1. Introduction

Bistability is a property of a system that exhibits two stable steady states and the system rests in one of those states depending on the history of this system. For example, a permanent magnet can be in one of two states with oppositely directed magnetization, depending on the sign of the magnetic field that had been applied and then removed. Two states thus can represent two values of a binary digit (bit) in magnetic storage. Similarly, resistance of many electronic devices, such as thyristors and memristors, can be low or high depending on whether a high current had or had not been applied. Bistability is the underlying principle for switching logic and storage applications. While electronic and magnetic bistable devices of various types are in wide use, when it comes to applications in the optical domain, the situation is less sanguine as no bistable elements with acceptable characteristics (fast speed, low power, wide optical bandwidth) have emerged, and not for the lack of trying. Most of the proposed and demonstrated bistable optical devices are based on a combination of a nonlinear optical material and a resonant cavity, require high optical powers for switching, and are often incompatible with integrated photonics. A different type of bistable device is an electro-absorptive self-electro-optic device (SEED), constructed of a semiconductor multiple quantum wells biased by an external voltage, which creates an bistable shifts of the absorbing wavelengths of the transmitted light and does not require optical cavity. However, the device requires electrical bias and hence cannot be considered all-optical, and its switching time is limited by the incorporated RC constant of the electrical circuit. All the devices operate within a relatively narrow optical bandwidth determined by either the resonance wavelength in nonlinear cavity-based systems or the excitonic electro-absorption bandwidth in in SEEDs. The lack of reliable bistable all-optical devices impedes technological progress in many fields of conventional and quantum optics, hindering particularly the development of all-optical neural networks that exhibit bistable dynamics.

2. New Flavor of Bistability with ENZ Waveguides

Here, we propose a fundamentally new concept of all-optical bistable devices that do not require a resonator, enabling thereby broadband operation, which can be integrated into a photonic waveguide and operate at sub-ps switching speeds. The proposed device is based on the unique properties of epsilon-near-zero (ENZ) materials (typically transparent conductive oxides (TCO) like indium tin oxide (ITO)) like indium tin oxide (ITO) (fast speed, low power, wide optical bandwidth) have emerged, and not for the lack of trying. Most of the proposed and demonstrated bistable optical devices are based on a combination of a nonlinear optical material and a resonant cavity, require high optical powers for switching, and are often incompatible with integrated photonics. A different type of bistable device is an electro-absorptive self-electro-optic device (SEED), constructed of a semiconductor multiple quantum wells biased by an external voltage, which creates an bistable shifts of the absorbing wavelengths of the transmitted light and does not require optical cavity. However, the device requires electrical bias and hence cannot be considered all-optical, and its switching time is limited by the incorporated RC constant of the electrical circuit. All the devices operate within a relatively narrow optical bandwidth determined by either the resonance wavelength in nonlinear cavity-based systems or the excitonic electro-absorption bandwidth in in SEEDs. The lack of reliable bistable all-optical devices impedes technological progress in many fields of conventional and quantum optics, hindering particularly the development of all-optical neural networks that exhibit bistable dynamics.

Bistable All-Optical Devices Based on Nonlinear Epsilon-Near-Zero (ENZ) Materials

Jacek Gosciniak,* Zibo Hu, Martin Thomaschewski, Volker J. Sorger, and Jacob B. Khurgin

Nonlinear and bistable optical systems are a key enabling technology for the next generation optical networks and photonic neural systems with many potential applications in optical logic and information processing. Here, a novel bistable resonator-free all-optical waveguide device based on indium tin oxide as nonlinear epsilon-near-zero material providing a cost-efficient and high-performance binarity photonic platform is proposed. The salient features of the proposed device are compatibility with silicon photonics, enabling sub-picosecond operation speeds with moderate switching power. The device can act as an optical analogue of memristor or thyristor and can become an enabling element of photonic neural networks not requiring OEO conversions.
negative at the so-called ENZ frequency $\omega_0 = \omega_p/\epsilon_\infty^{1/2}$ (identified as screened plasma frequency) with an imaginary part of the permittivity introducing absorption in the material, as shown in Figure 1a,b. Due to the large permittivity tunability, transparent conductive oxides have been used extensively for electro-optic modulation\cite{23,24} and other optoelectronic applications (e.g., in display technology or vertical-cavity surface-emitting lasers).\cite{25–27} But its versatility extends beyond existing applications as will be shown in this work.

Note that bistability had been explored theoretically in the so-called ENZ metamaterials—periodic metal dielectric structures,\cite{7,8,12,13} but that bistability was not intrinsic as it was necessary to introduce additional nonlinearity into the dielectric, and the performance dependent on the periodicity which led to a very narrow bandwidth. Furthermore, the bistability was predicted only in transverse geometry with oblique angles—a geometry hardly conducive to incorporating into photonic integrated circuits. In contrast, here we propose a bistable device based on natural ENZ material with intrinsic nonlinearity of TCO capable of operation in waveguide geometry with wide bandwidth, i.e., ideally suitable for incorporating into integrated photonic neural networks.

Consider the interface of ENZ material in the vicinity of $\omega_0$ with $Re(\epsilon(\omega_0)) < 0$ and a conventional dielectric with $\epsilon_d > 0$.

According to boundary conditions for light polarized normal to interface (TM), the displacement $D = eE$ is continuous across the interface, and therefore the electric field (and hence the energy density) is much higher in the ENZ material than in the conventional dielectric. If a thin layer of ENZ material is inserted in the slot of a waveguide (which can be all-dielectric or metallic) as shown in Figure 1b the closer is the wavelength to $\omega_0$, the higher is the field the slot. This is a well-studied slot waveguide effect\cite{28,29} that allows subwavelength field confinement in the low permittivity medium, taken to its ultimate $\epsilon \approx 0$ limit. Note that using metal waveguide, used here as an example, allows higher concentration of optical field and thus requires switching power, but one can equally well also use all-dielectric waveguides trading off switching power for lower insertion loss.

Unlike the dielectric cladding, ENZ material is highly absorptive and as the optical field concentrated inside the slot increases, the absorption increases. At the point of large electric fields, the nonlinear properties of the TCO-based ENZ materials come into play.\cite{20,30–32} When light is absorbed by the TCO material, it causes an increase of the electron temperature in the conduction band, that in turn leads to an increase of the effective mass and corresponding rise of the dielectric constant, according to Equation (1). If the carrier density is such that in the absence of light the real permittivity is $Re(\epsilon) < 0$, an increase of $Re(\epsilon)$ would move it closer to zero, which in turn will cause even higher concentration of energy inside the ENZ material layer. Thus, a positive feedback mechanism, essential for bistability, is established. As more energy is concentrated in the ENZ layer, the electron temperature in it rises causing an increase of the dielectric constant and further concentration of energy in the ENZ layer. Obviously, no stable solution can exist with positive feedback and switching occurs.
3. Mechanism of Bistability and Switching

The proposed concept of bistability is investigated using 2D finite element method (FEM) simulations at the telecom wavelength of 1550 nm using commercial software COMSOL.

The investigated arrangement consists of a typical metal–insulator–metal (MIM) structure where Al₂O₃ serves as an insulator and the ENZ material is placed in the center of the Al₂O₃ (Figure 1). Here, the ITO was used as the ENZ material as it is the most widely used TCO material in industry with very well-established fabrication process that shows the presence of the ENZ region at tolerable carrier concentrations. [26,28]

For the MIM plasmonic structure, the electric field of the propagating mode exhibits maxima at both metal–insulator interfaces and decays exponentially into the insulator (Figure 2b,c). The relative amount of energy in the metal and the insulator depends on the optical properties of the material and the waveguide geometry.

The thickness of Au was chosen to be \( h = 50 \) nm, while the thickness of Al₂O₃ and ITO is \( h_1 = 30 \) nm and \( h_2 = 10 \) nm, respectively. The width of the multilayer stack is \( w = 400 \) nm. For the calculations, the complex refractive index of Au is assumed to be \( n_{\text{Au}} = 0.596 + 10.923i \) while for the refractive index of Al₂O₃ is \( n_{\text{Al}_2\text{O}_3} = 1.621 + 0.00008i \). The calculations are performed for a thermalization time \( \tau = 100 \) fs, the Fermi energy \( E_F = 4.5 \) eV, and the carrier concentration \( N_c = 8 \times 10^{26} \) m⁻³.

Let us now quantify the switching and bistability in a specific waveguide shown in Figure 2. We model the waveguide absorption coefficient \( \alpha \) of light at a given wavelength \( \lambda = 1500 \) nm, \( \omega = 2\pi \times 200 \) THz as a function of \( \omega_p^2 \). As shown in Figure 3a, the maximum of absorption corresponds to \( \omega_0 = \omega_p/e_m = 2\pi \times 200 \) THz, i.e., to the minimum of \( \text{Re}(\varepsilon) \). It is obviously the
expected consequence of an electric field concentrated in the slot where absorption takes place. One can then estimate the power loss per unit volume of ITO. For an optical power \( P \) in the waveguide, the absorbed power density in the ITO filled slot is

\[ P_{\text{abs}} = \alpha P S_{\text{ITO}}, \]

where \( S_{\text{ITO}} \) is the cross section of the ITO layer. The density of energy transferred to the hot carriers in the ITO is

\[ \dot{E}_{\text{hot}} = P_{\text{abs}} \tau = \alpha P \tau S_{\text{ITO}}, \]

where \( \tau \) is the “cooling time,” i.e., the time it takes to transfer energy from the hot carriers to the lattice via electron–phonon scattering. This time is typically a few hundreds of femtoseconds.\(^{31}\) As shown in refs. \(^{30,31}\) when the hot carriers gain energy, they move higher into the conduction band of ITO where their effective mass is increased which introduces a change of \( \alpha_p \) impacting the effective mass as according to\(^{30}\)

\[ \frac{\delta m^*}{m^*} = -\frac{\delta \alpha_p}{\alpha_p} \approx \frac{U_{\text{hot}}}{\hbar v_F^2} \]

where \( k_F \) and \( v_F \) are, respectively, Fermi wave vector and velocity. The relationship to the Fermi energy for a parabolic band follows \( k_F v_F = 2E_F \) and for an extremely nonparabolic band (such as graphene) follows \( k_F v_F = E_F \), where \( E_F \) is Fermi energy.

The relationship to the Fermi energy for a parabolic band follows\(^{31}\)

\[ \omega_p = \omega_p^0 \left( 1 + \frac{\alpha \tau}{N E_F S} P \right) \]

where \( \omega_p^0 \) is the plasma frequency in the absence of optical power and heating. In our example the carrier density is \( 8 \times 10^{26} \text{ m}^{-3} \) and \( E_F = 4.5 \text{ eV} \). We can now plot a family of straight lines [Equation (3)] for different values of power \( P \) in Figure 3a. Depending on the value of \( P \) the curve \( \alpha(\omega_p^2) \) shows either one or three intersections with straight lines. In the case of three intersection points (on the right slope of the curve) the middle solution is unstable due to positive feedback. Therefore, gradually increasing power (right to left) leads to gradual increasing both effective mass and loss until the critical (switching) value \( P_{\text{up}} \). When the power line gets tangential to the curve \( \alpha(\omega_p^2) \) the positive feedback generates a stable solution to switch at a much higher value of the effective mass and loss. During switching, \( \text{Re}(\alpha) \) changes the sign from negative to positive and a further increase in power no longer causes an increase of the degree of carrier concentration in the ITO which induces a stable but gradually changing solution on the left slope of the curve. When the power is cycled down the high-loss regime, the solution remains stable until the second switching value \( P_{\text{down}} < P_{\text{up}} \) is reached, and the positive feedback enables switching into the low-loss state with \( \text{Re}(\alpha) < 0 \). Thus, a hysteresis region is established as shown in Figure 3b showing the absorptive loss per unit length as a function of the propagating power.

4. Evaluating the Performance of Bistable Switch

The hysteresis curve of Figure 3b is a local curve, i.e., which relates to the absorption \( \alpha(z) \) and propagating power \( P(z) \) at a given point \( z \). To find a relation between the input and output power at \( z = L \), we need to integrate the coupled system of equations

\[ a(z) = a(P(z)) \]

\[ \frac{dP(z)}{dz} = -\alpha(z) P(z) \]

with boundary condition \( P(0) = P_{\text{in}} \). The results are shown in Figure 4 for the case of \( L = 1 \text{ \mu m} \). The switching does not occur simultaneously over the entire length of the device but propagates shock-wave-like from the front as local \( P(z) \) reaches the switching power \( P_{\text{up}} \). As input power is ramped down, the “switching wave” propagates backward. Thus, the hysteresis curve has a tilted shape (Figure 4). Presence of hysteresis indicates that the proposed nonlinear device can serve as an optical analogue of memristor. Alternatively, if a separate switching pulse is coupled to the device, then it becomes optical analogue of thyristor.
The bistable characteristics of the proposed device providing fast and reliable all-optical binarity can have impactful applications in artificial neural networks with binary weights and nonlinear activations. Binarization in neural networks enjoy several hardware-friendly properties including memory saving, power efficiency and significant acceleration. For instance, computationally heavy matrix multiplication operations can be replaced with lightweight bitwise XNOR operations and Bitcount operations. Furthermore, it has been shown that binary neural networks significantly improve the robustness of the network by keeping the magnitude of the noise small, which can be integrated to an optical neural network with our proposed device providing low power consumption, high computational speed, and large information storage.

5. Bistable ENZ Device in Neural Networks

All-optical neural networks have been developed in recent years, and this field requires all-optical nonlinear activation functions. Many researchers are trying to develop nonlinear activation functions. Bistable devices can be used as nonlinear activation functions of all-optical neural networks. Such bistable devices can be easily implemented into integrated optical deep neural networks since they are all waveguide-based systems. It can also be implemented into free-space optical neural networks with pixel-wise modulation array as SLM. Also, if we need to deal with time series data like NLP tasks, we can match the system clock to the response time of the bistable material, we will take advantage of the temporal dynamic behavior of the bistable material, the activation function output depends on the previous state.

To explore the applications of the proposed bistable ENZ device in the all-optical neural networks, we simulate fully connected deep neural networks using our bistable characteristics as the activation function. In the simulation, we use MNIST, FashionMNIST and CIFAR-10 as test data sets. The neural network performance with bistable nonlinear activation function. We simulate a neural network with two hidden layers with four different activation functions, no activation, ReLu, sigmoid, and bistable nonlinear. The simulation shows our bistable nonlinear has similar performance with classic digital nonlinear activation functions ReLu and sigmoid, 98% for MNIST data set, ≈90% for FashionMNIST data set, and 52% for CIFAR-10 data set comparing to the linear activation function with 92%, 84%, and 38%, respectively.
consists of flat images as input layer, two hidden layers with 400 and 200 nodes, and 10 output nodes for classification (Figure 5). Simulations show that our bistable nonlinear activation function performs similarly to the ReLu and Sigmoid, the classic digital nonlinear activation functions, and they all show a prediction accuracy of 98% for MNIST, 89% for FashionMNIST, and ≈52% for CIFAR-10. Compared to a neural network without a nonlinear activation function, MNIST shows only 92%, FashionMNIST is 84%, and CIFAR-10 is 38% (https://github.com/Sorger-Lab/General/tree/master/Bistable).

6. Conclusions

In this work, we have proposed and analyzed a novel bistable device based on ENZ material in a slot waveguide and have shown that it shows high nonlinearity and significant bistability without optical cavity. Small footprint and high switching speed of the proposed device can lead to applications in various all-optical systems including neural networks where the performance of the devices was estimated and found to be excellent.

Acknowledgements

J.G. thanks the “ENSEMBLE3 - Centre of Excellence for Nanophotonics, advanced materials and novel crystal growth-based technologies” project (GA No. MAB/2020/14) carried out within the International Research Agendas program of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund and the European Union’s Horizon 2020 research and innovation program Team ing for Excellence (Grant Agreement No. 857543) for support of this work. J.K. would like to thank DARPA NLM program. V.S. was supported by European Union’s Horizon 2020 research and innovation program “Building Advanced Materials and Novel Crystal Growth-based Technologies” project Agendas program of the Foundation for Polish Science co-financed by the European Union’s Horizon 2020 research and innovation program Team ing for Excellence (Grant Agreement No. 857543) for support of this work. J.K. would like to thank DARPA NLM program. V.S. was supported by European Union’s Horizon 2020 research and innovation program Team ing for Excellence (Grant Agreement No. 857543) for support of this work. J.K. would like to thank DARPA NLM program.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

bistability, ENZ materials, neuromorphic computing, plasmonics, TCOs

Received: September 22, 2022
Revised: January 19, 2023
Published online: February 9, 2023

[1] R. Sessoli, D. Gatteschi, A. Caneschi, M. A. Novak, Nature 1993, 365, 141.
[2] A. Blicher, Thyristor Physics, Vol. 12, Springer Science & Business Media, Springer New York, NY 2013.

[3] D. B. Strukov, G. S. Snider, D. R. Stewart, R. S. Williams, Nature 2008, 453, 80.
[4] H. M. Gibbs, Optical Bistability: Controlling Light with Light, Academic, Orlando, FL, 1985, pp. xii, 471.
[5] H. M. Gibbs, S. L. McCall, T. N. C. Venkatesan, A. C. Gossard, A. Passner, W. Wiegmann, Appl. Phys. Lett. 1979, 35, 451.
[6] E. Abraham, S. D. Smith, Rep. Prog. Phys. 1982, 45, 815.
[7] A. Ciatti, C. Rizza, E. Palange, Opt. Lett. 2010, 35, 2130.
[8] A. Ciatti, C. Rizza, E. Palange, Phys. Rev. A 2011, 83, 043813.
[9] V. R. Almeida, M. Lipson, Opt. Lett. 2004, 29, 2387.
[10] A. Baas, J. P. Karr, H. Eleuch, E. Giacobino, Phys. Rev. A 2004, 69, 023809.
[11] T. K. Fryett, C. M. Dodson, A. Majumdar, Opt. Express 2015, 23, 16246.
[12] A. Ciatti, C. Rizza, E. Palange, Phys. Rev. A 2010, 81, 043839.
[13] A. Ciatti, C. Rizza, A. Marini, A. Di Falco, M. Scalara, Laser Photonics Rev. 2016, 10, 517.
[14] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, C. A. Burrell, Appl. Phys. Lett. 1984, 45, 13.
[15] G. D. Boyd, A. M. Fox, D. A. B. Miller, L. M. F. Chirovsky, L. A. D’Asaro, J. M. Kuo, R. F. Kopf, A. L. Lentine, Appl. Phys. Lett. 1990, 57, 1843.
[16] A. Sinigham, R. E. Evans, D. S. Sukachev, M. J. Burek, J. Borregaard, M. K. Bhaskar, C. T. Nguyen, J. L. Pacheco, H. A. Atikian, C. Meuwly, R. M. Camacho, F. Jelezko, E. Bielejec, H. Park, M. Lontar, M. D. Lukin, Science 2016, 354, 847.
[17] G. R. Steinbrecher, J. P. Olson, D. Englund, J. Carolan, npj Quantum Inf. 2019, 5, 60.
[18] M. Miscuglio, A. Mehrabian, Z. Hu, S. I. Azzam, J. George, A. V. Kildishev, M. Pelton, V. J. Sorger, Opt. Mater. Express 2018, 8, 3851.
[19] J. Gosciniak, IEEE Photonics J. 2022, 4, 3.
[20] N. Kinsey, J. Khurgin, Opt. Mater. Express 2019, 9, 2793.
[21] M. Z. Alam, I. D. Leon, R. W. Boyd, Science 2016, 352, 795.
[22] L. Caspani, R. P. M. Kaiparuth, M. Clerici, M. Ferrera, T. Roger, J. Kim, N. Kinsey, M. Pietrzyk, A. Di Falco, V. M. Shalaev, A. Boltasseva, D. Faccio, Phys. Rev. Lett. 2016, 117, 233901.
[23] V. J. Sorger, N. D. Lanzillotti-Kimura, R.-M. Ma, X. Zhang, Nanophotonics 2012, 1, 17.
[24] Y. L. Gui, B. M. Nouri, M. Miscuglio, R. Amin, H. Wang, J. B. Khurgin, H. Dalir, V. J. Sorger, Nanophotonics 2022, 11, 4001.
[25] J. B. Khurgin, M. Clerici, V. Bruno, L. Caspani, C. DeVault, J. Kim, A. Shaltout, A. Boltasseva, V. M. Shalaev, M. Ferrera, Optica 2020, 7, 226.
[26] H. Dalir, F. Koyama, Appl. Phys. Express 2014, 7, 022102.
[27] M. Ahmed, A. Bakry, M. S. Alghamdi, H. Dalir, F. Koyama, Opt. Express 2015, 23, 15365.
[28] P. Sanchis, J. Blasco, A. Martínez, J. Martí, J. Lightwave Technol. 2007, 25, 1298.
[29] L. Chen, J. Shakya, M. Lipson, Opt. Lett. 2006, 31, 2133.
[30] R. Secondo, J. Khurgin, N. Kinsey, Opt. Mater. Express 2020, 10, 1545.
[31] J. B. Khurgin, M. Clerici, N. Kinsey, Laser Photonics Rev. 2021, 15, 2000291.
[32] A. Ciatti, E. Spinozzi, Phys. Rev. A 2012, 85, 043806.
[33] M. S. Farhan, E. Zalnezhad, A. R. Bushroa, A. A. D. Sarhan, Int. J. Precision Eng. Manuf. 2013, 14, 1465.
[34] H. Qin, R. Gong, X. Liu, X. Bai, J. Song, N. Sebe, Pattern Recognit. 2020, 105, 107281.
[35] J. Gosciniak, arXiv preprint arXiv:2109.13562 2021.
[36] A. Jha, C. H. Huang, P. R. Prucnal, Opt. Lett. 2020, 45, 4819.
[37] J. George, R. Amin, A. Mehrabian, J. Khurgin, T. El-Ghazawi, P. R. Prucnal, V. J. Sorger, Signal Processing in Photonic Communications, Optica Publishing Group, Zurich, Switzerland 2018.
[38] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhuk, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, M. Soljačić, Nat. Photonics 2017, 11, 441.
[39] T. F. de Lima, H.-T. Peng, A. N. Tait, M. A. Nahmias, H. B. Miller, B. J. Shastri, P. R. Prucnal, J. Lightwave Technol. 2019, 37, 1515.
[40] T. Y. Teo, X. Ma, E. Pastor, H. Wang, J. K. George, J. K. W. Yang, S. Wall, M. Miscuglio, R. E. Simpson, V. J. Sorger, Nanophotonics 2022, 11, 4073.
[41] J. Gosciniak, AIP Adv. 2022, 12, 035321.
[42] X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, A. Ozcan, Science 2018, 361, 1004.
[43] Z. Hu, S. Li, R. L. T. Schwartz, M. Solyanik-Gorgone, B. M. Nouri, M. Miscuglio, P. Gupta, H. Dalir, V. J. Sorger, Proceedings Volume 12204, Emerging Topics in Artificial Intelligence (ETAI) 2022, 2022, 1220409.