Computing Optical Properties of Photonic Crystals by Using Multilayer Perceptron and Extreme Learning Machine

Adriano da Silva Ferreira, Gilliard Nardel Malheiros-Silveira, and Hugo Enrique Hernández-Figueroa, *Senior Member, IEEE*

***Abstract*—In this paper, dispersion relations (DRs) of photonic crystals (PhCs) are computed by multilayer perceptron (MLP) and extreme learning machine (ELM) artificial neural networks (ANNs). Bi- and tri-dimensional optimized structures presenting distinct DRs and photonic band gaps (PBGs) were selected for case studies. Optical properties of a set of PhCs with similar geometries and different dimensions were calculated by an electromagnetic solver in order to provide input data for ANN training and testing. We demonstrate that simple- and fast-training ANN models are capable of providing accurate DRs' curves in a very short time.**

***Index Terms*—Dispersion relation, extreme learning machine, multilayer perceptron, photonic crystal, photonic band gap.**

I. Introduction

PHOTONIC crystals are optical structures composed of a periodical distribution of dielectric materials by which light propagation is affected. Such structures may be designed for defining DRs presenting frequency ranges in which light cannot propagate in certain directions: a phenomenon called photonic band gap (PBG) [1], whose characteristics enable the engineering of technologies for light control [2].

PBG-based devices design generally relies upon electromagnetic solvers based on numerical methods such as finite element method [3] and block-iterative frequency-domain methods [4], among others, and becomes time and memory consuming when dealing with structure optimization or solving inverse problems. Such iterative analyses are particularly computationally-demanding for tri-dimensional PhCs, for which parallel numerical solutions and/or high-performance computing are frequently mandatory.

In order to provide an efficient alternative for computing PhCs DRs, we modeled MLP [5] and ELM [6] ANNs for inferring DRs of variants of bi-dimensional (2D) and tri-dimensional (3D) PhC structures taken as references. We focused on designing simple, fast-training ANN schemes through relatively-small datasets which may enable effective and efficient DR and PBG (whenever it exists) computations.

A number of photonics applications have employed and benefited from the ANN capabilities [7], such as prediction of gain and noise figures of PhC fiber amplifiers [8] and coupling efficiency of photonic couplers [9], as well as optimization of passive polarization rotator based on slanted rib waveguide and design of ultra-flattened zero dispersion PhC fiber [10]. Furthermore, the prediction of DRs and PBGs through ANNs was firstly studied in [11], where two distinct MLP models estimated, separately, fundamental modes and PBGs of 2D PhCs with distinct geometric and material properties values.

Here, we address a comprehensive study by modeling ANNs for building sets of modes of 2D and 3D PhCs. In effect, existing PBGs should be obtained from the predicted modes. A broader analysis of the capabilities of multilayer feed-forward ANNs [5] (the MLP and ELM's architectures) can be also observed in this problem along with a comparison of MLP and ELM performances.

This work is organized as follows. Section II describes the MLP and ELM's concepts and modeling as well as the creation of variant PhCs through geometric value changes for ANN training and testing. Section III presents and discusses the assessment of the modeled ANNs on testing PhCs by comparing their DR estimations and computing runtime with MIT Photonic Bands (MPB) [4], an electromagnetic solver that computes DRs and modes of periodic dielectric structures. Lastly, Section IV presents our conclusions and future works.

II. Materials and Methods

The MLP and ELM's architectures, neural model and training methods are introduced in this section along with the PhC types used for generating datasets. Variant PhCs are created by changing some geometric property values and their DRs were

Manuscript received April 8, 2018; revised July 4, 2018; accepted July 9, 2018. Date of publication July 16, 2018; date of current version August 6, 2018. The work of A. da Silva Ferreira and H. E. Hernández-Figueroa was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, in part by the Conselho Nacional de Desenvolvimento Tecnológico (CNPq) under Grant #312110/2016-2, and in part by the Fundação de Amparo à Pesquisa do Estado de São Paulo under Project 2015/24517-8. The work of G. N. Malheiros-Silveira was supported by CNPq under Grant #300594/2017-8. *(Corresponding author: Adriano da Silva Ferreira.)*

A. da Silva Ferreira and H. E. Hernández-Figueroa are with the School of Electrical and Computer Engineering, University of Campinas, Campinas SP 13083-852, Brazil (e-mail: adrianof@decom.fee.unicamp.br; hugo@decom.fee.unicamp.br).

G. N. Malheiros-Silveira is with the School of Electrical and Computer Engineering, University of Campinas, Campinas SP 13083-852, Brazil, with the São Paulo State University (UNESP), Campus of São João da Boa Vista, SP, 13876-750, Brazil, and also with the Centro de Tecnologia da Informação Renato Archer, Campinas SP 13092-599, Brazil (e-mail: gilliardn@yahoo.com.br).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2018.2856364

0733-8724 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

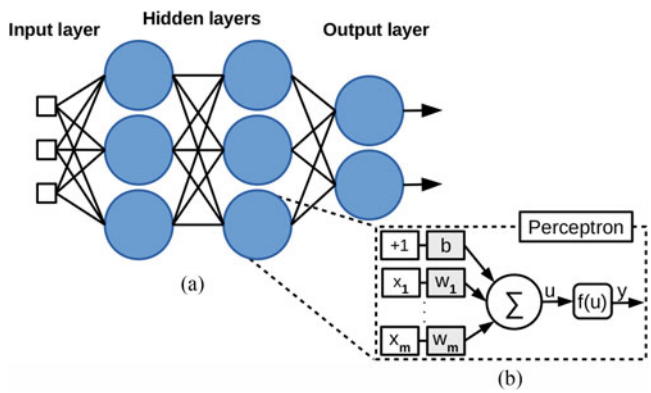

Fig. 1. Schematic of a (a) two-hidden-layer ANN and (b) a perceptron model with input $\{x_j\}_{j=1}^m$, weights $\{w_j\}_{j=1}^m$, bias b and activation function $y = f(u)$.

calculated from the irreducible Brillouin zone (IBZ) using the original mesh resolutions: 128/a and 32/a (where a is the constant lattice or the unit cell dimension) for 2D [12] and 3D PhCs [13], respectively. Finally, 16 points were interpolated between the IBZ's corner points in all simulations.

A. Artificial Neural Networks

As universal learning machines [14], MLP and ELM can be trained with a finite and appropriate labeled dataset—in our case, the set of parameters relating PhC geometric data (inputs) and DRs (target outputs)—in order to build models which generate suitable outputs for input data not provided during the training process (unknown inputs).

The MLP and ELM approximation capabilities are directly linked to their multilayer feed-forward architecture [14] (Fig. 1(a)): a structure where an input signal (input layer) is propagated forward through—and parallelly processed by—neurons of one or more hidden layers, until an answer is produced by neurons of the output layer. Adjacent layers are fully interconnected: all neurons of a layer process all neurons' outputs of the precedent layer and send their outputs to all neurons of the following layer. The processing unit is the perceptron, a neuron model composed of weights, bias, a summation, and an activation function (Fig. 1(b)).

An *m*-length input signal $\{x_j\}_{j=1}^m$ is processed by a perceptron of a given layer as

$$y = f\left(\sum_{j=1}^{m} w_j x_j + b\right), \tag{1}$$

where w_j is the weight for the input signal element x_j, b is the bias term, and $f(.)$ or y is the activation function, which can be linear or non-linear. By summing Eq. (1) of each output layer's perceptron, the network's output signal is generated and can be compared to a given target output for providing a performance measure for the current ANN configuration.

Considering a dataset $\{(s_i, d_i)\}_{i=1}^N$, where for a given PhC data s_i is expected the DR d_i, a total error energy can be obtained

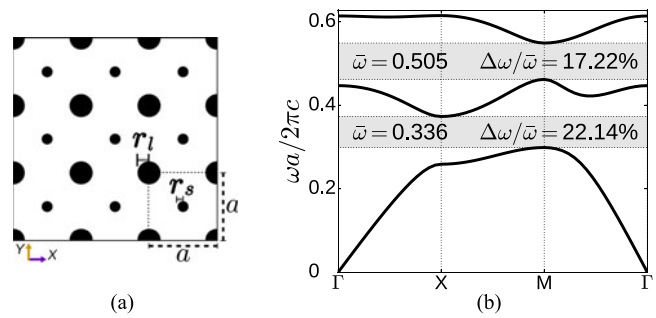

Fig. 2. (a) Square-lattice PhC type with small and large cylinders with radii r_s and r_l, respectively; and (b) its multiple PBGs between the 3 TM lowest modes. Γ, X, M, and Γ represent the IBZ's corner points, ω the angular frequency and c the speed of light in vacuum [12].

for the network as

$$E = \sum_{i=1}^{N} \sum_{n=1}^{q} (d_{in} - y_{in})^2, \tag{2}$$

where *q* is the total number of output neurons, d_{in} the *n*-th target output's attribute and y_{in} the *n*-th output neuron's answer regarding the *i*-th pattern. A mean-square error (MSE) is obtained by E/N in order to provide an average measure for the ANN performance regarding N patterns.

Before presenting the training methods by which we aim at designing low-MSE ANNs, we introduce the employed PhC types and describe the geometric changes for creating variant PhCs and the definition of simple patterns for modeling ANNs.

B. Datasets

This study considers transverse magnetic (TM) and transverse electric (TE) polarized 2D PhCs [12], and three distinct 3D ones [2], [13]. They are optimized structures that present single (3D PhCs) and multiple (2D PhCs) PBGs. By linearly varying the geometric attribute values of each PhC type through the proportions $p = [0.5, 0.6, \ldots, 0.9, \ldots, 1.2, 1.3]$, nine variant structures, including the original ones, were created to supply data for ANN training.

Once obtained the training PhCs' DRs, simple patterns can be generated for designing a simple ANN scheme. As a DR is formed by the set of q normalized frequencies of each wave vector element (**k** point), a pattern can be formed by relating the PhC's geometry and the **k** point and magnitude to the respective frequencies. An ANN design can then receive, as input pattern, an *l*-length geometric attributes set along with the **k** point (k_x, k_y, k_z) and magnitude $k_{mag}/2\pi$ as

$$s_i = [(g_1, g_2, \ldots, g_l), (k_x, k_y, k_z, k_{mag}/2\pi)], \tag{3}$$

for estimating the q normalized frequencies, at the respective **k** point, represented by the following target output:

$$d_i = [f_1, f_2, \ldots, f_q].$$

First, we present 2D square-lattice PhCs whose geometries are composed of doped GaAs ($\epsilon_r = 12.8$) and are embedded in air ($\epsilon_r = 1.0$) [12]. Fig. 2 shows such PhC type composed of two

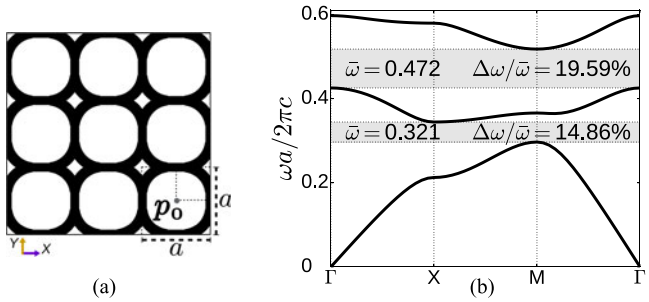

Fig. 3. (a) Square-lattice, connected-structure PhC type whose outer rounded corners are centralized in p_0; and (b) its multiple PBGs between the 3 TE first modes. Γ, X, M, and Γ represent the IBZ's corner points, ω the angular frequency and c the speed of light in vacuum [12].

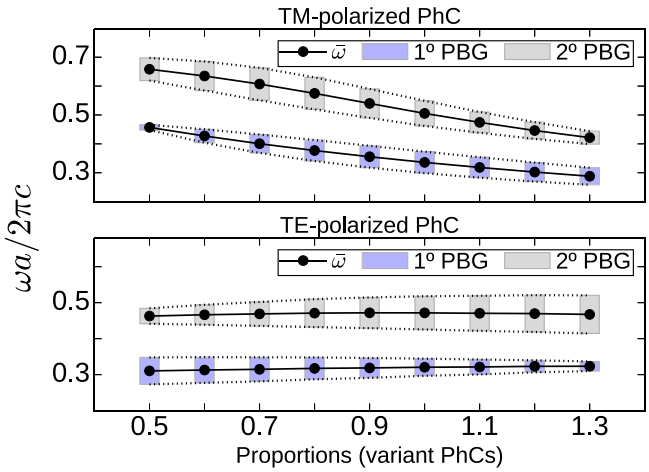

Fig. 4. PBGs' maps and central frequencies $\bar{\omega}$ of the training 2D PhCs.

different cylinders with radii $r_s = 0.08$/a and $r_l = 0.17$/a and its TM DR. Fig. 3 depicts a connected-structure PhC with outer rounded corners' centralized at $p_0(x = 0.113\text{/a}, y = 0.113\text{/a})$ and its TE DR. The first three bands were considered for both 2D PhC types, since they define the first and second PBGs (characterized by the fractional gap size $\Delta\omega/\bar{\omega}$, where $\Delta\omega$ is the absolute gap and $\bar{\omega}$ the central frequency).

As the radii of the small and large cylinders represent the geometric attributes of the TM-polarized PhC (TM-PhC), its input pattern can be expressed through Eq. (3) as

$$s_i = [(r_s, r_l), (k_x, k_y, k_z, k_{mag}/2\pi)],$$

where r_s and r_l are the radii of the small and large cylinders, respectively. On the TE-polarized PhC (TE-PhC), the outer rounded corners' center p_0 represents the geometric variable (the inner one was fixed). As each center's coordinate varies proportionally equal, we consider that p_0 can be represented by a unique value (x or y) so that the TE-PhC's input pattern can be given by

$$[p_0, (k_x, k_y, k_z, k_{mag}/2\pi)].$$

Each variant PhC possesses a different DR. Fig. 4 shows the PBGs' maps and respective central frequencies of the training 2D PhCs. All structures possess PBGs between 1–2 and 2–3 modes. On the TM-PhCs, the central frequencies decrease

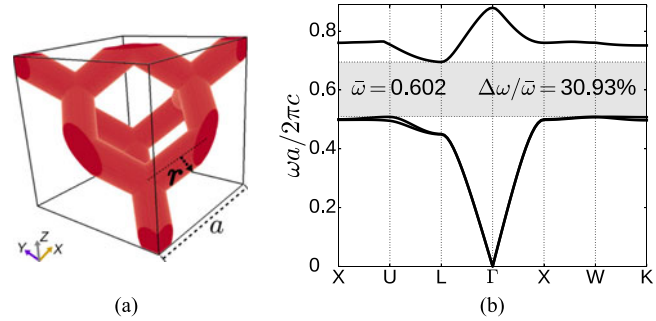

Fig. 5. (a) Fcc-lattice PhC type with dielectric cylinders' radius r; and (b) its PBG between the 2–3 modes. X, U, L, Γ, X, W, K correspond to the IBZ's corner points, w to the angular frequency and c to the speed of light in vacuum [13].

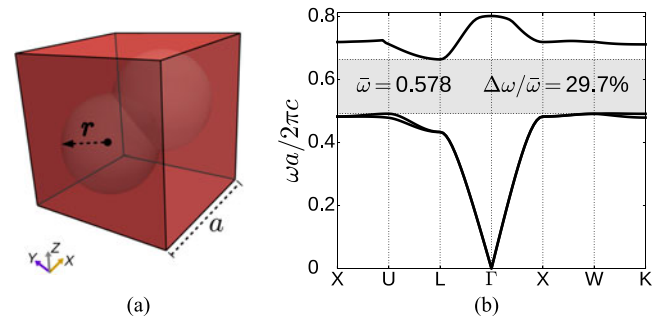

Fig. 6. (a) Fcc-lattice PhC type composed of air spheres with radius r and embedded in dielectric background; and (b) its PBG between the 2–3 modes. X, U, L, Γ, X, W, K represent the IBZ's corner points, w the angular frequency and c the speed of light in vacuum [2].

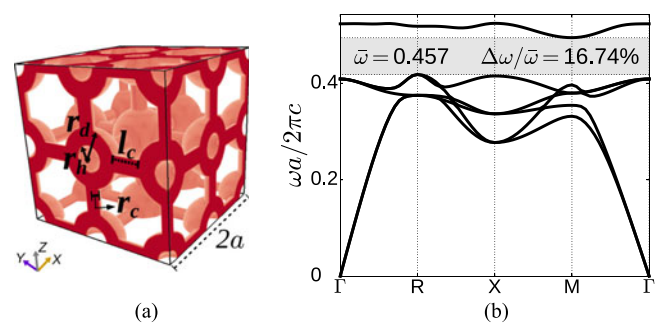

Fig. 7. (a) Simple-cubic lattice PhC type with hollow and dielectric spheres' radii r_h and r_d, and cylinders' radius and length r_c and l_c, respectively; and (b) its PBG between the 5–6 modes. Γ, R, X, M, Γ represent the IBZ's corner points, w the angular frequency and c the light speed in vacuum [13].

and both PBGs firstly increase and then decrease along the proportions' increase. On the TE-PhCs, a decrease of the first PBG is followed by an increase of the second PBG and the central frequencies vary slightly along the proportions' rise.

The employed 3D PhCs are based on simple and face-centered cubic (fcc) lattices, and their dielectric constants are $\epsilon_r = 12.96$ (which is similar to GaAs in the near infrared) and air ($\epsilon_r = 1.0$). Fig. 5(a) shows a fcc-lattice PhC whose geometry is formed by dielectric cylindrical bonds with radius $r = 0.1$/a and height $\sqrt{3/4}$ a [13]. Fig. 6(a) shows a fcc-lattice PhC composed of two air spheres with radius $r = 0.325$/a each, in a dielectric background [2]. Fig. 7(a) shows a simple-cubic lattice composed

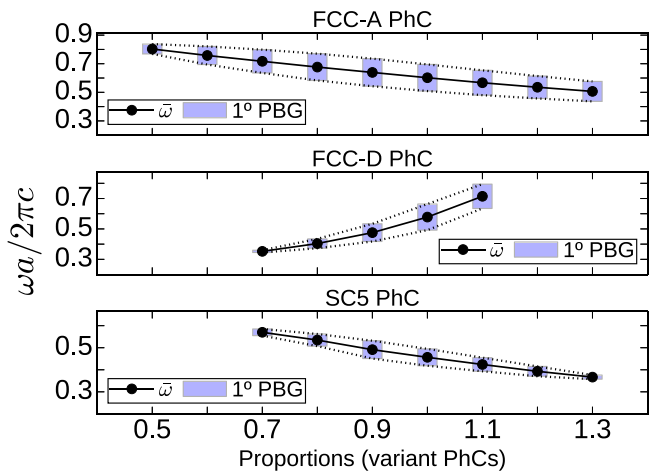

Fig. 8. PBGs' maps and central frequencies $\bar{\omega}$ of the training 3D PhCs.

of hollow and dielectric spheres with radii $r_h = 0.14$/a and $r_d = 0.36$/a, respectively, connected by dielectric cylindrical bonds with radius $r_c = 0.105$/a and length $l_c = a - 2r_d$ [13]. For both fcc-lattice structures, the first three bands were taken into account, since a $\Delta\omega/\bar{\omega} = 30.93\%$ appears between the 2–3 modes of the PhC with dielectric cylinders embedded in air (FCC-A PhC) and a $\Delta\omega/\bar{\omega} = 29.7\%$ appears between the 2–3 modes of the PhC with air spheres and dielectric background (FCC-D PhC) (Figs. 5(b) and 6(b), respectively). The simple-cubic structure (SC5 PhC) has a $\Delta\omega/\bar{\omega} = 16.74\%$ between the 5-6 modes (Fig. 7(b)), so that its first six modes were considered.

As the fcc-lattice PhCs are parameterized by the radii of the cylinders and spheres, their input patterns can be expressed as

$$[r, (k_x, k_y, k_z, k_{mag}/2\pi)].$$

The simple-cubic lattice structure is characterized by the radii of the hollow and dielectric spheres and the radius and length of the cylinders so that its input pattern is given by

$$[(r_h, r_d, r_c, l_c), (k_x, k_y, k_z, k_{mag}/2\pi)].$$

Fig. 8 illustrates how PBGs and central frequencies vary according to the training 3D PhCs. One may observe a slightly non-linear variation of the FCC-A and SC5 PhCs' absolute gaps—which increase as long as the lower proportions get closer to the original PhC ones, and decrease when the proportions get bigger than the original ones—around a nearly-linear decrease of their central frequencies. The PBGs and central frequencies of the FCC-D PhCs increase in a non-linear manner between the proportions $p = 0.7$ and $p = 1.1$. However, the FCC-D PhCs modified by $p = 0.5$, $p = 0.6$, $p = 1.2$ and $p = 1.3$, as well as the SC5 PhCs changed by $p = 0.5$ and $p = 0.6$, do not present PBGs.

C. PhC and DR Modeling With ANNs

The MLP and ELM modeling is accomplished through training processes by which the adjustment to their weights might lead to the minimization of Eq. (2) (or MSE). They are trained through different methods. The whole set of MLP weights have been obtained through the second order *Broyden-Fletcher-Goldfarb-Shanno* algorithm [5], whereas the ELM output neurons' weights have been analytically adjusted through the *Moore-Penrose* generalized inverse [15], based on the method of least squares (the weights of hidden neurons are randomly determined and not adjusted during ELM training).

TABLE I
MLP AND ELM ARCHITECTURES PER PHC DATASET

MLP property	TM-PhC	TE-PhC	FCC-A	FCC-D	SC5
Hidden layers (*hl*)	3	3	3	3	3
Neurons per *hl*	23	25	11	27	20
Output neurons	3	3	3	3	6
Transfer function*	*tanh*	*tanh*	*tanh*	*tanh*	*tanh*
ELM property	TM-PhC	TE-PhC	FCC-A	FCC-D	SC5
Hidden layers (*hl*)	1	1	1	1	1
Neurons per *hl*	80	85	88	298	100
Output neurons	3	3	3	3	6
Transfer function*	*multi*	*tanh*	*multi*	*multi*	*multi*

**tanh* and *multi*: hyperbolic tangent and multiquadric activation functions.

TABLE II
ANN TRAINING PERFORMANCE PARAMETERS

MLP	TM-PhC	TE-PhC	FCC-A	FCC-D	SC5
Iterations	360	288	662	1147	716
MSE	5.9e−6	2.4e−5	8.8e−6	1e−5	8.2e−6
Time(s)	1.8	1.4	1.47	3	2
ELM	TM-PhC	TE-PhC	FCC-A	FCC-D	SC5
MSE	1.1e−6	5.4e−6	1.2e−6	2.6e−6	5.2e−6
Time(s)	0.53	0.53	0.54	0.7	0.54

Both ANNs assumed different architectures for each PhC dataset. Although these architectures were not defined by a systematic process such as cross-validation [5], we designed ANNs with reasonable performances through an incremental number of neurons and layers. Table I presents the obtained ANN architectures for each PhC dataset. Output layer's neurons consist of linear activation functions in all ANN models.

Before training MLP and ELM, the input patterns of the datasets were normalized by removing the mean and scaling to unit variance. On MLP training, the learning rate and the maximum number of iterations were fixed to 1e–4 and 1.5 e3, respectively. Table II lists the training performance parameters of both ANNs for each training dataset. Fig. 9 shows the MLPs' MSEs along the 200-first training iterations for each PhC dataset (where the MSEs mostly varied). One may note sharp MSE falls during the first iterations, demonstrating that the training algorithm took big steps towards the local minimum of Eq. (2).

III. NUMERICAL RESULTS AND DISCUSSION

In this section, we validate the obtained MLP and ELM models by evaluating their DR outputs for testing PhCs and present their computing performances regarding MPB. The ANN models are tested through two unknown PhCs of each PhC type. Finally, the MPB's and ANNs' runtimes are compared, followed by a detailed discussion of the results.

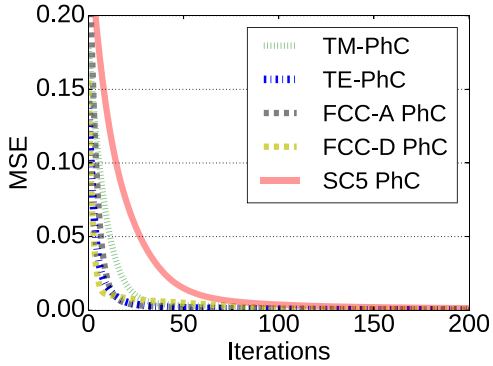

Fig. 9. MSE behavior along the MLP training for each PhC dataset.

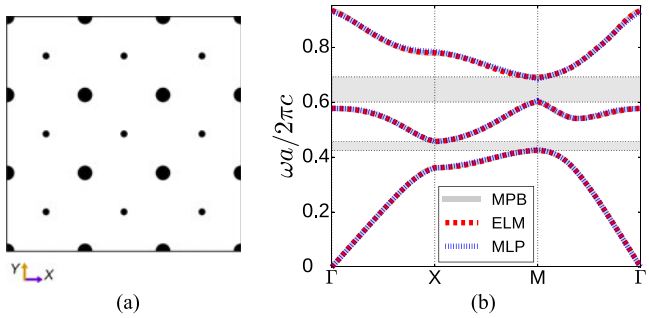

Fig. 10. (a) TM-PhC 1 with cylinders' radii $r_s = 0.044$/a and $r_l = 0.0935$/a, and (b) its DR computed by MPB, ELM, and MLP (PBG in grey color).

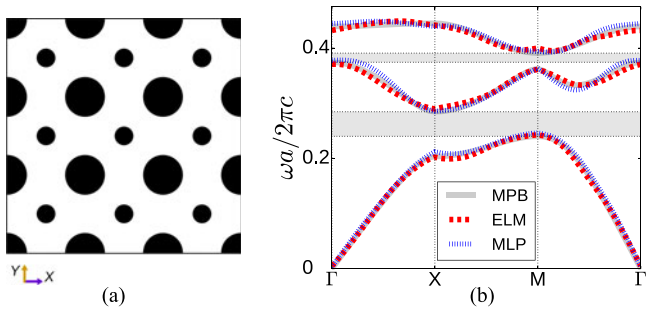

Fig. 11. (a) TM-PhC 2 with cylinders' radii $r_s = 0.12$/a and $r_l = 0.255$/a, and (b) its DR computed by MPB, ELM, and MLP (PBG in grey color).

A. TM-PhC

In this issue, the small and large dielectric cylinders of the test TM-PhCs are considerably different in size. The proportions $p = 0.55$ and $p = 1.5$ were applied to the original cylinders' radii in order to generate two test PhCs—denoted as TM-PhC 1 and TM-PhC 2 (Figs. 10(a) and 11(a))—whose cylinders notably define distinct filling factors in the unit cell. It resulted in distinct DRs of which their first 3 modes are confined within noticeably different ranges of frequencies. In TM-PhC 1, these 3 modes are defined between the normalized frequencies 0–0.9343 (ωa/2πc) (Fig. 10(b)), whereas in TM-PhC 2 they are between 0–0.4452 (ωa/2πc) (Fig. 11(b)). Both TM-PhCs also present PBGs between the 1–2 and 2–3 modes. Their PBGs are in accord to the TM gap map of Fig. 4, where PhCs from lower proportions possess a larger second PBG and PhCs from higher proportions possess a larger first PBG.

The trained MLP and ELM's computations agreed very well with MPB calculations for both test TM-PhCs. Although their predictions were more accurate for TM-PhC 1 (Fig. 10(b)) than for TM-PhC 2 (Fig. 11(b)), both ANN models suitably built the first 3 modes, defining the multiple PBGs which appear between the 1–2 and 2–3 TM modes. The MLP and ELM's MSEs were $7.47\text{e}{-6}$ and $1.42\text{e}{-6}$ for building the TM-PhC 1's DR, and $2.62\text{e}{-5}$ and $1.77\text{e}{-5}$ for the TM-PhC 2's one.

In general, ELM outperformed MLP in computing the underlying PBGs. As shown in Table III, ELM not only predicted more accurately the central frequencies of TM-PhC 1 but also the frequency ranges. For TM-PhC 2, both ANN models computed roughly $\bar{\omega}$ of the first PBG, but ELM provided a better precision of $\Delta\omega$, leading to a better approximation of $\Delta\omega/\bar{\omega}$. ELM also provided a better estimation of both $\Delta\omega$ and $\bar{\omega}$ of the second TM-PhC 2's PBG, providing a more accurate fractional gap size than MLP.

B. TE-PhC

In this case study, the geometric differences of both TE-test PhCs define DRs whose the first 3 modes are within slightly different frequency ranges. The first 3 modes of the PhC from the proportion $p = 0.45$ (TE-PhC 1) are defined between 0–0.6927 (ωa/2πc) (Fig. 12), and of the PhC from $p = 1.25$ (TE-PhC 2) are defined between 0–0.5705 (ωa/2πc) (Fig. 13). Their multiple PBGs respect the TE gap map of Fig. 4, where the first PBGs are larger than the second ones for PhCs from lower proportions and vice-versa.

MLP and ELM suitably built the first 3 TE modes of both test PhCs, delimiting rightly the existing multiple PBGs (Figs. 12(b) and 13(b)). The MLP and ELM's MSEs were $1.1\text{e}{-5}$ and $3.41\text{e}{-6}$ for obtaining the 3 lowest modes of the TE-PhC 1, and $2.63\text{e}{-6}$ and $5.22\text{e}{-7}$ for obtaining the 3 lowest modes of the TE-PhC 2. As shown in Table III, MLP was more accurate than ELM for estimating the first PBG of TE-PhC 1 and the second PBG of TE-PhC 2 (although ELM predicted slightly better $\bar{\omega}$ of the TE-PhC 2's second PBG, MLP provided a better approximation of $\Delta\omega$ and $\Delta\omega/\bar{\omega}$). ELM estimated better the second PBG of TE-PhC 1 and the first PBG of TE-PhC 2. Altogether, both ANN models provided akin performances for computing PBGs in this case.

C. FCC-A PhC

These fcc-lattice PhCs possess cylinders with noticeable different volume sizes, producing DRs with different profiles: the 3 lowest modes of the PhC from $p = 0.4$ (FCC-A PhC 1) are delimited to the frequency range 0–1.1874 (ωa/2πc) (Fig. 14), whereas the ones of the PhC from $p = 1.4$ (FCC-A PhC 2) are delimited to 0–0.6804 (ωa/2πc) (Fig. 15). Moreover, FCC-A PhC 1 does not possess PBG between these modes and FCC-A PhC 2 does with a $\Delta\omega/\bar{\omega} = 25.35\%$ between the 2–3 modes. Since all FCC-A training PhCs contain PBGs, such tests provide a different scenario which is adequate for assessing the generalization capabilities (suitable estimation for unknown inputs) of the trained ANNs.

TABLE III
MPB, MLP AND ELM'S PBG COMPUTATIONS (NORMALIZED FREQUENCY ($\omega a/2\pi c$)) FOR TEST PHCS

		MPB			MLP			ELM		
Devices	p	$\Delta\omega$	$\bar{\omega}$	$\Delta\omega/\bar{\omega}$	$\Delta\omega$	$\bar{\omega}$	$\Delta\omega/\bar{\omega}$	$\Delta\omega$	$\bar{\omega}$	$\Delta\omega/\bar{\omega}$
TM-PhC 1*	0.55	0.0344	0.4412	7.80%	0.0287	0.4428	6.48%	0.0318	0.4411	7.20%
		0.0918	0.6468	14.20%	0.0796	0.6458	12.33%	0.0878	0.6462	13.59%
TM-PhC 2*	1.5	0.0446	0.2625	16.99%	0.0392	0.2660	14.73%	0.0467	0.2664	17.53%
		0.0166	0.3829	4.33%	0.0143	0.3857	3.71%	0.0161	0.3810	4.23%
TE-PhC 1*	0.45	0.0622	0.3553	17.51%	0.0600	0.3553	16.90%	0.0597	0.3534	16.89%
		0.0548	0.5671	9.66%	0.0612	0.5624	10.88%	0.0537	0.5645	9.51%
TE-PhC 2*	1.25	0.0458	0.3184	14.37%	0.0414	0.3198	12.94%	0.0435	0.3193	13.61%
		0.0763	0.4560	16.74%	0.0762	0.4556	16.73%	0.0750	0.4559	16.44%
FCC-A PhC 2	1.4	0.1221	0.4818	25.35%	0.1274	0.4869	26.17%	0.1279	0.4880	26.21%
FCC-D PhC 2	1.15	0.0693	0.8	8.67%	0.0575	0.803	7.16%	0.0585	0.802	7.3%
SC5 PhC 2	1.35	0.0112	0.3578	3.14%	0.0159	0.3583	4.43%	0.0136	0.3542	3.84%

*PhCs using two rows for listing their two PBGs. FCC-A PhC 1 ($p = 0.4$), FCC-D PhC 1 ($p = 0.45$) and SC5 PhC 1 ($p = 0.55$) do not possess PBGs.

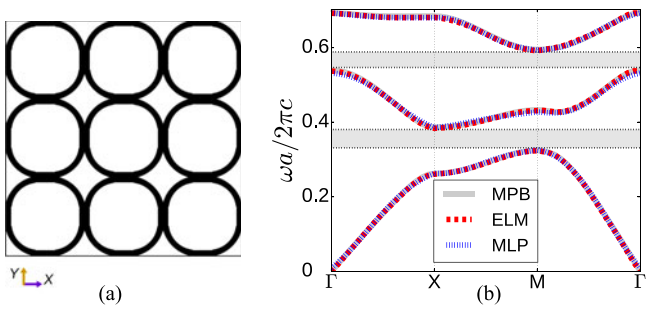

Fig. 12. (a) TE-PhC 1 with outer corners' center $p_0 = 0.05085$/a, and (b) its DR computed by MPB, ELM, and MLP (PBG in grey color).

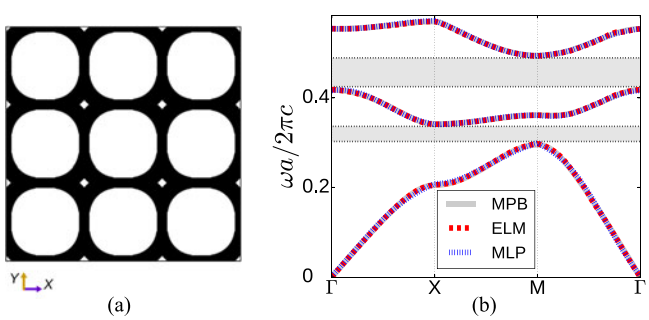

Fig. 13. (a) TE-PhC 2 with outer corners' center $p_0 = 0.14125$/a, and (b) its DR computed by MPB, ELM, and MLP (PBG in grey color).

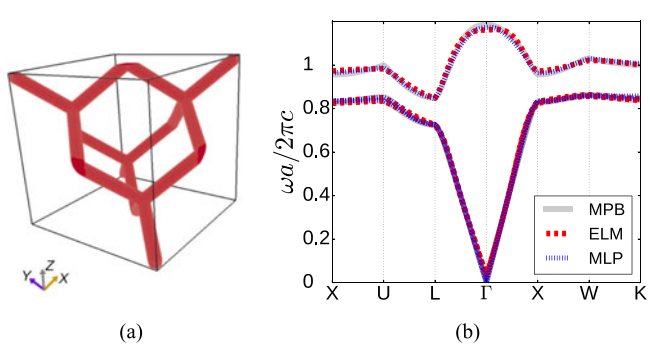

Fig. 14. (a) FCC-A PhC 1 with cylinders' radius $r = 0.04$/a, and (b) its DR computed by MPB, ELM, and MLP. This structure does not possess PBG.

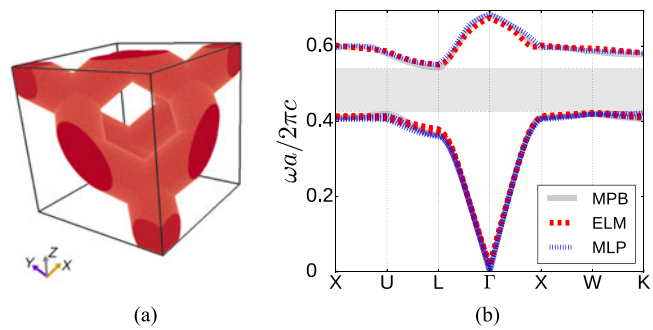

Fig. 15. (a) FCC-A PhC 2 with cylinders' radius $r = 0.14$/a, and (b) its DR computed by MPB, ELM, and MLP (PBG in grey color).

MLP and ELM succeed in calculating the underlying DRs of both test PhCs. Firstly, both ANN models suitably computed the 3 lowest modes of FCC-A PhC 1, determining a DR without PBGs between the considered modes (Fig. 14(b)). They also computed the FCC-A PhC 2's DRs with good precision, as illustrated in Fig. 15(b). MLP estimated slightly better the existing PBG, but both ANNs obtained good approximations in regard to the MPB's computed PBG, as shown in Table III. The MLP and ELM's MSEs were $4.97\text{e}{-5}$ and $2.98\text{e}{-5}$ for building the modes of both PhCs, respectively.

D. FCC-D PhC

In this case study, for sake of completeness, we analyze the ANN performance in a different context, once this PhC type is made of a different background material. Furthermore, the different radius values for both test PhCs defined geometries with either separated or joined spheres−PhC from $p = 0.45$ or FCC-D PhC 1 (Fig. 16(a)) and PhC from $p = 1.15$ or FCC-D PhC 2 (Fig. 17(a)), respectively. The geometric differences resulted in DRs with considerably different profiles: the 3 lowest modes of the FCC-D PhC 1 are delimited to the frequency range 0–0.4904 (ωa/2πc) (Fig. 16(b)), and the ones of the FCC-D PhC 2 are delimited to 0–1.0820 (ωa/2πc) (Fig. 17(b)). Moreover, the FCC-D PhC 1 does not possess PBGs between the

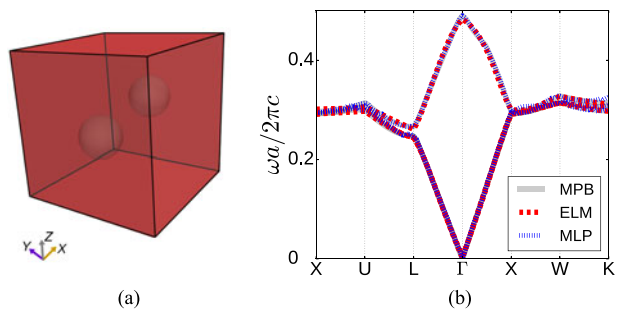

Fig. 16. (a) FCC-D PhC 1 with spheres' radius $r = 0.14625$/a, and (b) its DR computed by MPB, ELM, and MLP. This structure does not possess PBG.

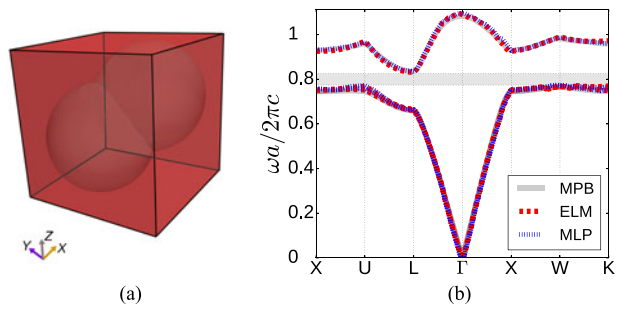

Fig. 17. (a) FCC-D PhC 2 with spheres' radius $r = 0.37375$/a, and (b) its DR computed by MPB, ELM, and MLP (PBG in grey color).

considered modes and the FCC-D PhC 2 does with a $\Delta\omega/\bar{\omega} = 8.67\%$ between the 2–3 modes. Also, one may note a reasonable difference between their 3th modes, where one assumed a sharp form (FCC-D PhC 1) and the other a rounded shape (FCC-D PhC 2) at the k-point $= \Gamma$, providing a challenging scenario for ANN evaluation considering the small number of samples for ANN training.

MLP and ELM satisfactorily computed the underlying DRs of these test PhCs. Firstly, both ANNs computed with good precision the 3 lowest modes of FCC-D PhC 1, building DRs without PBGs between the considered modes (Fig. 16(b)). They also reasonably estimated the FCC-D PhC 2's DRs, as illustrated in Fig. 17(b). In this case, ELM estimated slightly better $\Delta\omega$ and $\bar{\omega}$, as shown in Table III. The MLP and ELM's MSEs were $6.2\text{e}{-6}$ and $6.39\text{e}{-6}$ for estimating the FCC-D PhC 1' DR, and $5\text{e}{-5}$ and $1\text{e}{-4}$ for the FCC-D PhC 2' DR.

E. SC5 PhC

Finally, this case study provides a more challenging instance since MLP and ELM are assessed for computing higher modes than the previous tests. The first 6 modes of the SC5-based test PhCs also provide DRs without PBGs (PhC from $p = 0.55$ or SC5 PhC 1) (Fig. 18) and with a PBG between the 5–6 modes (PhC from $p = 1.35$ or SC5 PhC 2) (Fig. 19). The modes of these test PhCs are within distinguishable ranges of frequencies: 0–0.7138 (ωa/2πc) in SC5 PhC 1's DR (Fig. 18(b)) and 0–0.3840 (ωa/2πc) in SC5 PhC 2's one (Fig. 19(b)).

MLP and ELM were capable of building the first six modes of SC5 PhC 1 and 2. As shown in Fig. 18(b), both ANN models obtained DRs without PBGs for SC5 PhC 1, corresponding to the MPB calculations. Less precise were the ANN models'

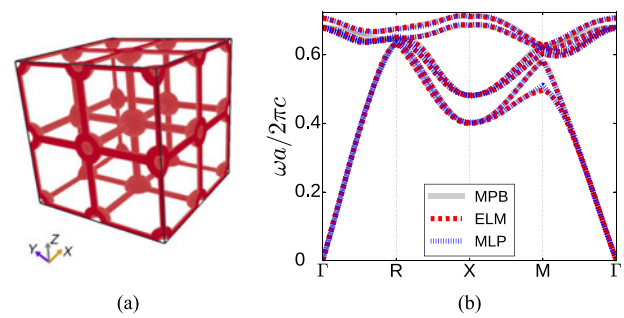

Fig. 18. (a) SC5 PhC 1 with hollow and dielectric spheres' radii $r_h = 0.077$/a and $r_d = 0.198$/a, respectively, and cylinders with radius $r_c = 0.05775$/a; and (b) its DR without PBG computed by MPB, ELM, and MLP.

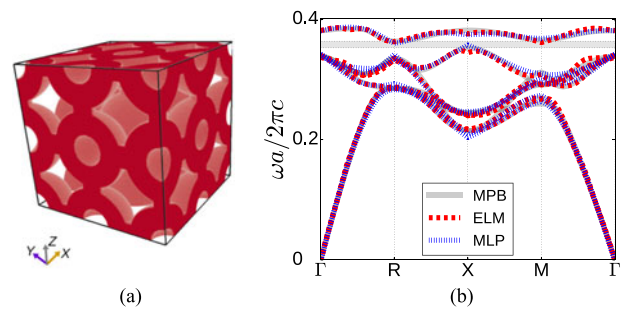

Fig. 19. (a) SC5 PhC 2 with hollow and dielectric spheres' radii $r_h = 0.189$/a and $r_d = 0.486$/a, respectively, and cylinders with radius $r_c = 0.14175$/a; and (b) its DR computed by MPB, ELM, and MLP (PBG in grey color).

TABLE IV
MPB, MLP, AND ELM MEAN ELAPSED REAL TIME (S) FOR COMPUTING A TEST-PHC DR

Method	TM-PhC	TE-PhC	FCC-A	FCC-D	SC5
MPB	14.98	24.98	165.87	121.3	192.80
MLP	2.50e−3	2.55e−3	2.45e−3	2.50e−3	2.51e−3
ELM	2.55e−3	2.55e−3	2.55e−3	2.53e−3	2.55e−3

computed modes for SC5 PhC 2 (Fig. 19(b)). However, MLP and ELM still calculated reasonably the frequencies which define the PBG, as shown in Table III. Despite predicting slightly less accurate the PBG's $\bar{\omega}$, ELM estimated better the PBG's $\Delta\omega$, leading to a more accurate $\Delta\omega/\bar{\omega}$. The MSEs of MLP and ELM were $1.43\text{e}{-5}$ and $1.22\text{e}{-6}$ for building the SC5 PhC 1' DR, and $9.22\text{e}{-6}$ and $7.1\text{e}{-6}$ for the SC5 PhC 2' DR.

F. Computing Performance

Table IV lists the mean elapsed real times of MPB, MLP, and ELM for computing the DR of a test sample of each PhC type. The measured times correspond to a mean of 5 calls of each procedure, based on serial processing. The employed computational platform was a laptop Intel Core i-5 2410M (2.3 GHz, 3 MB L3 cache), DDR3 RAM (6 GB, 1333 MHz), with Linux Ubuntu. MLP and ELM were coded in Python. The Linux wall-clock time function was used for time statistics.

One may note both ANNs are at least four orders of magnitude faster than MPB, taking a few milliseconds to compute the DRs of 2D and 3D PhCs. On the other hand, the DR computation for 3D structures hugely affects MPB performance, requiring a few minutes to be executed.

G. Discussion

MLP and ELM demonstrated an assured predicting performance by computing DRs of unknown PhCs presenting dissimilar geometric proportions and distinct DRs. Such results were accomplished by a simple ANN scheme−where ANNs only estimate the frequencies at a given **k** point−and using few training-PhC samples, which allowed the construction of small datasets. It culminated in relatively-small ANN models and fast ANN training (due also to efficient training algorithms). Furthermore, the results suggest both ANNs are able to compute at least the lowest modes simultaneously, providing means of gauging existing PBGs between them. Moreover, both ANNs presented similar MSE performances when processing test PhCs modified by proportions within and outside the training data range, showing that MLP and ELM were able to predict near but extrapolated geometric variations.

The greatest ANN advantage is the overall computing performance on a personal computer. MLP and ELM substantially reduced the runtime for processing PhC DRs regarding MPB. Combined with their fast training, both ANNs provided an efficient alternative for computing DRs, primarily for 3D structures which are generally computationally-costly and time-consuming. Although both ANNs presented similar predicting performances, ELM might be a more efficient procedure for this application, once its training has been faster, its MSE has been generally lower than MLP one, and its architecture has been simpler to define, given its unique hidden layer.

IV. Conclusion

In summary, MLP and ELM have potential to support electromagnetic solvers on the DR computation due to their generalization capabilities−prediction of unknown PhCs, especially structures whose geometric dimensions extrapolate the training data range−and extensibility for processing different 2D and 3D PhC types with distinct background materials. They could bring efficient alternatives for PhC design, once their relatively-small architectures and fast computing capabilities could, e.g., pave a way for a broader exploration of the search space in inverse problems, with substantial reduction of computational resources and runtime. In the future, the MLP and ELM's predicting capabilities should be investigated for higher order modes as well as their effectiveness for supporting adequately the design of photonic components.

References

[1] E. Yablonovitch, "Inhibited spontaneous emission in solid-state physics and electronics," *Phys. Rev. Lett.*, vol. 58, no. 20, pp. 2059–2062, 1987.

[2] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. Princeton, NJ, USA: Princeton Univ. Press, 2008.

[3] G. N. Malheiros-Silveira, V. F. Rodriguez-Esquerre, and H. E. Hernandez-Figueroa, "Strategy of search and refinement by GA in 2-D photonic crystals with absolute PBG," *IEEE J. Quantum Electron.*, vol. 47, no. 4, pp. 431–438, Apr. 2011.

[4] S. G. Johnson and J. D. Joannopoulus, "Block-iterative frequency-domain methods for Maxwell's equations in a plane-wave basis," *Opt. Express*, vol. 8, no. 3, pp. 173–190, 2001.

[5] S. S. Haykin, *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[6] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.

[7] D. Zibar, H. Wymeersch, and I. Lyubomirsky, "Machine learning under the spotlight," *Nature Photon.*, vol. 11, no. 12, pp. 749–751, 2017.

[8] L. Mescia, G. Fornarelli, D. Magarielli, F. Prudenzano, M. De Sario, and F. Vacca, "Refinement and design of rare earth doped photonic crystal fibre amplifier using an ANN approach," *Opt. Laser Technol.*, vol. 43, no. 7, pp. 1096–1103, 2011.

[9] A. da Silva Ferreira, C. H. da Silva Santos, M. S. Gonçalves, and H. E. Hernández Figueroa, "Towards an integrated evolutionary strategy and artificial neural network computational tool for designing photonic coupler devices," *Appl. Soft Comput.*, vol. 65, pp. 1–11, 2018.

[10] D. D. El-Mosalmy, M. F. O. Hameed, N. F. F. Areed, and S. S. A. Obayya, "Novel neural network based optimization approach for photonic devices," *Opt. Quantum Electron.*, vol. 46, no. 3, pp. 439–453, 2014.

[11] G. N. Malheiros-Silveira and H. E. Hernandez-Figueroa, "Prediction of dispersion relation and PBGs in 2-D PCs by using artificial neural networks," *IEEE Photon. Technol. Lett.*, vol. 24, no. 20, pp. 1799–1801, Oct. 2012.

[12] H. Men, N. C. Nguyen, R. M. Freund, K. M. Lim, P. A. Parrilo, and J. Peraire, "Design of photonic crystals with multiple and combined band gaps," *Phys. Rev. E*, vol. 83, no. 4, 2011, Art. no. 046703.

[13] H. Men, K. Y. K. Lee, R. M. Freund, J. Peraire, and S. G. Johnson, "Robust topology optimization of three-dimensional photonic-crystal band-gap structures," *Opt. Express*, vol. 22, no. 19, pp. 22632–22648, 2014.

[14] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Netw.*, vol. 4, no. 2, pp. 251–257, 1991.

[15] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.

Adriano da Silva Ferreira received the B.Sc. degree in information technology and the M.Sc. degree in electrical engineering in 2009 and 2013, respectively, from the University of Campinas, Campinas, Brazil, where he is currently working toward the Ph.D. degree in electrical engineering with the Department of Communications. His research interests include natural computing, electromagnetic simulators, and software engineering.

Gilliard Nardel Malheiros-Silveira was born in Caetité, Bahia, Brazil. He received the B.Sc. degree in system analysis from the Universidade Estadual da Bahia (UNEB), Bahia, Brazil, in 2007, the B.Sc. degree in electrical engineering from the Instituto Federal da Bahia, Salvador, Brazil, in 2008, and the M.Sc. degree in electrical engineering in 2010 and the Ph.D. degree in 2014, both from the University of Campinas, Campinas, Brazil. From 2015 to 2016, he was a Postdoctoral Researcher with CCH Optoelectronics Group, University of California at Berkeley, Berkeley, CA, USA. From 2017 to 2018, he was a researcher with the Centro de Tecnologia da Informação Renato Archer, Campinas, Brazil, an R&D unit of the Brazilian Ministry of Science, Technology, Innovation, and Communication. He became an Assistant Professor with São Paulo State University in 2018. He is an author and co-author of 14 papers in international journals, 2 book chapters, and 27 papers in international conferences. His current research interests include topics related to photonics, RF devices, and natural computing.

Hugo Enrique Hernández-Figueroa received the Ph.D. degree in physics from the Imperial College of Science, Technology, and Medicine, London, U.K., in 1992. Since 2005, he has been a Full Professor with the School of Electrical and Computer Engineering, University of Campinas. He has published more than 120 journal papers, 230 international conference papers, and 10 patents. His research interests concentrate on electromagnetic phenomena applied to integrated photonics, biosensors, nanophotonics, optical fibers, metamaterials, plasmonics, and antennas. He was an Associate Editor (Nanophotonics) for the IEEE Photonics Journal (March 2011–February 2018) and was an Associate Editor (Opto-Electronics/Integrated Optics) for the IEEE/OSA Journal of Lightwave Technology (January 2004–December 2009). He is a Co-Editor of the books *Localized Waves: Theory and Applications* (Wiley, 2008) and *Non-Diffracting Waves* (Wiley-VCH, 2013). He is a Fellow of the OSA class 2011, was a recipient of the IEEE Third Millennium Medal in 2000, and also the recipient of the Attilio Giarolas Medal in 2013 for outstanding research achievements from the Brazilian Microwaves and Optoelectronics Society (SBMO). Since 2014, he has been an Area Coordinator of Engineering with the prestigious Brazilian Agency FAPESP.