Elements of Hybrid Opto-superconducting Convolutional Neural Networks

A. E. Schegolev[1,2,3,a], N. V. Klenov[1,2,3,4,b], M. V. Tereshonok[2] and S. S. Adjemov[2]

[1]*Faculty of Physics, Lomonosov Moscow State University, 119991, Moscow, Russia*

[2]*Moscow Technical University of Communication and Informatics, 111024, Moscow, Russia*

[3]*Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, 119234, Moscow, Russia*

[4]*Moscow Institute of Physics and Technology, Dolgoprudny, 141701, Russia*

Keywords: Artificial Intelligent, Integrated Photonics, Convolutional Neural Network, Neuron, Superconductor, ReLU.

Abstract: In this paper authors proposed the concepts and principals of operating of basic nonlinear elements for hybrid opto–superconducting convolutional neural network. Optical elements in computing systems are usually designed to produce only linear mathematical operations. This is insufficient for complete neural network realization on chip, where non-linear operations like activation function calculations in neuron or transfer function of rectifier linear unit are needed. We have shown the opportunity of realization of elemental base for the hybrid neural network consists of optical and superconducting parts.

1 INTRODUCTION

The creation of the hybrid architecture of neural networks for physical and mathematical calculations is an intriguing area of research for today. In this area, further strengthening of the positions of alternative element bases for computing systems is observed. In particular, an attempt to combine optical and superconducting physical processes in a hybrid neural network was provided in 1990 by Harold H. Szu (Szu, 1990). He has developed a neural architecture in the form of lattice of superconducting wires, in which local currents (and magnetic fields) in superconductive matrix was governed by electromagnetic (optical) radiation. This invention was proposed as "switching" mechanism in digital or analog applications in a superconducting computation. In this paper, we propose updating the concept of hybrid opto-superconducting neural networks with a magnetic representation of information. Particular attention will be given below to nonlinear network elements optimized for the currently used version of neural networks.

2 CONVOLUTIONAL NEURAL NETWORK

The widely used architecture of artificial neural networks is convolutional neural networks, was proposed by Yann LeCun in 1988. The main purpose of these networks are the recognition and analysis of images, by identifying important key features and screening of insignificant ones. The idea of convolutional networks as well as conventional ANNs appeared thanks to the analysis of the structure of the visual cortex of the animals' brain. Individual cortical neurons respond to stimuli only in a limited area of the visual field known as receptive field (Matsugu, 2003). The receptive fields of different neurons partially overlap thus it leads to the covering of the entire field of view. This feature tried to implement in artificial convolutional neural networks (CNN).

The advantages of CNN over the conventional ANN architectures is that they use relatively little pre-processing compared to other image classification algorithms. It means that the network learns to use some filters, for which usual networks are manually configured. The ability of CNN to create filters itself

[a] https://orcid.org/0000-0002-5381-3297

[b] https://orcid.org/0000-0001-6265-3670

Schegolev, A., Klenov, N., Tereshonok, M. and Adjemov, S.
Elements of Hybrid Opto-superconducting Convolutional Neural Networks.
DOI: 10.5220/0009100101350139
In *Proceedings of the 8th International Conference on Photonics, Optics and Laser Technology (PHOTOPTICS 2020)*, pages 135-139
ISBN: 978-989-758-401-5; ISSN: 2184-4364

that separate the key features of the image is its key advantage.

The general architecture of CNN is not a secret and mainly consists of convolutional layers, an activating layer, a downsampling layer or pooling layer and a fully connected neural network layer (usually a perceptron type is used). We will mainly be interested in the activating layer the distinctive feature of which is a presence of some function, filtering coming to the input scalar coefficients of the convolutional layer. Rectifier linear unit (ReLU) is exactly that element of CNN performing the role of this feature (Hahnloser, 2000), which can be mathematically expressed as $f(x)=max\{0, x\}$. ReLU is a filter of negative values, which allows one to increase the nonlinear properties of the decision function and the network as a whole. It doesn't affect the receptive fields of the convolutional layer itself (Glorot, Bordes, Bengio, 2011). In addition, ReLU allows to train CNN in several times faster than the other functions (hyperbolic tangent function, sigmoid function) without compromising of the generalizing features of the network (Nair, Hinton, 2010). Moreover, this function and its modifications (Noisy ReLU, Leaky ReLU) are the most often used activation functions in deep learning networks, in particular, convolutional neural networks.

ReLU is an inherent element of the CNN and its implementation on a superconducting base will allow the creation of a hybrid opto-superconducting CNN. As a rule, the so-called softmax function or leaky ReLU (which allows for a small, non-zero gradient when the unit is saturated and not active) are used, which show the best network performance (Maas et al, 2013).

For superconducting ReLU realization we present in the paper superconducting neuron scheme, that was developed in (Schegolev et al, 2016; Soloviev et al, 2018; Klenov et al, 2018). The main idea of the proposed scheme is shown in the Figure 1a. Here l, l_{out} and l_a are inductances normalized as $l=2\pi LI_C/\Phi_0$, where I_C – critical Josephson junction, Φ_0 – magnetic flux quantum, and all φ like phases normalized ($\varphi=\Phi/\Phi_0$ *et cetera*). The neuron activation function is a nonlinear sigmoid function (Figure 1b), and its part can be used as transfer function of leaky ReLU (red line in Figure 1b), which consists of two linear parts and one nonlinear section. Study of this non-linear part will be devoted to this paper.

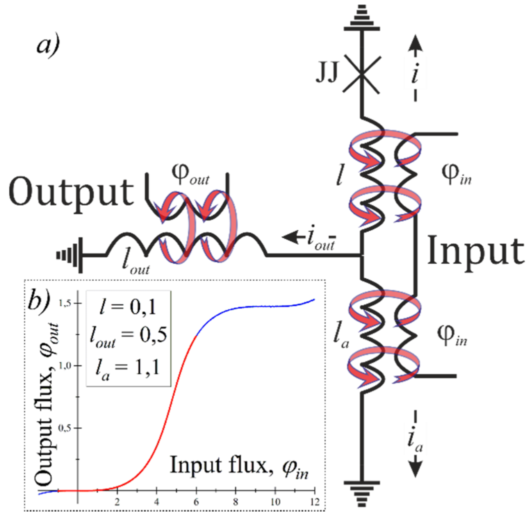

Figure 1: a) Principal scheme of superconducting neuron. b) Transfer function of neuron (blue line) and softmax function or leaky ReLU (red line) for $l=0.1$, $l_{out}=0.5$ and $l_a=1.1$.

3 ReLU TRANSFER FUNCTION

Before we begin to analyze the functioning of superconducting ReLU, it is necessary to determine the operating point of the transfer characteristic function with which we are going to work. To begin with, we should shift the transfer function so that the first linear section falls on the negative values of the external input flux, while the rest – on the positive part thus expected ReLU's characteristic should filtering almost all negative input meanings. For this, it is necessary to apply some additional constant magnetic signal into the input flux, the absolute value of which will shift transfer function of neuron to the left or right depending of the sign.

3.1 Mathematical ReLU

To study the opportunities of ReLU on filtering input signals, we have applied a harmonic, time-dependent signal $s(t)=A\times sin(\omega_0\times t)+shift$, where A – amplitude of the external flux and ω_0 – its frequency, to the input of this element, as shown in Figure 2. Such a choice of circuit's parameters is dictated by the type of transfer characteristic, which has significant linear sections and minimal non-linear transition between them.

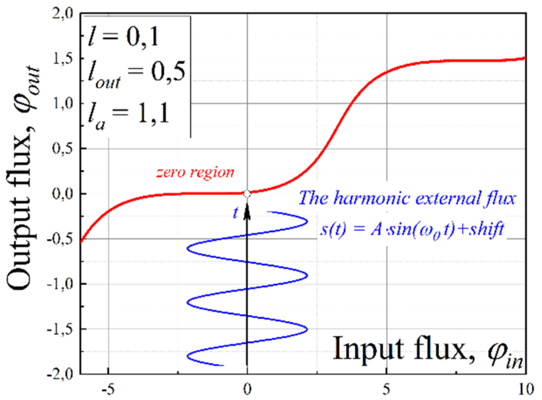

Figure 2: Illustration of applying a simple harmonic signal to the input of ReLU scheme after selection of an operating point at the end of the zero section.

It is in common knowledge that ideal filtering of such signal in the form of $f(x)=max\{0, s(t)\}$ have the following Fourier spectrum:

$$f(\omega)=-\frac{A}{\pi}\cdot\left\{\begin{matrix}\delta(\omega)+i\frac{\pi}{4}\cdot\begin{bmatrix}\delta(\omega_0-\omega)+\\ +\delta(\omega_0+\omega)\end{bmatrix}+\\ \sum_{k=1}^{\infty}\frac{1}{4k^2-1}\cdot\begin{bmatrix}\delta(2k\omega_0-\omega)-\\ -\delta(2k\omega_0+\omega)\end{bmatrix}\end{matrix}\right\}, \quad (1)$$

where $\delta(\omega)$ – is a Dirac delta-function. The filtering signal and its Fourier spectrum showed in the Figure 3a) and b).

It is seen that ReLU skips half the period of the harmonic signal and the spectrum of the output signal has the main and the next even harmonics.

3.2 Real ReLU

3.2.1 Zero Region

The transfer characteristic and spectrum of the “real” ReLU was analyzed for “zero region”, when the average meaning of external signal is equal to zero, and showed in the Figure 4, with amplitude A of external signal equal to 2 (this choice is explained by the requirement to stay within the working range of the transfer characteristic of real ReLU) and frequency ω_0 is equal to 0.01 (for ease of consideration). Also we should note that the value of the shift flux $\varphi_{shift}=0.5\pi$. It is clearly seen that the “real” ReLU filters the external signal a little worse than the mathematical ReLU, however, since the “real” ReLU has a nonlinear transfer characteristic, additional harmonics - odd ones - are present in the output signal spectrum. In addition, since initially the characteristic of the “real” ReLU was taken from the periodic activation function of the neuron, a limitation is placed on the amplitude of the incoming harmonic signal - when a certain value is exceeded, the signal “climbs” beyond the operating range and additional distortions appear in the output signal.

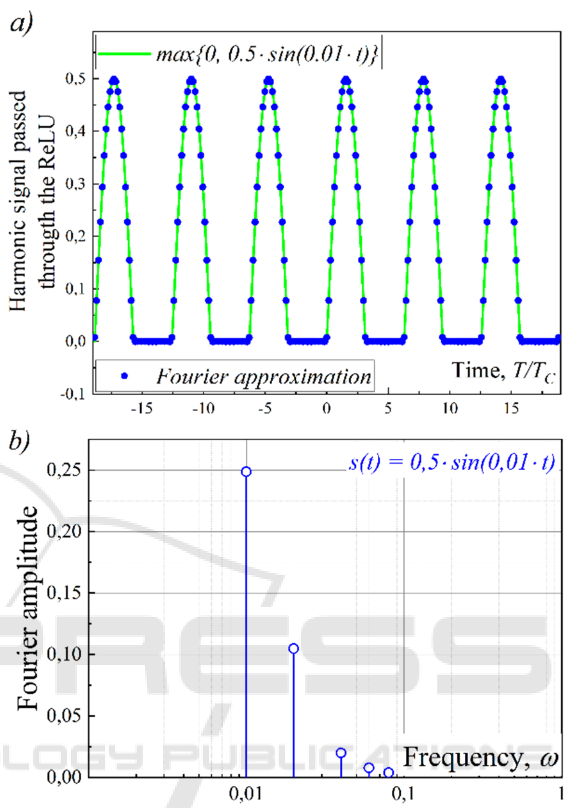

Figure 3: The result of passing through a mathematical ReLU a simple harmonic signal and its approximation using the Fourier series (a) and the Fourier spectrum of the output signal (b).

3.2.2. Linear Region

For completeness of the analysis of the proposed solution for the implementation of ReLU, it is also necessary to evaluate the degree of linearity of the second linear section, for which the operating point will be shifted so that the doubled amplitude of the input signal fits completely within. In this case the value of the shift flux $\varphi_{shift}=1.5\pi$, other parameters of the external signal stay the same (see Figure 5).

The result of transmitting the external harmonic signal through ReLU with an operating point lying on a linear section is shown below on the Figure 6. At the first sight, the signal is passed through without distortion and only multiplied by the corresponding weight of the rectifier. However, Fourier analysis shows that in the spectrum of the output signal, even with small amplitudes of the external signal, higher harmonics are still present and it is obvious that with an increase in the amplitude of the signal, their total contribution to the nonlinearity of the output signal also increases.

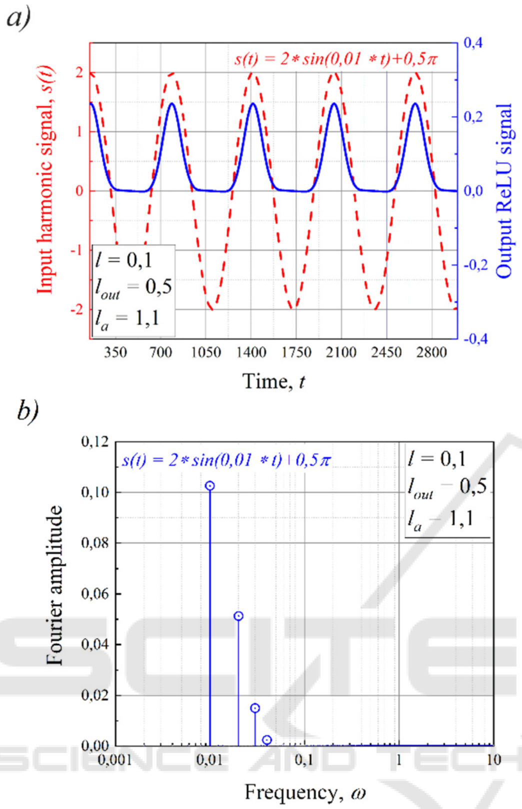

Figure 4: An example of transmitting a harmonic signal through the real ReLU when selecting an operating point at the zero part of the transfer characteristic (a) and the Fourier spectrum of the output signal (b).

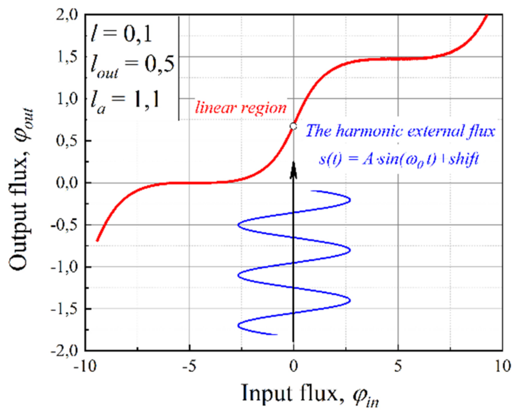

Figure 5: Illustration of applying a simple harmonic signal to the input of ReLU scheme when selecting an operating point in the middle of the linear section.

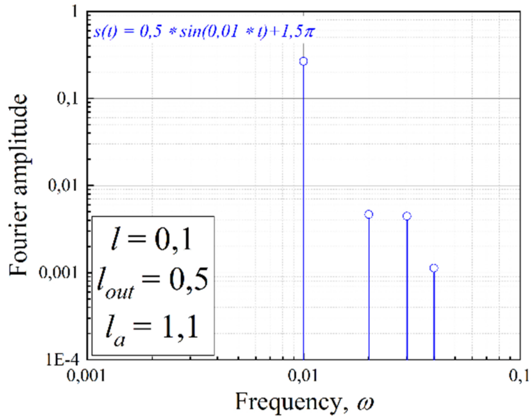

Figure 6: Fourier spectrum of the output signal from real ReLU transfer function during of transmitting a harmonic signal for the case of an operating point at the linear part of the transfer characteristic.

4 CONCLUSIONS

In conclusion, this article was devoted to the inherent element of convolutional neural networks – rectifier linear unit (ReLU) with single-clock “calculation” of transfer function as a non-linear part of hybrid opto-superconducting neural networks. The functionality of this cell was based on the superconducting neuron, investigated earlier (Schegolev et al, 2016; Soloviev et al, 2018), and the parameters of which were selected so that the transfer characteristic could be approximated as accurately as possible by a mathematical form of ReLU over a fairly wide range of changes in the external magnetic flux. Due to the physical features of the implementation of this element, it is not possible to accurately repeat the transfer characteristic of mathematical ReLU, however, it is not so necessary, while leaky ReLU copes with its task. The degree of suitability of the developed element was evaluated using the Fourier analysis apparatus. The numerical simulation methods were used and the spectrum of the output signal from “real” ReLU was obtained. A comparison was performed for the results obtained for mathematical and real ReLU, which showed a good correlation of its transfer characteristics.

How it was mentioned above, the developed cell should be considered from the perspective of using it as a leaky ReLU, performed a role of basic element in a complex of hybrid opto-superconducting neural networks, which does not cut off all the negative values of the input signal. The linear optical part of the computing system should be implemented as a network of waveguides on a chip (Shainline, 2017;

Shainline, 2019). Bidirectional optoelectronic interfaces can be made on the basis of superconducting single-photon detectors and cryogenic n-Trons (Buckley, 2017; Bogatskaya, 2018; Zheng, 2019).

The obtained characteristics of the "real" ReLU give reason to believe that the developed scheme can be suitable for the physical implementation of convolutional opto-superconducting neural networks.

ACKNOWLEDGEMENTS

The analytical study of the proposed concept was supported by RFBR (19-37-90020, 19-02-00981) and President Grant (MD-186.2020.8). Numerical calculations were done with support of Russian Science Foundation (18-72-10118). Also Schegolev is appreciative for the support to the Foundation for the advancement of theoretical physics and mathematics "BASIS".

REFERENCES

Soloviev, I. I., Schegolev, A. E., Klenov, N. V., Bakurskiy, S. V., Kupriyanov, M. Y., Tereshonok, M. V., Golubov, A. A. (2018). Adiabatic superconducting artificial neural network: Basic cells. Journal of applied physics, 124(15), 152113.

Klenov, N. V., Schegolev, A. E., Soloviev, I. I., Bakurskiy, S. V., Tereshonok, M. V. (2018). Energy Efficient Superconducting Neural Networks for High-Speed Intellectual Data Processing Systems. IEEE Transactions on Applied Superconductivity, 28(7), 1-6.

Schegolev, A. E., Klenov, N. V., Soloviev, I. I., Tereshonok, M. V. (2016). Adiabatic superconducting cells for ultra-low-power artificial neural networks. Beilstein journal of nanotechnology, 7(1), 1397-1403.

Szu H. H. (1990). Superconducting neural network computer and sensor array. U.S. Patent documents, US4943556A.

Matsugu, M., Mori, K., Mitari, Y., Kaneda, Y. (2003). Subject independent facial expression recognition with robust face detection using a convolutional neural network. Neural Networks, 16(5-6), 555-559.

Hahnloser, R. H., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J., Seung, H. S. (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. Nature, 405(6789), 947.

Glorot, X., Bordes, A., Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In Proceedings of the fourteenth international conference on artificial intelligence and statistics (pp. 315-323).

Nair, V., Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th international conference on machine learning (ICML-10) (pp. 807-814).

Maas, A. L., Hannun, A. Y., Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In Proc. icml (Vol. 30, No. 1, p. 3).

Shainline, J. M., Buckley, S. M., Mirin, R. P. and Nam S. W. (2017). Superconducting Optoelectronic Circuits for Neuromorphic Computing. Phys. Rev. Applied 7, 034013.

Shainline, J. M., Buckley, S. M., Nader, N., Gentry, C. M., Cossel, K. C., Cleary, J. W., Popović, M., Newbury, N. R., Nam, S. W. and Mirin, R. P. (2017). Room-temperature-deposited dielectrics and superconductors for integrated photonics. Optics Express 25(9), pp. 10322-10334.

Shainline, J. M., Buckley, S. M., McCaughan, A. N., Chiles, J. T., Salim, A. J., Beltran, M. C., Donnelly, C. A., Schneider, M. L., Mirin, R. P. and Nam S. W. (2019). Superconducting optoelectronic loop neurons. Journal of Applied Physics 126, 044902.

Buckley, S., Chiles, J., McCaughan, A. N., Moody, G., Silverman, K. L., Stevens, M. J., Mirin, R. P., Nam, S. W. and Shainline J. M. (2017). All-silicon light-emitting diodes waveguide-integrated with superconducting single-photon detectors. Appl. Phys. Lett. 111, 141101.

Bogatskaya, A. V., Klenov, N. V., Popov, A. M. and Tereshonok, M. V. (2018). Resonance tunneling of electromagnetic waves for enhancing the efficiency of bolometric photodetectors. Technical Physics Letters, 44(8):667–670.

Zheng, K., Zhao, Q.-Y., Kong, L.-D., Chen, S., Lu, H.-Y.-Bo, Tu, X.-C., Zhang, L.-B., Jia, X.-Q., Chen, J., Kang L. and Wu, P.-H. (2019). Characterize the switching performance of a superconducting nanowire cryotron for reading superconducting nanowire single photon detectors. Scientific Reports volume 9, Article number: 16345.