

Research Article

Grigorii Slinkov, Steven Becker, Dirk Englund and Birgit Stiller*

All-optical nonlinear activation function based on stimulated Brillouin scattering

<https://doi.org/10.1515/nanoph-2024-0513>

Received September 30, 2024; accepted January 8, 2025;

published online February 14, 2025

Abstract: Optical neural networks have demonstrated their potential to overcome the computational bottleneck of modern digital electronics. However, their development towards high-performing computing alternatives is hindered by one of the optical neural networks' key components: the activation function. Most of the reported activation functions rely on opto-electronic conversion, sacrificing the unique advantages of photonics, such as resource-efficient coherent and frequency-multiplexed information encoding. Here, we experimentally demonstrate a photonic nonlinear activation function based on stimulated Brillouin scattering. It is coherent and frequency selective and can be tuned all-optically to take LEAKYReLU, SIGMOID, and QUADRATIC shape. Our design compensates for the insertion loss automatically by providing net gain as high as 20 dB, paving the way for deep optical neural networks.

Keywords: optoacoustics; nonlinear optics; nonlinear activation function; photonic neuromorphic computing; Brillouin scattering; optical fiber; optical neural network

Grigorii Slinkov and Steven Becker contributed equally to this work.

*Corresponding author: **Birgit Stiller**, Max-Planck-Institute for the Science of Light, Staudtstr. 2, 91058 Erlangen, Germany; Department of Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstr. 7, 91058 Erlangen, Germany; and Institute of Photonics, Leibniz University Hannover, Welfengarten 1A, 30167 Hannover, Germany, E-mail: birgit.stiller@mpl.mpg.de.

<https://orcid.org/0000-0002-9868-9998>

Grigorii Slinkov and Steven Becker, Max-Planck-Institute for the Science of Light, Staudtstr. 2, 91058 Erlangen, Germany; and Department of Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstr. 7, 91058 Erlangen, Germany.

<https://orcid.org/0000-0003-1258-5854> (G. Slinkov).

<https://orcid.org/0000-0003-3323-8800> (S. Becker)

Dirk Englund, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

<https://orcid.org/0000-0002-1043-3489>

1 Introduction

Artificial neural networks (ANNs) have emerged as powerful instruments for solving difficult tasks that range from speech recognition to image processing in medicine. Thanks to their self-learning abilities and nonlinearity [1], they can provide creative solutions derived from their training on large data sets. After years of rapid scaling of model complexity, machine learning inference and training are close to reaching a bottleneck formed by the limitations of conventional Boolean logic processing hardware, especially with regard to power consumption, latency, and data movement.

Overcoming this “von Neumann” bottleneck has motivated the search for new ANN computing architectures based on fundamentally different principles. Transferring the linear algebraic operation – vector-matrix multiplication – of ANNs to the optical domain, in particular, yields a potential for fundamental improvements in energy consumption and latency [2]–[11]. Although these approaches have demonstrated the potential of optical neural networks (NNs), most of them achieve nonlinearity through opto-electro-optic conversion or digital post-processing [5]–[10], [12]–[14]. As the opto-electronic conversion of the signal at each neuron limits the power efficiency, the computing speed and the scalability of the system [15], all-optical activation functions are demanded [15]–[18].

We propose a list of desirable, but so far not demonstrated in one “package”, features which is aimed at maximizing the benefits of having an all-optical activation function. These features are: (i) programmable nonlinearity, (ii) low insertion loss, (iii) coherence, (iv) WDM capability, and (v) compatibility with on-chip designs. The aim of this list is to ensure that the activation function will not impose additional limitations on the optical neural network as well as add new capabilities, allowing to fully exploit the potential of an optical neural network. A programmable activation function (i) allows the optical NN to adapt better to a specific problem and can be used as an additional training parameter. It has been shown for digital NNs that this additional degree of freedom is beneficial for the NN's performance [19], [20]. Insertion losses (ii) limit the depth of the neural network, reducing the number of layers that

can be stacked before the signal will have to be amplified. A coherent activation function (iii) is not only beneficial for phase-based optical NN architectures, such as [7], [9], but can also allow to implement efficient training schemes [21]. The WDM compatibility (iv) is essential for applying resource-efficient frequency-basis information encoding, which is a unique feature of photonics. Due to the overall lack of a frequency-selective activation function, multi-frequency photonic machine learning architectures have been so far limited to vector-matrix multiplication [10], [22], [23].

In this work, we experimentally demonstrate an optoacoustic activation function that combines features (i)–(v) (see Figure 1). Our design is based on the nonlinear effect of stimulated Brillouin scattering (SBS) [24]–[26], which arises from the interplay between optical and acoustic fields. SBS is inherently frequency-selective [27], which makes it particularly suitable for resource-efficient frequency-basis information encoding. This means that our all-optical activation function treats different frequencies independently, while being coherent. Moreover, our approach amplifies the signal with a positive net gain, facilitating its use in deep

optical NNs. The nonlinear response is controlled all-optically and can be tuned continuously between different activation function shapes, including **LEAKYReLU**, **SIGMOID** and **QUADRATIC**, which are favored by the machine learning community [5], [20].

While Brillouin scattering has been traditionally used for lasers [28]–[30], sensing [31], [32], gyroscopes [33], [34], and microscopy [35]–[37], with our work we reveal new applications, going beyond what has been previously demonstrated for Brillouin-based signal processing [38]–[44]. Our approach is not limited to a specific platform as SBS can be observed in different waveguide types ranging from optical chips [45]–[53] and microresonators [54]–[58] to photonic crystal fiber [39], [59], [60] and fiber tapers [61], [62].

In the following, we demonstrate different nonlinear input-output mappings of the photonic activation function and its tunability. In addition, we apply it to a dual-frequency signal, with a 3 GHz channel separation, demonstrating its frequency selectivity. Finally, we simulate the performance of our activation function using a digital neural network.

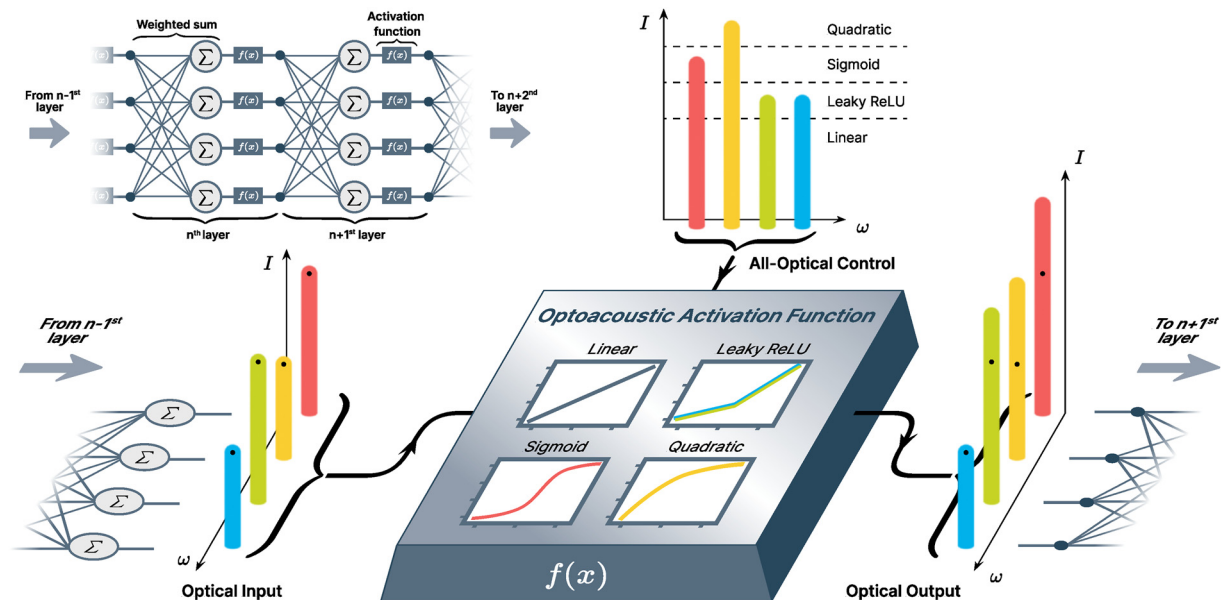


Figure 1: A schematic representation of how an optoacoustic activation function can be employed between layers $n - 1$ and $n + 1$ of an all-optical multi-frequency neural network. The result of the weighted summation of the $n - 1$ layer is encoded in the magnitudes of individual spectral components and sent to the optoacoustic activation function together with the multi-frequency control signal. The magnitudes of the control signal's spectral components define the type of activation function applied to the corresponding input signal component. The optical output of the optoacoustic activation function has the same frequency components as the input, with their magnitudes transformed nonlinearly depending on the type of activation function. The output is fed to the next layer of the neural network. The inset shows a conceptual scheme of a neural network. Each neuron performs two crucial operations: it takes a weighted sum of the inputs and applies a nonlinear activation function to the result.

2 Results

2.1 SBS-based nonlinear activation function

Stimulated Brillouin scattering (SBS) is a third-order nonlinear effect that couples a pair of counterpropagating optical waves with a traveling acoustic wave serving as a mediator between them [24]. It follows a strict phase-matching condition [27]: the frequencies of the optical waves, propagating in opposite directions, have to be separated by the acoustic wave's frequency Ω , which, for a given optical wavelength, is defined by the properties of the interaction medium. A schematic of an experimental realization is depicted in Figure 2A, where the probe wave a_{probe} is taken to be the one with the lower frequency ω and the pump wave a_{pump} oscillates with the frequency $\omega + \Omega$. The interaction between the fields a_{probe} , a_{pump} , and b can be described formally with the interaction Hamiltonian (1) [63]:

$$\hat{H}_{\text{int}} = \hbar g \int_{-\infty}^{\infty} dz \left(\hat{a}_{\text{pump}} \hat{a}_{\text{probe}}^{\dagger} \hat{b}^{\dagger} + \text{H.c.} \right), \quad (1)$$

with the optoacoustic coupling constant g and the time- and space-dependent wave packet operators $\hat{a}_{\text{probe}}(z, t)$, $\hat{a}_{\text{pump}}(z, t)$, $\hat{b}(z, t)$ of the probe, pump and acoustic field, respectively. This three-wave optoacoustic interaction holds at the quantum level [63]–[65] as well as in the classical regime [24], [26] and has been studied widely in

optical fibers and on-chip configurations [39], [45]–[62]. A detailed description of SBS can be found in the Supplementary material.

Our nonlinear activation function is based on a modified Brillouin amplifier scheme (see Figure 3A) that has been demonstrated to deliver high-gain and low-noise operation [66]–[68]. Under the undepleted pump assumption (UPA), the equations governing SBS admit a simple analytic solution. UPA implies that the pump is not affected by the Brillouin interaction. In this case, the relation between the optical powers of the probe P_{in} , the pump P_{pump} and the amplified probe P_{out} can be written as follows [25], [69]:

$$P_{\text{out}} = P_{\text{in}} \exp(gL \cdot P_{\text{pump}}), \quad (2)$$

where g is the Brillouin gain and L is the interaction length.

We consider now a slightly different situation, where the usually independent pump power is made dependant on the optical input of the Brillouin amplifier: $P_{\text{pump}} = P_{\text{pump}}(P_{\text{in}}) = \gamma P_{\text{in}}$. Such a relation can be established using the scheme shown in Figure 3B, given that the chosen optical amplifier provides linear amplification. This way the information encoded in the intensity of the input light will be transferred to both the probe and the pump of the Brillouin amplifier.

Hence, the equation [2] would write as:

$$P_{\text{out}} = P_{\text{in}} \exp(gL \cdot \gamma P_{\text{in}}). \quad (3)$$

Solving numerically the corresponding set of coupled mode equations for SBS allows us to study the behaviour of

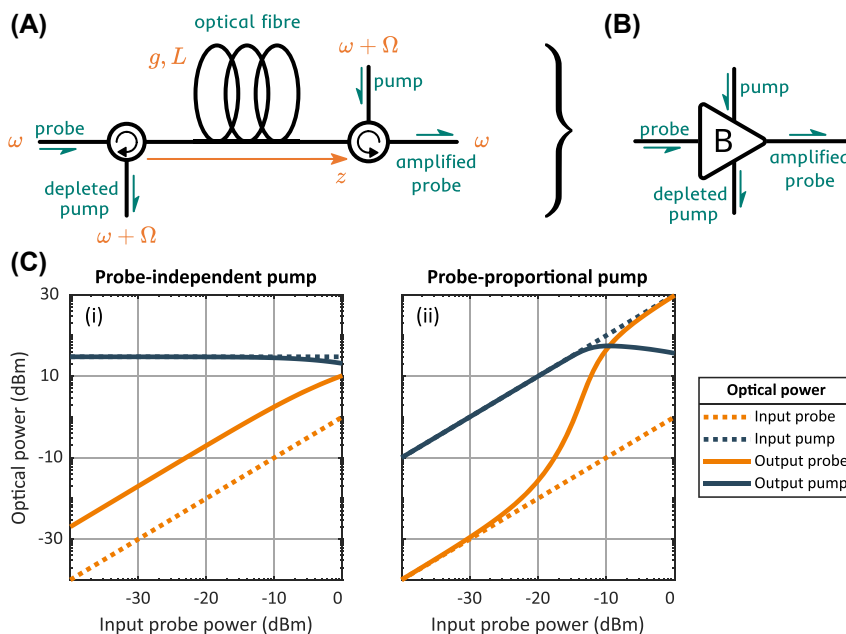


Figure 2: Brillouin amplifier schematics and modes of operation. (A) A schematic of an optical fiber-based Brillouin amplifier. (B) A convention introduced to represent the Brillouin amplifier. (C) Numerical simulation of SBS process, showing linear (i) and nonlinear (ii) relation between the input and the output probe optical power.

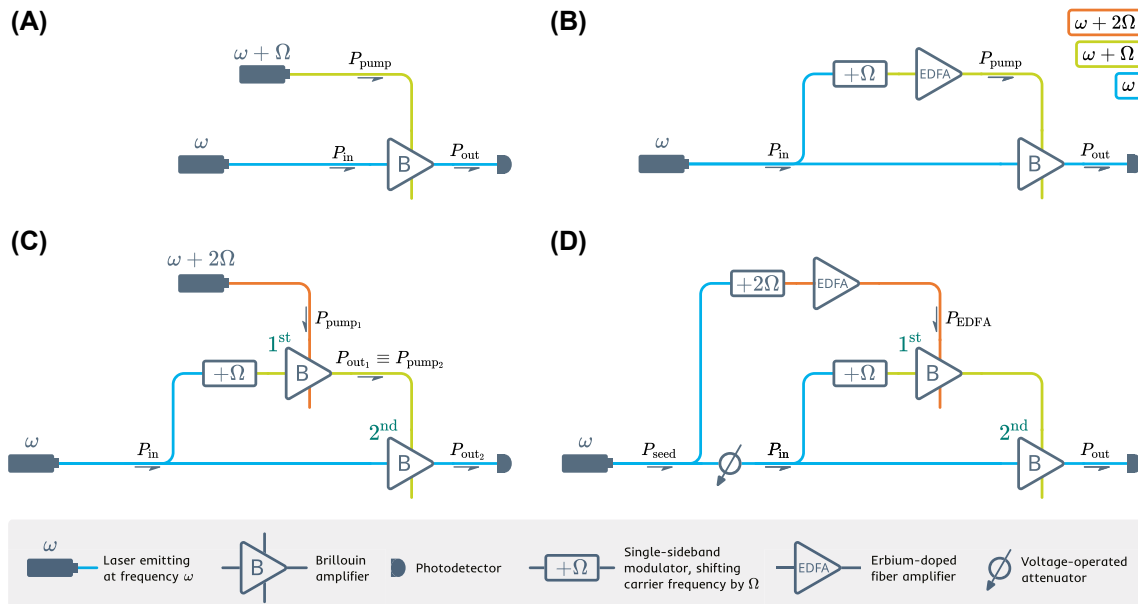


Figure 3: Schematic evolution of the activation function design. The steps that lead from the conventional Brillouin amplifier (A) and up to the double-stage scheme that is utilized in the experiment (D) are displayed.

an input-dependent Brillouin amplifier for applying it as a versatile nonlinear activation function for an optical neural network architecture. The corresponding equations can be found in the Supplementary material.

The simulation results are presented in Figure 2C. The $g \cdot L$ product is taken to be 0.1 in our simulation. In both panels, the input probe power $P_2(0)$ is swept from -40 to 0 dBm. In the left panel, the input pump power P_{pump} is kept constant, which results in a linear dependence of probe output P_{out} on the probe input P_{in} . As the output probe power surpasses -10 dBm, the pump starts to get depleted, and so the probe output growth rate decreases as well. In the right panel, both probe P_{in} and pump P_{pump} inputs are swept at the same rate. As one can see, this changes drastically the output probe dynamic: the dependence goes from linear to exponential growth and then back to linear, as we begin to observe the pump depletion. This nonlinear input-output relation in combination with the strict phase matching condition of SBS represents a perfect tool to implement a frequency-selective nonlinear activation function.

However, the scheme shown in Figure 3B had to be modified for the experiment. As can be seen in Figure 2C (ii), the nonlinear behavior of the output probe requires an input pump with high power and a high dynamic range. In order to satisfy this condition, we replace the EDFA with a Brillouin amplifier (see Figure 3C). In the created double-stage scheme, a Brillouin amplifier is pumped using the output of another Brillouin amplifier, both stages sharing

the same probe. In order to increase the stability of the scheme, we utilize the same light source to create the pump for the first stage Brillouin amplifier, which comprises the final design (see Figure 3D). The information is encoded in the intensity of the input light using a voltage-operated attenuator placed after a portion of input has been utilized to create the pump.

2.2 Single-wavelength operation

The measurement results are presented in Figure 4, where the output power is plotted against the input power for different 1st stage pump power levels provided by an erbium-doped fiber amplifier (EDFA). Panel A shows a selection of activation function shapes accessible with the setup; panel B demonstrates the complete family of activation function curves. The variation of the EDFA power allows to choose a specific curve, adjusting the activation function shape continuously. When the EDFA is turned off, the activation function is linear.

The selected curves in panel A are fitted with analytical activation functions. A pump power of 30.5 dBm corresponds to the LEAKY ReLU function. Its first section is a direct proportionality between the input and the output in the absence of SBS process – in this region the pump power is too low to be in the exponential regime of SBS. When the corresponding pump power exceeds the Brillouin threshold, the input light gets amplified. As a result, the input-output dynamic changes its slope as required for LEAKY ReLU. The

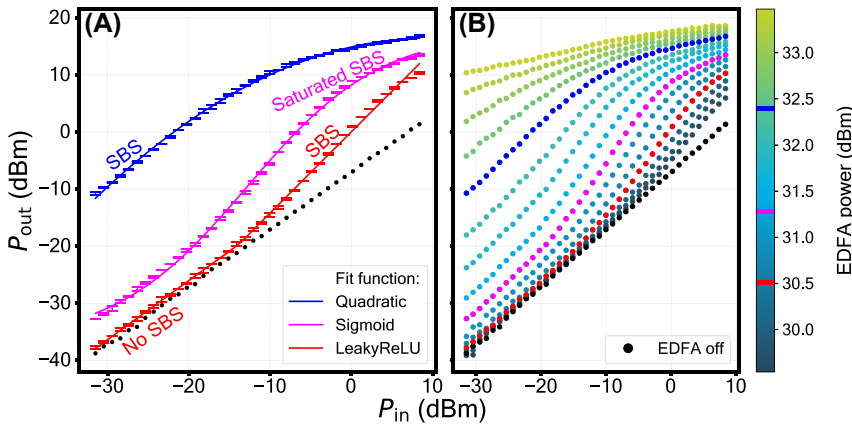


Figure 4: Nonlinear activation function shapes – the mapping between the optical input and the optical output of the setup. (A) A selection of curves fitted with conventional analytic activation functions. (B) The complete family of curves obtained at various 1st stage pump optical power levels provided by the EDFA.

curve can be fitted with a polynomial of degree 4 (fit parameters are presented in section S8 of the supplement). The next curve, obtained at pump power of 31.3 dBm, is fitted with SIGMOID. Its nonmonotonic growth can be split into three distinct sections. First, the absence of SBS in the beginning. Second, the amplification provided by SBS in the middle section of the plot. Lastly, the saturated SBS in the final section, where the SBS process becomes so intense that the pump starts to get depleted, resulting in the saturation of the growth. The curve can be fitted with a polynomial of degree 3 (fit parameters are presented in section S8 of the supplement). The last curve in Figure 4A, obtained at 32.4 dBm pump power, is fitted with a QUADRATIC function, formed by a SBS process that gives way to saturated SBS as the probe power is increased. The curve can be fitted with a polynomial of degree 2 (fit parameters are presented in section

S8 of the supplement). The optoacoustic activation function provides a gain of up to 2.4 dB, 8.8 dB, and 21.9 dB for the LEAKY ReLU, SIGMOID, and QUADRATIC case, respectively. Hence, the optoacoustic activation function can be used to compensate for losses induced by the preceding matrix operation. This is an essential feature for implementing deep optical NNs.

It is also useful to consider how the activation functions perform a linear scale. Figure 5 shows the same selection of curves as in Figure 4A: “Leaky ReLU” (30.5 dBm EDFA power), “Sigmoid” (31.3 dBm EDFA power) and “Quadratic” (32.4 dBm EDFA power). The experimental data is fitted with polynomial functions. The insets (panels B, D, F) are used to demonstrate the features found at low input power. The fit parameters can be found in the Supplementary material.

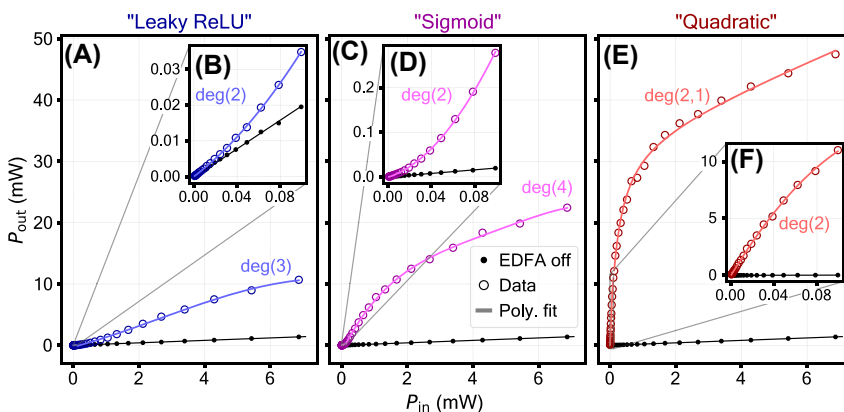


Figure 5: Nonlinear activation function shapes – the mapping between the optical input and the optical output of the setup. The polynomial fit degree p is specified as $\text{deg}(p)$. Panels A, C, E show the full range of the input power used in the experiment. In panels B, D, F (insets) the lower-input parts of the activation functions are fitted separately. In panel E a rational function of order 2,1 (second degree polynomial divided by a first degree polynomial) is used. This is due to a nearly-asymptotic drop at low input power which cannot be fitted reasonably with a polynomial.

2.3 Dual-wavelength operation

We demonstrate the feature of frequency selectivity by splitting the input into two wavelength-multiplexed channels, provided by two tunable lasers at the input. The frequency separation between the two is set to 3 GHz, limited in the experiment by the spectrum analyzing device resolution. Each of the two channels hosts its own variable attenuator, which allows us to control the power levels independently (see supplement for details). We use a Finisar Wave-Analyzer to perform a wavelength-selective measurement of optical power at the output. The measurement results are plotted in Figure 6. Panel **A** shows the case where the power in channel 1 is swept, while the power in channel 2 is kept constant, panel **B** is vice versa. In panel **C**, both channels are swept simultaneously. The reference case where both channels are swept, but the EDFA is turned off can be found in the Supplementary material. The presented selection shows that for a given channel neither the presence, nor the variation of a signal in the neighbouring channel affects the shape of the nonlinear activation function. The EDFA power required to achieve the SIGMOID shape in the dual-wavelength case (34.5 dBm, Figure 6) is higher than what is needed to achieve the same shape in the single-wavelength case (Figure 4): 31.5 dBm. This is due to EDFA distributing its output power equally between the pumps of the two frequency channels, yielding a 3 dB less power per channel.

2.4 Simulated performance in a neural network

In the following, we study how our activation functions could behave in an optical neural network. For

this, we take references [12]–[15], [70] as an example and equip a digital neural network with our activation function. We base our digital neural network architecture on experimentally demonstrated optical neural networks. More precisely, our digital network represents the size of one of the largest optical neural networks reported so far [9], which consists of three layers of six neurons each (see Figure 7A). In addition, we set the bias of all neurons to zero similar to the optical neural network. We evaluate the performance of the digital neural network using two tasks. Firstly, we perform a six class vowel recognition task as done in [9] (see Figure 7B). Secondly, we perform a binary classification task on a spiral dataset (see Figure 7C).

We perform each task with six different activation functions, one of which is a linear function (see Figure 7D). We take the digital implementations of the well-known ReLU and sigmoid functions and compare them to our LeakyReLU, Sigmoid, and quadratic realizations which are defined in the Supplementary material. We use a softmax activation function in the last layer of the digital neural network for all six cases.

To use our activation functions for the two tasks, we linearly transform the input data into the range $[p - \Delta/2, p + \Delta/2]$, where p is the center position of the optical input and Δ is the dynamic range of the simulated optical input. With Δ , we assess how our activation functions perform for different input dynamic ranges. In the following studies, we use $p = -13$ and $\Delta \in \{6, 10, 20\}$. More details about the underlying machine learning pipeline can be found in the Supplementary material.

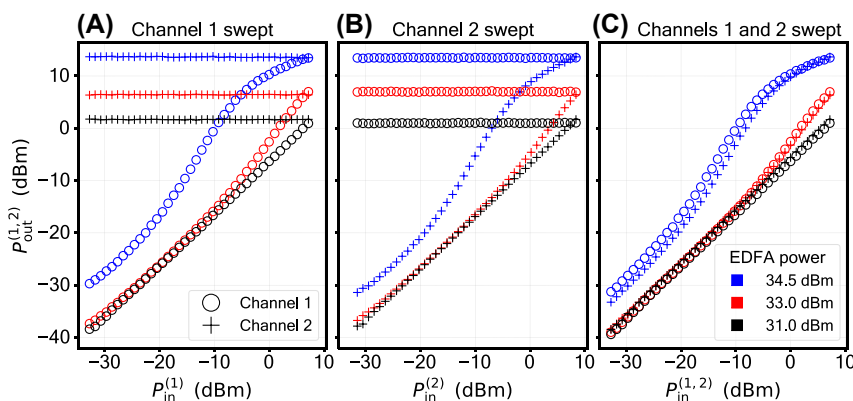


Figure 6: Demonstration of the frequency-selective operation. Circles and crosses correspond to a pair of wavelength-multiplexed channels, provided by two lasers at the input. The shape of the activation function in either of the channels is not affected by the SBS interaction taking place in the neighbouring channel. The instrumentation and the data acquisition process are identical between this measurement and the single-wavelength measurement, giving the same error level as in Figure 4.

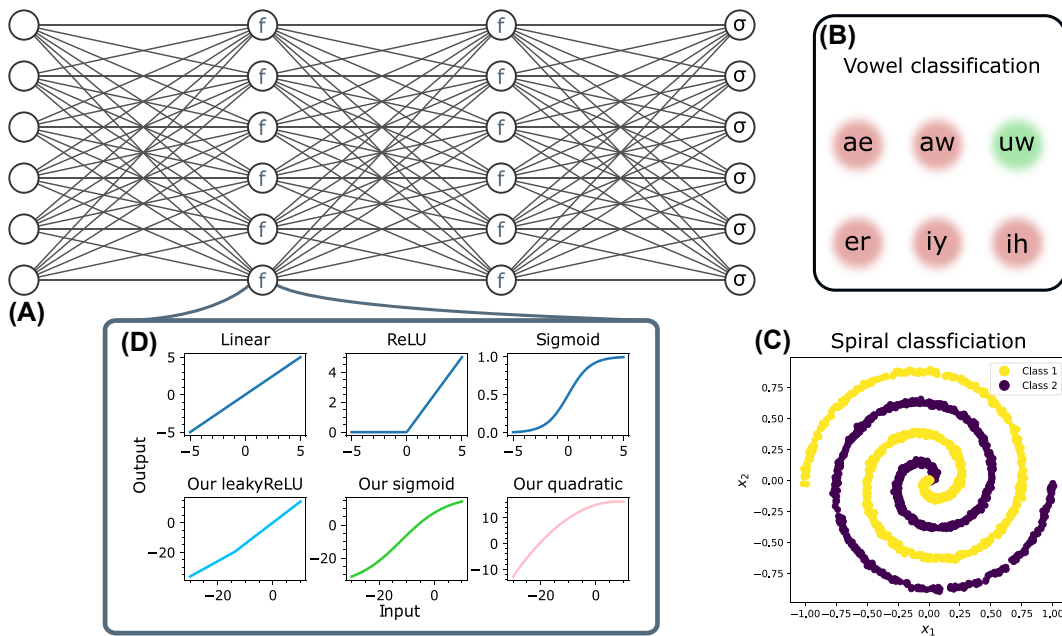


Figure 7: Testing our activation function in a digital neural network. (A) Three-layer neural network. We always use a softmax activation function in the output layer. Illustration of the neural network was created with [71]. (B) and (C) We apply the neural network for vowel and spiral classification, respectively. (D) We test the neural network with six different activation functions including the typical used ReLU and sigmoid functions.

2.4.1 Vowel recognition task results

The vowel dataset is the one used in [72] and taken from [73]. Analogous to [9], we extract the formant frequencies $F1$, $F2$, and $F3$ at steady-state and at half of the vowel duration from the provided dataset. In total, the dataset provides 1,667 samples. After standardizing the values, we feed them into our pipeline.

Figure 8 shows the highest achieved test accuracy for the different activation functions selected from all swept parameters. Our activation functions perform as well as the digital sigmoid function. In fact, the neural networks based on our Sigmoid and Quadratic function achieved the highest training accuracy of 98.4 %. Furthermore, the results show that for this task the Δ value has only a little effect on the accuracy. Interestingly, the linear activation functions lead

to a higher accuracy of 94.4 % than the ReLU activation function with 92.0 %.

One underlying reason for this could be the dying ReLU problem. This problem occurs when a neuron only outputs zero for any given input, which in turn leads to vanishing gradients [74]. In addition, it could be that the ReLU-based models are stuck in a local minimum for all hyper parameters configuration tested in this study. Consequently, this could lead to a lower accuracy as the one achieved by the models using a linear activation function.

As a result, the optimizer can no longer improve the neural network. Bandyopadhyay et al. achieved an accuracy of 92.7 % [9] which is slightly less than the results presented here. The reason for this could be that our digital neural network is not limited to unitary matrix operations. Details about the training dynamics can be found in the Supplementary material.

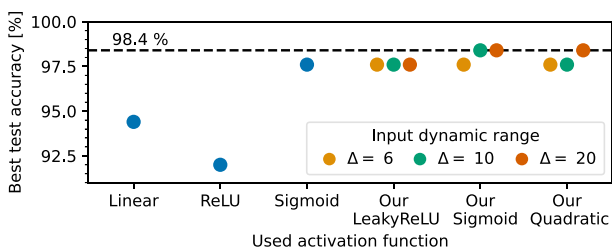


Figure 8: Best achieved test accuracy for the vowel task for the different activation functions.

2.4.2 Spiral classification task results

We generate the spiral data randomly starting by drawing an angle $\varphi_0 \in [0, 2\pi)$ from a uniform distribution. This angle is then converted to the spiral coordinate (x, y) using

$$\begin{aligned} \varphi &= v\varphi_0, & r &= \frac{s\varphi_0}{2\pi}, \\ x &= (r + \sigma_x) \cos(\varphi), & y &= (r + \sigma_y) \sin(\varphi), \end{aligned} \quad (4)$$

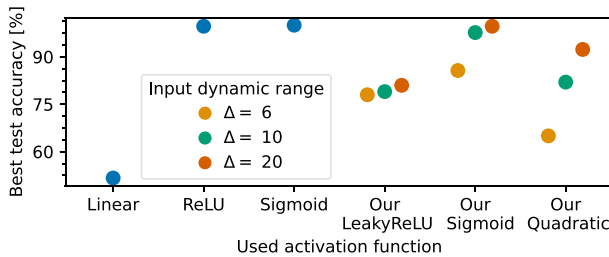


Figure 9: Best achieved test accuracy for the spiral task for the different activation functions.

with the number of spiral turns ν , the scale of the spiral s , and the noise on the x and y coordinates σ_x and σ_y , respectively. For our study, we set $\nu = 2$, $s = 1$, and generate the noise $\sigma_{x,y}$ with a Gaussian distribution with a mean value of 0 and with a standard deviation of 0.01. The resulting dataset is illustrated in Figure 7C. After sampling 1,000 data points, we follow the feature engineering approach of TENSORFLOW [75] and add the features x^2 , y^2 , and $x \cdot y$ to the dataset.

Figure 9 shows the highest achieved accuracies for the different activation functions. The study shows that this task cannot be solved with high accuracy by a neural network equipped with the linear activation function. The best test accuracy achieved by such a model is 55.0 %, which is slightly above the random guessing accuracy of 50 %. In fact, the networks equipped with ReLU and Sigmoid activation functions achieved test accuracies of almost 100 %. Only our Sigmoid activation function achieves a similar accuracy. Moreover, our activation functions benefit from higher Δ values. In particular, models with our Quadratic function achieve significantly higher accuracies for higher Δ values. The reason for this could be that a higher degree of nonlinearity is provided for an input dynamic range. This could also explain why networks with our LeakyReLU function are less sensitive to different Δ values. A potential reason is the chosen center position of the optical input p which is close to the P_0 parameter of our LeakyReLU function. Details about the training dynamics can be found in the Supplementary material.

3 Discussion

We have experimentally demonstrated for the first time, to the best of our knowledge, a nonlinear photonic activation function based on stimulated Brillouin scattering. Our activation function is coherent and frequency selective owing to the nature of SBS. A coherent activation function is the next step from the phase-reliant optical matrix multiplication

approaches [7], [9] to fully all-optical ANNs that would not require opto-electronic conversion, which imposes bandwidth limitations, introduces cross-talk, requires introduction of time delay and eliminates frequency selectivity.

The frequency selectivity opens up prospects of drastically increasing the data throughput by exploiting wavelength multiplexing techniques to, for example, distribute the neurons in a layer across the frequency domain. It is the first demonstration of a frequency sensitive activation function, which shows no correlation between frequency channels that hinders existing solutions, such as frequency-encoded deep neural networks [10]. As another example, one can upgrade the throughput rate of the scheme presented by Shen et al. [7] using multi-frequency operation if an EDFA with sufficient output power is used. Naturally, the single input laser must be replaced with a multi-frequency input source such as multiple lasers or a frequency comb source. Then in theory each wavelength channel could be used to execute an independent training and inference step.

As it has been shown, there is no cross-talk between neighbouring frequency channels at frequency separations as small as 3 GHz, which surpasses the telecommunication standard of 25 GHz. For the continuous wave case, the minimal frequency separation between the two channels is intrinsically limited by the linewidth of the optoacoustic gain function, which, for the commercial single-mode fiber at room temperature is about 26 MHz [76]. As shown in [44], [76], the linewidth can be decreased by lowering the temperature of the waveguide. In the pulsed case, the minimal frequency separation is dictated by the pulse length [27]. This limitation goes both ways, meaning that the minimal pulse length is dictated by the channel separation. The general rule of thumb for SBS-based applications is this: channel separation has to be higher than $\Delta\nu_{\text{laser}} + 1/\tau_{\text{pulse}} + \Delta\nu_{\text{B}}$, where $\Delta\nu_{\text{laser}}$ is the laser linewidth, τ_{pulse} is the pulse length and $\Delta\nu_{\text{B}}$ is the acoustic gain linewidth.

The activation function can be tuned by varying the 1st Brillouin amplifier's pump power to take such well-proven shapes as LEAKYReLU, SIGMOID, and QUADRATIC.

With a numerical study, we have shown that our activation function shapes can achieve similar results in a digital neural network as their digital counterparts. In addition, the external control of the dynamics of the 1st stage Brillouin amplifier opens the possibility to use the nonlinearity as an additional training parameter of an optical ANN.

For digital ANNs, this has been shown to be a powerful tool for boosting the ANN performance [19], [20]. It is also

feasible to engineer the 1st stage pump in such a way that different frequency channels would have different activation function shapes.

The output signal amplification that is inherent to our activation function design is suitable for compensating insertion and propagation losses. This should be particularly useful for designing deep optical NNs that comprise multiple neuron layers [21].

Though the experimental realisation presented in the paper relies on highly nonlinear optical fiber as the optoacoustic interaction medium, our approach is not limited to this platform. Conventional single-mode fiber (SMF) and photonic crystal core fiber (PCF) are also an eligible choice for SBS as well as integrated waveguides, including on-chip devices [41], [43], [77]–[79]. The choice of the platform combined with pulsed operation constitutes the way for improving the energy efficiency of the presented optoacoustic activation function. One needs to maximize the term $g \cdot P_{\text{pump}} L_{\text{eff}} - \alpha L$ in order to improve the energy efficiency. Here, P_{pump} is the pump power, L_{eff} is the effective interaction length, and α is the optical loss of the waveguide (see Suppl. for details).

In conclusion, our frequency selective and coherent photonic nonlinear activation fills a gap in the current landscape of photonic machine learning. It could therefore be the key to unlocking the full potential of photonic neuromorphic computing.

4 Materials and methods

We implement the optoacoustic nonlinear activation using a setup depicted schematically in Figure 10. We build what can be called a double-stage Brillouin amplifier: the output of one Brillouin amplifier (1st stage) is utilized as a pump for another (2nd stage). The 1st stage Brillouin amplifier is pumped with an Erbium-doped fiber amplifier (EDFA).

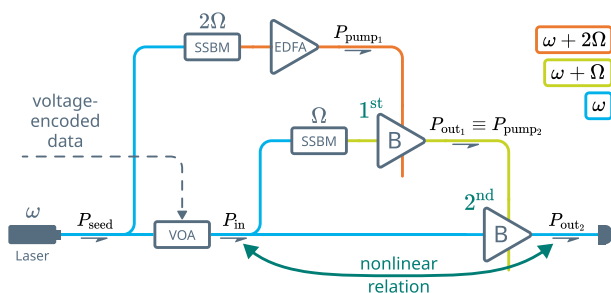


Figure 10: A principal scheme of the experimental setup: VOA – voltage-operated attenuator, SSBM – single-sideband modulator, EDFA – Erbium-doped fiber amplifier, B – Brillouin amplifier, as introduced in Figure 2B. The color of the connecting lines depicts the light frequency.

The reason for pumping the 2nd stage with a Brillouin amplifier, as opposed to applying an EDFA directly, is that conventional EDFAs operate in the saturated regime, providing input-independent output power. This way, replacing the 1st stage with an EDFA would have eliminated any possible relation between the input and the pump that is required by [3]. The setup is fed with a 1550.12 nm fiber-coupled laser. A voltage-controlled attenuator (VOA) is used in the experiment to test the nonlinear input-output behaviour of the setup, simulating the amplitude-encoded data from the previous neuron layer. The VOA is inserted before the light gets distributed between the two stages, which ensures that they receive the same amplitude variation, as required by [3]. Note that feeding the top branch of the setup with the same laser is a measure taken to enhance the stability of the Brillouin amplifier and is not an actual requirement.

A Brillouin frequency shift f_B (corresponding angular frequency $\Omega = 2\pi f_B$) is applied to the middle branch of the setup, satisfying the SBS phase matching condition for the 2nd stage. This requires the signal in the top branch to be up-shifted by the sum of the Brillouin frequencies of the fibers. The optical fiber used in both of the stages was of the same material and structure, yielding a 2Ω shift for the top branch.

We use highly nonlinear fiber (HNLF) with equal parameters for the two Brillouin amplifiers. The lengths of the HNLF fibers used for the first and the second stage are 20 m and 100 m, correspondingly. The Brillouin frequency of the fibers is $f_B = 9.730$ GHz. A couple of single sideband modulators (SSBMs) driven with two separate RF sources apply required frequency shifts to the top and the middle branches of the setup.

Acknowledgments: The authors thank Jakob Storp for fruitful discussions.

Research funding: We acknowledge funding from: DFG Project STI 792/1-1; Max Planck Society through the Independent Max Planck Research Groups scheme; Studienstiftung des deutschen Volkes.

Author contributions: Initial idea: SB, BS. Conceptualization: SB, GS, DE, BS. Methodology: SB, GS, BS. Investigation: SB, GS. Visualization: SB, GS. Supervision: BS. Writing – original draft: SB, GS, DE, BS. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: Authors state no conflicts of interest.

Data availability: The data that supports findings of this study are available from the corresponding author upon request.

References

- [1] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, “State-of-the-art in artificial neural network applications: A survey,” *Heliyon*, vol. 4, no. 11, p. e00938, 2018.
- [2] G. V. D. Sande, D. Brunner, and M. C. Soriano, “Advances in photonic reservoir computing,” *Nanophotonics*, vol. 6, no. 3, pp. 561–576, 2017.
- [3] W. Bogaerts, *et al.*, “Programmable photonic circuits,” *Nature*, vol. 586, no. 7828, pp. 207–216, 2020.
- [4] T. Wu, M. Menarini, Z. Gao, and L. Feng, “Lithography-free reconfigurable integrated photonic processor,” *Nat. Photonics*, vol. 17, no. 8, pp. 710–716, 2023.
- [5] C. M. Valensise, I. Grecco, D. Pierangeli, and C. Conti, “Large-scale photonic natural language processing,” *Photon. Res.*, vol. 10, no. 12, pp. 2846–2853, 2022.
- [6] A. Sludds, *et al.*, “Delocalized photonic deep learning on the internet’s edge,” *Science*, vol. 378, no. 6617, pp. 270–276, 2022.
- [7] Y. Shen, *et al.*, “Deep learning with coherent nanophotonic circuits,” *Nat. Photonics*, vol. 11, no. 7, pp. 441–446, 2017.
- [8] Z. Chen, *et al.*, “Deep learning with coherent VCSEL neural networks,” *Nat. Photonics*, vol. 17, pp. 723–730, 2023.
- [9] S. Bandyopadhyay, *et al.*, “Single-chip photonic deep neural network with forward-only training,” *Nat. Photonics*, vol. 18, pp. 1335–1343, 2024. <https://doi.org/10.1038/s41566-024-01567-z>
- [10] R. Davis III, Z. Chen, R. Hamerly, and D. Englund, “Frequency-encoded deep learning with speed-of-light dominated latency,” arXiv: 2207.06883[physics], 2022.
- [11] B. J. Shastri, *et al.*, “Photonics for artificial intelligence and neuromorphic computing,” *Nat. Photonics*, vol. 15, no. 2, pp. 102–114, 2021.
- [12] M. M. Pour Fard, *et al.*, “Experimental realization of arbitrary activation functions for optical neural networks,” *Opt. Express*, vol. 28, no. 8, p. 12138, 2020.
- [13] A. Jha, C. Huang, and P. R. Prucnal, “Reconfigurable all-optical nonlinear activation functions for neuromorphic photonics,” *Opt. Lett.*, vol. 45, no. 17, pp. 4819–4822, 2020.
- [14] I. A. D. Williamson, *et al.*, “Reprogrammable electro-optic nonlinear activation functions for optical neural networks,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 26, no. 1, pp. 1–12, 2020.
- [15] M. Miscuglio, *et al.*, “All-optical nonlinear activation function for photonic neural networks [Invited],” *Opt. Mater. Express*, vol. 8, no. 12, p. 3851, 2018.
- [16] G. Mourgas-Alexandris, A. Tsakyrdis, N. Passalis, A. Tefas, K. Vysokinos, and N. Pleros, “An all-optical neuron with sigmoid activation function,” *Opt. Express*, vol. 27, no. 7, pp. 9620–9630, 2019.
- [17] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, “All-optical spiking neurosynaptic networks with self-learning capabilities,” *Nature*, vol. 569, no. 7755, pp. 208–214, 2019.
- [18] Y. Shi, *et al.*, “Nonlinear germanium-silicon photodiode for activation and monitoring in photonic neuromorphic networks,” *Nat. Commun.*, vol. 13, p. 6048, 2022.
- [19] F. Agostinelli, M. Hoffman, P. Sadowski, and P. Baldi, “Learning activation functions to improve deep neural networks,” arXiv: 1412.6830[cs,stat], 2015.
- [20] S. R. Dubey, S. K. Singh, and B. B. Chaudhuri, “Activation functions in deep learning: A comprehensive survey and benchmark,” *Neurocomputing*, vol. 503, pp. 92–108, 2022.
- [21] V. Lopez-Pastor and F. Marquardt, “Self-learning machines based on Hamiltonian echo backpropagation,” *Phys. Rev. X*, vol. 13, p. 031020, 2023, arXiv: 2103.04992[nlin,physics:physics].
- [22] J. Feldmann, *et al.*, “Parallel convolutional processing using an integrated photonic tensor core,” *Nature*, vol. 589, no. 7840, pp. 52–58, 2021.
- [23] S. Buddhiraju, A. Dutt, M. Minkov, I. A. D. Williamson, and S. Fan, “Arbitrary linear transformations for photons in the frequency synthetic dimension,” *Nat. Commun.*, vol. 12, p. 2401, 2021.
- [24] C. Wolff, M. J. A. Smith, B. Stiller, and C. G. Poulton, “Brillouin scattering — theory and experiment: Tutorial,” *J. Opt. Soc. Am. B*, vol. 38, no. 4, p. 1243, 2021.
- [25] R. W. Boyd, *Nonlinear Optics*, 3rd ed., Amsterdam; Boston, Academic Press, 2008, p. 613.
- [26] A. Kobayakov, M. Sauer, and D. Chowdhury, “Stimulated Brillouin scattering in optical fibers,” *Adv. Opt. Photon.*, vol. 2, no. 1, p. 1, 2010.
- [27] B. Stiller, P. Dainese, and E. Verhagen, “Optoacoustics — advances in high-frequency optomechanics and Brillouin scattering,” *APL Photonics*, vol. 4, no. 11, p. 040802, 2019.
- [28] S. Gundavarapu, *et al.*, “Sub-hertz fundamental linewidth photonic integrated Brillouin laser,” *Nat. Photonics*, vol. 13, pp. 60–67, 2019.
- [29] N. T. Otterstrom, R. O. Behunin, E. A. Kittlaus, Z. Wang, and P. T. Rakich, “A silicon Brillouin laser,” *Science*, vol. 360, pp. 1113–1116, 2018.
- [30] N. Chauhan, *et al.*, “Visible light photonic integrated Brillouin laser,” *Nat. Commun.*, vol. 12, p. 4685, 2021.
- [31] C. A. Galindez-Jamioy and J. M. López-Higuera, “Brillouin distributed fiber sensors: An overview and applications,” *J. Sens.*, pp. 1–17, 2012, <https://doi.org/10.1155/2012/204121>.
- [32] A. Geilen, A. Popp, D. Das, *et al.*, “Extreme thermodynamics in nanolitre volumes through stimulated Brillouin–Mandelstam scattering,” *Nat. Phys.*, vol. 19, pp. 1805–1812, 2023.
- [33] J. Li, M.-G. Suh, and K. Vahala, “Microresonator Brillouin gyroscope,” *Optica*, vol. 4, no. 3, p. 346, 2017.
- [34] Y.-H. Lai, *et al.*, “Earth rotation measured by a chip-scale ring laser gyroscope,” *Nat. Photonics*, vol. 14, pp. 345–349, 2020.
- [35] G. Antonacci, *et al.*, “Recent progress and current opinions in Brillouin microscopy for life science applications,” *Biophys. Rev.*, vol. 12, pp. 615–624, 2020.
- [36] R. Prevedel, A. Diz-Muñoz, G. Ruocco, and G. Antonacci, “Brillouin microscopy: An emerging tool for mechanobiology,” *Nat. Methods*, vol. 16, pp. 969–977, 2019.
- [37] G. Scarcelli and S. H. Yun, “Confocal Brillouin microscopy for three-dimensional mechanical imaging,” *Nat. Photonics*, vol. 2, pp. 39–43, 2008.
- [38] S. Becker, D. Englund, and B. Stiller, “An optoacoustic field-programmable perceptron for recurrent neural networks,” arXiv: 2309.01543[physics], 2023.
- [39] X. Zeng, *et al.*, “Nonreciprocal vortex isolator via topology-selective stimulated Brillouin scattering,” *Sci. Adv.*, vol. 8, no. eabq6064, 2022.
- [40] D. Marpaung, J. Yao, and J. Capmany, “Integrated microwave photonics,” *Nat. Photonics*, vol. 13, no. 2, pp. 80–90, 2019.

- [41] B. J. Eggleton, C. G. Poulton, P. T. Rakich, M. J. Steel, and G. Bahl, “Brillouin integrated photonics,” *Nat. Photonics*, vol. 13, pp. 664–677, 2019.
- [42] Z. Zhu, D. J. Gauthier, and R. W. Boyd, “Stored light in an optical fiber via stimulated Brillouin scattering,” *Science*, vol. 318, pp. 1748–1750, 2007.
- [43] M. Merklein, B. Stiller, K. Vu, S. J. Madden, and B. J. Eggleton, “A chip-integrated coherent photonic-phononic memory,” *Nat. Commun.*, vol. 8, p. 574, 2017.
- [44] S. Becker, A. Geilen, and B. Stiller, “High-speed coherent photonic random-access memory in long-lasting sound waves,” arXiv: 2311.06219[physics], 2023.
- [45] R. Van Laer, B. Kuyken, D. Van Thourhout, and R. Baets, “Interaction between light and highly confined hypersound in a silicon photonic nanowire,” *Nat. Photonics*, vol. 9, no. 3, pp. 199–203, 2015.
- [46] E. A. Kittlaus, H. Shin, and P. T. Rakich, “Large Brillouin amplification in silicon,” *Nat. Photonics*, vol. 10, no. 7, pp. 463–467, 2016.
- [47] H. Shin, J. A. Cox, R. Jarecki, A. Starbuck, Z. Wang, and P. T. Rakich, “Control of coherent information via on-chip photonic-phononic emitter-receivers,” *Nat. Commun.*, vol. 6, no. 1, p. 6427, 2015.
- [48] M. Merklein, B. Stiller, K. Vu, P. Ma, S. J. Madden, and B. J. Eggleton, “On-chip broadband nonreciprocal light storage,” *Nanophotonics*, vol. 10, pp. 75–82, 2021.
- [49] D. Munk, et al., “Surface acoustic wave photonic devices in silicon on insulator,” *Nat. Commun.*, vol. 10, no. 1, p. 4214, 2019.
- [50] F. Gyger, et al., “Observation of stimulated Brillouin scattering in silicon nitride integrated waveguides,” *Phys. Rev. Lett.*, vol. 124, no. 1, p. 013902, 2020.
- [51] R. Botter, et al., “Guided-acoustic stimulated Brillouin scattering in silicon nitride photonic circuits,” *Sci. Adv.*, vol. 8, no. eabq2196, 2022.
- [52] K. Ye, et al., “Surface acoustic wave stimulated Brillouin scattering in thin-film lithium niobate waveguides,” arXiv: 2311.14697[physics], 2023.
- [53] C. C. Rodrigues, et al., “On-chip backward stimulated Brillouin scattering in lithium niobate waveguides,” arXiv: 2311.18135[physics], 2023.
- [54] G. Bahl, M. Tömes, F. Marquardt, and T. Carmon, “Observation of spontaneous Brillouin cooling,” *Nat. Phys.*, vol. 8, no. 3, pp. 203–207, 2012.
- [55] J. Kim, M. C. Kuzyk, K. Han, H. Wang, and G. Bahl, “Number: 3 publisher: Nature publishing group,” *Nat. Phys.*, vol. 11, pp. 275–280, 2015.
- [56] E. A. Cryer-Jenkins, et al., “Second-order coherence across the Brillouin lasing threshold,” *Optica*, vol. 10, no. 11, pp. 1432–1438, 2023.
- [57] G.ENZIAN, et al., “Non-Gaussian mechanical motion via single and multiphonon subtraction from a thermal state,” *Phys. Rev. Lett.*, vol. 127, p. 243601, 2021.
- [58] G.ENZIAN, et al., “Observation of Brillouin optomechanical strong coupling with an 11 GHz mechanical mode,” *Optica*, vol. 6, no. 1, pp. 7–14, 2019.
- [59] P. Dainese, et al., “Stimulated Brillouin scattering from multi-GHz-guided acoustic phonons in nanostructured photonic crystal fibres,” *Nat. Phys.*, vol. 2, no. 6, pp. 388–392, 2006.
- [60] M. Pang, et al., “Stable subpicosecond soliton fiber laser passively mode-locked by gigahertz acoustic resonance in photonic crystal fiber core,” *Optica*, vol. 2, no. 4, pp. 339–342, 2015.
- [61] J.-C. Beugnot, S. Lebrun, G. Pauliat, H. Maillotte, V. Laude, and T. Sylvestre, “Brillouin light scattering from surface acoustic waves in a subwavelength-diameter optical fibre,” *Nat. Commun.*, vol. 5, no. 1, p. 5242, 2014.
- [62] W. Xu, A. Iyer, L. Jin, S. Y. Set, and W. H. Renninger, “Strong optomechanical interactions with long-lived fundamental acoustic waves,” *Optica*, vol. 10, no. 2, pp. 206–213, 2023.
- [63] J. Zhang, C. Zhu, C. Wolff, and B. Stiller, “Quantum coherent control in pulsed waveguide optomechanics,” *Phys. Rev. Res.*, vol. 5, p. 013010, 2023.
- [64] J. Sipe and M. Steel, “A Hamiltonian treatment of stimulated Brillouin scattering in nanoscale integrated waveguides,” *New J. Phys.*, vol. 18, p. 045004, 2016.
- [65] V. Laude and J.-C. Beugnot, “Lagrangian description of Brillouin scattering and electrostriction in nanoscale optical waveguides,” *New J. Phys.*, vol. 17, p. 125003, 2015.
- [66] C. N. Pannell, P. S. J. Russell, and T. P. Newson, “Stimulated Brillouin scattering in optical fibers: The effects of optical amplification,” *JOSA B*, vol. 10, pp. 684–690, 1993.
- [67] L. Xing, L. Zhan, S. Luo, and Y. Xia, “High-power low-noise fiber Brillouin amplifier for tunable slow-light delay buffer,” *IEEE J. Quantum Electron.*, vol. 44, no. 12, pp. 1133–1138, 2008.
- [68] M. Pelusi, T. Inoue, and S. Namiki, “Brillouin amplifier noise characterization by a coherent receiver and digital signal processing,” *J. Lightwave Technol.*, vol. 38, no. 16, pp. 4221–4236, 2020.
- [69] F. S. Gökhan, H. Göktaş, and V. J. Sorger, “Analytical approach of Brillouin amplification over threshold,” *Appl. Opt.*, vol. 57, no. 4, p. 607, 2018.
- [70] H. Li, B. Wu, W. Tong, J. Dong, and X. Zhang, “All-optical nonlinear activation function based on germanium silicon hybrid asymmetric coupler,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 29, no. 2, pp. 1–6, 2023.
- [71] NN SVG, 2024. <https://alexlenail.me/NN-SVG/>.
- [72] J. Hillenbrand and R. T. Gayvert, “Vowel classification based on fundamental frequency and formant frequencies,” *J. Speech Lang. Hear. Res.*, vol. 36, no. 4, pp. 694–700, 1993.
- [73] Julian, julian3rd/hillenbrand-vowel-clustering, original-date: 2015-08-12T19:32:46Z (2024; <https://github.com/julian3rd/hillenbrand-vowel-clustering>).
- [74] L. Lu, Y. Shin, Y. Su, and G. E. Karniadakis, “Dying ReLU and initialization: Theory and numerical examples,” *Commun. Comput. Phys.*, vol. 28, no. 5, pp. 1671–1706, 2020, arXiv: 1903.06733[cs,math,stat], <https://doi.org/10.4208/cicp.oa-2020-0165>.
- [75] Google, Neural Net Spiral, Google for Developers, 2024. Available from: <https://developers.google.com/machine-learning/crash-course/DPE/tp-ss-neural-net-intro-spiral>.
- [76] S. Le Floch and P. Cambon, “Study of Brillouin gain spectrum in standard single-mode optical fiber at low temperatures (1.4–370 K) and high hydrostatic pressures (1–250 bars),” *Opt. Commun.*, vol. 219, no. 1–6, pp. 395–410, 2003.

- [77] A. Choudhary, *et al.*, “Advanced integrated microwave signal processing with giant on-chip Brillouin gain,” *J. Lightwave Technol.*, vol. 35, no. 4, pp. 846–854, 2017.
- [78] L. B. Martínez, P. Wiedemann, C. Zhu, A. Geilen, and B. Stiller, “Optoacoustic cooling of traveling hypersound waves,” arXiv: 2305.19823[physics,physics:quant-ph], 2023.
- [79] B. Morrison, *et al.*, “Compact Brillouin devices through hybrid integration on silicon,” *Optica*, vol. 4, no. 8, p. 847, 2017.

Supplementary Material: This article contains supplementary material (<https://doi.org/10.1515/nanoph-2024-0513>).