



OPEN Higher-order factorization machine for accurate surrogate modeling in material design

Sanghyo Hwang^{1,4}, Seongmin Kim^{2,4}, Zhihao Xu³, Tengfei Luo³✉ & Eungkyu Lee¹✉

Efficient and robust optimization is important in material science for identifying optimal structural parameters and enhancing material performance. Surrogate-based active learning algorithms have recently gained great attention for their ability to efficiently navigate large, high-dimensional design spaces. Among surrogate models, 2nd-order factorization machine (FM) models are widely employed as the surrogate model in active learning algorithms due to their balance between simplicity and effectiveness. However, their quadratic nature limits their capacity to capture complex, higher-order interactions among variables, often leading to suboptimal solutions. To overcome this limitation, we propose an active learning scheme integrating a 3rd-order FM model, capable of modeling three-variable interactions and more intricate relationships in material systems. We comprehensively evaluate the surrogate modeling performance of the 3rd-order FM case using various objective functions. Furthermore, we examine the optimization reliability and efficiency of the 3rd-order FM-based active learning in a real-world material design task (e.g., nanophotonic structures for transparent radiative cooling). Our study shows that the 3rd-order FM outperforms the 2nd-order model in both surrogate accuracy and optimization performance, highlighting higher-order models' promises for material design and optimization problems.

Keywords Factorization machine, Higher-order interactions, Optimization, Machine learning, Material design

Data-driven approaches, such as machine learning (ML), have become important tools for addressing complex design challenges in material sciences, particularly in high-dimensional complex optimization problems^{1,2}. ML models are particularly effective at uncovering hidden relationships within experimental or simulation data, enabling the construction of surrogate functions that approximate parametric spaces for optimization. However, building accurate surrogate models with ML often demands large amounts of training data, which can be costly and time-consuming to obtain^{3–5}. To mitigate this challenge, active learning algorithms have been proposed, which strategically collect training data, allowing for iterative optimization towards global optimal with a sparse initial dataset^{5–9}.

Among various ML models, the 2nd-order factorization machine (FM)¹⁰ is efficient in modeling systems. The FM model significantly reduces algorithmic complexity from polynomial to linear, enhancing computational efficiency in surrogate modeling tasks. As a result, FM-based surrogate models have demonstrated great performance in active learning frameworks, requiring fewer data points than conventional big-data-driven ML approaches while maintaining accuracy^{11–14}. While other ML models, such as random forests^{11,15}, neural networks¹², and Bayesian optimization (BO)¹⁶ can also formulate surrogate functions by capturing complex relationships between variables. These limitations can lead to suboptimal results due to the difficulty of exhaustively exploring optimization spaces. In contrast, the 2nd-order FM provides both computational efficiency and seamless integration with quantum computing, providing an effective solution for complex optimization tasks^{5,17,18}. In particular, quantum annealing (QA)^{19–21} has emerged as an effective solver for surrogate functions formulated by the 2nd-order FM.

However, cost functions defined by multi-dimensional design parameters often cannot be adequately interpreted by 2nd-order interactions alone. Consequently, the 2nd-order FM, which inherently excludes higher-order terms, may lead to underfitting and limited applicability in complex material optimization tasks. For instance, recent studies applying the 2nd-order FM in optical film design, such as digital metasurfaces², optical

¹Department of Electronic Engineering, Kyung Hee University, Yongin-Si, Gyeonggi-do 17104, Republic of Korea.

²National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA.

³Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA.

⁴Sanghyo Hwang and Seongmin Kim contributed equally to this work. ✉email: tluo@nd.edu; eleest@khu.ac.kr

diodes⁴, and transparent radiative cooling windows⁵, have shown that the optimized surrogate functions often rediscover structures already present in the training dataset. This highlights the limitation of the 2nd-order FM model, necessitating higher-order terms for a more accurate surrogate representation. A promising solution meeting such criteria is the use of higher-order FMs^{22–24}, which generalize a function analogous to polynomial regressions. However, their application into active learning schemes for material optimization remains unexplored.

It should be noted that the FM models will be compatible with QA, as demonstrated in previous work^{4,5}, which allows to solve high-dimensional design spaces. We agree that comparisons with other surrogate models are valuable. Random forest is an ensemble learning method that constructs multiple decision trees randomly and combines their outputs to make predictions. Although this model has shown great predictive performance in many applications, solving their surrogate functions remain challenging. In one study, a quantum genetic algorithm (QGA), which has higher performance than classical genetic algorithms, was used to explore surrogate functions obtained from the random forest model¹⁵. However, QGA was still less effective than QA in exploring large optimization spaces due to its limited capability to escape local minima¹⁵. BO¹⁶ is widely used for global optimization, but its computational cost significantly increases as datasets grow. The predictive complexity of the BO scales quadratically with the number of data points, which pose significant computational costs²⁵. Furthermore, BO requires solving acquisition functions using heuristic algorithms such as particle swarm optimization (PSO), and GAs, which can be intractable for high-dimensional spaces.

In this study, we investigate the use of the 3rd-order FM as a surrogate in active learning for material design tasks (nanophotonics as a practical demonstration). Our scheme incorporates three-variable interaction terms, which are essential for analyzing complex nanophotonics systems. We perform benchmarking studies to evaluate the surrogate modeling performance of 3rd-order FM with various objective functions, and compare it with the 2nd-order FM. For active learning, we focus on a representative design task (multilayered optical structure), evaluating optimization performance based on the distribution of the identified optimal structures. Our results show that the 3rd-order FM consistently outperforms the 2nd-order FM, enabling the active learning scheme to discover superior designs more reliably and effectively.

Theory

Factorization machine models

2nd-order factorization machine

The 2nd-order FM learns the linear and the quadratic coefficients of input variables from labeled training dataset. Here, linear terms characterize the individual self-interactions of each variable while quadratic terms capture pairwise interactions between two different variables. The 2nd-order FM is given by:¹⁰

$$\hat{y}_{2nd}(\mathbf{x}) = b_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \mathbf{v}_i \cdot \mathbf{v}_j x_i x_j, \quad (1)$$

where x_i indicates the i -th element of a binary vector x (length of n), and $\hat{y}_{2nd}(\mathbf{x})$ is the output of the model, representing the corresponding label of \mathbf{x} . The model parameters b_0 , w , and $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$, respectively, denote the global bias, linear coefficients, and quadratic coefficients. $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$ is the dot product of i -th, j -th raw vectors of $\mathbf{V}^{2nd} \in \mathbb{R}^{n \times k}$, where k is the latent space size. The complexity of the learning process of the 2nd-order FM can be substantially reduced by factorizing quadratic interactions with the latent matrix $(v_{i,j})$, as expressed:

$$\sum_{i=1}^n \sum_{j=i+1}^n \mathbf{v}_i \cdot \mathbf{v}_j x_i x_j = \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{i,f} x_i \right)^2 - \sum_{i=1}^n v_{i,f}^2 x_i^2 \right). \quad (2)$$

Equation (2) reformulates the quadratic interactions into linear combination of the coefficients, effectively reducing the complexity from $O(kn^2)$ to $O(tkn)$, where t is the order of FM ($=2$, for 2nd-order FM), thereby improving training efficiency of FM.

3rd-order factorization machine

The 3rd-order FM^{26,27} is capable of capturing interactions between three different variables. This enhanced model incorporates additional parameters to represent 3rd-order interactions, which can be expressed by:

$$\hat{y}_{3rd}(\mathbf{x}) = \hat{y}_{2nd}(\mathbf{x}) + \sum_{j_3 > j_2 > j_1}^n \mathbf{v}_{j_1} \cdot \mathbf{v}_{j_2} \cdot \mathbf{v}_{j_3} x_{j_1} x_{j_2} x_{j_3}, \quad (3)$$

where $\langle \mathbf{v}_{j_1}, \mathbf{v}_{j_2}, \mathbf{v}_{j_3} \rangle$ represents the 3rd-order interaction coefficients, which are the dot product of three raw vectors of $\mathbf{V}^{3rd} \in \mathbb{R}^{n \times k}$. In this study, we set the latent space size of \mathbf{V}^{2nd} and \mathbf{V}^{3rd} to 8. Similar to quadratic coefficients, the algorithmic complexity for model training can be reduced by factorizing 3rd-order coefficients, as expressed in the following equation:

$$\sum_{j_3 > j_2 > j_1}^n \mathbf{v}_{j_1} \cdot \mathbf{v}_{j_2} \cdot \mathbf{v}_{j_3} x_{j_1} x_{j_2} x_{j_3} = \frac{1}{6} \left[H^3(\mathbf{v}, \mathbf{x}) - 3D^{2,1}(\mathbf{v}, \mathbf{x}) + 2D^3(\mathbf{v}, \mathbf{x}) \right], \quad (4)$$

where a polynomial kernel (H) and an operator (D) acting on p and x are respectively defined by:^{26,27}

$$H^3(\mathbf{v}, \mathbf{x}) = \langle \mathbf{v}, \mathbf{x} \rangle^3 = \sum_{j_1=1}^n \sum_{j_2=1}^n \sum_{j_3=1}^n v_{j_1} x_{j_1} v_{j_2} x_{j_2} v_{j_3} x_{j_3}, \quad (5)$$

$$D^m(\mathbf{v}, \mathbf{x}) := \sum_{j=1}^n (v_j x_j)^m \text{ and } D^{m,n}(\mathbf{v}, \mathbf{x}) := D^m(\mathbf{v}, \mathbf{x}) D^n(\mathbf{v}, \mathbf{x}). \quad (6)$$

This formulation also transforms the 3rd-order interaction models into linear ones like the 2nd-order interaction representations, greatly reducing the complexity from $O(kn^3)$ to $O(tkn)$, where t is 3.

Active learning scheme

We validate our approach through the optimization of planar multilayered (PML) structures as a testbed for a transparent radiative cooler (TRC)^{28–30} coating, which has a complex optimization objective (Section "Active learning study with the design task of optical coating"). The active learning algorithm accumulates data points converging toward optimal points through an iterative process (Fig. 1):

1. *Generating an initial dataset* including 25 randomly selected data points.
2. *Training an FM model* (either 2nd or 3rd-order FM) to build a surrogate function.
3. *Identifying an optimal structure (binary vector)* that has the best figure-of-merit (FOM) of the given surrogate function. Here, several methods can be used including QA, simulated annealing, and brute force search.
4. *Calculating FOM* for the selected structure (binary vector) using a numerical method. In this work, the transfer matrix method (TMM)³¹ is used; TMM is a wave-optics simulation that uses matrices to analyze wave transmission and reflection in multilayer structures.
5. *Updating the dataset* by including the new data point (i.e., selected binary vector and calculated FOM). When a newly proposed binary vector already exists in the dataset, we add a random, previously unlabeled binary vector instead.

This approach enables the identification of the optimum point in the complex parametric space.

Benchmarking study with analytical function

Analytical function

We first conduct a benchmarking study to evaluate the models of FM performance in capturing the complexity of a target function. Since most of the realistic design tasks we try to optimize are often black box functions. Thus, in order to systematically study at the performance of the FM models, we used an analytical function that is relatively simple and smooth parametric landscape. The analytical function, which yields complex landscape, is employed as a target objective:

$$f(d_x) = 50 \log(d_x + 1) \sin\left(\frac{1}{20}(d_x - 75)\right) + \frac{10^{-7}(d_x - 2.2)(d_x - 4)(d_x - 30)(d_x - 45)}{9}, \quad (7)$$

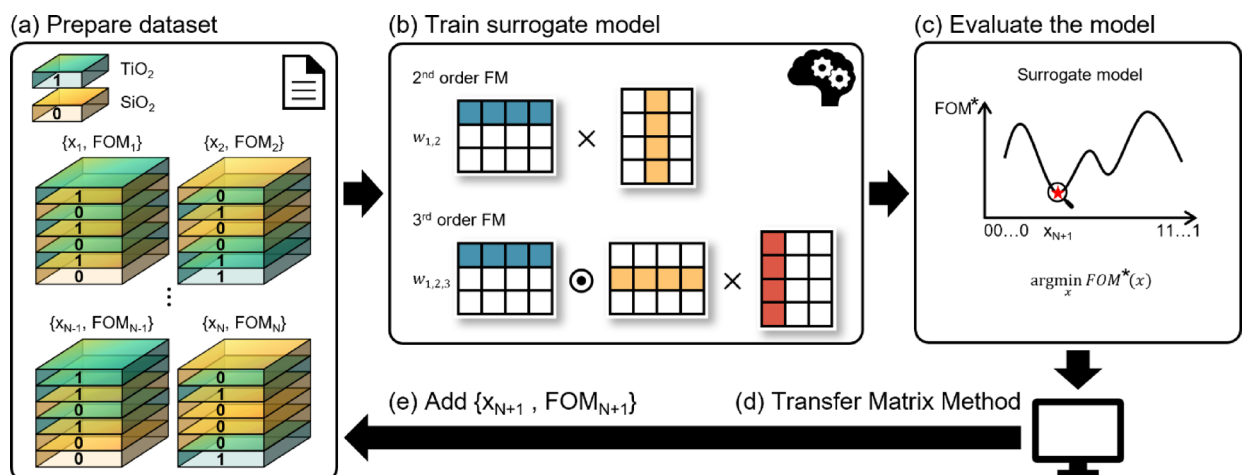


Fig. 1. Active learning employing the 2nd and 3rd-order FMs. **(a)** Preparing a training dataset consisting of input binary vectors (\mathbf{x}_i) and associated outputs FOM_i . **(b)** Formulating a surrogate function with the training dataset. **(c)** Evaluating the surrogate function to identify the optimal binary vector. **(d)** Calculating FOM for the optimal binary vector. **(e)** Updating the training dataset with the identified binary vector and the corresponding FOM.

where d_x is a decimal representation of a binary vector x with a length of 10, where the domain of d_x is $0 \leq d_x \leq 1023$. The output $f(d_x)$ consists of 2^{10} (= 1024) discrete data points. We randomly sample N data points from these 2^{10} points to train both 2nd and 3rd-order FM. These FM models formulate surrogate functions by optimizing their trainable parameters. We train both 2nd and 3rd-order FM models using datasets containing varying numbers of sampled data points and evaluate their surrogate modeling performance by comparing their ability to approximate the target function. A similar study is conducted for a binary vector x of length 20. In this case, FM models are trained using datasets containing 2^{16} (= 65,536) sample points.

Surrogate modeling performance on analytical function

The landscape of a surrogate function can be visualized by evaluating all possible states (i.e., d_x from 0 to $2^n - 1$, Fig. 2 and Fig. S1). If the FM model successfully learns the relationship between input vectors and their corresponding labels, the landscape of the surrogate function should closely resemble (ideally match) the target function with small RMSE values. Figure 2 shows that both the 2nd and 3rd-order FM models can capture the overall trend. Notably, surrogate functions trained with more data points can more accurately describe the analytical function (Figs. 2a–f), as indicated by lower RMSE values. With a smaller training data, the RMSE of the 2nd-order FM is 587.67, which is smaller than that of 3rd-order FM (963.96). This indicates that the 2nd-order FM is more effective in approximating the target function when data points are sparse. However, as the size of the training dataset increases to 200, the 3rd-order FM becomes more accurate in describing the data, as evidenced by a significantly reduced RMSE (126.32), while the RMSE of the 2nd-order FM remains large (503.39). Overfitting refers to the phenomenon where a model fits the training data too closely, resulting in poor generalization to unseen data. In Fig. 2e, although the 3rd-order FM was trained on only 200 data points, it generalizes well. Hence, the model shows great performance for remaining ~ 800 unseen data points, where the predicted values are well matched with the actual values. This indicates good generalization capability of the FM model rather than overfitting. Furthermore, parity plots with R^2 scores (Fig. S2) highlight demonstrate the 3rd-order FM have great performance especially when enough data points are used for training. These results are consistently observed as the system size increases to $n = 20$ (Fig. 2c, f). These results indicate that with sufficient training data, the 3rd-order coefficients learn the complex interactions within the dataset, which is hard to capture by quadratic interactions. This capability of exploiting larger datasets in the 3rd-order FM suggests that the accuracy of surrogate function is improved near local and non-quadratic parametric space including optimal points as the active learning proceeds through additional optimization cycles.

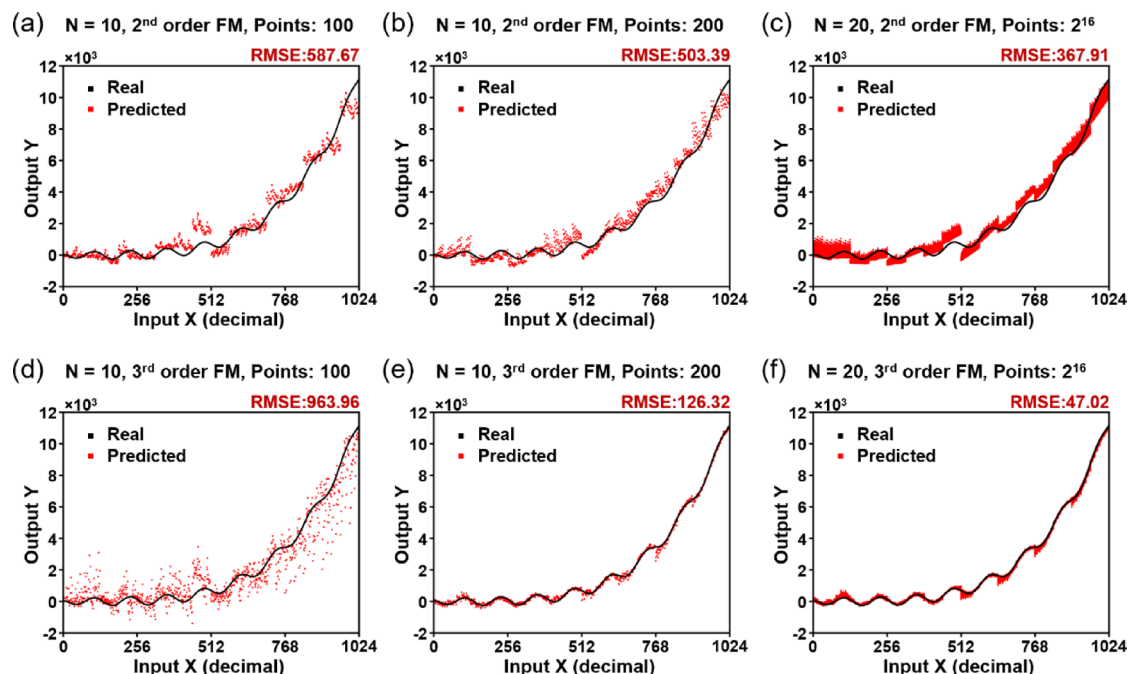


Fig. 2. The surrogate modeling performance of the 2nd and 3rd-order FMs for an analytical function. The output of surrogate functions as a function of input (d_x) after training (a–c) 2nd-order FM and (d–f) 3rd-order FM. Black dots are the output directly evaluated from the analytical function. (a, c) Dataset size = 100, binary vector length = 10. (b, d) Dataset size = 200, binary vector length = 10. (e, f) Dataset size = 2^{16} , binary vector length = 20.

Active learning study with the design task of optical coating

Numerical cost function

We investigate the performance of 2nd and 3rd-order FMs with a practical optimization problem to understand their effectiveness in handling real-world optimization scenarios. For this purpose, we select the optimization task of TRC coating.

A TRC coating for windows is designed to save energy consumption for cooling by emitting thermal radiation through an atmosphere window ($8 < \lambda < 13 \mu\text{m}$), while blocking UV and NIR and selectively transmitting visible photon within solar spectrum (Fig. 3). In this study, a TRC coating consists of a multilayered structure on the glass substrate with a thermal radiative polymer layer (polydimethylsiloxane; PDMS, $\sim 40 \mu\text{m}$ thick) placed on the top of the PML^{32–35}. PML, subject of optimization (total thickness: 1,200 nm), selectively transmits visible photons while reflecting photons in UV and NIR regions (Fig. 3a). PML is divided into n pseudo-layers, where each layer is one of two dielectric materials: silicon dioxide (SiO_2) or titanium dioxide (TiO_2). The configuration of pseudo-layers is encoded as a binary vector, where each digit corresponds to the material selection for each layer: '0' for SiO_2 , and '1' for TiO_2 (Fig. 3b). The structure presented in Fig. 3b is an example of the TRC structure, and as the length of the binary vector increases, the overall thickness is made constant, and the thickness of one layer is defined thinner.

To evaluate the performance of the designed PML at a given binary vector, we define a FOM, which quantifies the squared difference between the transmitted irradiance of the ideal TRC and that of the designed structure:

$$\text{FOM} = \frac{10 \int_{300}^{2500} (\text{TI}_{\text{ideal}}(\lambda) - \text{TI}_{\text{designed}}(\lambda))^2 d\lambda}{\int_{300}^{2500} S(\lambda)^2 d\lambda}. \quad (8)$$

In this equation, $S(\lambda)$ represents the solar spectral irradiance, and $\text{TI}(\lambda) (=S(\lambda) \times T(\lambda))$, where $T(\lambda)$ is transmission efficiency) represents transmitted irradiance through the ideal or designed PML. The transmission efficiency of the PML is calculated by TMM. We note that the smaller the FOM indicates the better design.

Surrogate modeling performance on numerical cost function

A numerical cost function can be formed with input binary vectors (number of pseudo layers, $n=12$) and the corresponding FOMs as the outputs. The landscape of the function is analyzed by evaluating all possible binary vectors using TMM, with the results plotted as black dots in Fig. 4. Surrogate functions are then constructed using 2nd-order and 3rd-order FMs, trained on 100 randomly sampled data points. The landscapes of these surrogate functions are shown as red dots in Fig. 4.

With the sparse training dataset, the RMSE of the 2nd-order FM is 1.16 (Fig. 4a), which is lower than that of the 3rd-order FM (1.32, Fig. 4c). This can be visually confirmed by the distribution density of the outputs. However, as the training dataset size increases to 500, the surrogate function from the 3rd-order FM shows better accuracy than that from the 2nd-order FM (Figs. 4b, d). Notably, the output density distribution of the 3rd-order FM is closely matched with that of the numerical function, indicating its improved ability to capture the target landscape. This result aligns well with our earlier observations in Fig. 2, confirming that incorporating 3rd-order interactions enables the FM model to effectively learn complex optimization spaces when sufficient data is available. While the 2nd-order FM has been widely employed in photonic structure design by employing individual features and pairwise interactions^{4,5,9,17,36,37}, our results highlight the importance of considering higher-order interactions.

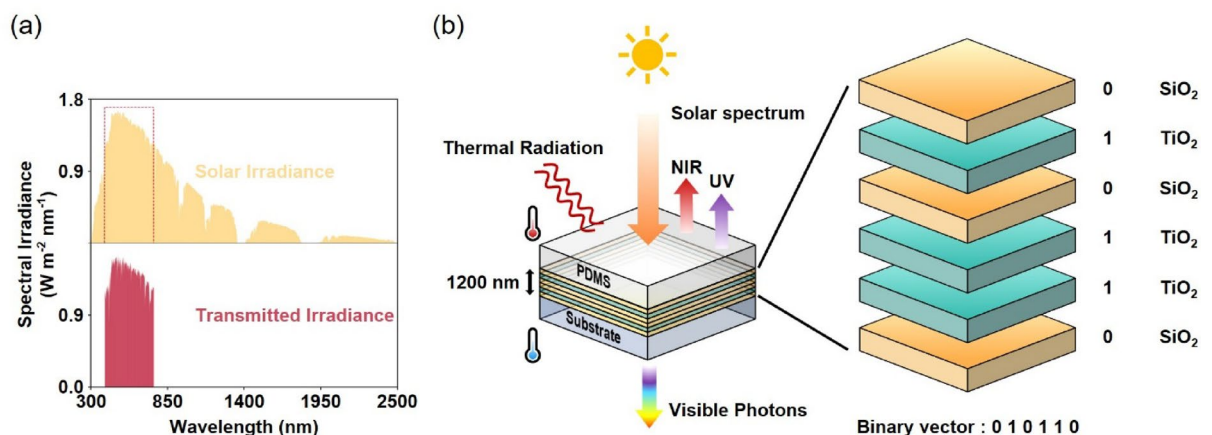


Fig. 3. Design task of optical coating using an active learning scheme. **(a)** Solar irradiance (air mass 1.5 global), transmission efficiency (highlighted by a red-dotted line), and the corresponding transmitted irradiance of an ideal TRC structure as a function of wavelength. **(b)** Schematic of the TRC structure, where the PML structure is encoded into a binary vector for input training data in ML. The schematics are created by Microsoft PowerPoint.

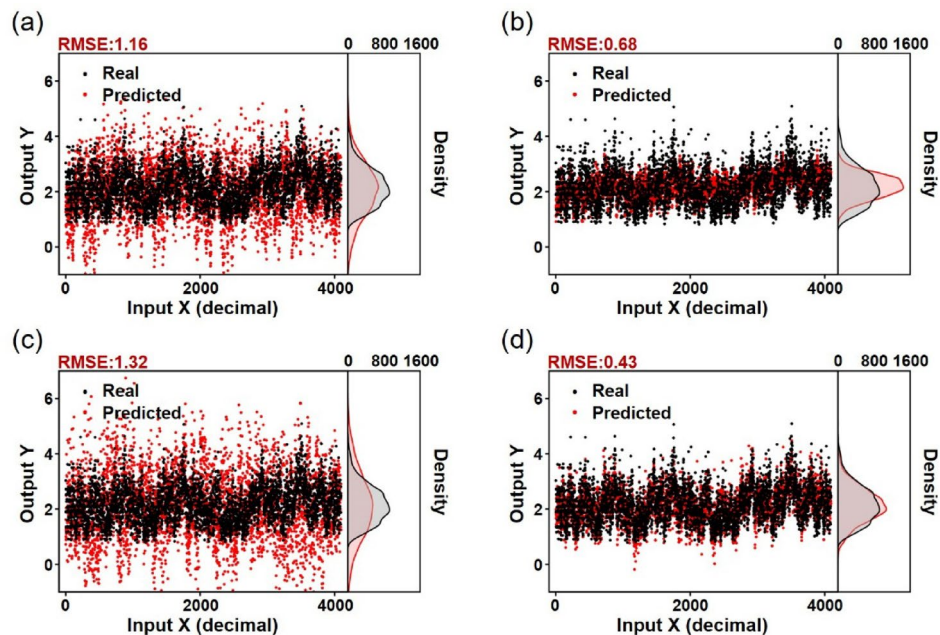


Fig. 4. The surrogate modeling performance of 2nd and 3rd-order FMs with the numerical cost function. The outputs of surrogate functions as a function of binary vectors (length = 12) after training with (a, b) 2nd-order FM and (c, d) 3rd-order FM. Black solid dots represent the outputs directly evaluated from the numerical function. Training dataset size: (a, c) 100. (b, d) 500.

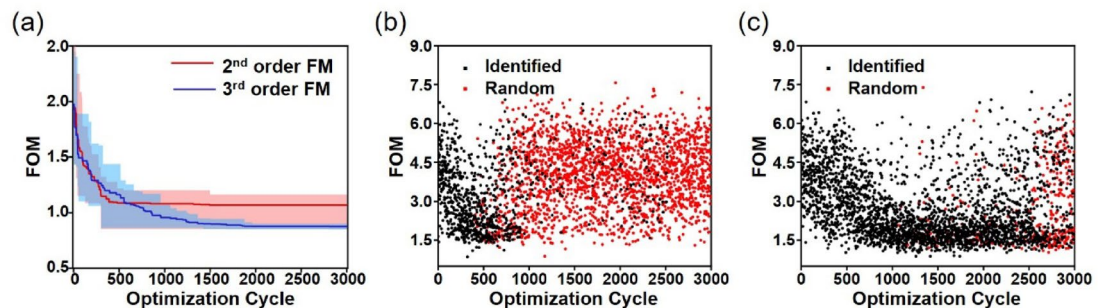


Fig. 5. Active learning with 2nd-order and 3rd-order FMs. (a) The statistical analysis of the identified smallest FOMs as a function of optimization cycles for (red) the 2nd-order and (blue) 3rd-order FMs. Solid lines represent the average values with different initial training sets, and shaded regions indicate the minimum–maximum. (b, c) FOM of the identified binary vector as a function of optimization cycles for (b) the 2nd-order FM and (c) the 3rd-order FM.

Performance analysis of the active learning algorithm

For the active learning study, we select the PML optimization task ($n=20$). To ensure reliable results, active learning is repeated 10 times using different training datasets, and we analyzed the statistics (average, minimum, and maximum) of the identified smallest FOMs (Fig. 5a). The results show that the 2nd-order FM-based active learning converges at ~400 optimization cycles, while the 3rd-order FM continues to reduce the FOM until ~2000 optimization cycles. Notably, active learning with the 3rd-order FM identifies better structures with lower FOMs. For the 3rd-order FM case, the gap between the maximum and the minimum values continuously narrows as the optimization cycle increases, achieving convergence after 2000 cycles. In contrast, for the 2nd-order FM case, the gap remains unchanged after 300 cycles.

Figures 5b and 5c show the FOM of the identified binary vector from the 2nd-order and the 3rd-order FM-based active learning as a function of optimization cycles. During the early optimization cycles (< 500), the 2nd-order FM case can continuously find better structures indicated by lowering FOMs as cycles proceed. However, after that, it usually fails to identify new structures, repeatedly selecting those already present in the training set, indicating convergence to a local minimum. In this scenario, adding these repeated structures to the dataset does not help escape the local minimum. Hence, random structures (red dots) are included into the dataset instead of the identified structures when duplicates appear (Fig. 5b). Since our ultimate goal is to maximize

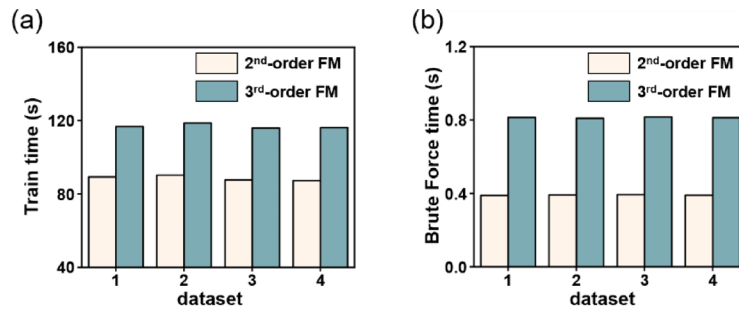


Fig. 6. Computational cost per optimization cycle in active learning. **(a)** Computational time to train FM models and **(b)** time required to find the optimal binary vector from FM-based surrogate functions using brute force search.

model improvement per cycle, adding a duplicate point can waste computational costs without enhancing the model accuracy. Hence, we add an unused point when the active learning proposes a data point that is already in dataset. The high frequency of random substitution indicates that the active learning algorithm is repeatedly proposing the points already explored, suggesting the model is approaching convergence. The active learning process is terminated at 3000th optimization cycle in this study. In contrast, the accumulation of additional data points for the 3rd-FM training progressively enhances the accuracy of the surrogate function. Therefore, the 3rd-order FM can formulate more reliable surrogate functions that are descriptive for the optimization landscape. Hence, the 3rd-order FM case achieves better optimization outcomes by finding lower FOMs (Fig. 5c).

Figure 6 compares the computational time per optimization cycle for the 2nd-order and 3rd-order FM models during active learning. The training time required for the 3rd-order FM is 31% higher than that for the 2nd-order FM due to the additional space for the three-variable interaction terms (Fig. 6a). Figure 6b presents the time to identify the optimal binary vector of the given surrogate functions through brute force search. This time is also comparable between the two FM models, although there is a slight increase for the 3rd-order FM. Since brute force search exhaustively explores the entire optimization space, we utilize it to establish the ground truth for each optimization cycle. However, as the problem size scales up, the exponential growth in computational complexity makes brute force search impractical. In such cases, QA offers a promising alternative, exploiting quantum principles, such as tunneling and superposition, to explore complex energy landscapes more efficiently than classical methods. This makes QA particularly advantageous for large-scale optimization tasks that are otherwise intractable^{5,20,37,38}. Overall, the results demonstrate that although the 3rd-order FM requires slightly higher computational cost in active learning, it more accurately captures the complex structure of the parametric space, leading to improved optimization performance.

Conclusion

In this paper, we performed a benchmarking study of the 3rd-order FM model using analytical and numerical cost functions and then evaluated its performance within an active learning scheme to design TRC coating. Our results indicate that with limited training data, the 2nd-order FM achieves a lower RMSE than the 3rd-order FM. However, as the dataset size increases, the 3rd-order FM significantly improves in accuracy. These characteristics are consistently found in both the analytical function and the numerical cost function. Despite their different parametric landscapes, this benchmarking study demonstrates the robustness of the 3rd-order FM in capturing complex design spaces. This highlights the trade-off between model complexity and data availability. While higher-order models such as the 3rd-order FM can achieve superior performance when sufficient data is available, simpler models like the 2nd-order FM may be more appropriate in data-scarce scenarios, especially when data acquisition is costly. These considerations allow practitioners to choose a suitable model depending on practical constraints. Likewise, in the active learning results, the 3rd-order FM performs better with identifying better optimization outcomes with low FOM values. These results suggest that the 3rd-order FM-based active learning can optimize any discretized structural parameters, such as transparent radiative cooler (TRC)⁵, optical diode⁴, and frequency selective surface (FSS)². These studies have demonstrated its effectiveness for structures with up to 378 design variables (TRC: $n=48$, optical diode: $n=40$, and FSS: $n=378$). As design spaces grow, the corresponding higher-order binary optimization (HOB) problems become increasingly difficult to solve via brute-force search used in this study. In such regimes, classical heuristics, such as swarm optimization (PSO), simulated annealing (SA), can be applied, but their efficiency decreases due to the exponential growth in optimization spaces. To address this limitation, quantum algorithms such as quantum annealing (QA), hybrid quantum annealing (HQA), or quantum approximate optimization algorithm (QAOA) can be used to solve such HOBs. This makes it a promising surrogate modeling strategy for challenging material optimization problems, such as the design of multilayer photonic coatings, metasurfaces, and advanced nanophotonic device architectures. Looking ahead, integrating the 3rd-order FM model with quantum computing techniques, particularly QA, may offer further performance enhancement.

Data availability

All data generated and analyzed during the study are available from the corresponding author upon reasonable request.

Received: 24 June 2025; Accepted: 8 September 2025

Published online: 09 October 2025

References

- Han, J., Jentzen, A. & Weinan, E. Solving high-dimensional partial differential equations using deep learning. *Proc. Natl. Acad. Sci.* **115**, 8505–8510 (2018).
- Kim, Y.-B. et al. W-band frequency selective digital metasurface using active learning-based binary optimization. *Nanophotonics* **14**, 1597–1606 (2025).
- Guan, Q., Raza, A., Mao, S. S., Vega, L. F. & Zhang, T. Machine learning-enabled inverse design of radiative cooling film with on-demand transmissive color. *ACS Photonics* **10**, 715–726 (2023).
- Kim, S. et al. Quantum annealing-aided design of an ultrathin-metamaterial optical diode. *Nano Conver.* **11**, 16 (2024).
- Kim, S. et al. High-performance transparent radiative cooler designed by quantum computing. *ACS Energy Lett.* **7**, 4134–4141 (2022).
- Chen, X., Zhou, H. & Li, Y. Effective design space exploration of gradient nanostructured materials using active learning based surrogate models. *Mater. Des.* **183**, 108085 (2019).
- Kim, J.-H. et al. Wide-angle deep ultraviolet antireflective multilayers via discrete-to-continuous optimization. *Nanophotonics* **12**, 1913–1921 (2023).
- Rao, Z. et al. Machine learning-enabled high-entropy alloy discovery. *Science* **378**, 78–85 (2022).
- Yu, J.-S. et al. Ultrathin Ge-YF₃ antireflective coating with 0.5% reflectivity on high-index substrate for long-wavelength infrared cameras. *Nanophotonics* **13**, 4067–4078 (2024).
- Rendle, S. In *2010 IEEE International conference on data mining*. 995–1000 (IEEE, 2010).
- Breiman, L. Random forests. *Mach. learn.* **45**, 5–32 (2001).
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- Song, Y.-Y. & Ying, L. Decision tree methods: Applications for classification and prediction. *Shanghai Arch. Psychiatry*. **27**, 130 (2015).
- Varatharajan, R., Manogaran, G. & Priyan, M. Retracted article: A big data classification approach using LDA with an enhanced SVM method for ECG signals in cloud computing. *Multimed. Tools Appl.* **77**, 10195–10215 (2018).
- Xu, Z. et al. Quantum-inspired genetic algorithm for designing planar multilayer photonic structure. *npj Comput. Mater.* **10**, 257 (2024).
- Frazier, P. I. In *Recent advances in optimization and modeling of contemporary problems* 255–278 (Informs, 2018).
- Kim, S., Jung, S., Bobbitt, A., Lee, E. & Luo, T. Wide-angle spectral filter for energy-saving windows designed by quantum annealing-enhanced active learning. *Cell Rep. Phys. Sci.* **5**(3), 101847 (2024).
- Kim, S. et al. A review on machine learning-guided design of energy materials. *Prog. Energy* **6**, 042005 (2024).
- Baldassi, C. & Zecchina, R. Efficiency of quantum vs. classical annealing in nonconvex learning problems. *Proc. Natl. Acad. Sci.* **115**, 1457–1462 (2018).
- Chapuis, G., Djidjev, H., Hahn, G. & Rizk, G. Finding Maximum Cliques on a Quantum Annealer. In *Proceedings of 22nd ACM International Conference on Computing Frontiers (CF '17)*. ACM. 63–70 (2017).
- Pastorello, D. & Blanzieri, E. Quantum annealing learning search for solving QUBO problems. *Quantum Inf. Process.* **18**, 303 (2019).
- An, Z. & Joe, I. TMH: Two-tower multi-head attention neural network for CTR prediction. *PLoS ONE* **19**, e0295440 (2024).
- Cai, S. et al. In *Proceedings of the 2021 International Conference on Management of Data*. 207–220.
- Meng, Z. et al. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1298–1307.
- Eggensperger, K., Hutter, F., Hoos, H. H. & Leyton-Brown, K. In *MetaSel@ 24–31 (ECAI)*
- Blondel, M., Fujino, A., Ueda, N. & Ishihata, M. Higher-order factorization machines. *Adv. neural inf. Process. Syst.* **29** (NeurIPS 2016), 3351–3359 (2016).
- Blondel, M., Ishihata, M., Fujino, A. & Ueda, N. In *International Conference on Machine Learning*. 850–858 (PMLR).
- Guo, J., Ju, S., Lee, Y., Gunay, A. A. & Shiomi, J. Photonic design for color compatible radiative cooling accelerated by materials informatics. *Int. J. Heat Mass Transf.* **195**, 123193 (2022).
- Jin, Y., Jeong, Y. & Yu, K. Infrared-reflective transparent hyperbolic metamaterials for use in radiative cooling windows. *Adv. Func. Mater.* **33**, 2207940 (2023).
- Wang, S. et al. Scalable thermochromic smart windows with passive radiative cooling regulation. *Science* **374**, 1501–1504 (2021).
- Mackay, T. G. & Lakhtakia, A. *The transfer-matrix method in electromagnetics and optics*. (Springer Nature, 2022).
- Guo, J. & Shiomi, J. Advances in materials informatics for tailoring thermal radiation: A perspective review. *Next Energy* **2**, 100078 (2024).
- Lee, K. W. et al. Visibly clear radiative cooling metamaterials for enhanced thermal management in solar cells and windows. *Adv. Func. Mater.* **32**, 2105882 (2022).
- Zhou, L. et al. A polydimethylsiloxane-coated metal structure for all-day radiative cooling. *Nat. Sustain.* **2**, 718–724 (2019).
- Zhou, Z., Wang, X., Ma, Y., Hu, B. & Zhou, J. Transparent polymer coatings for energy-efficient daytime window cooling. *Cell Rep. Phys. Sci.* **1**(11), 100231 (2020).
- Inoue, T. et al. Towards optimization of photonic-crystal surface-emitting lasers via quantum annealing. *Opt. Express* **30**, 43503–43512 (2022).
- Kitai, K. et al. Designing metamaterials with quantum annealing and factorization machines. *Phys. Rev. Res.* **2**, 013319 (2020).
- Kim, S. et al. Quantum annealing for combinatorial optimization: A benchmarking study. *npj Quantum Inf.* **11**, 1–8 (2025).

Acknowledgements

This research was supported by the Quantum Computing Based on Quantum Advantage Challenge Research (RS-2023-00255442) through the National Research Foundation of Korea (NRF). This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. Notice: This manuscript has in part been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a non-exclusive, paid up, irrevocable, world-wide license to publish or reproduce the published form of the manuscript, or allow others to do so, for U.S. Gov-

ernment purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-publicaccess-plan>).

Author contributions

S.H. and S.K. contributed equally. S.H. and S.K. developed the code. S.H. conducted experiments, and prepared all figures. All authors discussed the results and reviewed the manuscript.

Funding

National Research Foundation of Korea, RS-2023-00255442, U.S. Department of Energy, DE-AC05-00OR22725

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-19270-6>.

Correspondence and requests for materials should be addressed to T.L. or E.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025