[49] E. L. Hill, Hamilton's principle and the conservation theorems of mathematical physics, Rev. Mod. Phys. **23**, 253 (1951).

[50] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Photons and Atoms* (Wiley, New York, 1997).

[51] P. J. Martin, B. G. Oldaker, A. H. Miklich, and D. E. Pritchard, Bragg scattering of atoms from a standing light wave, Phys. Rev. Lett. **60**, 515 (1988).

[52] E. Noether, Invariante Variationsprobleme, Nachr. Ges. Wiss. Göttingen **1918**, 235 (1918).

[53] P. A. Braun, Discrete semiclassical methods in the theory of Rydberg atoms in external fields, Rev. Mod. Phys. **65**, 115 (1993).

[54] P. Reineker, M. Schulz, B. M. Schulz, and R. Walser, *Theoretische Physik*, 2nd ed. (Wiley-VCH, Weinheim, 2021).

[55] S. Raghavan, A. Smerzi, S. Fantoni, and S. R. Shenoy, Coherent oscillations between two weakly coupled Bose-Einstein condensates: Josephson effects, π oscillations, and macroscopic quantum self-trapping, Phys. Rev. A **59**, 620 (1999).

[56] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, NIST Digital Library of Mathemtical Functions, https://dlmf.nist.gov/.

[57] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain, Psychol. Rev. **65**, 386 (1958).

[58] A. Hirose, *Complex-Valued Neural Networks*, 2nd ed., Studies in Computational Intelligence Vol. 400 (Springer, Berlin, 2012).

[59] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).

[60] J. Bassey, L. Qian, and X. Li, A survey of complex-valued neural networks, arXiv:2101.12249.

[61] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. (Pearson Prentice Hall, New York, 2009).

[62] M. Minsky and S. A. Papert, *Perceptrons* (MIT Press, Cambridge, MA, 1969).

[63] T. Nitta, Solving the XOR problem and the detection of symmetry using a single complex-valued neuron, Neural Networks **16**, 1101 (2003).