Inverse design and optimization of graphene metamaterial for multi-peak plasmon induced transparency based on machine learning and evolutionary algorithms

Tian Zhang, Qi Liu, Yihang Dan, Shuai Yu, Jian Dai, Xu Han, Xiaojuan Sun, and Kun Xu

*State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

Huawei Technologies Co., Ltd, Shenzhen 518129, Guangdong, China

School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

*Corresponding author: xukun@bupt.edu.cn

Article history: Received XX XXXXX 2019: Received in revised form XX XXXXXX 2019: Accepted XX XXXXXXX 2019: Available online XX XXXXX 2019

ABSTRACT

In this article, we propose an intelligent approach to achieve inverse design and performance optimization for the graphene metamaterial (GM) structure which consists of double layers graphene nanoribbons. Simulation results reveal that the multi-peak plasmon induced transparency (PIT) effect with wide bandwidth and large extinction ratios emerges in the transmission spectrum. And the simulated PIT effect has good agreement with the theoretical results based on transfer matrix method. More importantly, several simple regression algorithms (k nearest neighbour, decision tree, random forest and extremely randomized trees) based on machine learning have been applied in the spectrum prediction and inverse design for the GM structure. The comparison results demonstrate that the simple regression algorithms, such as random forest, have advantage in accuracy and efficiency compared with the artificial neural networks which have used to design the photonic devices in recent years. Besides, both single-objective optimization and multi-objective optimization (non-dominated sorting genetic algorithm-II) are employed in the performance optimization for the GM structure. Compared with previous works, we find that simple regression algorithms rather than artificial neural networks are more suitable for the design of photonic devices and multi-objective optimization can take many different performance metrics of photonic devices into consideration synthetically.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Owing to the wonderfully electronic and optical properties of high thermal conductivity [1], high damage threshold [2], strong third-order nonlinearities [3], wide bandwidth response [4], flexible tenability [5] and stable transparency (2.3%) in the visible range [6], graphene, a typical two-dimensional material, has been applied in an impressively large number of photonic devices, such as optical modulator [8], photoelectric detector [9], sensor [10], absorber [11], switching [12], polarization controller [13], diode [14] and so on. Especially, it has been demonstrated that graphene can support the excitation and propagation of surface plasmon polaritons (SPPs) with low propagation loss and high confinement [15-17]. Until now, based on unique properties of graphene and optical characteristics of SPPs, various nanostructures, such as graphene metamaterial (GM) [18-21], graphene nanoribbons [22-24] and graphene waveguide [25-27], have been proposed to construct ultra-compact plasmonic filters [25], perfect absorbers [23], sensor [24], logic gates [27] and so on. In these nanostructures, the periodically spaced graphene nanoribbons (GNRs) have attracted widespread attention because their relatively simple fabrication technique [22-24]. Moreover, in order to achieve steep optical characteristics, the electromagnetic induced transparency effect is combined with SPPs to introduce a new concept, plasmon induced transparency (PIT) effect, which has been demonstrated to construct high-performance photonic devices, for example optical filter, switching, polarization-insensitive sensor (parallel GNRs) and perfect absorber (crossed GNRs) [28-30]. It should be noticed that when we design the graphene nanostructures, the physical parameters of graphene should be determined to obtain desired performance. For example, we often consider the influence of the physical parameters (the chemical potential, relaxation times and thickness of graphene) on the transmission spectrum. However, the absence of the empirical relationship between the physical parameters of graphene ribbons and corresponding electromagnetic responses often enforces utilization of the time-consuming brute force search. It takes significant computational time and resources to compute the electromagnetic responses for all physical parameters of graphene by using numerical simulation methods, such as finite-difference time-domain (FDTD) and finite element method (FEM) [29, 30]. Actually, we can construct theoretical models to analyze the physical mechanisms behind the electromagnetic responses [31]. Then the physical parameters of graphene are determined to predict the electromagnetic responses of the graphene structure. However, for the complicated graphene nanostructures, the theoretical models are difficult to construct because the physical mechanisms are hard to interpret. In addition, although inverse design and performance optimization methods have been applied in the design of mode multiplexer [32], wavelength multiplexer [31], polarization beam splitter [33], polarization rotator [34], power splitter [35] and so on, previous researches pay little attention to graphene nanostructures, especially for the determination and optimization of the chemical potential for graphene nanoribbons.
Artificial neural networks (ANNs), deep learning in particular, have attracted a great deal of research attentions for an impressively large number of applications, such as image processing [36], natural language processing [37], acoustical signal processing [38], time series processing [39], self-driving [40], game [41], robot [42] and so on. Recently, ANNs have applications in the forward predict, inverse design and performance optimization of photonic devices by approximating electromagnetic responses instead of simulation methods [43-50]. For example, J. Peurifoy et al. found that the ANNs could be used to simulate the light scattering and inversely determine the thicknesses of multilayer nanosphere [43]. Here, the electromagnetic responses for all parameters of nanospheres were trained by using a small sampling of simulation results. The ANNs were appealing for designing and optimizing the photonic devices largely due to their ultra-fast prediction speed [44]. Many ANNs with different network architectures, such as shallow networks [43, 44], deep networks (such as ResNet-50) [45], tandem networks [46], bidirectional networks [47] and generative adversarial networks [48] had been proposed to design and optimize the complex photonics devices, such as metamaterials [47], gratings [44], meta-surface [48], power splitter [45], photonic crystals [49] and plasmonic devices [50]. In order to speed up the search process of training sets and optimize the network architectures of ANNs, ANNs were combined with the evolution algorithms to design the plasmonic waveguide system and micro-to-nano photonic couplers [50, 51]. And other machine learning algorithms, such as bayesian optimization and reinforcement learning, were used to design wavelength-selective thermal radiator and optical couplers [52, 53]. However, although the ANNs provided an effective approximation approach to replace the simulation methods, it had disadvantages in training time compared with simple machine learning methods, such as support vector machines (SVM) and random forest (RF). It had been proven that simple machine learning methods were more effective in some uncomplicated applications, especially for the modeling of the physical phenomena with less parameters [54, 55]. There was a lack of comprehensive analysis report for applications of simple machine learning methods in design and optimization of graphene structures. On the other hand, inverse design and performance optimization of photonic devices could be solved by some optimization algorithms that fell into two categories: gradient based methods and gradient free methods [50]. As a representative method of gradient based methods, although adjoint method could not only design linear optical devices but also optimize nonlinear devices in the frequency domain, it required physical background to derive the gradient of objective function [56, 57]. Evolution algorithms (such as genetic algorithms and particle swarm optimization) and search algorithms (such as direct-binary search), as gradient free methods, were used in inversely designing and optimizing of many photonic devices [33-35]. Evolution algorithms had advantages in simplicity and effectiveness, but they easily fell into local optimum and demanded significant computational time [58]. And no matter gradient based methods and gradient free methods, they usually optimized for a single performance metric, for example the coupling efficiency and transmittance at a particular wavelength [31-35]. Multi-objective optimization algorithms were rare to search for the most suitable structure parameters by considering multiple performance metrics synthetically.

In this article, we propose an intelligent approach to achieve inverse design and performance optimization for the GM structure.

The structure parameters of the GM structure are well-designed to obtain multi-peak PIT effect in the transmission spectrum. The theoretically calculated results based on transfer matrix method agree well with the simulated results. In addition, several simple regression algorithms based on machine learning are used to achieve spectrum prediction and inverse design for the GM structure. Compared with the ANNs which have been utilized to design the photonic devices in recent years, the simple regression algorithms have advantage in accuracy and efficiency. Moreover, the multi-objective optimization have been successfully employed in the performance optimization for the GM structure by taking many different performance metrics into consideration synthetically.

2. Device design and simulation results

As shown in Fig. 1, our proposed GM structure consists of two layer GNRs with alternative chemical potential. This double-layers GNRSs structure is periodically arranged and infinite along x (y) axis. The thin conductive layer covered on the bottom and top of silica layer ($n_{air}=1.45$) forms as electrodes to alternatively apply voltage $V_{1}$ ($V_{2}$) and $V_{2}$ ($V_{1}$) on the GNRS 1 (2), leading to the graphene ribbons of two GNRSs with alternative chemical potential ($\mu_{1}$ and $\mu_{2}$ for GNRS 1, $\mu_{3}$ and $\mu_{4}$ for GNRS 2). It should be noticed the thickness of the silica insulated layer embedded in the structure is set large enough, so that the top and bottom conductive layers don’t affect the optical transmission property under normal incidence. As a result, we can leave the top and bottom conductive layer out of account in our simulation about analyzing the electromagnetic wave transmission property. Here, the period of the GNRSs (GNRS 2) is set as $A_{1}$=400 nm ($A_{2}$=200 nm), and the width of the graphene ribbon in GNRS 1 (GNRS 2) is $w_{1}$=350 nm ($w_{2}$=175 nm), leading to a duty cycle of $r_{1}$=0.875 ($r_{2}$=0.875). In the simulation of the GM structure, we employ the Kubo formula to model the conductivity of the monolayer graphene ribbons [17]:

$$
\sigma_g = i\frac{e^2k_BT}{\pi\hbar^2}(\omega+ir\tau)^{-1} \left[ \frac{\mu_x}{k_BT} + 2\ln\left(\exp\left(-\frac{e}{k_BT}\right) + 1\right) \right] 
+ i\frac{\hbar}{4\pi\hbar} \left[ \frac{2}{\mu_x} \right] \right] 
$$

where $k_B$, $T$ ($=300$ K), $\hbar$, $\tau$ ($=0.5$ ps), $\mu_x$, $e$, and $\omega$ represent the Boltzmann’s constant, temperature, reduced Planck’s constant, relaxation time, chemical potential, electron charge and angular frequency, respectively. For a few layers (<6) of graphene nanoribbons, the conductivity can be expressed as $\sigma_{g}=N\sigma_{g}$, where $N$ is the number of layers [13]. In the analytical wavelength
region, the simplified conductivity can be achieved by considering the domination of the interband electron-photon process and $\mu_0 \gg k_0 T$

$$\sigma_n = i \frac{Ne^2 \mu_e}{\pi \hbar c (\omega + i \tau^{-1})}$$  (2)

In order to analyze the excitation condition of the SPPs in proposed GM structure, the dispersion equation of SPPs is retrieved based on the Maxwell equation and continuous boundary condition [17]

$$\frac{\varepsilon_1 \sqrt{\beta_{SPP}^2 - \varepsilon_1 \omega^2}}{c^2} + \frac{\varepsilon_2 \sqrt{\beta_{SPP}^2 - \varepsilon_2 \omega^2}}{c^2} = \frac{i \sigma_{SPP}}{\omega \varepsilon_0}$$  (3)

where $\beta_{SPP}$ relates to the propagation constant of SPPs, $c$ represents the light speed in vacuum, $\omega$ is the dielectric constant of free space, $\varepsilon_1$ and $\varepsilon_2$ are the effective permittivity of the medium on each side of graphene layers ($\varepsilon_1$ and $\varepsilon_2$ are both equal to $\varepsilon_{SiO2} = 3.9$, $\varepsilon_{Si} = 12.35$), because the GNRs are surrounded by the same medium in our proposed GM structure). Here, since the solution satisfies $\beta_{SPP} > \omega / c$, the effective refractive index of SPPs deduced from Eq. (3) is given by [16]

$$n_{SPP} = \frac{\beta_{SPP}}{k_0} = \frac{2\varepsilon_0 \varepsilon_{SPP} \pi \hbar c}{Ne^2 \mu_e} \left( \omega + i \tau^{-1} \right)$$  (4)

where $k_0$ relates to the wave vector of vacuum. As shown in Fig. 2(a), the dispersion curves (solid lines) of few layers graphene nanoribbons surrounded by silica match well with the dispersion data (circle marks) obtained from the mode solution simulation. In our simulation, monolayer and multilayer graphene are both treated as a surface with electric conductivity $\sigma_{GSM}$ since the graphene nanoribbons are ultrathin. It can be found that both the real part and imaginary part of effective refractive indices for the SPPs on graphene with different layer numbers decrease with the increasing of wavelength. Thus, when increasing the wavelength of incident light, the field confinement of the SPPs on graphene layers become weaker and the propagation loss of SPPs becomes smaller. Moreover, the field confinement of the SPPs on graphene layers become weaker with the increasing of layer number of graphene. In order to increase the interaction of upper and lower GNRs, we set the layer number of graphene layer as $N=4$ in the following article.

Firstly, as shown in Fig. 2(c), we analyze the transmission of a single grating layer composed of graphene ribbons. For the grating with small value of duty cycle ($w/l$), the graphene plasmons (GPs) on a graphene ribbon can hardly interact with that in the adjacent graphene ribbons. As a result, the propagation of GPs in the grating can be equivalently substituted by that in a single graphene ribbon. It has been reported that the GPs in a single graphene ribbon are nearly totally reflected at the boundary together with a phase jump of $\phi = 0.27 \pi$ [59]. Thus, it can be considered that the GPs excited on graphene ribbons are caused by Fabry-Perot (FP) like resonances, which requires that

$$\text{Re}(n_{SPP}) = k_0 w + \phi = m \pi, \quad m = 1, 2, 3, \ldots$$  (5)

Substituting Eq. (4) into Eq. (5), the resonance wavelengths of GPs on graphene ribbon can be achieved as following,

$$\omega = \sqrt{\frac{(m - \phi)Ne^2 \mu_e}{2\varepsilon_0 \varepsilon_{SPP} \hbar^2 w}}, \quad m = 1, 2, 3, \ldots$$  (6)

As shown in Fig. 2(d), the resonance curves (three dashed lines) for three modes of a graphene grating agree with the simulated results calculated by the FDTD method (the absorption contour patterns). Obviously, the comparison results verify the effectiveness of both the theoretical FP model and the FDTD simulated method. Here, we only calculated the resonance curves for the odd modes in Fig. 2(d) since the even modes cannot be excited by normal incident wave [59].

In order to analyze the mechanism of the multi-peak PIT effect that would emerge in transmission spectrum, the optical responses of the GM structure which only contains the upper and lower GNRs
are calculated by using the FDTD method, respectively. We employ the 2D simulations whose x direction and other boundaries are set as periodic boundary conditions and perfectly matched layer to model the GM. The Fermi levels of the graphene ribbons are set to \( \mu_{\text{GC}} = 0.7 \) eV, \( \mu_{\text{GPs}} = 0.5 \) eV, \( \mu_{\text{Ge}} = 0.15 \) eV and \( \mu_{\text{Ge}} = 0.75 \) eV in our simulations. When TM polarized light normally illuminates on the GM, it can be found two remarkable dips emerge in the transmission spectrum of the GM that includes only the upper (blue dashed line) GNRs and lower (green dashed line) GNRs [Fig. 3(a)]. Here, the reasons for the appearances of the dips whose central wavelength are 7.04 (5.30) \( \mu \)m and 10.40 (13.16) \( \mu \)m are related to the excitation of the SPPs modes on the upper (lower) GNRs. Next, we proceed to consider the optical characteristics of the complete GM that includes double layer GNRs as shown in Fig. 1. From the red line shown in Fig. 3(a), two pronounced transmission peaks respectively located between two dips emerge in the transmission spectrum, which indicates the appearance of the multi-peak PIT effect [4]. Here, the multi-peak PIT effect have wide bandwidth and large extinction ratios thus it can be applied in the optical switching and slow light [59, 60]. The optical characteristics of the dips in the multi-peak PIT effect are similar to those of the single layer GNRs, which suggests that the appearance of the dips are attributed to the excitation of the SPPs mode on graphene ribbons. Here, the formation of the multi-peak PIT effect is explained by the normalized magnetic field distribution of the transmission peaks (B and E) and dips (A, C, D and F). As shown in Fig. 3, it can be observed that the appearances of the dips A, C, D and F are related to the excitation of SPPs on the graphene ribbons 1, 2, 3 and 4, respectively. While it’s the coupling between the SPPs mode on graphene ribbons 1 (2) and 3 (4) gives rise to the transmission peaks B (D) in the multi-peak PIT effect. In order to model the dynamic transmission of the double GNRs, we employ the transfer matrix method to explain the physical phenomenon. The transfer matrix can be defined as [60]

\[
H = M_1 S_{12} M_2 \tag{7}
\]

where \( M_1, S_{12} \) and \( M_2 \) represent the matrices of GRA1, silica layer and GRA2, respectively. They are governed by

\[
S_{12} = \begin{pmatrix}
\cos \phi & 0 \\
0 & e^{-i \phi}
\end{pmatrix}, \quad M_q = \frac{1}{r_{12}} \begin{pmatrix}
t_{12} t_{21} & r_{21} t_{21} \\
-r_{12} & 1
\end{pmatrix}, \quad q = 1, 2 \tag{8}
\]

Here, under light normally illuminates on the GM, the Fresnel coefficients in matrix \( M_q \) are expressed as \( t_{12} = t_{21} = 2n_{\text{SiO2}}/2n_{\text{SiO2}} + Z_{i\sigma}Z_{o\sigma} \), \( r_{12} = r_{21} = Z_{i\sigma}Z_{o\sigma}/(2n_{\text{SiO2}} + Z_{i\sigma}Z_{o\sigma}) \), where \( Z_0 = 367.73 \Omega \) represents the vacuum impedance and \( \phi = d_{\text{SiO2}} \) is the phase difference between the GRA1 and GRA2. Under the condition of quasi-static approximation, the average sheet conductivity \( \sigma_q \) is given by

\[
\sigma_q = \frac{1}{2} \begin{pmatrix}
r_{12} e^{i \mu_{\text{GC}}} N_{\omega} \\
r_{12} e^{-i \mu_{\text{GC}}} N_{\omega}
\end{pmatrix} \quad + \quad \frac{1}{2} \begin{pmatrix}
p_2 e^{i \mu_{\text{GC}}} N_{\omega} \\
p_2 e^{-i \mu_{\text{GC}}} N_{\omega}
\end{pmatrix} \quad + \quad \frac{1}{2} \begin{pmatrix}
p_2 e^{i \mu_{\text{GC}}} N_{\omega} \\
p_2 e^{-i \mu_{\text{GC}}} N_{\omega}
\end{pmatrix} \quad + \quad \frac{1}{2} \begin{pmatrix}
p_2 e^{i \mu_{\text{GC}}} N_{\omega} \\
p_2 e^{-i \mu_{\text{GC}}} N_{\omega}
\end{pmatrix}
\]

(9)

Here, \( \omega_0 \) is the resonance frequency which is calculated by using Eq. (6) for graphene ribbons with different \( \mu_{\text{GPs}} \) (j = 1, 2, 3, 4). And the GPs resonance width of graphene ribbon \( \Gamma_j \) is usually 10% lower than the Drude scattering width \( \Gamma_j = e v F (\mu_{\text{GPs}}) \) in unpatterned graphene, where \( v_F = 300 \) is the Fermi velocity and \( \mu = 10000 \) cm\(^2\)V\(^{-1}\) is the DC mobility [60]. The phase factor \( \Phi = m \phi \) (m = 1, 2, 3, 4...) is a fitting parameter deduced from the FDTD simulation. According to Eqs. (7)-(9), the transmittance of the GM structure can be expressed as

\[
T = \left| \frac{4n_{\text{SiO2}}^2}{2n_{\text{SiO2}} + Z_0 \sigma_q} \right|^2 \tag{10}
\]

According to Eq. (10), the theoretical transmission spectrum of the complete GM structure is shown by the purple dashed line in Fig. 3. It can be found that the theoretical transmission spectrum has good agreement with the simulated transmission spectrum (solid red line) when the fitting parameters \( \phi_j \) are fitted as \( \phi_j = 3.77, \phi_j = 0.85, \phi_j = 5, \phi_j = 0.45 \).

We also analyze the influence of the structure parameters in the GM on the transmission spectrum, and the comparative results are exhibited in Fig. 4. As shown in Fig. 4(a), the increasing of chemical potential \( \mu_{\text{GC}} \) and \( \mu_{\text{GPs}} \) lead to the decreasing of resonant wavelengths of the dips C and D in the transmission spectrum. Here, since the generations of the dips C and D are mainly originated from the SPPs modes on the GNRs with \( \mu_{\text{GC}} \) and \( \mu_{\text{GPs}} \), which can be confirmed by the normalized magnetic distribution of the points C and D in Fig. 3. And the increasing of \( \mu_{\text{GC}} \) will cause the decreasing of the effective refractive index of the SPPs mode on graphene ribbons, leading to the decreasing of resonance wavelength according to Eq. (6) [60]. These two reasons can clearly explain the variation of transmission spectrums in Fig. 4(a). Similarly, as shown in Fig. 4(b), the blue-shifting of the dips A and D with the increasing of \( \mu_{\text{GC}} \) are induced by the SPPs resonance on the GNRs with \( \mu_{\text{GPs}} \) and \( \mu_{\text{GPs}} \), respectively. In addition, as shown in Fig. 4(c), when the gap \( d_b \) between two layer GNRs is increased from 100 nm to 300 nm, the dips C and F move to the short-wavelength direction and are more sensitive than the dips A and D. As the field distributions of dips shown in Fig. 3, the SPPs resonance of the dip C (F) interact with the vertical graphene ribbons more strongly compared with that of the dip A (D), leading to a significant influence of the gap \( d_b \) on the resonant wavelength of
the dip C (F) than that of the dip A (D). Moreover, as shown in Fig. 4(d), the resonant wavelengths of the transmission peak E and dips D, F increase significantly when the duty cycle increases from 0.85 to 0.90. The reason for this phenomenon is as following: as shown in the field distribution in Fig. 3, for the graphene ribbons with the same width in the GNRs 1, the resonant SPPs mode for the dip D interacts with the horizontal graphene ribbon stronger than that for the dip C. As for the graphene ribbons with the same width in the GNRs 2, the SPPs mode for the dip F [Re(\(\Delta \omega\))=22.95] is confined more tightly than that for the dip A [Re(\(\Delta \omega\))=10.83]. As a result, the resonant wavelengths of the dips D and F are more sensitive to the duty cycle than that of the dips A and C. And according to Eq. (6), the resonant wavelengths of all the dips red-shift with the increasing of the duty cycle.

3. Spectrum prediction and inverse design of the GM structure

As shown in Fig 4, it can be found that the slight changes of the structures parameters in the GM structure have significant influence on the transmission spectrum. If we want to discover the potential relationship between the structure parameters and transmission spectrum, it requires a high computational cost to traverse hundreds of the structure parameters. Although we have used the 2D FDTD simulation to reduce the calculation time, it requires several minutes to guarantee the convergence of algorithm for a group of structure parameters. Actually, we can use Monte Carlo method or interval sampling method to reduce simulation time, but it leads to the loss of accuracy due to interpolation and fitting. Another way to improve the efficiency is to train a model based on the machine learning algorithms by using a small part of simulation results [43-50]. It has been demonstrated that the ANNs-based models could equivalently replace the electromagnetic simulation method for some photonic structures [43-50]. It should be noted that the inference time of the ANNs-based model is far less than that of the simulation method once the model is constructed [43-50]. However, the collection of the training data and data cleaning also need considerable time and effort. Here, the prediction process for electromagnetic response based on the data-driven model is known as ‘forward prediction’ [50]. On the other hand, we often face the inverse design problems in the design of photonic devices [43]. The goal of inverse design is to discover the most suitable structure parameter for a particular transmission spectrum. Generally speaking, the inverse design problems can be transferred to optimization problems which can be solved by using the gradient based method (such as adjoint method) or gradient free method (such as genetic algorithm (GA)) [50]. The machine learning algorithms, especially for the ANNs, are also used in solving the inverse design problems for photonic devices. Similar to forward prediction, a machine learning based model is trained to predict the corresponding structure parameters according to a given transmission spectrum [50]. It has been demonstrated that the ANNs-based methods are competitive with the gradient free methods for an achievable transmission spectrum in reality [50]. Notably, if the targeted transmission spectrum is not achievable, the gradient free method can return the structure parameters that generate similar transmission spectrum, while the ANNs-based method may return a relatively absurd result.

Notably, the principles behind the forward prediction and inverse design based on machine learning are date regression between the structure parameters and transmission spectrum. It indicates that the labels of the training data are continuous variables rather than discrete variables. There are several machine learning algorithms can be used in date regression except for the ANNs. Compared with the other simple machine learning algorithms, such as support vector machine (SVM) and random forest (RF), the training cost of the ANNs-based method is much higher because the ANNs-based model is more complex. However, until now, only the ANNs-based methods are utilized in the design and optimization for photonics devices. There is no doubt that for the complicated photonic devices, such as a resonator with random distribution [45], the ANNs-based method, especially for the deep ANNs (deep learning) is an effective method to construct the complex relationship between the input and output. However, for the uncomplicated application scenarios, such as the design and optimization of a photonics device that contains several structure parameters (<15), the ANNs-based method may not be the best choice. For example, it has been demonstrated that the SVM-based model performs better than ANNs-based model in the trend prediction of soil organic carbon and river flow [54, 55]. In addition, the selection of hyper-parameters in the training of an ANNs-based model (layers, activation function, solver, learning rate, batch size and so on) is more complex than that of the SVM-based model and RF-based model. In particular, we may even need to find the optimal hyper-parameters for ANNs by using the evolutionary algorithms [50]. Moreover, the training and inference time of the ANNs significantly exceeds that of the simple machine learning algorithms [54]. The reason for this is attributed to the training of a ANNs-based model usually requires a complex process consisting of data forward-propagation, error back-propagation and stochastic gradient decent of gradients, while the training of simple machine learning algorithms is relative easy [55]. In order to overcome the defects of the ANNs-based methods as mentioned above, we use several different regression algorithms based on machine learning to complete the forward prediction and inverse design of the GM structure. For example, Similar to k nearest neighbour (KNN) classification, KNN regression calculates the distances between the targeted instance and each training instance and then selects the most similar k data as the candidate set to determine the results [61]. And three kinds of tree-based regression methods, decision tree (DT), RF, extremely randomized trees (ERT), are also used in the design of the GM structure. These tree-based regression algorithms includes the same steps, such as selecting splits and selecting the optimal tree [62]. And the difference is that the RF regression is an ensemble algorithm based on bagging method that involves several regression trees [63]. Compared with the RF regression, the split of features for the ERT is more random, leading to the reduction of the variance for the trained model [64].

First of all, we attempt to use the regression algorithms to replace the traditional FDTD simulation in the forward prediction. From Fig. 5(a), it can be found that the regression algorithms take the structure parameters of the GM structure as algorithmic input and predict the transmission spectrums correspondingly. For instance, the potential relationships between the chemical potentials of graphene ribbons \(\mu_{a1}, \mu_{a2}, \mu_{a3}\) and \(\mu_{a4}\) and the transmittances in transmission spectrum are taken into consideration here. In order to train the regression models, we use the repeatable 2D FDTD simulations combined with the MC method to generate training sets because the regression algorithms belong to supervised learning [65]. Here, each instance in 20,000 training instances includes 4 structure parameters (\(\mu_{a1}, \mu_{a2}, \mu_{a3}, \mu_{a4}\) and 200 transmittances which are unevenly sampled from
training set to train several other regression algorithms. Fig. 5(c) exhibits the training time and accuracies for different regression algorithms. Surprisingly, it can be observed that the scores of all regression algorithms are greater than 91, which indicates that other regression algorithms are competitive with the ANNs in forward spectrum prediction. And although the model of the ANNs is more complex intuitively, the accuracy (score) of the RF (96) outperforms that of the ANNs (95). In order to illustrate the effectiveness of the regression models vividly, we compare the transmission spectrums predicted by the regression algorithms and simulated by the FDTD method. We randomly select a group of structure parameter in the test set and calculate the transmission spectrums based on machine learning and electromagnetic simulation. As shown in Fig. 5(d), the transmission spectrums predicted by the regression algorithms agree with the FDTD simulation results obviously. And compared with the transmission spectrum predicted by the ANNs, the results predicted by the RF are closer to the ground truth (the FDTD simulated transmission spectrum). More importantly, the training cost of the ANNs (36 seconds) far exceeds those of other regression algorithms (<5 seconds) when the hyper-parameters for regression algorithms are determined. With a comprehensive consideration of training cost and accuracy, the RF regression algorithm is a more appropriate method to complete the forward spectrum prediction for the GM structure compared with the ANNs obviously.

Similar to the spectrum prediction, the regression algorithms mentioned above can be employed in the inverse design for the GM structure. Contrary to the spectrum prediction, Fig. 6(a) shows the diagram of the regression algorithms used in the inverse design for the GM structure. It can be found that the inputs (outputs) of the models trained by the regression algorithms are the transmittances in transmission spectrum (structure parameters of the GM structure). It should be noted that there is no need to generate new training instances, we use the same training instances to train the regression algorithms by converting the inputs (outputs) to outputs (inputs) in the spectrum prediction reversely. The training time and accuracies for all regression algorithms in the inverse design are exhibited in Fig. 6(b). It can be observed that the all regression algorithms can achieve excellent performance (scores for all regression algorithms are more than 90) and the score of the DT (93.0) is lower than that of the ANNs (97), ERT (96), KNN (96.5) and RF (98). In order to validate the effectiveness of the regression algorithms in the inverse design, we randomly select a transmission spectrum from the test set and input it into the model trained by the regression algorithms. The structure parameters (chemical potentials $\mu_1, \mu_2, \mu_3$ and $\mu_4$ for the graphene ribbons in the GM structure) predicted by regression algorithms and the ground truth are shown in Fig. 6(c). We can observe that the predicted chemical potentials $\mu_1, \mu_2, \mu_3$ and $\mu_4$ are close to the real chemical potentials (red dashed line) in the GM structure, conforming the effectiveness of the regression algorithms. In addition, we also use the chemical potentials predicted by the regression algorithms to simulate the GM structure based on the FDTD methods. As shown in Fig. 6(d), the accuracy of the RF regression algorithm outperforms the accuracies of the DT, ANNs, ERT and KNN because of the high similarity. More importantly, the training time of the RF regression algorithm (6 seconds) is lower than that of the ANNs (34 seconds). Obviously, the calculated results shown in Fig. 6(b)-(d) indicate that the ANNs is not the best choice in the inverse design for the GM structure. And the RF regression algorithm outperforms ANNs in accuracy and efficiency.
where \( \lambda, \lambda_{\text{min}} (\lambda_{\text{max}}), S(\lambda) (S(\hat{\lambda})) \) are the wavelength, minimum (maximum) wavelength of the wavelength range and targeted (optimized) transmission spectrum, respectively. And after that, the individuals of population are sorted according to the fitness in descending order. (iii) Trying to generate a new population by using the standard selection, crossover and mutation procedures. In the selection process, two parent individuals are selected from the previous generation based on the roulette-wheel selection method or tournament strategy [18]. Here, the structure parameters with better fitness are selected with higher probability. To maintain the diversity of population and keep some superior individuals, some percentage of the superior (inferior) structure parameters are kept in the next generation. In the crossover process, the structure parameters are converted into the binary values firstly. It should be noted that the conversion of decimal to binary is likely to result in the loss of digital precision. The optimization variables (structure parameters) of parent individuals cross over to generate a new population based on the uniform crossover algorithm or single-point crossover (xovsp) [67]. In the mutation process, each element in the binary number has 5% probability to flip from 0 (1) to 1 (0). After converting the optimization variables (structure parameters) from binary number to decimal number, a new population is generated. (iv) The fitness of the newly generated population are evaluated to determine the optimization process whether stop or not. If the generation of the structure parameters evolve for 1000 times or the optimization objective remain unchanged for more than 5 generations, the GA stops, otherwise, proceeds to Step (ii). Quantum genetic algorithm (QGA) is a new parallel evolutionary algorithm which combines with the traditional GA and quantum algorithms [68]. In the QGA, the encode method for the variable is quantum bit rather than binary number. And in the crossover and mutation processes, QGA uses the quantum rotation gate to update the individual. Similar to the GA, the PSO is an evolutionary algorithm which is suitable for the decimal number rather than binary number [35]. The generation of initial population for the PSO is the same as that of the GA. However, the generation of new population for the PSO is not through selection, crossover and mutation procedures. It means that there no need to convert the decimal number to binary number in the PSO which can effectively avoid the loss digital precision. For the PSO, the individuals in the population depend on the globally optimal individual and historically optimal record for each individual to search for the optimal solution [35]. Similarly, when we use the PSO to optimize the GM structure, each individual in the population searches for the optimal structure parameters by synthetically considering the currently optimal structure parameters and individually optimal structure parameters. The evolution of the structure parameters is controlled by a specified velocity [35]:

\[
V_i^{k+1} = W V_i^k + c_1 r_1 \left( p_b^k - X_i^k \right) + c_2 r_2 \left( g_{b_i}^k - X_i^k \right)
\]

where \( i \) is the \( i \)th structure parameter in the population, \( k \) is the iteration number, \( W \) is variable inertia weight, \( c_1, c_2 = 1.49445 \) are acceleration constants, \( r_1 (r_2) \) is random value between 0 and 1, \( g_{b_i} \) is the globally optimal structure parameters, \( X_i \) and \( p_b^i \) are current structure parameter and individually optimal structure parameter for the \( i \)th structure parameter in the \( k \)th iteration, respectively. The \( i \)th structure parameter in the population is updated according to following equation:

4. Optimization of the GM structure

Similar to the inverse design, the machine learning algorithms can also be used in optimizing the GM structure. On the one hand, as same as the inverse design, the transmission spectrum with wide wavelength range can be comprehensively optimized by inputting into the machine learning algorithms. On the other hand, we can optimize the GM structure for single or several performance metrics, such as the transmittance at a given wavelength and the bandwidth of a transparency window. As the typical gradient free methods, the GA and particle swarm optimization (PSO) have been applied in optimizing for various photonic devices because of simplicity and effectiveness [34, 35]. Here, we compare the optimization effects between different evolutionary algorithms. And the gradient based methods, such as adjacent method [57] and objective first method [56], are not considered in this article. The algorithmic details of the GA are outlined as follows: (i) randomly generating an initial population consisted of N=40 individuals. Each individual has four structure parameters, namely, the chemical potentials of graphene ribbons (\( \mu_1, \mu_2, \mu_3, \mu_4 \)). Here, all structure parameters are initialized in different ranges specified by minimum and maximum values 0.6 eV<\( \mu_i \)<0.8 eV, 0.4 eV<\( \mu_i \)<0.6 eV, 0.05 eV<\( \mu_i \)<0.25 eV and 0.6 eV<\( \mu_i \)<0.8 eV. (ii) For the generated N group of structure parameters, the transmission spectrums are simulated by using the FDTD method. And different performance metrics, such as the transmittance at a given wavelength, are regarded as the fitness or the optimization objective for GA. If we want to optimize the transmission spectrum with wide bandwidth, the fitness can be defined as

\[ F = \sum_{\lambda_{\text{min}}}^{\lambda_{\text{max}}} [S(\lambda) - S(\lambda)] \]
\[ X_{i}^{t+1} = X_{i}^{t} + V_{i}^{t+1} \]  

(7)

In order to avoid the premature problem, the velocities of evolution are limited to a certain range (-1~1). Finally, the evaluation of the newly generated population is the same as that of the GA. If the population does not meet the termination conditions, the velocities of all structure parameters are calculated based on Eq. (6) in the next iteration. To compare the optimization effects of the GA, QGA and PSO, we randomly select a complete transmission spectrum (red dashed line in Fig. 7(c)) in test dataset as the optimization objective. For all optimization algorithms, the degree of approximation (Eq. (5)) between the targeted transmission spectrum and the optimized transmission spectrum is treated as the fitness for the GA, QGA and PSO. Fig. 7(a) exhibits the fitnesses of the GA, QGA and PSO for different generations in the optimization of the targeted transmission spectrum. It can be observed that the fitnesses of the GA, QGA and PSO are gradually close to 0, which indicates these single-objective optimization algorithms are convergence. And the convergence speeds of the GA and PSO are faster than that of the QGA. Here, in the 100th generation, we select the optimized chemical potentials for all optimization algorithms and compare them with the ground truth. It can be found in Fig. 7(b) that the chemical potentials optimized by the GA, QGA and PSO agree well with the targeted chemical potentials. In Fig. 7(c), the green solid line and blue solid line are the optimized transmission spectrums in the first generation and 100th generation, respectively. Obviously, the optimized transmission spectrums in the first generation (green solid lines) are randomly generated and those in the 100th generation (blue solid lines) are close to the targeted transmission spectrums.

Finally, the GM structure is also optimized for multi-objective metrics, such as several transmittances at different wavelengths. The steep degree of the PIT effect is a critical performance indicator which affects the bandwidth, group index, figure of merit and so on. To achieve more steep optical characteristics, we use a famous multi-objective optimization algorithm, non-dominated sorting genetic algorithm-II (NSGA-II), to optimize the transmittances at a certain number of wavelengths. Compared with other multi-objective optimization algorithm, NSGA-II finds the pareto optimal solution based on the fast nondonated sorting method (FNSM) and elitist strategy [69]. In the NSGA-II, the crowding distances of the individuals and the levels calculated by the FNSM are combined to jointly determine the order of individuals [69]. For all performance indicators, the individuals in the lower level are better than those in the higher level, while the individuals in the same level are incommensurable. In our simulations, the steep degree of the PIT effect is simply characterized as the differences of the transmittances between the transmission peaks and dips. The algorithmic details of the NSGA-II are outlined as follows: (i) the generation of initial population for the NSGA-II is the same as that of the GA, QGA and PSO. Here, each individual has seven structure parameters, namely, the chemical potentials of graphene ribbons (\( \mu_{c1}, \mu_{c2}, \mu_{c3}, \mu_{c4} \)), the duty cycle of graphene ribbons (\( r_1, r_2 \)) and the distance \( d_p \) between the GNRF and GNR2. Here, all structure parameters are initialized in different ranges 0.6 eV<\( \mu_{c1} <0.8 \) eV, 0.4 eV<\( \mu_{c2} <0.6 \) eV, 0.05 eV<\( \mu_{c3} <0.25 \) eV, 0.6 eV<\( \mu_{c4} <0.8 \) eV, 0.7<\( r_1 <0.9 \), 0.7<\( r_2 <0.9 \) and 100 nm<\( d_p <500 \) nm. (ii) The differences between the transmittances at different wavelengths are regarded as the fitness for the NSGA-II. It indicates that two differences and four differences between the transmission peaks and dips are calculated for the single PIT effect and double PIT effects, respectively. Unlike the GA, QGA and PSO, the levels of the individuals in the population for the NSGA-II are determined by using the FNSM. And the crowding distances are calculated for the individuals in the same level to maintain the diversity of the population. The individuals in the population are sorted according to the levels and crowding distances [69]. (iii) The generation process of a new population for the NSGA-II is the same as that of the GA, QGA and PSO. (iv) The individuals in the newly generated population are placed into the old population to generate a large population. And the individuals in the large population are sorted based on the FNSM and crowding distances. Finally, top \( N \) individuals are selected to generate the new population for the next iteration based on the elitist strategy. (v) The evaluation of the newly generated population for the NSGA-II is similar to the GA and PSO. And the best individual in the pareto front is selected as the solution of the NSGA-II. Fig. 7(c) and Fig. 7(d) exhibit the multi-objective optimization results for single-peak PIT effect and multi-peak PIT effect, respectively. The optimization objective for single-peak PIT effect is two differences between a transmission peak and two dips, while that for multi-peak PIT effect is four differences between two transmission peak and four dips. After 100 iterations, we can found the differences between the transmission peaks and dips reach to 0.76 and 0.97 (0.87, 0.83, 0.79 and 0.69) for the single PIT effect (multi-peak PIT effect), indicating the NSGA-II is effective for the optimization of the GM structure. Obviously, the multi-objective optimization can be used to achieve steep optical characteristics by synthetically considering several performance metrics.

---

**Fig. 7.** (a) The fitnesses of the GA, QGA and PSO for different generations in the reverse design. (b) Optimization results of the chemical potentials for the GA, QGA and PSO in the 100th iteration. (c) The optimized transmission spectrums of the GA, QGA and PSO algorithms in the first iteration (green line) and the 100th iteration (blue line). (d) The multi-objective optimization results for two differences between one peak (8161 nm) and two dips (7659 nm and 11620 nm). (f) The multi-objective optimization results for four differences between two peaks (6110 nm and 12620 nm) and four dips (5150 nm, 6890 nm, 10310 nm and 13220 nm).
5. Conclusion

In this article, we propose an intelligent approach to achieve spectrum prediction, inverse design and performance optimization for the GM structure. The structure parameters of the GM structure are well-designed to obtain multi-peak PIT effect in the transmission spectrum. And the theoretically analyzed results based on transfer matrix method agree well with the simulation results. In addition, several simple regression algorithms based on machine learning are used to achieve spectrum prediction and inverse design for the GM structure. Compared with the ANNs which have been utilized to design the photonic devices in recent years, the simple regression algorithms have advantage in accuracy and efficiency. Moreover, the multi-objective optimization have been successfully employed in the performance optimization for the GM structure by taking many different performance metrics into consideration synthetically. This work not only paves a new way towards the realization of intelligent design for graphene based devices, but also has important applications in other advanced materials and metamaterial.

Acknowledgements

This work is supported by the national Natural Science Foundation of China (Grant No. 61625104, No. 61431003); The Beijing Municipal Science & Technology Commission (Grant No. Z181100008918011); The Fundamental Research Funds for the Central Universities (Grant No. 2019RC15, No. 2018XKJC02); National Key Research and Development program (Grant No. 2016YFA03010300).

6. references

[1] A. A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, and C. N. Lau, “Superior thermal conductivity of single-layer graphene,” Nano Lett. 8, 902-907 (2008).
[2] A. Roberts, D. Cormode, C. Reynolds, and T. Newhouse-Illige, “Response of graphene to femtosecond high-intensity laser irradiation,” Appl. Phys. Lett. 99, 197 (2011).
[3] E. Hendry, P. J. Hale, J. Moger, A. Savchenko, and S. Mikhailov, “Coherent nonlinear optical response of graphene,” Phys. Rev. Lett. 105, 097401 (2010).
[4] Q. Bao and K. P. Loh, “Graphene photonics, plasmonics, and broadband optoelectronic devices,” ACS nano 6, 3677-3694 (2012).
[5] T. Zhang, J. Dai, Y. Dai, Y. Fan, X. Han, J. Li, F. Yin, Y. Zhou, and K. Xu, “Tunable plasmon induced transparency in a metallo-dielectric grating coupled with graphene metamaterials,” J. Lightwave Technol. 35, 5142-5149 (2017).
[6] R. R. Nair, P. Blake, A. N. Grigorenko, K. S. Novoselov, T. J. Booth, T. Stauber, N. M. Peres, and A. K. Geim, “Fine structure constant defines visual transparency of graphene,” Science 320, 1308-1308 (2008).
[7] A. K. Geim and K. S. Novoselov, “The rise of graphene,” Nat. Mater. 6, 183-191 (2007).
[8] M. Liu, X. Yin, E. Ulin-Avila, B. Geng, T. Zentgraf, L. J. Fu, W. Wang, and X. Zhang, “A graphene-based broadband optical modulator,” Nature 474, 64-67 (2011).
[9] T. J. Echtermeyer, P. Nene, M. Trushin, R. V. Gorbachev, A. L. Eiden, S. Milana, Z. Sun, J. Schliemann, E. Lidorikis, and K. S. Novoselov, “Photothermoelectric and photoclectric contributions to light detection in metal–graphene–metal photodetectors,” Nano Lett. 14, 3733-3742 (2014).
[10] S.-H. Bae, Y. Lee, B. K. Sharma, H.-J. Lee, J.-H. Kim, and J.-H. Ahn, “Graphene-based transparent strain sensor,” Carbon 51, 236-242 (2013).
[11] M. Amin, M. Farhat, and H. Bağcı, “An ultra-broadband multilayered graphene absorber,” Opt. Express 21, 29938-29948 (2013).

[12] X. Han, T. Wang, X. Li, S. Xiao, and Y. Zhu, “Dynamically tunable plasmon induced transparency in a graphene-based nanoribbon waveguide coupled with graphene rectangular resonators structure on sapphire substrate,” Opt. Express 23, 31945-31955 (2015).
[13] T. Zhang, X. Yin, L. Chen, and X. Li, “Ultra-compact polarization beam splitter utilizing a graphene-based asymmetrical directional coupler,” Opt. Lett. 41, 356-359 (2016).
[14] H.-Y. Kim, K. Lee, N. McEvoy, C. Yim, and G. S. Duesberg, “Chemically modulated graphene diodes,” Nano Lett. 13, 2182-2188 (2013).
[15] A. V. Zayats, I. I. Smolyaninov, and A. A. Maradudin, “Nanoptics of surface plasmon polaritons,” Phys. Rep. 408, 131-314 (2005).
[16] M. Jablan, B. Buljan, and M. Soljačić, “Plasmonics in graphene at infrared frequencies,” Phys. Rev. B 80, 245435 (2009).
[17] T. Zhang, L. Chen, and X. Li, “Graphene-based tunable broadband hyperlens for far-field subdiffraction imaging at mid-infrared frequencies,” Opt. Express 21, 20888-20899 (2013).
[18] A. Vakil and N. Engheta, “Transformation optics using graphene,” Science 332, 1291-1294 (2011).
[19] L. Ju, B. Geng, J. Hongr, C. Girint, M. Martin, Z. Hao, H. A. Bechtel, X. Liang, A. Zettl, and Y. R. Shen, “Graphene plasmonics for tunable terahertz metamaterials,” Nat. Nanotechnol. 6, 630-634 (2011).
[20] M. A. Othman, C. Gucha, and F. Capolino. "Graphene-based tunable hyperbolic metamaterials and enhanced near-field absorption," Opt. Express 21, 7614-7632 (2013).
[21] S. Xiao, T. Wang, T. Liu, X. Yan, Z. Li, and C. Xu, "Active modulation of electromagnetically induced transparency analogue in terahertz hybrid metal-graphene metamaterials," Carbon 126, 271-278 (2018).
[22] R. Alae, M. Farhat, C. Rockstuhl, and F. Lederer, "A perfect absorber made of a graphene micro-ribbon metamaterial," Opt. Express 20, 28017-28024 (2012).
[23] D. Rodrigo, O. Limaj, D. Janner, D. Etezadi, F. J. G. de Abajo, V. Pruneri, and H. Altug, "Mid-infrared plasmonic biosensing with graphene," Science 349, 165-168 (2015).
[24] T. Zhang, L. Chen, B. Wang, and X. Li, "Tunable broadband plasmonic field enhancement on a graphene surface using a normal-incidence plane wave at mid-infrared frequencies," Sci. Rep. 5, 11195 (2015).
[25] H. Li, L. Wang, J. Liu, Z. Huang, B. Sun, and X. Zhai, "Investigation of the graphene based planar plasmonic filters," Appl. Phys. Lett. 103, 211104 (2013).
[26] A. Y. Nikitin, F. Guinea, F. J. García-Vidal, and L. Martin-Moreno, "Edge and waveguide terahertz surface plasmon modes in graphene microribbons," Phys. Rev. B 84, 161407 (2011).
[27] T. Zhang, Z. Zhou, J. Dai, Y. Dai, X. Han, J. Li, F. Yin, Y. Zhou, and K. Xu, "Plasmon induced absorption in a graphene-based nanoribbon waveguide system and its applications in logic gate and sensor," J. Phys. D: Appl. Phys. 51, 055103 (2018).
[28] R. D. Kekatpure, E. S. Barnard, W. Cai, and M. L. Brongersma, "Phase-controlled plasmon-induced transparency," Phys. Rev. Lett. 104, 243902 (2010).
[29] S.-X. Xia, X. Zhai, L.-L. Wang, and S.-C. Wen, "Plasmonically induced transparency in double-layered graphene nanoribbons," Photon. Res. 6, 692-702 (2018).
[30] H. Li, C. Ji, Y. Ren, J. Hu, M. Qin, and L. Wang, "Investigation of multiband plasmonic metamaterial perfect absorbers based on graphene ribbons by the phase-coupled method," Carbon 141, 481-487 (2019).
[31] A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," Nat. Photonics 9, 374-377 (2015).
[32] L. F. Freilenz, Y. Ding, O. Sigmund, and L. H. Frandsen, "Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides," Opt. Express 24, 16866-16873 (2016).
[33] B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with 2.4 × 2.4 µm² footprint," Nat. Photonics 9, 378-382 (2015).
organic carbon stocks across an Afromontane landscape," Ecological Indicators 52, 394-403 (2015).

A. Ahmad, M. Hassan, M. Abdullah, H. Rahman, F. Hussin, H. Abdullah, and R. Saidur, "A review on applications of ANN and SVM for building electrical energy consumption forecasting," Renewable and Sustainable Energy Reviews 33, 102-109 (2014).

Z. Lin, X. Liang, M. Lončar, S. G. Johnson, and A. W. Rodriguez, "Cavity-enhanced second-harmonic generation via nonlinear-overlap optimization," Optica 3, 233-238 (2016).

T. W. Hughes, M. Minkov, I. A. Williamson, and S. Fan, "Adjoint method and inverse design for nonlinear nanophotonic devices," ACS Photonics 5, 4781-4787 (2018).

Y. Xing, D. Spina, A. Li, T. Dhaene, and W. Bogaerts, "Stochastic collocation for device-level variability analysis in integrated photonics," Photon. Res. 4, 93-100 (2016).

L. Du, D. Tang, and X. Yuan, "Edge-reflection phase directed plasmonic resonances on graphene nano-structures," Opt. Express 22, 22689-22698 (2014).

C. Zeng, J. Guo, and X. Liu, "High-contrast electro-optic modulation of spatial light induced by graphene-integrated Fabry-Pérot microcavity," Appl. Phys. Lett. 105, 121103 (2014).

M. Maltamo and A. Kangas, "Methods based on k-nearest neighbor regression in the prediction of basal area diameter distribution," Canadian Journal of Forest Research 28, 1107-1115 (1998).

A. Swetapadma and A. Yadav, "A novel decision tree regression-based fault distance estimation scheme for transmission lines," IEEE Transactions on Power Delivery 32, 234-245 (2016).

A. Liaw and M. Wiener, "Classification and regression by randomForest," R news 2, 18-22 (2002).

P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," Machine learning 63, 3-42 (2006).

M. C. Bishop, Pattern recognition and machine learning (springer, 2006).

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg, "Scikit-learn: Machine learning in Python," J. Mach. Learn. Res. 12, 2825-2830 (2011).

A. Chipperfield and P. Fleming, "The MATLAB genetic algorithm toolbox," From IEE Colloquium on Applied Control Techniques Using MATLAB Digest No. 1995/014 (1995).

K.-H. Han and J.-H. Kim, "Genetic quantum algorithm and its application to combinatorial optimization problem," in Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No. 00TH8512), (IEEE, 2000), 1354-1360.

K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," IEEE transactions on evolutionary computation 6, 182-197 (2002).