

Engineering nonlinear activation functions for all-optical neural networks via quantum interference

RUBEN CANORA,^{1,†}  XINZHE XU,^{1,†} ZIQI NIU,^{1,†}  HADISEH ALAEIAN,^{1,2,3,4}  AND SHENGWANG DU^{1,2,3,5} 

¹Elmore Family School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA

²Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana 47907, USA

³Purdue Quantum Science and Engineering Institute, Purdue University, West Lafayette, Indiana 47907, USA

⁴halaeian@purdue.edu

⁵dusw@purdue.edu

[†]These authors contributed equally to this work.

Abstract: All-optical neural networks (AONNs) promise transformative gains in speed and energy efficiency for artificial intelligence (AI) by leveraging light's intrinsic parallelism and wave nature. However, their scalability has been fundamentally limited by the high power requirements of conventional nonlinear optical elements. Here, we present a low-power nonlinear activation scheme based on a three-level quantum system driven by dual laser fields. This platform introduces a two-channel nonlinear activation matrix with self- and cross-nonlinearity, enabling true multi-input, multi-output optical processing. The system supports tunable activation behaviors, including sigmoid and ReLU functions, at ultralow power levels (17 μW per neuron). We validate our approach through theoretical modeling and experimental demonstration in rubidium vapor cells, showing the feasibility of scaling to deep AONNs with millions of neurons operating under 20 W of total optical power. Crucially, we demonstrate the all-optical generation of gradient-like signals with backpropagation, paving the way for all-optical training. These results mark a significant advance toward scalable, high-speed, and energy-efficient optical AI hardware.

© 2025 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. introduction

All-optical neural networks (AONNs) have emerged as a promising photonic computing paradigm for artificial intelligence (AI) applications, with potential in leveraging the wave nature of light to achieve intrinsic parallelism, ultrafast speed, and low energy consumption [1–9]. These attributes position AONNs as attractive alternatives to electronic platforms such as CPUs, GPUs [10–12], TPUs, FPGAs [13,14], and ASICs [15,16], particularly as the scaling limits of Moore's law become increasingly apparent [17,18]. Despite their advantages, the scalability of deep AONNs remains limited by the challenge of implementing nonlinear activation functions using purely optical mechanisms. These nonlinearities are essential for enabling deep learning but typically require high optical power, limiting the practicality of large-scale, energy-efficient AONN systems. While linear operations can be performed efficiently even at the few-photon level [19], the absence of robust low-power optical nonlinear elements restricts the depth and computational expressiveness of current architectures. As a result, most large-scale ONNs have focused primarily on linear transformations, often relying on hybrid optical-electronic designs to compensate for the missing nonlinearity [20–22].

Several approaches have been explored to introduce optical nonlinearities into AONNs. For instance, phase-change materials have enabled chip-scale nonlinear functions, supporting

demonstrations of only single-layer optical neurons [23], but the implementations remain limited in scalability. Another effort employed electromagnetically induced transparency (EIT) [24,25] in a laser-cooled atomic ensemble to realize nonlinear activation in a two-layer AONN [2]. While conceptually significant, the large system footprint and operational complexity associated with laser cooling of atoms hindered its practical deployment. Recently, saturated absorptions of thermal atomic vapor cells found applications in free-space AONNs [26–28], though their nonlinearities are usually very weak with limited tunability. To date, experimental realizations of multi-layer AONNs have typically been constrained to systems with only a few hundred hidden neurons [29], underscoring the need for low-power, scalable solutions that can support deep architectures.

In this work, we present a quantum-optical strategy for implementing nonlinear activation functions using a three-level quantum medium driven by two laser fields. In contrast to conventional designs that implement single-channel scalar nonlinearities, our method enables a two-channel activation function that simultaneously exhibits self- and cross-nonlinear behavior. These nonlinearities arise from quantum interference between atomic transition pathways. Through a combination of theoretical modeling in both homogeneous and Doppler-broadened regimes, as well as experimental demonstration with a rubidium (Rb) vapor cell, we show that the system can emulate rectified linear unit (ReLU) and sigmoid activation functions at optical power levels as low as 17 μ W per neuron. This approach enables the design of deep AONNs containing millions of neurons operating with only less than 100-watt-level laser sources. We further show that the scheme allows for backpropagation all-optical training. Our numerical simulations and experimental measurements confirm the scalability and energy efficiency of this architecture, offering a compelling pathway toward practical, high-speed, all-optical computing.

This article is organized as follows. Section 2 introduces the general formulation of a three-level quantum system, including the steady-state density matrix and light propagation equations, focusing on lifetime-broadened optical media. Section 3 extends the analysis to Doppler-broadened media, confirmed by experimental measurements. Section 4 summarizes the work and suggests directions for future research.

2. Lifetime-broadened medium

Conventional non-linear optical processes typically operate at frequencies far detuned from resonances to avoid significant linear absorption [30,31]. The n^{th} order of nonlinear susceptibility can be expressed as $\chi^{(n)} \propto \frac{1}{\Delta_1 + i\gamma_1} \frac{1}{\Delta_2 + i\gamma_2} \cdots \frac{1}{\Delta_n + i\gamma_n}$, where Δ_m and γ_m denote the frequency of detuning and dephasing of the m -photon process, respectively. To mitigate resonant absorption, Δ_m is generally required to be much larger than the absorption linewidth. Consequently, non-linear optics in solid-state materials requires high light intensities due to their substantial absorption bandwidths, typically on the order of terahertz (THz) [30].

On the other hand, both real and artificial atoms, such as quantum dots [32] and defects in solids [33], with discrete energy levels, exhibit much narrower resonance linewidths, primarily determined by lifetime, making them promising candidates for achieving large nonlinear susceptibilities. Moreover, phenomena such as EIT [24,25] and coherent population trapping [34] have demonstrated that resonant linear absorption can be suppressed or even eliminated through quantum interference in a multilevel system, thereby significantly improving nonlinear optical effects [35,36]. We recognize that quantum interference originating from internal transitions in these systems offers a controllable approach to engineering nonlinear activation functions for optical computing. While early demonstrations of AONNs using EIT have shown conceptual promise [2], their nonlinear activation functions have remained highly specialized, and a general framework for engineering diverse nonlinearities suited for machine learning remains lacking.

Here, we systematically investigate a three-level lifetime-broadened atomic medium and demonstrate the realization of both sigmoid and ReLU-like activation functions at ultralow optical

power via quantum interference, surpassing the limitations of previous EIT work [2]. Although our numerical analysis is based on ^{87}Rb atoms, the conclusions are broadly applicable to other lifetime-broadened systems.

Figure 1(a) presents the energy level diagram of a three-level system, where $|1\rangle$ and $|2\rangle$ are two long-lived hyperfine ground states and $|3\rangle$ is an excited state. We define $\omega_{31} = (E_3 - E_1)/\hbar$ as the resonance frequency for the $|1\rangle \rightarrow |3\rangle$ transition and $\omega_{32} = (E_3 - E_2)/\hbar$ as the resonance frequency for the $|2\rangle \rightarrow |3\rangle$ transition. Optical field 1 (depicted in blue) couples the $|1\rangle \rightarrow |3\rangle$ transition with a Rabi frequency Ω_1 and one-photon detuning $\Delta_1 = \omega_1 - \omega_{31}$, where ω_1 is the angular frequency of the first laser. Similarly, the optical field 2 (depicted in red) couples the transition $|2\rangle \rightarrow |3\rangle$ with the Rabi frequency Ω_2 and the detuning $\Delta_2 = \omega_2 - \omega_{32}$, where ω_2 is the frequency of the second laser. Γ_{31} and Γ_{32} represent the population decay rates from $|3\rangle$ to $|1\rangle$ and $|2\rangle$, respectively. In addition, Γ_{12} and Γ_{21} denote incoherent population transfer rates between the two ground states $|1\rangle$ and $|2\rangle$, typically arising from collisions. Quantum interference among dipole transitions driven by external fields gives rise to diverse nonlinear optical responses at low light levels. In an EIT system, for instance, a resonant control field Ω_2 induces destructive interference that renders the probe transition Ω_1 transparent [24,25].

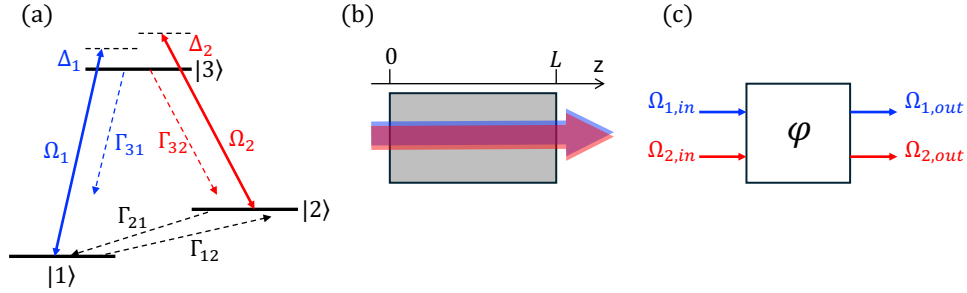


Fig. 1. Schematic of the three-level nonlinear optical medium. (a) The energy level diagram of a three-level quantum system with two driving lasers and decay mechanisms, which are considered in this work. (b) Optical setup showing the alignment of two laser beams and their propagation in the medium. (c) Simplified circuit diagram for the two-channel (2-input \times 2-output) nonlinear activation function unit.

In the Markovian limit, the atomic density matrix $\hat{\rho}$ evolves according to the following master equation

$$\dot{\hat{\rho}} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] + \mathcal{D}(\hat{\rho}). \quad (1)$$

The system Hamiltonian \hat{H} accounting for the coherent dynamics in the rotating reference frame is:

$$\hat{H} = -\hbar(\Delta_1 - \Delta_2)|2\rangle\langle 2| - \hbar\Delta_1|3\rangle\langle 3| - \frac{1}{2}\hbar(\Omega_1|3\rangle\langle 1| + \Omega_2|3\rangle\langle 2| + \text{H.C.}). \quad (2)$$

The dissipative term $\mathcal{D}(\hat{\rho})$ describes decoherence via jump operators \hat{L}_m with rates Γ_m :

$$\mathcal{D}(\hat{\rho}) = \sum_m \frac{\Gamma_m}{2} \left(2\hat{L}_m \hat{\rho} \hat{L}_m^\dagger - \{\hat{L}_m^\dagger \hat{L}_m, \hat{\rho}\} \right). \quad (3)$$

For lifetime-broadened media, decoherence is dominated by spontaneous emission, modeled by $\hat{L}_1 = \sqrt{\Gamma_{31}}|1\rangle\langle 3|$ and $\hat{L}_2 = \sqrt{\Gamma_{32}}|2\rangle\langle 3|$. Collisional dephasing is included via $\hat{L}_3 = \sqrt{\Gamma_{12}}|1\rangle\langle 2|$.

The steady-state elements ρ_{mn} of the density matrix are obtained by solving Eq. (1):

$$0 = \dot{\rho}_{11} = -\frac{i}{2}(\Omega_1\rho_{13} - \Omega_1^*\rho_{31}) - \Gamma_{12}\rho_{11} + \Gamma_{21}\rho_{22} + \Gamma_{31}\rho_{33}, \quad (4)$$

$$0 = \dot{\rho}_{22} = -\frac{i}{2}(\Omega_2\rho_{23} - \Omega_2^*\rho_{32}) + \Gamma_{12}\rho_{11} - \Gamma_{21}\rho_{22} + \Gamma_{32}\rho_{33}, \quad (5)$$

$$0 = \dot{\rho}_{21} = -i(\Delta_2 - \Delta_1 - i\gamma_{12})\rho_{21} - \frac{i}{2}(\Omega_1\rho_{23} - \Omega_2^*\rho_{31}), \quad (6)$$

$$0 = \dot{\rho}_{31} = \frac{i}{2}[\Omega_1(\rho_{11} - \rho_{33}) + \Omega_2\rho_{21}] + i(\Delta_1 + i\gamma_{13})\rho_{31}, \quad (7)$$

$$0 = \dot{\rho}_{32} = \frac{i}{2}[\Omega_2(\rho_{22} - \rho_{33}) + \Omega_1\rho_{12}] + i(\Delta_2 + i\gamma_{23})\rho_{32}. \quad (8)$$

where $\Gamma_3 = \Gamma_{31} + \Gamma_{32}$, and γ_{12} , γ_{13} , and γ_{23} are the decoherence rates of non-diagonal elements. For cold atoms, $\gamma_{13} = (\Gamma_3 + \Gamma_{12})/2$ and $\gamma_{23} = (\Gamma_3 + \Gamma_{21})/2$. The ground-state dephasing rate can be expressed as $\gamma_{12} = \gamma_{12,0} + (\Gamma_{12} + \Gamma_{21})/2$, where $\gamma_{12,0}$ is the residual dephasing. Note that we only wrote the equations of motion for distinct density matrix entries, taking into account the conservation of particle numbers $\rho_{11} + \rho_{22} + \rho_{33} = 1$ and the hermiticity of the density matrix, i.e. $\rho_{mn} = \rho_{nm}^*$.

Figure 1(b) shows the optical setup, where the two spatially overlapped beams propagate along the z direction through a medium of length L . The beam propagation is governed by the slowly varying envelope equations:

$$\frac{\partial}{\partial z}\Omega_n = i\frac{k_n}{\varepsilon_0\hbar}N|\mu_{n3}|^2\rho_{3n}, \quad (9)$$

where $n = 1, 2$, $k_n = \omega_n/c$ is the wavenumber of the corresponding beam, N is the atomic density, and μ_{n3} is the transition dipole moment between $|n\rangle$ and $|3\rangle$.

Given input fields $\Omega_{1(2),in}$ at $z = 0$, and after formally solving Eqs. (4)–(9), the output fields can be written as

$$\Omega_{n,out} = \varphi_n(\Omega_{1,in}, \Omega_{2,in}). \quad (10)$$

where φ_n with $n = 1, 2$ denote nonlinear activation functions of $\Omega_{1,in}$ and $\Omega_{2,in}$. Figure 1(c) illustrates the simplified icon for this two-channel (two-input, two-output) nonlinear activation function unit. In the extreme case of a single input (i.e., $\Omega_2 = 0$), the system reduces to the well-known self-saturation effect of two-level atoms [37–39]. When a second laser field Ω_2 with moderate power is applied, it modifies the population distribution and consequently alters the nonlinear response of Ω_1 . As Ω_2 becomes stronger, the Ω_1 response transitions into the EIT regime [24,25]. In the following, we analyze different configurations through detailed theoretical modeling and numerical simulations.

2.1. EIT in the weak-probe limit

We begin with the simplest case of electromagnetically induced transparency (EIT) in a cold atomic ensemble, under the weak-probe approximation where $|\Omega_{1,in}| \ll |\Omega_{2,in}|$, and neglect collisional effects, i.e., $\Gamma_{12} = \Gamma_{21} = 0$. Under these conditions, the population remains predominantly in state $|1\rangle$, i.e., $\rho_{11} \approx 1$, and the non-depletion approximation applies. Since states $|2\rangle$ and $|3\rangle$ are barely populated, the control field remains unaffected: $\Omega_{2,out} = \Omega_{2,in} = \Omega_2$. Consequently,

Eq. (9) reduces to a linear propagation equation for the probe field, yielding

$$\Omega_{1,\text{out}} = \Omega_{1,\text{in}} e^{ik_1 L \sqrt{1+\chi_1}} \cong \Omega_{1,\text{in}} e^{ik_1 L} e^{ik_1 \chi_1 L/2}, \quad (11)$$

where the linear susceptibility $|\chi_1| \ll 1$ is given by

$$\chi_1(\Delta_1) = \frac{4N|\mu_{13}|^2(\Delta_1 - \Delta_2 + i\gamma_{12})/(\epsilon_0 \hbar)}{|\Omega_2|^2 - 4(\Delta_1 + i\gamma_{13})(\Delta_1 - \Delta_2 + i\gamma_{12})}. \quad (12)$$

In the ideal case with no ground-state dephasing ($\gamma_{12} = 0$), the susceptibility vanishes at two-photon resonance ($\Delta_1 = \Delta_2$), yielding $\chi_1(0) = 0$, and thus full transparency for the probe field 1. Notably, this transparency is independent of the strength of the control field, as even an arbitrarily weak control field can render the medium transparent for a much weaker probe field. To enable nonlinear activation, we consider a finite dephasing rate $\gamma_{12} \neq 0$, which introduces an imaginary component to the susceptibility:

$$\chi_1(0) = \frac{i4N|\mu_{13}|^2\gamma_{12}/(\epsilon_0 \hbar)}{|\Omega_2|^2 + 4\gamma_{13}\gamma_{12}}. \quad (13)$$

Substituting Eq. (13) into Eq. (11), the output field becomes

$$\Omega_{1,\text{out}}(\Delta_1 = 0) = \Omega_{1,\text{in}} e^{ik_1 L} e^{-\text{OD} \frac{2\gamma_{12}\gamma_{13}}{|\Omega_2|^2 + 4\gamma_{13}\gamma_{12}}}, \quad (14)$$

or equivalently

$$|\Omega_{1,\text{out}}(\Delta_1 = 0)|^2 = |\Omega_{1,\text{in}}|^2 e^{-\text{OD} \frac{4\gamma_{12}\gamma_{13}}{|\Omega_2|^2 + 4\gamma_{13}\gamma_{12}}}, \quad (15)$$

where $\text{OD} = N\sigma_{13}L$ is the optical depth for the $|1\rangle \rightarrow |3\rangle$ transition without the control field, and $\sigma_{13} = k_1|\mu_{13}|^2/(\epsilon_0 \hbar \gamma_{13})$ is the resonant absorption cross section. Equation (15) shows that the probe output power is nonlinearly controlled by the control field intensity $|\Omega_2|^2$, forming the basis of an optical activation function with $\Omega_{2,\text{in}}$ as input and $\Omega_{1,\text{out}}$ as output, as previously demonstrated [2,29]. Here, we call field 1 the probe field and field 2 the control field.

To illustrate this nonlinear mechanism, we simulate transmission through a cold ^{87}Rb atomic ensemble, as in Refs. [2,29]. We use states $|1\rangle = |5S_{1/2}, F=1\rangle$, $|2\rangle = |5S_{1/2}, F=2\rangle$, and $|3\rangle = |5P_{1/2}, F=2\rangle$, setting $\Delta_2 = 0$, $\Gamma_3 = 2\pi \times 6$ MHz, and $\gamma_{13} = \gamma_{23} = \Gamma_3/2$. To estimate the required laser powers, we assume the medium length $L = 2$ cm. We take a beam radius of $71 \mu\text{m}$ such that the diffraction length, for the wavelength of the ^{87}Rb D1 line of 795 nm, is $b = 4$ cm, which is much longer than the medium length.

Figure 2(a) shows the probe transmission spectrum for $\text{OD} = 50$ and three dephasing rates: $\gamma_{12,a} = 0.0001\Gamma_3$, $\gamma_{12,b} = 0.1\Gamma_3$, and $\gamma_{12,c} = 0.5\Gamma_3$. At resonance ($\Delta_1 = 0$), the transmission increases with decreasing γ_{12} and increasing control field power. As shown in Fig. 2(b), the EIT peak saturates at unity for large $|\Omega_2|^2$. The slope of the transmission curve depends on γ_{12} , consistent with Eq. (15) and prior experiments [2,29].

To expand beyond the resonance EIT explored in prior work, we demonstrate that the turning point of the nonlinear activation function can be tuned by the probe detuning Δ_1 . As shown in Fig. 2(c), for $\Delta_1 \neq 0$, the probe initially experiences absorption. As the control field increases, the EIT bandwidth broadens and eventually encompasses the probe detuning, resulting in transparency. Figure 2(d) shows the resulting sigmoid-like activation functions with tunable thresholds, enabling flexible reconfiguration—a capability not previously demonstrated experimentally.

2.2. Two-channel nonlinear optical activation functions

In the above EIT-based nonlinear activation function described above, the strong control field (field 2) serves as the input, and the weak probe field (field 1) is the output, with atomic population

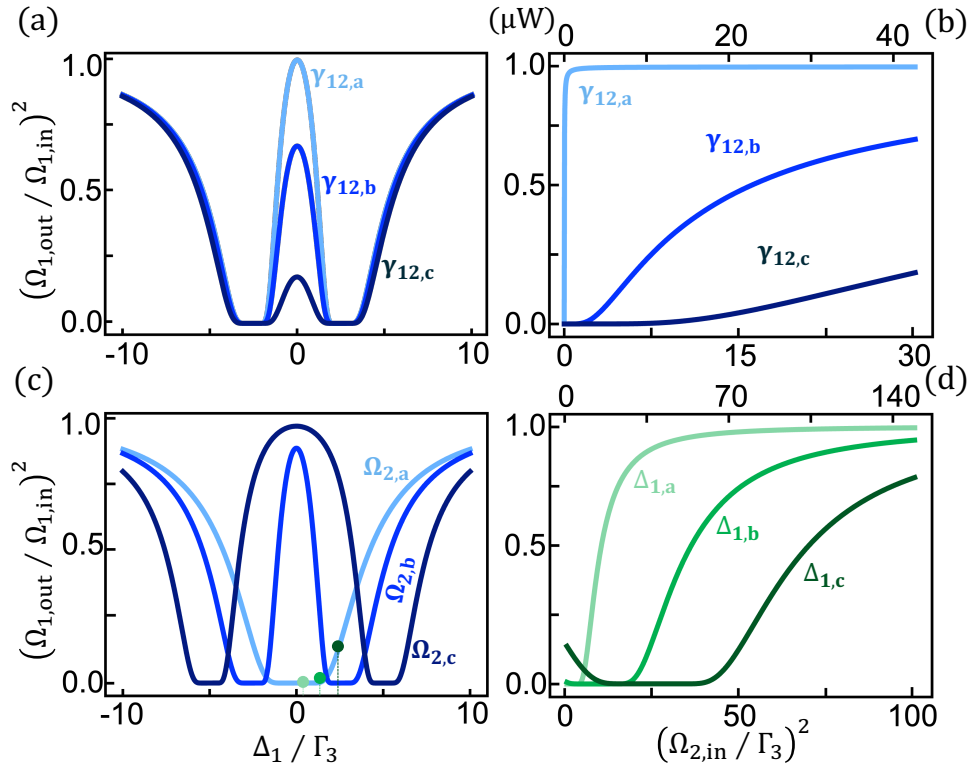


Fig. 2. EIT in the weak probe regime. $\Delta_2 = 0$. (a) EIT transmission spectra for varying ground-state decoherence rates, with $(\gamma_{12,a}, \gamma_{12,b}, \gamma_{12,c}) = (0.0001, 0.1, 0.5) \Gamma_3$, and a fixed control field $\Omega_2 = 5 \Gamma_3$. (b) EIT nonlinear activation functions with different ground-state decoherence rates γ_{12} , expressed as the cross-nonlinearity between the probe output $|\Omega_{1,out}|^2$ and the control input $|\Omega_2|^2$ [Eq. (15)]. (c) EIT transmission spectra for different control Rabi frequencies $(\Omega_{2,a}, \Omega_{2,b}, \Omega_{2,c}) = (0, 5, 10) \Gamma_3$ at fixed $\gamma_{12} = 0.03 \Gamma_3$. Green markers denote selected probe detunings $(\Delta_{1,a}, \Delta_{1,b}, \Delta_{1,c}) = (0.5, 1.5, 2.5) \Gamma_3$. (d) Corresponding nonlinear activation functions for the three detuning values shown in (c), illustrating tunable sigmoid-like response.

primarily in state $|1\rangle$. In a multilayer AONN architecture, however, the roles of control and probe fields must alternate between layers. This creates a challenge: the output probe field from one layer becomes the control field for the next, while a new, even weaker probe field must be introduced. Due to the EIT condition $|\Omega_p| \ll |\Omega_c|$, the signal power diminishes with increasing network depth. In an ideal EIT system with zero ground-state dephasing ($\gamma_{12} = 0$), such a layered scheme remains functional as long as the weak-probe condition holds. In practice, however, a finite ground-state dephasing rate $\gamma_{12} \neq 0$ imposes a power threshold on the control field, requiring $|\Omega_c|^2 > 4 \text{OD} \gamma_{12} \gamma_{13}$ to maintain transparency. This limits the maximum achievable depth of the AONN.

To overcome this limitation, we propose a two-field scheme in which both laser fields have comparable intensities. Although no analytical expression exists for this case, a qualitative understanding can be drawn. When $\Omega_1 \approx \Omega_2$, the atomic population distribution between states $|1\rangle$ and $|2\rangle$ becomes dependent on the relative field strengths. As Ω_2 is increased from zero and remains weaker than Ω_1 , atoms are increasingly pumped into state $|1\rangle$, enhancing the absorption of field 1. We refer to this process as *electromagnetically induced absorption* (EIA). As Ω_2 exceeds Ω_1 , EIT becomes dominant, and absorption of field 1 decreases. This

transition from EIA to EIT leads to a sigmoid-like nonlinear activation function, with the turning point determined by the crossover. Figure 3 shows simulated resonant ($\Delta_1 = \Delta_2 = 0$) probe transmissions $(\Omega_{1,\text{out}}/\Omega_{1,\text{in}})^2$ for various input probe intensities. Panel (a) corresponds to a small ground-state population transfer rate $\Gamma_{12} = 0.1\Gamma_3$, while panel (b) uses $\Gamma_{12} = \Gamma_3$. In both cases, the sigmoid nonlinearity breaks down for $\Omega_{1,\text{in}} > 3\Gamma_3$ but remains effective at moderate intensities ($\Omega_{1,\text{in}} \leq 3\Gamma_3$), comparable to Ω_2 . The presence of finite Γ_{12} enhances absorption at lower coupling powers, effectively mitigating the degradation in sigmoid behavior due to increasing probe intensity. However, above a certain power threshold, the nonlinear response becomes ineffective.

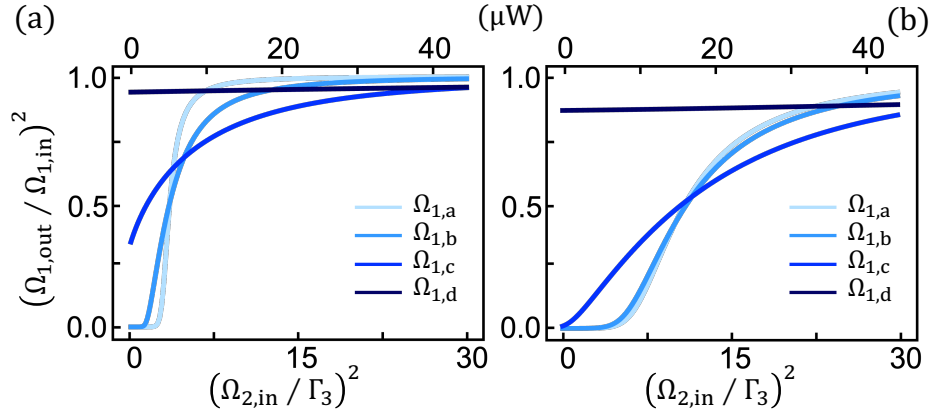


Fig. 3. Power effect on the resonant probe transmission ($\Delta_1 = \Delta_2 = 0$). Control-probe nonlinear functions with different probe inputs $(\Omega_{1,a}, \Omega_{1,b}, \Omega_{1,c}, \Omega_{1,d}) = (10^{-5}, 1, 3, 10)\Gamma_3$ for (a) $\Gamma_{12} = 0.1\Gamma_3$ and (b) $\Gamma_{12} = \Gamma_3$.

We next examine the self-nonlinearity of field 1 in the presence of a fixed field 2. Initially, when $|\Omega_1| < |\Omega_2|$, the atomic population resides primarily in state $|1\rangle$, resulting in high absorption. As Ω_1 increases, atoms are pumped out of state $|1\rangle$, reducing absorption and giving rise to a ReLU-like activation function. Unlike simple saturation in a two-level system, this three-level quantum interference allows for tunable nonlinear responses, as verified in numerical simulations.

Figure 4 presents numerical results for nonlinear activation functions with comparable fields. Panel (a) shows sigmoid-type responses—output probe power $|\Omega_{1,\text{out}}|^2$ as a function of control field power $|\Omega_{2,\text{in}}|^2$ —for different probe detunings Δ_1 , under $\Gamma_{12} = \Gamma_{21} = 0$ and $\Delta_2 = 0$. Similar to the EIT case in Fig. 2(d), the turning point of the nonlinearity shifts to higher control powers as Δ_1 increases. A narrow transmission spike near $\Omega_{2,\text{in}} = 0$ is caused by the pumping effect of the strong field 1, and is suppressed when finite Γ_{12} is included, as shown in panel (b). Panel (c) shows ReLU-type self-nonlinearity activation functions for field 1, under varying control powers. The turning point of the response shifts with $|\Omega_{2,\text{in}}|^2$, offering dynamic control. When field 2 is on resonance ($\Delta_2 = 0$), it induces a dressed-state splitting of level $|3\rangle$ into two levels separated by $\hbar\Omega_2$. The optimal detuning for field 1 is then $\Delta_1 = \Omega_2/2$ to maximize interaction. Panel (d) illustrates how different values of Γ_{12} affect the response at this detuning.

With both fields at comparable strengths, we further demonstrate two-channel (2-input \times 2-output) nonlinear activation functions, suitable for multi-input, multi-output (MIMO) photonic architectures. Figure 5 displays the output intensities of each beam as functions of their own and the other beam's input intensities under symmetric detunings $\Delta_1 = -\Delta_2 = \Gamma_3/3$. The self-nonlinearity responses in panels (a) and (d) are similar for both fields. The cross-nonlinearity in panels (b) and (c) confirm that either beam can modulate the other, enabling flexible design of reconfigurable, dual-channel activation units. We note that the cross-nonlinearity requires

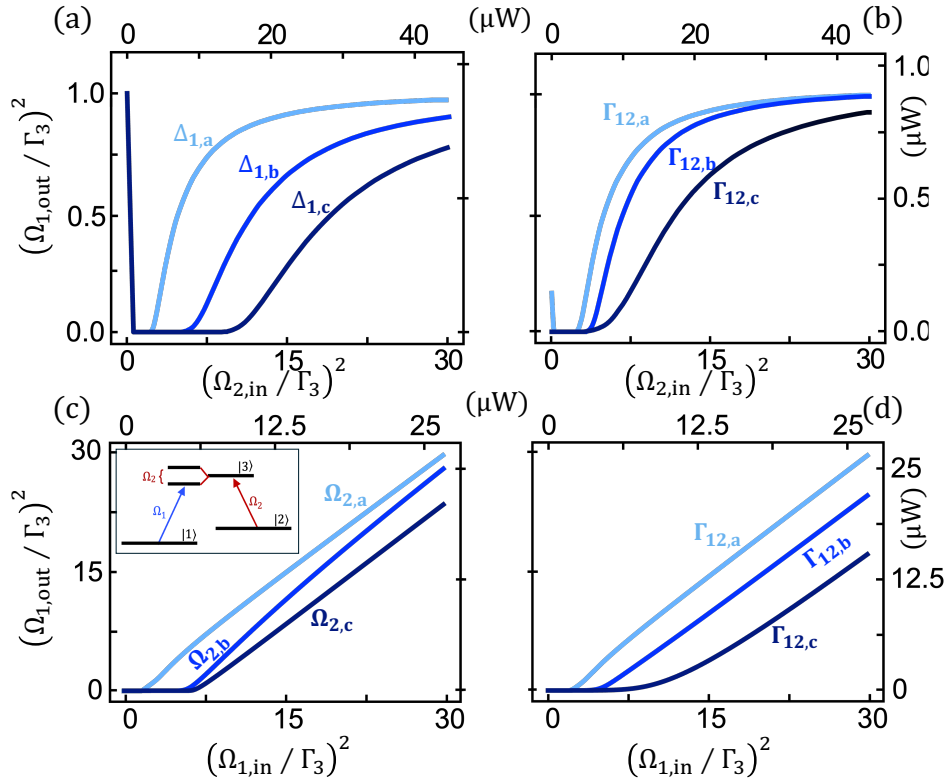


Fig. 4. Nonlinear optical activation functions with comparable driving fields. $\Delta_2 = 0$. (a) Control-probe cross nonlinear activation functions between $|\Omega_{1,out}|^2$ and $|\Omega_{2,in}|^2$ with different probe detuning: $\Delta_1 = (1/3, 2/3, 1)\Gamma_3$. Other parameters: $\Omega_{1,in} = \Gamma_3$, and $\Gamma_{12} = 0$. (b) Control-probe cross nonlinear activation functions between $|\Omega_{1,out}|^2$ and $|\Omega_{2,in}|^2$ with different ground state population transfer rate: $\Gamma_{12} = (0.01, 0.1, 1)\Gamma_3$. Other parameters: $\Omega_1 = \Gamma_3$, and $\Delta_1 = 1/3\Gamma_3$. (c) Probe self nonlinear activation functions with different control powers: $(\Omega_{2,a}, \Omega_{2,b}, \Omega_{2,c}) = (1, 3, 10)\Gamma_3$. Other parameters: $\Delta_1 = \Omega_2/2$, and $\Gamma_{12} = 0$. (d) Probe self nonlinear activation functions with different ground state population transfer rate: $(\Gamma_{12,a}, \Gamma_{12,b}, \Gamma_{12,c}) = (0.01, 0.1, 1)\Gamma_3$. Other parameters are $\Omega_2 = \Gamma_3$, and $\Delta_1 = \Omega_2/2$.

a significantly higher power in the control beam [Fig. 5(b) and (c)], which may prevent its application in a very deep AONN where the roles of control and probe beams need to be exchanged. On the other side, the self-nonlinearity can go much deeper: as the light goes deeper in an AONN, we can tune the turning point of the ReLU function to a lower power to compensate for the optical loss. In constructing a deep AONN, the dual-channel self-nonlinearity are employed as the nonlinear activation elements. In contrast, the cross-nonlinearity are used exclusively to fine-tune the threshold points of the self-nonlinearity ReLU responses. Therefore, we expect the two-channel ReLU functions from self-nonlinearity to play more important roles in deep AONNs.

3. Doppler-broadened medium

The preceding results highlight the versatility and tunability of nonlinear activation functions in lifetime-broadened media. While our simulations were based on cold atoms, the underlying principles are broadly applicable to other atom-like platforms, such as quantum dots and solid-state defect centers [32,33]. In parallel, thermal atomic vapors—despite exhibiting Doppler

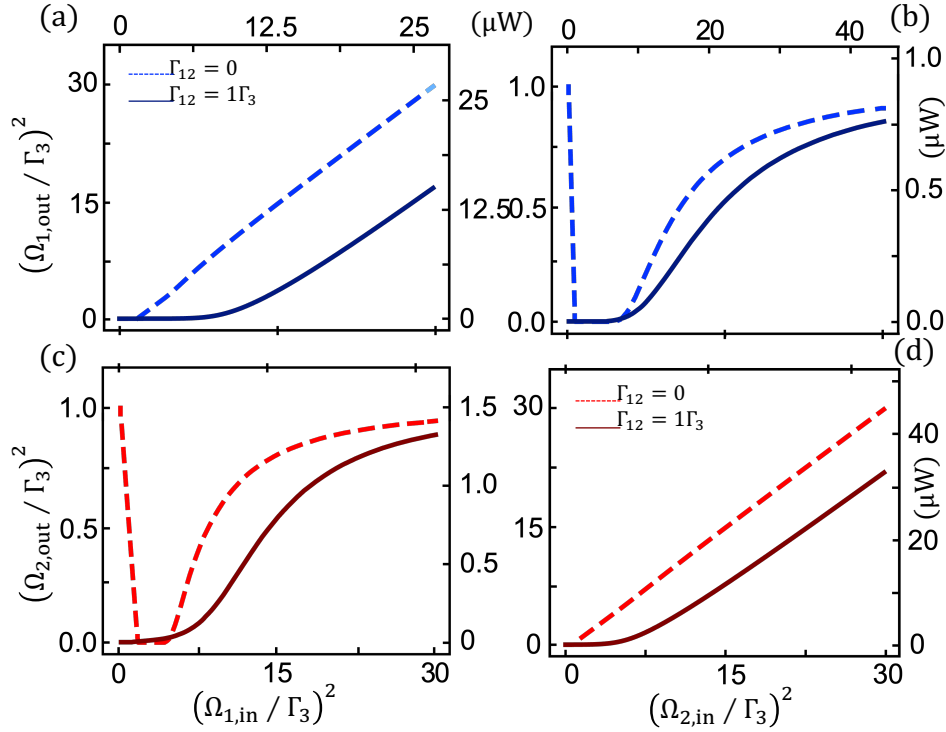


Fig. 5. Two-channel (2-input \times 2-output) nonlinear activation functions with self- and cross-nonlinearity in a lifetime-broadened atomic medium. (a) Self- and (b) cross-nonlinearity of input 1, respectively. (c) Cross- and (d) self-nonlinearity of input 2, respectively. We set $\Omega_{1,\text{in}} = \Gamma_3$ in (b) and $\Omega_{2,\text{in}} = \Gamma_3$ in (c). $\Delta_1 = -\Delta_2 = 1/3\Gamma_3$.

broadening, elevated dephasing, and collisional effects—remain attractive candidates for scalable implementations due to their simplicity and integrability. Recent advances in coupling thermal vapors with nanophotonic devices open pathways for hybrid architectures that combine nonlinear activation with on-chip linear photonic circuits [40–42]. To evaluate the robustness of the aforementioned results to decoherence and Doppler broadening, we investigate the achievable nonlinearity with a thermal vapor platform.

With the same configuration shown in Fig. 1(a),(b), along the beams, atoms follow the one-dimensional Maxwell-Boltzmann velocity distribution

$$f(v) = \frac{1}{v_{\text{th}}\sqrt{\pi}} e^{-(v/v_{\text{th}})^2}, \quad (16)$$

where v_{th} is the most probable thermal speed related to the atomic mass m and temperature T as

$$v_{\text{th}} = \sqrt{2k_B T/m}, \quad (17)$$

with k_B the Boltzmann constant. An atom moving at velocity v experiences a Doppler shift, such that a laser field with nominal detuning Δ_0 acquires an effective detuning:

$$\Delta_n(v) = \Delta_0 - k_n v, \quad (18)$$

for $k_n = \omega_n/c$ being the wave number of field n . The averaged $\bar{\rho}_{mn}$ can then be written as the weighted integral of the density matrices of all velocity classes as

$$\bar{\rho}_{mn} = \int_{-\infty}^{+\infty} dv \rho_{mn}(\Omega, \Delta_n(v)) f(v), \quad (19)$$

where $\rho_{mn}(v)$ is the steady-state solution of the density matrix determined through Eqs. (4)–(8).

To experimentally verify our theoretical predictions, we use a cylindrical glass cell filled with pure ^{87}Rb vapor, with a length of 2 cm and a diameter of 2 cm. Both laser beams are focused into the cell with a Gaussian waist radius (at $1/e^2$ intensity) measured to be $75 \mu\text{m}$.

3.1. ReLU nonlinear activations and transient responses

Figure 6(a) shows the measured ReLU-type nonlinear activation function, where the output power $P_{1,\text{out}}$ of beam 1 is plotted against its input power $P_{1,\text{in}}$ at $T = 75^\circ\text{C}$ for three different input powers of beam 2: $P_{2,\text{in},a} = 1.3 \mu\text{W}$ (blue squares), $P_{2,\text{in},b} = 13.3 \mu\text{W}$ (blue triangles), and $P_{2,\text{in},c} = 150 \mu\text{W}$ (blue circles). As expected, increasing the control power shifts the turning point of the ReLU function toward higher $P_{1,\text{in}}$. Solid curves represent numerical simulations fitted with ground-state population transfer rates $\Gamma_{12} = 2\pi \times (0.09, 0.132, 0.6) \text{ MHz}$, respectively. Notably, a single constant Γ_{12} cannot reproduce all experimental curves simultaneously. In particular, the highest-power case ($P_{2,\text{in},c}$) deviates from the model, likely due to the simplified three-level approximation neglecting hyperfine and Zeeman substructure. Nonetheless, the experimental data show good qualitative agreement with theoretical predictions. While the simplified three-level model presented here serves primarily as a guideline by capturing the essential physical mechanism, these quantitative deviations do not affect the scalability or consistency of nonlinear responses in practical applications, since the measured nonlinear activation functions are directly used to design and train real AONNs.

The nonlinear activation response is also tunable via temperature, as shown in Fig. 6(b), where ReLU curves are measured at $T = 55^\circ\text{C}$, 65°C , and 75°C . As temperature increases, the turning point shifts to higher power, consistent with the scaling of saturation power with optical density. At lower temperatures, activation thresholds can be as low as $10 \mu\text{W}$.

In addition to the tunability, the system's temporal response is critical for evaluating its computational performance. Figure 6(c) shows the transient dynamics of the output signal for laser 1 at various input powers, corresponding to selected points in Fig. 6(b). The characteristic response time is approximately $1 \mu\text{s}$, with faster transitions observed at higher powers due to increased optical pumping rates. We find that temperature also plays a key role in modulating response speed. Figure 6(d) compares the transient response times at $T = 55^\circ\text{C}$, 65°C , 70°C , and 75°C , with $P_{1,\text{out}} = 5 \mu\text{W}$. While the response time remains nearly constant between 55°C and 65°C , a marked change occurs between 65°C and 75°C . This temperature dependence, which contrasts with the stronger temperature sensitivity of the ReLU turning point itself, provides additional control for optimizing device performance. Depending on the specific requirements of various applications, optimal settings for temperature and power can be selectively determined.

The system's intrinsic response time imposes a fundamental limit on the operating frequency of AONNs based on atomic media. While classical electronic processors operate at a gigahertz clock speed, a general computing process normally takes many clock cycles; our system operates on the order of megahertz, several orders of magnitude slower in clock speed. However, the overall optical response time of an AONN scales with only the square root of the network depth but is independent of its transverse size. Thus, for sufficiently large-scale networks, the parallelism inherent to AONNs can enable overall processing speeds that outperform electronic counterparts. Moreover, we find that the falling response time of a rectangular pulse is as short as $1\sim 2 \text{ ns}$ (limited by the detector resolution), as shown in Fig. 6(c), which is determined by the bandwidth of Doppler broadening. This finding suggests that we can operate the optical

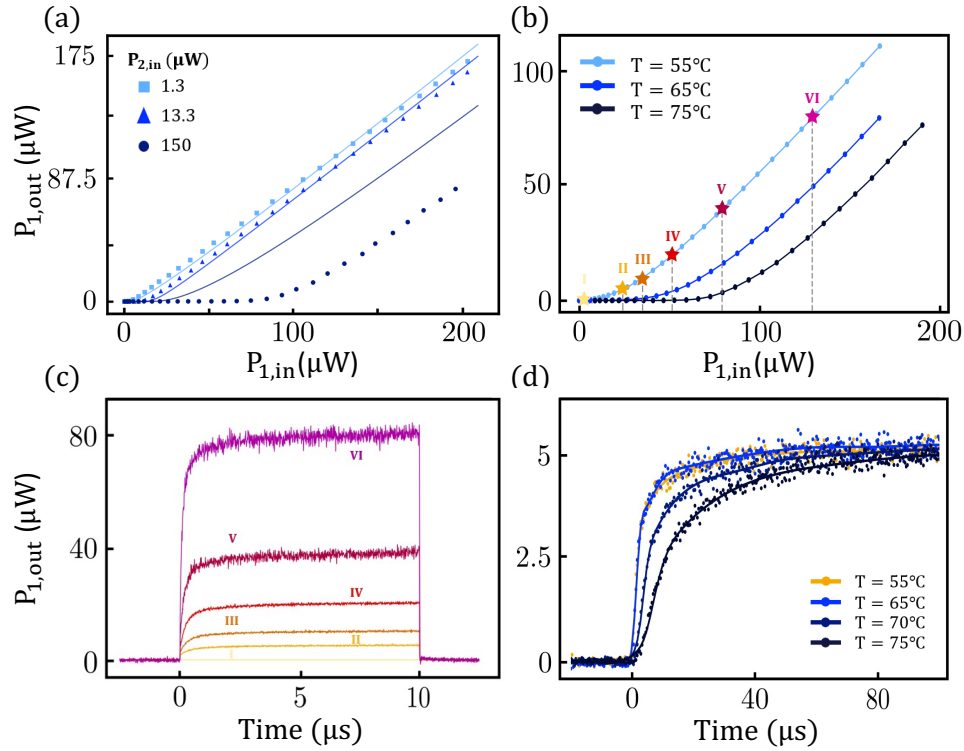


Fig. 6. ReLU nonlinear activation functions for channel 1 with a Rb vapor cell. All measurements were performed with detunings set to $\Delta_1 = \Omega_{2,\text{in}}/2$ and $\Delta_2 = 0$. (a) Measured ReLU activation functions (dots) for varying control beam powers at fixed temperature $T = 75^\circ\text{C}$. Solid lines represent numerical simulations fitted with $\Gamma_{12} = 2\pi \times (0.09, 0.132, 0.6)$ MHz. (b) Temperature-dependent ReLU curves recorded at a fixed control power of $150 \mu\text{W}$. Asterisks I–VI mark input powers used in transient measurements. (c) Transient response of the system under $10 \mu\text{s}$ rectangular probe pulses corresponding to the power levels marked in (b). Control power is fixed at $150 \mu\text{W}$. (d) Measured rise times at different cell temperatures, recorded at the ReLU turning point using $400 \mu\text{s}$ pulses to ensure saturation.

neurons in the transient response mode (such as about 100 MHz repetition rate) for high-speed computing instead of steady-state nonlinearity. As such, the observed saturation and rise times do not impose a practical limitation on the scalability or applicability of the proposed scheme. The vapor cell used in this study contains pure ^{87}Rb . Introducing specific buffer gases may help reduce the response time [43], although this effect warrants further investigation. We also note that temperature variations can substantially influence both the ReLU nonlinear functions [Fig. 6(b)] and their transient responses [Fig. 6(d)]. Thus, practical implementation of a real AONN necessitates active cell-temperature stabilization to suppress noise induced by thermal fluctuations.

3.2. Sigmoid-like nonlinear activation functions

Figure 7 presents the nonlinear activation behavior arising from EIT-based cross-nonlinearity. In panel (a), we show EIT transmission spectra under various control field detunings $\Delta_{2,a,b,c}$. The corresponding nonlinear activation functions are shown in panel (b), where the probe transmission is plotted versus control power for three probe detunings $\Delta_{1,a,b,c}$. At small detuning $\Delta_{1,a}$, saturation is reached at approximately 10 mW. For larger detunings $\Delta_{1,b}$ and $\Delta_{1,c}$, higher

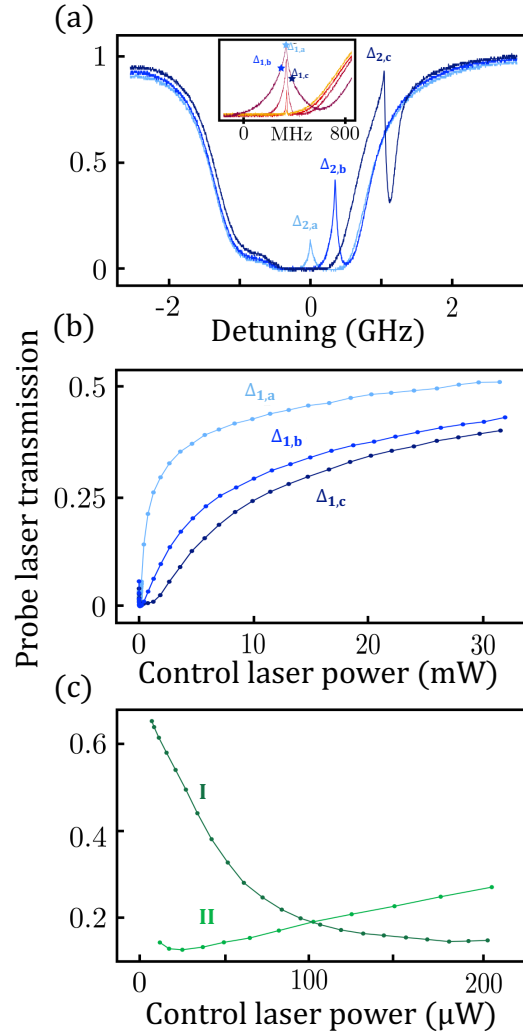


Fig. 7. Sigmoid-like nonlinear activation functions via cross-nonlinearity in a rubidium vapor cell. (a) EIT transmission spectra for probe laser 1 as a function of detuning Δ_1 , with fixed input power $P_{1,\text{in}} = 1.3 \mu\text{W}$. Control field detunings are set to $\Delta_{2,a} = 0$ MHz, $\Delta_{2,b} = 2\pi \times 352$ MHz, and $\Delta_{2,c} = 2\pi \times 1024$ MHz. The inset shows EIT responses at control powers $P_{2,\text{in}} = 0.1, 0.5, 5, \text{ and } 20$ mW (yellow to crimson). Asterisks mark the detunings used in panel (b). (b) Sigmoid activation functions measured at $\Delta_{1,a} = 0$ MHz, $\Delta_{1,b} = 2\pi \times (-20)$ MHz, and $\Delta_{1,c} = 2\pi \times 20$ MHz, relative to a fixed control detuning $\Delta_{2,b} = 2\pi \times 352$ MHz. (c) Nonlinear transmission behavior in the low control power regime. Curve I shows self-nonlinearity (ReLU-type) of laser 1 at $T = 75^\circ\text{C}$. Curve II illustrates cross-nonlinearity dominated by EIA at low optical depth, followed by a transition to EIT at higher control powers. This transition enables low-power switching between ReLU- and sigmoid-like responses.

coupling powers in the 20–30 mW range are required to achieve similar levels of transparency. While the resulting sigmoid-like curves resemble those observed in lifetime-limited media [Fig. 2(d)], Doppler broadening alters key characteristics such as steepness, threshold location, and peak transmission. These differences highlight the distinct activation dynamics of a thermal atomic ensemble.

Beyond conventional sigmoid and ReLU functions, additional nonlinear profiles can be engineered, as illustrated in Fig. 7(c). At elevated temperatures, where EIA dominates for low input power $P_{2,\text{in}}$, the system exhibits a reversed ReLU response (e.g., curve I). As input power increases, the activation behavior transitions to that shown in Fig. 7(b). At lower temperatures, we observe a clear evolution from EIA to EIT with increasing control power. This transition offers an alternative route to implement ReLU-type activation within a low-power regime ($\leq 200 \mu\text{W}$) and enables convenient switching between sigmoid and ReLU nonlinearities. While our AONN implementation focuses on conventional activation functions, these EIA-to-EIT transitions may enable additional reconfigurable architectures and merit further investigation for advanced photonic neural network designs.

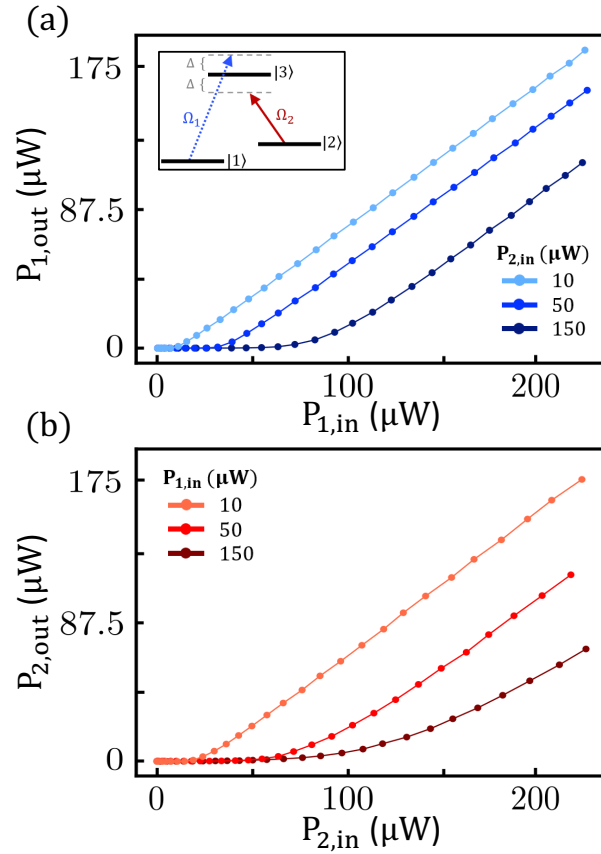


Fig. 8. Symmetric two-channel nonlinear activation functions in a rubidium vapor cell. Measurements are performed with detunings $\Delta_1 = -\Delta_2 = 2\pi \times 2 \text{ MHz}$. (a) ReLU-type activation behavior in channel 1 for varying input powers in channel 2 ($P_{2,\text{in}}$). (b) ReLU-type activation behavior in channel 2 for varying input powers in channel 1 ($P_{1,\text{in}}$).

3.3. Two-channel nonlinear activations

Using two optical fields with comparable powers not only improves energy efficiency but also enables the realization of symmetric dual-channel ($2\text{-input} \times 2\text{-output}$) nonlinear activation functions. Experimental validation of this concept is shown in Fig. 8. To achieve symmetric performance between the two channels, we set the detunings to $\Delta_1 = -\Delta_2 = 2\pi \times 2\text{ MHz}$. Under these conditions, both channels exhibit ReLU-like activation behavior in response to variations in the other channel's input, without the need to change laser frequencies. This symmetric, tunable nonlinearity offers new opportunities for multi-channel AONN architectures, as discussed in the Outlook section. Importantly, because the ReLU response here is mediated by population transfer rather than EIT, the behavior is robust against Doppler broadening and closely resembles that of cold atom systems shown in Fig. 5.

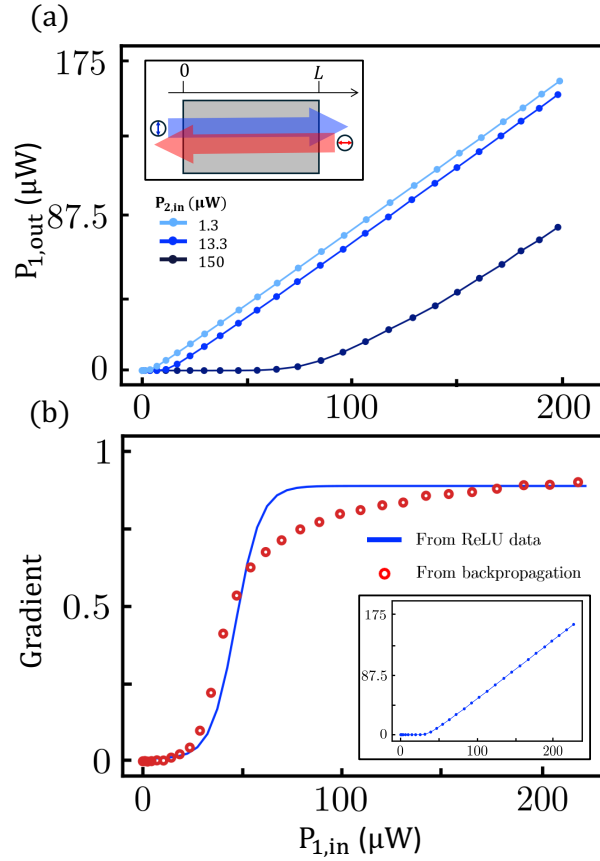


Fig. 9. Nonlinear activation and gradient measurement with counter-propagating beams. (a) ReLU activation functions measured under a counter-propagating geometry, using the same parameters as Fig. 6(a). Inset: schematic of the counter-propagating setup with orthogonal linear polarizations. (b) Gradient extracted from a backward-propagating weak probe input ($P_{\text{probe}} = 5\text{ }\mu\text{W}$, red circles) agrees well with the derivative of the measured ReLU function (inset), demonstrating direct optical gradient readout for backpropagation.

3.4. Counter-propagation and optical gradient measurement

Although our theoretical framework assumes two co-propagating and spatially overlapped laser beams, we experimentally confirm that ReLU-type nonlinear activation functions are also

preserved under a counter-propagating configuration. As shown in Fig. 9(a), the measured activation curve closely matches the results from the co-propagating geometry presented in Fig. 6(a). The ability to implement both co-propagating and counter-propagating geometries provides flexibility in optical system design and is particularly advantageous for realizing optical backpropagation training in AONNs [28,44,45]. To demonstrate this capability, we introduce a weak ($5 \mu\text{W}$) backward-propagating probe beam in the co-propagating setup. Its transmitted signal directly measures the local gradient of the activation function. As shown in Fig. 9(b), the experimentally measured gradient (red circles) agrees well with the derivative extracted from the ReLU curve, confirming the feasibility of optical gradient computation. Minor discrepancies at higher powers may arise from imperfect beam alignment or polarization mismatch.

3.5. Beam size and power reduction

In the preceding experiments, the laser beam waist was set to $75 \mu\text{m}$ to approximate a plane-wave configuration, thereby aligning with our theoretical model. However, further power reduction is possible by decreasing the beam size, even if the plane-wave assumption no longer strictly holds. Figure 10 compares ReLU-type nonlinear activation functions measured using beam waists of $75 \mu\text{m}$ and $45 \mu\text{m}$. In both cases, input powers of laser beam 2 were adjusted to maintain a constant Rabi frequency $\Omega_{2,in}$. As shown in the figure, the activation threshold shifts significantly: the turning point is reduced from approximately $80 \mu\text{W}$ to $40 \mu\text{W}$ with the smaller beam waist. This demonstrates that smaller beams can substantially lower the power required per optical neuron. Although a reduced beam size increases diffraction, this can be compensated by shortening the cell length and increasing cell temperature to preserve sufficient optical depth. This trade-off offers a practical path toward implementing ultra-low-power AONNs with compact physical footprints.

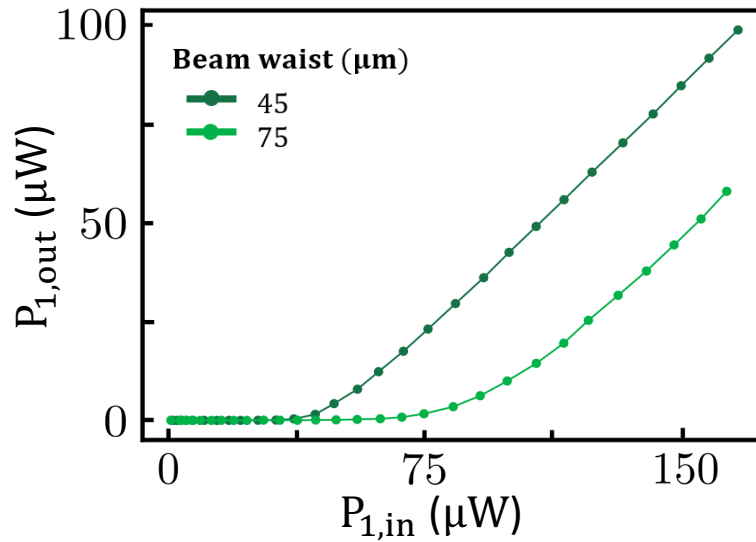


Fig. 10. ReLU nonlinear activation functions for different beam waists. Comparison of activation curves for beam waist radii of $75 \mu\text{m}$ and $45 \mu\text{m}$. Detunings are set to $\Delta_1 = -\Delta_2 = 2\pi \times 2 \text{ MHz}$. In both cases, input powers of laser beam 2 were adjusted to maintain a constant Rabi frequency $\Omega_{2,in}$. A smaller beam size yields a lower activation threshold due to enhanced field intensity.

4. Discussion and outlook

In this work, we demonstrate the engineering of low-power optical nonlinear activation functions using quantum interference in three-level atomic systems, including both lifetime- and Doppler-broadened media. These activation functions serve as core components for implementing AONNs. Compared with earlier approaches based on two-level saturable absorbers [28], our scheme offers significantly greater tunability and reconfigurability. Most notably, we propose and experimentally demonstrate two-channel nonlinear optical activation functions for the first time.

These dual-channel functions enable more expressive transformations while preserving energy efficiency and bandwidth—surpassing the capabilities of conventional electronic and hybrid photonic-electronic architectures. Our results confirm that atomic vapor media offer a compact, tunable, and reconfigurable platform for implementing scalable nonlinear activation. Furthermore, their compatibility with integrated photonic systems positions them as a promising candidate for edge computing and real-time optical processing.

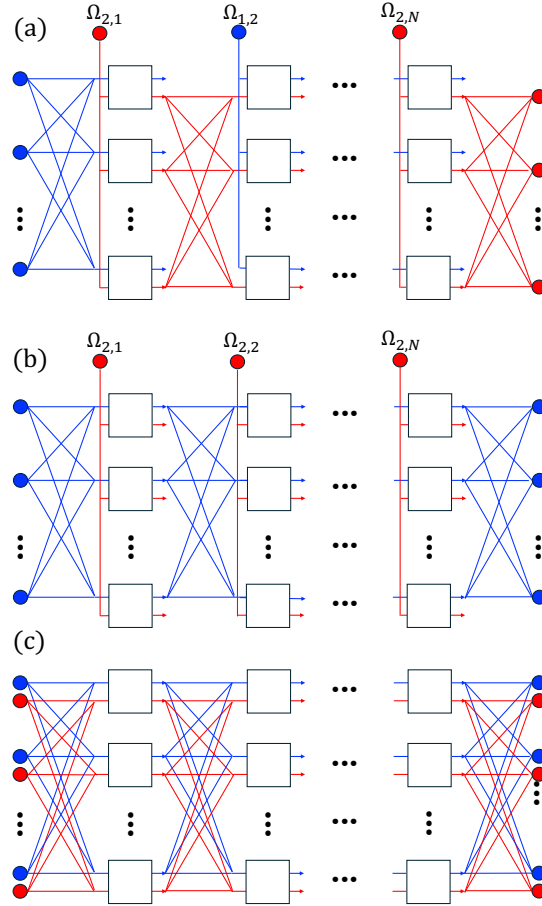


Fig. 11. Architectures for multilayer deep AONNs. (a) AONN based on single-channel control-probe sigmoid nonlinear activation functions. (b) AONN using single-channel ReLU-type self-nonlinear activation functions. (c) MIMO AONN architecture enabled by two-channel (2-input \times 2-output) nonlinear activation functions.

Outlook. Multichannel activation functions, especially those with multiple-input multiple-output (MIMO) capabilities, provide significant advantages by capturing correlations between

different input and output signals. This expressiveness is valuable in a wide range of domains. In image super-resolution, MIMO activations enable convolutional neural networks to process color channels jointly, yielding more accurate reconstructions [46]. In denoising autoencoders, they improve robustness by modeling inter-channel correlations [47]. In natural language processing and transformer models, MIMO activations support simultaneous processing of multiple embeddings, enhancing contextual understanding [48]. In robotics, autonomous vehicles, and IoT systems, where multimodal sensor fusion is critical, MIMO activations help capture relationships between signals from diverse sources [49].

Figure 11 illustrates candidate AONN architectures. Figures 11(a) and (b) show traditional single-channel implementations based on sigmoid and ReLU nonlinearities, respectively. Figure 11(c) introduces a MIMO AONN enabled by two-channel nonlinear optical activation. This multichannel architecture reduces the required network depth and expands the repertoire of nonlinear operations while maintaining the energy efficiency of quantum-interference-based mechanisms.

Experimental results from Figs. 2–9 indicate that the activation functions typically operate at optical powers on the order of 100 μW . Assuming 10% loss per layer, this power level can support up to seven sequential neurons, resulting in approximately 17 μW per neuron. Extrapolating this to a full network, one could support one million optical neurons with only 17 W of total optical power.

Despite this promise, challenges remain—particularly in scaling to larger networks, managing thermal effects, and ensuring stability and noise resilience. Continued progress in photonic integration, quantum materials, and system-level co-design will be critical to overcoming these limitations. As AI workloads increasingly demand high-speed, low-power computation, AONNs based on atomic vapor or atom-like materials such as quantum dots and solid-state defects represent a compelling path forward. This study lays the foundation for future exploration of optimized activation designs, alternative materials, and integrated system implementations that will unlock the full potential of optical neural computing.

Funding. National Science Foundation (2211989).

Disclosures. The authors declare the following competing interests: S.D., H.A., R.C., and X.X. are listed as inventors on US provisional patent application 63/781,585 filed by Purdue University. All remaining authors declare no competing interests.

Data availability. Experimental data and code for numerical simulations are available at: [50]

References

1. B. J. Shastri, A. N. Tait, T. Ferreira de Lima, *et al.*, “Photonics for artificial intelligence and neuromorphic computing,” *Nat. Photonics* **15**(2), 102–114 (2021).
2. Y. Zuo, B. Li, Y. Zhao, *et al.*, “All-optical neural network with nonlinear activation functions,” *Optica* **6**(9), 1132–1137 (2019).
3. M. Matuszewski, A. Prystupik, and A. Opala, “Role of all-optical neural networks,” *Phys. Rev. Appl.* **21**(1), 014028 (2024).
4. A. Montes McNeil, Y. Li, A. Zhang, *et al.*, “Fundamentals and recent developments of free-space optical neural networks,” *J. Appl. Phys.* **136**(3), 030701 (2024).
5. T. Fu, J. Zhang, R. Sun, *et al.*, “Optical neural networks: progress and challenges,” *Light: Sci. Appl.* **13**(1), 263 (2024).
6. X. Sui, Q. Wu, J. Liu, *et al.*, “A review of optical neural networks,” *IEEE Access* **8**, 70773–70783 (2020).
7. I. Shariv and A. Friesem, “All-optical neural network with inhibitory neurons,” *Opt. Lett.* **14**(10), 485–487 (1989).
8. C. Isil, T. Gan, F. O. Ardic, *et al.*, “All-optical image denoising using a diffractive visual processor,” *Light: Sci. Appl.* **13**(1), 43 (2024).
9. S. Chen, Y. Li, Y. Wang, *et al.*, “Optical generative models,” *Nature* **644**(8078), 903–911 (2025).
10. B. Li, E. Zhou, B. Huang, *et al.*, “Large scale recurrent neural network on gpu,” in *2014 International Joint Conference on Neural Networks (IJCNN)*, (IEEE, 2014), pp. 4062–4069.
11. X. She, Y. Long, and S. Mukhopadhyay, “Fast and low-precision learning in gpu-accelerated spiking neural network,” in *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, (IEEE, 2019), pp. 450–455.
12. S. W. Keckler, W. J. Dally, B. Khailany, *et al.*, “Gpus and the future of parallel computing,” *IEEE Micro* **31**(5), 7–17 (2011).

13. K. Wang, Z. Liu, Y. Lin, *et al.*, “HAQ: hardware-aware automated quantization with mixed precision,” in *CVPR*, (Computer Vision Foundation / IEEE, 2019), pp. 8612–8620.
14. C. Zhang, D. Wu, J. Sun, *et al.*, “Energy-efficient cnn implementation on a deeply pipelined fpga cluster,” in *Proceedings of the 2016 International Symposium on Low Power Electronics and Design*, (ACM, 2016), pp. 326–331.
15. E. Park, D. Kim, and S. Yoo, “Energy-efficient neural network accelerator based on outlier-aware low-precision computation,” in *ISCA*, (IEEE Computer Society, 2018), pp. 688–698.
16. T. Geng, T. Wang, C. Wu, *et al.*, “O3BNN: an out-of-order architecture for high-performance binarized neural network inference with fine-grained pruning,” in *ICS*, (ACM, 2019), pp. 461–472.
17. S. E. Thompson and S. Parthasarathy, “Moore’s law: the future of si microelectronics,” *Mater. Today* **9**(6), 20–25 (2006).
18. L. B. Kish, “End of moore’s law: thermal (noise) death of integration in micro and nano electronics,” *Phys. Lett. A* **305**(3–4), 144–149 (2002).
19. T. Wang, S.-Y. Ma, L. G. Wright, *et al.*, “An optical neural network using less than 1 photon per multiplication,” *Nat. Commun.* **13**(1), 123 (2022).
20. R. Hamerly, L. Bernstein, A. Sludds, *et al.*, “Large-scale optical neural networks based on photoelectric multiplication,” *Phys. Rev. X* **9**(2), 021032 (2019).
21. J. George, R. Amin, A. Mehrabian, *et al.*, “Electrooptic nonlinear activation functions for vector matrix multiplications in optical neural networks,” in *Signal Processing in Photonic Communications, SPPCom 2018*, (Optica Publishing Group (formerly OSA), 2018), Optics InfoBase Conference Papers.
22. J. Spall, X. Guo, and A. I. Lvovsky, “Hybrid training of optical neural networks,” *Optica* **9**(7), 803–811 (2022).
23. J. Feldmann, N. Youngblood, C. D. Wright, *et al.*, “All-optical spiking neurosynaptic networks with self-learning capabilities,” *Nature* **569**(7755), 208–214 (2019).
24. S. E. Harris, “Electromagnetically induced transparency,” *Phys. Today* **50**(7), 36–42 (1997).
25. M. Fleischhauer, A. Imamoglu, and J. P. Marangos, “Electromagnetically induced transparency: Optics in coherent media,” *Rev. Mod. Phys.* **77**(2), 633–673 (2005).
26. M. Yang, E. Robertson, L. Esguerra, *et al.*, “Optical convolutional neural network with atomic nonlinearity,” *Opt. Express* **31**(10), 16451–16459 (2023).
27. A. Ryou, J. Whitehead, M. Zhelyeznyakov, *et al.*, “Free-space optical neural network based on thermal atomic nonlinearity,” *Photonics Res.* **9**(4), B128–B134 (2021).
28. J. Spall, X. Guo, and A. I. Lvovsky, “Training neural networks with end-to-end optical backpropagation,” *Adv. Photonics* **7**(01), 016004 (2025).
29. Y. Zuo, Y. Zhao, Y.-C. Chen, *et al.*, “Scalability of all-optical neural networks based on spatial light modulators,” *Phys. Rev. Appl.* **15**(5), 054034 (2021).
30. R. W. Boyd, *Nonlinear Optics* (Academic Press, 2003), 3rd ed.
31. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley-Interscience, 1984).
32. P. Michler, ed., *Single Quantum Dots: Fundamentals, Applications and New Concepts* (Springer, 2003).
33. I. Aharonovich, D. Englund, and M. Toth, “Solid-state single-photon emitters,” *Nat. Photonics* **10**(10), 631–641 (2016).
34. E. Arimondo, “Coherent population trapping in laser spectroscopy,” in *Progress in Optics*, vol. 35 E. Arimondo and W. D. Phillips, eds. (Elsevier, 1996), pp. 257–354.
35. H. Schmidt and A. Imamoglu, “Giant kerr nonlinearities obtained by electromagnetically induced transparency,” *Opt. Lett.* **21**(23), 1936–1938 (1996).
36. S. E. Harris and Y. Yamamoto, “Photon switching by quantum interference,” *Phys. Rev. Lett.* **81**(17), 3611–3614 (1998).
37. M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, 1997). See Sec. 7.2 “Two-Level Atom and Saturation” and Fig. 7.1.
38. R. W. Boyd, *Nonlinear Optics* (Academic Press, New York, 2020), 4th ed. See Sec. 4.2 “Saturable Absorption” for the practical and nonlinear optics perspective.
39. C. Cohen-Tannoudji and S. Reynaud, “Dressed-atom description of resonance fluorescence and saturation in two-level systems,” *J. Phys. B: At. Mol. Phys.* **10**(3), 345–363 (1977).
40. H. Alaeian, R. Ritter, M. Basic, *et al.*, “Cavity QED based on room temperature atoms interacting with a photonic crystal cavity: a feasibility study,” *Appl. Phys. B* **126**(2), 25 (2020).
41. A. Skljarow, N. Gruhler, W. Pernice, *et al.*, “Integrating two-photon nonlinear spectroscopy of rubidium atoms with silicon photonics,” *Opt. Express* **28**(13), 19593–19607 (2020).
42. A. Skljarow, H. Kübler, C. S. Adams, *et al.*, “Purcell-enhanced dipolar interactions in nanostructures,” *Phys. Rev. Res.* **4**(2), 023073 (2022).
43. S. Krishnamurthy, Y. Wang, Y. Tu, *et al.*, “High-speed modulation in ladder transitions in rb atoms using high-pressure buffer gas,” *Opt. Express* **23**(9), 11470–11482 (2015).
44. X. Guo, T. D. Barrett, Z. M. Wang, *et al.*, “Backpropagation through nonlinear units for the all-optical training of neural networks,” *Photonics Res.* **9**(3), B71–B80 (2021).
45. T. W. Hughes, M. Minkov, Y. Shi, *et al.*, “Training of photonic neural networks through in situ backpropagation and gradient measurement,” *Optica* **5**(7), 864–871 (2018).

46. X. Wang, K. Yu, S. Wu, *et al.*, “Esrgan: Enhanced super-resolution generative adversarial networks,” in *Computer Vision – ECCV 2018 Workshops*, (Springer International Publishing, Cham, 2019), pp. 63–79.
47. K. Zhang, W. Zuo, Y. Chen, *et al.*, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE Trans. on Image Process.* **26**(7), 3142–3155 (2017).
48. P. Ramachandran, B. Zoph, and Q. V. Le, “Searching for activation functions,” (2017).
49. Z. Huang, C. Lv, Y. Xing, *et al.*, “Multi-modal sensor fusion-based deep neural network for end-to-end autonomous driving with scene understanding,” *IEEE Sens. J.* **21**(10), 11781–11790 (2021).
50. X. Xu, R. Canora, and Z. Niu, “Engineering nonlinear activation functions for all-optical neural networks via quantum interference,” GitHub (2025), <https://github.com/AndersonXu99/AONN-Nonlinearity>.