



Cite this: *Digital Discovery*, 2024, **3**, 381

Multi-BOWS: multi-fidelity multi-objective Bayesian optimization with warm starts for nanophotonic structure design†

Jungtaek Kim, ^a Mingxuan Li, ^a Yirong Li, ^a Andrés Gómez, ^b Oliver Hinder ^a and Paul W. Leu*^a

The design of optical devices is a complex and time-consuming process. To simplify this process, we present a novel framework of multi-fidelity multi-objective Bayesian optimization with warm starts, called Multi-BOWS. This approach automatically discovers new nanophotonic structures by managing multiple competing objectives and utilizing multi-fidelity evaluations during the design process. We employ our Multi-BOWS method to design an optical device specifically for transparent electromagnetic shielding, a challenge that demands balancing visible light transparency and effective protection against electromagnetic waves. Our approach leverages the understanding that simulations with a coarser mesh grid are faster, albeit less accurate than those using a denser mesh grid. Unlike the earlier multi-fidelity multi-objective method, Multi-BOWS begins with faster, less accurate evaluations, which we refer to as “warm-starting,” before shifting to a dense mesh grid to increase accuracy. As a result, Multi-BOWS demonstrates 3.2–89.9% larger normalized area under the Pareto frontier, which measures a balance between transparency and shielding effectiveness, than low-fidelity only and high-fidelity only techniques for the nanophotonic structures studied in this work. Moreover, our method outperforms an existing multi-fidelity method by obtaining 0.5–10.3% larger normalized area under the Pareto frontier for the structures of interest.

Received 7th September 2023
Accepted 22nd November 2023

DOI: 10.1039/d3dd00177f
rsc.li/digitaldiscovery

Introduction

Electrodynamic simulations play an essential role in the design of optical devices for a range of applications, including waveguides, photonic crystals, lenses, plasmonics, solar cells, and nanophotonics.^{1–3} These simulations involve solving Maxwell’s equations to examine the interaction of electromagnetic waves with different materials and structures. This process allows us to compute various optical properties and understand how to control light, which is relevant for optical devices. However, several challenges are associated with the design of optical devices. These include defining parameters to optimize, considering multiple objectives, and balancing evaluation time and accuracy.

Designing an optical device requires defining a parametric design space for these devices and identifying specific objective functions to optimize. However, the design process of an optical device can be complex due to the need to balance several distinct competing objectives over many parameters. For instance, in lens design, factors such as resolution, wavelength

range, and field angles need to be considered.⁴ Antireflection coating design requires the minimization of reflection at multiple wavelengths and angles.^{5,6} For light-emitting diodes, considerations include efficiency, color rendering, lifetime, and thermal management.⁷

We can use one of several electrodynamic methods, such as rigorous coupled-wave analysis,⁸ finite element method,⁹ or finite-difference time-domain method,¹⁰ to simulate an optical device. Interestingly, these simulation methods involve different levels of fidelity, such as mesh resolution, frequency domain decomposition, and time step. Different nanophotonic structures can be evaluated at lower fidelity, which is less expensive but more prone to noise, or at higher fidelity, which is costlier but yields more accurate results. Both low-fidelity and high-fidelity evaluations are valuable due to their unique properties related to accuracy and time efficiency.

To efficiently design an optimal optical device, we propose a framework of multi-fidelity multi-objective Bayesian optimization with warm starts, called Multi-BOWS. This framework combines multiple objectives and multi-fidelity evaluations in the design process of optical devices with electrodynamic simulations. We utilize Bayesian optimization,^{11–13} a sample-efficient technique for black-box optimization, which has been shown to effectively automate structure discovery.^{14–24} This automatic discovery process allows us to investigate a high-

^aUniversity of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA. E-mail: pleu@pitt.edu

^bUniversity of Southern California, Los Angeles, California 90089, USA

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d3dd00177f>



dimensional search space of disparate nanophotonic structures, reducing human intervention in the design process. Specifically, we use the Pareto frontier of low-fidelity evaluations to kickstart the high-fidelity Bayesian optimization by providing better initial points and thereby accelerating the optimization process.

We demonstrate the effectiveness of our Multi-BOWS method in the specific context of designing optical devices for transparent electromagnetic shielding. This requires a structure with high visible transparency and efficient electromagnetic shielding. Our findings show that Multi-BOWS outperforms several approaches that use low-fidelity evaluations only, high-fidelity evaluations only, or a multi-fidelity approach that uses a mix of both.²⁵ In particular, our method achieves 3.2–89.9% larger normalized area under the Pareto frontier (AUPF) than the low-fidelity only and high-fidelity only techniques for the nanophotonic structures investigated in our work. Moreover, it achieves 0.5–10.3% larger AUPF than the earlier multi-fidelity method for the structures studied in this work.

Preliminaries

In this section, we delve into the challenges of optical device design for transparent electromagnetic shielding. Then, we discuss the nanophotonic structures under consideration and the Bayesian optimization strategy that will be used to discover novel structures.

Electromagnetic shielding is crucial for safeguarding electronic devices and circuits by mitigating electromagnetic interference.^{26–31} This has been a major research focus for a variety of applications such as protecting RFID chips from radio-frequency interference and shielding medical implants from electromagnetic waves. Besides reducing interference, some applications such as consumer electronics, automotive and aviation, medical devices, and building windows need to fulfill additional design objectives like visible transparency. The simultaneous consideration of several different factors complicates the task of identifying the most effective structure.

Formally, suppose that we have an objective for transparency, denoted as f_{tr} , and another for shielding effectiveness (SE), denoted as f_{se} . An optimal structure for transparency \mathbf{x}_{tr}^* and one for SE \mathbf{x}_{se}^* can be defined by solving the following equations:

$$\mathbf{x}_{tr}^* = \arg \max_{\mathbf{x} \in \mathcal{X}} f_{tr}(\mathbf{x}), \quad (1)$$

$$\mathbf{x}_{se}^* = \arg \max_{\mathbf{x} \in \mathcal{X}} f_{se}(\mathbf{x}), \quad (2)$$

where \mathbf{x} represents a nanophotonic structure and \mathcal{X} is the search space for optical device design. It is important to consider the trade-off between these two objectives – we want to devise a nanophotonic structure that maximizes both f_{tr} and f_{se} .^{29,32} However, optimizing both eqn (1) and (2) is a complex task as the optimal solutions \mathbf{x}_{tr}^* and \mathbf{x}_{se}^* are not likely to coincide.

In addition to the aforementioned complexities of multi-objective optimization, a specific expression of f_{tr} cannot be

explicitly obtained for many structures and requires electrodynamic simulations. Simulating a nanophotonic structure to evaluate f_{tr} is a time-consuming task because an accurate evaluation requires a dense mesh grid. These challenges make a compelling case for employing a black-box optimization technique for a costly function. Notably, Bayesian optimization is a sample-efficient black-box optimization strategy that stands as a suitable candidate to tackle this problem.^{11–13} Furthermore, by utilizing the nature of mesh-based simulations, we can evaluate less expensive, albeit noisier functions using a coarse mesh grid, rather than more expensive but more accurate functions using a dense mesh grid. Hence, we can define a low-fidelity multi-objective function $[f_{tr}^{low}, f_{se}^{low}]$ and a high-fidelity multi-objective function $[f_{tr}^{high}, f_{se}^{high}]$. These functions, with varying degree of accuracy, help us to select an optimal structure concerning both objectives from a multi-fidelity optimization standpoint. This approach allows us to strike a balance between evaluation accuracy and time.

Nanophotonic structures for transparent electromagnetic shielding

Transparent electromagnetic shielding, which allows for efficient transmission of visible light, is crucial for various optoelectronic applications. Metal meshes have been widely explored in the pursuit of high transparency and low sheet resistance, which are essential for electromagnetic shielding.³¹ Meanwhile, to enhance the visible transmission of silver films, many researchers have investigated the encapsulation of the silver layer with high-index dielectric materials. ITO/Ag–Cu/ITO structures have achieved 96.5% transmittance and 26 dB SE,²⁹ while ZnO/Ag/ZnO sandwich structures have shown 88.9% transmittance in the visible range and 35 dB SE.³³ Furthermore, to improve the performance of sandwich structures, nanocone structures have been proposed. These structures enhance the antireflection effect by using a graded refractive index. Double-sided nanocone sandwiches demonstrate 90.8% average visible transmittance with 41.2 dB SE and 95.1% average visible transmittance with 35.6 dB SE.³² Suggestions have been made to explore different cone geometries to break traditional performance limits and to understand the fabrication sensitivity of these structures better.^{34–37} It is noteworthy that these nanocone structures could be fabricated by maskless reactive ion etching^{5,6,38} or nanosphere lithography combined with etching.³⁹ However, designing nanocone structures introduce the need to optimize over many parameters, necessitating a large number of structure evaluations.

Automatic structure discovery

Automatic structure discovery, the pursuit of an optimal structure, has been actively studied in diverse research fields. These include protein structure discovery,^{19,20} drug discovery,²¹ neural architecture search,^{22,23} and causal discovery.^{24,40,41} All these fields share the challenge of seeking optimal outcomes in a vast landscape of possible structures, akin to finding a needle in a haystack.



To overcome this challenge, it is necessary to define three key elements:

- structure representation: this is the process of expressing a structure of interest as a specific type of input, such as discrete variables,
- evaluation function: this is used to assess the performance of a particular structure, and
- decision-making policy: this is a strategy used to identify potential optimal structures based on previously evaluated structures and their corresponding evaluations.

In this paper, we carefully design structure representations, taking into account the structures described in this section and the feasibility of structures. The evaluation function is then defined based on the structure representation to measure specific properties. Our framework considers the multiple objectives of transparency and SE. Moreover, this evaluation function of transparency is inherently black-box, as it cannot be explicitly expressed as a function. Lastly, the decision-making process incorporates both the structure representation and the evaluation function, to sequentially recommend optimal structure candidates.

On the other hand, topology optimization can be employed in the design of photonic structures, leveraging gradient information with respect to these structures.^{42–44} Previous research has shown that combining adjoint methods with topology optimization is a powerful approach for tackling inverse design problems in photonics.^{45–48} These methods rely on gradients and typically use gradient-based optimization techniques to find solutions. However, these approaches may be limited when objective functions are complex and it is important to find a global optimum as opposed to local optima.

Bayesian optimization

Bayesian optimization^{11–13} has been reported in various studies as a powerful method for identifying optimal solutions for black-box functions^{49–51} where evaluations are costly.^{14–18,49–53} It is important to note that the efficacy of this method may diminish as the number of parameters increases and managing a surrogate model becomes increasingly complex. However, Bayesian optimization has been shown to perform well compared to other competitors for black-box optimization, such as DIRECT and evolutionary algorithms.^{54–57}

Its strengths have been validated in attractive real-world problems, including optimizing chemical reactions,¹⁸ battery charging protocols,¹⁶ automatic chemical design,¹⁷ and automated machine learning.^{52,53} Building on this work, Bayesian optimization is particularly well-suited for optimizing nanophotonic structures where a structure representation and evaluation functions are already provided. Specifically, it excels in optimizing objectives when categorical and discrete variables are present.^{51,58,59}

Suppose that we do not know an objective function f and can only evaluate a d -dimensional query point $\mathbf{x} \in \mathcal{X}$ from f , where \mathcal{X} is a d -dimensional search space, *i.e.*, typically a hypercube. Bayesian optimization sequentially optimizes f by selecting a solution candidate at each iteration. Initially, we construct

a surrogate function, often using probabilistic regression, based on the points already evaluated and their evaluations. Gaussian process regression is a popular surrogate function in the Bayesian optimization community,⁶⁰ though other models such as random forests,⁶¹ tree-based surrogate models,⁵⁹ and Bayesian neural networks⁶² can also be used. For our problem, we utilize a Gaussian process-based surrogate model. Using the surrogate function, we define an acquisition function a to select the next query point. Various acquisition functions exist, including the probability of improvement,¹¹ expected improvement,⁶³ Gaussian process upper confidence bound,⁶⁴ and a portfolio of existing acquisition functions.⁶⁵ This work uses the expected improvement, aligning with numerous studies that attest to its robustness.^{18,49,50,66}

Recent research in Bayesian optimization has explored multi-fidelity methods, which seek a balance in evaluations across varying levels of fidelity.^{67–69} In parallel, multi-objective Bayesian optimization has been developed to optimize multiple objectives simultaneously.^{70–72} Recent research efforts have sought to combine these two concepts into multi-fidelity multi-objective Bayesian optimization by the introduction of continuous fidelity levels as an optimizeable parameter^{73,74} or aiming to maximize information gain per unit cost of resources.²⁵

Structure specifications

In this section, we delve into the specific nanophotonic structures studied in this work, as illustrated in Fig. 1. In particular, we examine four following structures:

- three-layer structure,
- matched-period double-sided nanocone structure,
- unmatched-period double-sided nanocone structure, and
- meta-structure.

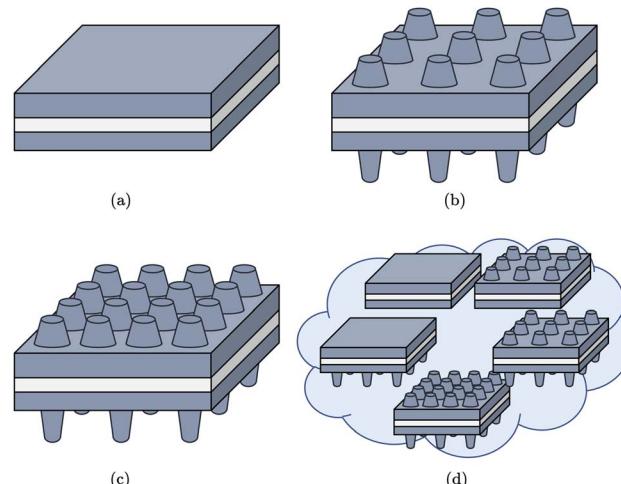


Fig. 1 Schematics of nanophotonic structures studied. Each structure is composed of silver (Ag, represented by white) and titanium dioxide (TiO_2 , shown in dark blue). (a) Three-layer structure. (b) Matched-period double-sided nanocone structure. (c) Unmatched-period double-sided nanocone structure. (d) Meta-structure.



The three-layer and matched-period double-sided nanocone structures have been previously explored.³² In this paper, we introduce two new structures: (c) the unmatched-period double-sided nanocone structure and (d) the meta-structure.

Fig. 2 provides a depiction of how the parameters of a structure are defined. The structure's parameters include the silver-layer thickness t_s , upper-layer thickness t_u , lower-layer thickness t_l , heights for upper and lower cones h_u , h_l , radii for upper cones r_{ub} , r_{ut} , radii for lower cones r_{lb} , r_{lt} , pitches for upper and lower cones a_u , a_l , and the number of upper and lower cones n_u , n_l . Depending on the specific structure, some parameters may not be applicable. For instance, in a three-layer structure, the parameters related to the upper and lower cones are disregarded, and only t_s , t_u , and t_l are utilized. For a matched-period double-sided nanocone structure, n_u and n_l are dismissed and a_u is equal to a_l .

Table 1 provides a detailed description of the parameter ranges and constraints. All parameters except for a_u and a_l are discretized to integers. Several constraints are applied as follows: $r_{ut} < r_{ub}$, $r_{lt} > r_{lb}$, $2r_{ub} \leq a_u$, $2r_{lt} \leq a_l$, and $n_u a_u = n_l a_l$. For easier management of the constraints, $r_{ut} < r_{ub}$ and $r_{lt} > r_{lb}$, we introduce new variables q_{ru} and q_{rl} as follows: $q_{ru} = r_{ut}/r_{ub}$ and $q_{rl} = r_{lb}/r_{lt}$ where q_{ru} and q_{rl} are both variables $\in [0, 1]$. Moreover, the next acquired point is only sampled over the region of the parameter space that is known to be feasible. If the proposed solution of an acquisition function violates a constraint, then it is instead evaluated at the boundary of that constraint.

Besides the three basic structures – three-layer, matched-period double-sided nanocone, and unmatched-period double-sided nanocone structures – each defined by a specific set of parameters, we introduce a new type called a meta-structure. This is a generalized structure and it is introduced to optimize the structure from the perspective of automatic structure discovery. To accommodate the meta-structure, we include an extra parameter – structure selection parameter – that allows the selection of one structure among various

structures. In our study, we consider five types of structures: three-layer, single-sided (upper) nanocone, single-sided (lower) nanocone, matched-period double-sided nanocone, and unmatched-period double-sided nanocone structures. It is worth noting that the types of structures can be easily expanded by altering the potential choices for the structure selection parameter.

Methodology

We address the issue of automatic structure discovery with multi-fidelity multi-objective Bayesian optimization with warm starts, named Multi-BOWS, which effectively incorporates knowledge from multi-fidelity evaluations and multiple objective functions. It is inspired by the methodologies previously presented.^{75,76}

Before explaining the details of Multi-BOWS, we enumerate the high-level procedure of our algorithm:

- (i) selection of initial points for low-fidelity multi-objective Bayesian optimization,
- (ii) execution of low-fidelity multi-objective Bayesian optimization, constrained by a time budget allocated for the low-fidelity Bayesian optimization,
- (iii) identification of the Pareto frontier (*i.e.*, optimal solutions) from low-fidelity evaluations,
- (iv) warm-starting of high-fidelity multi-objective Bayesian optimization using the identified Pareto frontier,
- (v) execution of high-fidelity multi-objective Bayesian optimization, constrained by a time budget allocated for this high-fidelity Bayesian optimization, and
- (vi) identification of the Pareto frontier from high-fidelity evaluations.

This procedure is visually outlined in Fig. 3. Steps (i) and (iv) act as initialization steps, Steps (ii) and (v) are considered as optimization phases, and Steps (iii) and (vi) focus on identifying the Pareto frontiers.

Firstly, a certain number of initial points are randomly sampled in Step (i), using uniform distributions or low-discrepancy sequences like the Sobol' sequence.⁷⁷ However, unlike Step (i), Step (iv) uses the Pareto frontier from low-fidelity evaluations as initial points for high-fidelity multi-objective Bayesian optimization. If the number of points on the Pareto frontier exceeds the predefined number of initial points, we randomly select the required number of points from the Pareto frontier. Steps (ii) and (v), similar to the standard Bayesian optimization algorithm, sequentially determine the query points based on the allocated time budgets.

In order to determine the next point, we first create a Gaussian process regression model to serve as a surrogate function. Given a set of data points $\mathbf{X} \in \mathbb{R}^{n \times d}$ and their corresponding responses $\mathbf{y} \in \mathbb{R}^n$, a posterior predictive distribution over $\mathbf{x} \in \mathbf{X}$ is defined by the following:

$$p(y|\mathbf{x}, \mathbf{X}, \mathbf{y}) = \mathcal{N}(\mu(\mathbf{x}|\mathbf{X}, \mathbf{y}), \sigma^2(\mathbf{x}|\mathbf{X}, \mathbf{y})), \quad (3)$$

where $\mu(\mathbf{x}|\mathbf{X}, \mathbf{y})$ is the posterior mean function and $\sigma^2(\mathbf{x}|\mathbf{X}, \mathbf{y})$ is the posterior variance function. The specific definitions for the

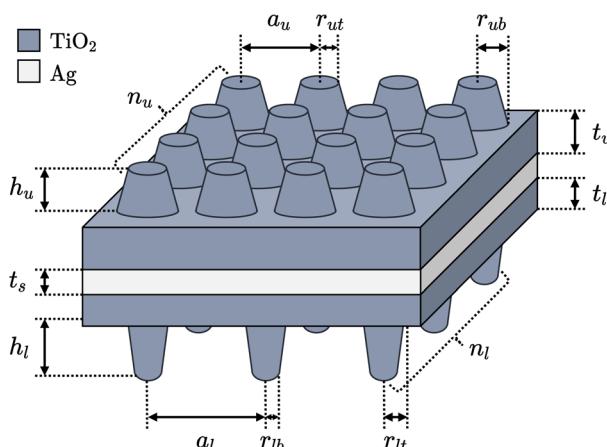


Fig. 2 Schematic of nanophotonic structures. The parameters used in this diagram apply across all structures studied. As an example, for a three-layer structure, the parameters for upper and lower cones are not used, while the other parameters remain applicable.



Table 1 Definitions, notations, ranges, and constraints applicable to parameters in nanophotonic structures. All values are in nanometers, with the exception of the last two parameters, which are unitless

Parameter	Symbol	Range	Constraints
Silver-layer thickness	t_s	{3, 4, ..., 20}	—
Upper-layer thickness	t_u	{5, 6, ..., 100}	—
Lower-layer thickness	t_l	{5, 6, ..., 100}	—
Height of upper cones	h_u	{50, 51, ..., 400}	—
Height of lower cones	h_l	{50, 51, ..., 400}	—
Pitch for upper cones	a_u	[20, 400]	$n_u a_u = n_l a_l, 2r_{ub} \leq a_u$
Pitch for lower cones	a_l	[20, 400]	$n_u a_u = n_l a_l, 2r_{lt} \leq a_l$
Bottom radius of upper cones	r_{ub}	{10, 11, ..., 100}	$r_{ut} < r_{ub}, 2r_{ub} \leq a_u$
Top radius of upper cones	r_{ut}	{1, 2, ..., 99}	$r_{ut} < r_{ub}$
Bottom radius of lower cones	r_{lb}	{1, 2, ..., 99}	$r_{lt} > r_{lb}$
Top radius of lower cones	r_{lt}	{10, 11, ..., 100}	$r_{lt} > r_{lb}, 2r_{lt} \leq a_l$
The number of upper cones	n_u	{1, 2, ..., 10}	$n_u a_u = n_l a_l$
The number of lower cones	n_l	{1, 2, ..., 10}	$n_u a_u = n_l a_l$

posterior mean and variance functions are given by the following equations:

$$\mu(\mathbf{x}|\mathbf{X}, \mathbf{y}) = \mathbf{k}(\mathbf{x}, \mathbf{X})(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}, \quad (4)$$

$$\sigma^2(\mathbf{x}|\mathbf{X}, \mathbf{y}) = k(\mathbf{x}, \mathbf{x}) - \mathbf{k}(\mathbf{x}, \mathbf{X})(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}(\mathbf{x}, \mathbf{X})^T, \quad (5)$$

where k , \mathbf{k} , and \mathbf{K} are covariance functions over two points, one point and one array of points, and two arrays of points, respectively, σ_n^2 is a noise variance, and $\mathbf{I} \in \mathbb{R}^{n \times n}$ is an identity matrix. For example, an exponentiated quadratic kernel $k(\mathbf{x}, \mathbf{x}') = s^2 \exp(-\|\mathbf{x} - \mathbf{x}'\|_2^2/2l^2)$ can be employed where s^2 is a signal scale and l is a length scale. As we aim to optimize two objectives, transparency and SE, at both low fidelity and high fidelity, surrogate functions should be constructed for both low fidelity and high fidelity. In particular, $(\mu_{tr}^{\text{low}}, \sigma_{tr}^{\text{low}})$ and $(\mu_{se}^{\text{low}}, \sigma_{se}^{\text{low}})$, as featured in eqn. (1) and (2), define the surrogate functions for low fidelity, and $(\mu_{tr}^{\text{high}}, \sigma_{tr}^{\text{high}})$ and $(\mu_{se}^{\text{high}}, \sigma_{se}^{\text{high}})$ are used to define surrogate functions for high fidelity.

By using four surrogate functions, we are able to define the corresponding acquisition functions: a_{tr}^{low} , a_{se}^{low} , a_{se}^{high} , and a_{se}^{high} . This assumes the use of the expected improvement acquisition function:⁶³

$$a(\mathbf{x}|\mathbf{X}, \mathbf{y}) = \begin{cases} \sigma(\mathbf{x}|\mathbf{X}, \mathbf{y})(z(\mathbf{x})\Phi(z(\mathbf{x})) + \phi(z(\mathbf{x}))) & \text{if } \sigma^2(\mathbf{x}) > 0, \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

where $z(\mathbf{x}) = (\mu(\mathbf{x}|\mathbf{X}, \mathbf{y}) - \max \mathbf{y})/\sigma(\mathbf{x}|\mathbf{X}, \mathbf{y})$, Φ is the cumulative distribution function of standard normal distribution, and ϕ is the probability density function of standard normal distribution.

To handle multiple objective functions, we use a random scalarization technique:⁷⁸

$$a^{\text{low}}(\mathbf{x}|\mathbf{X}, \mathbf{y}) = a_{tr}^{\text{low}}(\mathbf{x}|\mathbf{X}, \mathbf{y}) + 10^{\lambda^{\text{low}}} a_{se}^{\text{low}}(\mathbf{x}|\mathbf{X}, \mathbf{y}), \quad (7)$$

$$a^{\text{high}}(\mathbf{x}|\mathbf{X}, \mathbf{y}) = a_{tr}^{\text{high}}(\mathbf{x}|\mathbf{X}, \mathbf{y}) + 10^{\lambda^{\text{high}}} a_{se}^{\text{high}}(\mathbf{x}|\mathbf{X}, \mathbf{y}), \quad (8)$$

where $\lambda^{\text{low}}, \lambda^{\text{high}} \sim \mathcal{U}(\alpha, \beta)$. The coefficients λ^{low} and λ^{high} are sampled every iteration of Bayesian optimization, in order to

efficiently identify Pareto frontiers. In this paper, we set $\alpha = -2$ and $\beta = 2$. We then optimize eqn (1) and (2) to determine a query point:

$$\mathbf{x}_*^{\text{low}} = \arg \max_{\mathbf{x} \in \mathcal{X}} a^{\text{low}}(\mathbf{x}|\mathbf{X}, \mathbf{y}), \quad (9)$$

$$\mathbf{x}_*^{\text{high}} = \arg \max_{\mathbf{x} \in \mathcal{X}} a^{\text{high}}(\mathbf{x}|\mathbf{X}, \mathbf{y}), \quad (10)$$

for Steps (ii) and (v), respectively. Given time budgets for low fidelity and high fidelity, T^{low} and T^{high} , we repeat eqn (1) and (2) until the allotted time budget is exhausted. Then, as described in Step (iii), the Pareto frontier of the query points acquired by eqn (9), denoted as \mathcal{P}^{low} , is used as the initial points of the high-fidelity multi-objective Bayesian optimization:

$$\mathcal{P}^{\text{low}} = \left\{ \left[y_{i,tr}^{\text{low}}, y_{i,se}^{\text{low}} \right] : \left\{ \left[y_{j,tr}^{\text{low}}, y_{j,se}^{\text{low}} \right] : i \neq j, \left[y_{i,tr}^{\text{low}}, y_{i,se}^{\text{low}} \right] < \left[y_{j,tr}^{\text{low}}, y_{j,se}^{\text{low}} \right] \right\} \right\} \\ = \emptyset, \quad (11)$$

where $[y_{i,tr}^{\text{low}}, y_{i,se}^{\text{low}}]$ is the i -th low-fidelity evaluation by two objectives and $[y_{i,tr}^{\text{low}}, y_{i,se}^{\text{low}}] < [y_{j,tr}^{\text{low}}, y_{j,se}^{\text{low}}]$ implies that both $y_{i,tr}^{\text{low}} < y_{j,tr}^{\text{low}}$ and $y_{i,se}^{\text{low}} < y_{j,se}^{\text{low}}$ are satisfied. Similarly, the Pareto frontier of high-fidelity multi-objective Bayesian optimization can be readily computed using the query points acquired by eqn (10).

Simulations

We conduct electrodynamic simulations on the aforementioned nanophotonic structures. Our goal is to compare our Multi-BOWS framework to existing methods.²⁵ We carry out each simulation on a machine with an Intel Xeon Gold 6126 CPU. For modeling and simulating nanophotonic structures, we employ the finite-difference time-domain method through Ansys Lumerical 2022 R2.1 and its Python API.

We execute a low-fidelity multi-objective function $[f_{tr}^{\text{low}}, f_{se}^{\text{low}}]$ and a high-fidelity multi-objective function $[f_{tr}^{\text{high}}, f_{se}^{\text{high}}]$ using uniform mesh sizes of 40 nm and 2 nm, respectively. The meshes are overridden at the silver and



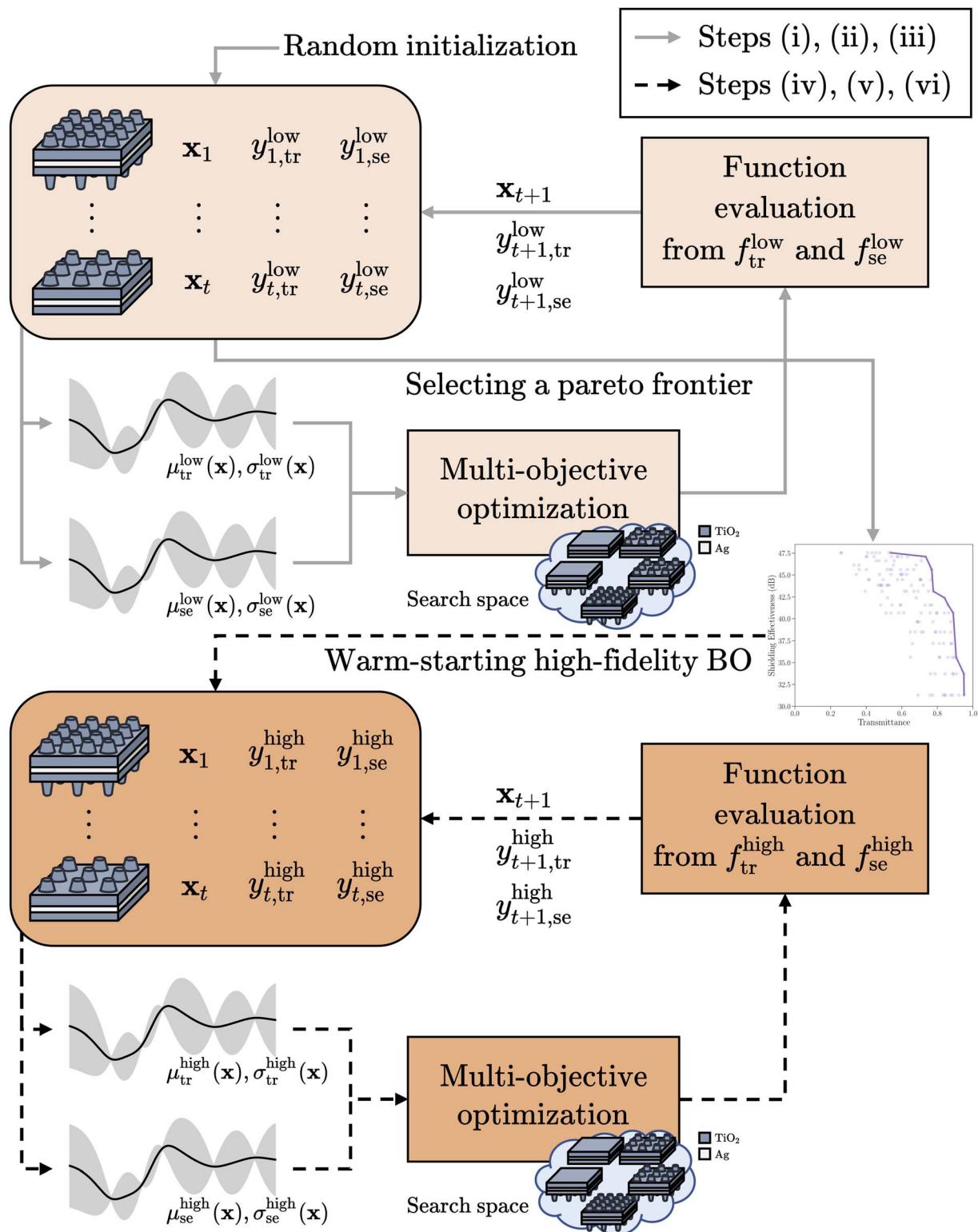


Fig. 3 Multi-BOWS framework. Initially, low-fidelity multi-objective Bayesian optimization is performed with randomly selected initial points (light orange). Following that, high-fidelity multi-objective Bayesian optimization is run, utilizing the Pareto frontier derived from the low-fidelity Bayesian optimization (dark orange) to suggest optimal structure candidates.

titanium oxide interfaces in order to capture the effect of small thickness. These mesh sizes are selected to ensure an appropriate simulation time. As expected, the evaluations of

$[f_{tr}^{\text{high}}, f_{se}^{\text{high}}]$ are slower but more accurate than the ones of $[f_{tr}^{\text{low}}, f_{se}^{\text{low}}]$. Notably, the evaluations of f_{tr}^{low} can be larger than 1, which is physically impossible. Due to the lower accuracy of low-



Table 2 Quantitative results on our simulations. The AUPF of the low-fidelity only method indicates the result obtained by re-evaluating the Pareto frontiers of low-fidelity evaluations using a high-fidelity function. The standard errors of the sample mean are presented

Structure	AUPF			
	Single-fidelity algorithm		Multi-fidelity algorithm	
	Low-fidelity only	High-fidelity only	Multi-fidelity	Multi-BOWS
Three-layer	0.4529 ± 0.0529	0.7939 ± 0.0058	0.7795 ± 0.0090	0.8600 ± 0.0014
Matched-period	0.7231 ± 0.0204	0.8344 ± 0.0049	0.8392 ± 0.0049	0.8579 ± 0.0022
Unmatched-period	0.7088 ± 0.0273	0.7727 ± 0.0072	0.7941 ± 0.0100	0.8141 ± 0.0039
Meta-structure	0.7157 ± 0.0104	0.8286 ± 0.0054	0.8509 ± 0.0035	0.8551 ± 0.0021

fidelity evaluations, we do not report the results of low-fidelity evaluations in this section. Instead, we evaluate the final Pareto frontier acquired by low-fidelity Bayesian optimization using $[f_{tr}^{high}, f_{se}^{high}]$. Moreover, to compare Bayesian optimization algorithms, we normalize the evaluations of f_{se}^{high} with min-max scaling. This way, the AUPF is confined with the range $[0, 1]$. The AUPF is computed as follows:

$$AUPF = \sum_{i=1}^{|P^{high}|} \frac{|P^{high}| \left(y_{i,tr}^{high} - y_{i-1,tr}^{high} \right) (y_{i,se}^{high} - y_{min,se}^{high})}{y_{max,se}^{high} - y_{min,se}^{high}}, \quad (12)$$

where P^{high} is retrieved to satisfy $y_{i-1,tr}^{high} \leq y_{i,tr}^{high}$ for $i \in \{1, \dots, |P^{high}|\}$, $y_{0,tr}^{high} = 0$ is assumed, $y_{min,se}^{high}$ is the minimum of SE, and $y_{max,se}^{high}$ is the maximum of SE. The AUPF is defined within a two-dimensional space, where it serves the same metric as the normalized version of the hypervolume measure. Lastly, to measure f_{tr}^{low} or f_{tr}^{high} , the average transparency of visible incident light with wavelengths between 400 to 700 nm is used.

In our Multi-BOWS approach, we employ Gaussian process regression utilizing the Matérn 5/2 kernel as a surrogate function.⁶⁰ We choose the expected improvement policy as an acquisition function,⁶³ and this function is optimized using multi-started L-BFGS-B by following the work.⁷⁹ For the time budget, we allocate 20% to T^{low} and the remaining 80% to T^{high} . The low-fidelity only or high-fidelity only multi-objective Bayesian optimization initializes with 10 points, while the multi-fidelity multi-objective Bayesian optimization starts with a total of 10 points, out of which 8 are evaluated by a low-fidelity function and the other 2 points are by a high-fidelity function. Moreover, for the low-fidelity Bayesian optimization of Multi-BOWS, we start with 8 initial points. If the size of the Pareto frontier of low-fidelity evaluations exceeds 10, we randomly select 10 points from the Pareto frontier of the low-fidelity evaluations. For the existing methods, we employ the official implementation of the recent work.^{25†}

We investigate four following structures: the three-layer structure, matched-period double-sided nanocone structure, unmatched-period double-sided nanocone structure, and meta-structure. The AUPF is calculated for each structure with four variations: low-fidelity only, high-fidelity only, and multi-fidelity

multi-objective Bayesian optimization, and Multi-BOWS. Using the qualitative results in Table 2, we compare four algorithms by computing X/Y where X and Y are the AUPF results. We obtain those results by assuming the uncorrelated non-central normal ratio for a ratio distribution.

We find that the Multi-BOWS approach discovers superior structures more rapidly compared to other methods and is successful in identifying structures that exhibit higher SE and visible transmittance compared to other methods, as presented in Fig. 4 and Table 2. Our method delivers an AUPF that is $89.9 \pm 22.2\%$ and $8.3 \pm 0.8\%$ larger in the three-layer structure, $18.6 \pm 3.4\%$ and $2.8 \pm 0.7\%$ larger in the matched-period double-sided nanocone structure, $14.9 \pm 4.5\%$ and $5.4 \pm 1.1\%$ larger in the unmatched-period double-sided nanocone structure, and $19.5 \pm 1.8\%$ and $3.2 \pm 0.7\%$ larger in the meta-structure compared to the low-fidelity only and high-fidelity only methods, respectively. Interestingly, the earlier multi-fidelity multi-objective Bayesian optimization technique tends to outperform the low-fidelity only and high-fidelity only methods except for one case between the high-fidelity only and multi-fidelity methods for the three-layer structure. Furthermore, our Multi-BOWS shows $10.3 \pm 1.3\%$, $2.2 \pm 0.7\%$, $2.5 \pm 1.4\%$, and $0.5 \pm 0.5\%$ larger AUPF than the existing multi-fidelity algorithm for four structures, respectively.

It is important to note that the number of initial points is identical across all experiments as previously mentioned. Moreover, the number of evaluations varies significantly across nanophotonic structures because the simulation time is dependent on the size of the simulation cell. For example, the low-fidelity only Bayesian optimization evaluates 462.5000 ± 6.8007 structures for the three-layer structure, 1218.2000 ± 16.1728 structures for the matched-period double-sided nanocone structure, 2648.6000 ± 98.2692 structures for the unmatched-period double-sided nanocone structure, and 2948.0000 ± 16.7571 structures for the meta-structure, and the high-fidelity only Bayesian optimization method evaluates 397.6000 ± 1.7436 structures for the three-layer structure, 230.3000 ± 42.5888 structures for the matched-period double-sided nanocone structure, 43.1000 ± 12.3810 structures for the unmatched-period double-sided nanocone structure, and 205.7000 ± 59.3701 structures for the meta-structure.

Therefore, we can remark two main messages here. Firstly, the use of multi-fidelity evaluations improves Bayesian

† It is available at <https://github.com/belakaria/mf-osemo>.



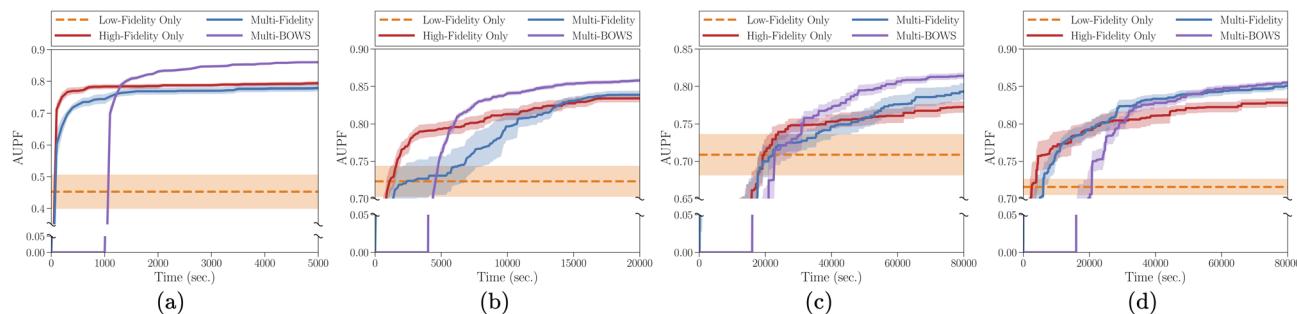


Fig. 4 AUPF versus execution time for electrodynamic simulations, based on 10 repeated experiments. AUPF is all reported based on high-fidelity simulations. The mean (solid) and the standard error (shaded areas) are shown. (a) Three-layer structure. (b) Matched-period double-sided nanocone structure. (c) Unmatched-period double-sided nanocone structure. (d) Meta-structure.

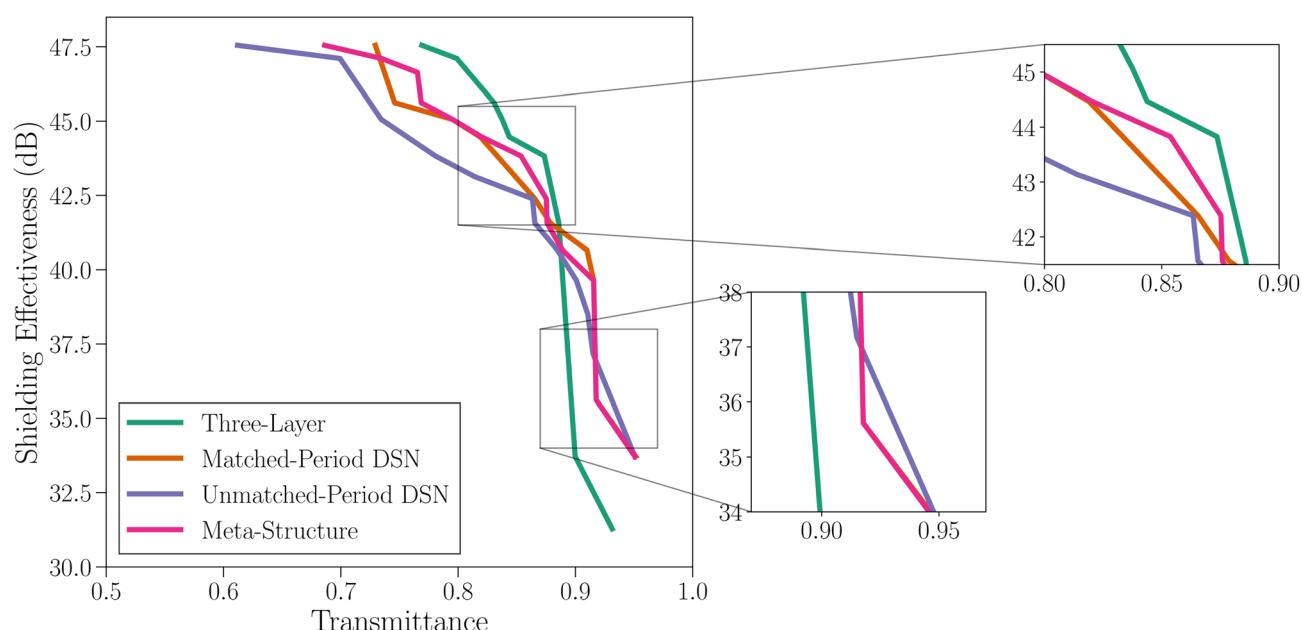


Fig. 5 Plot of the aggregated Pareto frontiers for the structures we study using Multi-BOWS, based on 10 repeated experiments. To alleviate the effects of the number of evaluations, we sample the first 100 evaluations from each simulation, except in the case of the unmatched-period double-sided nanocone structure. Those sampled evaluations of 10 repeated experiments are aggregated in order to show the best structures found. Note that DSN stands for double-sided nanocone.

optimization's ability to find high quality structures. Secondly, while the performance gain in higher-dimensional problems is smaller than in lower-dimensional problems, our Multi-BOWS approach remains effective for all the structures compared to the other methods.

We observe interesting characteristics in the structures identified by our method, as shown in Fig. 5. The structures with nanocones – both matched-period and unmatched-period double-sided nanocone structures – exhibit greater visible transparency than the three-layer structures in the region of high transparency. The lines plotted in the bottom right box of Fig. 5 show better performance in the region of high transparency. In particular, the results for the unmatched-period double-sided nanocone structure are comparable to or better than the results for the matched-period double-sided nanocone structure, even though the number of evaluations for the

unmatched-period structure is less than the number of evaluations for the matched-period structure. Additionally, Fig. 5 shows that the meta-structure favors high-transmission structures in the region of high SE, thus achieving similar performance to the three-layer structure. It implies that the meta-structure allows the optimization algorithm to actively seek diverse structures without thorough domain knowledge in optical device design. By leveraging this feature, we can systematically address the problem of optical device design by defining a more generic search space and employing a Bayesian optimization strategy, such as our Multi-BOWS framework.

Conclusion

In this paper, we have introduced a novel method Multi-BOWS, aimed at addressing challenges in optical device design. This

problem involves optimizing multiple conflicting objectives while taking into account the fidelity of evaluations. To address this, we compared various existing Bayesian optimization methods with Multi-BOWS. Our results demonstrate that Multi-BOWS outperforms the existing baseline methods in terms of the AUPF, yielding 3.2–89.9% larger AUPF than the low-fidelity only and high-fidelity only methods for the nanophotonic structures studied, and demonstrating 0.5–10.3% larger AUPF than the existing multi-fidelity method for the investigated structures. Additionally, we note interesting characteristics of the nanophotonic structures discovered by our method, indicating its potential in uncovering more effective solutions.

Data availability

The code for Multi-BOWS implementation and simulations can be found at <https://github.com/jungtaekkim/Multi-BOWs>.

Author contributions

Conceptualization: JK, ML, YL, and PWL. Formal Analysis: JK, ML, OH, and PWL. Funding acquisition: PWL. Methodology: JK, AG, OH, and PWL. Project administration: PWL. Software: JK. Supervision: OH and PWL. Visualization: JK. Writing – original draft: JK. Writing – review & editing: JK, ML, OH, and PWL.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This research was partly funded by the National Science Foundation (NSF) under grant AM 1930582. The authors acknowledge support from the MDS-Rely Center to conduct this research. The MDS-Rely Center is supported by the NSF's Industry-University Cooperative Research Center (IUCRC) Program under award EEC-2052662 and EEC-2052776. Also, this research was supported in part by the University of Pittsburgh Center for Research Computing through the resources provided. Specifically, this work used the H2P cluster, which is supported by NSF award number OAC-2117681. OH was supported by the NSF and United States-Israel Binational Science Foundation (NSF-BSF) program under grant 2239527 and from the Airforce Office of Scientific Research (AFOSR) grant FA9550-23-1-0242. In addition, we thank Mehdi Zarei for helpful discussions on materials and electromagnetic shielding.

References

- 1 M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Elsevier, 2013.
- 2 T. Gao, E. Stevens, J.-K. Lee and P. W. Leu, *Opt. Lett.*, 2014, **39**, 4647–4650.
- 3 T. Gao, S. Haghafifar, M. G. Lindsay, P. Lu, M. I. Kayes, B. D. Pafchek, Z. Zhou, P. R. Ohodnicki and P. W. Leu, *Adv. Opt. Mater.*, 2018, **6**, 1700829.
- 4 R. Kingslake and R. B. Johnson, *Lens design fundamentals*, Academic Press, 2009.
- 5 S. Haghafifar, M. McCourt, B. Cheng, J. Wuenschell, P. Ohodnicki and P. W. Leu, *Mater. Horiz.*, 2019, **6**, 1632–1642.
- 6 S. Haghafifar, M. McCourt, B. Cheng, J. Wuenschell, P. Ohodnicki and P. W. Leu, *Optica*, 2020, **7**, 784–789.
- 7 E. F. Schubert, *Light-Emitting Diodes*, Cambridge University Press, 2006.
- 8 M. G. Moharam and T. K. Gaylord, *J. Opt. Soc. Am.*, 1981, **71**, 811–818.
- 9 J.-M. Jin, *The finite element method in electromagnetics*, John Wiley & Sons, 2015.
- 10 A. Taflove, *IEEE Trans. Electromagn. Compat.*, 1980, **22**, 191–202.
- 11 H. J. Kushner, *J. Basic Eng.*, 1964, **86**, 97–106.
- 12 J. Mockus, *Optimization Techniques IFIP Technical Conference*, 1975, pp. 400–404.
- 13 D. R. Jones, M. Schonlau and W. J. Welch, *J. Global Optim.*, 1998, **13**, 455–492.
- 14 M. M. R. Elsawy, S. Lanteri, R. Duvigneau, G. Brière, M. S. Mohamed and P. Genevet, *Sci. Rep.*, 2019, **9**, 17918.
- 15 P.-I. Schneider, X. G. Santiago, V. Soltwisch, M. Hammerschmidt, S. Burger and C. Rockstuhl, *ACS Photonics*, 2019, **6**, 2726–2733.
- 16 P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, P. K. Herring, M. Aykol, S. J. Harris, R. D. Braatz, S. Ermon and W. C. Chueh, *Nature*, 2020, **578**, 397–402.
- 17 R.-R. Griffiths and J. M. Hernández-Lobato, *Chem. Sci.*, 2020, **11**, 577–586.
- 18 B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams and A. G. Doyle, *Nature*, 2021, **590**, 89–96.
- 19 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, *Nature*, 2021, **596**, 583–589.
- 20 K. Tunyasuvunakool, J. Adler, Z. Wu, T. Green, M. Zielinski, A. Žídek, A. Bridgland, A. Cowie, C. Meyer, A. Laydon, S. Velankar, G. J. Kleywegt, A. Bateman, R. Evans, A. Pritzel, M. Figurnov, O. Ronneberger, R. Bates, S. A. A. Kohl, A. Potapenko, A. J. Ballard, B. Romera-Paredes, S. Nikolov, R. Jain, E. Clancy, D. Reiman, S. Petersen, A. W. Senior, K. Kavukcuoglu, E. Birney, P. Kohli, J. Jumper and D. Hassabis, *Nature*, 2021, **596**, 590–596.
- 21 M. Popova, O. Isayev and A. Tropsha, *Sci. Adv.*, 2018, **4**, eaap7885.

22 B. Zoph and Q. V. Le, *Proceedings of the International Conference on Learning Representations*, ICLR, 2017.

23 K. Kandasamy, W. Neiswanger, J. Schneider, B. Póczos and E. P. Xing, *Advances in Neural Information Processing Systems*, NeurIPS, 2018, pp. 2016–2025.

24 S. Zhu, I. Ng and Z. Chen, *Proceedings of the International Conference on Learning Representations*, ICLR, 2020.

25 S. Belakaria, A. Deshwal and J. R. Doppa, *Proceedings of the AAAI Conference on Artificial Intelligence*, AAAI, 2020, pp. 10035–10043.

26 S. Geetha, K. K. Satheesh Kumar, C. R. K. Rao, M. Vijayan and D. C. Trivedi, *J. Appl. Polym. Sci.*, 2009, **112**, 2073–2086.

27 M. Li, M. Zarei, A. J. Galante, B. Pilsbury, S. B. Walker, M. LeMieux and P. W. Leu, *Prog. Org. Coat.*, 2023, **179**, 107506.

28 M. Li, S. Sinha, S. Hannani, S. B. Walker, M. LeMieux and P. W. Leu, *ACS Appl. Electron. Mater.*, 2022, **5**, 173–180.

29 H. Wang, C. Ji, C. Zhang, Y. Zhang, Z. Zhang, Z. Lu, J. Tan and L. J. Guo, *ACS Appl. Mater. Interfaces*, 2019, **11**, 11782–11791.

30 A. Iqbal, P. Sambyal and C. M. Koo, *Adv. Funct. Mater.*, 2020, **30**, 2000883.

31 M. Li, M. Zarei, K. Mohammadi, S. B. Walker, M. LeMieux and P. W. Leu, *ACS Appl. Mater. Interfaces*, 2023, **15**, 30591–30599.

32 M. Li, M. J. McCourt, A. J. Galante and P. W. Leu, *Opt. Express*, 2022, **30**, 33182–33194.

33 C. Yuan, J. Huang, Y. Dong, X. Huang, Y. Lu, J. Li, T. Tian, W. Liu and W. Song, *ACS Appl. Mater. Interfaces*, 2020, **12**, 26659–26669.

34 X. Zhao, X. Meng, H. Zou, Z. Wang, Y. Du, Y. Shao, J. Qi and J. Qiu, *Adv. Funct. Mater.*, 2023, **33**, 2209207.

35 S. Yalamanchili, E. Verlage, W.-H. Cheng, K. T. Fountaine, P. R. Jahelka, P. A. Kempler, R. Saive, N. S. Lewis and H. A. Atwater, *Nano Lett.*, 2019, **20**, 502–508.

36 H. Chen, X. Li, Y. Wang, Y. Li, Y. Yu, H. Li and B. Shentu, *ACS Omega*, 2022, **7**, 46769–46776.

37 W. Zhang, J. Zhang, P. Wu, G. Chai, R. Huang, F. Ma, F. Xu, H. Cheng, Y. Chen, X. Ni, L. Qiao and J. Duan, *ACS Appl. Mater. Interfaces*, 2020, **12**, 23340–23346.

38 S. Haghifar, P. Lu, M. I. Kayes, S. Tan, K.-J. Kim, T. Gao, P. Ohodnicki and P. W. Leu, *J. Mater. Chem. C*, 2018, **6**, 9191–9199.

39 M. I. Kayes, M. Zarei, F. Feng and P. W. Leu, *Nanotechnology*, 2023, **35**, 025102.

40 J. Pearl, *Causality*, Cambridge University Press, 2009.

41 J. Peters, D. Janzing and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*, The MIT Press, 2017.

42 J. S. Jensen and O. Sigmund, *Laser Photonics Rev.*, 2011, **5**, 308–321.

43 S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković and A. W. Rodriguez, *Nat. Photonics*, 2018, **12**, 659–670.

44 R. E. Christiansen and O. Sigmund, *J. Opt. Soc. Am. B*, 2021, **38**, 496–509.

45 L. F. Frellsen, Y. Ding, O. Sigmund and L. H. Frandsen, *Opt. Express*, 2016, **24**, 16866–16873.

46 D. Sell, J. Yang, S. Doshay, R. Yang and J. A. Fan, *Nano Lett.*, 2017, **17**, 3752–3757.

47 A. M. Hammond, A. Oskooi, S. G. Johnson and S. E. Ralph, *Opt. Express*, 2021, **29**, 23916–23938.

48 A. M. Hammond, J. B. Slaby, M. J. Probst and S. E. Ralph, *ACS Photonics*, 2022, **10**, 808–814.

49 E. Brochu, V. M. Cora and N. de Freitas, *arXiv*, 2010, preprint, arXiv:1012.2599, pp. 1–49.

50 B. Shahriari, K. Swersky, Z. Wang, R. P. Adams and N. de Freitas, *Proc. IEEE*, 2016, **104**, 148–175.

51 R. Garnett, *Bayesian Optimization*, Cambridge University Press, 2023.

52 M. Feurer, A. Klein, K. Eggensperger, J. T. Springenberg, M. Blum and F. Hutter, *Advances in Neural Information Processing Systems*, NeurIPS, 2015, pp. 2962–2970.

53 F. Hutter, L. Kotthoff and J. Vanschoren, *Automated Machine Learning: Methods, Systems, Challenges*, Springer Nature, 2019.

54 A. Borji and L. Itti, *Advances in Neural Information Processing Systems*, NeurIPS, 2013, pp. 55–63.

55 M. McLeod, M. A. Osborne and S. J. Roberts, *Proceedings of the International Conference on Machine Learning*, ICML, 2018, pp. 3440–3449.

56 R. Turner, D. Eriksson, M. McCourt, J. Kiili, E. Laaksonen, Z. Xu and I. Guyon, *Proceedings of the NeurIPS Competition and Demonstration Track*, 2020, pp. 3–26.

57 D. Eriksson and M. Poloczek, *Proceedings of the International Conference on Artificial Intelligence and Statistics*, AISTATS, 2021, pp. 730–738.

58 R. Baptista and M. Poloczek, *Proceedings of the International Conference on Machine Learning*, ICML, 2018, pp. 462–471.

59 J. Kim and S. Choi, *Proceedings of the International Conference on Artificial Intelligence and Statistics*, AISTATS, 2022, pp. 4359–4375.

60 C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006.

61 F. Hutter, H. H. Hoos and K. Leyton-Brown, *Proceedings of the International Conference on Learning and Intelligent Optimization*, LION, 2011, pp. 507–523.

62 J. T. Springenberg, A. Klein, S. Falkner and F. Hutter, *Advances in Neural Information Processing Systems*, NeurIPS, 2016, pp. 4134–4142.

63 J. Močkus, V. Tesis and A. Žilinskas, *Towards Global Optimization*, 1978, vol. 2, pp. 117–129.

64 N. Srinivas, A. Krause, S. Kakade and M. Seeger, *Proceedings of the International Conference on Machine Learning*, ICML, 2010, pp. 1015–1022.

65 M. Hoffman, E. Brochu and N. de Freitas, *Proceedings of the Annual Conference on Uncertainty in Artificial Intelligence*, UAI, 2011, pp. 327–336.

66 C. Qin, D. Klabjan and D. Russo, *Advances in Neural Information Processing Systems*, NeurIPS, 2017, pp. 5382–5392.

67 K. Kandasamy, G. Dasarathy, J. B. Oliva, J. Schneider and B. Póczos, *Advances in Neural Information Processing Systems*, NeurIPS, 2016, pp. 1000–1008.



68 K. Kandasamy, G. Dasarathy, J. Schneider and B. Póczos, *Proceedings of the International Conference on Machine Learning*, ICML, 2017, pp. 1799–1808.

69 S. Takeno, H. Fukuoka, Y. Tsukada, T. Koyama, M. Shiga, I. Takeuchi and M. Karasuyama, *Proceedings of the International Conference on Machine Learning*, ICML, 2020, pp. 9334–9345.

70 D. Hernández-Lobato, J. M. Hernández-Lobato, A. Shah and R. P. Adams, *Proceedings of the International Conference on Machine Learning*, ICML, 2016, pp. 1492–1501.

71 S. Belakaria, A. Deshwal and J. R. Doppa, *Advances in Neural Information Processing Systems*, NeurIPS, 2019, pp. 7825–7835.

72 S. Daulton, D. Eriksson, M. Balandat and E. Bakshy, *Proceedings of the Annual Conference on Uncertainty in Artificial Intelligence*, UAI, 2022, pp. 507–517.

73 F. Irshad, S. Karsch and A. Döpp, *arXiv*, 2021, preprint, arXiv:2112.13901, pp. 1–17.

74 F. Irshad, S. Karsch and A. Döpp, *Phys. Rev. Res.*, 2023, 5, 013063.

75 M. Poloczek, J. Wang and P. I. Frazier, *Proceedings of the Winter Simulation Conference*, 2016, pp. 770–781.

76 J. Kim, S. Kim and S. Choi, *arXiv*, 2017, preprint, arXiv:1710.06219, pp. 1–14.

77 I. M. Sobol', *USSR Comput. Math. Math. Phys.*, 1967, 7, 784–802.

78 B. Paria, K. Kandasamy and B. Póczos, *Proceedings of the Annual Conference on Uncertainty in Artificial Intelligence*, UAI, 2019, pp. 766–776.

79 J. Kim and S. Choi, *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, ECML-PKDD, 2020, pp. 675–690.

