ISSN 0021-3640, JETP Letters, 2021, Vol. 114, No. 6, pp. 321–325. © Pleiades Publishing, Inc., 2021.
Russian Text © The Author(s), 2021, published in Pis'ma v Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki, 2021, Vol. 114, No. 6, pp. 360–364.

OPTICS
AND LASER PHYSICS

Optimization of Multilayer Photonic Structures using Artificial Neural Networks to Obtain a Target Optical Response

K. R. Safronov[a], *, V. O. Bessonov[a], and A. A. Fedyanin[a]

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

**e-mail: safronov@nanolab.phys.msu.ru*

Received August 11, 2021; revised August 12, 2021; accepted August 12, 2021

A new deep machine learning method is proposed for the task of selecting the parameters of a multilayer photonic structure to obtain a target optical spectrum of the reflection coefficient. The proposed training method is based on the connection of an artificial neural network for solving the inverse problem and the analytical transfer matrix method. This approach allows achieving high accuracy of the network. The developed method can be applied to the design of a structure that takes the derivative of the coordinate for an incident optical signal.

DOI: 10.1134/S0021364021180119

Nanophotonic devices provide complex functionalities through structuring materials at the micro- and nanoscale [1–3]. One of the simplest, yet versatile nanophotonic structures is an array of alternating layers with different thicknesses and refractive indices. The optical response of such a multilayer structure (MS) is controlled by the parameters of the layers and can be adjusted to the requirements of a specific task. Due to this flexibility, MSs are successfully used in such areas as topological photonics [4], nonlinear optics [5, 6], magneto-optics [7], optical computing [8], polaritonics [9], etc. [10–12]. However, realization of all these possibilities requires developing methods for selecting the parameters (design) of MS with a target optical response.

The MS can be designed by iterating through all possible configurations and calculating the optical response for each set of parameters. Such a calculation can be carried out by one of many well-developed methods, including the finite element method, the finite difference in the time domain method, the transfer matrix method (TMM), etc. The disadvantage of this approach is its large time and computational costs. Another approach is based on solving the inverse problem, i.e., determining the MS parameters directly from a given response [13]. Many optimization methods have been developed to solve the inverse problem [14–17]. However, they are iterative and therefore computationally expensive, which makes them unsuitable for large-scale and complex structures. In addition, even a minor modification of the target optical response forces to start over the optimization procedure.

To overcome this obstacle, it was recently proposed to use an approach based on artificial neural networks (ANNs) [18, 19]. The ANN can be used in one of two ways. The first option is to predict the optical response of the structure with particular geometric parameters [19–21]. In this case, the ANN replaces expensive computational modeling and increases the speed of standard optimization methods. The second option is to train the ANN to predict the MS with the required response [22, 23], i.e., to solve the inverse problem. This allows designing structures with a given optical response in a fraction of a second. The main problem of the second way of using the ANN is the nonuniqueness of the inverse problem solution, which leads to data inconsistency and poor convergence of the ANN [18]. Several approaches have been proposed to solve this problem. One of the most common of them is two-stage ANN training: first, one network is trained to predict the optical response of MS, and then this pre-trained ANN is connected to another network solving the inverse problem to form a tandem of networks [18, 23]. In this case, the ANN is not forced to any specific solution of the inverse problem and the network chooses one of the many solutions that leads to better convergence. Another notable approach is based on generative adversarial networks. In this case, two ANNs (generator and discriminator) are trained simultaneously [24]. The third approach uses reducing the dimension of the parameter space using an autoencoder [25]. It is assumed that the inverse problem has a unique solution in a space of smaller dimension. One common problem in all the proposed methods is

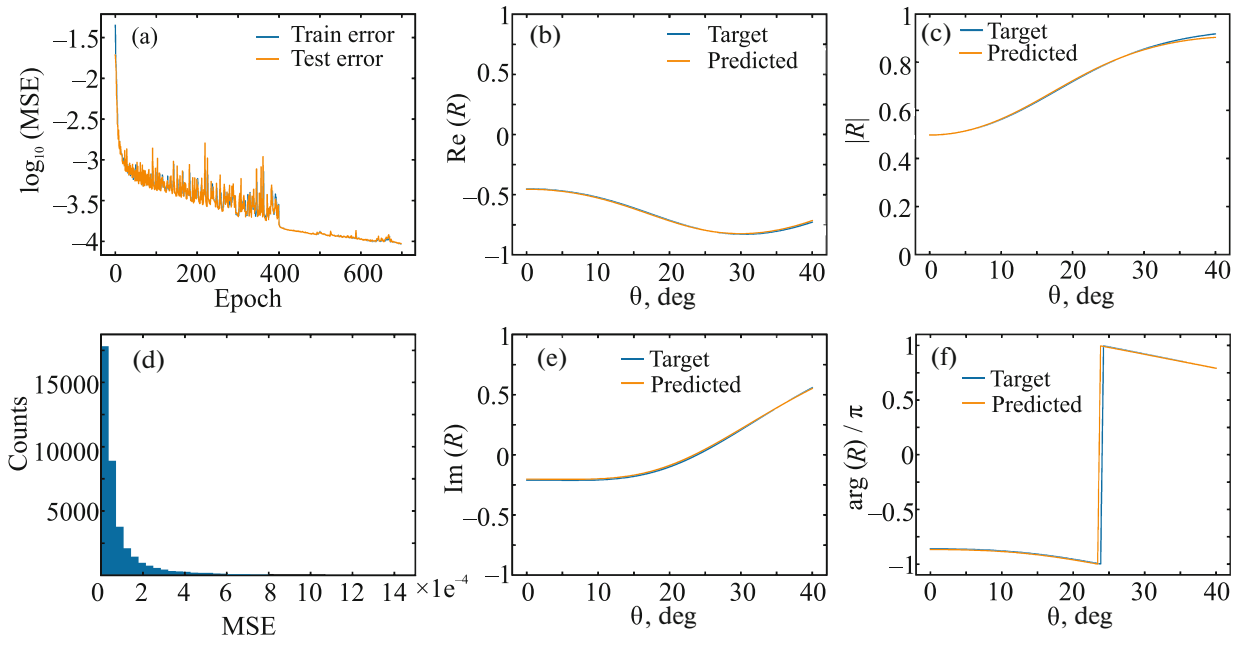

Fig. 1. (Color online) (a) Decimal logarithm of the rms error (MSE) for the (orange line) training and (blue line) test datasets versus the epoch number. (b) Real part, (c) amplitude, (e) imaginary part, and (f) phase divided by π of the reflection coefficient versus the angle of incidence. (d) Histogram of the distribution of rms error over the test dataset. The target reflection coefficient is shown by the blue line; the reflection coefficient of the multilayer structure designed by the artificial neural network is shown by the orange line.

the need to train several ANNs, which inevitably leads to additional errors.

In this paper, an alternative approach is proposed that uses only one ANN to solve the inverse problem of the MS design, which can significantly improve the accuracy of the design. It is shown that the trained ANN successfully bypasses the problem of non-uniqueness of the inverse problem solution. This approach can be used to determine the parameters of various nanophotonic devices.

The paper considers an MS consisting of 10 alternating layers of SiO_2 ($n = 1.45$)/Ta_2O_5 ($n = 2.08$) with a thickness from 0 to $d_{max} = 350$ nm, deposited on a glass substrate ($n = 1.52$). The goal is to determine the thicknesses of the layers d_i ($i = 1...10$) of MS with a given optical response. The input data for the ANN is the angular spectrum of the amplitude reflection coefficient for TE-polarized radiation at $\lambda = 800$ nm in the range of incidence angles θ from 0° to 40°. The spectrum is sampled by 100 points. The input of the ANN receives an array consisting of the real r_i' and imaginary r_i'' parts of the reflection coefficient for each angle of incidence; i.e., the array has a length of 200. The real and imaginary parts of the reflection spectrum were chosen instead of the phase and amplitude, since phase can have discontinuities that limit the accuracy of the ANN [26]. The output data of the ANN is an array D of 10 normalized MS thicknesses d_i/d_{max}. The material of each layer is considered known and fixed. The layer adjacent to air is made of SiO_2.

The TMM was used to create a dataset containing the reflection coefficient spectra of randomly generated 10-layer structures. This ensures that the target spectra can be achieved using the one-dimensional MS considered in this paper. Then the dataset is divided into a training part (used for ANN training) and a test part (to check the ANN accuracy). Our method of ANN training is based on connecting the network with the TMM [27]. During the training process, the target angular spectrum $R_{target}(\theta)$ is used as the input of the ANN, and the array of normalized thicknesses D is the output of the ANN. Then D is used in the TMM to calculate the angular spectrum $R_{predicted}(\theta)$ of the MS generated by the ANN. Then the rms error (MSE) between the spectra $R_{predicted}$ and R_{target} is calculated. After that, the weights and biases of the ANN layers are updated using backpropagation of the error to minimize the MSE. In order to TMM not to become a bottleneck in the training of ANN, an optimized code was developed in which TMM is implemented by multiplying multidimensional matrices (tensors). TMM naturally fits into the process of ANN training since it is based on matrix multiplication, which is the main operation of the deep learning. In addition, the TMM states the analytical relationship between the layer thickness and the reflection

spectrum and can be used for backpropagation of the error, i.e., calculating the MSE gradient relative to the weights and biases of the ANN. In the described training process, the MS parameters from the training dataset are not used in any way. In this case, the ANN learns the patterns inherent in the TMM instead of simply interpolating data from the training dataset.

The ANN training method was tested on the example of a network containing four hidden layers of 500 neurons each. The ANN was trained on 100000 samples from the training dataset for 700 epochs. The MSE sharply drops after 10 epochs, and then gradually decreases to 10^{-4} (Fig. 1a). The distribution of MSE in the test dataset is localized near zero, which proves a good quality of training (Fig. 1d). Figures 1b and 1e demonstrate the operation of our network using a random example from a test dataset. Our network successfully solve inverse problem. Figures 1c and 1f demonstrate the amplitude $|R|$ and phase $\arg(R)$ of the reflection coefficient, which are calculated from the real r_i' and imaginary r_i'' parts of the reflection coefficient using the formulas:

$$|R_i| = \sqrt{(r_i')^2 + (r_i'')^2}, \quad (1)$$

$$\arg(R_i) = \arctan(r_i''/r_i'), \quad (2)$$

where $i = 1...100$. Since $\arg(R)$ lies in the range from $-\pi$ to π, a jump of 2π in the reflection coefficient spectrum can be observed (Fig. 1f).

An important question is how the developed ANN deals with the problem of non-uniqueness of the inverse problem solution. Figures 2a and 2c show two similar spectra from completely different MS from the test dataset (Fig. 2b). The average deviation of the layer thicknesses, determined by the formula:

$$\Delta d = \sum_{i=1}^{N} \frac{d_i^{\text{target}_1} - d_i^{\text{target}_2}}{N},$$

where $N = 10$, was used as a measure of the similarity of MSs. For selected samples, $\Delta d = 96$ nm. Figures 2a and 2c demonstrate that ANN successfully designs MS with the target spectra. However, the average difference in the layer thicknesses Δd for the generated MSs is only 3 nm (Fig. 2d). Next, the correlation between the difference in spectra ΔR and the difference in layer thicknesses Δd was estimated for samples from the test dataset and for MSs generated by the ANN. The Pearson coefficient $\rho_{(\Delta R,\Delta d)}$ was used as a measure of correlation. For the test dataset $\rho^{\text{test}}_{(\Delta R,\Delta d)} = -0.05$, and for the MSs designed by the ANN $\rho^{NN}_{(\Delta R,\Delta d)} = 0.48$. It can be seen that in the test dataset there is no correlation between the difference in spectra and the difference in thicknesses; i.e., a similar optical response can be demonstrated by MSs with completely different parameters. However, there is a

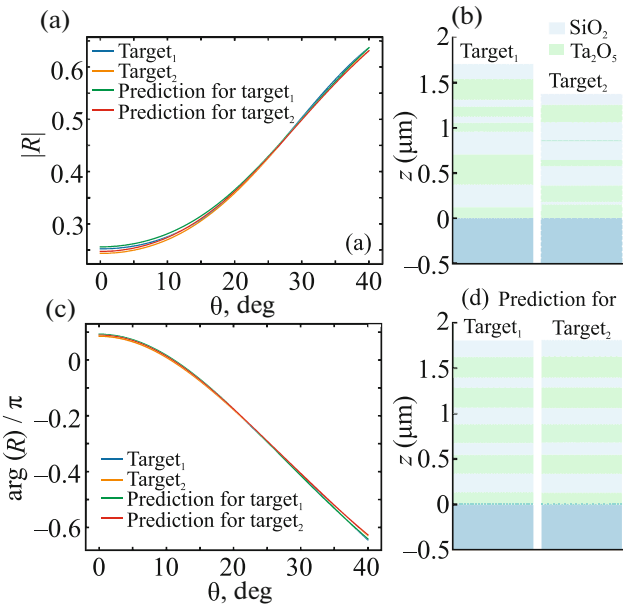

Fig. 2. (Color online) (a) Amplitude and (c) phase divided by π of the reflection coefficient versus the angle of incidence for two similar spectra from the test dataset (blue and orange lines) and the spectra of multilayer structures designed by the artificial neural network (green and red lines). (b, d) Multilayer structures with close spectra from the (b) test dataset and (d) designed by the artificial neural network.

fairly strong positive correlation between the difference in spectra and the difference in thicknesses for samples designed by the network. This means that for these MSs, similar spectra are produced by similar structures. Thus, during training, the ANN smooths the parameter space and converges to one of the many possible solutions of the inverse problem.

As an example application of our network, we design MS, which is an optical spatial differentiator. In order to implement a first-order spatial derivative, such a device must have a transfer function (in our case, the reflection coefficient) $R \approx ik_x$ [28], where $k_x = k_0 \sin\theta$ is the tangential component of the wave vector and θ is the angle of incidence. Thus, the amplitude of the reflection coefficient of the differentiator should have a form similar to the modulus x, and the phase should experience a π jump. To demonstrate the operation of the network, an MS was designed to perform differentiation at an angle of incidence of 20°. The MS designed by the ANN has a spectrum similar to the response of an ideal differentiator (Fig. 3a). The operation of the developed device was studied using custom TMM, which allows us to calculate the reflection not only for plane waves, but also for focused Gaussian beams. This is achieved by expansion the Gaussian beam into a set of plane waves [29]. The quality of differentiation was studied for a 3-μm Gaussian beam focused on the MS surface. Since

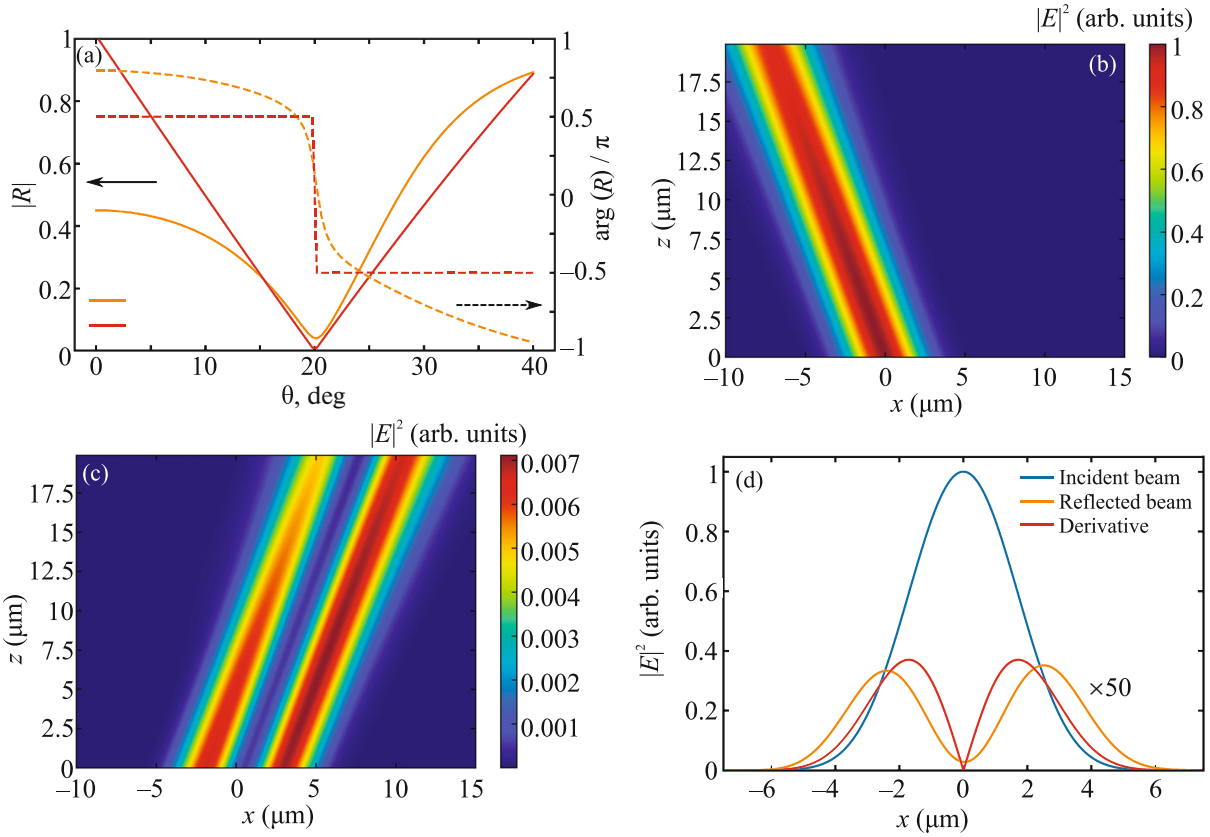

Fig. 3. (Color online) (a, left axis) Amplitude and (a, right axis) phase divided by π of the reflection coefficient versus the angle of incidence for (red lines) the ideal differentiator and (orange lines) multilayer structure designed by artificial neural network. (b, c) Intensities of the (b) incident and (c) reflected beams from the multilayer structure versus the coordinate. (d) Cross sections of the intensity of the (blue line) incident and (orange line) reflected beams and (red line) the analytically calculated derivative of the intensity of the incident beam.

the reflection in the region of 20° is close to 0, the incident (Fig. 3b) and the reflected (Fig. 3c) beams are visualized separately for visual clarity. Figure 3d shows the cross section of the incident and reflected beams, as well as the analytically calculated first-order derivative. Indeed, our ANN designed MS with a response similar to the ideal differentiator, but not identical to it. Increasing the number of layers can provide a better quality of differentiation.

In conclusion, in this paper, we demonstrate a method for training a single artificial neural network to solve the inverse problem of designing a multilayer photonic structure with a given angular spectrum of the reflection coefficient. The proposed method makes it possible to achieve a root-mean-square error of less than 10^{-4}, and thereby design MS that have a target optical response with a high fidelity. The trained ANN successfully solves the problem of non-uniqueness of the inverse problem solution by smoothing the parameter space. In addition, the ANN is suitable for designing nanophotonic devices, for example, an optical differentiator. The proposed method of ANN training can be transferred to the frequency spectrum of the reflection coefficient, which will open up even more opportunities for creating nanophotonic devices with different functionalities.

FUNDING

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (project no. 14.W03.008.31, development of an optical differentiator, and project no. 075-15-2020-801, development of a neural network), by the Russian Foundation for Basic Research (project no. 19-32-90225, development of an optimized transfer matrix method), and in part by the Quantum Technology Center, Moscow State University and the nonprofit Foundation for the Development of Science and Education Intelligence.

REFERENCES

1. A. F. Koenderink, A. Alu, and A. Polman, Science (Washington, DC, U. S.) **348**, 516 (2015).
2. A. D. Gartman, M. K. Kroichuk, A. S. Shorokhov, and A. A. Fedyanin, JETP Lett. **112**, 693 (2020).

3. A. M. Chernyak, M. G. Barsukova, A. S. Shorokhov, A. I. Musorin, and A. A. Fedyanin, JETP Lett. **111**, 46 (2020).
4. P. S. Pankin, B.-R. Wu, J.-H. Yang, K.-P. Chen, I. V. Timofeev, and A. F. Sadreev, Commun. Phys. **3**, 1 (2020).
5. B. I. Afinogenov, A. A. Popkova, V. O. Bessonov, B. Lukyanchuk, and A. A. Fedyanin, Phys. Rev. B **97**, 115438 (2018).
6. B. I. Afinogenov, V. O. Bessonov, I. V. Soboleva, and A. A. Fedyanin, ACS Photon. **6**, 844 (2019).
7. M. N. Romodina, I. V. Soboleva, A. I. Musorin, Y. Nakamura, M. Inoue, and A. A. Fedyanin, Phys. Rev. B **96**, 081401 (2017).
8. W. Wu, W. Jiang, J. Yang, S. Gong, and Y. Ma, Opt. Lett. **42**, 5270 (2017).
9. F. Barachati, A. Fieramosca, S. Hafezian, J. Gu, B. Chakraborty, D. Ballarini, L. Martinu, V. Menon, D. Sanvitto, and S. Kéna-Cohen, Nat. Nanotechnol. **13**, 906 (2018).
10. I. S. Kriukova, V. A. Krivenkov, P. S. Samokhvalov, and I. R. Nabiev, JETP Lett. **112**, 537 (2020).
11. A. E. Schegolev, A. M. Popov, A. V. Bogatskaya, P. M. Nikiforova, M. V. Tereshonok, and N. V. Klenov, JETP Lett. **111**, 371 (2020).
12. K. R. Safronov, D. N. Gulkin, I. M. Antropov, K. A. Abrashitova, V. O. Bessonov, and A. A. Fedyanin, ACS Nano **14**, 10428 (2020).
13. S. D. Campbell, D. Sell, R. P. Jenkins, E. B. Whiting, J. A. Fan, and D. H. Werner, Opt. Mater. Express **9**, 1842 (2019).
14. R. L. Haupt and D. H. Werner, *Genetic Algorithms in Electromagnetics* (Wiley, New Jersey, 2007).
15. J. Robinson and Y. Rahmat-Samii, IEEE Trans. Antennas Propag. **52**, 397 (2004).
16. J. S. Jensen and O. Sigmund, Laser Photon. Rev. **5**, 308 (2011).
17. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, Nat. Photon. **12**, 659 (2018).
18. D. Liu, Y. Tan, E. Khoram, and Z. Yu, ACS Photon. **5**, 1365 (2018).
19. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, Sci. Adv. **4**, eaar4206 (2018).
20. C. C. Nadell, B. Huang, J. Malof, and W. J. Padilla, Opt. Express **27**, 27523 (2019).
21. I. Sajedian, J. Kim, and J. Rho, Microsyst. Nanoeng. **5**, 1 (2019).
22. J. He, C. He, C. Zheng, Q. Wang, and J. Ye, Nanoscale **11**, 17444 (2019).
23. S. So, J. Mun, and J. Rho, ACS Appl. Mater. Interfaces **11**, 24264 (2019).
24. Z. Liu, D. Zhu, S. P. Rodrigues, K. T. Lee, and W. Cai, Nano Lett. **18**, 6570 (2018).
25. Y. Kiarashinejad, S. Abdollahramezani, and A. Adibi, Npj Comput. Mater. **6**, 1 (2020).
26. S. An, C. Fowler, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, J. Ding, A. M. Agarwal, C. Rivero-Baleine, K. A. Richardson, T. Gu, J. Hu, and H. Zhang, ACS Photon. **6**, 3196 (2019).
27. C. C. Katsidis and D. I. Siapkas, Appl. Opt. **41**, 3978 (2002).
28. A. Pors, M. G. Nielsen, and S. I. Bozhevolnyi, Nano Lett. **15**, 791 (2015).
29. F. I. Baida and M.-P. Bernal, Commun. Phys. **3**, 1 (2020).