

Enhancing overall performance of thermophotovoltaics via deep reinforcement learning-based optimization



Cite as: J. Appl. Phys. 136, 023101 (2024); doi: 10.1063/5.0213211

Submitted: 9 April 2024 · Accepted: 13 June 2024 ·

Published Online: 8 July 2024



Shilv Yu,¹ , Zihe Chen,¹ , Wentao Liao,¹ Cheng Yuan,² , Bofeng Shang,^{3,a)} , and Run Hu^{1,4,a)}

AFFILIATIONS

¹School of Energy and Power Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

²Wuhan Fiberhome Fuhua Electric Co., Ltd, Wuhan 430074, China

³School of Physics and Microelectronics, Zhengzhou University, Zhengzhou 450001, China

⁴Department of Applied Physics, Kyung Hee University, 1732 Deogyong-daero, Giheung-gu, Yongin-Si, Gyeonggi-do 17104, Republic of Korea

^{a)}Authors to whom correspondence should be addressed: shangbofeng@zzu.edu.cn and hurun@hust.edu.cn

ABSTRACT

Thermophotovoltaic (TPV) systems can be used to harvest thermal energy for thermoelectric conversion with much improved efficiency and power density compared with traditional photovoltaic systems. As the key component, selective emitters (SEs) can re-emit tailored thermal radiation for better matching with the absorption band of TPV cells. However, current designs of the SEs heavily rely on empirical design templates, particularly the metal-insulator-metal (MIM) structure, and lack of considering the overall performance of TPV systems and optimization efficiency. Here, we utilized a deep reinforcement learning (DRL) method to perform a comprehensive design of a 2D square-pattern metamaterial SE, with simultaneous optimization of material selections and structural parameters. In the DRL method, only the database of refractory materials with gradient refraction indexes needs to be prepared in advance, and the whole design roadmap will automatically output the SE with optimal Figure-of-Merit (FoM) efficiently. The optimal SE is composed of a novel material combination of TiO₂, Si, and W substrate, with its thickness and structure precisely optimized. Its emissivity spectra match well with the external quantum efficiency curve of the GaSb cell. Consequently, the overall performance of TPV is significantly enhanced with an output power density of 5.78 W/cm², an energy conversion efficiency of 38.26%, and a corresponding FoM of 2.21, surpassing most existing designs. The underlying physics of optimal SE is explained by the coupling effect of multiple resonance modes. This work advances the practical application potential of TPV systems and paves the way for addressing other multi-physics optimization problems and metamaterial designs.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>). <https://doi.org/10.1063/5.0213211>

INTRODUCTION

Thermophotovoltaic (TPV) systems, with the aid of selective emitters (SEs), convert various energy sources, including solar radiation, fuels, and industrial waste heat, into thermal radiation that matches the bandgap of TPV cells, so as to be capable for outputting electricity with higher conversion efficiency and power density.^{1–3} TPV systems are anticipated to emerge as the subsequent advanced power generation technology and has promising applications in thermal energy recycling, hybrid electric vehicles, deep space exploration, and other pertinent fields.^{4–6}

A schematic of a typical TPV system is illustrated in Fig. 1(a), which consists of three components, including an absorber, an emitter, and a TPV cell. The absorber is responsible for capturing solar energy, fuel, or other energy sources to heat the emitter. Since the TPV cells can only convert the photons above the bandgap for photoelectric conversion, and the photons below the bandgap are wasted, selective emitters, as the core component of TPV systems, are meticulously engineered to regulate the emissivity spectra to align with the bandgap of TPV cells. Therefore, the finer optimization design of SEs has become the research focus to enhance the performance of TPV systems.⁷ Owing to the rapid development of

14 July 2024 13:42:37

nanophotonics, a variety of structural forms of SEs have been developed for TPV systems, such as multilayers,^{1,8,9} photonic crystals,¹⁰ metal-insulator-metal (MIM) structure,^{11–13} cavities,¹⁴ and metasurfaces.^{15–17} Among them, MIM and metasurfaces, due to their nature of resonant, can generate sharper resonance bands and higher emissivity compared with multilayers and, thus, have the potential to maximize the performance of TPV systems. In addition, they demonstrate greater design flexibility and more stable in structure at high temperature and are, therefore, chosen by a majority of researchers for the TPV emitters. However, there are two issues that need addressing. For one thing, conventional manual design heavily relies on past templates and experience, resulting in inefficient design and poor performance.¹⁸ Although computational methods, including stochastic gradient descent, Bayesian optimization, and genetic algorithm, are utilized to enhance the performance of SEs in the existing research studies, most of them still primarily employ fixed material combinations and solely optimize structural parameters.^{19–21} The potential for improved material combinations and structures remains unexplored, limiting further improvements in the performance of SEs and TPV systems. In addition, due to the time-consuming simulations of MIM structures or metasurfaces in the design process, the efficiency of evolutionary algorithms that were originally not high becomes even lower in optimizing such structures. For another thing, previous works have shown that energy conversion efficiency and output power density are obviously competitive but equally important as two evaluation metrics for TPV systems.²² Most of the existing works only consider the simple objective of emissivity when designing SEs, without accurately modeling the TPV systems. As a result, the optimized SEs fail to accurately match the system and neglect the comprehensive enhancement of energy conversion efficiency and output power density. Therefore, it is still a strong demand to further explore more material combinations and achieve more efficient optimization of structural parameters of the SEs based on precise TPV system modeling, in order to enhance the comprehensive performance of the TPV systems.

In this work, a deep reinforcement learning (DRL) method is adopted to design a two-dimensional square metasurface SE, which can autonomously select the most appropriate materials from a self-built material library and efficiently optimize the structural parameters. The design process is guided by the comprehensive indicators of the precisely modeled TPV system. The resulting SE is composed of a completely new material combination and achieves a good match between its emissivity spectrum and the external quantum efficiency (EQE) of the GaSb cell. Equipped with the designed SE, the system achieves significant performance enhancement, yielding an energy conversion efficiency of up to 38.26%, coupled with an output power density of 5.78 W/cm².

METHOD AND MODELING

Deep reinforcement learning method for design and optimizing

DRL, which is a reinforcement learning based on deep learning, differs significantly from supervised or unsupervised deep learning.^{23,24} The neural networks serve as agents that interact with an environment, updating its state based on actions taken, rather

than learning direct data mappings between input–output pairs, where the state of the environment and the actions are the input and output of the neural network, respectively. What guides the agent to make the appropriate decision is the expected accumulated value of the Figure of Merit (FoM), namely, the predicted Q-value by the agent. When the agent modifies the current state, the resulting state must be evaluated to determine its FoM value, which is then fed back to the agent for calculating the true Q-value. The mean square error between the true Q-value and the predicted Q-value serves as the loss function for backpropagation in a neural network to update the weights of neurons, as follows:²⁵

$$\text{loss} = \text{MSE}[Q_{\text{true}}(s_{t+1}, a; w) - Q_{\text{pre}}(s, a; w)], \quad (1)$$

where s and s_{t+1} are the current and updated state, respectively. a is the action and w is the weight of the neural network. Subsequently, the updated state is utilized as the input for the next iteration. During each iteration, the input state, the output decision, and the resulting FoM value are recorded in a buffer as the datasets for training the neural network, making the agent to learn how to make more appropriate decisions to maximize the expected cumulative FoM.²⁶

To be specific, the environment here is regarded as the whole optimization process of the SEs, with the design parameters of the SEs representing the state of the environment. As shown in Fig. 1(b), the SE is set to a periodic 2D metamaterial with a square pattern on the top. To differentiate from the three-layer classic MIM structure and further expand the design space, we added one more layer into the middle layer between the top square pattern and the substrate. The material types for the top pattern and middle layers are used as design parameters and selected by DRL model subsequently, while the substrate material is fixed as tungsten (W) for support, thermal conduction, and resonant excitation due to its good mechanical, thermal, and optical properties.⁷ Additionally, there are four more design parameters that need to be optimized by DRL, including three corresponding layer thicknesses and the fill rate of the top pattern. Consequently, there are seven design parameters in total, and DRL performs the materials selection and the structural design simultaneously during the optimization. Three candidate materials were selected by the DRL model from a self-built material library, which contains five commonly used high-temperature resistant materials, including SiO₂, W, Si, TiO₂, and HfO₂. These five materials were then encoded as No. 1–5 so that they can be accepted by the neural network. Their optical properties are referred to Palik's work²⁷ and other research works.^{28,29} The thicknesses of the top pattern and two intermediate layers are restrained in the range of 20–200 nm and are designed with the precision of 10 nm. The fill rate of the top pattern (the ratio of the side length to the period) is varied from 0.1 to 0.9, with a step size of 0.1. Consequently, the design space comprises about 7.7×10^6 candidates, which renders manual design impractical and presents significant challenges to conventional evolution-based methods.^{23,30}

Subsequently, the seven design parameters are encoded and fed into the neural network, as shown in Fig. 1(c). The neural network, which consists of three fully connected layers with the neuron number of 24, 48, and 24, respectively, will output an action to update the current design parameters, such as changing

14 JULY 2024 13:42:37

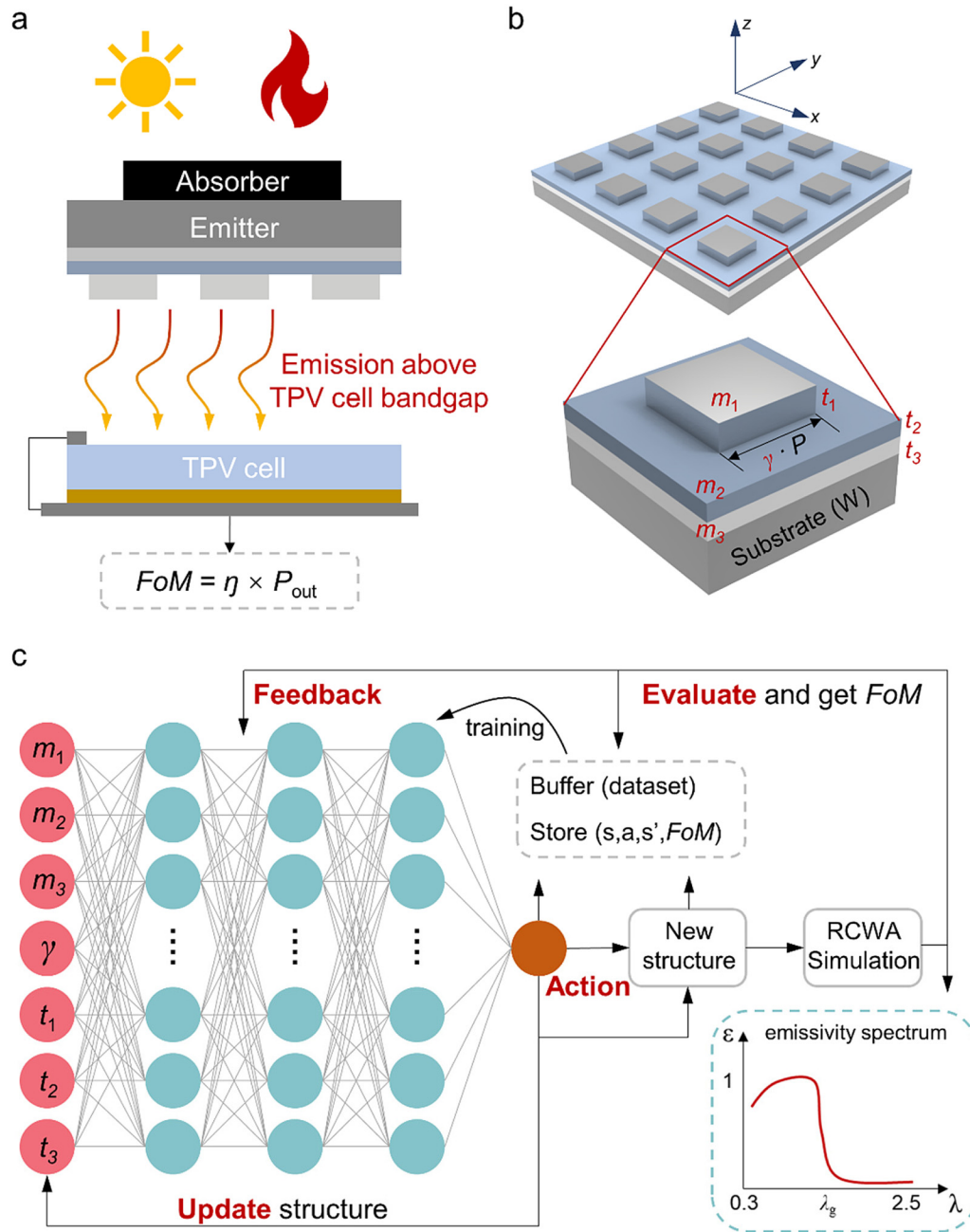


FIG. 1. (a) Schematic diagram of a TPV system. (b) Schematic diagram of the structure of a two-dimensional square pattern emitter; seven design parameters are presented. (c) Roadmap of designing selective emitters using deep reinforcement learning.

the materials or adjusting the thicknesses. The correspondence between the action numbers and the policies for updating states can be referred to Table I. Then, the updated design parameters are converted to a physical model, and their emissivity spectra are obtained utilizing a rigorous coupled wave analysis (RCWA)

algorithm.³¹ To evaluate the performance of the SE and TPV systems, the simulated emissivity spectrum is combined with the accurately modeled TPV system to obtain the energy conversion efficiency and the output power density of the system. The product of two metrics is defined as the FoM, which is as follows:

TABLE I. Policies corresponding to each action.

Action no.	Policy
0	Decrease the material ID of m_1 by 1 (min 1)
1	Increase the material ID of m_1 by 1 (max 5)
2	Decrease the material ID of m_2 by 1 (min 1)
3	Increase the material ID of m_2 by 1 (max 5)
4	Decrease the material ID of m_3 by 1 (min 1)
5	Increase the material ID of m_3 by 1 (max 5)
6	Decrease γ by 0.1 (min 0.1)
7	Increase γ by 0.1 (max 0.9)
8	Decrease t_1 by 10 (min 20)
9	Increase t_1 by 10 (max 200)
10	Decrease t_2 by 10 (min 20)
11	Increase t_2 by 10 (max 200)
12	Decrease t_3 by 10 (min 20)
13	Increase t_3 by 10 (max 200)

$$\text{FoM} = \eta \times P_{\text{out}}. \quad (2)$$

The FoM is then fed back to the neural network, and the next iteration is started with the updated design parameters used as inputs for the neural network. Through iterative processes, the neural network, acting as the agent for continuously designing SE, gradually learns how to make an appropriate action to update the SE's design parameters, enhance the emitter's performance, and ultimately improve the overall system performance. Eventually, the design parameters converge to the optimal state.

Accurate modeling of the TPV system

Since the ultimate goal of designing a SE is to improve the performance of the TPV system, an accurately modeled TPV system plays a crucial role in guiding the optimization process of the emitter. Specifically, the accurate calculation of energy conversion efficiency and output power density is of utmost importance. We focus on the energy radiation of the emitter and the photoelectric conversion process of the TPV cell, while disregarding energy losses resulting from convection and conduction. Here, we choose GaSb, a low bandgap semiconductor cell, as the photoelectric conversion element of the TPV system, which has a bandgap of 0.726 eV with the corresponding bandgap wavelength of 1.708 μm . After the emitter is heated by various heat sources, the energy it emits to the TPV cell can be represented as P_{in} for the GaSb cell, which is calculated as follows:

$$P_{\text{in}} = \int_{\lambda_1}^{\lambda_2} \varepsilon(\lambda) I_{\text{BB}}(\lambda, T_e) d\lambda, \quad (3)$$

where λ_1 and λ_2 are the lower and upper boundaries of the studied wavelength band, taking 0.3 and 5 μm , respectively. $\varepsilon(\lambda)$ is the emissivity spectrum of the SE. $I_{\text{BB}}(\lambda, T_e)$ is the spectral radiance of the blackbody at temperature T_e and wavelength λ , which is given

by Planck's law,

$$I_{\text{BB}} = \frac{2hc^2}{\lambda^5} \frac{1}{\exp(hc/\lambda k_B T_e) - 1}, \quad (4)$$

where h is Planck's constant, k_B is the Boltzmann constant, and c is the speed of the light. In a subsequent design process, T_e is set to 1700 K, which is the optimal adaptation temperature of the GaSb cell according to the bandgap and the Wien displacement law. For a TPV cell, the calculation for the maximum output power density can be denoted by¹

$$P_{\text{out}} = V_{\text{oc}} J_{\text{sc}} \phi_{\text{FF}}, \quad (5)$$

where V_{oc} is the open-circuit voltage, J_{sc} is the short-circuit current density, and ϕ_{FF} is the fill factor. However, the TPV cells are not able to fully utilize all photons larger than the bandgap for photoelectric conversion. Some photons are still lost due to reflection, scattering, and absorption. The introduction of external quantum efficiency (EQE) effectively measures the photovoltaic conversion ability of batteries. Considering this, the J_{sc} can be calculated as³²

$$J_{\text{sc}} = e \int_{\lambda_1}^{\lambda_g} \frac{I_{\text{BB}} \varepsilon(\lambda) \eta_{\text{EQE}}(\lambda) d\lambda}{hc/\lambda}, \quad (6)$$

where e is the elementary charge and λ_g is the bandgap wavelength of the TPV cell. To obtain the V_{oc} in Eq. (4), when considering a relatively ideal TPV cell, the relationship between the current and voltage is given by³³

$$J(V) = J_{\text{sc}} - J_0 \left[\exp\left(\frac{eV}{k_B T_c}\right) - 1 \right], \quad (7)$$

where T_c is the temperature of the TPV cell, which is set to a suitable operating temperature of 300 K.^{2,4} J_0 is the sum of the reverse-saturation current of the P-N junction, which can be given by³⁴

$$J_0 = \beta T_c^3 \exp\left(-\frac{E_g}{k_B T_c}\right), \quad (8)$$

where β is a constant that combines the dimensions, the doping, and the material parameters of a TPV cell. For a GaSb cell, it equals $3.165 \times 10^{-4} \exp(2.5 \times E_g)$. E_g is the bandgap of the cell, which is 0.726 eV (1.708 μm) for GaSb. When solving Eq. (5) for $J(V) = 0$, we can obtain V_{oc} , as follows:

$$V_{\text{oc}} = \frac{k_B T_c}{e} \ln\left(\frac{J_{\text{sc}}}{J_0} + 1\right). \quad (9)$$

The fill factor in Eq. (4) can be expressed as³⁵

$$\phi_{\text{FF}} = \frac{V^* - \ln(V^* - E_g)}{V^* + 1}, \quad (10)$$

where V^* is the normalized open-circuit voltage, which can be

obtained as follows:³⁵

$$V^* = \frac{e}{k_B T_c} V_{oc}. \quad (11)$$

After the above calculation, we can finally obtain the expression of the energy conversion efficiency of the TPV system as follows:

$$\eta = \frac{P_{out}}{P_{in}}. \quad (12)$$

RESULTS AND DISCUSSION

Performance of the optimized SE and TPV system

By utilizing the aforementioned DRL optimization model in conjunction with the accurately modeled TPV system, after 1000 iterations, the FoM is finally converged. Consequently, we obtained the optimal SE that maximizes the comprehensive performance of the TPV system, as shown in Fig. 2(a). Interestingly, the materials of the top square pattern and the two intermediate layers are

selected by DRL automatically as TiO₂, Si, and W, respectively. Since the material of the substrate is also W, the same stacked materials can be considered a single material, resulting in a three-layer optimal structure. Although the optimal SE is structurally similar to MIM, the combination of the TiO₂ top pattern, the Si intermediate layer, and the W substrate has not been previously reported. This can be attributed to the strong capability of the DRL optimization model, which accurately identifies optimal material combinations by selecting the most suitable materials from a wide range of options. The thickness of the TiO₂ and Si layers is precisely determined by DRL to 110 and 40 nm, respectively, and the fill ratio of the top pattern is 0.7. Figure 2(a) also presents the emissivity spectrum of the optimal emitter, which exhibits quite high emissivity in the range above the bandgap of the GaSb cell and a sharp decline in emissivity below the bandgap. To validate the reliability of the optimization results, we employed the finite-difference time-domain (FDTD) method to simulate the emissivity of the optimized SE. The obtained results exhibited consistency with those obtained from the RCWA algorithm. Equipped with the optimal SE, the corresponding TPV system demonstrates an output power density of 5.78 W/cm² and a high energy conversion

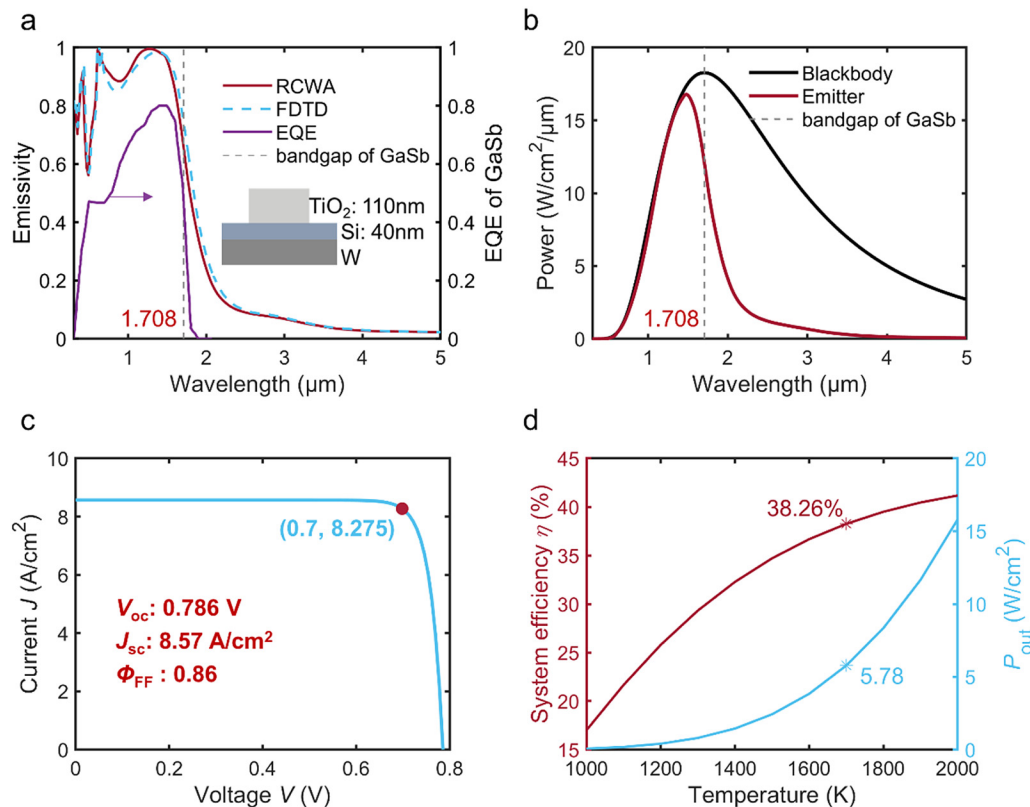


FIG. 2. (a) Emissivity spectra of the optimal SE simulated by RCWA (red) and FDTD (dashed blue). EQE of a GaSb cell is represented by a purple line. The inset is the schematic diagram of the optimal SE. The bandgap of the GaSb cell is represented in dashed gray. The schematic of optimal SE is shown in the inset. (b) Radiation energy spectra of the blackbody and the optimal SE. (c) J - V curve of the TPV system equipped with optimal SE. (d) System efficiency and output power density as a function of SE temperature.

14 July 2024 13:42:37

efficiency of 38.26%, resulting in the FoM of 2.21. To elucidate the reasons behind the improved performance of the TPV system due to the optimized SE, the EQE of the GaSb cell as a function of wavelength is also plotted in the figure.³⁶ It can be obviously seen that the emissivity spectrum of the optimized SE exhibits a close resemblance to the trend of the EQE curve, peaking at slightly shorter wavelengths than the cell's bandgap wavelength. This high matching effectively suppresses emission below bandgap energy and minimizes energy loss, resulting in enhancing the conversion efficiency. Additionally, the broad emission spectrum with high emissivity guarantees a remarkable output power density.

To provide a clearer depiction of the SE's suppression of thermal emission below the bandgap of the GaSb cell, Fig. 2(b) displays the radiation energy spectrum of the blackbody and the designed selective emitter at 1700 K. It is evident that only a small amount of energy below the bandgap is emitted, while the majority of energy above the bandgap is emitted to the GaSb cell for photoelectric conversion. Furthermore, the sudden decrease in emissivity at $0.5\text{ }\mu\text{m}$ does not significantly affect the radiation energy, given that the energy density at that wavelength is already negligible.

Figure 2(c) illustrates the current-voltage curve of the TPV system, indicating that at a voltage of 0.7 V and a current of 8.275 A/cm^2 , the system achieves a maximum power output of 5.78 W/cm^2 . Additionally, the corresponding open-circuit voltage, the short-circuit current density, and the fill factor of the system are 0.786 V, 8.57 A/cm^2 , and 0.86, respectively. This high-power density positions the TPV system as a competitive contender among the various power generation methods available today.

Since the temperature of the SE is fixed to 1700 K during the optimization, to show the performance of the optimized TPV system at different operating temperatures, Fig. 2(d) presents the system efficiency and the output power density as a function of the temperature of SE. It can be observed that both the system efficiency and the output power density increase with the temperature. This can be easily explained by Planck's law. As the temperature increases, the radiation peak shifts to blue, causing an increase in energy greater than the bandgap. As a result, the energy conversion

efficiency of the system is improved, and the output power density also increases. The performance of the system at 1700 K is marked with an asterisk, indicating superior overall performance.

Additionally, for an excellent performance TPV emitter, the angular independence of emission is another crucial feature. Consequently, Fig. 3(a) presents the emissivity spectra of the optimal SE at different incident angles. It is evident that the emissivity demonstrates good angular independence under the angles of 70° . It can also be observed that the emissivity around $0.7\text{ }\mu\text{m}$ slightly decreases as the incident angle exceeds 60° . Moreover, emission above the bandgap energy experiences a significant reduction when the incident angle exceeds 80° . To further the impact of the incident angle on the performance of the system, the energy conversion efficiency, the output energy density, and the FoM as a function of incident angle are shown in Fig. 3(b). With the incident angle increasing, the energy conversion efficiency shows a slight improvement due to the reduction in the incident energy received by the cell exceeds the reduction in the output energy. More specifically, the reduction in the incident energy stems from both inside and outside the cell bandgap, while the decrease in the output energy only occurs within the bandgap. Additionally, the system maintains excellent performance within the range of 70° according to the FoM curve. Although a significant decrease is observed for angles greater than 70° , it is anticipated that the impact of reduction in large-angle emission on system performance will be weakened due to the typical parallel arrangement of the SEs and TPV cells.

Optimization efficiency and optimality verification

For the optimal design of the TPV emitter, it is crucial to achieve high design efficiency while maximizing its performance. Therefore, we undertake a quantitative assessment of the efficiency of designing SEs when adopting the DRL method. To illustrate this evaluation, Fig. 4(a) plots the FoM of the SE against the percentage of calculated structures. It can be seen that only 0.692% candidates were calculated to find a SE with a high FoM up to 2.18, which is about 98.6% of the maximum FoM. Likewise, only 3.19%

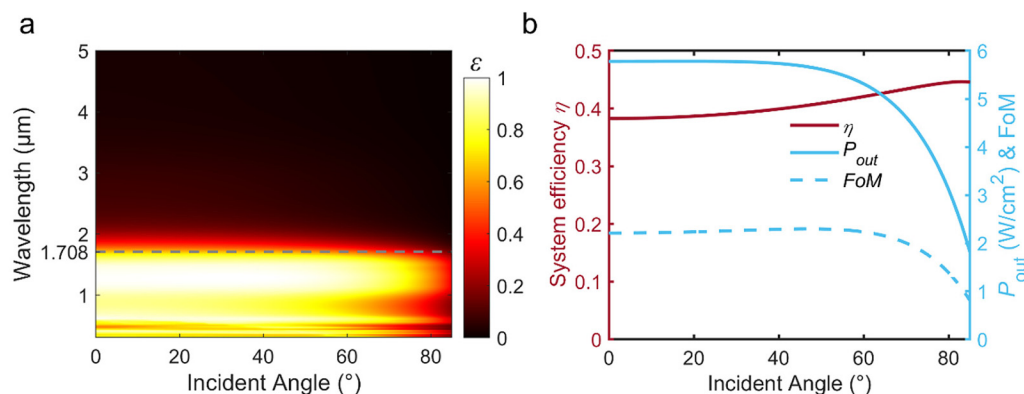


FIG. 3. (a) Emissivity of the optimal SE as a function of wavelength and incident angle. (b) System energy conversion efficiency, output power density, and FoM as a function of incident angle.

14 JULY 2024 13:42:37

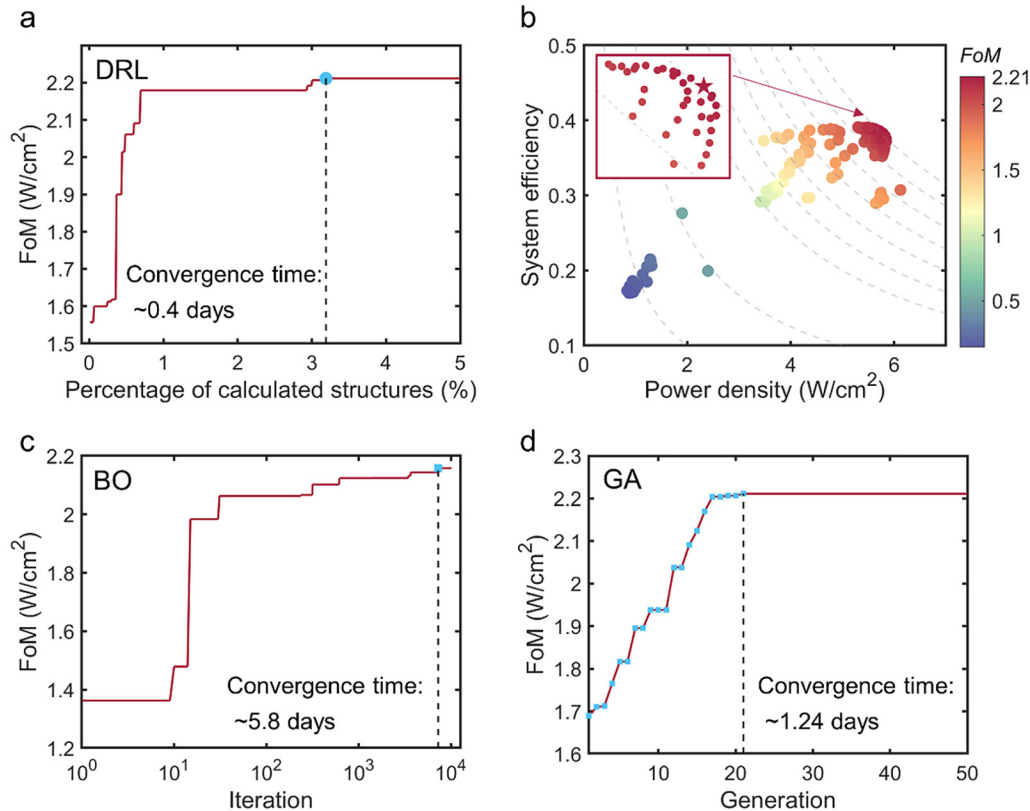


FIG. 4. (a) Tracing of FoM vs percentage of calculated structures when using a DRL method. The convergence time is annotated inside. (b) System efficiency and output power density of the intermediate structure during the optimization process. The color bar denotes the value of FoM. (c) Tracing of FoM vs iterations when using BO. (d) Tracing of FoM vs generations when using GA.

14 July 2024 13:42:37

candidates were calculated to find a structure with the maximum FoM. Such tracings vividly demonstrate the high efficiency and effectiveness of the DRL method in finding the optimal structure when simultaneously considering the material selection and structure optimization. In addition, the intermediate structure of the optimization process is plotted in Fig. 4(b). It can be observed that the FoM of most structures is concentrated above 1.5. Also due to the extremely fast convergence speed, the number of the intermediate structures is not large. Additionally, it can be seen that the structure with the maximum FoM, which can optimize the overall performance of the system, is not optimal in terms of system efficiency or output power density unilateral contrast. Therefore, this also verifies the competitive relationship between the output power density and the system efficiency. Of course, structures with high conversion efficiency or power density can also be further screened from the intermediate structures in the optimization process.

In order to demonstrate the efficiency of the DRL method at the methodological level, we adopted two existing commonly used optimization algorithms for comparison. Given that our design parameters include different types: material species and structural parameters, gradient-based optimization methods are not applicable. Therefore, Gaussian process-based Bayesian Optimization (BO)³⁷

and evolution-based Genetic algorithm (GA)³⁸ were utilized, and the execution details can be found in [supplementary material](#) Note 1 and 2. Since the Gaussian process of BO is not well suited to the optimization problem of huge parameter space, the performance of the optimal SE is still worse than that of obtained by DRL in the case of consuming more time, as shown in Fig. 4(c). While the set number of iteration steps of BO may not be insufficient for evaluating convergence, it has been able to assess the pros and cons of BO and DRL, considering the aspect of time. The optimal SE obtained by GA is consistent with the structure obtained by DRL. However, the genetic algorithm must simulate all the individuals in the initial population and the offspring population obtained in each iteration, while the optimization process of DRL can be regarded as updating and iterating only for the optimal individual. In the case that the simulation of the 2D structure is quite time-consuming, it results in the GA being less efficient than DRL, as illustrated in Fig. 4(d).

Additionally, to further show the optimality of the TiO₂-Si-W structure compared with MIM structures, we performed another two rounds of optimization by DRL, fixing the materials as W-TiO₂-W and W-Si-W, respectively. The optimization parameters include the fill rate and thickness of the top W pattern and the thickness of the middle layer. Figures 5(a) and 5(b) display the emissivity and

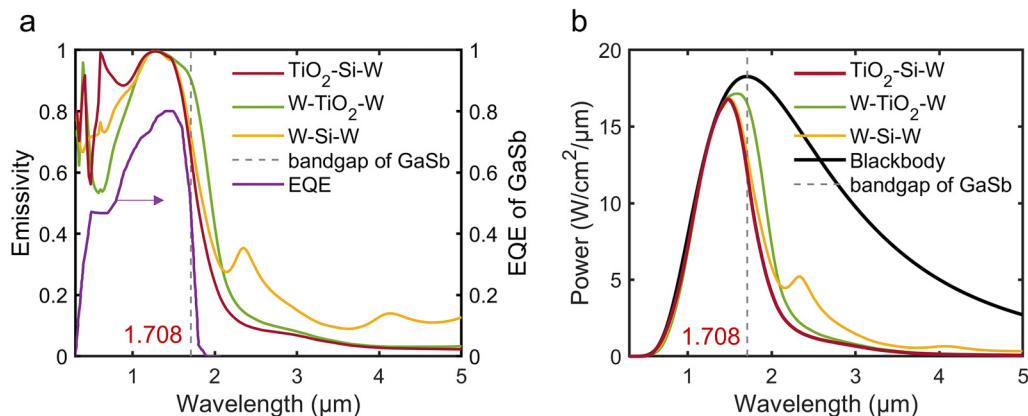


FIG. 5. (a) Emissivity spectra of three optimized SEs, including the optimal structure (TiO₂-Si-W) and two MIM structures (W-TiO₂-W and W-Si-W). (b) Radiation energy spectra of three optimized SEs.

radiation energy spectra of the three optimized structures. It can be obviously seen that the TiO₂-Si-W structure exhibits the highest radiation suppression below the bandgap, while maintaining high emission above the bandgap, resulting in the highest system efficiency, as illustrated in Table II. With the emissivity of the W-TiO₂-W structure slightly higher than that of TiO₂-Si-W at around 1.7 μm, it leads to a slightly higher output power density compared to TiO₂-Si-W. However, the W-Si-W structure exhibits less remarkable performance. In addition, to demonstrate the optimality of the top square pattern, we replaced the pattern of optimal SE with a cylinder to serve as another control group. The overall performance of SE with a square pattern is slightly better than that of SE with a cylinder pattern, as shown in Fig. S1 and Table S1 of the [supplementary material](#). In conclusion, the TiO₂-Si-W emitter, based on the completely new material combination, exhibits superior radiation performance compared to the MIM structure, thus supporting the TPV system with the optimal comprehensive performance.

Additionally, we explored the effects of different thicknesses of the top and middle layers on the performance under the condition of fixing their materials, as shown in Figs. S2 and S3 of the [supplementary material](#), which demonstrate the structural optimality when considering the overall performance of the TPV system.

Mechanism of the optimal SE

The mechanism behind the optimal SE to show such excellent performance requires further explanation. From Fig. 2(a), it can be

TABLE II. Performance index of the corresponding TPV system of three kinds of SEs.

Structure	System efficiency (%)	Output power density (W/cm ²)	FoM
W-TiO ₂ -W	34.11	5.98	2.04
W-Si-W	31.47	5.80	1.83
TiO ₂ -Si-W	38.26	5.78	2.21

observed that the SE generates two emission peaks above the bandgap energy, similar to the classical MIM structure emitter. Therefore, the electromagnetic fields inside the emitter at two emission peaks are plotted in Fig. 6, with the white line outlining the SE's structure. In the MIM structure, the two emission peaks can be explained by surface plasmon polariton (SPP) and magnetic polariton (MP) resonance, respectively.^{17,39} However, our structure replaces the material of the top pattern from metal to TiO₂, resulting in different mechanisms for peak generation and making it difficult to determine the triggering resonance. At the first emission peak (0.606 μm), the magnetic field is enhanced at the W-Si junction and the bottom of TiO₂, while the electric field is enhanced in Si and the top of TiO₂, which can be explained by the coupling effect of MP resonance and Mie resonance. However, at the second emission peak (1.39 μm), the magnetic field is strongly enhanced at the interface between W and Si, and the electric field is enhanced throughout TiO₂, resulting in a stronger emission. This can be attributed to the SPP resonance and the stronger Mie resonance. Consequently, both emission peaks originate from the combined action of multiple resonance modes. The DRL method promotes the generation of structures with elusive mechanisms, breaking through the fixed design template of MIM structures.

Finally, in order to demonstrate the superior performance of our designed SE, it is compared with the exiting related works. In addition, the multilayer structure with the filling rate of the top pattern set to 1 is also utilized as one of the comparison groups. The comparison results are shown in Fig. 7, where the x axis represents the output power density of the TPV system with the corresponding emitter, the y axis represents the energy conversion efficiency of the system, and the color depth of the scatterplot indicates the product of two indices, namely, FoM. Additionally, several FoM's climbing lines (dashed lines) are also plotted in the figure for an intuitive comparison. It can be obviously seen that the TPV system with the emitter optimized here exhibits the optimal comprehensive performance, while other works do not consider the improvement of the system's overall performance, leading to a blind pursuit of enhancing one index at the expense of another.

14 JULY 2024 13:42:37

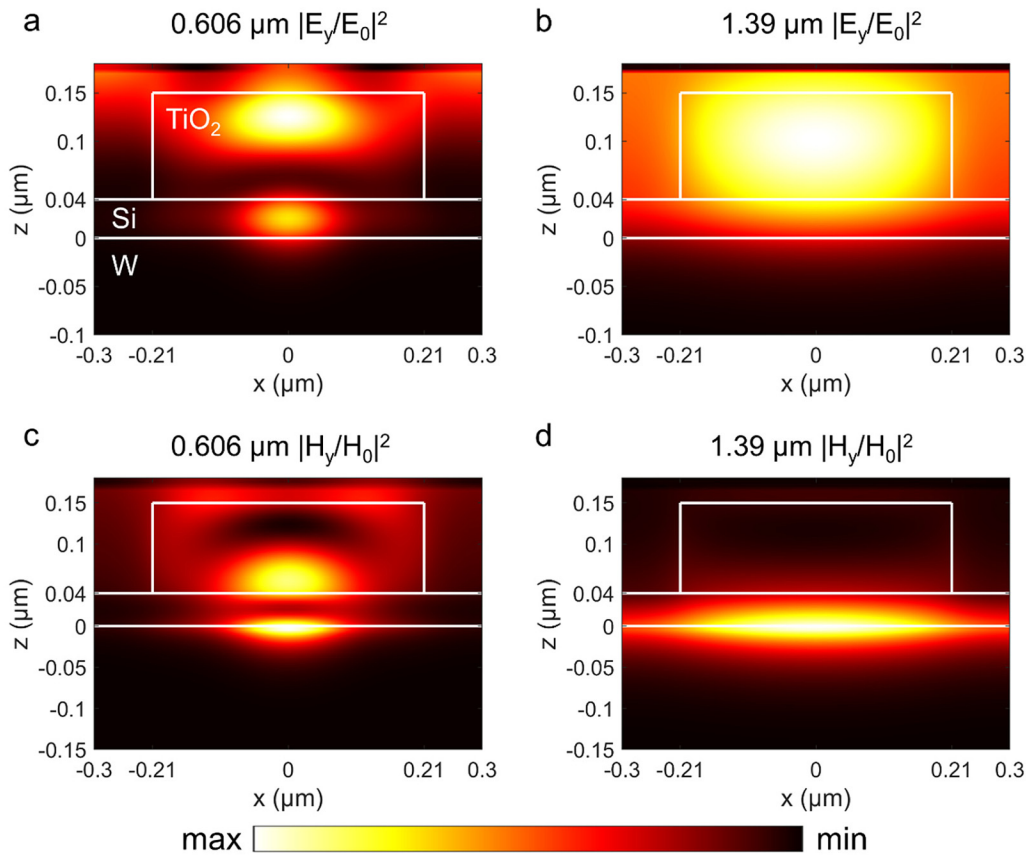


FIG. 6. Electromagnetic fields inside the optimal SE at two emission peaks. (a) Electric field at $0.606\ \mu\text{m}$. (b) Electric field at $1.39\ \mu\text{m}$. (c) Magnetic field at $0.606\ \mu\text{m}$. (d) Magnetic field at $1.39\ \mu\text{m}$.

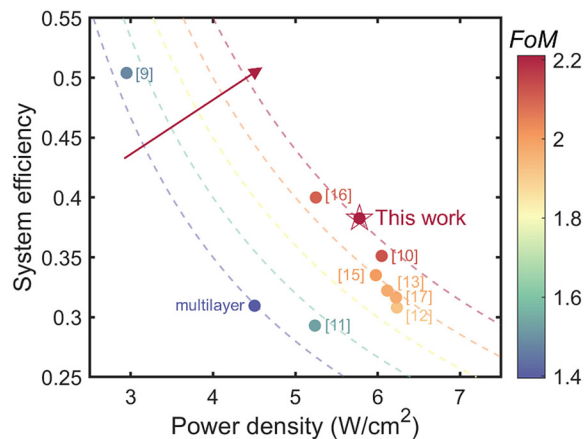


FIG. 7. Comparison of the performance of SEs between this work and existing works. The color bar denotes the value of FoM. The 1D multilayer includes Refs. 9 and 10. MIM structures or metasurfaces include Refs. 11–13 and 15–17.

CONCLUSION

In summary, we achieved a significant improvement in the overall performance of the TPV system. Specifically, the deep reinforcement learning approach was utilized to perform a comprehensive and efficient optimization design of SE in a huge design parameter space, including material selection and structure optimization. The materials of SEs were autonomously selected by an artificial neural network from a self-built material library, while the structural parameters were precisely optimized. The optimal SE consists of a 110 nm TiO_2 top square pattern, a 40 nm Si intermediate layer, and the W substrate, which is a completely new structural combination that breaks through the design template of a previous MIM structure. Additionally, the optimal SE exhibits excellent radiation properties, with high emissivity above the bandgap of the GaSb cell and significantly suppressed emission within the out-of-band. Through the accurately modeled TPV system, the energy conversion efficiency is calculated as 38.26%, the output power density is $5.78\ \text{W}/\text{cm}^2$, and the corresponding FoM is 2.21, which exceeds most existing works. The mechanism behind the excellent radiation characteristics of the optimal SE

is explained by the electromagnetic field, which can be interpreted as multiple resonant modes. Overall, this work propels advancements in the TPV and thermal radiation regulation and provides innovative ideas for the optimal design of metasurfaces and metamaterials.

SUPPLEMENTARY MATERIAL

The [supplementary material](#) contains the execution details of the Bayesian optimization and genetic algorithm and additional supporting tables and figures for demonstrating the optimality of designed SE.

ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support by the National Natural Science Foundation of China (Nos. 52211540005 and 52076087), the Science and Technology Program of Hubei Province (No. 2023AFA072), the Open Project Program of Wuhan National Laboratory for Optoelectronics (No. 2021WNLOK004), and the Wuhan Knowledge Innovation Shuguang Program.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Shilv Yu: Conceptualization (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal). **Zihe Chen:** Data curation (equal); Validation (equal). **Wentao Liao:** Validation (equal). **Cheng Yuan:** Validation (equal). **Bofeng Shang:** Supervision (equal); Writing – review & editing (equal). **Run Hu:** Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

REFERENCES

- ¹A. Lenert, D. M. Bierman, Y. Nam, W. R. Chan, I. Celanović, M. Soljačić, and E. N. Wang, “A nanophotonic solar thermophotovoltaic device,” *Nat. Nanotechnol.* **9**(2), 126–130 (2014).
- ²A. LaPotin *et al.* “Thermophotovoltaic efficiency of 40%,” *Nature* **604**, 287–291 (2022).
- ³A. Datas, A. López-Ceballos, E. López, A. Ramos, and C. del Cañizo, “Latent heat thermophotovoltaic batteries,” *Joule* **6**(2), 418–443 (2022).
- ⁴T. Burger, C. Sempere, B. Roy-Layinde, and A. Lenert, “Present efficiencies and future opportunities in thermophotovoltaics,” *Joule* **4**(8), 1660–1680 (2020).
- ⁵Y. Wang, H. Liu, and J. Zhu, “Solar thermophotovoltaics: Progress, challenges, and opportunities,” *APL Mater.* **7**(8), 080906 (2019).
- ⁶Z. Omair, G. Scranton, L. M. Pazos-Outón, T. P. Xiao, M. A. Steiner, V. Ganapati, P. F. Peterson, J. Holzrichter, H. Atwater, and E. Yablonovitch, “Ultraefficient thermophotovoltaic power conversion by band-edge spectral filtering,” *Proc. Natl. Acad. Sci. U.S.A.* **116**(31), 15356–15361 (2019).
- ⁷Z. Wang, D. Kortge, Z. He, J. Song, J. Zhu, C. Lee, H. Wang, and P. Bermel, “Selective emitter materials and designs for high-temperature thermophotovoltaic applications,” *Sol. Energy Mater. Sol. Cells* **238**, 111554 (2022).
- ⁸P. N. Dyachenko, S. Molesky, A. Y. Petrov, M. Störmer, T. Krekeler, S. Lang, M. Ritter, Z. Jacob, and M. Eich, “Controlling thermal emission with refractory epsilon-near-zero metamaterials via topological transitions,” *Nat. Commun.* **7**(1), 11809 (2016).
- ⁹Q. Wang, Z. Huang, J. Li, G.-Y. Huang, D. Wang, H. Zhang, J. Guo, M. Ding, J. Chen, Z. Zhang, Z. Rui, W. Shang, J.-Y. Xu, J. Zhang, J. Shiomi, T. Fu, T. Deng, S. G. Johnson, H. Xu, and K. Cui, “Module-level polaritonic thermophotovoltaic emitters via hierarchical sequential learning,” *Nano Lett.* **23**(4), 1144–1151 (2023).
- ¹⁰W. Zhang, B. Wang, and C. Zhao, “Selective thermophotovoltaic emitter with aperiodic multilayer structures designed by machine learning,” *ACS Appl. Energy Mater.* **4**(2), 2004–2013 (2021).
- ¹¹B. Zhao, L. Wang, Y. Shuai, and Z. M. Zhang, “Thermophotovoltaic emitters based on a two-dimensional grating/thin-film nanostructure,” *Int. J. Heat Mass Transf.* **67**, 637–645 (2013).
- ¹²J. Lyu, G. Cui, L. Shi, L. Gao, M. Bai, and L. Jiang, “FDTD method study on the effects of geometric parameters on the W-Al₂O₃ nano-structure thermal emitter,” *Nanotechnology* **32**(8), 085706 (2021).
- ¹³Y. Khorrami and D. Fathi, “Broadband thermophotovoltaic emitter using magnetic polaritons based on optimized one- and two-dimensional multilayer structures,” *J. Opt. Soc. Am. B* **36**(3), 662–666 (2019).
- ¹⁴J.-M. Kim, K.-H. Park, D.-S. Kim, B. Hwang, S.-K. Kim, H.-M. Chae, B.-K. Ju, and Y.-S. Kim, “Design and fabrication of spectrally selective emitter for thermophotovoltaic system by using nano-imprint lithography,” *Appl. Surf. Sci.* **429**, 138–143 (2018).
- ¹⁵J. Song, M. Si, Q. Cheng, and Z. Luo, “Two-dimensional trilayer grating with a metal/insulator/metal structure as a thermophotovoltaic emitter,” *Appl. Opt.* **55**(6), 1284 (2016).
- ¹⁶T. C. Huang, B. X. Wang, and C. Y. Zhao, “A novel selective thermophotovoltaic emitter based on multipole resonances,” *Int. J. Heat Mass Transf.* **182**, 122039 (2022).
- ¹⁷X. J. Liu, C. Y. Zhao, B. X. Wang, and J. M. Xu, “Tailorable bandgap-dependent selective emitters for thermophotovoltaic systems,” *Int. J. Heat Mass Transf.* **200**, 123504 (2023).
- ¹⁸Y. Xuan, X. Chen, and Y. Han, “Design and analysis of solar thermophotovoltaic systems,” *Renew. Energy* **36**(1), 374–387 (2011).
- ¹⁹D. Li and Y. Xuan, “Design and evaluation of a hybrid solar thermophotovoltaic-thermoelectric system,” *Sol. Energy* **231**, 1025–1036 (2022).
- ²⁰Y. Zhao, F. Yang, J. Song, and R. Hu, “Bayesian-optimized infrared grating for tailoring thermal emission to boost thermophotovoltaic performance,” *J. Appl. Phys.* **133**(12), 124904 (2023).
- ²¹M. He, J. R. Nolen, J. Nordlander, A. Cleri, N. S. McIlwaine, Y. Tang, G. Lu, T. G. Folland, B. A. Landman, J.-P. Maria, and J. D. Caldwell, “Deterministic inverse design of tamm plasmon thermal emitters with multi-resonant control,” *Nat. Mater.* **20**(12), 1663–1669 (2021).
- ²²R. Hu, J. Song, Y. Liu, W. Xi, Y. Zhao, X. Yu, Q. Cheng, G. Tao, and X. Luo, “Machine learning-optimized tamm emitter for high-performance thermophotovoltaic system with detailed balance analysis,” *Nano Energy* **72**, 104687 (2020).
- ²³W. Ma, Z. Liu, Z. A. Kudyshv, A. Boltasseva, W. Cai, and Y. Liu, “Deep learning for the design of photonic structures,” *Nat. Photonics* **15**(2), 77–90 (2021).
- ²⁴H. Wang, Z. Zheng, C. Ji, and L. Jay Guo, “Automated multi-layer optical design via deep reinforcement learning,” *Mach. Learn.: Sci. Technol.* **2**(2), 025013 (2021).
- ²⁵V. Mnihi, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, “Human-level control through deep reinforcement learning,” *Nature* **518**(7540), 529–533 (2015).
- ²⁶S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, “Inverse design in nanophotonics,” *Nat. Photonics* **12**(11), 659–670 (2018).

- ²⁷E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, 1998).
- ²⁸T. Siefke, S. Kroker, K. Pfeiffer, O. Puffky, K. Dietrich, D. Franta, I. Ohlídal, A. Szeghalmi, E.-B. Kley, and A. Tünnermann, “Materials pushing the application limits of wire grid polarizers further into the deep ultraviolet spectral range,” *Adv. Opt. Mater.* **4**(11), 1780–1786 (2016).
- ²⁹T. J. Bright, J. I. Watjen, Z. M. Zhang, C. Muratore, and A. A. Voevodin, “Optical properties of HfO₂ thin films deposited by magnetron sputtering: From the visible to the far-infrared,” *Thin Solid Films* **520**(22), 6793–6802 (2012).
- ³⁰S. Yu, P. Zhou, W. Xi, Z. Chen, Y. Deng, X. Luo, W. Li, J. Shiomi, and R. Hu, “General deep learning framework for emissivity engineering,” *Light: Sci. Appl.* **12**(1), 291 (2023).
- ³¹J. Hugonin and P. Lalanne, “RETICOLO software for grating analysis,” [arXiv:2101.00901](https://arxiv.org/abs/2101.00901) (2021).
- ³²L. G. Ferguson and L. M. Fraas, “Theoretical study of GaSb PV cells efficiency as a function of temperature,” *Sol. Energy Mater. Sol. Cells* **39**(1), 11–18 (1995).
- ³³W. Shockley and H. J. Queisser, “Detailed balance limit of efficiency of p-n junction solar cells,” *J. Appl. Phys.* **32**(3), 510–519 (2004).
- ³⁴M. E. Nell and A. M. Barnett, “The spectral p-n junction model for tandem solar-cell design,” *IEEE Trans. Electron Devices* **34**(2), 257–266 (1987).
- ³⁵M. A. Green, *Solar Cells: Operating Principles, Technology, and System Applications* (Englewood Cliffs, NJ, Prentice-Hall, Inc., 1982) p. 288.
- ³⁶X. Liu, T. Tyler, T. Starr, A. F. Starr, N. M. Jokerst, and W. J. Padilla, “Taming the blackbody with infrared metamaterials as selective thermal emitters,” *Phys. Rev. Lett.* **107**(4), 045901 (2011).
- ³⁷W. Xi, Y.-J. Lee, S. Yu, Z. Chen, J. Shiomi, S.-K. Kim, and R. Hu, “Ultrahigh-efficient material informatics inverse design of thermal metamaterials for visible-infrared-compatible camouflage,” *Nat. Commun.* **14**(1), 4694 (2023).
- ³⁸Y. Shi, W. Li, A. Raman, and S. Fan, “Optimization of multilayer optical films with a memetic algorithm and mixed integer programming,” *ACS Photonics* **5**(3), 684–691 (2018).
- ³⁹W.-W. Zhang, H. Qi, Y.-M. Yin, and Y.-T. Ren, “Tailoring radiative properties of a complex trapezoidal grating solar absorber by coupling between SPP and multi-order MP for solar energy harvesting,” *Opt. Commun.* **479**, 126416 (2021).

Supplementary Material for

Enhancing Overall Performance of Thermophotovoltaics via Deep Reinforcement Learning-Based Optimization

Shilv Yu ¹, Zihe Chen ¹, Wentao Liao ¹, Cheng Yuan ², Bofeng Shang ^{3,*}, Run Hu ^{1,*}

¹ School of Energy and Power Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

² Wuhan Fiberhome Fuhua Electric Co., Ltd, Wuhan 430074, China

³ School of Physics and Microelectronics, Zhengzhou University, Zhengzhou 450001, China

* Corresponding author

E-mail address: shangbofeng@zzu.edu.cn (B. Shang), hurun@hust.edu.cn (R. Hu)

Content

The details of Bayesian Optimization and Genetic Algorithm for optimizing selective emitters (SEs) in thermophotovoltaic (TPV) system. (Note 1–2)

The impact of thickness of the top and intermediate layers on the overall performance of the SE and corresponding TPV system. (Fig. S1–S3)

Note 1: Bayesian Optimization for optimizing SE

The Bayesian Optimization (BO) was performed with the help of the open-source Python library PHYSBO. Before performing BO, a list of all candidates needs to be created, which is consistent with the design parameter space when using DRL method. The number of iterations was set as 10149 to ensure that the total number of calculated structures is the same as that of DRL, of which 100 iterations are random searches for obtaining training data. Although there is significant difference between the set number of iterations (10149) and the number of candidate structures (7716375), our main purpose is to compare the optimization efficiency of BO and DRL. Therefore, the number of set iterations can meet the comparison condition. In addition, the random feature map and Thompson sampling are utilized to avoid the computationally expensive training process. The results are shown in Fig. 4(c). It can be seen that the FoM of the optimal SE obtained by BO is 2.1575, which is worse than that of obtained by DRL. More seriously, it takes longer for FoM to reach convergence.

Note 2: Genetic Algorithm for optimizing SE

The Genetic Algorithm (GA) was performed in MATLAB. The size of the population is set to 100 and the number of iterations (generations) is 50. Similarly, the population size and the number of iterations are small relative to the total candidate structures, but sufficient in order to achieve the comparison of optimization efficiency with DRL method. The tracing of FoM vs the generation is shown in Fig. 4(d). The optimal structure of SE is consistent with that obtained by our DRL method. However, the convergence time of GA is longer than that of DRL, which demonstrate the high efficiency of the DRL method. DRL, BO and GA all run on a local server (Intel Xeon CPU E5-2690).

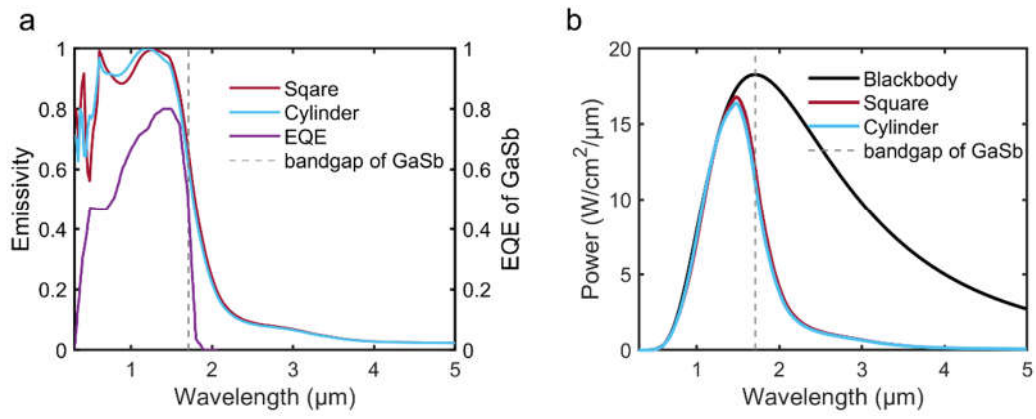


Fig. S1 (a) Emissivity spectra of optimal square topped SE and cylinder topped SE. (b) Radiation energy spectra of two SEs.

Table S1 Performance index of the corresponding TPV system of SEs with top square and cylinder pattern

Structure	System efficiency (%)	Output power density (W/cm^2)	FoM
Square	38.26%	5.78	2.21
Cylinder	38.65%	5.68	2.19

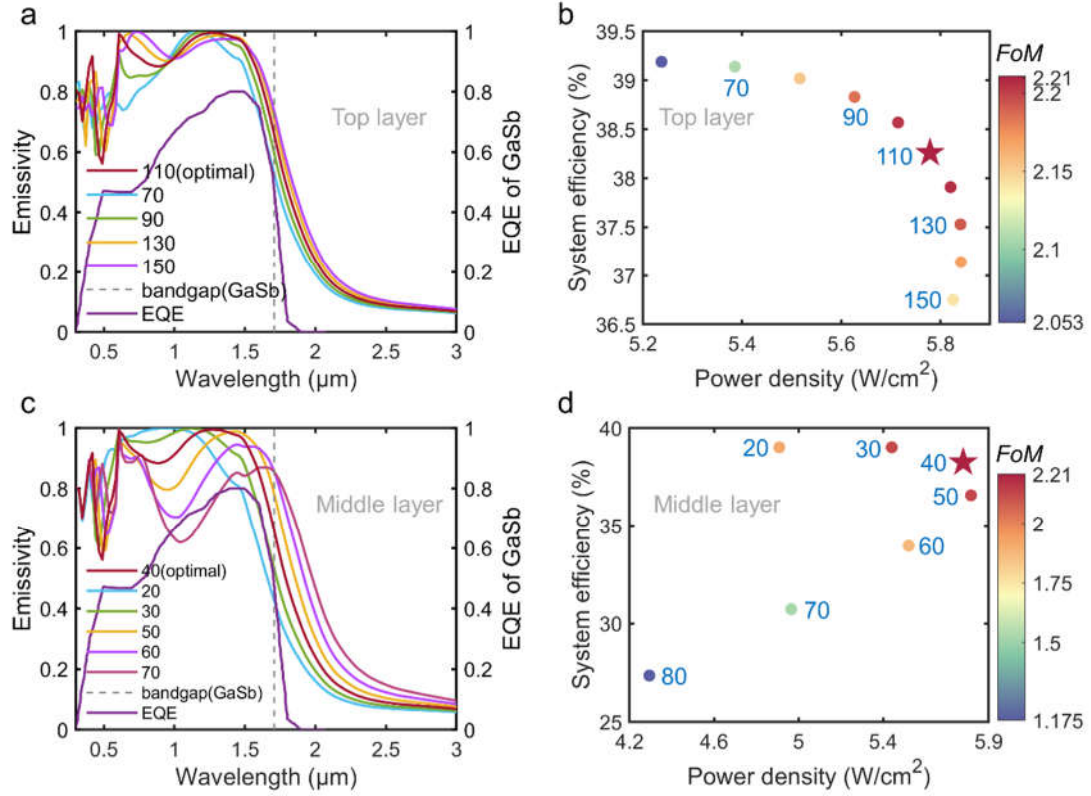


Fig. S2 Effects of top and middle layer thickness on SE and TPV system performance when the materials of top and middle layer are fixed as TiO_2 and Si. (a) The emissive spectra of SEs with different top layer thicknesses. (b) Effects of top layer thickness on TPV system performance. (c) The emissive spectra of SEs with different middle layer thicknesses. (d) Effects of middle layer thickness on TPV system performance.

When the thickness of the top layer decreases, the emissivity decreases in the ranges of both above and below the bandgap wavelength of the GaSb cell (1.708 μm), resulting in a slight increase in the efficiency, but a significant decrease in the power density, as shown in Fig. S2 (a) and (b). When the thickness of the top layer is slightly increased, the above trends are reversed. When the thickness exceeds 130 nm, both the efficiency and power density are reduced. Therefore, Fig. S2 (b) can be considered as the Pareto front of the system performance under different SEs' top layer thickness. Considering two performance indexes of the system, the structure optimized by our DRL model is demonstrated to be the optimal one.

When the thickness of the middle layer decreases, the emissivity spectrum is blue shifted, resulting in slightly increase in system efficiency, as shown in Fig. S2 (c) and (d). However, the reduction in emissivity at wavelengths ($\sim 1.5 \mu\text{m}$) slightly smaller than the cell bandgap leads to a significant reduction in the power density. When the thickness increases, the emissivity spectrum is redshifted and the in-band ($< 1.708 \mu\text{m}$) emissivity is significantly reduced, resulting in the performance degradation of the system. Consequently, the optimal performance is achieved when the thickness of intermediate Si layer is 40 nm, which demonstrates the effectiveness of our DRL model.

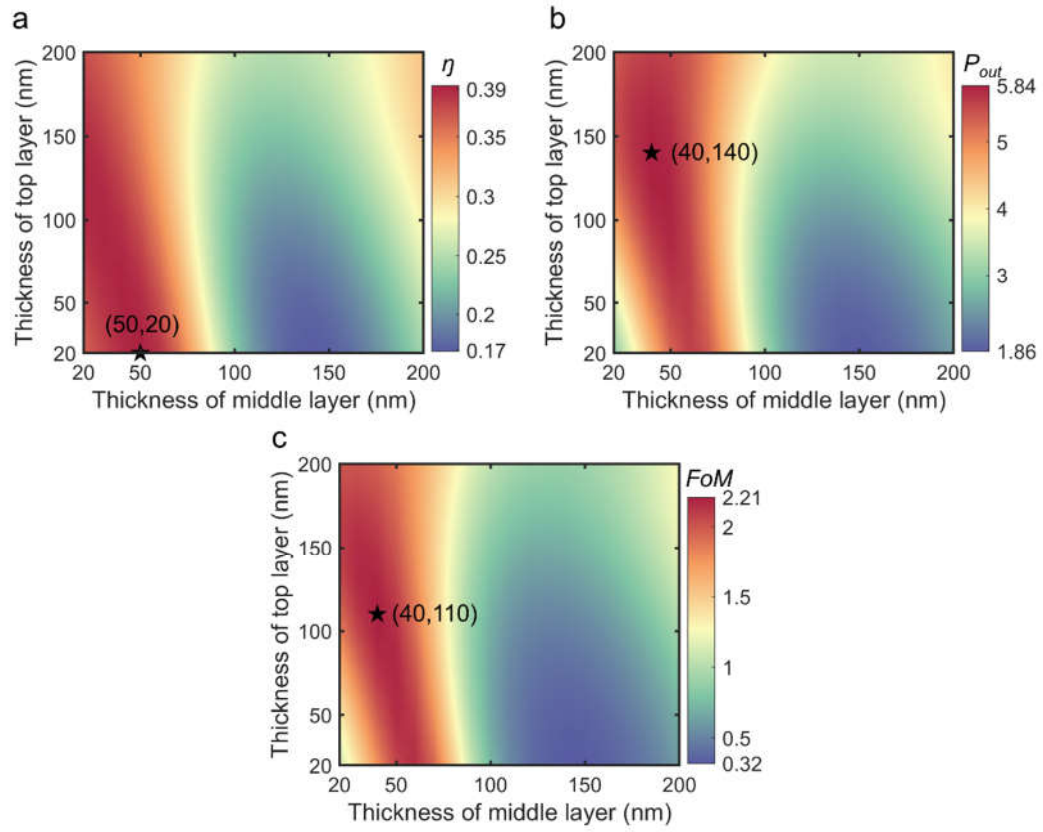


Fig. S3 Effects of top and middle layer thickness on TPV system performance. (a) System efficiency. (b) Output power density. (c) FoM.

Although it can be observed that the system efficiency and power density reach their optimal values at different thickness combinations, respectively, we achieved the optimal enhancement of overall TPV system performance.