Optica Applicata, Vol. LIV, No. 1, 2024
DOI: 10.37190/oa240101

Design of a photonic unitary neural network based on MZI arrays

Ye Zhang[1,*], Ruiting Wang[2,3], Yejin Zhang[2,3], Yanmei Su[2], Pengfei Wang[2,3], Guangzhen Luo[2,3], Xuliang Zhou[2,3], Jiaoqing Pan[2,3,*]

[1] School of Automation, Beijing Information Science and Technology University, Beijing, China

[2] Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

[3] College of Materials Science and Opto-Electronic Technology, University of Chinese Academy of Sciences, Beijing, China

[*] Correspondening authors: zhangyethu@163.com (Y.Z.); jqpan@semi.ac.cn (J.P.)

In recent years, optical neural networks have attracted widespread attention, due to their advantages of high speed, high parallelism, high bandwidth, and low power consumption. Photonic unitary neural network is a kind of neural networks that utilize the principles of unitary matrices and photonics to perform computations. In this paper, we design a photonic unitary neural network based on Mach–Zehnder interferometer arrays. The results show that the network has a good performance on both triangular and circular binary classification datasets, where most of the data points are correctly classified. The accuracies achieve 97% and 95% for triangular and circular datasets, with the loss function values of 0.023 and 0.046, respectively.

Keywords: optical neural network, unitary matrix, Mach–Zehnder interferometer.

1. Introduction

The explosive growth of data in the information age, the innovation of algorithms, and the improvement of computing power have promoted the rapid development of artificial intelligence (AI) technology. In recent years, AI technology has been widely used in face recognition [1], speech recognition [2], natural language processing [3] and related other fields. The rise of smart cities, intelligent manufacturing, intelligent medical care, intelligent education and other fields indicates that people are gradually entering a more intelligent era. Neural network is an important computing model in the field of AI, which greatly promotes the development of AI technology. Neural networks are a type of machine learning algorithm that are designed to mimic the way the human brain works [4]. They consist of multiple layers of interconnected nodes that can learn to recognize patterns in data, and can be trained to perform a wide range of tasks, including image and speech recognition [2,5], natural language processing [3], and decision-making [6]. The development of neural networks has been a key driver of the advancement of AI technology, because they have enabled the creation of more powerful and sophisticated machine learning systems.

The development of neural network models is inseparable from the computing power support of chips. However, with the slowdown of the development of Moore's law [7], the development of electronic chips has also encountered serious bottlenecks [8], which has been difficult to meet the demand for computing power of large-scale neural networks. Besides, most electric chips adopt the von Neumann architecture, which extracts data from the memory to the processor during calculation, and then writes the data back to the memory after processing. When the computing power of the chip reaches a certain level, the access speed of the memory is much smaller than the computing speed of the processor, forming a bottleneck of the von Neumann architecture [9,10], which could affect the performance of the chip seriously. Thus, new archtecures of computing chips are urgently needed.

Among numerous new computing chip solutions, optical neural network (ONN) chips have attracted widespread attention from the academic community, due to their advantages of high speed, high parallelism, high bandwidth, and low power consumption [11-13]. ONNs are an emerging class of computing architectures that use light to perform complex computations. In recent years, there has been a growing interest in developing ONNs for various machine learning tasks. One approach to implementing ONNs is to use integrated photonics platforms, such as silicon photonics. These platforms can be used to build compact and scalable ONNs that can perform complex computations using light. Several recent studies have demonstrated the use of ONNs for various machine learning tasks, such as image classification [14-16] and speech recognition [17].

Photonic unitary neural networks are a class of neural networks that utilize the principles of unitary matrices and photonics to perform computations. In photonic unitary neural networks, the neural network weights are implemented as optical components that operate on optical signals, resulting in fast and efficient computation [18].

In this paper, we design a photonic unitary neural network based on Mach–Zehnder interferometer (MZI) arrays. The performance of the designed photonic unitary neural network is evaluated through a binary classification problem.

2. Architecture

The photonic unitary neural networks are special neural networks whose weight matrix is a unitary matrix. One of the key advantages of photonic unitary neural networks is their ability to perform complex computations in parallel, with low power consumption and high accuracy. This makes them well-suited for applications in fields such as image and signal processing, pattern recognition, and machine learning. Photonic unitary neural networks represent a promising direction for the development of efficient and high-performance computing systems, with potential applications in a wide range of fields.

MZI is a type of optical device that utilizes the interference of light waves to perform various functions, such as filtering [19], modulation [20], and sensing [21]. It consists of two arms with a common input and output port [22]. The light entering the input port is split into two paths, which recombine at the output port. The relative phase difference between the two paths determines the interference pattern at the output port, which can be used to extract information about the input signal.

In optical neural networks, MZIs are commonly used as basic building blocks for performing matrix-vector multiplication [17]. By encoding the input signals as the phase and amplitude of optical waves, and properly designing the optical path lengths and coupling coefficients of the MZIs, the network can perform the necessary computations to achieve the desired neural network functionality [17].

One of the key advantages of using MZIs in optical neural networks is their inherent parallelism [23,24]. Many MZIs can be operated simultaneously, allowing for the processing of large amounts of data in parallel [23,24]. Additionally, MZIs can operate at very high speeds [25], which makes them particularly useful for applications that require real-time processing.

The basic computing unit of the photonic unitary neural network we designed is MZI computing unit, as shown in Fig. 1. Each unit consists of two 50 : 50 beam splitters and two tunable single-mode phase shifters, with phase shift value θ and φ, respectively.

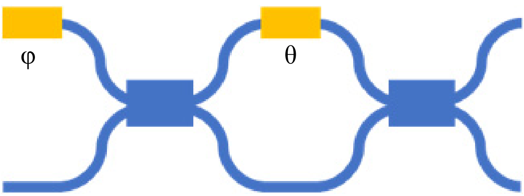

Fig. 1. MZI computing unit.

The MZI computing unit can realize unitary operation as follows [18,26,27]:

$$U(\theta, \varphi) = \begin{bmatrix} \exp(i\varphi)\sin\dfrac{\theta}{2} & \cos\dfrac{\theta}{2} \\ \exp(i\varphi)\cos\dfrac{\theta}{2} & -\sin\dfrac{\theta}{2} \end{bmatrix} \tag{1}$$

In this paper, based on the MZI unit, we design a photonic unitary neural network using cascaded MZI arrays, as shown in Fig. 2. A single-layer photonic unitary neural network is shown in the figure, where the unitary weight matrix and the input of the

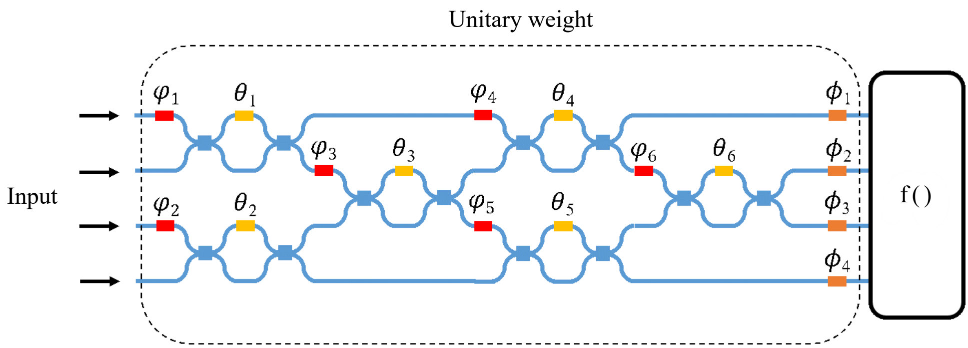

Fig. 2. Photonic unitary neural network.

neural network are multiplied and added, followed by a nonlinear function calculation.

Figure 2 shows a designed 4×4 photonic unitary neural network, with six cascaded MZIs and four phase shifters. The cascaded MZI arrays could implement unitary transformation matrix U_4, while the phase shifters could implement a diagonal matrix Σ. Thus the structure could realize a 4×4 linear transformation matrix of

$$W_{4\times 4} = \Sigma U_4 = \begin{bmatrix} u_1 & 0 & 0 & 0 \\ 0 & u_2 & 0 & 0 \\ 0 & 0 & u_3 & 0 \\ 0 & 0 & 0 & u_4 \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} & U_{13} & U_{14} \\ U_{21} & U_{22} & U_{23} & U_{24} \\ U_{31} & U_{32} & U_{33} & U_{34} \\ U_{41} & U_{42} & U_{43} & U_{44} \end{bmatrix} \tag{2}$$

3. Results

The photonic unitary neural network we designed is of simple structure, which uses a few amount of MZI arrays and could be fabricated easily. Thus it can effectively reduce architecture power consumption and complexity. In order to evaluate the performance of the photonic unitary neural network, we use it for a classification problem of specific datasets.

Two binary classification datasets are constructed for the training process, as shown in Fig. 3. Figure 3(a) shows a linearly separable triangular binary dataset, consisting of 1000 data points distributed on both sides of $y = -x$. Points with coordinates $y < -x$ are labeled as “0”, while points at coordinates $y > -x$ are labeled as “1”. In order to increase the complexity of the dataset, a small number of noise points are added. For example, points whose coordinates match $y < -x$, but have a label of “1”. Figure 3(b) is a linearly indivisible circular binary dataset with data points distributed on either side of $x^2 + y^2 = 0.5$. Points with coordinates $x^2 + y^2 < 0.5$ are labeled as “0”, while points at coordinates $x^2 + y^2 > 0.5$ are labeled as “1”. The circular binary dataset also consists of 1000 data points and small amount of noise points. In Figure 3, data points with label “0” are marked in red, while that with label “1” are marked in purple. By

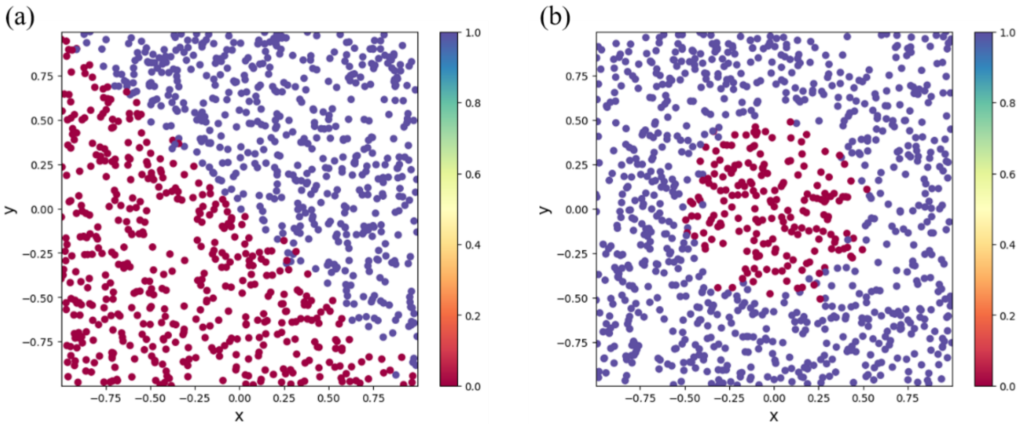

Fig. 3. (a) Triangular and (b) circular binary classification datasets.

using these two datasets, we can demonstrate the ability of the photonic unitary neural network to handle both linear and non-linear classification problems.

We constructed a three-layer fully connected neural network based on the chip-level photonic computing framework Neuroptica [28] to classify triangular and circular binary datasets. Each layer of the photonic unitary network contains four neurons, as shown in Fig. 2. The nonlinear function of the hidden layer is the electro-optic activation function, and the electro-optic activation function of the output layer is replaced by the detection layer to simulate the optical power measurement of the photodetector. The photonic unitary neural network is trained using the mean squared error (MSE) loss function [29] and the adaptive moment estimation (Adam) [30] algorithm. Accuracy is commonly used to evaluate ONNs, since it is a straightforward and easy-to-understand metric, which measures the percentage of correctly classified samples out of the total number of samples.

The training processes and the classification results of the designed photonic unitary neural network on two binary classification datasets are shown in the following two figures, respecitively, where Fig. 4 shows the results on the triangular binary dataset, while Fig. 5 shows them on the circular binary dataset. As shown in Fig. 4(a), after 100 epochs, the accuracy of the triangular binary classification dataset reaches 97%, with the loss function value of 0.023. Figure 4(b) shows the classification results on the same triangular binary classification dataset, where the red and purple areas represent the classification results with label “0” and “1”, respectively, and the yellow line stands for the decision boundary. From the picture we can see that most of the data points are correctly classified.

While Fig. 5 shows the training processes and the classification results of the designed photonic unitary neural network on the circular binary dataset, Fig. 5 (a) shows the training processes with accuracy and loss function. As we can see from Fig. 5 (a), after 100 epochs, the photonic unitary neural network achieves 95% accuracy, and the loss function value is 0.046. Figures 5 (b) show the classification result. It can be seen

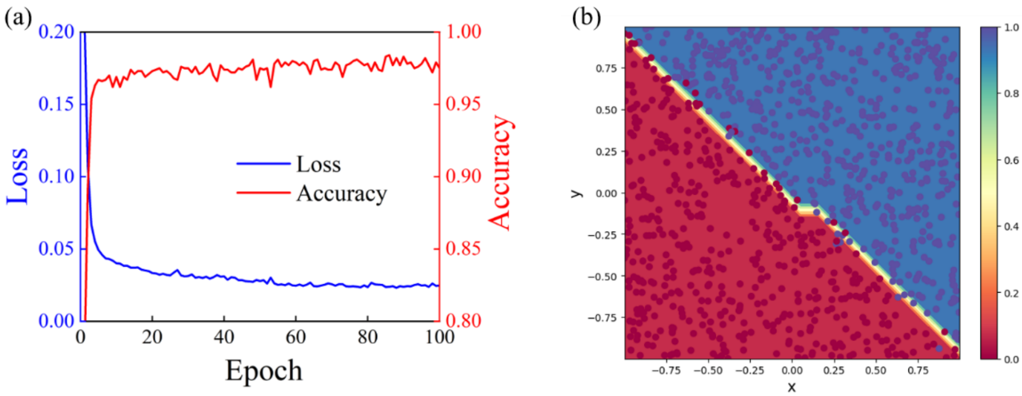

Fig. 4. (a) Training process and (b) classification results of photonic unitary neural network on the triangular dataset.

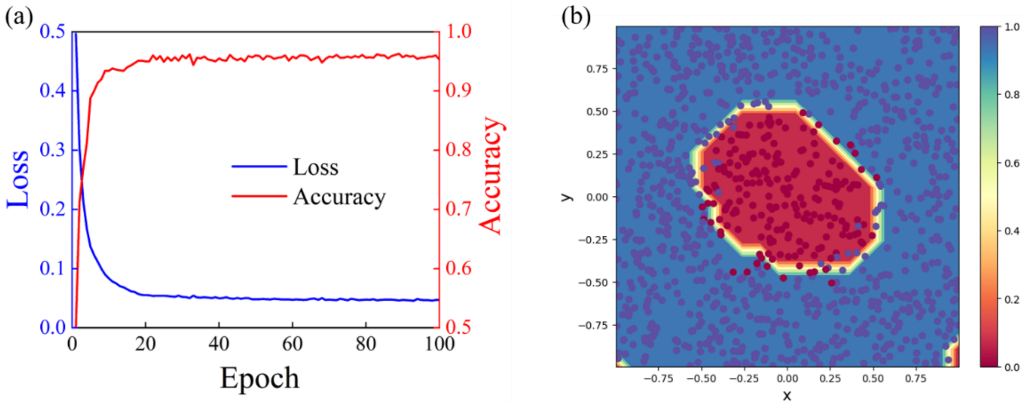

Fig. 5. (a) Training process and (b) classification results of photonic unitary neural network on the circular dataset.

from the figure that the correct classification results have been achieved for most of the data points.

The significant advantages of ONNs compared to traditional architectures lie in their high speed and low power consumption. The ONN system could run at 100 GHz rate, which is two orders of magnitude faster than electronic system. The floating point operations (FLOPs) per second of our ONN could achieve $R = 9.6 \times 10^{12}$ FLOPs. Meanwhile, the power consumption is primarily determined by the power required to activate an optical nonlinearity and photodetectors. The typical energy consumption of an ONN chip reaches femtojoule per FLOP [17], which is three to five orders of magnitude [31] better than conventional computers. Overall, the excellent speed and power advantages of ONNs make them promising candidates for accelerating artificial intelligence and machine learning tasks.

4. Conclusion

In conclusion, we design a photonic unitary neural network based on MZI arrays. The basic computing unit is composed of MZI computing unit, which can realize unitary operation. In order to evaluate the performance of the photonic unitary neural network, two binary classification datasets are used. Both triangular and circular datasets consist of 1000 data points with a small number of noise points. Neuroptica computing framework is used for the experiment, while MSE loss function and Adam algorithm are adopted for the training process. The experimental results show that after 100 epochs, the proposed photonic unitary neural network achieves 97% and 95% accuracy on triangular and circular datasets, respectively, with the loss function values of 0.023 and 0.046, respectively. Meanwhile, correct classification results have been achieved for most of the data points, which indicate the good performance of the network.

Acknowledgements
This work is supported by the Beijing Natural Science Foundation (Z200006), the National Natural Science Foundation of China (62090053, 61934007, 61974141), the Beijing Municipal Science and Technology Project (Z191100004819011), the National Key R&D Program of China (2018YFE0203103), and the Supplementary and Supportive Project for Teachers at the Beijing Information Science and Technology University (5029011103).

References

[1] Zhao W., Chellappa R., Phillips P.J., Rosenfeld A., *Face recognition: A literature survey*, ACM Computing Surveys **35**(4), 2003: 399-458. https://doi.org/10.1145/954339.954342

[2] Wang D., Chen J., *Supervised speech separation based on deep learning: An overview*, IEEE/ACM Transactions on Audio, Speech, and Language Processing **26**(10), 2018: 1702-1726. https://doi.org/10.1109/TASLP.2018.2842159

[3] Young T., Hazarika D., Poria S., Cambria E., *Recent trends in deep learning based natural language processing*, IEEE Computational Intelligence Magazine **13**(3), 2018: 55-75. https://doi.org/10.1109/MCI.2018.2840738

[4] Bishop C.M., *Neural networks and their applications*, Review of Scientific Instruments **65**(6), 1994: 1803-1832. https://doi.org/10.1063/1.1144830

[5] Krizhevsky A., Sutskever I., Hinton G.E., *ImageNet classification with deep convolutional neural networks*, Communications of the ACM **60**(6), 2017: 84-90. https://doi.org/10.1145/3065386

[6] Hill T., Marquez L., O'Connor M., Remus W., *Artificial neural network models for forecasting and decision making*, International Journal of Forecasting **10**(1), 1994: 5-15. https://doi.org/10.1016/0169-2070(94)90045-0

[7] Khan H.N., Hounshell D.A., Fuchs E.R.H., *Science and research policy at the end of Moore's law*, Nature Electronics **1**(1), 2018: 14-21. https://doi.org/10.1038/s41928-017-0005-9

[8] Williams R.S., *What's Next? [The end of Moore's law]*, Computing in Science & Engineering **19**(2), 2017: 7-13. https://doi.org/10.1109/MCSE.2017.31

[9] Ielmini D., Wong H.-S.P., *In-memory computing with resistive switching devices*, Nature Electronics **1**(6), 2018: 333-343. https://doi.org/10.1038/s41928-018-0092-2

[10] Sebastian A., Le Gallo M., Khaddam-Aljameh R., Eleftheriou E., *Memory devices and applications for in-memory computing*, Nature Nanotechnology **15**(7), 2020: 529-544. https://doi.org/10.1038/s41565-020-0655-z

[11] HAMERLY R., BERNSTEIN L., SLUDDS A., SOLJAČIĆ M., ENGLUND D., *Large-scale optical neural networks based on photoelectric multiplication*, Physical Review X **9**(2), 2019: 021032. https://doi.org/10.1103/PhysRevX.9.021032

[12] SHASTRI B.J., TAIT A.N., FERREIRA DE LIMA T., PERNICE W.H.P., BHASKARAN H., WRIGHT C.D., PRUCNAL P.R., *Photonics for artificial intelligence and neuromorphic computing*, Nature Photonics **15**(2), 2021: 102-114. https://doi.org/10.1038/s41566-020-00754-y

[13] WU J., LIN X., GUO Y., LIU J., FANG L., JIAO S., DAI Q., *Analog optical computing for artificial intelligence*, Engineering **10**, 2022: 133-145. https://doi.org/10.1016/j.eng.2021.06.021

[14] ASHTIANI F., GEERS A.J., AFLATOUNI F., *An on-chip photonic deep neural network for image classification*, Nature **606**(7914), 2022: 501-506. https://doi.org/10.1038/s41586-022-04714-0

[15] FU T., ZANG Y., HUANG Y., DU Z., HUANG H., HU C., CHEN M., YANG S., CHEN H., *Photonic machine learning with on-chip diffractive optics*, Nature Communications **14**(1), 2023: 70. https://doi.org/10.1038/s41467-022-35772-7

[16] ZHANG H., GU M., JIANG X.D., THOMPSON J., CAI H., PAESANI S., SANTAGATI R., LAING A., ZHANG Y., YUNG M.H., SHI Y.Z., MUHAMMAD F.K., LO G.Q., LUO X.S., DONG B., KWONG D.L., KWEK L.C., LIU A.Q., *An optical neural chip for implementing complex-valued neural network*, Nature Communications **12**(1), 2021: 457. https://doi.org/10.1038/s41467-020-20719-7

[17] SHEN Y., HARRIS N. C., SKIRLO S., PRABHU M., BAEHR-JONES T., HOCHBERG M., SUN X., ZHAO S., LAROCHELLE H., ENGLUND D., SOLJAČIĆ M., *Deep learning with coherent nanophotonic circuits*, Nature Photonics **11**(7), 2017: 441-446. https://doi.org/10.1038/nphoton.2017.93

[18] PAI S., BARTLETT B., SOLGAARD O., MILLER D.A.B., *Matrix optimization on universal unitary photonic devices*, Physical Review Applied **11**(6), 2019: 064044. https://doi.org/10.1103/PhysRevApplied.11.064044

[19] DING Y., PU M., LIU L., XU J., PEUCHERET C., ZHANG X., HUANG D., OU H., *Bandwidth and wavelength-tunable optical bandpass filter based on silicon microring-MZI structure*, Optics Express **19**(7), 2011: 6462-6470. https://doi.org/10.1364/OE.19.006462

[20] PALMER R., ALLOATTI L., KORN D., SCHINDLER P. C., SCHMOGROW R., HENI W., KOENIG S., BOLTEN J., WAHLBRINK T., WALDOW M., YU H., BOGAERTS W., VERHEYEN P., LEPAGE G., PANTOUVAKI M., VAN CAMPENHOUT J., ABSIL P., DINU R., FREUDE W., KOOS C., LEUTHOLD J., *Silicon-organic hybrid MZI modulator generating OOK, BPSK and 8-ASK signals for up to 84 Gbit/s*, IEEE Photonics Journal **5**(2), 2013: 6600907. https://doi.org/10.1109/JPHOT.2013.2258142

[21] MUMTAZ F., DAI Y., ASHRAF M.A., *Inter-cross de-modulated refractive index and temperature sensor by an etched multi-core fiber of a MZI structure*, Journal of Lightwave Technology **38**(24), 2020: 6948-6953. https://doi.org/10.1109/JLT.2020.3014857

[22] PRIETO F., SEPÚLVEDA B., CALLE A., LLOBERA A., DOMÍNGUEZ C., LECHUGA L.M., *Integrated Mach–Zehnder interferometer based on ARROW structures for biosensor applications*, Sensors and Actuators B: Chemical **92**(1-2), 2003: 151-158. https://doi.org/10.1016/S0925-4005(03)00257-0

[23] KIM J.-Y., KANG J.-M., KIM T.-Y., HAN S.-K., *All-optical multiple logic gates with XOR, NOR, OR, and NAND functions using parallel SOA-MZI structures: Theory and experiment*, Journal of Lightwave Technology **24**(9), 2006: 3392-3399. https://doi.org/10.1109/JLT.2006.880593

[24] MISHINA K., NISSANKA S. M., MARUTA A., MITANI S., ISHIDA K., SHIMIZU K., HATTA T., KITAYAMA K., *All-optical modulation format conversion from NRZ-OOK to RZ-QPSK using parallel SOA-MZI OOK/BPSK converters*, Optics Express **15**(12), 2007: 7774-7785. https://doi.org/10.1364/OE.15.007774

[25] SINGH S., LOVKESH, *Ultrahigh speed optical signal processing logic based on an SOA-MZI*, IEEE Journal of Selected Topics in Quantum Electronics **18**(2), 2012: 970-977. https://doi.org/10.1109/JSTQE.2011.2155623

[26] SPENGLER C., HUBER M., HIESMAYR B.C., *Composite parameterization and Haar measure for all unitary and special unitary groups*, Journal of Mathematical Physics **53**(1), 2012: 013501. https://doi.org/10.1063/1.3672064

[27] MILLER D.A.B., *Perfect optics with imperfect components*, Optica **2**(8), 2015: 747-750. https://doi.org/10.1364/OPTICA.2.000747

[28] WILLIAMSON I.A.D., HUGHES T.W., MINKOV M., BARTLETT B., PAI S., FAN S., *Reprogrammable electro-optic nonlinear activation functions for optical neural networks*, IEEE Journal of Selected Topics in Quantum Electronics **26**(1), 2020: 7700412. https://doi.org/10.1109/JSTQE.2019.2930455

[29] KÖKSOY O., YALCINOZ T., *Mean square error criteria to multiresponse process optimization by a new genetic algorithm*, Applied Mathematics and Computation **175**(2), 2006: 1657-1674. https://doi.org/10.1016/j.amc.2005.09.011

[30] ZHOU J., WANG H., WEI J., LIU L., HUANG X., GAO S., LIU W., LI J., YU C., LI Z., *Adaptive moment estimation for polynomial nonlinear equalizer in PAM8-based optical interconnects*, Optics Express **27**(22), 2019: 32210-32216. https://doi.org/10.1364/OE.27.032210

[31] HOROWITZ M., *Computing's energy problem (and what we can do about it)*, [In] *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, San Francisco, CA, USA, 2014: 10-14. https://doi.org/10.1109/ISSCC.2014.6757323

Received April 25, 2023
in revised form June 6, 2023