



## Automated design of nonreciprocal thermal emitters via Bayesian optimization

Bach Do <sup>a,1</sup>, Sina Jafari Ghalekohneh <sup>b,1</sup>, Taiwo Adebiyi <sup>a</sup>, Bo Zhao <sup>b,\*</sup>, Ruda Zhang <sup>a,\*</sup>

<sup>a</sup> Department of Civil and Environmental Engineering, University of Houston, Houston, TX, 77204, USA

<sup>b</sup> Department of Mechanical and Aerospace Engineering, University of Houston, Houston, TX, 77204, USA

### ARTICLE INFO

#### Keywords:

Nonreciprocity  
Thermal emitters  
Gaussian process  
Bayesian optimization  
Reparameterization

### ABSTRACT

Nonreciprocal thermal emitters that break Kirchhoff's law of thermal radiation promise exciting applications for thermal and energy applications. The design of the bandwidth and angular range of the nonreciprocal effect, which directly affects the performance of nonreciprocal emitters, typically relies on physical intuition. In this study, we present a general numerical approach to maximize the nonreciprocal effect. We choose doped magneto-optic materials and magnetic Weyl semimetal materials as model materials and focus on pattern-free multilayer structures. The optimization randomly starts from a less effective structure and incrementally improves the broadband nonreciprocity through the combination of Bayesian optimization and reparameterization. Optimization results show that the proposed approach can discover structures that can achieve broadband nonreciprocal emission at wavelengths from 5 to 40  $\mu\text{m}$  using only a few layers, significantly outperforming current state-of-the-art designs based on intuition in terms of both performance and simplicity.

### 1. Introduction

Over the past two decades, the control of thermal radiation transfer has gained significant interest due to its vital applications in spacecraft, manufacturing, thermal management, and energy conversion [1–4]. Typically, the design and simulation of these systems assume reciprocity. According to Kirchhoff's law [5,6], for reciprocal thermal emitters, emissivity ( $\epsilon$ ) and absorptivity ( $\alpha$ ) are equal for a given direction, frequency, and polarization. This reciprocal relationship between emission and absorption represents a significant constraint on controlling the radiative heat flow [7,8], and prevents radiative energy harvesting technologies from reaching their thermodynamic limits [9–14].

Recent research advances [2,15–17] have suggested that reciprocity is not a requirement of thermodynamics, and it breaks down for nonreciprocal thermal emitters which allow  $\epsilon$  and  $\alpha$  to be separately controlled. Nonreciprocal thermal emission provides the pathway to record-breaking high-efficiency radiative energy harvesting techniques [9,13,14,18–20], thermal regulation systems [21–24], and mechanical propulsion with radiative heat flow [25,26]. Importantly, to achieve the ultimate performance in these applications, especially for far-field scenarios, nonreciprocal thermal radiative properties should be achieved over a broad wavelength and angular range. Designing nonreciprocal

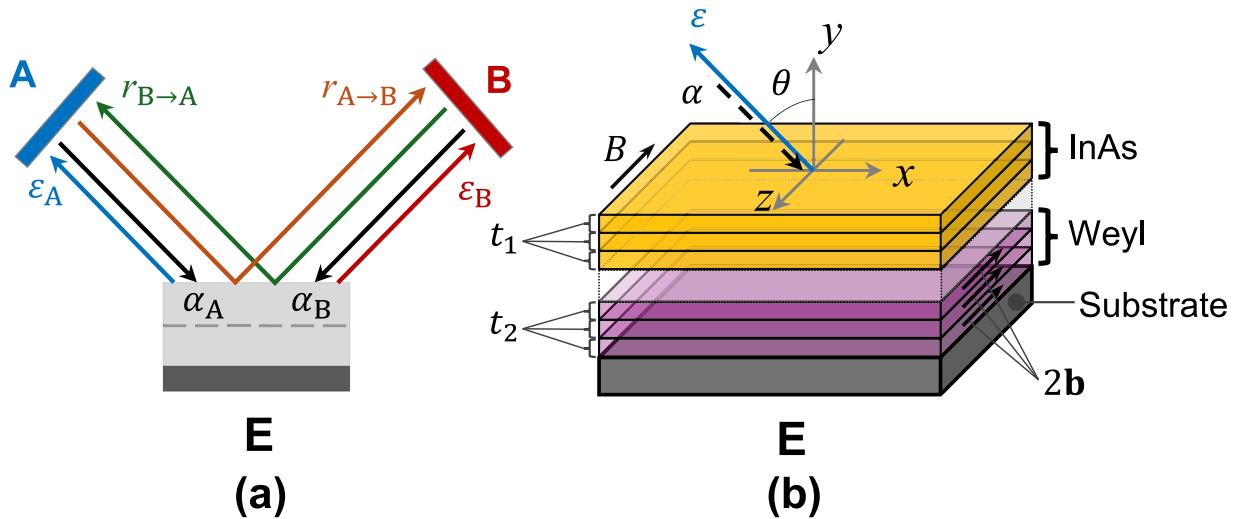
structures with desired properties is a new challenge compared to designing reciprocal emitters [27,28], since one needs to codesign emissivity and absorptivity simultaneously.

Early work has demonstrated that enhanced nonreciprocal properties [16,29] could be obtained in semiconductors and semimetals at the excitation of resonances in the wavelength range where the dielectric function crosses zero (i.e., epsilon-near-zero or ENZ range). Based on this principle, current approaches to achieving broadband nonreciprocity rely on multilayered structures in which each layer exhibits strong nonreciprocity at a different ENZ wavelength. Combined, the structure can trigger a number of these resonances simultaneously and yield broadband nonreciprocity. For example, Liu et al. [22] demonstrated broadband nonreciprocal emission at wavelengths from 20 to 40  $\mu\text{m}$  using a gradient-doped multilayer InAs structure, where each layer shows strong nonreciprocity at a specific resonance wavelength under an external magnetic field. Similarly, Zhang and Zhu [30] proposed broadband nonreciprocity at wavelengths from 10 to 20  $\mu\text{m}$  using a multilayer magnetic Weyl semimetal structure with a gradient chemical potential. However, since the material and the associated parameters, like carrier concentration and geometric parameters, can all be free to choose, designing high-performance structures of this kind is quite challenging and typically relies on physical intuition. As a result, the

\* Corresponding authors.

E-mail addresses: [bzhao8@uh.edu](mailto:bzhao8@uh.edu) (B. Zhao), [rudaz@uh.edu](mailto:rudaz@uh.edu) (R. Zhang).

<sup>1</sup> Equal contribution.



**Fig. 1.** (a) Illustration of radiative heat exchange between two blackbodies (A and B) and an emitter (E) under thermal equilibrium. (b) Illustration of a multilayer structure consisting of InAs layers on top of Weyl semimetal layers on a reflective substrate.

final designs are oftentimes suboptimal; for example, requiring a high number of layers that are challenging to fabricate. Additionally, current studies usually focus on optimizing the absolute value of the contrast between emissivity and absorptivity [31]. We make sure the sign of the contrast is respected in the design algorithm since it is crucial to the direction of photon transport [32] and can be important in various applications where the net photon flow direction is significant [13,33]. Despite the effectiveness of multilayer structures in achieving broadband nonreciprocal thermal emission, an effective approach to optimize the performance is lacking.

Here we propose a numerical method based on Bayesian optimization (BO) [34–39] and reparameterization to optimize thermal emitters with a large number of design parameters. BO is used because it can address the high computational demand presented by repeated calculations of absorptivity and emissivity during the optimization process, which hinders the use of any population-based optimizers (e.g., genetic algorithm or particle swarm optimization) [27,40], gradient-based optimizers (e.g., gradient descent) [41], or deep learning approaches that require a large dataset [28,42]. The reparameterization imposes user-defined constraints on the structures' profiles and further reduces the number of optimization parameters, thereby reinforcing the efficiency of BO which is well-suited for small to medium optimization problems [38]. Our optimization approach is applicable to both reciprocal and nonreciprocal emitters. In this work, we focus on demonstrating its application in achieving enhanced broadband nonreciprocal radiative properties with the sign of the contrast being maintained. We aim to find structures with a much fewer number of layers and yet can achieve nonreciprocity in an even broader wavelength range than current designs.

## 2. Absorptivity and emissivity of nonreciprocal emitters

For opaque structures, the emissivity and absorptivity can be computed from the reflectivity, as discussed in previous studies [43,44]. Here, we briefly review the rationale behind the approach to facilitate our later discussion. Consider a system in thermal equilibrium that consists of an opaque thermal emitter (E) exchanging energy with two blackbodies, A and B, (Fig. 1(a)). The emission from blackbody A is either absorbed by the thermal emitter E ( $\alpha_A$ ), or reflected and absorbed by blackbody B ( $r_{A \rightarrow B}$ ), leading to the following relationship:

$$\alpha_A + r_{A \rightarrow B} = 1, \quad (1)$$

where  $\alpha_A$  represents the absorptivity in direction A at a specific wavelength, and  $r_{A \rightarrow B}$  is the reflectivity from A to B at that wavelength.

Moreover, each blackbody absorbs and emits an equal amount of energy at thermal equilibrium, resulting in:

$$\epsilon_A + r_{B \rightarrow A} = 1. \quad (2)$$

Combining Eqs. (1) and (2), we have:

$$\alpha_A - \epsilon_A = r_{B \rightarrow A} - r_{A \rightarrow B}. \quad (3)$$

For reciprocal emitters with  $r_{A \rightarrow B} = r_{B \rightarrow A}$  [45], it follows that  $\alpha_A = \epsilon_A$ , which is consistent with Kirchhoff's law of radiation [46]. However, for nonreciprocal emitters, we have  $r_{A \rightarrow B} \neq r_{B \rightarrow A}$  [43] and thus  $\alpha_A \neq \epsilon_A$ , which violates Kirchhoff's law of radiation. The absorptivity and emissivity are given as:

$$\alpha_A = 1 - r_{A \rightarrow B}, \quad \epsilon_A = 1 - r_{B \rightarrow A}. \quad (4)$$

For linear and static thermal emitters we consider here [47,48], one needs to start with a material that breaks the reciprocity to obtain nonreciprocal thermal radiation. Magneto-optical material InAs [22,29,43,49] and magnetic Weyl semimetals [16,50–52] have been used for this purpose. As shown from their dielectric functions discussed later, the former works more effectively in the longer wavelength in the mid-infrared range, whereas the latter can provide nonreciprocity for shorter wavelengths. We note that, though each material has been used individually for nonreciprocal emitter design, it is not clear whether and how these materials can be combined to achieve even broader band of nonreciprocal thermal emission. Therefore, we choose these materials as our model nonreciprocal materials and focus on structures shown in Fig. 1(b) that have InAs layers on top of Weyl semimetal layers on a reflective substrate, where we use silver as substrate in this study. The Weyl semimetals provide an ENZ effect in the shorter wavelength range. They have strong metallic behavior in the longer wavelength range that could result in strong reflection and thus are not suitable for top layers. With this structure, we rule out the type of structures that possess alternating InAs and Weyl semimetal layers since they can be very challenging to fabricate in practice. Regardless of the challenges in fabrication, alternating layers may bring additional benefits. For example, one may form a photonic crystal structure by alternating InAs and Weyl semimetal layers that supports other nonreciprocal modes for engineering radiative properties. We will also seek to optimize the structures containing InAs only, so that we can compare with the state-of-the-art design [22] and showcase the significant improvement through using our approach.

We consider transverse magnetic (TM) waves with a magnetic field along the  $-z$  direction, and consider the Voigt configurations [29]. In

this way, we trigger the nonreciprocity for the radiative properties at different polar angle  $\theta$  in the  $x$ - $y$  plane. For InAs, we apply an external magnetic field ( $B$  field) along the  $-z$  direction [53,54]. In this case, the permittivity tensor of each InAs layer reads [15]:

$$\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} & 0 \\ \epsilon_{yx} & \epsilon_{yy} & 0 \\ 0 & 0 & \epsilon_{zz} \end{bmatrix}. \quad (5)$$

The components of the permittivity tensor for InAs can be expressed as:

$$\begin{aligned} \epsilon_{xx} = \epsilon_{yy} &= \epsilon_{\infty} - \frac{\omega_p^2(\omega + i\Gamma)}{\omega[(\omega + i\Gamma)^2 - \omega_c^2]}, \\ \epsilon_{xy} = -\epsilon_{yx} &= \frac{i\omega_p^2\omega_c}{\omega[(\omega + i\Gamma)^2 - \omega_c^2]}, \\ \epsilon_{zz} &= \epsilon_{\infty} - \frac{\omega_p^2}{\omega(\omega + i\Gamma)}, \end{aligned} \quad (6)$$

where  $\omega$  is the angular frequency,  $\epsilon_{\infty}$  is the high-frequency permittivity,  $\Gamma$  is the damping rate,  $\omega_p = \sqrt{n_e e^2/(m^* \epsilon_0)}$  is the plasma frequency, and  $\omega_c = eB/m^*$  is the cyclotron frequency. Here,  $n_e$  is the carrier concentration,  $e$  is the elementary charge,  $m^*$  is the effective electron mass, and  $\epsilon_0$  is the vacuum permittivity. We focus on the ENZ point where the real part of  $\epsilon_{xx}$  crosses zero [15,55–58] and brings a significant enhancement to  $\epsilon_{xy}/\text{Re}(\epsilon_{xx})$  and the nonreciprocal effect [29]. By adjusting the carrier concentration  $n_e$ , which we can experimentally control while fabricating InAs among other parameters in Eq. (6), the ENZ point shifts, allowing us to control the wavelength at which the ENZ region occurs.

The nonreciprocal effect in magnetic Weyl semimetal layers is intrinsic and does not require an external magnetic field [16]. The momentum separation of the Weyl cones,  $2b$ , acts similarly to an applied magnetic field in magneto-optical systems. We set  $\mathbf{b} = -b\mathbf{k}_z$  to be also along the  $-z$  direction (Fig. 1(b)) following the Voigt configuration and, in this case, the permittivity tensor of Weyl semimetal has the same format as Eq. (5) [16]. Similar to InAs, the nonreciprocal effect for Weyl is significantly enhanced when the diagonal element  $\epsilon$  crosses zero at ENZ point.

The diagonal elements of the permittivity of each Weyl semimetal layer read  $\epsilon_{xx} = \epsilon_{yy} = \epsilon_{zz} = \epsilon_b + \frac{i\sigma}{\omega\epsilon_0}$ , where  $\epsilon_b$  is the background permittivity,  $\omega$  is the radiation frequency, and  $\sigma$  is the bulk conductivity obtained from [16]:

$$\begin{aligned} \sigma &= \frac{\epsilon_0 r_s g E_F \Omega G}{6\hbar} \left( \frac{E_F \Omega}{2} \right) + i \frac{\epsilon_0 r_s g E_F}{6\pi\hbar} \left\{ \frac{4}{\Omega} \left[ 1 + \frac{\pi^2}{3} \left( \frac{k_B T}{E_F} \right)^2 \right] \right. \\ &\quad \left. + 8\Omega \int_0^{\xi_c} \frac{G(E_F \xi) - G\left(\frac{E_F \Omega}{2}\right)}{\Omega^2 - 4\xi^2} \xi d\xi \right\}. \end{aligned} \quad (7)$$

Here,  $r_s = e^2/(4\pi\epsilon_0\hbar v_F)$  is the effective fine structure constant,  $\hbar$  is the reduced Planck constant, and  $v_F$  is the Fermi velocity.  $\Omega = \hbar(\omega + i\tau^{-1})/E_F$  is the normalized complex frequency,  $\tau^{-1}$  is the scattering rate,  $E_F$  is the Fermi level, and  $T$  is the temperature, which we assume is 300 K for this study.  $G(E) = n(-E) - n(E)$ , where  $n(E)$  is the Fermi-Dirac distribution function,  $g$  is the number of Weyl points, and  $k_B$  is the Boltzmann constant.  $\xi_c = E_C/E_F$  is the normalized cutoff energy, where  $E_C$  is the cutoff energy beyond which the band dispersion is no longer linear. In addition, the off-diagonal element of the permittivity tensor of each Weyl semimetal layer is calculated as:

$$\epsilon_{xy} = -\epsilon_{yx} = i \frac{be^2}{2\pi^2\hbar\omega}. \quad (8)$$

By tuning the Fermi level  $E_F$  — a parameter similar to  $n_e$  for InAs that can be adjusted in experiments — the ENZ point of the Weyl semimetal can be modified, which shifts the wavelength at which the associated resonance occurs.

### 3. Optimization problem

For the structure shown in Fig. 1(b), each layer exhibits a distinct ENZ point, resulting in enhanced nonreciprocity across different wavelengths. Due to the multilayer design, with each layer having a unique ENZ point, broadband nonreciprocity can be achieved over a wide wavelength range. To optimize the contrast between absorptivity and emissivity across this broad bandwidth, it is necessary to solve an optimization problem to determine the optimal values of design parameters: the carrier concentrations for InAs layers, the Fermi levels for Weyl semimetal layers, and the thickness of each layer in the multilayer structure. In this study, we only try to optimize the mentioned design parameters, while assuming other parameters of the dielectric functions of InAs and Weyl semimetals are fixed. For the InAs, the high-frequency permittivity and the effective electron mass of InAs are  $\epsilon_{\infty} = 12.37$  and  $m^* = 0.033 m_e$  [59], respectively, where  $m_e$  is the electron mass. The damping rate of InAs is set as  $\Gamma = 5.9 \times 10^{12}$  rad/s [22]. The background permittivity, the number of Weyl points, the Fermi velocity, the relaxation time, and the cutoff energy of Weyl semimetals are  $\epsilon_b = 6.2$ ,  $g = 2$ ,  $\tau = 10^{-12}$  s,  $v_F = 0.83 \times 10^5$  m/s, and  $E_C = 0.45$  eV, respectively [16].

Our optimization starts by defining an objective function. Denote wavelength  $\lambda \in \mathbb{R}_+$ , angle of incidence  $\theta \in \mathbb{R}_+$ , and emitter design parameters  $\mathbf{p} \in \mathbb{R}_+^p$ , where  $\mathbb{R}_+ = (0, \infty)$  and  $p$  is the number of design parameters. Let  $\alpha(\mathbf{p}, \lambda, \theta)$  and  $\epsilon(\mathbf{p}, \lambda, \theta)$  be the local functions for the absorptivity and emissivity of this emitter, respectively. The local function for the contrast between the absorptivity and emissivity is defined as:

$$\eta(\mathbf{p}, \lambda, \theta) = \alpha(\mathbf{p}, \lambda, \theta) - \epsilon(\mathbf{p}, \lambda, \theta), \quad (9)$$

where  $\eta \in [-1, 1]$ . Achieving maximum and minimum contrast is equivalent, as the sign of the contrast can be flipped by reversing the magnetic field. Fig. 2 shows the design parameters  $\mathbf{p}$  for a nonreciprocal thermal emitter that is the combination of  $n_1$  doped InAs layers on top of  $n_2$  Weyl semimetal layers. For this structure,  $\mathbf{p}$  has  $n_1$  distinct carrier concentrations  $n_{e,i}$  ( $i = 1, \dots, n_1$ ) for the InAs layers, the thickness  $t_1$  of each InAs layer,  $n_2$  distinct Fermi levels  $E_{F,j}$  ( $j = 1, \dots, n_2$ ) for the Weyl semimetal layers, and the thickness  $t_2$  of each Weyl semimetal layer. Here we simplify the optimization problem by using the same thickness for layers of each type of material.

Maintaining the sign of the contrast, we formulate the following optimization problem for the nonreciprocal emitter, aimed at finding an optimal structure that achieves negative contrast over broad wavelength and angular ranges:

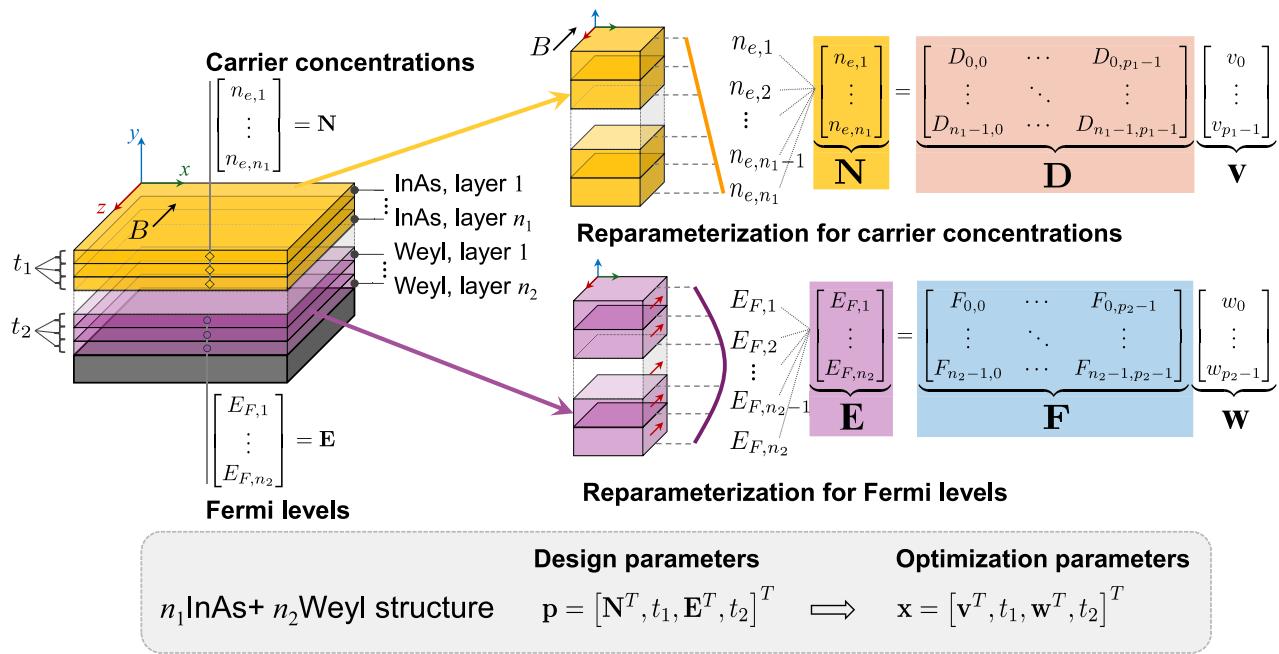
$$\underset{\mathbf{p} \in \mathcal{P}}{\text{minimize}} \quad \tilde{\eta}(\mathbf{p}), \quad (10)$$

where  $\tilde{\eta} \in [-1, 1]$  represents the normalized contrast, which serves as the figure of merit for the broadband nonreciprocity, and  $\mathcal{P} \subseteq \mathbb{R}_+^p$  is the feasible domain of  $\mathbf{p}$ . The normalized contrast  $\tilde{\eta}$  represents the contrast averaged across the considered ranges of wavelengths and angles, which is defined as:

$$\tilde{\eta}(\mathbf{p}) = \frac{\int_{\lambda_L}^{\lambda_U} \int_0^{\pi/2} \eta(\mathbf{p}, \lambda, \theta) \sin \theta \cos \theta d\theta d\lambda}{\int_{\lambda_L}^{\lambda_U} \int_0^{\pi/2} \sin \theta \cos \theta d\theta d\lambda}, \quad (11)$$

where  $\lambda_L$  and  $\lambda_U$  are the lower and upper values of the considered wavelength range, respectively. A numerical quadrature method for efficiently computing the integrals in the numerator of  $\tilde{\eta}(\mathbf{p})$  is given in the Supplementary Material.

We further apply a reparameterization strategy to the objective function  $\tilde{\eta}(\mathbf{p})$  to enforce desired design constraints on the emitter structure and/or other constraints on the optimization problem. The design constraints allow for experimental control over the profiles (e.g., linear or nonlinear) of carrier concentrations and Fermi levels inside the multilayer structures which profoundly affect the contrast value [60]. Meanwhile, the reparameterization ensures that the number



**Fig. 2.** Design parameters  $\mathbf{p}$  for a nonreciprocal emitter of interest and their reparameterization into  $\mathbf{x}$ . The structure consists of  $n_1$  InAs layers with carrier concentrations  $\mathbf{N}$  and thickness  $t_1$ , and  $n_2$  Weyl semimetal layers with Fermi levels  $\mathbf{E}$  and thickness  $t_2$ . After reparameterizing  $\mathbf{N}$  and  $\mathbf{E}$ , the number of parameters reduces from  $n_1 + n_2 + 2$  to  $p_1 + p_2 + 2$ , where  $p_1 \leq n_1$  and  $p_2 \leq n_2$  are the numbers of columns of  $\mathbf{D}$  and  $\mathbf{F}$ , respectively.

of optimization parameters remains manageable for the optimizer, which is crucial when there are many design parameters.

The reparameterization strategy transforms  $\tilde{\eta}(\mathbf{p})$  over the design parameter space  $\mathbf{p} \in \mathcal{P}$  into  $\tilde{\eta}(\mathbf{x})$  over the space of optimization parameters  $\mathbf{x} = \phi(\mathbf{p}) \in \mathbb{R}^d$ , with the reparameterization map  $\phi(\cdot) : \mathbb{R}_+^p \mapsto \mathbb{R}^d$  encapsulating all desired constraints. As a result, the optimization problem after reparameterization reads:

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} \quad \tilde{\eta}(\mathbf{x}), \quad (12)$$

where  $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^d : \phi^{-1}(\mathbf{x}) \in \mathcal{P}\}$  and  $\tilde{\eta}(\mathbf{x}) = \tilde{\eta}(\phi^{-1}(\mathbf{x}))$ . Thus, instead of solving problem (10) directly, we solve problem (12) and recover an optimal set of  $\mathbf{p}$  from the resulting optimal set of  $\mathbf{x}$ .

Given a specific value of the optimization parameters  $\mathbf{x}$  and the reparameterization map  $\phi(\cdot)$ , computing  $\tilde{\eta}(\mathbf{x})$  is straightforward. We first recover the design parameter  $\mathbf{p} = \phi^{-1}(\mathbf{x})$ . We then compute  $\tilde{\eta}(\mathbf{p})$  and set  $\tilde{\eta}(\mathbf{x})$  as  $\tilde{\eta}(\mathbf{p})$ .

**Fig. 2** details the reparameterization technique used in this study. This technique transforms the carrier concentrations  $\mathbf{N} = [n_{e,1}, \dots, n_{e,n1}]^T \in \mathbb{R}^{n1}$  for the InAs layers and Fermi levels  $\mathbf{E} = [E_{F,1}, \dots, E_{F,n2}]^T \in \mathbb{R}^{n2}$  for the Weyl semimetal layers into optimization parameters  $\mathbf{v} \in \mathbb{R}^{p1}$  and  $\mathbf{w} \in \mathbb{R}^{p2}$ , respectively, using two linear maps  $\mathbf{D} \in \mathbb{R}^{n1 \times p1}$  and  $\mathbf{F} \in \mathbb{R}^{n2 \times p2}$  where  $p_1 \leq n_1$  and  $p_2 \leq n_2$ . By carefully designing the column spaces of  $\mathbf{D}$  and  $\mathbf{F}$ , we can reduce the number of parameters for our optimization problem, which is now formulated for  $\mathbf{v}$  and  $\mathbf{w}$ . This reduction is crucial when  $n_1$  or  $n_2$  is large, as the performance of the optimization algorithm strongly depends on the problem's input dimension. Bayesian optimization, for example, is most effective for problems of less than 20 dimensions [38]. By further designing the entries of each column of  $\mathbf{D}$  or  $\mathbf{F}$ , we are able to impose specific geometrical properties on the emitter, for example, linear profiles of carrier concentrations or quadratic profiles of Fermi levels (see Fig. 2). The detailed computations for  $\mathbf{D}$  and  $\mathbf{F}$  are provided in the Supplementary Material.

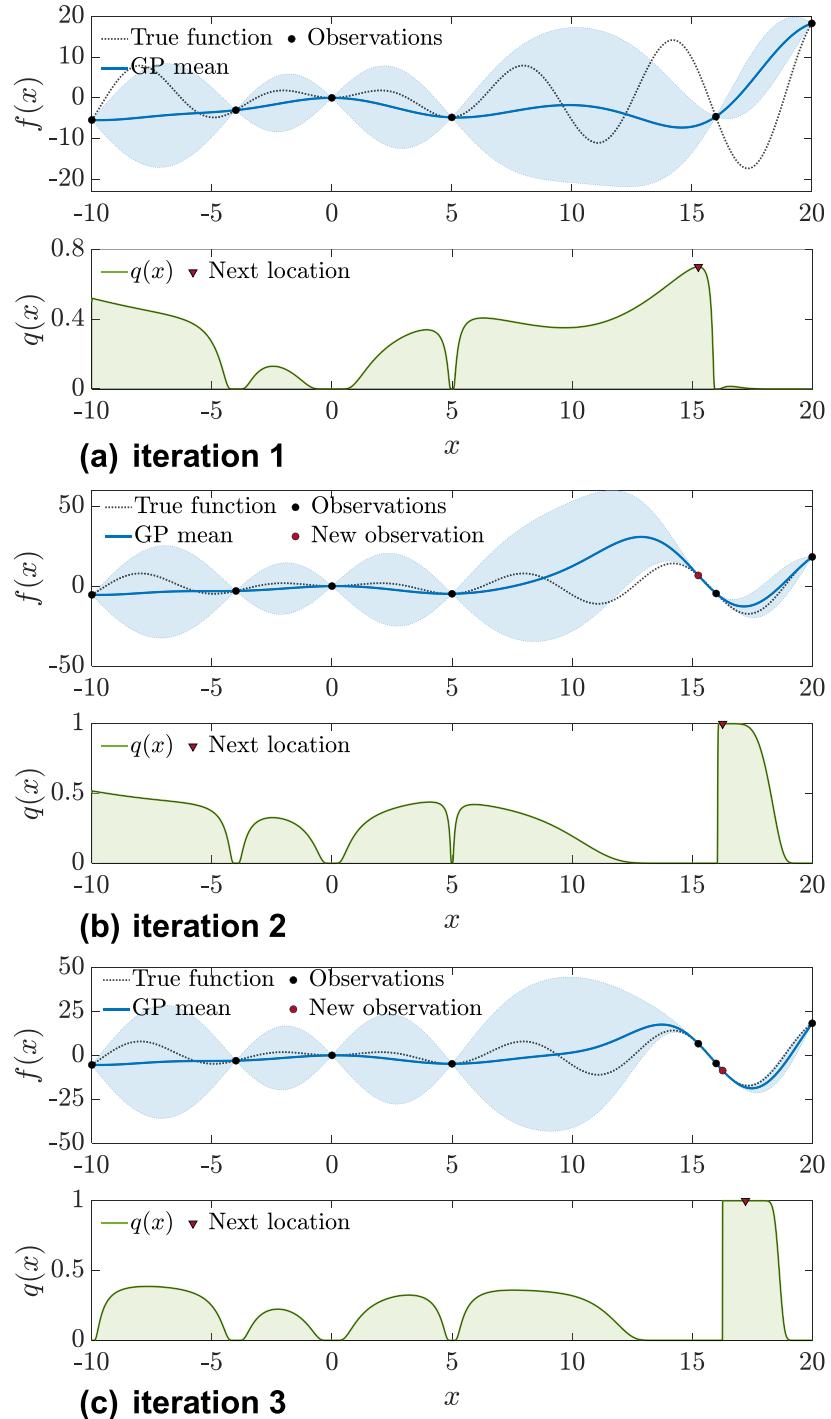
#### 4. Optimization algorithm

Bayesian optimization (BO) [34–39] is a successful approach to optimizing functions that are costly to evaluate, cannot be evaluated exactly, have no analytical expressions, or offer no efficient way to compute their derivatives [39]. It finds applications in diverse domains of science and engineering, including photonic curing processes [61], adaptive experimental design [62,63], accelerator physics [64], and material design [65–67], to name a few. BO also enables the integration of physical and mathematical insights into the optimization process [68–70].

In this study, we use BO to solve problem (12). This is justified because the objective function  $\tilde{\eta}(\mathbf{x})$  lacks a useful analytical expression and is computationally expensive. In the following, we provide an overview of BO and explain how it can be used to incrementally improve the broadband nonreciprocity of nonreciprocal emitters starting from ineffective ones.

Given a few observations for a specific optimization problem, BO solves the problem by repeatedly: (1) constructing a probabilistic model for the objective function to represent our prior knowledge and the observations; (2) formulating an acquisition function to define what we value in the current dataset given the probabilistic model; and (3) maximizing the acquisition function to guide the optimization process. This iterative process typically terminates as it reaches a prespecified number of observations, which reflects our computational budget. **Fig. 3** illustrates three consecutive iterations of BO for minimizing a univariate objective function  $f(x)$ . BO starts with six observations of  $f(x)$  and locates the true global minimizer of  $f(x)$  after only three iterations (bottom panel of **Fig. 3(c)**).

A Gaussian process (GP) [71] model often serves as the probabilistic model for BO due to its flexibility and tractability. Typically, the GP prior is specified simply by a zero mean function and a covariance function with a closed form and a few hyperparameters. Given a dataset of the objective function observations, the posterior can be derived by conditioning the prior on the dataset. This GP posterior then serves as the probabilistic model that represents our beliefs about the structure of the costly objective function. The reader is referred to the



**Fig. 3.** Schematic illustration of three consecutive iterations of BO for optimizing a univariate function  $f(x)$ . In each iteration, BO constructs a probabilistic model for the objective function  $f(x)$  using the current data, formulates an acquisition function  $q(x)$  from the model, and maximizes  $q(x)$  to identify a new location to query the objective function.

Supplementary Material for the mathematical foundation of GP and the analytical formula of the posterior.

Once the posterior model has been constructed, we formulate an acquisition function to define an optimization policy representing what we wish to exploit from this model. The acquisition function maps each point in the design variable space to a score on its ability to benefit the optimization process, considering our imperfect knowledge of the objective function [34,72]. By maximizing this acquisition function, we thus maximize the potential to obtain a better new design. We do so numerically by using, for example, a global optimization algorithm, and the resulting solution is used in the next iteration of BO for a

new evaluation of the objective function. Maximizing the acquisition function is typically simpler and less computationally expensive than optimizing the costly objective function directly.

In Algorithm 1, we outline the use of BO for optimizing a non-reciprocal thermal emitter of interest. We begin with specifying the design parameter domain  $\mathcal{P}$ , the number  $N$  of initial observations of the normalized contrast  $\bar{\eta}(x)$ , and the threshold  $K$  for the number of BO iterations, followed by defining the reparameterization map  $\phi(\cdot)$  (Line 1). To construct an initial dataset  $\mathcal{D}_0$  (Line 6), which is necessary for initializing BO, we randomly generate a set of  $N$  samples for the optimization parameters  $\mathbf{x}$  using Latin hypercube sampling [73] (Line

**Algorithm 1** Bayesian optimization with noiseless observations

---

```

1: input: domain  $\mathcal{P}$  of design parameters  $\mathbf{p}$ , number of initial ob-
   servations  $N$ , threshold for the number of BO iterations  $K$ ,
   reparameterization map  $\phi(\cdot)$ 
2: Generate  $N$  initial observations of optimization parameters  $\mathbf{x}$ 
3: for  $i = 1 : N$  do  $\triangleright$  Generate initial observations of the normalized
   contrast
4:    $y_i \leftarrow \bar{\eta}(\mathbf{x}_i)$ 
5: end for
6:  $\mathcal{D}_0 \leftarrow \{\mathbf{x}_i, y_i\}_{i=1}^N$   $\triangleright$  Dataset of initial observations
7:  $\{\mathbf{x}_{\min}, y_{\min}\} \leftarrow \min\{y_i, i = 1, \dots, N\}$   $\triangleright$  The best observation found
   so far
8: for  $i = N + 1 : N + K$  do
9:    $k \leftarrow i - N$ 
10:  Build a GP posterior  $\hat{\eta}_k(\mathbf{x})|\mathcal{D}_{k-1}$ 
11:  Formulate an acquisition function  $q(\mathbf{x}|\mathcal{D}_{k-1})$  from  $\hat{\eta}_k(\mathbf{x})|\mathcal{D}_{k-1}$ 
12:   $\mathbf{x}_k \leftarrow \arg \max_{\mathbf{x}} q(\mathbf{x}|\mathcal{D}_{k-1})$  s.t.  $\phi^{-1}(\mathbf{x}) \in \mathcal{P}$ ;  $\mathbf{x} \notin \mathcal{D}_{k-1}$   $\triangleright$  Maximize
   the acquisition function
13:   $y_k \leftarrow \bar{\eta}(\mathbf{x}_k)$   $\triangleright$  Obtain a new observation of the normalized
   contrast
14:   $\mathcal{D}_k \leftarrow \mathcal{D}_{k-1} \cup \{\mathbf{x}_k, y_k\}$   $\triangleright$  Update the dataset of observations
15:   $\{\mathbf{x}_{\min}, y_{\min}\} \leftarrow \min\{y_{\min}, y_k\}$   $\triangleright$  The best observation found so far
16: end for
17: return  $\{\mathbf{x}_{\min}, y_{\min}\}$  and  $\mathbf{p}_{\min} \leftarrow \phi(\mathbf{x}_{\min})$ 

```

---

2), and subsequently evaluate  $\bar{\eta}(\mathbf{x})$  at the generated samples using rigorous coupled-wave analyses [74,75] (Line 4). Here, we assume that the evaluated values of  $\bar{\eta}(\mathbf{x})$  are noiseless, which means that the numerical results perfectly capture the distribution of nonreciprocal thermal radiation in the emitter without introducing any numerical errors.

The for-loop of BO (Line 8 to Line 16) commences with constructing a GP posterior  $\hat{\eta}_k(\mathbf{x})$  for  $\bar{\eta}(\mathbf{x})$  from the current dataset  $\mathcal{D}_{k-1}$  (Line 10), where  $k = 1, \dots, K$  represents the index variable for the loop. It then formulates an acquisition function  $q(\mathbf{x})$  based on  $\hat{\eta}_k(\mathbf{x})$  (Line 11) and maximizes  $q(\mathbf{x})$  for a new observation  $\mathbf{x}_k$  of the optimization parameters (Line 12). To avoid reselecting points already seen in the current dataset  $\mathcal{D}_{k-1}$ , the constraint  $\mathbf{x} \notin \mathcal{D}_{k-1}$  is imposed to the maximization of  $q(\mathbf{x})$  (Line 12). Finally, the algorithm evaluates the normalized contrast  $y_k = \bar{\eta}(\mathbf{x}_k)$  at the new point  $\mathbf{x}_k$ , and updates the dataset with  $\mathbf{x}_k$  and  $y_k$  for use in the next iteration (Lines 13 and 14). The final optimal solution is the best observation found among points of the dataset recommended by the algorithm (Lines 15 and 17).

## 5. Results and discussion

We present the optimization results for two multilayer structures of nonreciprocal thermal emitters: a 3-layer InAs structure and a 6-layer 3InAs+3Weyl structure (i.e., three layers of InAs on top of three layers of Weyl semimetal). By considering these structures, we aim to investigate how a combination of InAs and Weyl semimetal can improve broadband nonreciprocity. For this purpose, we also optimize other two structures, namely a 5-layer InAs structure and an 8-layer 5InAs+3Weyl structure. The optimization results for these structures are provided in the Supplementary Material. To evaluate the performance of the obtained optimal structures, we compare the contrast between absorptivity and emissivity of them with that of the state-of-the-art 10-layer InAs structure proposed by Liu et al. [22].

### 5.1. Experiment setup

We consider wavelengths  $\lambda$  ranging from 5 to 40  $\mu\text{m}$ , which covers the thermal infrared range and is much broader than that in existing

**Table 1**

Carrier concentrations of InAs and layer thickness for the best 3-layer InAs structures from LCB and PI.

Parameter	LCB	PI
$n_{e,1} (\times 10^{17} \text{ atoms/cm}^3)$	3.353	3.367
$n_{e,2} (\times 10^{17} \text{ atoms/cm}^3)$	5.212	5.181
$n_{e,3} (\times 10^{17} \text{ atoms/cm}^3)$	7.353	7.367
$t_1 (\text{nm})$	1243	1247

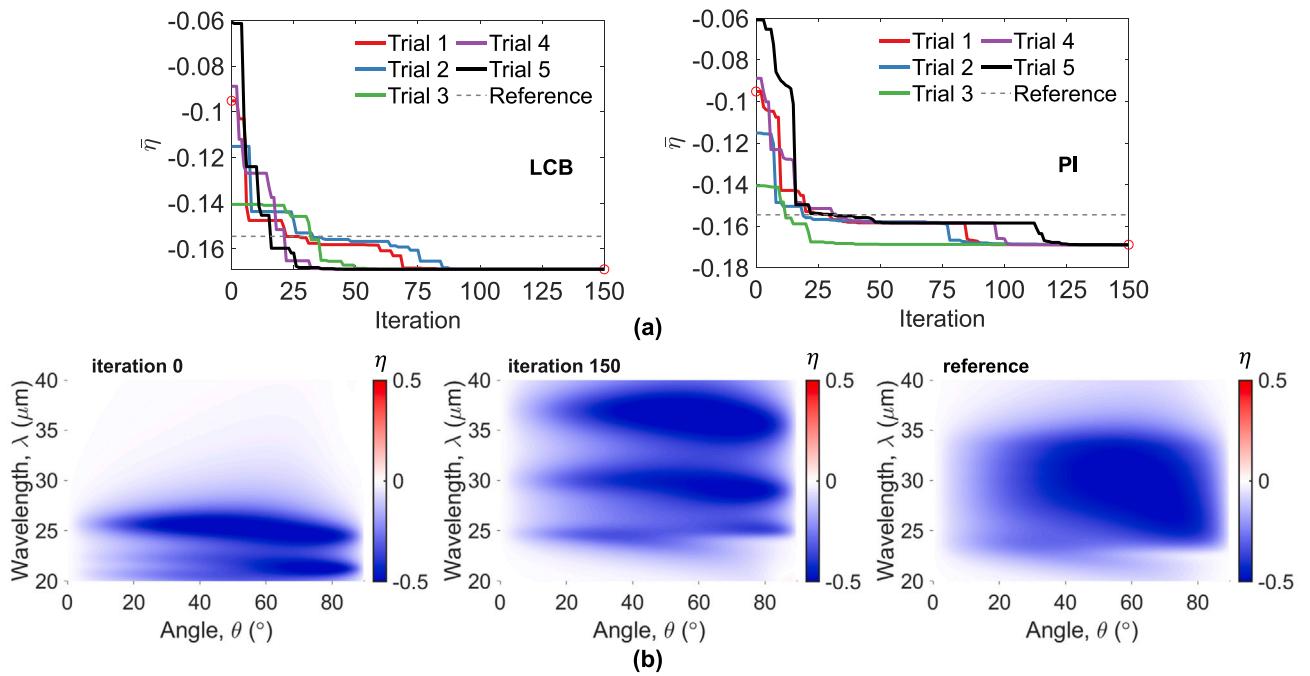
works. We analyze TM waves under an applied magnetic field of  $B = 1.5$  T. The reparameterization is carefully designed for each structure to ensure that the carrier concentrations of InAs and the Weyl Fermi levels across the layers conform to a polynomial of up to second order. As a result, there are four optimization parameters for 3-layer and 5-layer InAs structures, and eight parameters for 6-layer 3InAs+3Weyl and 8-layer 5InAs+3Weyl structures. For 5-layer InAs and 8-layer 5InAs+3Weyl structures, the number of optimization parameters is less than the number of design parameters. Further details on the reparameterization scheme are provided in the Supplementary Material.

There are several settings for BO and the optimizer used for maximizing the acquisition function at each BO iteration. The GP prior is determined by a zero-mean function and the squared exponential covariance function (see Eq. (S2) of the Supplementary Material). The carrier concentrations of InAs layers, the Fermi levels of Weyl semimetal layers, and the layer thicknesses are selected such that  $n_{e,i} \in [1, 10] \times 10^{17} \text{ atoms/cm}^3$ ,  $E_{F,j} \in [10^{-3}, 0.5] \text{ eV}$ , and  $t_1, t_2 \in [100, 5000] \text{ nm}$ . The number of initial observations of the normalized contrast is  $N = 5d$ , where  $d$  is the number of optimization parameters  $\mathbf{x}$ . The threshold for the number of BO iterations is set as  $K = 150$  for 3-layer InAs structure, and  $K = 1000$  for 6-layer 3InAs+3Weyl structure. We use two canonical improvement-based acquisition functions for BO: lower confidence bound (LCB) and probability of improvement (PI). Their analytical formulas are provided in the Supplementary Material. To maximize each of these acquisition functions, we use a multi-start local optimization algorithm with 500 random starting points over the space of  $\mathbf{x}$ . The tolerance for the first-order optimality measure and the upper bound on the magnitude of any constraint functions are set at  $10^{-16}$ . The algorithm is implemented in MATLAB and executed on the Carya Cluster at the University of Houston.

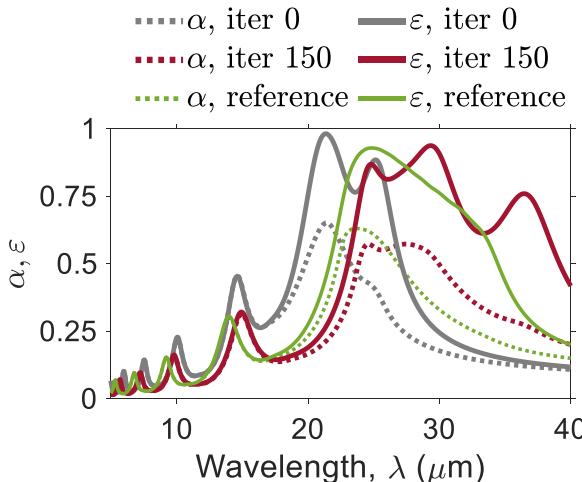
### 5.2. Optimized 3-layer InAs structure

Fig. 4 show the optimization results for the 3-layer InAs structure obtained from five different BO trials of LCB and five different trials of PI, with each set of trials utilizing distinct initial datasets. Despite starting from different initial values, all BO trials converge to a unique normalized contrast value of  $\bar{\eta} = -16.9\%$  after 100 iterations for LCB and 125 iterations for PI, see Fig. 4(a). Thus, each LCB trial requires 20+100 objective function calls to reach the optimal design, while each PI trial needs about 20+125 calls. Fig. 4(b) compares the contrast values of 3-layer InAs structures associated with the first and last iterations of the first BO trial of LCB with that from the state-of-the-art 10-layer InAs structure, which has  $\bar{\eta} = -15.4\%$ . We see that, although BO starts from a less effective initial structure (left panel), it can provide an optimal structure (middle panel) with negative contrast better than that of the state-of-the-art structure (right panel), confirming its crucial role in enhancing the nonreciprocity of the 3-layer InAs structure. Moreover, it shows that we can achieve substantial nonreciprocal effects with only a few layers.

Table 1 lists the optimal parameters for the best 3-layer InAs structures obtained from LCB and PI. The optimal structures from these acquisition functions are almost identical. Notably, the carrier concentrations of these optimal structures exhibit a linear increase from the top to the bottom layers.



**Fig. 4.** Optimization results for 3-layer InAs structure in comparison with the state-of-the-art 10-layer InAs structure (reference) [22]. (a) Optimization histories from LCB and PI. (b) Comparison of contrast values of the initial and final structures from the first and last iterations of the first LCB trial, and the state-of-the-art 10-layer InAs structure.



**Fig. 5.** Absorptivity ( $\alpha$ ) and emissivity ( $\epsilon$ ) of the initial and final 3-layer InAs structures from the first and last iterations of the first LCB trial, and the state-of-the-art 10-layer InAs structure (reference) [22].

Fig. 5 shows the absorptivity and emissivity over the considered range of incidence angles for the 3-layer InAs structures from the first and last iterations of the first BO trial using LCB, as well as those of the state-of-the-art structure. We see that BO enhances the nonreciprocity of the 3-layer InAs structure by increasing the absorptivity at wavelength values from 25 to 40  $\mu\text{m}$ , which correspond to the upper region of the considered spectrum.

### 5.3. Optimized 6-layer 3InAs+3Weyl structure

Fig. 6 presents the optimization results for the 6-layer 3InAs+3Weyl structure. Although different BO trials for this structure using LCB and PI do not converge to a single contrast value after 1000 iterations, they

**Table 2**

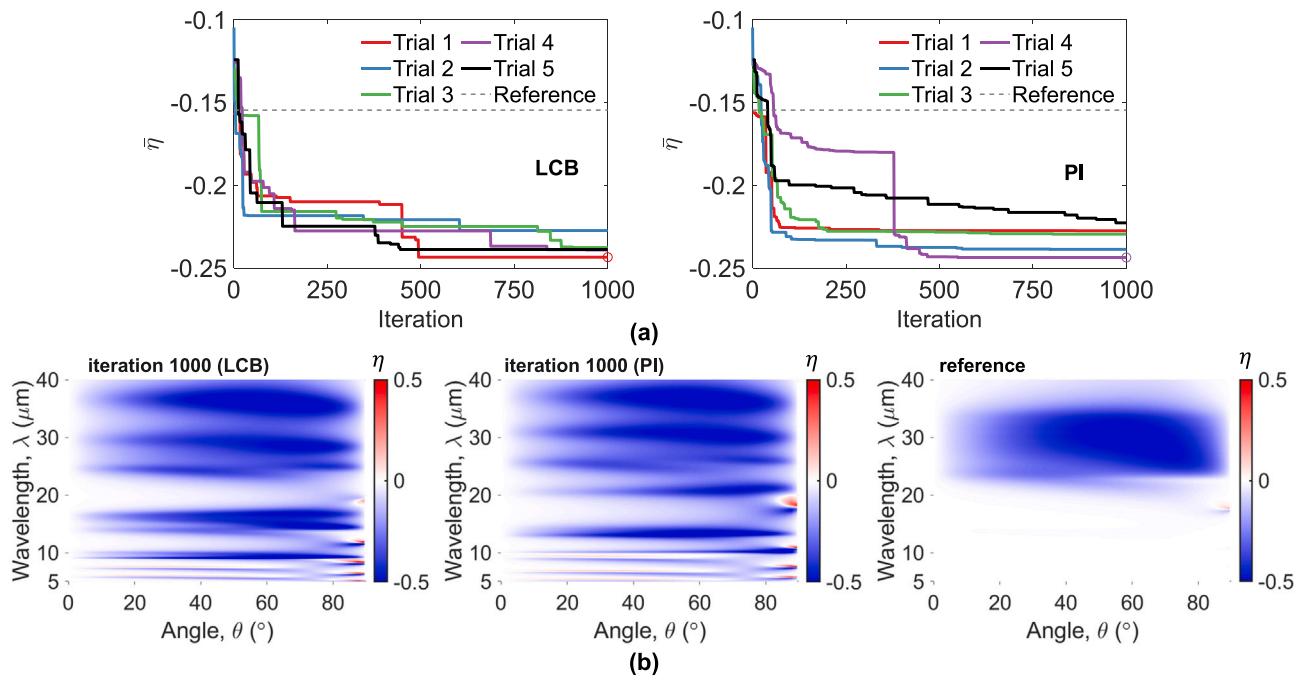
Carrier concentrations of InAs, Weyl Fermi levels, and layer thicknesses for the best 6-layer 3InAs+3Weyl structures from LCB and PI.

Parameter	LCB	PI
$n_{e,1}$ ( $\times 10^{17}$ atoms/cm $^3$ )	3.416	3.308
$n_{e,2}$ ( $\times 10^{17}$ atoms/cm $^3$ )	5.454	4.861
$n_{e,3}$ ( $\times 10^{17}$ atoms/cm $^3$ )	7.416	6.777
$t_1$ (nm)	1203	1066
$E_{F,1}$ (eV)	0.097	0.061
$E_{F,2}$ (eV)	0.148	0.103
$E_{F,3}$ (eV)	0.097	0.132
$t_2$ (nm)	100	100

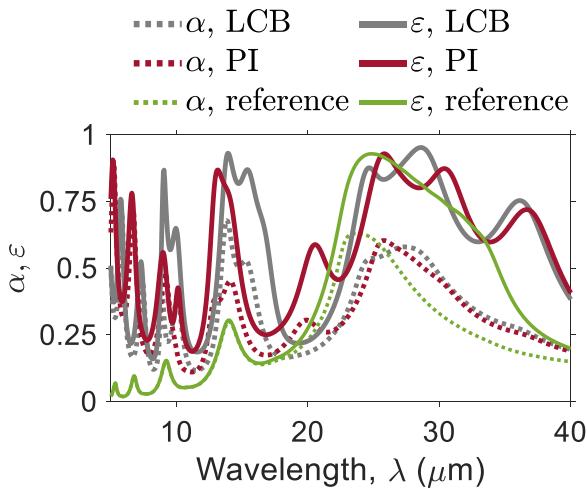
demonstrate significant improvements in the nonreciprocity. The best trial of both LCB and PI requires 500 BO iterations to provide the best structures, which correspond to 40 + 500 objective function calls. The best structures recommended by LCB and PI achieve the same negative contrast value of  $\bar{\eta} = -24.4\%$ , which is notably superior to that of the optimal 3-layer InAs structure with  $\bar{\eta} = -16.9\%$  and that of the state-of-the-art 10-layer InAs structure with  $\bar{\eta} = -15.4\%$ .

As shown in Figs. 6(b) and 7, the use of three InAs layers atop three Weyl semimetal layers enhances the nonreciprocity at wavelengths in both the lower and upper regions of the considered spectrum. Specifically, the top InAs layers improve the absorptivity at wavelength values from 25 to 40  $\mu\text{m}$ , while the bottom Weyl semimetal layers focus on improving the absorptivity at wavelengths from 5 to 15  $\mu\text{m}$ .

Table 2 lists the optimal parameters for the best 6-layer 3InAs+3Weyl structures obtained from LCB and PI. While the two acquisition functions yield the same optimal contrast value ( $\bar{\eta} = -24.4\%$ ), they provide two different sets of optimal parameters. Nevertheless, the carrier concentrations of the InAs layers of these optimal structures still increase linearly from the top to the bottom. Additionally, the thickness of the Weyl semimetal layers reaches the lower bound of its defined domain, indicating that further improvement may be possible with even thinner layers.



**Fig. 6.** Optimization results for 6-layer 3InAs+3Weyl structure in comparison with the state-of-the-art 10-layer InAs structure (reference) [22]. (a) Optimization histories from LCB and PI. (b) Comparison of contrast values of the best structures from LCB and PI, and the state-of-the-art 10-layer InAs structure.



**Fig. 7.** Absorptivity ( $\alpha$ ) and emissivity ( $\epsilon$ ) of the best 6-layer 3InAs+3Weyl structures from LCB and PI, and the state-of-the-art 10-layer InAs structure (reference) [22].

#### 5.4. Discussion

While we choose two common acquisition policies for BO — i.e., LCB and PI — it is beyond the scope of this work to provide a comprehensive performance comparison of different BO algorithms or alternative methods. Other BO acquisition policies, such as Thompson sampling [76–78], may have a better balance between exploitation and exploration and may be less sensitive to the initial dataset.

Alternative optimization methods have been used in photonics design as well, but each has its own limitations. Gradient-based optimizers (e.g., gradient descent) search for a local minimum instead of the global minimum, and they require computing the gradient of the objective function, which is not available for most solvers and is expensive to estimate numerically. Evolutionary algorithms (e.g., genetic algorithm), population-based metaheuristics (e.g., particle swarm optimization),

and other metaheuristic optimizers (e.g., simulated annealing) can find good solutions to global optimization problems, but they usually require significantly more objective function evaluations than BO, which is computationally prohibitive. Some of these can work well on discrete or combinatorial domains, which can be challenging for BO.

Comparing the performance of BO algorithms with these alternatives would help gain deeper insights into the strengths and limitations of each method. To ensure fair comparisons, gradient-based algorithms should start at the best-found solution from the initial dataset of each BO method and expend the same number of objective function calls as required by each BO method, which include those for evaluating the derivatives of the objective function. Population-based algorithms should start at a population that includes the samples from the initial dataset of each BO method and should also use the same number of objective function evaluations as that of each BO method.

#### 6. Conclusion

We present an optimization approach combining BO and reparameterization that demonstrates optimal broadband nonreciprocal thermal emitter performance, surpassing state-of-the-art broadband nonreciprocal effects in the infrared range of thermal radiation using doped InAs and Weyl semimetal materials. Starting from a less effective structure, the proposed approach incrementally improves the broadband nonreciprocity of the structure by repeatedly reparameterizing the normalized contrast, constructing a probabilistic surrogate model for the reparameterized contrast, maximizing a cheap-to-compute acquisition function formulated from the constructed surrogate model to identify a promising new structure, and updating the surrogate model with the new structure. The optimal structure is the best structure among those recommended by the optimization algorithm upon its termination. Optimization results indicate that our approach can propose an optimal structure of only three InAs layers that outperforms the current state-of-the-art 10-layer InAs structure [22]. Additionally, the broadband nonreciprocal effect considerably increases when using InAs and Weyl semimetals, which shows the feasibility of combining different nonreciprocal materials for enhanced nonreciprocity. The significant

improvements in the designed emitters highlight the role of numerical optimization in advancing practical nonreciprocal thermal emitter development. Our approach can also be adapted to optimize more general structures containing patterns. In an ongoing study, we conduct experiments to validate the performance of the optimized structures. We then combine the experimental and numerical observations within the framework of multi-fidelity BO [69,70,79–81] to further enhance our design approach.

## CRediT authorship contribution statement

**Bach Do:** Conceptualization, Data curation, Methodology, Writing – original draft. **Sina Jafari Ghalekohneh:** Conceptualization, Methodology, Software, Writing – original draft. **Taiwo Adebiyi:** Data curation, Resources, Writing – review & editing. **Bo Zhao:** Conceptualization, Funding acquisition, Software, Supervision, Writing – review & editing. **Ruda Zhang:** Conceptualization, Data curation, Funding acquisition, Methodology, Software, Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

The authors acknowledge the funding from the University of Houston through the SEED program and the National Science Foundation, USA under Grant No. CBET-2314210, and the support of the Research Computing Data Core at the University of Houston, USA for assistance with the calculations carried out in this work.

## Appendix A. Supplementary data

Supplementary Material contains the mathematical foundation of Gaussian process models, the expressions of LCB and PI acquisition functions, the quadrature method for computing the contrast between absorptivity and emissivity at a particular value of design parameters, the details of the reparameterization scheme, and additional results for 5-layer InAs and 8-layer 5InAs+3Weyl structures.

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jqsrt.2024.109260>.

## Data availability

Data will be made available on request.

## References

- Greffet J-J, Carminati R, Joulain K, Mulet J-P, Mainguy S, Chen Y. Coherent emission of light by thermal sources. *Nature* 2002;416(6876):61–4.
- Fan S. Thermal photonics and energy applications. *Joule* 2017;1(2):264–73.
- Baranov DG, Xiao Y, Nechepurenko IA, Krasnok A, Alù A, Kats MA. Nanophotonic engineering of far-field thermal emitters. *Nature Mater* 2019;18(9):920–30.
- Li Y, Li W, Han T, Zheng X, Li J, Li B, Fan S, Qiu C-W. Transforming heat transfer with thermal metamaterials and devices. *Nat Rev Mater* 2021;6(6):488–507.
- Kirchhoff G. In: Schöpf H-G, editor. Von kirchhoff bis planck: Theorie der wärmestrahlung in historisch-kritischer darstellung. Vieweg+Teubner Verlag; 1978, p. 131–51.
- Planck M. The theory of heat radiation. Blakiston; 1914.
- Ries H. Complete and reversible absorption of radiation. *Appl Phys B* 1983;32:153–6.
- Snyder QC, Wan Z, Li X. Thermodynamic constraints on reflectance reciprocity and Kirchhoff's law. *Appl Opt* 1998;37(16):3464–70.
- Green MA. Time-asymmetric photovoltaics. *Nano Lett* 2012;12(11):5985–8.
- Buddhiraju S, Santhanam P, Fan S. Thermodynamic limits of energy harvesting from outgoing thermal radiation. *Proc Natl Acad Sci* 2018;115(16):E3609–15.
- Li W, Buddhiraju S, Fan S. Thermodynamic limits for simultaneous energy harvesting from the hot sun and cold outer space. *Light: Sci Appl* 2020;9(1):68.
- Park Y, Zhao B, Fan S. Reaching the ultimate efficiency of solar energy harvesting with a nonreciprocal multijunction solar cell. *Nano Lett* 2021;22(1):448–52.
- Jafari Ghalekohneh S, Zhao B. Nonreciprocal solar thermophotovoltaics. *Phys Rev Appl* 2022;18:034083.
- Park Y, Omair Z, Fan S. Nonreciprocal thermophotovoltaic systems. *ACS Photonics* 2022;9(12):3943–9.
- Zhu L, Fan S. Near-complete violation of detailed balance in thermal radiation. *Phys Rev B* 2014;90:220301.
- Zhao B, Guo C, Garcia CAC, Narang P, Fan S. Axion-field-enabled nonreciprocal thermal radiation in Weyl semimetals. *Nano Lett* 2020;20(3):1923–7.
- Shayegan KJ, Biswas S, Zhao B, Fan S, Atwater HA. Direct observation of the violation of Kirchhoff's law of thermal radiation. *Nat Photonics* 2023;17(10):891–6.
- Park Y, Zhao B, Fan S. Reaching the ultimate efficiency of solar energy harvesting with a nonreciprocal multijunction solar cell. *Nano Lett* 2022;22(1):448–52.
- Zhang Z, Zhu L. Nonreciprocal thermal photonics for energy conversion and radiative heat transfer. *Phys Rev Appl* 2022;18:027001.
- Yang S, Liu M, Zhao C, Fan S, Qiu C-W. Nonreciprocal thermal photonics. *Nat Photonics* 2024;18(5):412–24.
- Zhu L, Fan S. Persistent directional current at equilibrium in nonreciprocal many-body near field electromagnetic heat transfer. *Phys Rev Lett* 2016;117:134303.
- Liu M, Xia S, Wan W, Qin J, Li H, Zhao C, Bi L, Qiu C-W. Broadband mid-infrared non-reciprocal absorption using magnetized gradient epsilon-near-zero thin films. *Nature Mater* 2023;22(10):1196–202.
- Dong J, Zhang W, Liu L. Nonreciprocal thermal radiation of nanoparticles via spin-directional coupling with reciprocal surface modes. *Appl Phys Lett* 2021;119(2):021104.
- Fan L, Guo Y, Papadakis GT, Zhao B, Zhao Z, Buddhiraju S, Orenstein M, Fan S. Nonreciprocal radiative heat transfer between two planar bodies. *Phys Rev B* 2020;101:085407.
- Khandekar C, Buddhiraju S, Wilkinson PR, Gimzewski JK, Rodriguez AW, Chase C, Fan S. Nonequilibrium lateral force and torque by thermally excited nonreciprocal surface electromagnetic waves. *Phys Rev B* 2021;104:245433.
- Gelbwaser-Klimovsky D, Graham N, Kardar M, Krüger M. Near field propulsion forces from nonreciprocal media. *Phys Rev Lett* 2021;126:170401.
- Shi Y, Li W, Raman A, Fan S. Optimization of multilayer optical films with a memetic algorithm and mixed integer programming. *ACS Photonics* 2018;5(3):684–91.
- Yu S, Zhou P, Xi W, Chen Z, Deng Y, Luo X, Li W, Shiomi J, Hu R. General deep learning framework for emissivity engineering. *Light: Sci Appl* 2023;12(1):291.
- Zhao B, Shi Y, Wang J, Zhao Z, Zhao N, Fan S. Near-complete violation of Kirchhoff's law of thermal radiation with a 0.3 T magnetic field. *Opt Lett* 2019;44(17):4203–6.
- Zhang Z, Zhu L. Broadband nonreciprocal thermal emission. *Phys Rev Appl* 2023;19:014013.
- Wang L, García de Abajo FJ, Papadakis GT. Maximal violation of Kirchhoff's law in planar heterostructures. *Phys Rev Res* 2023;5(2):L022051.
- Jafari Ghalekohneh S, Du C, Zhao B. Controlling the contrast between absorptivity and emissivity in nonreciprocal thermal emitters. *Appl Phys Lett* 2024;124(10):101104.
- Zhu L, Fan S. Persistent directional current at equilibrium in nonreciprocal many-body near field electromagnetic heat transfer. *Phys Rev Lett* 2016;117(13):134303.
- Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. *J Global Optim* 1998;13(4):455–92.
- Brochu E, Cora VM, de Freitas N. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. 2010, arXiv preprint arXiv:1012.2599.
- Shahriari B, Swersky K, Wang Z, Adams RP, de Freitas N. Taking the human out of the loop: A review of Bayesian optimization. *Proc IEEE* 2016;104(1):148–75.
- Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. In: Pereira F, Burges C, Bottou L, Weinberger K, editors. Advances in neural information processing systems. Vol. 25. Curran Associates, Inc.; 2012, p. 2951–9.
- Frazier PI. Recent advances in optimization and modeling of contemporary problems. In: Gel E, Ntaimo L, editors. In: INFORMS tutORials in operations research, INFORMS; 2018, p. 255–78, Chapter Bayesian optimization.
- Garnett R. Bayesian optimization. Cambridge, UK: Cambridge University Press; 2023.
- Gold H, Pajovic S, Mukherjee A, Boriskina SV. GAGA for nonreciprocal emitters: genetic algorithm gradient ascent optimization of compact magnetophotonic crystals. *Nanophotonics* 2024;13(5):773–92.
- Hughes TW, Minkov M, Williamson IAD, Fan S. Adjoint method and inverse design for nonlinear nanophotonic devices. *ACS Photonics* 2018;5(12):4781–7.
- Kudyshev ZA, Kildishev AV, Shalaev VM, Boltasseva A. Machine learning-assisted global optimization of photonic devices. *Nanophotonics* 2021;10(1):371–83.

[43] Zhu L, Fan S. Near-complete violation of detailed balance in thermal radiation. *Phys Rev B* 2014;90(22):220301.

[44] Zhao B, Wang J, Zhao Z, Guo C, Yu Z, Fan S. Nonreciprocal thermal emitters using metasurfaces with multiple diffraction channels. *Phys Rev Appl* 2021;16(6):064001.

[45] Zhao Z, Guo C, Fan S. Connection of temporal coupled-mode-theory formalisms for a resonant optical system and its time-reversal conjugate. *Phys Rev A* 2019;99(3):033839.

[46] Zhang ZM, Zhang ZM, Luby. *Nano/microscale heat transfer*, vol. 410, Springer; 2007.

[47] Ghanekar A, Wang J, Fan S, Povinelli ML. Violation of Kirchhoff's law of thermal radiation with space-time modulated grating. *ACS Photonics* 2022;9(4):1157–64.

[48] Ghanekar A, Wang J, Guo C, Fan S, Povinelli ML. Nonreciprocal thermal emission using spatiotemporal modulation of graphene. *ACS Photonics* 2022;10(1):170–8.

[49] Shayegan KJ, Hwang JS, Zhao B, Raman AP, Atwater HA. Broadband nonreciprocal thermal emissivity and absorptivity. *Light: Sci Appl* 2024;13(1):176.

[50] Pajovic S, Tsurimaki Y, Qian X, Chen G. Intrinsic nonreciprocal reflection and violation of Kirchhoff's law of radiation in planar type-I magnetic Weyl semimetal surfaces. *Phys Rev B* 2020;102(16):165417.

[51] Park Y, Asadchy VS, Zhao B, Guo C, Wang J, Fan S. Violating Kirchhoff's law of thermal radiation in semitransparent structures. *ACS Photonics* 2021;8(8):2417–24.

[52] Picardi MF, Moerbeek VI, Pascale M, Papadakis GT. Nonreciprocity in transmission mode with planar structures for arbitrarily polarized light. *Opt Mater Express* 2024;14(9):2201–9.

[53] Pershan PS. Magneto-optical effects. *J Appl Phys* 1967;38(3):1482–90.

[54] Aers GC, Boardman AD. The theory of semiconductor magnetoplasmon-polariton surface modes: Voigt geometry. *J Phys C* 1978;11(5):945.

[55] Halterman K, Alidoust M, Zyuzin A. Epsilon-near-zero response and tunable perfect absorption in Weyl semimetals. *Phys Rev B* 2018;98:085109.

[56] Campione S, Brener I, Marquier F. Theory of epsilon-near-zero modes in ultrathin films. *Phys Rev B* 2015;91:121408.

[57] Kinsey N, DeVault C, Boltasseva A, Shalaev VM. Near-zero-index materials for photonics. *Nat Rev Mater* 2019;4(12):742–60.

[58] Xu J, Mandal J, Raman AP. Broadband directional control of thermal emission. *Science* 2021;372(6540):393–7.

[59] Madelung O. *Semiconductors: Data handbook*. Springer Science & Business Media; 2004.

[60] Du C, Zhao B. Controlling the ENZ profile for broadband nonreciprocal thermal emitters with high contrast between emissivity and absorptivity. In: Proceedings of the 10th international symposium on radiative transfer. 2023.

[61] Xu W, Liu Z, Piper RT, Hsu JW. Bayesian optimization of photonic curing process for flexible perovskite photovoltaic devices. *Sol Energy Mater Sol Cells* 2023;249:112055.

[62] Talapatra A, Boluki S, Duong T, Qian X, Dougherty E, Arróyave R. Autonomous efficient experiment design for materials discovery with Bayesian model averaging. *Phys Rev Mater* 2018;2:113803.

[63] Greenhill S, Rana S, Gupta S, Vellanki P, Venkatesh S. Bayesian optimization for adaptive experimental design: A review. *IEEE Access* 2020;8:13937–48.

[64] Roussel R, Edelen AL, Boltz T, Kennedy D, Zhang Z, Ji F, Huang X, Ratner D, Garcia AS, Xu C, Kaiser J, Pousa AF, Eichler A, Lübsen JO, Isenberg NM, Gao Y, Kuklev N, Martinez J, Mustapha B, Kain V, Mayes C, Lin W, Luzzo SM, St. John J, Streeter MJV, Lehe R, Neiswanger W. Bayesian optimization algorithms for accelerator physics. *Phys Rev Accel Beams* 2024;27:084801.

[65] Frazier PI, Wang J. *Information science for materials discovery and design*. Springer International Publishing; 2016, p. 45–75, Chapter Bayesian optimization for materials design,

[66] Vangelatos Z, Sheikh HM, Marcus PS, Grigoropoulos CP, Lopez VZ, Flamourakis G, Farsari M. Strength through defects: A novel Bayesian approach for the optimization of architected materials. *Sci Adv* 2021;7(41):eabk2218.

[67] Khatamsaz D, Neuberger R, Roy AM, Zadeh SH, Otis R, Arróyave R. A physics informed Bayesian optimization approach for material design: application to NiTi shape memory alloys. *Npj Comput Mater* 2023;9(1):221.

[68] Astudillo R, Frazier PI. Thinking inside the box: a tutorial on grey-box Bayesian optimization. In: Proceedings of the winter simulation conference. WSC '21, IEEE Press; 2022, p. 1–15.

[69] Do B, Zhang R. Multi-fidelity Bayesian optimization: A review. 2023, arXiv preprint arXiv:2311.13050.

[70] Zhang R, Alemazkoor N. Multi-fidelity machine learning for uncertainty quantification and optimization. *J Mach Learn Model Comput* 2024;5(4):77–94.

[71] Rasmussen CE, Williams CKI. *Gaussian processes for machine learning*. Massachusetts, USA: The MIT Press; 2006.

[72] Jones DR. A taxonomy of global optimization methods based on response surfaces. *J Global Optim* 2001;21(4):345–83.

[73] Owen AB. A central limit theorem for Latin hypercube sampling. *J R Stat Soc Ser B Stat Methodol* 1992;54(2):541–51.

[74] Moharam MG, Gaylord TK. Rigorous coupled-wave analysis of planar-grating diffraction. *J Opt Soc Am* 1981;71(7):811–8.

[75] Yang C, Zhao B, Cai W, Zhang ZM. Mid-infrared broadband circular polarizer based on Weyl semimetals. *Opt Express* 2022;30(2):3035–46.

[76] Adebiyi T, Do B, Zhang R. Gaussian process Thompson sampling via rootfinding. In: NeurIPS 2024 workshop on Bayesian decision-making and uncertainty. 2024, URL <https://openreview.net/forum?id=1pRLTVblaV>.

[77] Adebiyi T, Do B, Zhang R. Optimizing posterior samples for Bayesian optimization via rootfinding. 2024, <http://dx.doi.org/10.48550/arXiv.2410.22322>, arXiv.

[78] Do B, Adebiyi T, Zhang R. Epsilon-greedy Thompson sampling to Bayesian optimization. *J Comput Inf Sci Eng* 2024;24(12):121006. <http://dx.doi.org/10.1115/1.4066858>.

[79] Huang D, Allen TT, Notz WI, Miller RA. Sequential Kriging optimization using multiple-fidelity evaluations. *Struct Multidiscip Optim* 2006;32(5):369–82.

[80] Forrester AJ, Sobester A, Keane AJ. Multi-fidelity optimization via surrogate modelling. *Proc R Soc Lond Ser A Math Phys Eng Sci* 2007;463(2088):3251–69.

[81] Kandasamy K, Dasarathy G, Schneider J, Póczos B. Multi-fidelity Bayesian optimisation with continuous approximations. In: Precup D, Teh YW, editors. *Proceedings of the 34th international conference on machine learning*. vol. 70, PMLR; 2017, p. 1799–808.

# Supporting Information

## Gaussian process

A Gaussian process (GP) model<sup>1</sup> is a non-parametric, probabilistic model that enables the prediction of an unknown function of interest  $f(\mathbf{x})$  for any  $\mathbf{x} \in \mathbb{R}^d$  and quantifies the uncertainty in that prediction. In this work,  $f(\mathbf{x})$  represents the normalized contrast function  $\bar{\eta}(\mathbf{x})$ , see problem (12) in **Optimization problem**. The GP model can also be viewed as a distribution over functions, extending the concept of finite-dimensional Gaussians to an infinite-dimensional space. It assumes that the function values  $\mathbf{f} = \{f(\mathbf{x}_i)\}_{i=1}^N$  evaluated at any finite set  $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^N$  of input variables are distributed according to a multivariate Gaussian. This assumption is encoded in the following GP prior:

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), \kappa(\mathbf{x}, \mathbf{x}'|\boldsymbol{\theta})), \quad (\text{S1})$$

where  $m(\mathbf{x})$  and  $\kappa(\mathbf{x}, \mathbf{x}'|\boldsymbol{\theta})$  are the mean and covariance functions, respectively. The mean function  $m(\mathbf{x})$  determines the expected value of  $f(\mathbf{x})$  at any location  $\mathbf{x}$  and is often set as  $m(\mathbf{x}) = 0$ . The covariance function  $\kappa(\mathbf{x}, \mathbf{x}'|\boldsymbol{\theta})$  defines the covariance between  $f(\mathbf{x})$  and  $f(\mathbf{x}')$  and is characterized by a set of parameters  $\boldsymbol{\theta}$ , known as hyperparameters. Hereafter, we omit the dependence of  $\kappa$  on  $\boldsymbol{\theta}$  for clarity.

The covariance function profoundly impacts important properties of a GP model such as continuity and differentiability of its sample paths.<sup>2</sup> These properties, in turn, affect the uniqueness and existence of a global solution. A popular choice is the squared exponential (SE) covariance function, which reads

$$\kappa(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2} \sum_{k=1}^d \frac{(x_k - x'_k)^2}{l_j^2}\right), \quad (\text{S2})$$

where  $\sigma_f$  is the scaling standard deviation and  $l_j$  are the characteristic length scales. The hyperparameters for the SE covariance function is  $\boldsymbol{\theta} = [\sigma_f, l_1, \dots, l_d]^T$ .

We now wish to find a most-likely set for the hyperparameters  $\boldsymbol{\theta}$  given the observations  $\mathbf{y} = \{y(\mathbf{x}_i)\}_{i=1}^N$  for the function values  $\mathbf{f} = \{f(\mathbf{x}_i)\}_{i=1}^N$ . We let  $p(\mathbf{f}|\mathbf{X}) = \mathcal{N}(\mathbf{0}, \mathbf{K}(\mathbf{X}, \mathbf{X}))$  represent the zero-mean GP prior from Eq. (S3), where  $p(\cdot)$  represents a probability density function,  $\mathcal{N}(\cdot)$  denotes a Gaussian, and the  $(i, j)$ th entry of  $\mathbf{K}$  is  $\mathbf{K}_{[ij]} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$  ( $i, j = 1, \dots, N$ ). We further assume that  $\mathbf{y}$  are generated from an observation model  $p(\mathbf{y}|\mathbf{f}, \mathbf{X}) = \mathcal{N}(\mathbf{f}, \sigma_n^2 \mathbf{I})$  known as the likelihood, where  $\sigma_n^2$  and  $\mathbf{I}$  represent the observation variance and the identity matrix, respectively. Therefore, we can derive the marginal likelihood (for a zero-mean function)  $p(\mathbf{y}|\mathbf{X}) = \int p(\mathbf{f}|\mathbf{X})p(\mathbf{y}|\mathbf{f}, \mathbf{X})d\mathbf{f} = \mathcal{N}(\mathbf{0}, \mathbf{C}(\mathbf{X}, \mathbf{X}))$  and compute the associated log-marginal likelihood  $\log p(\mathbf{y}|\mathbf{X})$ , where  $\mathbf{C}(\mathbf{X}, \mathbf{X}) = \mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}$ . We then maximize  $\log p(\mathbf{y}|\mathbf{X})$  for a set of most-likely hyperparameters  $\boldsymbol{\theta}$  given the observations, which is called the maximum likelihood estimate of the hyperparameters.

Once the hyperparameters have been chosen, the posterior  $p(\mathbf{f}|\mathbf{X}, \mathbf{y})$  can be obtained from Bayes' rule, such that

$$p(\mathbf{f}|\mathbf{X}, \mathbf{y}) = \frac{p(\mathbf{f}|\mathbf{X})p(\mathbf{y}|\mathbf{f}, \mathbf{X})}{p(\mathbf{y}|\mathbf{X})}. \quad (\text{S3})$$

In this case, the posterior  $p(\mathbf{f}|\mathbf{X}, \mathbf{y})$  is a Gaussian because  $p(\mathbf{f}|\mathbf{X})$  and  $p(\mathbf{y}|\mathbf{f}, \mathbf{X})$  are Gaussians. Thus, the posterior is also a GP, which reads<sup>1</sup>

$$\widehat{f}(\mathbf{x}) \sim \mathcal{GP}(\widehat{m}(\mathbf{x}), \widehat{\kappa}(\mathbf{x}, \mathbf{x}')), \quad (\text{S4})$$

where the conditional mean  $\widehat{m}(\mathbf{x}) = \mathbf{K}_{\mathbf{x}}^T \mathbf{C}^{-1} \mathbf{y}$  and the conditional covariance  $\widehat{\kappa}(\mathbf{x}, \mathbf{x}') = \kappa(\mathbf{x}, \mathbf{x}') - \mathbf{K}_{\mathbf{x}}^T \mathbf{C}^{-1} \mathbf{K}_{\mathbf{x}'}$  with  $\mathbf{K}_{\mathbf{x}} = [\kappa(\mathbf{x}, \mathbf{x}_1), \dots, \kappa(\mathbf{x}, \mathbf{x}_N)]^T$  and  $\mathbf{K}_{\mathbf{x}'} = [\kappa(\mathbf{x}', \mathbf{x}_1), \dots, \kappa(\mathbf{x}', \mathbf{x}_N)]^T$ . The prediction  $f_*$  at a new point  $\mathbf{x}_*$  can be calculated from the GP posterior as  $p(f_*|\widehat{f}, \mathbf{x}_*) = \mathcal{N}(\widehat{\mu}(\mathbf{x}_*), \widehat{\sigma}^2(\mathbf{x}_*))$ , where  $\widehat{\mu}(\mathbf{x}_*) = \widehat{m}(\mathbf{x}_*)$  and  $\widehat{\sigma}^2(\mathbf{x}_*) = \widehat{\kappa}(\mathbf{x}_*, \mathbf{x}_*)$  are the predictive mean and variance, respectively.

## Lower confidence bound and probability of improvement

In this section, we describe the lower confidence bound (LCB) and probability of improvement (PI) acquisition functions BO uses for optimizing the normalized contrast  $\bar{\eta}(\mathbf{x})$  in **Initialization**.

The LCB acquisition function balances greedy optimization introduced by minimizing the predictive mean function  $\hat{\mu}(\mathbf{x})$  (i.e., maximizing  $-\hat{\mu}(\mathbf{x})$ ) with uncertainty reduction employed by maximizing the predictive standard deviation function  $\hat{\sigma}(\mathbf{x})$ . It is a variant of the GP-lower confidence bound,<sup>3</sup> which has been proposed for maximization problems. The LCB acquisition function reads

$$q(\mathbf{x}) = -(\hat{\mu}(\mathbf{x}) - \beta\hat{\sigma}(\mathbf{x})), \quad (\text{S5})$$

where  $\beta \geq 0$  is a tuning parameter trading off between greedy optimization and uncertainty reduction.

The PI acquisition function<sup>4</sup> measures the probability that the one-step lookahead observation of the objective function  $f(\mathbf{x})$  recommended by BO is better than the best value  $f_{\min}$  it has observed so far. This is equivalent to measuring the chance of having a solution improvement in the next iteration of BO. Conditioning PI on the current GP posterior  $\hat{f}(\mathbf{x})$ , it can be analytically written as

$$q(\mathbf{x}) = \mathbb{P}[f(\mathbf{x}) < f_{\min} | \hat{f}(\mathbf{x})] = \Phi\left(\frac{f_{\min} - \hat{\mu}(\mathbf{x})}{\hat{\sigma}(\mathbf{x})}\right), \quad (\text{S6})$$

where  $\Phi(\cdot)$  represents the standard normal cumulative distribution function.

## Quadrature for computing contrast values

We describe here the numerical quadrature method used to compute the contrast  $\eta(\mathbf{p})$  for a specific value of design parameters  $\mathbf{p}$ . Recall that

$$\begin{aligned}\eta(\mathbf{p}) &= \int_{\lambda_L}^{\lambda_U} \int_0^{\pi/2} \eta(\mathbf{p}, \lambda, \theta) \sin \theta \cos \theta d\theta d\lambda \\ &= \int_{\lambda_L}^{\lambda_U} \int_0^{\pi/2} (\alpha(\mathbf{p}, \lambda, \theta) - \varepsilon(\mathbf{p}, \lambda, \theta)) \sin \theta \cos \theta d\theta d\lambda,\end{aligned}\tag{S7}$$

where  $\alpha(\mathbf{p}, \lambda, \theta)$  and  $\varepsilon(\mathbf{p}, \lambda, \theta)$  represent the local functions for the absorptivity and emissivity. Hereafter, we omit the dependence of  $\eta$ ,  $\alpha$ , and  $\varepsilon$  on  $\mathbf{p}$  for clarity.

Computing the integrals in Eq. (S7) requires evaluations of  $\alpha$  and  $\varepsilon$  at many different  $(\lambda, \theta)$  locations, which is costly. To improve computational efficiency and accuracy, we apply a numerical quadrature scheme for the integrals. A quadrature scheme approximates an integral  $I = \int_a^b f(x)dx$  with a numerical formula  $I_n = \sum_{k=1}^n w_k f(x_k)$ , where  $x_1, \dots, x_n$  are the nodes at which the integrand is evaluated and  $w_1, \dots, w_n$  are the quadrature weights.

First we separate the integrals of  $\alpha(\lambda, \theta)$  and  $\varepsilon(\lambda, \theta)$ , because these two functions are less fluctuating than  $\eta(\lambda, \theta)$  and thus easier to integrate. Our quadrature method applies identically to both  $\alpha$  and  $\varepsilon$ , so we only present the calculation for  $\alpha$  below. Starting with the inner integral of Eq. (S7) with respect to the angle of incidence  $\theta$ , we define

$$\alpha(\lambda) = \int_0^{\pi/2} \alpha(\lambda, \theta) \sin \theta \cos \theta d\theta.\tag{S8}$$

To approximate this integral, we use Gauss quadrature with  $n_\theta$  nodes and the corresponding weights. For the details of Gauss quadrature, see Chapter 19 in this textbook.<sup>5</sup>

To simplify the computation of the outer integral, we transform wavelength  $\lambda$  to wavenumber  $\nu$ . This is because  $\alpha(\nu)$  and  $\eta(\nu)$  are smoother than  $\alpha(\lambda)$  and  $\eta(\lambda)$ , respectively. Define

$$\alpha_\lambda = \int_{\lambda_L}^{\lambda_U} \alpha(\lambda) d\lambda.\tag{S9}$$

Let  $\lambda_1 = s_\lambda \lambda$  and  $\nu_1 = \lambda_1^{-1}$ , where the scaling factor  $s_\lambda = 10^5$  for  $\lambda$  in meters. This transformation maps the wavelength  $\lambda \in [\lambda_L, \lambda_U]$  m to  $\lambda_1 \in [\lambda_{1L}, \lambda_{1U}] \times 10^{-5}$  m, and subsequently to the wavenumber  $\nu_1 \in [\nu_{1L}, \nu_{1U}]$ . Thus, we have

$$\alpha_\lambda = \int_{\lambda_L}^{\lambda_U} \alpha(\lambda) d\lambda = s_\lambda^{-1} \alpha_{\lambda_1}, \quad (S10)$$

where

$$\alpha_{\lambda_1} = \int_{\lambda_{1L}}^{\lambda_{1U}} \alpha(\lambda_1) d\lambda_1 = \int_{\nu_{1L}}^{\nu_{1U}} \alpha(g(\nu_1)) g'(\nu_1) d\nu_1. \quad (S11)$$

Here,  $g(\nu_1) = 1/\nu_1$  and  $g'(\nu_1) = -1/\nu_1^2$ .

We compute the integral in Eq. (S11) with a transformed Gauss quadrature.<sup>6</sup> It starts with a linear change of variables from  $\nu_1 \in [\nu_{1L}, \nu_{1U}]$  to  $x \in [-1, 1]$ . This allows us to utilize the following relationship between the integrals of a function  $k(s)$  over  $[-1, 1]$  and over a strip map  $l(x)$  with  $x \in [-1, 1]$  to compute  $\alpha_{\lambda_1}$ :

$$I(k) = \int_{-1}^1 k(s) ds = \int_{-1}^1 k(l(x)) l'(x) dx, \quad (S12)$$

where  $k$  and  $l$  play the same roles as  $\alpha$  and  $g$  in Eq. (S11), respectively. The transformed integral is computed using Gauss quadrature with  $n_s$  nodes and the corresponding weights.

All optimization results presented in this study are generated by  $n_\theta = 24$  and  $n_s = 140$ . The significantly higher value of  $n_s$  compared to  $n_\theta$  is due to the substantial fluctuations in contrast along the wavelength direction. Based on numerical results, this quadrature scheme achieves a relative error of  $\mathcal{O}(10^{-5})$ .

## Details of the reparameterization scheme

This section details the reparameterization strategy we propose for optimizing the nonreciprocal emitters presented in this work. In **Optimization problem**, we transform the design parameters governing the carrier concentrations  $\mathbf{N} = [n_{e,1}, \dots, n_{e,n_1}]^T$  of  $n_1$  InAs layers and

Fermi levels  $\mathbf{E} = [E_{F,1}, \dots, E_{F,n_2}]^T$  of  $n_2$  Weyl layers into optimization parameters  $\mathbf{v}$  and  $\mathbf{w}$  using two linear maps  $\mathbf{D}$  and  $\mathbf{F}$ , respectively. Specifically, we have (see Figure 2)

$$\begin{aligned}\mathbf{N} = \mathbf{D}\mathbf{v} &= \sum_{i=0}^{p_1-1} v_i \mathbf{D}_{[i]}, \\ \mathbf{E} = \mathbf{F}\mathbf{w} &= \sum_{j=0}^{p_2-1} w_j \mathbf{F}_{[j]},\end{aligned}\tag{S13}$$

where  $\mathbf{N} \in \mathbb{R}^{n_1}$ ,  $\mathbf{E} \in \mathbb{R}^{n_2}$ ,  $\mathbf{D} \in \mathbb{R}^{n_1 \times p_1}$ ,  $\mathbf{v} \in \mathbb{R}^{p_1}$ ,  $\mathbf{F} \in \mathbb{R}^{n_2 \times p_2}$ ,  $\mathbf{w} \in \mathbb{R}^{p_2}$ ,  $\mathbf{D}_{[i]} \in \mathbb{R}^{n_1}$  is the  $(i+1)$ th column of  $\mathbf{D}$  ( $i = 0, \dots, p_1-1$ ), and  $\mathbf{F}_{[j]} \in \mathbb{R}^{n_2}$  is the  $(j+1)$ th column of  $\mathbf{F}$  ( $j = 0, \dots, p_2-1$ ). To increase the efficiency of BO when optimizing structures of a large number of layers, we set  $p_1 < n_1$  and  $p_2 < n_2$ .

To further control the profiles of the InAs carrier concentrations and the Weyl Fermi levels inside the multilayer structure, we define the  $k$ th entry of vector  $\mathbf{D}_{[i]}$  and the  $l$ th entry of vector  $\mathbf{F}_{[j]}$  using layer-wise integration of Chebyshev polynomials of the first kind, as

$$\begin{aligned}\mathbf{D}_{[ik]} &= \int_{u_{k-1}}^{u_k} T_i(u) du, \quad i = 0, \dots, p_1-1, \quad k = 1, \dots, n_1, \\ \mathbf{F}_{[jl]} &= \int_{u_{l-1}}^{u_l} T_l(u) du, \quad j = 0, \dots, p_2-1, \quad l = 1, \dots, n_2,\end{aligned}\tag{S14}$$

where  $T_i(u)$  represents the  $i$ th order Chebyshev polynomial of the first kind evaluated at any  $u \in [-1, 1]$ ,  $u_k = 2k/n_1 - 1$  for  $k = 1, \dots, n_1$ , and  $u_l = 2l/n_2 - 1$  for  $l = 1, \dots, n_2$ . The first three Chebyshev polynomials of the first kind are  $T_0 = 1$ ,  $T_1(u) = u$ , and  $T_2(u) = 2u^2 - 1$ . Higher order polynomials can be obtained from the recurrence relation  $T_{i+1}(u) = 2uT_i(u) - T_{i-1}(u)$  for  $i \geq 2$ .

We set  $p_1 = 3$  for optimizing the 3-layer InAs structure, and  $p_1 = p_2 = 3$  for the 6-layer 3InAs+3Weyl structure presented in **Results and discussion**. For the 5-layer InAs and 8-layer 5InAs+3Weyl structures presented in **Additional results** below, we also set  $p_1 = 3$  and  $p_1 = p_2 = 3$ , respectively. These settings allow the InAs carrier concentrations and Weyl Fermi levels over the layers to conform to a polynomial of up to second order, see Fig. S1.

## Additional results

This section provides optimization results for 5-layer InAs and 8-layer 5InAs+3Weyl structures. Figure S2 shows the results for the 5-layer InAs structures obtained from five different trials of LCB and those of PI. Figure S3 plots the absorptivity and emissivity values over the considered range of incidence angles for the initial and final 5-layer InAs structures from the first and last iterations of the first LCB trial, as well as those for the state-of-the-art 10-layer InAs structure.<sup>7</sup> Figure S4 shows the optimization results for the 8-layer 5InAs+3Weyl structures. The absorptivity and emissivity values over the considered range of incidence angles for the best 8-layer 5InAs+3Weyl structures from LCB and PI are provided in Fig. S5.

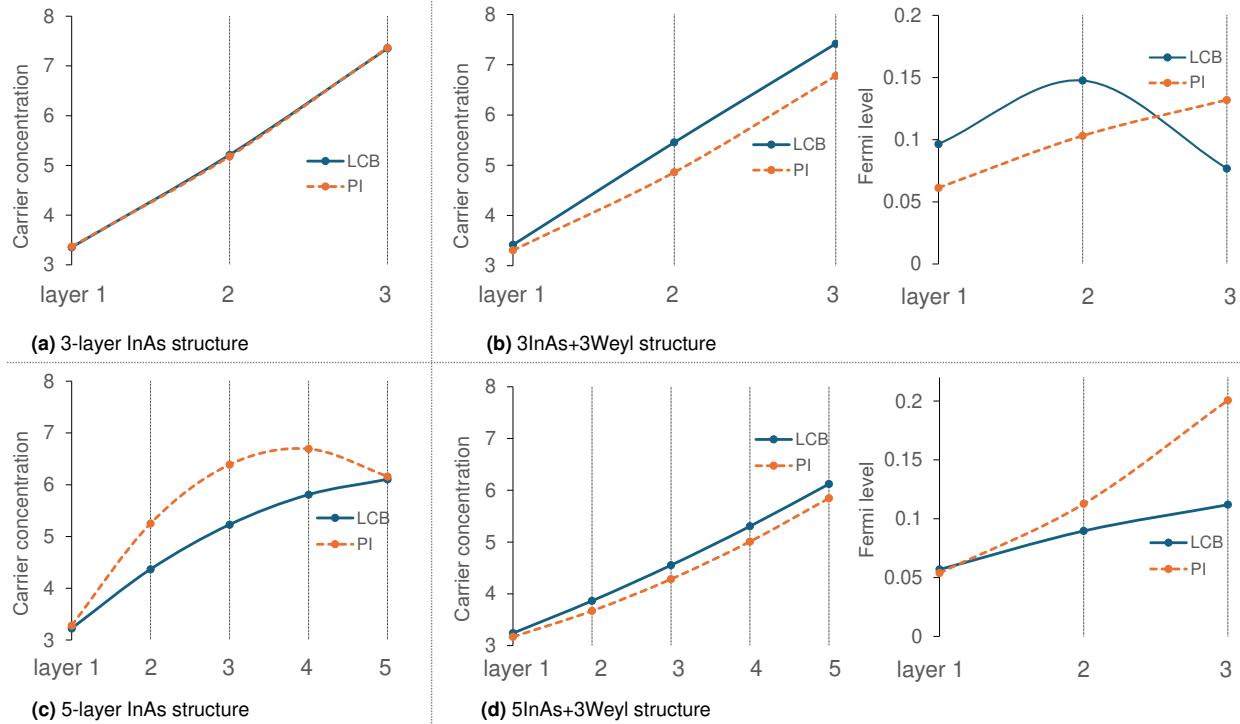


Figure S1: Profiles of InAs carrier concentrations and Weyl Fermi levels inside the optimal structures by LCB and PI.

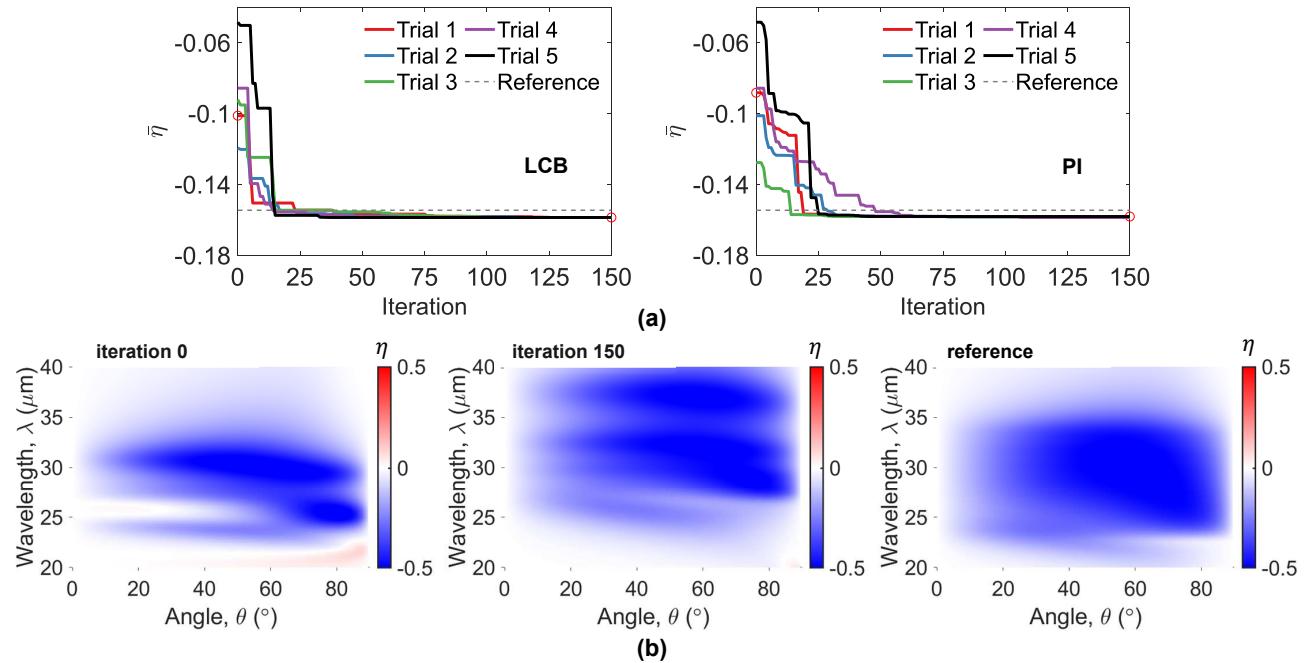


Figure S2: Optimization results for 5-layer InAs structure in comparison with the state-of-the-art 10-layer InAs structure (reference).<sup>7</sup> (a) Optimization histories from LCB and PI. (b) Comparison of contrast values of the initial and final structures from the first and last iterations of the first LCB trial, and the state-of-the-art 10-layer InAs structure.

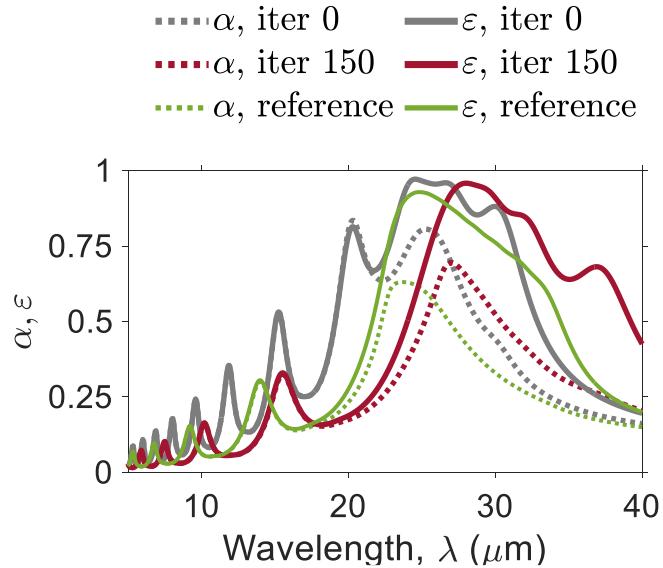


Figure S3: Absorptivity ( $\alpha$ ) and emissivity ( $\varepsilon$ ) of the initial and final 5-layer InAs structures from the first and last iterations of the first LCB trial, and the state-of-the-art 10-layer InAs structure (reference).<sup>7</sup>

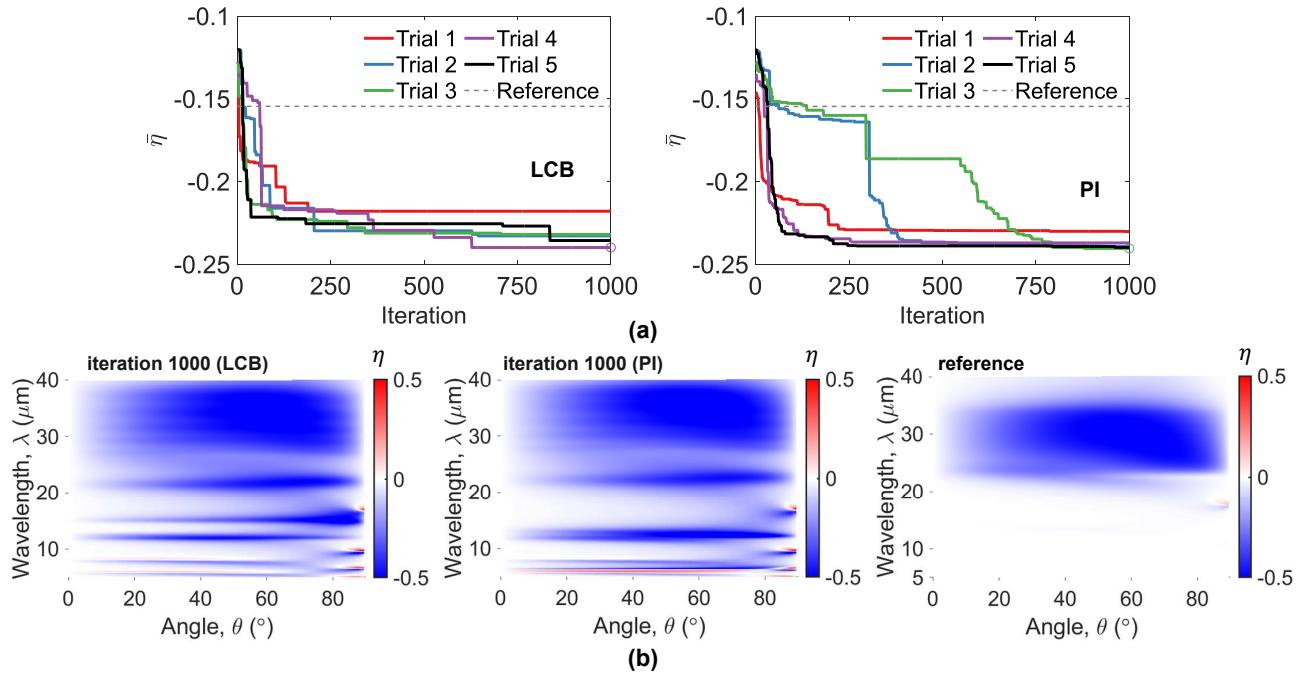


Figure S4: Optimization results for 8-layer 5InAs+3Weyl structure in comparison with the state-of-the-art 10-layer InAs structure (reference).<sup>7</sup> (a) Optimization histories from LCB and PI. (b) Comparison of contrast values of the best structures from LCB and PI, and the state-of-the-art 10-layer InAs structure.

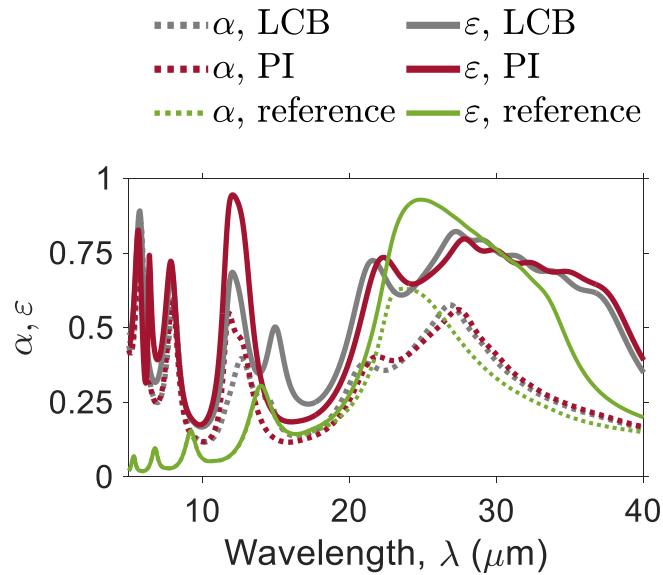


Figure S5: Absorptivity ( $\alpha$ ) and emissivity ( $\varepsilon$ ) of the best 8-layer 5InAs+3Weyl structures from LCB and PI, and the state-of-the-art 10-layer InAs structure (reference).<sup>7</sup>

## References

- (1) Rasmussen, C. E.; Williams, C. K. I. *Gaussian processes for machine learning*; The MIT Press: Massachusetts, USA, 2006.
- (2) Garnett, R. *Bayesian optimization*; Cambridge University Press: Cambridge, UK, 2023.
- (3) Srinivas, N.; Krause, A.; Kakade, S. M.; Seeger, M. Gaussian process optimization in the bandit setting: No regret and experimental design. Proceedings of the 27th International Conference on Machine Learning. 2010; pp 1015–1022.
- (4) Kushner, H. J. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal Basic Engineering* **1964**, *86*, 97–106.
- (5) Trefethen, L. N. *Approximation Theory and Approximation Practice, Extended Edition*; Other Titles in Applied Mathematics; Society for Industrial and Applied Mathematics: Philadelphia, PA, 2019; Vol. 164; First edition, 2013; Extended edition, 2019.
- (6) Hale, N.; Trefethen, L. N. New quadrature formulas from conformal maps. *SIAM Journal on Numerical Analysis* **2008**, *46*, 930–948.
- (7) Liu, M.; Xia, S.; Wan, W.; Qin, J.; Li, H.; Zhao, C.; Bi, L.; Qiu, C.-W. Broadband mid-infrared non-reciprocal absorption using magnetized gradient epsilon-near-zero thin films. *Nature Materials* **2023**, *22*, 1196–1202.