



HAL
open science

Global optimization of metasurface designs using statistical learning methods

Mahmoud M R Elsawy, Stéphane Lanteri, Régis Duvigneau, Mohamed Sabry Mohamed, Gauthier Brière, Patrice Genevet

► **To cite this version:**

Mahmoud M R Elsawy, Stéphane Lanteri, Régis Duvigneau, Mohamed Sabry Mohamed, Gauthier Brière, et al.. Global optimization of metasurface designs using statistical learning methods. 2019. hal-02156881v2

HAL Id: hal-02156881

<https://hal.science/hal-02156881v2>

Preprint submitted on 27 Jul 2019 (v2), last revised 1 Dec 2019 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Global optimization of metasurface designs using statistical learning methods

Mahmoud M. R. Elsayy,¹ Stéphane Lanteri,¹ Régis Duvigneau,¹
Mohamed Sabry Mohamed,² Gauthier Brière,² and Patrice Genevet^{2,*}

¹Université Côte d’Azur, Inria, CNRS, LJAD, 06902 Sophia Antipolis Cedex, France

²CNRS, CRHEA, Université Côte d’Azur, rue Bernard Gregory, Sophia Antipolis, Valbonne, France

(Dated: July 27, 2019)

Optimization of the performance of flat optical components, also dubbed *metasurfaces*, is a crucial step towards their implementation in realistic optical systems. Yet, most of the design techniques, which rely on large parameter search to calculate the optical scattering response of elementary building blocks, do not account for near-field interactions that strongly influence the device performance. In this work, we exploit two advanced optimization techniques based on statistical learning and evolutionary strategies together with a fullwave high order Discontinuous Galerkin Time-Domain (DGTD) solver to optimize phase gradient metasurfaces. We first review the main features of these optimization techniques and then show that they can outperform most of the available designs proposed in the literature. Statistical learning is particularly interesting for optimizing complex problems containing several global minima/maxima. We then demonstrate optimal designs for GaN semiconductor phase gradient metasurfaces operating at visible wavelengths. Our numerical results reveal that rectangular and cylindrical nanopillar arrays can achieve more than respectively 88% and 85% of diffraction efficiency for TM polarization and both TM and TE polarization respectively, using only 150 fullwave simulations. To the best of our knowledge, this is the highest blazed diffraction efficiency reported so far at visible wavelength using such metasurface architectures.

PACS numbers: 78.67.Pt, 02.30.Zz, 02.60.Pn, 82.20.Wt

INTRODUCTION

Metasurfaces have been studied extensively in the past few years due to their exceptional abilities in achieving arbitrary light control in a very short propagation distance, and due to their simplified fabrication procedures with respect to bulk metamaterials [1–5]. Metasurfaces consist of assemblies of nanoresonators with spatially varying geometrical parameters and separated by sub-wavelength distances, made of plasmonic [6] and/or high dielectric refractive index materials [5, 7]. Unlike the conventional optical components that provide a full control of the light properties over long propagation distances, metasurfaces can introduce highly resolved phase, amplitude, and polarization changes on the incoming wavefront over very short propagation distances, typically in the order of the wavelength [3–8]. Owing to the versatility and the capabilities of metasurfaces, many exotic and peculiar optical phenomena ranging from negative refraction [9], sub-diffraction optical microscopy [10], and broadband achromatic lenses [11, 12] have been demonstrated recently using ultrathin and compact devices. Most of these designs have been engineered by considering a brute force approach. The latter consists in performing an extensive and costly parametric search to obtain the optical response of individual building blocks. Although simulations are generally performed considering array of nanostructures, this direct approach does not properly consider potential coupling effects between neighboring elements having different shapes. Remarkably, complex designs

such as broadband and multiplexed interfaces exclusively rely on sub-units of near-field coupled antennas to achieve the required scattering responses. The inherent complexity of the latter designs result in poorly efficient components, indicating that direct modelling approaches are becoming substantially insufficient and are failing to achieve designs of realistic devices [13, 14]. New and advanced methods, such as inverse design techniques, are becoming mandatory to further exploit metasurface capabilities in highly demanding applications [14, 15]. To this end, several optimization methodologies have been developed and demonstrated in the recent years, including local and global search methods. The former is suitable to rapidly convergence to local maxima/minima and thus strongly depends on the initial parameter guess [16, 17]. This category includes topology optimization [18–22] and so-called *objective-first* algorithms [23–25].

The second approach, performing global parameter optimization includes stochastic search techniques such as genetic algorithms [26–28] and evolutionary algorithms [29, 30]. These are general methods which are very efficient for large parameter space optimization. The downside of the latter methods is that they all require a large numbers of forward solver calls and are thus impractical when combined with costly (three-dimensional (3D) time-domain simulations.

In the last two years, artificial neural networks have been used to develop innovative modelling strategies for several nanoscale light-matter interaction problems including light scattering problems from spherical nanoshells for example [31]. Artificial neural networks have also been utilized recently to design efficient metasurfaces [32, 33]. As a general rule, training an artificial network requires numerous training data before it be-

* patrice.genevet@crhea.cnrs.fr

comes capable of achieving the optimized design based on a specific input target. Thus, efficient neural networks, capable of generating practical designs, require thousands of training data using a fullwave electromagnetic solver. [34, 35]. The computation cost could become important, especially considering 3D complex problems. Another common problem of neural networks arises when the system under investigation has many diverse parameters, i.e. when several parameter sets could give approximately the same response. In this case, the performance of the network reduces dramatically [33, 34].

The main goal of our work is to introduce to the nanophotonics community novel and significantly more advanced evolutionary optimization strategies based on derandomized and statistical learning, evolution strategies. Using practical designs of semiconductor GaN phase gradient metasurfaces, we demonstrate that these techniques can outperform most of optimization techniques used in the inverse design of metasurfaces. These are especially useful when one is considering complex 3D problems. We adopt a parametric shape optimization viewpoint as opposed to a topology optimization approach, enabling faster convergence to a global minima/maxima even for large parameter space with regards to our setting. In addition, as shown inhere, our methods are capable of achieving effectively different global minima/maxima for the same value of the objective function, involving different parameter values. We first consider an analytical example to help readers gaining insights into our inverse design tools. The second goal of this paper is to apply theses techniques to the case of 3D GaN phase gradient metasurfaces made of nanopillars of different shapes, targeting maximum light deflection efficiency at a wavelength of $\lambda = 600 \text{ nm}$. The deliberate choice of GaN semiconductor has been made after a careful consideration of several factors such as the optical losses in the visible regime, its high refractive index in the visible regime, and the current ease of micro/nanofabrication technology in the industry of this materials, yielding ideal nanoresonators (phase-shifters) for metasurface designs and fabrication [7, 36]. Over the last couple of years, several example of light deflecting metasurfaces have been realized both numerically and experimentally for visible wavelength applications involving GaN [36, 37] or at near infrared [21, 29, 38] using silicon or hydrogenated amorphous silicon. For the latter, the absorption losses of silicon in the visible make them less efficient than, for example, TiO_2 [19] or c-Si [19, 39]. Various metagratings have been used to demonstrate efficient light deflection at visible regime but the performance does not exceed 80% [40]. Here, we provide optimized 3D metasurface designs with record efficiency above 87%. To the best of our knowledge, this is the highest expected performance reported in the literature at visible regime for 3D gradient metasurfaces. Our global optimization techniques rely respectively on advanced evolutionary strategies and statistical learning, coupled with a high order Discontinuous Galerkin Time-Domain (DGTD) solver from the

DIOGENE $\text{\textcircled{S}}$ software suite dedicated to computational nanophotonics [41]. We apply the optimization solution to design optimal phase gradient metasurfaces made of scatterers of different shaped, such as rectangular and cylindrical nanopillars. Our calculations target maximum diffraction efficiency ($\eta(n, m)$, where n, m are the mode indices) at $\lambda = 600 \text{ nm}$. For rectangular shaped nanopillars, we achieved more than 87% of diffraction efficiency at $\lambda = 600 \text{ nm}$ for TM polarized waves, while cylindrically shaped nanopillar interfaces lead to more than 85% efficiency for both TM and TE light polarization at $\lambda = 600 \text{ nm}$.

FULLWAVE TIME-DOMAIN SOLVER

The fullwave numerical modelling of light interaction with nanometer scale structures generally relies on the solution of the system of 3D time-domain Maxwell equations. In the computational nanophotonics literature, a large number of studies are devoted to Finite Difference Time-Domain (FDTD) type discretization methods based on Yee's scheme [42]. As a matter of fact, the FDTD [43] method is a widely used approach for solving the systems of partial differential equations modeling nanophotonic applications. In this method, the whole computational domain is discretized using a structured (cartesian) grid. However, in spite of its flexibility and second-order accuracy in a homogeneous medium, the Yee scheme suffers from serious accuracy degradation when used to model curved objects or when treating material interfaces, especially if one is interested in assessing near field effects. In general, this requires a refinement of the underlying grid, which incurs a substantial increase of the simulation time and memory footprint. During the last twenty years, numerical methods formulated on unstructured meshes have drawn a lot of attention with the aim of dealing with irregularly shaped structures and heterogeneous media. In particular, the Discontinuous Galerkin Time-Domain (DGTD) method [44] has met an increased interest. The DGTD method can be considered as a finite element method where the continuity constraint at an element interface is released. While it keeps almost all the advantages of the finite element method (large spectrum of applications, complex geometries, etc.), the DGTD method has other nice properties which explain the interest it recently gained in the computational electromagnetics community:

- It is naturally adapted to a high order approximation of the unknown field. Moreover, one may increase the degree of the approximation in the whole mesh as easily as for spectral methods but, with a DGTD method, this can also be done locally *i.e.* at the mesh cell level. In most cases, the approximation relies on a polynomial interpolation method (that we shall denote as \mathbb{P}_p where p is the interpolation degree) but the method also offers the flexibility of applying local approximation strategies that

best fit to the intrinsic features of the modelled physical phenomena.

- When the discretization in space is coupled to an explicit time integration method, the DG method leads to a block diagonal mass matrix independently of the form of the local approximation (e.g the type of polynomial interpolation). This is a striking difference with classical, continuous FETD (Finite Element Time-Domain) formulations. Moreover, the mass matrix is diagonal if an orthogonal basis is chosen.
- It easily handles complex meshes. The grid may be a classical conforming finite element mesh, a non-conforming one or even a hybrid mesh made of various elements (tetrahedra, prisms, hexahedra, etc.). The DGTD method has been proven to work well with highly locally refined meshes. This property makes the DGTD method more suitable to the design of a *hp*-adaptive solution strategy (i.e. where the characteristic mesh size h and the interpolation degree p changes locally wherever it is needed).
- It is flexible with regards to the choice of the time stepping scheme. One may combine the DG spatial discretization with any global or local explicit time integration scheme, or even implicit, provided the resulting scheme is stable.
- It is naturally adapted to parallel computing. As long as an explicit time integration scheme is used, the DGTD method is easily parallelized. Moreover, the compact nature of method is in favor of high computation to communication ratio especially when the interpolation order is increased.

As in a classical finite element framework, a discontinuous Galerkin formulation relies on a weak form of the continuous problem at hand. However, due to the discontinuity of the global approximation, this variational formulation has to be defined at the element level. Then, a degree of freedom in the design of a discontinuous Galerkin scheme stems from the approximation of the boundary integral term resulting from the application of an integration by parts to the element-wise variational form. In the spirit of finite volume methods, the approximation of this boundary integral term calls for a numerical flux function which can be based on either a centered scheme or an upwind scheme, or a blend of these two schemes. In this study, we exploit a recently developed high order DGTD-based solver that has been specifically designed for the simulation of nanoscale light-matter interaction problems [45]. This DGTD method is implemented in the DIOGENeS [41] software suite, which is programmed in Fortran 2008 and is adapted to high performance computing systems.

OPTIMIZATION METHODS

We are here leveraging on two different efficient global optimization techniques respectively based on advanced evolutionary strategies and statistical learning to perform inverse metasurface designs. The first optimization approach is the so-called "Covariance Matrix Adaptation Evolution Strategy" (CMA-ES) [46], which belongs to the family of evolutionary algorithms. Similarly as genetic algorithms (GAs) [47], CMA-ES mimics natural evolution principles to maximize an objective function. It has been tested on several academic benchmarks as well as industrial problems and reported as one of the most efficient optimization algorithms for a broad class of problems. Recently, it has been used to optimize infrared broadband quarter-wave and half-wave plates Bézier metasurfaces [48], reconfigurable metasurface absorbers [49], acoustic metamaterial [50], and for in optimizing apochromatic singlets metasurface-augmented grin lenses [51]. This is of critical importance for the design of complex assemblies of arrays of 3D nanostructure requiring expensive simulations with a large number of parameters, thus increasing the chance of finding several designs with relatively well-optimized performances.

As other Evolution Strategies (ES), CMA-ES is based on a sequence of random searches, ruled by a normally distributed sampling. At each iteration, the characteristics of the distribution, and in particular its covariance matrix, are adapted to account for the latest obtained observations, in order to accelerate the convergence towards the maximum/minimum of the fitness function. More precisely, the algorithm starts with an initial (some-time random) distribution mean (initial design) and the identity classical normal distribution as covariance matrix scaled by a user-defined scalar variance. According to this distribution, a set of N samples are generated randomly (*mutation* step as shown in Fig. 1(a)) and the corresponding designs are simulated to evaluate their fitness. Then, the best N_{best} designs among the N ones are selected for the evolution (*selection* step) and are used to update the mean of the distribution (*recombination* step). Finally, the covariance matrix is updated accounting for a principal component analysis of the best points N_{best} , whereas the scalar variance is modified using the path of the mean point (*step-size control*). The whole procedure is repeated using the updated distribution, until convergence. The detailed algorithm is described in [46, 52] and illustrated in Fig. 1(a). In the electromagnetic community, GAs are more commonly used [15, 26, 27, 47, 53], although they suffer from some well-known drawbacks, such as the necessity to calibrate several numerical parameters and their inability to tackle anisotropic behaviors of the fitness function, i.e. the ability to change/adapt the search shape distribution during the optimization in order to tackle the complexity of the objective function for faster convergence. Consequently, GAs require a substantial amount of computational time when applied to high-performance 3D time/frequency do-

main electromagnetic solvers. Therefore, due to the huge number of parameters required to ensure the flexibility of modern devices, more advanced evolutionary strategies like CMA-ES are needed, that could 1) offer a faster convergence, i.e. require fewer fullwave solver runs (due to its ability to update the shape and the size of the distribution during the optimization), and 2) provide more accurate results.

Despite the advantages of CMA-ES over other classical evolutionary algorithms, it still requires a large number of computationally expensive simulations, which increases with the number of optimization parameters (*curse of dimensionality*). Therefore, we also consider an alternative optimization strategy based on the iterative construction of surrogate models. It is known as Efficient Global Optimization (EGO) [54–56] and belongs to the class of Bayesian optimization methods. Contrary to evolutionary algorithms, EGO is not based on adaptive sampling, but on a surrogate model built on the basis of available fitness observations, which is employed to decide which new design should be tested next, by maximizing a statistical criterion related to the optimization goal (*merit function*). More precisely, this approach proceeds in two main steps. First, a Design Of Experiments (DOE) is carried out, which consists in exploring randomly the admissible design space using a uniform sampling strategy (e.g. Latin Hypercube Sampling) [57]. After simulation of the corresponding designs, all fitness values obtained are considered as observations and are stored in a database. In the second step, this database is iteratively refined towards the most promising areas. This is performed using a Gaussian Process (GP) model, constructed using all the database points. This is basically an interpolating or approximating model, whose internal parameters are calibrated according to a maximum likelihood principle [58]. Once this GP model is defined, one can estimate at any point of the design space the fitness value (model mean) and an uncertainty value (model variance). Both are used to define a statistical merit function, e.g. the expected improvement, whose maximum defines the next design parameters set to evaluate. After simulation of this new point, the database is updated accounting for this new observation and the second step of the algorithm repeats until convergence. The algorithm is illustrated in Fig. 1(b). As explained, this approach based on the iterative construction of a database and an associated model can be considered as a statistical learning strategy, like for instance Artificial Neural Networks (ANNs). Its main characteristic is the use of internal uncertainty estimation (variance) to drive both the search for the optimum and the improvement of the model accuracy. Contrary to the approaches based on ANNs, that often aims at constructing a model accurate in the whole design space before optimization, EGO focuses on the most promising areas regarding the optimization criterion. It is therefore far less expensive in terms of solver calls, making it very well suited to the context of expensive simulations. In practice, only a few hundreds of simulations are typically

required for EGO, whereas the database used to train ANNs usually requires tens of thousands of solver calls. As an example, in Ref. [32] a database of 90 000 simulations has been achieved in order to train an ANN to optimize 16 parameters. For EGO, as illustrated below, we need only 80 points for the initial design of experiments (DoE) and 150 solver calls for the iterative enrichment, to optimize structures with 12 and 8 parameters (rectangular and cylindrical shaped antennas).

In order to highlight the different optimization behaviors and performance between CMA-ES and EGO and to give the reader more insights about their properties, we provide here a 2D analytical example which is designed to find the global minimum of a 2D Branin function, defined by:

$$\begin{cases} \underset{x,y}{\text{minimize}} f(x,y) = \left(y - \frac{5.1}{4\pi^2}x^2 + \frac{5}{\pi}x - 6 \right)^2 \\ \quad + 10 \left(1 - \frac{1}{8\pi} \right) \cos(x) + 10, \\ \text{subject to } -5 \leq x \leq 10 \text{ and } 0 \leq y \leq 15. \end{cases} \quad (1)$$

The Branin function $f(x, y)$ depicted in Fig. 2(a)) has three global minima, i.e three different combinations of parameters x and y in the range of $-5 \leq x \leq 10$ and $0 \leq y \leq 15$ generate the same value of the objective function, as indicated by the black arrows in Fig. 2(a). Using this simple example, we highlight the different behaviors of the two algorithms. We compared the convergence and the computational efficiency of the two algorithms in Figs. 2(b-c). They depict the function value obtained with respect to the number of solver calls. As seen, both methods reach a very satisfactory function value, CMA-ES being clearly more expensive converging after about 180 solver calls, compared to 30 (including training data) for the EGO method. Below, we demonstrate the evolution of the two methods.

For CMA-ES, the yellow points shown in Figs. 2(d-g) show the evolution of all samples tested at generations ranging from 0 to 30, the red points corresponding to the last generation (size $N = 6$). CMA-ES sampling points are essentially exploring only the vicinity of the population mean, which progressively moves towards the best objective function values to converge on a local minimum.

For EGO, Figs. 2(h-k) show both the Gaussian Process model and the underlying database generated from iteration 0 to 30. At iteration 0, the database is composed of 6 points obtained from the DOE phase. Then, the database is enriched in most promising areas, yielding an improvement of the corresponding model. At iteration 30, the three regions around the global minima have been found. Moreover, a Gaussian Process model has been constructed, which is very close to the true function seen in Fig. 2(a), in which the three global points have been identified by the EGO (see black points in Fig. 2(k) and the analytical function given in Fig. 2(a)).

This simple analytical example illustrate the different optimization mechanisms involved in CMA-ES and EGO

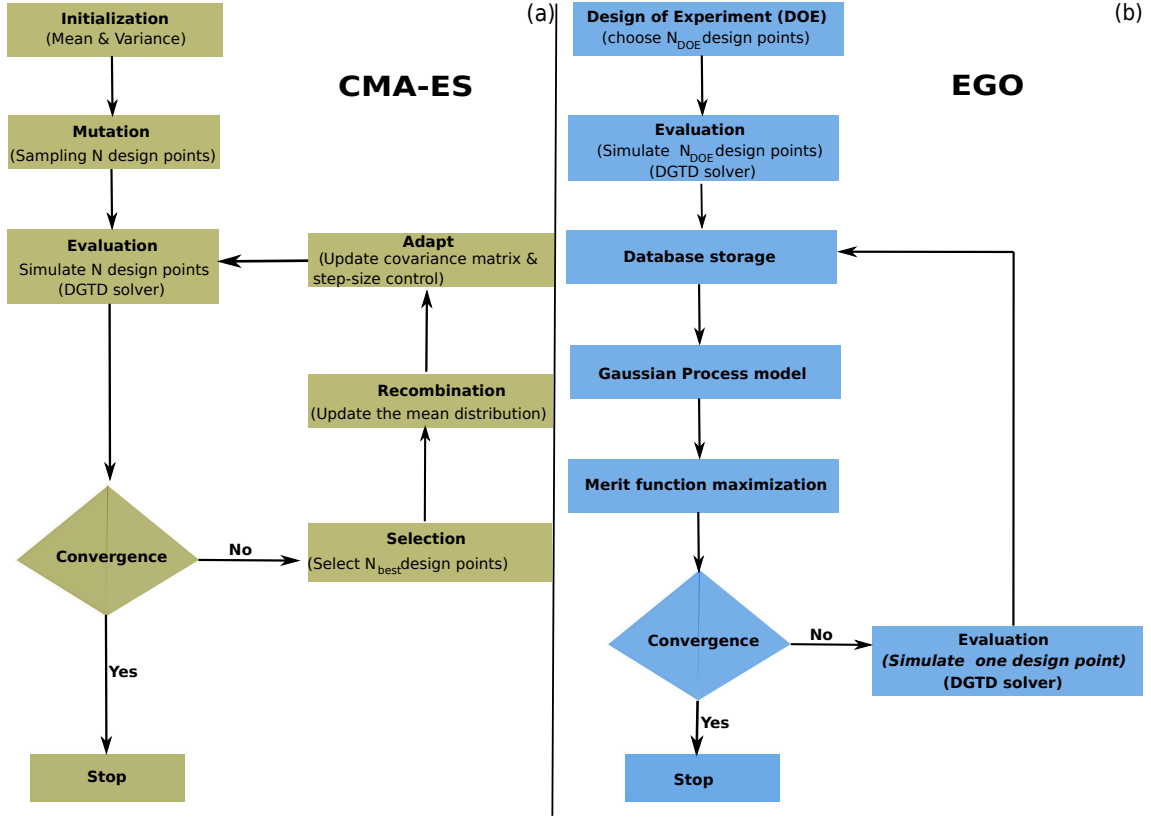


FIG. 1. For the CMA-ES (a), the optimization starts with an initial design and a given mean and variance. The second step consists in generating a population of N designs that will be evaluated using the fullwave DGTD solver. If the convergence is not satisfied, the algorithm chooses the N_{best} designs according to the objective function, and use them to update the mean of the distribution, size and the covariance matrix. N new designs are thereafter calculated and repeated steps until the convergence criterion is reached. For the EGO (b) instead, the algorithm starts with an initial design of experiments composed of N_{DOE} designs, that will be simulated to estimate the corresponding objective function. The results are used to construct a surrogate model, of interest to search for the next design. The latter is simulated using the fullwave DGTD solver, and the corresponding objective value enrich the database, repeating these steps until the convergence is obtained. The major difference between the CMA-ES and EGO methods essentially relies on the utilization of a surrogate model for EGO that drastically reduces the number of evaluations N_{DOE} by an order of magnitude.

algorithms. In the following, we exploit these two methods to optimize metasurface geometries for improved light deflection efficiency at $\lambda = 600 \text{ nm}$.

NUMERICAL RESULTS

As a first example, in Fig. 3(a), we consider a phase gradient metasurface made of rectangular GaN semiconductor (dark-red regions) placed over a semi-infinite substrate made of Al_2O_3 (shown in green). We consider a normal incident plane wave with electric field polarized in the y-direction, and we aim to maximize the diffraction efficiency of the first order mode $\eta(0, -1)$ (deflect light in the same plane of incidence y-z plane) at wavelength of $\lambda = 600 \text{ nm}$. To avoid diffraction inside the substrate, we consider sub-wavelength period in the x-direction (300 nm) and a period of phase gradient in the y-direction, which is the dimension along which we are

optimizing the geometries of subwavelength nanopillars, to be $\Gamma = 1500 \text{ nm}$, as shown in Fig. 3(a).

In the first design, we optimize the design of rectangular antennas, considering optimization of the following parameters: the height, which is chosen equal for all pillars to comply with nanofabrication techniques; the positions of each ridges in y direction; the thicknesses in x and y directions, leading to 12 optimized parameters that are represented by the red circles in Fig. 3(a). It is worth mentioning that we took into account the experimental constraints during the optimization process, in which the minimum feature size is set to 90 nm and the height of the ridges is set between 600 nm and 800 nm . Fig. 3(b-c) summarize the optimization results obtained for this first metasurface example. The CMA-ES results are shown in Fig. 3(b), in which the objective function evaluation at each iteration is shown in dark-yellow points, and the best values of the objective function evaluation obtained during the optimization process is represented by the pur-

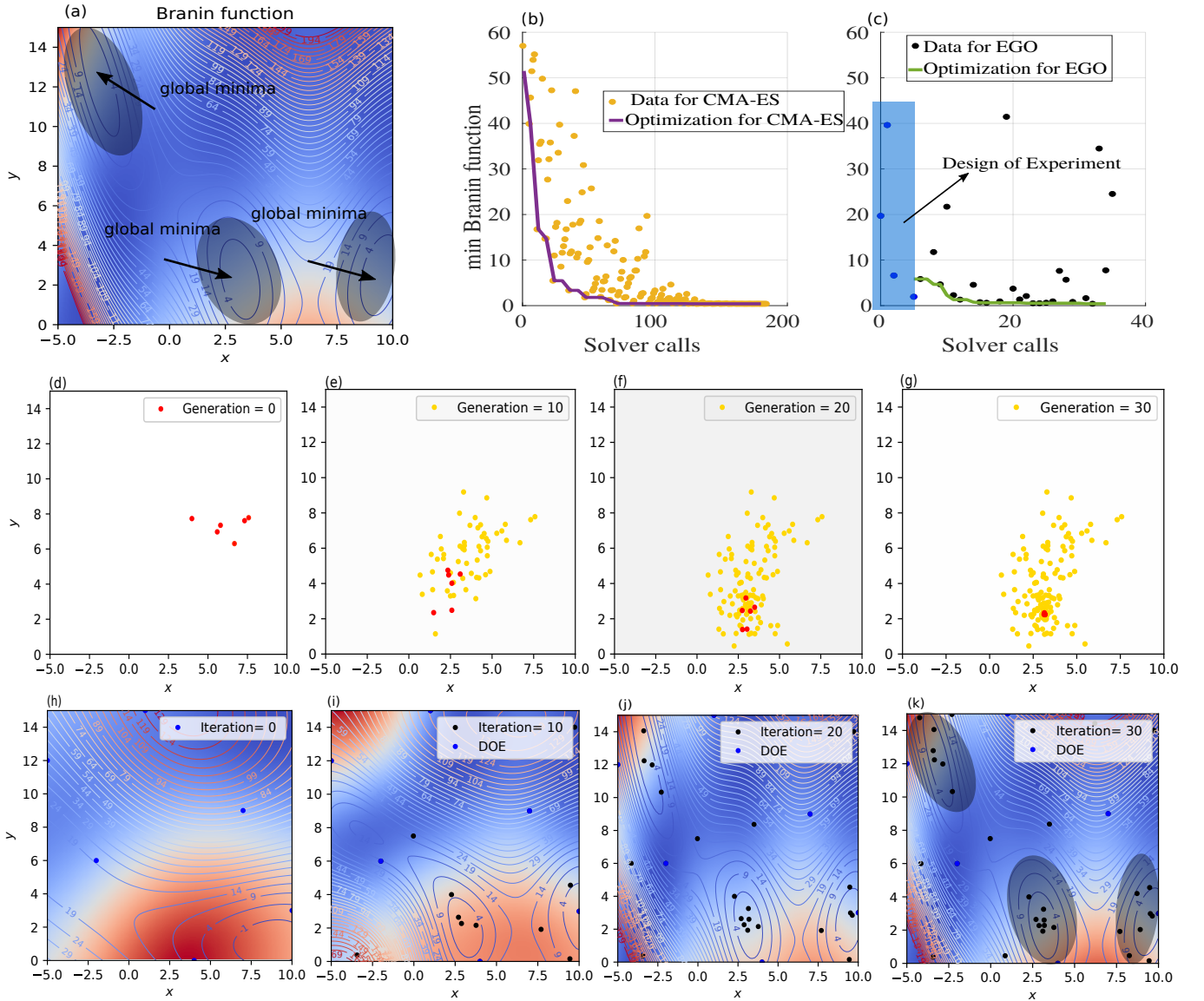


FIG. 2. A representative 2D analytical example to illustrate the different behaviors of CMA-ES and EGO methods. The problem considered here consists in minimizing an analytical function, known as the Branin function, characterized by the presence of 3 global minima, as indicated by the black arrows in (a). The evolution of CMA-ES as a function of the solver calls is provided in (b) in which the yellow points represent the objective function values at each iteration, the purple curve indicates the best value at each iteration. (c): similar to (b) except that here the blue points represent the DOE phase $N_{DOE} = 6$ (only 6 in this example). These points are used only for the initial training (blue shaded region in (b)) while the black dots represent the data generated during the optimization phase. The green line represents the best, optimized, results obtained during optimization phase, as explained in Fig. 1(b). (d-g): evolution of the points tested by CMA-ES, as a function of the generation numbers (for each generation we simulate 6 designs). The yellow points represent all samples evaluated so far, the red points correspond to the last generation of size $N = 6$. This illustrates the search by progressive sampling and convergence. (h-k): evaluation of the Gaussian process model (surrogate model) and the underlying database generated from iteration 0 to 30. Notice that the model converge to the analytical function shown in (a) after 30 iterations. The design points from the DOE phase are shown in blue, the black points represent the database progressively enriched during the optimization. Note that all minima are detected. As background, the GP model is plotted, which converges progressively towards the true cost function (a).

ple solid curve.

Note that the evolutionary model evaluates at each iteration exactly 11 metasurface designs, keeping the best value among the different realizations only if the latter

outperform those obtained during the previous iteration (purple curve in Fig.3(b)). After nearly 550 iterations, we obtain a global minimum such that $1 - \eta^{TM}(0, -1) \approx 0.119$ at $\lambda = 600 \text{ nm}$, which is corresponding to diffrac-

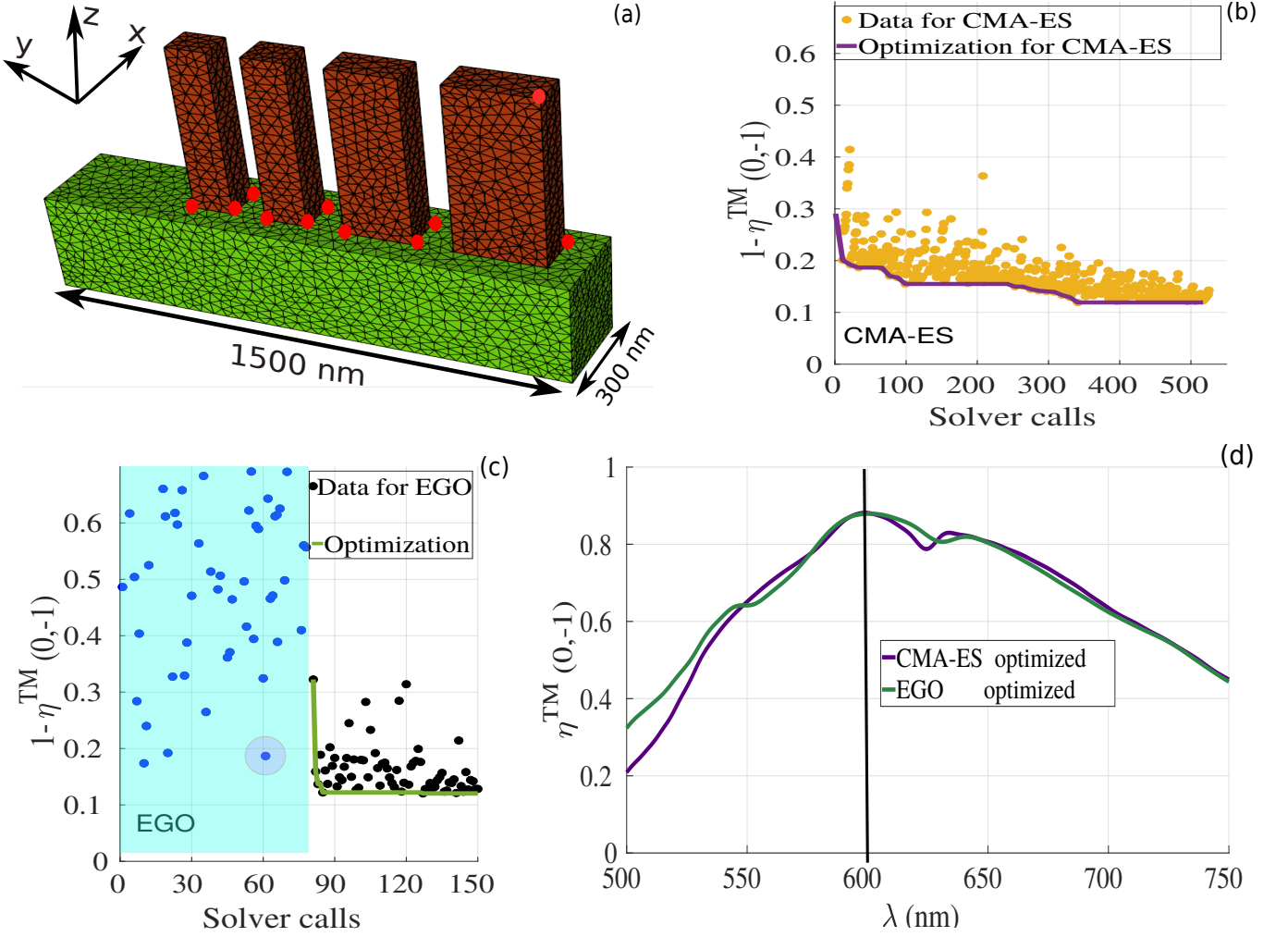


FIG. 3. Results of the optimizations for rectangular nanoantenna arrays. (a) the geometry under consideration with rectangular nano ridges made of GaN (dark-red ridges) on top of a semi-infinite substrate made of Al₂O₃ (green region). The 12 red circles represent the optimization parameters. (b) optimization process using CMA-ES as a function of the number of fullwave solver calls. Dark-yellow points represent the value of the objective function at each iteration, the solid purple line highlights the best point achieved during the optimization. In other words, for each generation we keep only the best point that minimizes the objective function. These best points that are obtained from each generation along the optimization process, are represented by the solid purple curve. (c) optimization realized with the EGO solver as a function of the number of fullwave solver calls. The blue points represent the DOE (shaded region), the black points represent the value of the objective function at each optimization iteration, and the green solid line indicates the optimized data. (d) comparison between the diffraction efficiency for the first order mode as a function of the wavelength for the TM polarized wave. We use purple and green colors for the CMA-ES and EGO optimized geometries, respectively; the corresponding parameter values shown in Tab.I, for CMA-ES, and EGO optimized parameters.

tion efficiency of approximately 88.10% for TM polarized waves. The evolution of the diffraction efficiency $\eta^{TM}(0, -1)$ as a function of the wavelength is represented by the purple curve in Fig. 3(d), indicating that a maximum is achieved at $\lambda = 600$ nm. The corresponding parameter values are shown in Tab.I, where dx and dy give the thicknesses in x and y directions, respectively for each rectangular element. From the result of the CMA-ES method, the optimal height of the ridges is $h \approx 800$ nm. As could be expected, a linear phase gradient, which leads to light deflection in y - z plane, is

achieved by increasing the effective mode index of the nanostructures, i.e. by gradually increasing the thicknesses in y direction as a function of the position of the nanoantenna within the period. Note also that the optimized height of the ridges is nearly 800 nm, which according to previous direct simulations and experimental works is sufficiently tall to provide high transparency windows and sufficient phase delay for increasing cross sections[12, 36, 37, 59]. The electromagnetic field distributions for the $Re(H_x)$ and $Re(E_y)$ at $\lambda = 600$ nm are shown in Fig. 4(a) and Fig. 4(b), respectively, which

clearly indicate light deflection behavior. It is worth mentioning that during the optimization process, accurate simulation results was obtained using a coarse mesh of only 3000 cells with a DGTD- \mathbb{P}_4 solver (i.e. with fourth order polynomial interpolation of the components of the electromagnetic field within each mesh cell). One full-wave simulation with these mesh and solver configurations cost 26 minutes using 48 cores. A numerical convergence is provided in Fig. 8(a) in the supplementary information section. The results shown in Fig. 4(a) and Fig. 4(b) have been obtained using a finer mesh and a DGTD- \mathbb{P}_2 solver.

The same optimization has also been realized using the EGO method, still considering the rectangular nanopillar setting shown in Fig. 3(a). As explained previously, this optimization approach build a DOE database in a first phase. In this example, we built a database with 80 design points represented by the blue points in Fig. 3(c). In a second phase, based on these 80 design points, a surrogate model is constructed and used during the optimization process to find a global minimum, i.e. that could achieve a better result than the actual best point found in the DOE process (represented by the pink point in Fig. 3(c)). From the green curve and its associated data (black points just above the green curve in Fig. 3(c)), one can note that the convergence is obtained after only a few number of iterations (approx. 150 iterations), beyond which the efficiency barely improves, thus indicating that the best point has been obtained. Compared to the CMA-ES, EGO minimization of the cost function optimizes the 12 parameters and converges to a global minimum in which $1 - \eta(0, -1) \approx 0.12$, corresponding to a deflection efficiency of 88.0% at $\lambda = 600 \text{ nm}$ after nearly 150 iterations, i.e. about 4 times faster than CMA-ES method.

Observing the optimized parameters, we realize that the geometries obtained by the EGO method (related to a specific stopping criteria) are different from those obtained with CMA-ES method (see Tab. I), for the optimized parameters obtained using the CMA-ES and EGO methods), although both could provide seemingly identical diffraction efficiency of about 88% at $\lambda = 600 \text{ nm}$, see Fig. 3(d).

Interestingly, the optimized parameters obtained by the EGO method are going against the intuitive designs for which linear phase gradients rely on increasing effective mode index of the antennas as a function of the pillar transverse section. Indeed, the expected gradient in the d_y thicknesses of the ridges is not fully satisfied as the d_y of the last ridge is larger than the previous pillar. These results are sensibly different to those obtained using the CMA-ES method. Note also that the global thicknesses of the ridges in y direction are slightly thicker than their counterparts obtained using the CMA-ES method. In addition, the height of the nanoridges is also much shorter than the optimized height found by the CMA-ES method by about nearly one wavelength. Similarly to the analytical example shown in Fig. 2, the noticeable differences in

the parameters indicates that EGO has reached another global optimum, which was not found by the CMA-ES method. The difference in structural parameters is also reflected in the field maps provided in Fig. 4, in which we clearly see that the number of longitudinal modes for the optimized geometry found by the EGO method shown in Figs. 4(a-b) are different from the ones obtained by the CMA-ES method shown in Figs. 4(c-d). The possibility of operating in the optimal transmission regime, also related to the Kerker condition in Mie theory, has been discussed recently. Since the height of nanopillar controls essentially the longitudinal resonance, similar light deflection response can be achieved by varying the transverse cross sections of nanopillar arrays having two distinct heights, thus explaining why we could achieve at least two different global minima in this configuration. We emphasize that in our unit cell, which is made of 4 nanoridges, optimization is performed by modifying all antennas parameters at once, thus properly optimizing the near-field coupling between the different modes, as illustrated in Fig. 4.

For the second example, we optimize a phase gradient metasurface made of cylindrical nanopillars (see Fig. 5(a)) to maximize the diffraction efficiency in the first order mode. Due to the symmetry of the nanopillars, we are targeting polarization insensitive properties. For this example, we optimized the structures by modifying 8 parameters (see Fig. 5(a)) including the pillars diameter (thick white arrows), the height, and the position of the nanopillars as indicated by the red points. We also kept experimental constraints identical to those for the rectangular case. We begin by analyzing again the optimization of the CMA-ES case, as shown in Fig. 5(b), for which a global point with diffraction efficiency around 85% for TM polarized waves (results for TE polarization is also around 85%, data not shown) is obtained after 300 fullwave solver calls. For the EGO method, we consider a DOE database with 50 design points (see blue points in Fig. 5(b)). Then, a surrogate model is constructed and a global optimum is clearly obtained after nearly 180 iterations with efficiency 85%. For this example, both optimization methods lead to structures with increasing gradient in the diameter, but the two set of parameters obtained by CMA-ES and EGO are slightly different as indicated in Tab. II.

The DGTD method allows us to obtain the wavelength dependance of the deflection efficiency for both CMA-ES and EGO, essentially showing similar behavior. We denote a slight difference, which can be attributed to the fact that both structures are not operating on the exact same longitudinal modes, see the field maps in Fig. 6 obtained at $\lambda = 600 \text{ nm}$. The CMA-ES optimum is achieved for $h \approx 735. \text{ nm}$, for which nearly half of the last mode field lobe is located in the first ridge, while EGO optimization results gives an optimal height $h \approx 708. \text{ nm}$.

To capture the influence of the height on the performance of the designed metasurfaces, we decided to compare the CMA-ES results for a fixed height $h = 800 \text{ nm}$

TABLE I. CMA-ES and EGO optimized parameters for the rectangular shaped antennas shown in Fig.3(a). The optimized heights obtained from CMA-ES and EGO are $h \approx 800 \text{ nm}$ and $h \approx 666 \text{ nm}$, respectively.

CMA-ES: element dimensions($h \approx 800 \text{ nm}$)			EGO: element dimensions ($h \approx 666 \text{ nm}$)		
Element	$d_x(\text{nm})$	$d_y(\text{nm})$	Element	$d_x(\text{nm})$	$d_y(\text{nm})$
1	108.5	248.7	1	120.0	254.1
2	93.3	199.6	2	106.3	203.3
3	96.6	116.5	3	103.6	109.1
4	108.6	108.1	4	107.3	130.6
CMA-ES: distance between elements			EGO: distance between elements		
between (1 – 2)	121.4 <i>nm</i>		between (1 – 2)	98.5 <i>nm</i>	
between (2 – 3)	114.6 <i>nm</i>		between (2 – 3)	91.5 <i>nm</i>	
between (3 – 4)	103.8 <i>nm</i>		between (3 – 4)	101.6 <i>nm</i>	

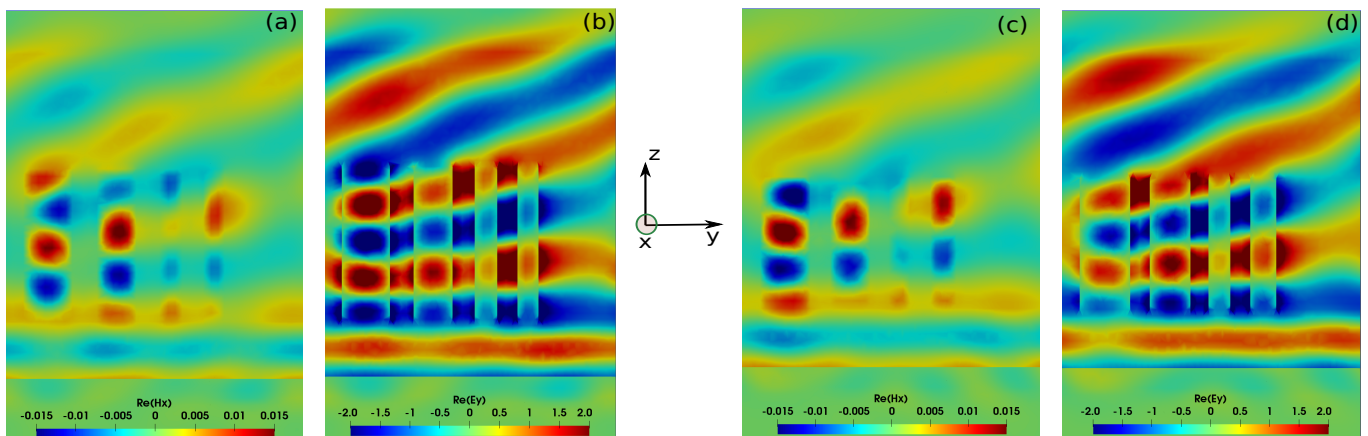


FIG. 4. Field maps of $\Re(E_x)$ and $\Re(E_y)$ obtained for rectangular nanopillars for the optimized geometries at $\lambda = 600 \text{ nm}$. (a-b): obtained from the CMA-ES method with height $h \approx 800 \text{ nm}$; (c-d): obtained from the EGO method with $h \approx 666 \text{ nm}$.

with those obtained when including the height in the set of optimization parameters, see Fig. 6 (b)). The comparison in Fig. 7(a) clearly illustrates that the results obtained with a fixed height $h = 800 \text{ nm}$ (dark orange curve) approach those obtained with varying the height, including the set of parameters obtained from both structures that converge to each other, see Tab. II and Tab. III for the reference case with $h \approx 735 \text{ nm}$ and the fixed height case $h = 800 \text{ nm}$, respectively. This indicates that the maximum of efficiency found in this configuration is quite robust, being resilient to significant height variation. To properly account for uncertainties due to the meshing of the nanostructures, we have realized a convergence proof considering the cylindrical nanopillar array, see Fig. 8(b), in which we investigated different mesh sizes and different polynomial orders. It is worth mentioning that during the optimization process, we used a coarse mesh of 10 000 cells with fourth order polynomial \mathbb{P}_4 , which proved sufficient to get accurate results (one fullwave simulation with this mesh and interpolation order specifications takes about 40 minutes using 48 cores), as it is depicted in Fig. 8(b). To present accurate field distributions and improve the visualization, the results obtained for the optimized geometries shown in Figs. 6(a-

d) and Fig. 7 have been obtained with a finer mesh and a DGTD- \mathbb{P}_2 solver. As shown in Fig. 8(b) and Fig. 9, it is possible to further improve the numerical accuracy using higher order curvilinear elements in combination with higher order polynomials. However, current experimental realization of metasurfaces using state of the art nanofabrication facilities do not reach such fine level of details and it is therefore not essential to further improve the nanostructure discretization to even higher numerical resolution. At this point, it is instead more interesting to point out that after various attempts to achieve close to unitary deflection efficiency, all optimized structures converge towards roughly the same number of 0.85.

This limitation of the deflection efficiency is expected as demonstrated in Refs. [60, 61]. According to these works, the efficiency of the phase gradient metasurfaces is limited due to impedance mismatching between the incident and the desired wavefront, leading to an increase of the scattering in the other modes. In particular, theoretical proofs have shown that a unitary light deflection to an arbitrary angle cannot be achieved using linear gradient phase profiles [60].

This effect is sufficiently important to avoid full unitary efficiency [60]. In order to overcome this limitations

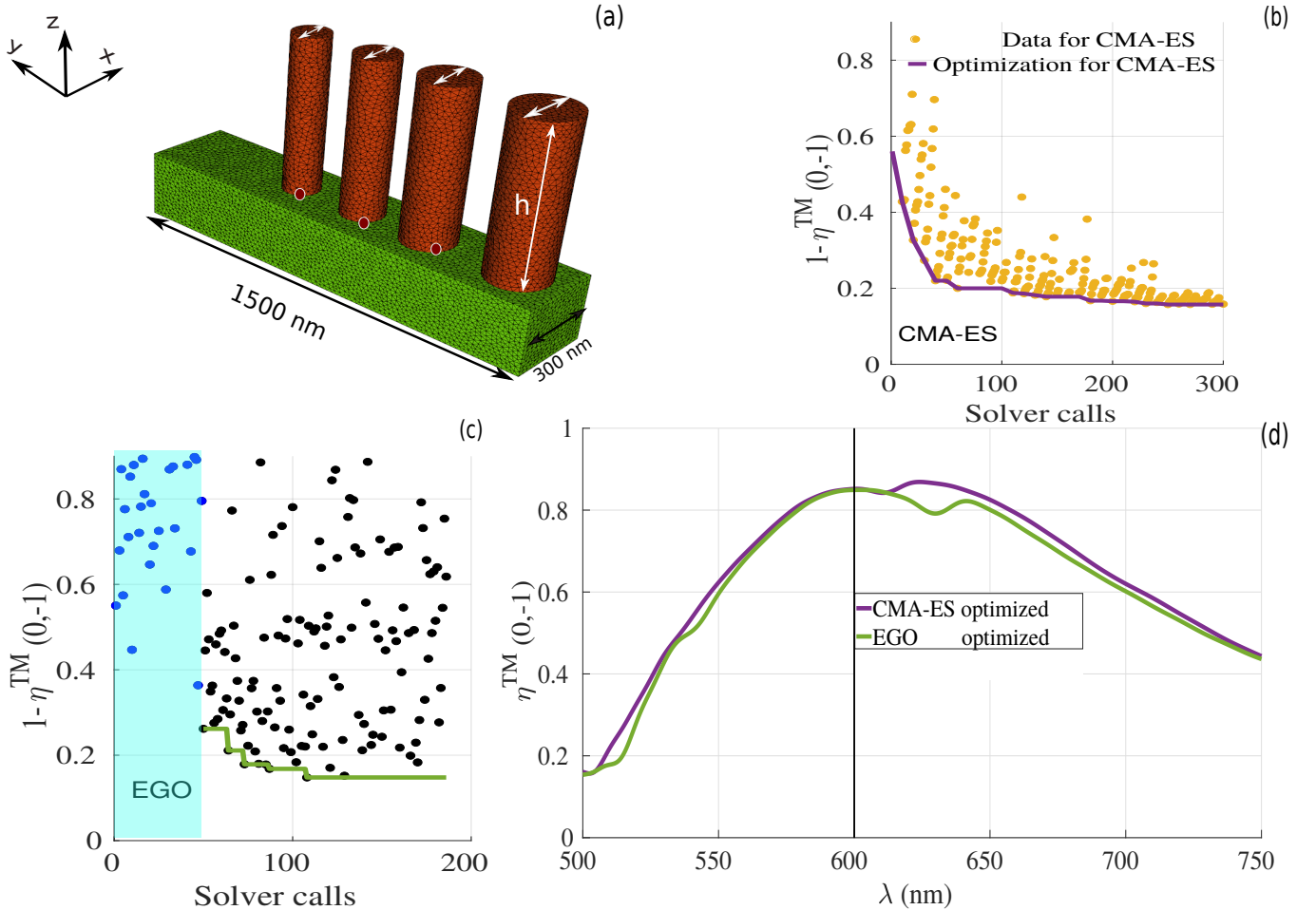


FIG. 5. Optimization results for cylindrical nanopillars arrays. Similarly to Fig. 3, (a) is the geometry under consideration with cylindrical nanoridges made of GaN (dark-red ridges) on top of a semi-infinite substrate made of Al_2O_3 (green region). The 8 red circles represent the optimization parameters. (b) optimization process using CMA-ES as a function of the fullwave solver calls. Dark-yellow points represent the value of the objective function at each iteration, the solid purple line indicates the best point during the optimization up to the current iteration. (c) optimization using the EGO method as a function of the fullwave solver calls. The blue points represent the DOE (shaded region), the black points represent the value of the objective function at each optimization iteration, and the green solid line indicates the optimized data. (d): comparison between the diffraction efficiency for the first order mode as a function of the wavelength for the TM polarized wave. Purple and green colors for the CMA-ES and EGO optimized geometries, respectively; the corresponding parameter values shown in Tab. II, for the CMA-ES and EGO results. The field maps for $\Re(H_x)$ and $\Re(E_y)$ for the optimized geometries at $\lambda = 600 \text{ nm}$ are shown in Fig. 6

and achieve a unitary efficiency for light deflection, two different approaches have been proposed. The first approach is based on designing a metasurface with a surface impedance profile that takes into account the impedance matching into consideration. A balance between gain and loss elements (active metasurfaces), or nonlocal effects must be considered. This first approach is relatively difficult to be considered from the fabrication point of view, it requires complex design fabrication, and deep sub-wavelength resolution. Another technique is introduced in Ref.[61]. In this work, the authors have shown that by using a lossless bi-anisotropic resonators with coupled magnetic and electric field responses located above a ground plane (the distance between the resonators and

the ground plane must be optimized in order to achieve the maximum efficiency). Even if the authors focused on the reflective case with no losses (ideal case), they mention that this work can be extended to the case of transmitting metasurfaces, by considering a two-layer metasurface. This latest technique requires using resonators with strong magnetic and electric responses at visible frequencies.

Besides the fact that we observed a limitation in the overall efficiency, as expected from previous theoretical predictions [60, 61], these studies for cylindrical and rectangular nanoantennas allow us to confirm that the optimization strategies based on the CMA-ES and EGO methods that we have considered here are capable of *un-*

TABLE II. CMA-ES and EGO optimized parameters for the cylindrical nanopillars shown in Fig. 5(a). The optimized heights obtained from CMA-ES and EGO are $h \approx 735 \text{ nm}$ and $h \approx 708 \text{ nm}$, respectively.

CMA-ES: element dimensions ($h \approx 735 \text{ nm}$)		EGO: element dimensions ($h \approx 705 \text{ nm}$)	
Element	Diameter(nm)	Element	Diameter(nm)
1	209.7	1	210.0
2	176.2	2	189.0
3	149.1	3	154.5
4	117.6	4	126.0
CMA-ES: distance between elements (center to center)		EGO: distance between elements (center to center)	
between (1 – 2)	308.6 nm	between (1 – 2)	293.9 nm
between (2 – 3)	255.7 nm	between (2 – 3)	279.4 nm
between (3 – 4)	249.3 nm	between (3 – 4)	236.9 nm

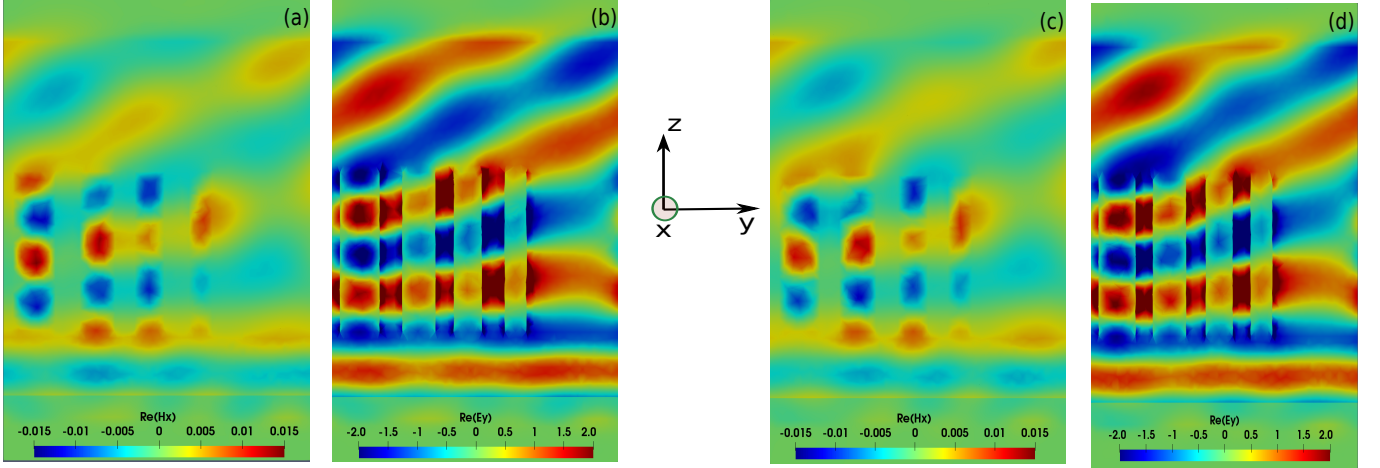


FIG. 6. Field maps of $\Re(H_x)$ and $\Re(E_y)$ obtained for cylindrical nanopillars for the optimized geometries at $\lambda = 600 \text{ nm}$. (a-b): obtained from the CMA-ES method with height $h \approx 735 \text{ nm}$, (c-d): obtained from the EGO method with $h \approx 708 \text{ nm}$.

CMA-ES: element dimensions		CMA-ES: distance between elements (center to center)	
Element	Diameter(nm)		
1	205.1	between (1 – 2)	328.0 nm
2	176.4	between (2 – 3)	264.3 nm
3	142.3	between (3 – 4)	238.5 nm
4	128.2		

TABLE III. CMA-ES optimized parameters for the cylindrical nanopillars shown in Fig. 5(a) with fixed height $h = 800 \text{ nm}$ during the optimization.

derstanding which key parameter influence the most on the minimization process. Indeed, for all the results presented herein, we figured out that both optimizers set the thickness of the first ridge (d_y in case of rectangular nanopillars, and diameter in case of cylindrical nanopillars) to the largest values, for which the effective refractive index is the largest. The height of the ridges is another interesting parameters as it essentially relates to the different longitudinal modes inside the nanopillars, leading to relatively well optimized diffraction efficiency at different height.

CONCLUSION

Two advanced optimization strategies have been introduced to optimize 3D semiconductor-based metasurfaces. We proved that our optimization methods are very efficient in obtaining different global minima/maxima, requiring only a few hundreds of electromagnetic solver calls. We applied these methods to maximize the light deflection efficiency at $\lambda = 600 \text{ nm}$ using GaN based metasurfaces. Our numerical results reveal that one can obtain 85% of deflection efficiency for both TM and TE polarization using cylindrical nanoantennas. In addition, we show that using rectangular nanoantennas, one might

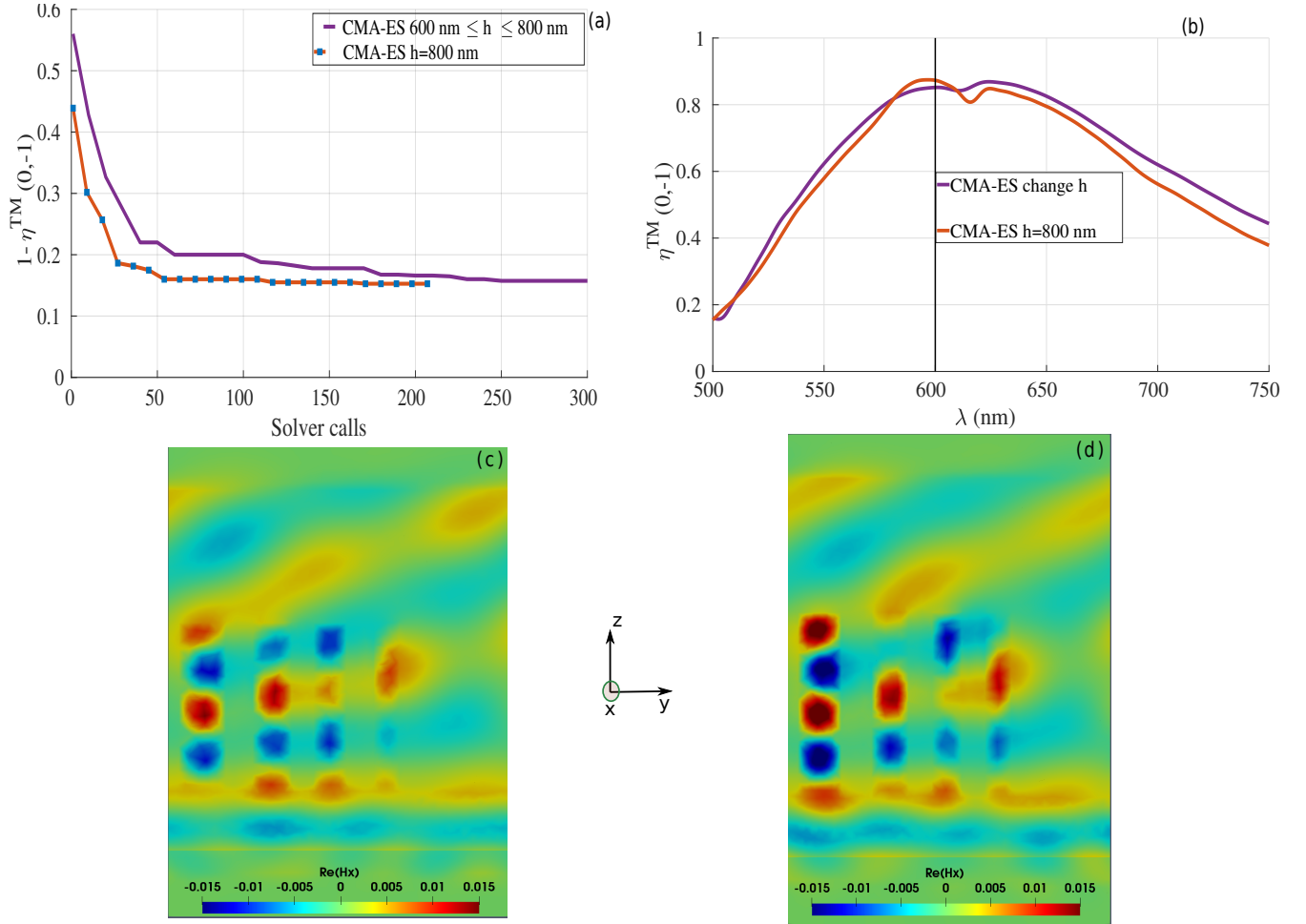


FIG. 7. (a): comparison between results obtained with CMA-ES (for the cylindrical nanopillars) as a function of height (purple curve, exactly as in Fig. 5(b) in which the height is optimized to $h \approx 735 \text{ nm}$), and the ones obtained with CMA-ES with fixed $h = 800 \text{ nm}$ (dark orange curve) as a function of the number of fullwave solver calls. The corresponding optimized parameter values for the case with $h = 800 \text{ nm}$ (dark orange curve) can be found in Tab. III. (b): diffraction efficiency as a function of the wavelength. (c) and (d) field maps of $\Re(H_x)$ at $\lambda = 600 \text{ nm}$ for the reference case with $h \approx 735 \text{ nm}$ (purple curve in (a)), and the case with $h = 800 \text{ nm}$ (dark orange curve in (a)).

obtain more than 88% of deflection efficiency for TM polarization. These methods, which are widely used in the computational fluid dynamics community, could have significant implications in the design of efficient metasurfaces, notably in view of their utilization in real world applications, for which reaching highly efficient designs in reasonable computation time is one of the most important figure of merit. Our future works will aim at improving the computational performances of these optimization strategies by leveraging the different levels of parallelism underlying their algorithmic structures one one hand, and extending their capabilities in view of dealing with multi-objective problems.

ACKNOWLEDGEMENTS

This research was funded by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement no. 639109). The authors acknowledge support from French defence procurement agency under the ANR ASTRID Maturation program, grant agreement number ANR-18-ASMA-0006-01.

AUTHOR CONTRIBUTIONS STATEMENT

M.M.R.E conducted the simulation for the DGTD and the optimization. R.D developed the optimization methods. P.G and S.L proposed the structure. M.M.R.E, R.D, and S.L analysed the optimization results. M.M.R.E and P.G analysed the physical results. M.M.R.E, R.D, S.L,

and P.G wrote the manuscript. G.B and M.S conducted the validation using FDTD solver. All authors reviewed the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

NUMERICAL VALIDATION

In this section, we discuss about the numerical validation of some of the results obtained in our work. First, we start by studying the influence of the mesh size/type using our DGTD fullwave solver from the DIOGENeS software suite [41] on some of the optimized designs. In Fig. 8(a), we study the influence of the mesh and the polynomial order on the optimized solution obtained in the rectangular nanoantennas case (the results obtained from the CMA-ES method are shown in Fig. 3). As it can be seen in Fig. 8(a), using fourth order polynomial order \mathbb{P}_4 with a coarse mesh (only 3 000 cells) provides the same results than when we consider second order polynomial order \mathbb{P}_2 with a finer mesh (with 147 000 cells). In Fig. 8(b), we show the convergence for the cylindrical nanoantennas case obtained using the CMA-ES method with $h \approx 735 \text{ nm}$ (the corresponding parameters can be found in Tab. II) using both linear tetrahedral elements and also curvilinear tetrahedral elements in order to achieve a high order approximation of the cylindrical geometry of the nanopillars. As it can be noticed, the convergence is obtained with less cells in the case of the mesh with curvilinear elements compared to the case with linear ones. To conclude on this point, using our DGTD solver, we are able to demonstrate the convergence for the optimized geometries using different mesh size and/or types, which is not trivial especially for the case of cylindrical elements that require higher order curvilinear mesh type for more accurate results, otherwise, a finer mesh must be used with the classical linear elements.

In Fig. 9, we provide another validation of our numerical results by simulating the optimized structure given in Tab. III using the FDTD fullwave solver from Lumerical commercial software, and comparing the results with the ones obtained from our DGTD solver. As it can be seen from Fig. 9, we have a very good agreement between the results obtained using the two methods for both the total transmission and the deflection efficiency of the first order mode. However, in the case of the DGTD method, we show the results using curvilinear elements with fourth order polynomial using only 13 000 cells (same results can be obtained with classical linear elements with 109000 cells as we have seen in Fig. 7). On the other hand, for the FDTD results, we consider a very fine mesh with nearly 10^5 cells in order to get accurate results that can be compared with the DGTD method.

Next, we show one example in order to demonstrate that our optimization techniques outperform the classical approach to phase gradient metasurface design. In Fig. 10(a), we present the geometry obtained using the classical approach. In this approach, one calculates the phase gradient needed with 1500 nm period in y -direction at wavelength $\lambda = 600 \text{ nm}$ and place the pillars at the right y positions in order to introduce the needed phase shifts to maximize the light deflection of the (0,-1) mode. The phase shifts and transmission introduced by each single element is calculated before using a classical forward simulation by changing the radius of the pillar and compute the corresponding phase and transmission (we consider a fixed height $h = 800 \text{ nm}$). The corresponding deflection efficiency for the first order mode can be seen in the red curve in Fig. 10(c), in which the maximum efficiency at $\lambda = 600 \text{ nm}$ is around 74%. On the other hand the optimized geometry obtained using the CMA-ES method presented in Fig. 10(b), provides nearly 85% of deflection efficiency at $\lambda = 600 \text{ nm}$ (data can be found in Tab. III). This discrepancy between the two results is clearly linked to the near field coupling as it is shown in Figs. 10(d-e). In the classical design approach (see the corresponding geometry in Fig. 9(a)), we neglect the effect of the near field coupling when we place the nanopillars together to construct the required phase shift, while in the optimization techniques presented in this paper, the near field coupling is taking into account during all optimization steps, which give us an optimized geometry as shown in Fig. 10(b).

-
- [1] Simion Astilean, Philippe Lalanne, Pierre Chavel, Edmond Cambril, and Huguette Launois. High-efficiency subwavelength diffractive element patterned in a high-refractive-index material for 633 nm. *Optics letters*, 23(7):552–554, 1998.
 - [2] Philippe Lalanne, Simion Astilean, Pierre Chavel, Edmond Cambril, and Huguette Launois. Design and fabrication of blazed binary diffractive elements with sampling periods smaller than the structural cutoff. *JOSA*

- A*, 16(5):1143–1156, 1999.
- [3] Nanfang Yu, Patrice Genevet, Mikhail A. Kats, Francesco Aieta, Jean-Philippe Tetienne, Federico Capasso, and Zeno Gaburro. Light propagation with phase discontinuities: Generalized laws of reflection and refraction. *Science*, 334(6054):333–337, 2011.
- [4] Shulin Sun, Qiong He, Shiyi Xiao, Qin Xu, Xin Li, and Lei Zhou. Gradient-index meta-surfaces as a bridge linking propagating waves and surface waves. *Nature mate-*

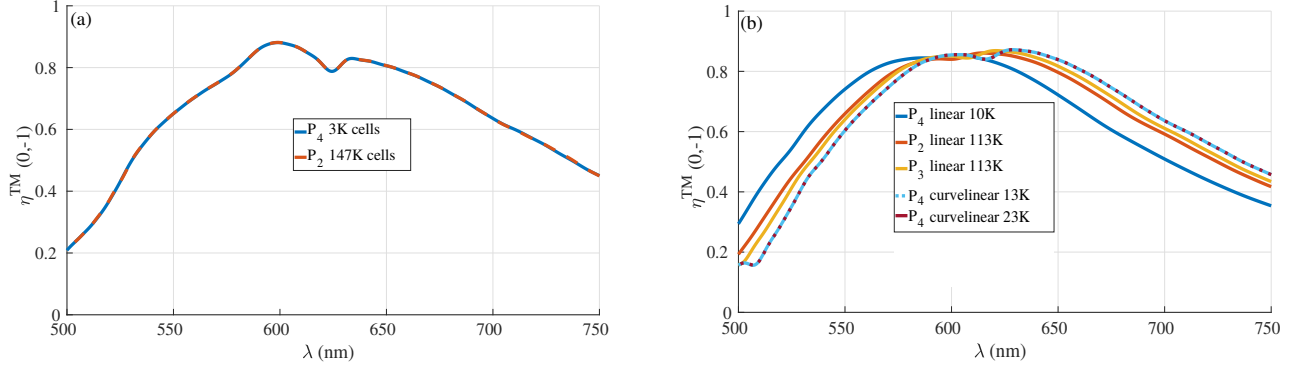


FIG. 8. Convergence study using the DGTD method (a): for the rectangular shaped antenna (for CMA-ES results, the corresponding parameters can be found in Tab. I) using coarse mesh with 3000 cells with fourth order polynomial P_4 order (blue curve) and with finer mesh size (147 000 cells) with second order polynomial P_2 order (purple curve). (b): convergence results for the cylindrical shaped antenna using the data shown in Tab. II for the CMA-ES results using different mesh sizes and/or types, different polynomial orders.

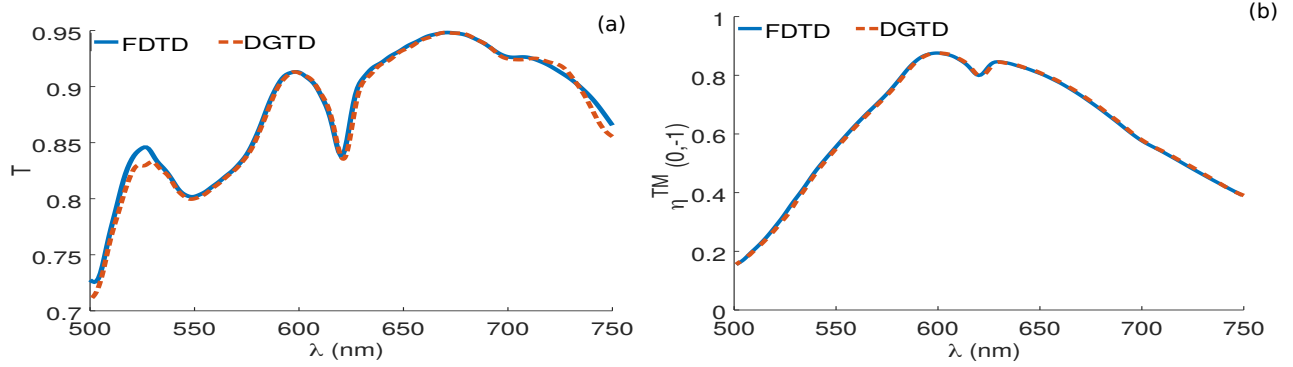


FIG. 9. Comparison between results obtained with the FDTD method (blue solid curves) and the ones obtained with the DGTD method (orange dashed curves) for the optimized cylindrical nanopillars shown in Tab. III. (a) for the total transmission and (b) for the deflection efficiency for the first order mode. The FDTD results are obtained with nearly 10^5 cells, however in our DGTD solver, we used only 13 000 cells with higher order polynomial order and curvilinear elements.

rials, 11(5):426, 2012.

- [5] Dianmin Lin, Pengyu Fan, Erez Hasman, and Mark L. Brongersma. Dielectric gradient metasurface optical elements. *Science*, 345(6194):298–302, 2014.
- [6] Alexander Y Zhu, Arseniy I Kuznetsov, Boris Luk'yanchuk, Nader Engheta, and Patrice Genevet. Traditional and emerging materials for optical metasurfaces. *Nanophotonics*, 6(2):452–471, 2017.
- [7] Patrice Genevet, Federico Capasso, Francesco Aieta, Mohammadreza Khorasaninejad, and Robert Devlin. Recent advances in planar optics: from plasmonic to dielectric metasurfaces. *Optica*, 4(1):139–152, Jan 2017.
- [8] Mane-Si Laure Lee, Ph Lalanne, JC Rodier, P Chavel, E Cambril, and Y Chen. Imaging with blazed-binary diffractive elements. *Journal of Optics A: Pure and Applied Optics*, 4(5):S119, 2002.
- [9] D. R. Smith, Willie J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz. Composite medium with simultaneously negative permeability and permittivity. *Phys. Rev. Lett.*, 84:4184–4187, May 2000.
- [10] Alessandro Salandrino and Nader Engheta. Far-field sub-diffraction optical microscopy using metamaterial crystals: Theory and simulations. *Phys. Rev. B*, 74:075103, Aug 2006.
- [11] Shuming Wang, Pin Chieh Wu, Vin-Cent Su, Yi-Chieh Lai, Mu-Ku Chen, Hsin Yu Kuo, Bo Han Chen, Yu Han Chen, Tzu-Ting Huang, Jung-Hsi Wang, et al. A broadband achromatic metalens in the visible. *Nature nanotechnology*, 13(3):227, 2018.
- [12] Ren Jie Lin, Vin-Cent Su, Shuming Wang, Mu Ku Chen, Tsung Lin Chung, Yu Han Chen, Hsin Yu Kuo, Jia-Wern Chen, Ji Chen, Yi-Teng Huang, et al. Achromatic metalens array for full-colour light-field imaging. *Nature nanotechnology*, 14(3):227, 2019.
- [13] Kenneth Diest, editor. *Numerical Methods for Metamaterial Design*, volume 127 of *Topics in Applied Physics*. Springer Netherlands, Dordrecht, 2013.
- [14] Kan Yao, Rohit Unni, and Yuebing Zheng. Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale. *Nanophotonics*, 2018.
- [15] Sean Molesky, Zin Lin, Alexander Y Piggott, Weiliang Jin, Jelena Vuckovic, and Alejandro W Rodriguez. Outlook for inverse design in nanophotonics. *arXiv preprint arXiv:1801.06715*, 2018.

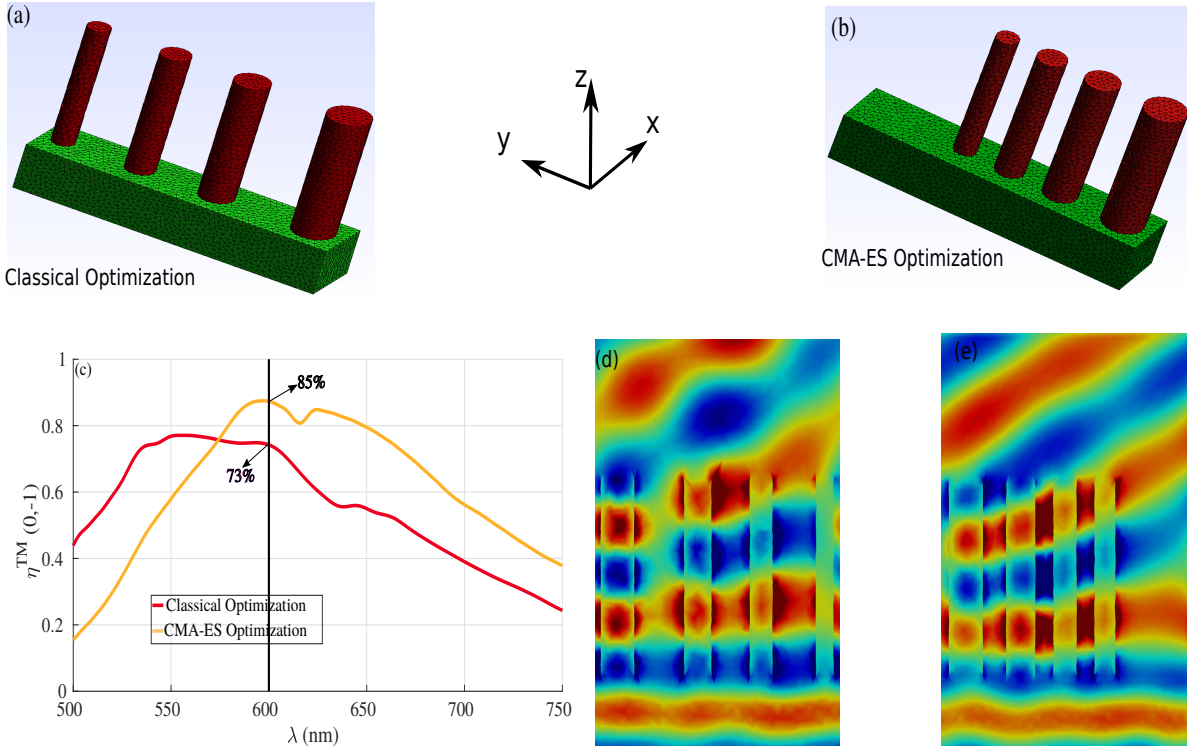


FIG. 10. Comparison between the classical approach to phase gradient metasurface design and our optimized geometries for the cylindrical nanopillars with $h = 800 \text{ nm}$. (a) and (b): the geometry obtained using the classical approach in which each nanopillar is optimized manually by changing the diameter and finally placed together in order to obtain the desired phase shift needed to maximize the light deflection for the first order mode at $\lambda = 600 \text{ nm}$ with period 1500 nm in y -direction. (c): results obtained using the CMA-ES for the cylindrical nanopillars (see Tab. III) for the corresponding parameters. (c): comparison between the deflection efficiency for the first order mode obtained using the classical (red curve) and the CMA-ES (orange curve). (d) and (e) represent field maps of $\Re\epsilon(E_y)$ obtained using the classical optimization design and the CMA-ES results, respectively.

- [16] Jianji Yang and Jonathan A. Fan. Topology-optimized metasurfaces: impact of initial geometric layout. *Optics Letters*, 42(16):3161, August 2017.
- [17] Jun Rong Ong, Hong Son Chu, Valerian Hongjie Chen, Alexander Yutong Zhu, and Patrice Genevet. Freestanding dielectric nanohole array metasurface for mid-infrared wavelength applications. *Opt. Lett.*, 42(13):2639–2642, Jul 2017.
- [18] Jakob Søndergaard Jensen and Ole Sigmund. Topology optimization for nano-photonics. *Laser & Photonics Reviews*, 5(2):308–321, 2011.
- [19] David Sell, Jianji Yang, Sage Doshay, Kai Zhang, and Jonathan A Fan. Visible light metasurfaces based on single-crystal silicon. *ACS Photonics*, 3(10):1919–1925, 2016.
- [20] David Sell, Jianji Yang, Sage Doshay, and Jonathan A Fan. Periodic dielectric metasurfaces with high-efficiency, multiwavelength functionalities. *Advanced Optical Materials*, 5(23):1700645, 2017.
- [21] David Sell, Jianji Yang, Evan W Wang, Thaibao Phan, Sage Doshay, and Jonathan A Fan. Ultra-high-efficiency anomalous refraction with dielectric metasurfaces. *ACS Photonics*, 5(6):2402–2407, 2018.
- [22] Thaibao Phan, David Sell, Evan W Wang, Sage Doshay, Kofi Edee, Jianji Yang, and Jonathan A Fan. High-efficiency, large-area, topology-optimized metasurfaces. *Light: Science & Applications*, 8(1):48, 2019.
- [23] Alexander Y. Piggott, Jesse Lu, Thomas M. Babinec, Konstantinos G. Lagoudakis, Jan Petykiewicz, and Jelena Vučković. Inverse design and implementation of a wavelength demultiplexing grating coupler. *Scientific Reports*, 4(1), May 2015.
- [24] F. Callewaert, V. Velez, P. Kumar, A. V. Sahakian, and K. Aydin. Inverse-Designed Broadband All-Dielectric Electromagnetic Metadevices. *Scientific Reports*, 8(1), December 2018.
- [25] Logan Su, Rahul Trivedi, Neil V. Sapra, Alexander Y. Piggott, Dries Vercruyssen, and Jelena Vučković. Fully-automated optimization of grating couplers. *Optics Express*, 26(4):4023, February 2018.
- [26] Victor Egorov, Michal Eitan, and Jacob Scheuer. Genetically optimized all-dielectric metasurfaces. *Optics express*, 25(3):2583–2593, 2017.
- [27] Samad Jafar-Zanjani, Sandeep Inampudi, and Hossein Mosallaei. Adaptive genetic algorithm for optical metasurfaces design. *Scientific reports*, 8(1):11040, 2018.
- [28] Alexej V Pogrebnjakov, Jeremy A Bossard, Jeremiah P Turpin, J David Musgraves, Hee Jung Shin, Clara Rivero-Baleine, Nikolas Podraza, Kathleen A Richardson, Douglas H Werner, and Theresa S Mayer. Reconfigurable near-ir metasurface based on ge 2 sb 2 te 5 phase-change material. *Optical Materials Express*, 8(8):2264–

- 2275, 2018.
- [29] Krupali D. Donda and Ravi S. Hegde. Rapid design of wide-area heterogeneous electromagnetic metasurfaces beyond the unit-cell approximation. *Progress In Electromagnetics Research M*, 60:1–10, 2017.
- [30] Peter R Wiecha, Arnaud Arbouet, Christian Girard, Aurélie Lecestre, Guilhem Larrieu, and Vincent Paillard. Evolutionary multi-objective optimization of colour pixels based on dielectric nanoantennas. *Nature nanotechnology*, 12(2):163, 2017.
- [31] John Peurifoy, Yichen Shen, Li Jing, Yi Yang, Fidel Cano-Renteria, Brendan G DeLacy, John D Joannopoulos, Max Tegmark, and Marin Soljačić. Nanophotonic particle simulation and inverse design using artificial neural networks. *Science advances*, 4(6):eaar4206, 2018.
- [32] Sandeep Inampudi and Hossein Mosallaei. Neural network based design of metagratings. *Applied Physics Letters*, 112(24):241102, 2018.
- [33] Zhaocheng Liu, Dayu Zhu, Sean Rodrigues, Kyu-Tae Lee, and Wenshan Cai. A generative model for the inverse design of metasurfaces. *Nano letters*, 2018.
- [34] Dianjing Liu, Yixuan Tan, Erfan Khoram, and Zongfu Yu. Training deep neural networks for the inverse design of nanophotonic structures. *ACS Photonics*, 5(4):1365–1369, 2018.
- [35] Mohammad H Tahersima, Keisuke Kojima, Toshiaki Koike-Akino, Devesh Jha, Bingnan Wang, Chungwei Lin, and Kieran Parsons. Deep neural network inverse design of integrated nanophotonic devices. *arXiv preprint arXiv:1809.03555*, 2018.
- [36] Nikolai Schmitt, Niklas Georg, Gauthier Brière, Dimitrios Loukrezis, Sébastien Héron, Stéphane Lanteri, Charalambos Klitis, Marc Sorel, Ulrich Römer, Herbert De Gerssem, Stéphane Vézian, and Patrice Genevet. Optimization and uncertainty quantification of gradient index metasurfaces. *Opt. Mater. Express*, 9(2):892–910, Feb 2019.
- [37] Gauthier Brière, Peinan Ni, Sébastien Héron, Sebastien Chenot, Stéphane Vézian, Virginie Brändli, Benjamin Damilano, Jean-Yves Duboz, Masanobu Iwanaga, and Patrice Genevet. An etching-free approach toward large-scale light-emitting metasurfaces. *Advanced Optical Materials*, page 1801271, 2019.
- [38] David Sell, Jianji Yang, Sage Doshay, Rui Yang, and Jonathan A Fan. Large-angle, multifunctional metagratings based on freeform multimode geometries. *Nano letters*, 17(6):3752–3757, 2017.
- [39] Zhenpeng Zhou, Juntao Li, Rongbin Su, Beimeng Yao, Hanlin Fang, Kezheng Li, Lidan Zhou, Jin Liu, Daan Stellinga, Christopher P Reardon, et al. Efficient silicon metasurfaces for visible light. *ACS Photonics*, 4(3):544–551, 2017.
- [40] Philippe Lalanne, Simion Astilean, Pierre Chavel, Edmond Cambriil, and Huguette Launois. Blazed binary subwavelength gratings with efficiencies larger than those of conventional échelette gratings. *Optics letters*, 23(14):1081–1083, 1998.
- [41] *Diogenes: A Discontinuous-Galerkin based software suite for nano-optics*. <https://diogenes.inria.fr/>.
- [42] K.S. Yee. Numerical solution of initial boundary value problems involving Maxwell’s equations in isotropic media. *IEEE Trans. Antennas and Propag.*, 14(3):302–307, 1966.
- [43] A. Taflove and S.C. Hagness. *Computational electrodynamics: the finite-difference time-domain method - 3rd ed*. Artech House Publishers, 2005.
- [44] L. Fezoui, S. Lanteri, S. Lohrengel, and S. Piperno. Convergence and stability of a discontinuous Galerkin time-domain method for the 3D heterogeneous Maxwell equations on unstructured meshes. *ESAIM: Math. Model. Numer. Anal.*, 39(6):1149–1176, 2005.
- [45] J. Viquerat. *Simulation of electromagnetic waves propagation in nano-optics with a high-order discontinuous Galerkin time-domain method*. PhD thesis, University of Nice-Sophia Antipolis, Dec 2015.
- [46] N. Hansen, S.D. Muller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18, 2003.
- [47] Randy L Haupt and Douglas H Werner. *Genetic algorithms in electromagnetics*. John Wiley & Sons, 2007.
- [48] Peter E Sieber and Douglas H Werner. Infrared broadband quarter-wave and half-wave plates synthesized from anisotropic bézier metasurfaces. *Optics express*, 22(26):32371–32383, 2014.
- [49] Idellyse Martinez, Anastasios H Panaretos, Douglas H Werner, Giacomo Oliveri, and Andrea Massa. Ultra-thin reconfigurable electromagnetic metasurface absorbers. In *2013 7th European Conference on Antennas and Propagation (EuCAP)*, pages 1843–1847. IEEE, 2013.
- [50] Bei Huang, Qiang Cheng, Gang Yong Song, and Tie Jun Cui. Design of acoustic metamaterials using the covariance matrix adaptation evolutionary strategy. *Applied Physics Express*, 10(3):037301, 2017.
- [51] J Nagar, SD Campbell, and DH Werner. Apochromatic singlets enabled by metasurface-augmented grin lenses. *Optica*, 5(2):99–102, 2018.
- [52] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2), 2001.
- [53] R.L. Haupt. An introduction to genetic algorithms for electromagnetics. *IEEE Antennas and Propagation Magazine*, 37(2):7–15, April 1995.
- [54] D.R. Jones. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4), 1998.
- [55] D.R. Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, 21:345–383, 2001.
- [56] Carlo Forestiere, Yanyan He, Ren Wang, Robert M Kirby, and Luca Dal Negro. Inverse design of metal nanoparticles’ morphology. *ACS Photonics*, 3(1):68–78, 2015.
- [57] J. Sacks, W.J. Welch, T.J. Mitchell, and H.P. Wynn. Design and analysis of computer experiments. *Statistical Science*, 4(4):409–435, 1989.
- [58] David J.C. MacKay. Bayesian interpolation. *Neural Computation*, 4(415–447), 1991.
- [59] Bo Han Chen, Pin Chieh Wu, Vin-Cent Su, Yi-Chieh Lai, Cheng Hung Chu, I Chen Lee, Jia-Wern Chen, Yu Han Chen, Yung-Chiang Lan, Chieh-Hsiung Kuan, et al. Gan metalens for pixel-level full-color routing at visible light. *Nano letters*, 17(10):6345–6352, 2017.
- [60] Nasim Mohammadi Estakhri and Andrea Alù. Wavefront transformation with gradient metasurfaces. *Physical Review X*, 6(4):041008, 2016.
- [61] Younes Ra’di, Dimitrios L Sounas, and Andrea Alù. Metagratings: Beyond the limits of graded metasur-

faces for wave front control. *Physical review letters*,
119(6):067404, 2017.