

Physics-informed Kolmogorov-Arnold networks for multi-material elasticity problems in electronic packaging



Yanpeng Gong^{ID a}, Yida He^a, Yue Mei^{b,*}, Xiaoying Zhuang^{ID c,d,*}, Fei Qin^{ID a},
Timon Rabczuk^e

^a Institute of Electronics Packaging Technology and Reliability, Department of Mechanics, Beijing University of Technology, 100124, Beijing, China

^b Department of Engineering Mechanics, International Research Center for Computational Mechanics, State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, 116023, Dalian, China

^c Chair of Computational Science and Simulation Technology, Institute of Photonics, Department of Mathematics and Physics, Leibniz University Hannover, 30167, Hannover, Germany

^d Department of Geotechnical Engineering, College of Civil Engineering, Tongji University, 200092, Shanghai, China

^e Institute of Structural Mechanics, Bauhaus University Weimar, 99423, Weimar, Germany

ARTICLE INFO

Keywords:

Physics-Informed neural network
Kolmogorov-Arnold networks
Multi-material structures
Electronic packaging structures
Deep energy method.

ABSTRACT

This paper proposes a Physics-Informed Kolmogorov-Arnold Network for analyzing elasticity problems in multi-material electronic packaging structures. The method replaces traditional Multi-Layer Perceptrons with Kolmogorov-Arnold Networks within an energy-based Physics-Informed Neural Network framework. By constructing admissible displacement fields satisfying essential boundary conditions and optimizing network parameters through numerical integration, the proposed method effectively handles material property discontinuities. Unlike traditional methods that require domain decomposition and interface constraints for multi-material problems, Kolmogorov-Arnold Networks' trainable B-spline activation functions provide inherent piecewise characteristics. This capability stems from B-splines' local support, which enables effective approximation of discontinuities despite their individual smoothness. Consequently, this approach enables accurate approximation across the entire domain using a single network and simplifying the computational framework. Numerical experiments demonstrate that the proposed method achieves excellent accuracy and robustness in multi-material elasticity problems, validating its practical potential for electronic packaging analysis. Source codes are available at <https://github.com/yanpeng-gong/PIKAN-MultiMaterial>.

1. Introduction

Electronic packaging provides critical mechanical support, electrical interconnection, and thermal management for high-density integrated electronic devices [1]. Its reliability analysis involves challenges including multi-material systems, multi-scale phenomena, and interface discontinuities. Numerical simulation techniques, benefiting from easy implementation, high accuracy, and environment-independent operation, have become mainstream methods for electronic packaging reliability analysis, including finite element method (FEM) [2,3], boundary element method (BEM) [4,5], coupled algorithms [6,7], finite difference method (FDM) [8], and phase-field models [9,10], among others, with each method advancing through specific applications. However, facing the growing design complexity and computational efficiency demands of modern packaging structures, traditional methods still have inherent

* Corresponding authors.

E-mail addresses: yanpeng.gong@bjut.edu.cn (Y. Gong), meiyue@dlut.edu.cn (Y. Mei), zhuang@iop.uni-hannover.de (X. Zhuang).

limitations in computational cost, local field resolution, and physical modeling, creating urgent demand for more efficient and flexible computational frameworks.

The development of artificial intelligence (AI) technology and enhanced computational resources have provided opportunities for deep learning as a new approach to electronic packaging reliability problems. Researchers have successfully applied machine learning to construct surrogate models for rapid thermal-mechanical coupling behavior assessment [11], geometric parameter co-design [12], and process prediction [13], demonstrating significant computational efficiency advantages. However, these data-driven methods typically require substantial high-confidence training data; in scenarios with limited engineering data, their predictions may deviate from physical laws, affecting model generalization and physical consistency [14]. Therefore, combining machine learning with traditional computational methods offers tremendous potential for transforming modeling and simulation in engineering and scientific domains, opening new pathways for next-generation computer-aided analysis and design [15].

In recent years, deep learning methods have provided new numerical paradigms for solving solid mechanics problems by avoiding traditional mesh generation and improving computational efficiency [16,17]. To reduce neural networks' dependence on high-quality data, researchers embed physical laws into loss functions to enhance models through physics-informed constraints. Physics-Informed Neural Networks (PINNs), proposed by Raissi et al. [18] in 2019, became the pioneering work in this direction. The implementation of PINNs benefits from the mature application of automatic differentiation functionality in machine learning frameworks such as PyTorch and TensorFlow. Compared to traditional numerical algorithms, PINNs leverage the powerful approximation capabilities of neural networks to provide a new paradigm for solving complex problems that conventional numerical methods struggle to address [19]. Furthermore, PINNs demonstrate significant potential in handling high-dimensional problems, as their loss functions are typically constructed through Monte Carlo sampling, which is inherently an effective tool for high-dimensional integration. PINNs encompass two forms: strong-form original PINNs and energy-form PINNs. Samaniego et al. [20] introduced the Deep Ritz Method proposed by E and Yu [21] into computational mechanics, forming energy-form PINNs called the Deep Energy Method (DEM). Strong-form PINNs construct loss functions by directly combining PDEs through weighted residual methods, while DEM constructs loss functions based on variational principles such as the minimum potential energy principle in mechanics. The flexibility and potential of the PINN framework continue to drive its development and application in various cutting-edge engineering fields [16,22–27,64]. In electronic packaging structure analysis, although there are limited studies using traditional data-driven surrogate models based on physical simulation data [28], research directly applying PINNs to such simulations has rarely been reported.

Within the standard PINN framework, multi-layer perceptrons have been widely adopted due to their structural simplicity and universal approximation properties [29]. However, the framework is inherently flexible, and the approximation function is not limited to MLPs. Research has successfully extended to various architectures such as RBFs [30], Wavelets [31], and others. Recently, Kolmogorov-Arnold Networks (KANs), proposed by Liu et al. [32], employ trainable one-dimensional activation functions that simultaneously replace linear weights and provide nonlinear activation behavior, with nodes simplified to perform only input summation, have also been introduced into the PINN framework, rapidly spawning multiple improved studies focusing on fair comparisons with MLPs [33], architectural innovations [34,35], and other aspects. When dealing with PDE problems involving high-dimensional or complex solution spaces, MLPs often require extremely wide or deep networks to achieve the required accuracy, leading to an enormous number of parameters. This increases optimization difficulty and computational cost [36], while KANs provide an alternative parameterization approach through their spline-based learnable activation functions. Meanwhile, the spectral bias problem in MLP training may lead to insufficient learning capability for high-frequency components of PDE solutions, whereas the flexible function representation provided by KANs has been indicated to be more conducive to capturing multi-scale features of solutions and alleviating spectral bias phenomena [37,38]. This provides a new potential avenue for solving PDE problems with multi-scale characteristics or solutions with sharp transitions. It should be noted that the completeness of the universal approximation theory for classical KANs is still being refined, with only a few variants rigorously proven to date [39]. Nevertheless, the excellent computational performance demonstrated by this architecture in a wide range of scientific computing problems motivates us to further explore its application potential in PDE solving.

Electronic packaging structures are typically composed of heterogeneous materials with different properties, exhibiting multi-material, multi-interface structural characteristics. For multi-material structures with sharp interfaces, the meshfree nature of PINNs enables them to avoid the mesh distortion issues that may arise in FEM. When applying PINNs to solve multi-material problems, the current mainstream approach uses different neural networks to describe physical fields in different material domains. Diao et al. [40] proposed a strong-form PINNs framework for solving multi-material solid mechanics problems by partitioning the computational domain through domain decomposition, utilizing independent sub-networks to characterize mechanical behavior of different materials, and introducing interface regularization terms to handle contact relationships and discontinuities. Wang et al. [41] focused on variational problems and proposed a conservative energy method based on subdomain neural networks for solving heterogeneous material and complex geometry problems. However, PINNs based on subdomain partitioning have inherent limitations. To accurately describe interface contact relationships, additional hyperparameters must be introduced, which typically rely on empirical parameter tuning. Although research has explored hyperparameter optimization strategies [42,43], obtaining optimal parameter combinations remains challenging when solving specific problems. Furthermore, the domain decomposition method requires independent optimization of neural networks for each subdomain, which increases training complexity and often causes the optimization process to converge to local rather than global optima.

Although Physics-Informed Kolmogorov-Arnold Networks (PIKANs) have received increasing attention [32–35,44–46], the distinctive contribution of this study lies in developing a straightforward PIKAN solution framework specifically tailored for multi-material solid mechanics problems with various interface configurations commonly encountered in electronic packaging structures, integrating KANs with the DEM to replace traditional MLPs. To our knowledge, we present and validate for the first time this integrated

approach for electronic packaging multi-material structure analysis, focusing on the persistent challenge of modeling multi-material structures with inherent interface discontinuities. Within the PINN framework, the strong form has broader generality, but its application typically requires extensive hyperparameter tuning. In contrast, the Deep Energy Method, based on variational principles, transforms the governing equations into an extremum problem of energy functionals. This not only reduces the number of required hyperparameters, but also typically achieves higher computational accuracy and efficiency due to the reduced order of functional derivatives [47], though its generality is usually weaker than the strong form. For elastic analysis problems of multi-material structures, this study aims to develop a more concise solution paradigm, and therefore adopts the DEM with fewer hyperparameters. It should be noted that theoretical analysis of numerical convergence and stability for PINNs and their variational (energy-based) forms has made certain progress [48,49]. For instance, rigorous convergence proofs have been established for specific higher-order multi-scale PINN methods in composite material mechanics [50]. However, for the emerging KAN architecture and its applications in physics-informed learning, rigorous convergence theory for KAN/PIKAN architectures remains an active research frontier.

Based on the above research background, this paper employs the PIKAN method to solve multi-material problems in solid mechanics and extends its application to electronic packaging multi-material structure analysis. This work constructs admissible displacement fields [41,42] that strictly satisfy essential boundary conditions. The main contributions of our work include:

- **Development of PIKAN methodology** that replaces traditional MLPs with KANs in deep energy method, leveraging KANs' inherent piecewise function characteristics to naturally accommodate material property discontinuities.
- **Elimination of domain decomposition requirements** by using a single KAN to approximate displacement fields across the entire computational domain without subdomain partitioning or interface continuity constraints.
- **Comprehensive validation through diverse examples** including cantilever beams with different interface geometries and other electronic packaging models, demonstrating superior accuracy compared to traditional domain decomposition approaches.
- **Systematic investigation of numerical integration schemes and hyperparameter sensitivity analysis**, with complete algorithm implementation details and source code to be provided in our GitHub repository, facilitating practical application, comparison, and further development of PIKANs for multi-material problems.

This paper is organized as follows: [Section 2](#) presents the DEM for elasticity problems in solid mechanics and the construction method for admissible displacement fields. [Section 3](#) provides a detailed description of the proposed PIKAN method. [Section 4](#) presents specific strategies for applying PIKAN to solve multi-material problems. [Section 5](#) validates the effectiveness of the proposed method through multi-material model examples and applies it to electronic packaging multi-material structure analysis. Finally, [Section 6](#) summarizes the research conclusions.

2. Deep energy method-based computational solid mechanics

This section introduces the Deep Energy Method (DEM) for solving elasticity problems. Then we present the construction of admissible displacement fields that naturally satisfy essential displacement boundary conditions.

2.1. Deep energy method for solving elasticity problems

For elastic systems, based on the principle of minimum potential energy in variational elasticity theory, the Deep Energy Method (DEM) constructs the total potential energy functional of the system as a loss function for optimization [20]. For problems without body forces, the specific mathematical expression of this loss function is

$$\mathcal{L} = \Psi_{\text{in}} - \Psi_{\text{ex}} = \int_{\Omega} \frac{1}{2} \sigma_{ij} \varepsilon_{ij} d\Omega - \int_{\Gamma_{\sigma}} \bar{p}_i u_i d\Gamma \quad (1)$$

where Ψ_{in} and Ψ_{ex} represent the internal strain energy and external force potential energy, respectively; σ_{ij} and ε_{ij} denote the stress and strain tensors, respectively; u_i represents the i -th displacement component; and \bar{p}_i is the traction applied on the natural boundary. DEM shares similarities with FEM in adopting energy-based variational weak forms, but its key distinction lies in the fact that DEM directly optimizes the total potential energy functional, thereby eliminating the steps of preliminary derivation and discretization of the weak form of governing equations required by the FEM. For the electronic packaging multi-material structures of interest, we focus on mechanical responses dominated by external mechanical loads and material interface effects. The body force term is omitted in the energy functional construction, based on engineering simplification appropriate for the scale and loading characteristics of the studied problems [51,52]. The proposed framework can straightforwardly incorporate body force terms to handle more general scenarios when needed.

Among all kinematically admissible displacement fields $\mathbf{u}(\mathbf{x})$ that satisfy the essential boundary conditions, we seek the solution that minimizes the total potential energy functional

$$\mathbf{u}(\mathbf{x}) = \arg \min_{\mathbf{u}(\mathbf{x})} \mathcal{L}(\mathbf{u}(\mathbf{x})) = \arg \min_{\mathbf{u}(\mathbf{x})} \left\{ \int_{\Omega} \psi(\varepsilon(\mathbf{u}(\mathbf{x}))) d\Omega - \int_{\Gamma_{\sigma}} \bar{\mathbf{p}} \cdot \mathbf{u}(\mathbf{x}) d\Gamma \right\} \quad (2)$$

where $\mathbf{u}(\mathbf{x})$ is approximated through a neural network as $\mathbf{u}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}; \theta)$, $\mathbf{F}(\cdot)$ denotes the neural network mapping, and $\psi(\varepsilon(\mathbf{u}(\mathbf{x})))$ is the strain energy density function, which is a function of the strain tensor $\varepsilon(\mathbf{u}(\mathbf{x}))$.

2.2. Construction of admissible displacement fields

The DEM selects the displacement field that minimizes the potential energy among all admissible displacement fields. The key to realizing admissible displacement fields lies in constructing displacement trial functions that naturally satisfy essential boundary conditions and embedding this constraint directly into the network architecture to ensure that the displacement field consistently satisfies essential boundary conditions throughout the optimization process.

The general structure of an admissible displacement field can be expressed as

$$\mathbf{u}^{\text{pred}}(\mathbf{x}) = \mathbf{P}(\mathbf{x}) + \mathbf{D}(\mathbf{x}) \odot \mathbf{F}(\mathbf{x}; \theta) \quad (3)$$

where $\mathbf{P}(\mathbf{x})$ is the boundary function that satisfies essential boundary conditions. Note that for fixed boundaries, the prescribed essential boundary condition value is $\mathbf{0}$, i.e., $\mathbf{P}(\mathbf{x}) = \mathbf{0}$. $\mathbf{F}(\mathbf{x}; \theta)$ is the generalized network obtained through training by minimizing the potential energy loss function. \odot denotes element-wise multiplication. $\mathbf{D}(\mathbf{x})$ is the distance function, representing the shortest distance from the current coordinate point to the essential boundary Γ_u

$$\mathbf{D}(\mathbf{x}) = \min_{\mathbf{y} \in \Gamma_u} \|\mathbf{x} - \mathbf{y}\| \quad (4)$$

The boundary and distance functions have flexible implementation forms; in addition to neural networks, fitting methods such as radial basis functions can also be employed [41]. The key advantage of this construction is that when the coordinate point is located on the boundary Γ_u , the distance function $\mathbf{D}(\mathbf{x}) = \mathbf{0}$, and the admissible displacement field obtained from Eq. (3) naturally satisfies the enforced essential boundary conditions. In the validation examples presented in Section 5 of this study, given that the essential boundary geometries are relatively simple, $\mathbf{D}(\mathbf{x})$ consistently adopts the analytical distance function form [53], thereby avoiding additional errors introduced by neural network approximation while accurately enforcing essential boundary conditions. This method enables the admissible displacement field to naturally satisfy boundary constraints during generalized network training, without requiring the introduction of boundary condition penalty terms and associated hyperparameter tuning in its loss function. For detailed discussion on admissible displacement fields, please refer to [53].

The neural network parameters of the generalized network $\mathbf{F}(\mathbf{x}; \theta)$ are trained and optimized by minimizing the potential energy loss function

$$\begin{aligned} \mathbf{F}(\mathbf{x}; \theta) &= \arg \min_{\theta} \mathcal{L}(\mathbf{u}^{\text{pred}}(\mathbf{x}; \theta)) \\ &= \arg \min_{\theta} \left\{ \int_{\Omega} \psi(\epsilon(\mathbf{u}^{\text{pred}}(\mathbf{x}; \theta))) d\Omega - \int_{\Gamma_\sigma} \bar{\mathbf{p}} \cdot \mathbf{u}^{\text{pred}}(\mathbf{x}; \theta) d\Gamma \right\} \end{aligned} \quad (5)$$

Note that $\mathbf{F}(\mathbf{x}; \theta)$ represents the generalized network output, while the actual displacement field is given by Eq. (3).

3. Methodology of PIKAN

The previous section provided a detailed introduction to using DEM for solving elasticity problems. In this work, instead of using the MLP traditionally employed in DEM, we adopt KAN as the neural network architecture. This section first introduces the fundamental principles of KAN, then presents the PIKAN method and finally describes the sample point distribution strategies employed in the analysis.

3.1. Kolmogorov-Arnold networks (KANs)

KANs are based on the Kolmogorov-Arnold representation theorem and feature a fully connected structure. Traditional MLPs achieve layer-wise transformations by applying fixed nonlinear activation functions to linear combinations of inputs at each node, where weights and biases are learned parameters. During backpropagation training, MLPs update these parameters by computing gradients of the loss function with respect to weights and biases.

In contrast, KANs move activation functions from nodes to edges, where each edge contains learnable univariate functions parameterized as B-spline functions. Instead of applying fixed activation functions at nodes after linear transformations (as in MLPs), KANs apply learnable activation functions on each edge connection. Nodes in KANs perform only summation operations on the outputs from incoming edges. Their core advantage lies in adaptive learning capability for complex functional relationships and the resulting potentially higher parameter efficiency (i.e., achieving comparable accuracy with fewer parameters) [32]. It must be objectively noted that choosing KANs over MLPs involves trade-offs. In many benchmark tests, their accuracies are often comparable [33]. The primary advantage of KANs lies in potentially mitigating spectral bias and more efficiently learning high-frequency features [37,38], but the more complex gradient computations of B-spline functions during training increase computational cost [32]. The piecewise function approximation characteristics of KANs are better suited for handling solution field discontinuities at multi-material interfaces, which traditional MLPs struggle to naturally capture. The KAN architecture is illustrated in Fig. 1.

The core idea of the Kolmogorov-Arnold Representation Theorem (KART) is that any continuous multivariate function f defined on a bounded domain can be expressed as a composition of finite univariate continuous functions and addition operations

$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) = \sum_{h=1}^{2n+1} \Phi_h \left(\sum_{g=1}^n \phi_{h,g}(x_g) \right) \quad (6)$$

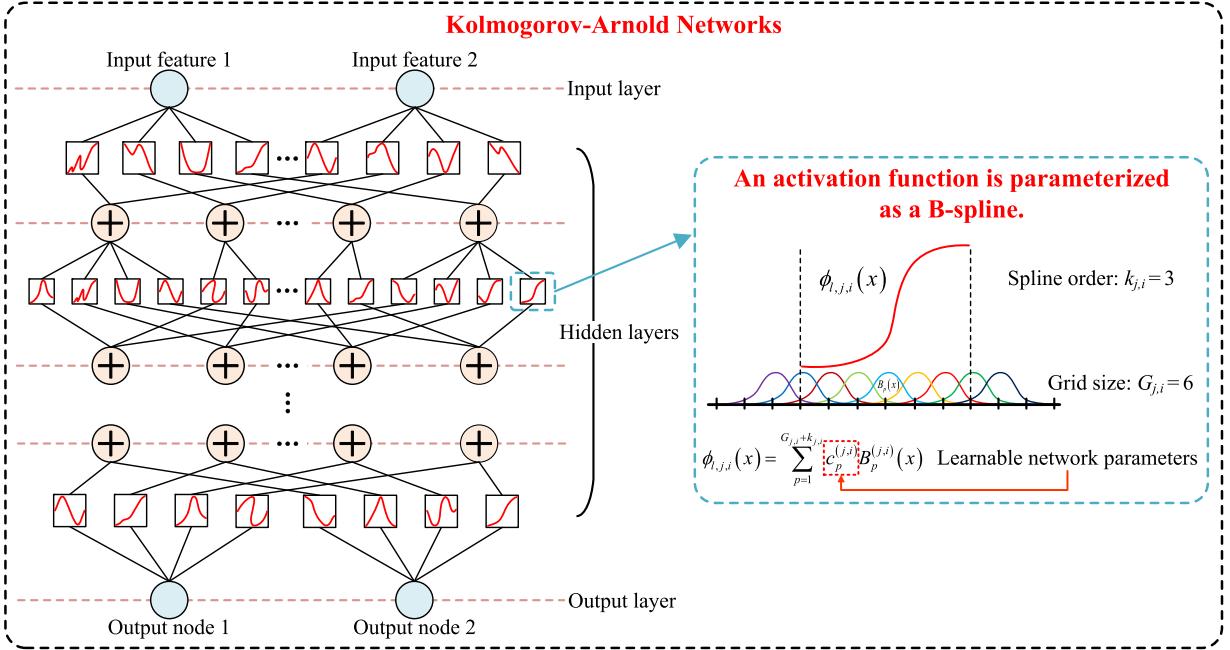


Fig. 1. Kolmogorov-Arnold network schematic (adapted and redrawn from the original KAN architecture[32]).

This provides theoretical inspiration for expressing multivariate functions through univariate function compositions. The design of KANs parameterizes the univariate functions ϕ and Φ in the theorem as learnable B-spline functions and extends them into nested, multi-layer network structures to enhance expressiveness. However, this deep KAN architecture goes beyond the scope of the original KART. As a new network paradigm, its theoretical validity is established and explored by the approximation theorem proposed and developed in the original paper itself [32]. This work focuses on the application of this architecture, and the following will introduce its network structure and implementation details in detail.

Define a KAN layer with N_{in} input neurons and N_{out} output neurons

$$\Phi = \{\phi_{h,g}\}, \quad g = 1, 2, 3, \dots, N_{\text{in}}; \quad h = 1, 2, 3, \dots, N_{\text{out}} \quad (7)$$

where $\phi_{h,g}$ has trainable parameters. The original Eq. (6) represents a combination of two KAN layers: the inner function with $N_{\text{in}} = n$, $N_{\text{out}} = 2n + 1$, and the outer function with $N_{\text{in}} = 2n + 1$, $N_{\text{out}} = 1$. By stacking more KAN layers, deeper KANs can be obtained. The shape of a KAN is represented by an integer array $[N_0, N_1, N_2, \dots, N_L]$, where N_i is the number of neurons in the i -th layer. Between the l -th and $(l+1)$ -th layers, there are $N_l N_{l+1}$ activation functions. The activation function connecting the i -th neuron (l, i) in the l -th layer and the j -th neuron ($l+1, j$) in the $(l+1)$ -th layer is denoted as

$$\phi_{l,j,i}, \quad l = 0, 1, \dots, L-1; \quad i = 1, 2, \dots, N_l; \quad j = 1, 2, \dots, N_{l+1} \quad (8)$$

The activation value of the $(l+1)$ -th layer is computed as

$$x_{l+1,j} = \sum_{i=1}^{N_l} \phi_{l,j,i}(x_{l,i}), \quad j = 1, 2, \dots, N_{l+1} \quad (9)$$

In vector form, this can be written as

$$\mathbf{x}_{l+1} = \begin{bmatrix} \sum_{i=1}^{N_l} \phi_{l,1,i}(x_{l,i}) \\ \sum_{i=1}^{N_l} \phi_{l,2,i}(x_{l,i}) \\ \vdots \\ \sum_{i=1}^{N_l} \phi_{l,N_{l+1},i}(x_{l,i}) \end{bmatrix} = \Phi_l(\mathbf{x}_l) \quad (10)$$

where Φ_l represents the function matrix of the l -th KAN layer.

Since the activation functions are parameterized as B-spline functions, each activation function $\phi_{l,j,i}$ can be expressed as

$$\phi_{l,j,i}(x) = \sum_{p=1}^{G_{j,i}+k_{j,i}} c_p^{(j,i)} B_p^{(j,i)}(x) \quad (11)$$

where $G_{j,i}$ and $k_{j,i}$ are the grid size and spline order for the activation function connecting neuron (l, i) to neuron $(l+1, j)$, $c_p^{(j,i)}$ are the trainable B-spline coefficients, and $B_p^{(j,i)}(x)$ are the corresponding B-spline basis functions.

Table 1

Trainable parameters of KAN layers.

Parameter	Type	Quantity per layer	Description
$c_p^{(j,i)}$	B-spline coefficients	$N_l \times N_{l+1} \times (G_{j,i} + k_{j,i})$	Coefficients for B-spline basis functions
$m_{j,i}$	Scaling factors	$N_l \times N_{l+1}$	Element-wise scaling factors for activation functions
$w_{j,i}$	Linear weights	$N_l \times N_{l+1}$	Weights for the nonlinear activation function $\sigma(\cdot)$

The total number of trainable parameters in layer l is

$$N_{\text{params},l} = \sum_{i=1}^{N_l} \sum_{j=1}^{N_{l+1}} (G_{j,i} + k_{j,i}) \quad (12)$$

To enhance the function approximation capability, scaling coefficients are introduced for each activation function [34]. For each activation function $\phi_{l,j,i}$, a corresponding scaling coefficient $m_{j,i}$ is defined. The scaling coefficient matrix \mathbf{M}_l is

$$\mathbf{M}_l = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,N_l} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,N_l} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N_{l+1},1} & m_{N_{l+1},2} & \cdots & m_{N_{l+1},N_l} \end{bmatrix} \quad (13)$$

The scaled activation functions become

$$\phi_{l,j,i}^*(x) = m_{j,i} \cdot \phi_{l,j,i}(x) \quad (14)$$

Additionally, a linear transformation matrix \mathbf{W}_l and nonlinear activation function $\sigma(\cdot)$ are introduced to further enhance the network's expressive and fitting capabilities. The final output of the KAN layer is computed as

$$x_{l+1,j} = \sum_{i=1}^{N_l} m_{j,i} \cdot \phi_{l,j,i}(x_{l,i}) + \sum_{i=1}^{N_l} w_{j,i} \cdot \sigma(x_{l,i}), \quad j = 1, 2, \dots, N_{l+1} \quad (15)$$

where $w_{j,i}$ are elements of the weight matrix \mathbf{W}_l . The introduction of $\sigma(\cdot)$ enables the network to capture smoother solutions, complementing the B-spline functions. The scaling coefficients $m_{j,i}$ and weights $w_{j,i}$ are both trainable parameters optimized during training.

Following common practice to simplify implementation and training, this work adopts a uniform configuration where all activation functions within each KAN layer share the same grid size $G_{j,i}$ and B-spline order $k_{j,i}$. The trainable parameters in KAN layers are given in Table 1.

For a KAN network with L layers, given an input vector $\mathbf{x}^{(0)} \in \mathbb{R}^{N_0}$, the forward propagation through the network is computed sequentially as

$$\mathbf{x}^{(L)} = \Phi_{L-1}^* (\Phi_{L-2}^* (\dots \Phi_0^* (\mathbf{x}^{(0)}) \dots)) \quad (16)$$

where $\Phi_l^*(\mathbf{x}^{(l)}) = \Phi_l(\mathbf{x}^{(l)}) \odot \mathbf{M}_l + \mathbf{W}_l \cdot \sigma(\mathbf{x}^{(l)})$, $0 \leq l \leq (L-1)$ and $\mathbf{x}^{(L)}$ is the network's final output.

3.2. Physics-informed Kolmogorov-Arnold networks (PIKANs)

As discussed in the Introduction, PIKAN has emerged as a widely explored new paradigm in the field of physics-informed machine learning [32–35,44–46]. This section presents the specific PIKAN framework adopted in this study. As illustrated in Fig. 2, KAN replaces the MLP traditionally used in the Deep Energy Method for computational solid mechanics.

Since KANs employ B-spline activation functions defined on preset grids ([0,1] or [-1,1]), and the algorithm takes sample point coordinates within the computational domain as neural network inputs, while actual computational models have arbitrary geometric dimensions that may exceed the preset grid range, normalization of network inputs is necessary. Geometric normalization is crucial for ensuring training stability and convergence speed [44,45,54], and its theoretical necessity has been thoroughly explored through rigorous analyses based on frameworks such as Neural Tangent Kernel (NTK), particularly in studies of Chebyshev KANs [45,46]. This work implements this standardization strategy in specific problems based on these established theoretical foundations.

As shown in the upper left corner of Fig. 2, we consider a 2D geometric domain with characteristic length L and characteristic height H , where $L \geq H$, and point o is the coordinate origin. For any point $A(x_A, y_A)$ within the domain, we perform coordinate normalization

$$\tilde{x}_A = \frac{x_A}{L}, \quad \tilde{y}_A = \frac{y_A}{L} \quad (17)$$

The normalized coordinates serve as actual inputs to the KAN, ensuring B-spline activation functions operate effectively within their definition domain.

To mitigate training instability caused by KAN layer outputs potentially exceeding the definition domain of subsequent spline functions, this work adopts a validated approach of applying tanh activation functions to constrain intermediate layer outputs [33,45]

$$x_{l+1,j}^{\text{new}} = \tanh(x_{l+1,j}) = \tanh\left(\sum_{i=1}^{N_l} m_{j,i} \cdot \phi_{l,j,i}(x_{l,i}) + \sum_{i=1}^{N_l} w_{j,i} \cdot \sigma(x_{l,i})\right) \quad (18)$$

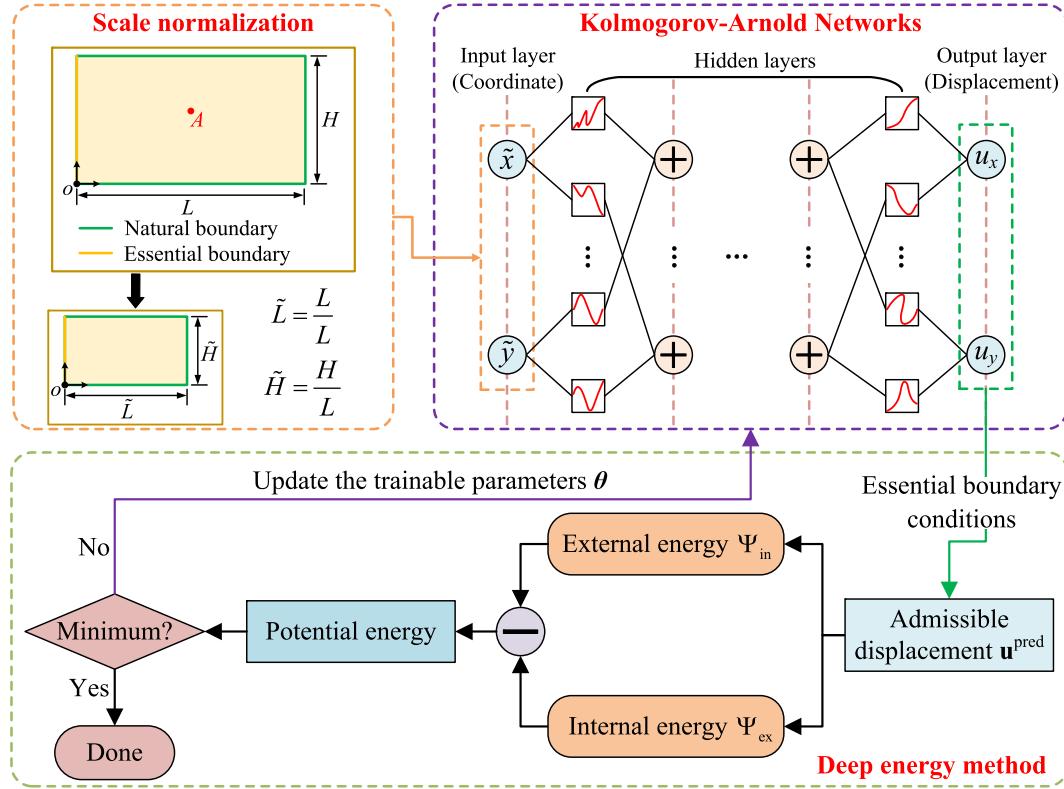


Fig. 2. PIKAN schematic: Deep energy method using KAN instead of MLP.

for $j = 1, 2, \dots, N_{l+1}$ and layers $l = 0, 1, \dots, L - 2$. Note that the final layer ($L - 1$) does not require tanh activation, as the output displacement field represents physical quantities with inherently unbounded ranges that should not be artificially constrained.

3.3. Sample point distribution strategies

We employ various numerical integration schemes to approximate the energy function, including Monte Carlo, trapezoidal, and Simpson integration. For these schemes, sample points follow uniform distribution throughout the computational domain, maintaining consistent spacing to form a regular grid. Additional boundary points can be incorporated along curved boundaries to better capture domain geometry, as illustrated by the purple points in Figs. 3 and 4.

Additionally, we implement two triangular mesh-based integration methods for computing strain energy in the loss function. The first method employs triangular integration rules by subdividing the computational domain into triangular meshes and using triangle centroids as sample points for neural network training. Each triangle's area serves as the control area for its corresponding centroid. The triangular integration formula is

$$\int_{\Omega} F(\mathbf{x}) d\Omega = \sum_{k=1}^N F(\mathbf{x}_k) \cdot S_k \quad (19)$$

where N is the number of triangular elements, \mathbf{x}_k is the centroid of the k -th triangle, S_k is its area, and F is the integrand function.

Fig. 3 illustrates the triangular integration scheme, where purple points represent the original uniform grid, red points are triangle centroids serving as neural network training points, and the blue triangle area in Fig. 3(b) shows the control area S_k for sample point \mathbf{x}_k . The mesh density should be chosen based on problem complexity.

The second method, shown in Fig. 4, uses the same integral formula as Eq. (19) but employs the uniform grid points (purple) as training points rather than centroids. Each triangle is subdivided by connecting its centroid to the midpoints of its edges, creating three equal areas distributed to adjacent training points. For instance, in Fig. 4(b), the green area is assigned to the bottom-right sample point. Each training point's control area comprises one-third of all surrounding triangles. This approach is termed Delaunay integration, where the red-outlined regions in Fig. 4(c) represent the control area S_k for point \mathbf{x}_k .

The choice of sampling strategy and numerical integration scheme directly affects loss function computation. For classical numerical integration schemes based on regular sample point distributions (Monte Carlo, trapezoidal, Simpson), this study employs their standard-defined weight coefficients to compute integral approximations. For Delaunay triangulation-based numerical integration,

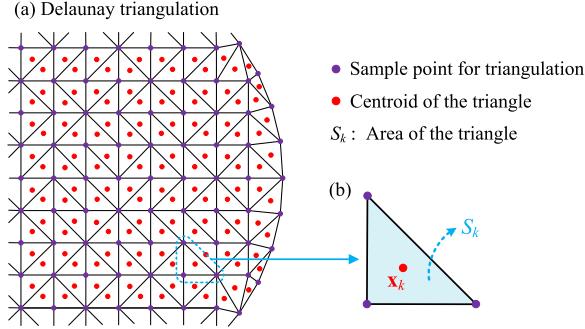


Fig. 3. Triangular integration scheme: sample points and control areas.

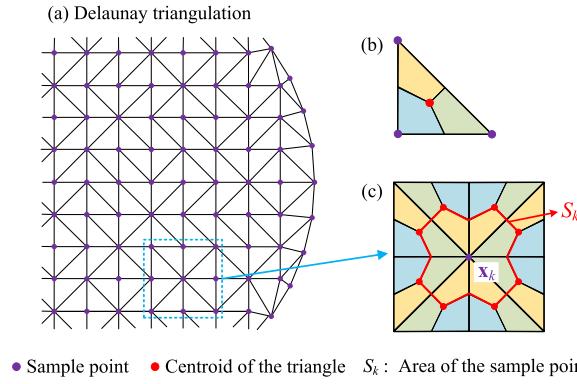


Fig. 4. Delaunay integration scheme: sample points and control areas.

depending on different sample points (triangle centroids or uniformly distributed points), corresponding area coefficients are selected to approximate the energy integral. Section 5 systematically compares computational results across different schemes.

4. Multi-material elasticity problems

Recent work has demonstrated KAN's effectiveness in solving PDEs [32] and heterogeneous problems [34]. Based on spline interpolation theory, B-splines naturally exhibit piecewise polynomial characteristics with controllable smoothness properties. In multi-material elasticity problems, displacement fields remain continuous across material interfaces while stress fields (and consequently displacement gradients) may exhibit discontinuities due to material property changes. The piecewise nature of B-spline functions makes KAN theoretically well-suited for capturing such solution behaviors without explicit domain decomposition.

Therefore, when solving multi-material problems, this work abandons the traditional domain decomposition approach that assigns separate neural networks to different material regions. Instead, we employ a single KAN network to predict displacement fields across the entire computational domain without requiring subdomain division. We construct a unified energy-based loss function for the complete multi-material model and optimize all trainable KAN parameters simultaneously, obtaining a neural network capable of predicting physically consistent displacement fields that naturally accommodate material property variations.

Fig. 5 illustrates the PIKAN framework for solving multi-material elasticity problems. As shown in the upper left, sample point coordinates are generated throughout different material domains according to geometric features and material distribution. The training point arrangement follows the integration schemes introduced in Section 3.3, with arrangements varying by integration method. Simultaneously, boundary condition points are generated, and all normalized coordinates serve as neural network inputs. It should be noted that since we employ a single KAN network to predict the entire computational domain, no interface continuity penalty terms are required in the loss function, eliminating the need for explicit interface point generation.

A Kolmogorov-Arnold network $\mathbf{F}(\mathbf{x}; \theta)$ is constructed to predict displacement fields across all material domains simultaneously. Material properties are incorporated during energy calculation, where strain energy density is computed using the appropriate material constants for each domain based on point locations. This approach constructs a unified admissible displacement field $\mathbf{u}^{\text{pred}}(\mathbf{x})$ that naturally accommodates material discontinuities through KAN's piecewise function characteristics.

The strain energies from all material domains are summed to obtain total strain energy, and subtracting external work yields the potential energy of the entire computational domain. Using the principle of minimum potential energy, an optimizer minimizes the

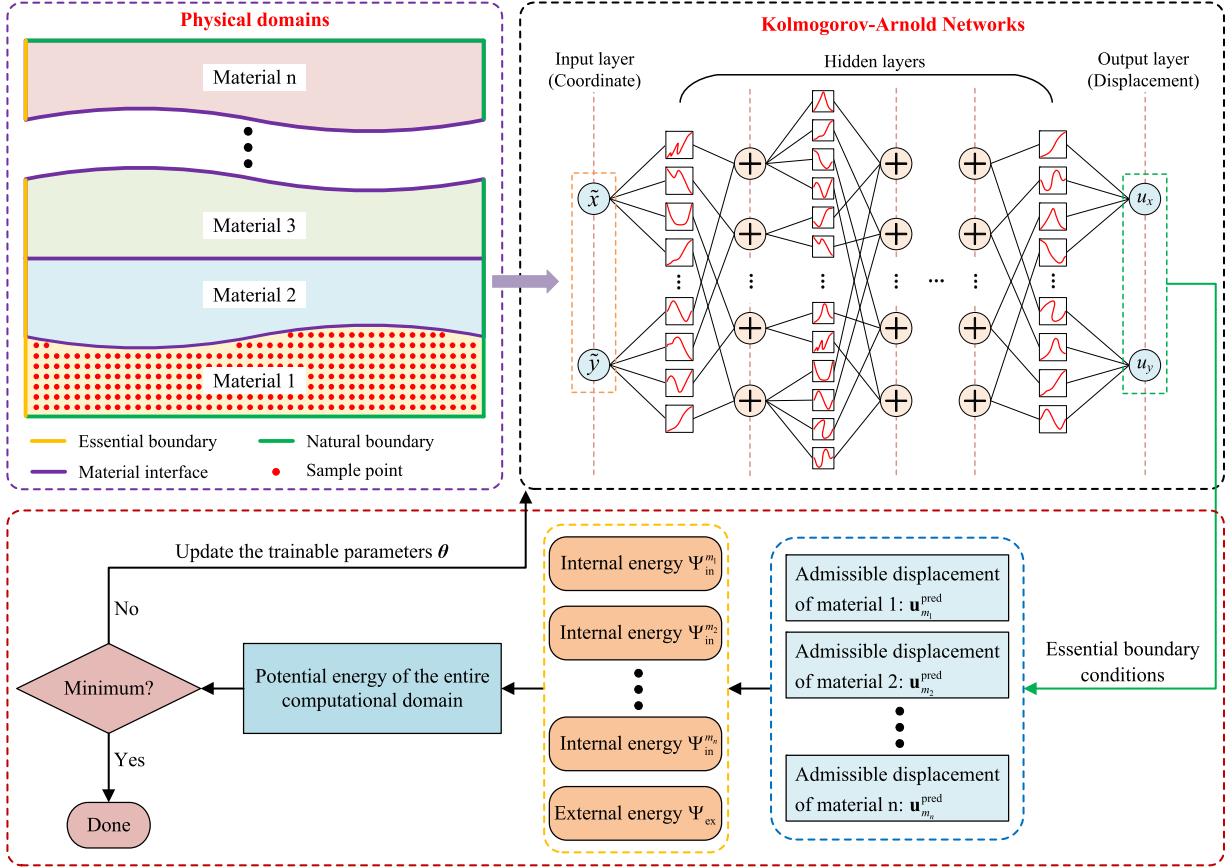


Fig. 5. PIKAN flowchart for multi-material elasticity problems.

loss function with respect to KAN parameters θ

$$\mathbf{F}(\mathbf{x}; \theta) = \arg \min_{\theta} \mathcal{L}_{\text{PIKAN}} = \arg \min_{\theta} \left\{ \sum_{i=1}^n \Psi_{in}^{m_i} - \Psi_{ex} \right\} \quad (20)$$

until convergence is achieved. $\Psi_{in}^{m_i}$ represents the strain energy in the i -th material domain and Ψ_{ex} is the total external work. The complete PIKAN solution procedure is summarized in [Algorithm 1](#) in the [Appendix A](#).

5. Numerical examples

This section demonstrates the accuracy of PIKAN through several numerical examples. The first two examples explore PIKAN's application to multi-material problems, while the latter two analyze electronic packaging structures. Unless otherwise specified, all geometric dimensions in the following examples are given in millimeters (mm). To assess accuracy and convergence, we consider two error metrics: absolute error and relative L_2 norm error. The absolute error is defined as

$$\text{Abs} = |\mathbf{u}_{\text{pred}} - \mathbf{u}_{\text{ref}}| \quad (21)$$

where \mathbf{u}_{pred} represents the admissible displacement solution constructed from the neural network output and \mathbf{u}_{ref} is the reference solution. This metric evaluates prediction accuracy at each sample point. The relative L_2 norm error is defined as

$$L_2 = \frac{\sqrt{\sum_{i=1}^N |u_{\text{pred}}^i - u_{\text{ref}}^i|^2}}{\sqrt{\sum_{i=1}^N |u_{\text{ref}}^i|^2}} \quad (22)$$

where N represents the total number of sample points. Since analytical solutions are unavailable for these examples, we use FEM results as reference solutions to validate the proposed method's accuracy. In the FEM setup, four-node bilinear quadrilateral elements with reduced integration (CPS4R) were employed, and mesh convergence studies were conducted for all analyses to ensure the validity of the reference solutions.

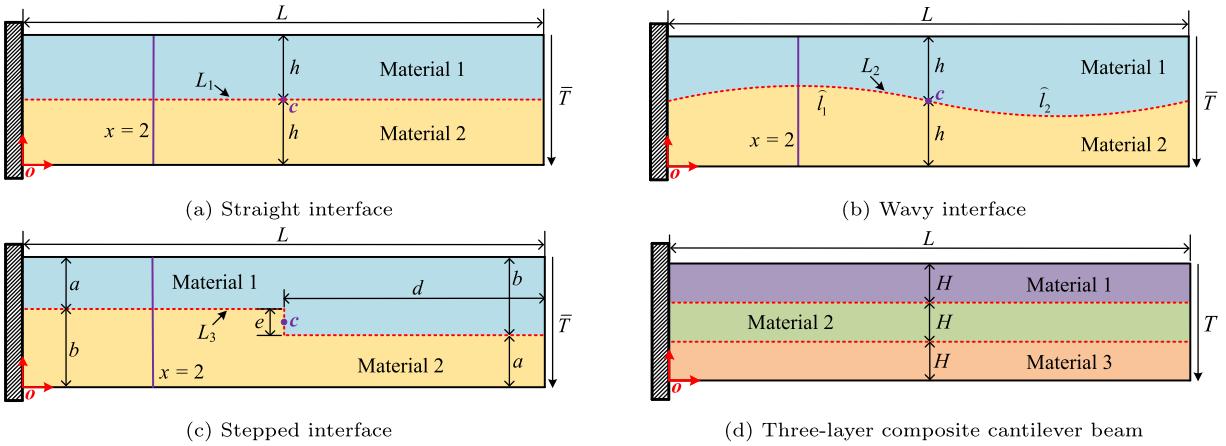


Fig. 6. Cantilever beam models with different material interface geometries.

All examples employ a consistent computational framework. We use the L-BFGS optimizer with initial learning rate 0.001 and constrained to a maximum of 20 iterations per optimization step. The KAN network architectural hyperparameters and optimizer settings employed in this study primarily follow the experience ranges widely adopted in the original KAN paper and recent physics-informed learning applications [32,34]. These were manually fine-tuned through iterative adjustments to achieve accurate solution predictions. Complete hyperparameter specifications are provided for each example to ensure reproducibility. The total trainable parameter counts reported for all examples in this section refer exclusively to the KAN network parameters as defined in [Table 1](#), and do not include any parameters from boundary condition treatment functions. Neural network training was performed on an NVIDIA GeForce RTX 4060 Laptop GPU.

5.1. Cantilever beam analysis across material interfaces

We investigate PIKAN's performance through three cantilever beam bending problems with different material interface geometries, as shown in [Fig. 6](#). Each beam has length $L = 8$ and height $2h = 2$, with the coordinate origin at the bottom-left corner. The straight interface ([Fig. 6\(a\)](#)) is located at distance $h = 1$ from the bottom edge. The wavy interface ([Fig. 6\(b\)](#)) consists of two arc segments \hat{l}_1 and \hat{l}_2 that are antisymmetric about the beam's center point $c(4, 1)$. Arc \hat{l}_1 belongs to a circle centered at $(2, -8)$ and arc \hat{l}_2 to a circle centered at $(6, 10)$, both with radius $r = \sqrt{85}$. The stepped interface ([Fig. 6\(c\)](#)) has geometric parameters $a = 1.2$, $b = 0.8$, $d = 4$, and $e = 0.4$. As shown in [Fig. 6](#), all models are subjected to a shear traction $\bar{T} = 6$ N/mm applied at the right boundary. The left end is fully constrained.

The material parameters are: Material 1 with Young's modulus $E_1 = 8,500$ MPa and Material 2 with $E_2 = 43,000$ MPa, both having Poisson's ratio $\nu = 0.3$. The method uses a KAN structure of $[2, 5, 5, 5, 2]$ with grid size 10, B-spline order 3, and grid range $[0, 1]$, totaling 1050 trainable parameters. The material interfaces are depicted as red dashed lines L_1 , L_2 , and L_3 in [Fig. 6](#), representing straight, wavy, and stepped configurations for the three cantilever beam models.

[Fig. 7](#) illustrates the distribution of training points used in the analysis of different models. Green points on the right boundary represent natural boundary condition points, with 101 points uniformly distributed at spacing $dy = 0.02$ (this boundary point distribution is consistent across all cantilever beam models). For the centroid sampling strategy, the interior point counts are: 79,200 points for the straight interface beam ([Fig. 7\(a\)](#)), 80,278 points for the wavy interface beam ([Fig. 7\(c\)](#)), and 79,163 points for the stepped interface beam ([Fig. 7\(e\)](#)). In contrast, the uniform sampling shown in [Figs. 7\(b\)](#), [7\(d\)](#), and [7\(f\)](#) employ identical distribution patterns with 40,501 interior points for each beam configuration.

[Fig. 8](#) shows the evolution of loss functions during training for different numerical integration methods: triangular, Simpson, trapezoidal and Monte Carlo schemes. The triangular scheme uses the point distribution from [Fig. 7\(a\)](#), [\(c\)](#), [\(e\)](#), while Simpson, trapezoidal and Monte Carlo schemes use the uniform distribution from [Fig. 7\(b\)](#), [\(d\)](#), [\(f\)](#). All schemes exhibit convergent behavior, with the triangular integration method generally achieving faster convergence and lower final energy values. Despite different interface geometries, all models demonstrate similar convergence trends with increasing training iterations, indicating the robustness of the PIKAN approach across various geometric configurations.

[Fig. 9](#) presents a detailed comparison of displacement fields between PIKAN results (using triangular integration) and FEM reference solutions. To provide a quantitative and intuitive assessment of PIKAN's prediction accuracy, [Fig. 10](#) presents pointwise absolute error plots for cantilever beam models with different material interface configurations. PIKAN accurately reproduces both u_x and u_y displacement components across all interface configurations, demonstrating excellent agreement with FEM solutions. The method successfully captures complex displacement near material interfaces, particularly evident in the wavy and stepped configurations where interface geometry creates significant stress concentrations and discontinuities in material properties.

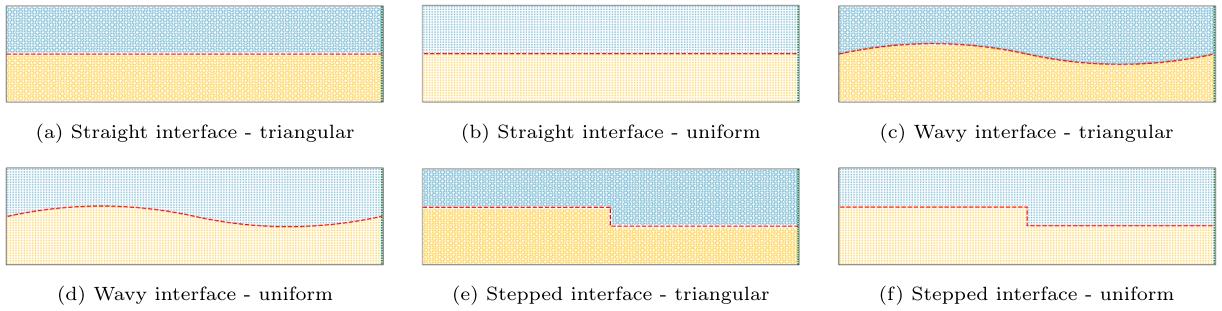


Fig. 7. Training point distributions: (a,c,e) Triangular scheme; (b,d,f) Uniform scheme.

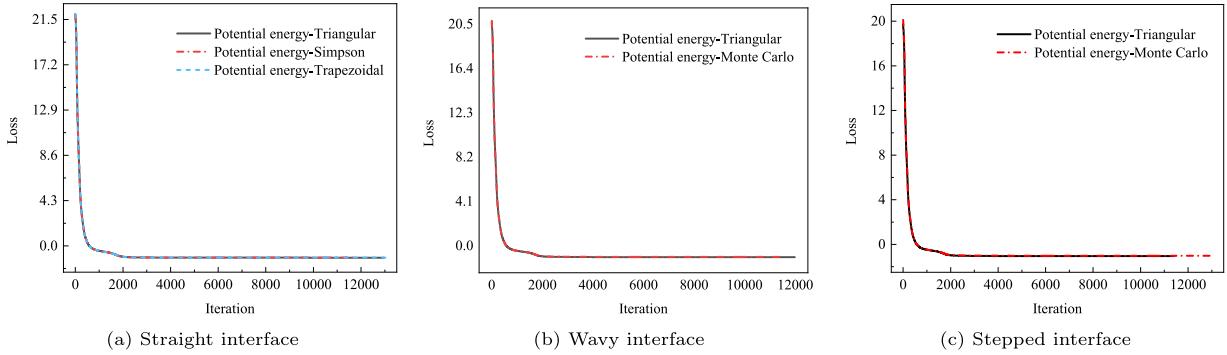


Fig. 8. Evolution of energy-based loss functions for different interface geometries.

To further validate PIKAN's computational accuracy, we compare results with the Conservative Energy Neural Network (CENN) [41] method using domain decomposition. CENN applies traditional DEM with separate MLPs for different subdomains, requiring interface continuity conditions that introduce additional penalty terms and hyperparameters in the loss function. The CENN loss function takes the form

$$\mathcal{L}_{\text{CENN}} = \sum_{i=1}^n \Psi_{\text{in}}^i - \Psi_{\text{ex}} + \beta \sum_{j=1}^{N_{\text{inter}}} \left\| \mathbf{u}_{\text{inter}}^+(\mathbf{x}_{\text{inter},j}) - \mathbf{u}_{\text{inter}}^-(\mathbf{x}_{\text{inter},j}) \right\|^2 \quad (23)$$

Compared to single material domain energy-based loss functions, Eq. (23) includes additional interface continuity penalty terms, where $\mathbf{u}_{\text{inter}}^+$ and $\mathbf{u}_{\text{inter}}^-$ represent the admissible displacement fields from different material subnetworks at the interface, and β is the penalty parameter controlling interface condition enforcement. The penalty parameter is calculated using

$$\beta = -\lambda \cdot \ln [\tanh(N_{\text{inter}} / N_{\text{dom}})] \quad (24)$$

where $\lambda = 1000$, N_{inter} is the number of interface sample points, and N_{dom} is the total number of domain sample points. CENN employs identical MLP architectures for each material domain, each consisting of 5 hidden layers with 20 neurons per layer, yielding a total of 3564 trainable parameters. The configuration uses tanh activation functions and the same triangular integration scheme for training sample point generation.

Fig. 11 compares u_x displacement profiles along $x = 6$ for all three cantilever beam models. PIKAN results using different integration schemes are compared with CENN predictions and FEM reference solutions. The figures show PIKAN's superior performance across all interface geometries, closely matching FEM solutions with excellent accuracy and stability. Among the integration schemes, triangular, Simpson and trapezoidal methods show the best agreement with FEM solutions, while Monte Carlo integration exhibits slightly reduced precision due to its stochastic nature. In contrast, CENN produces noticeable displacement discontinuities at material interfaces due to its subdomain decomposition approach. Under complex interface geometries (particularly wavy and stepped configurations), it shows computational instability with significant localized errors, highlighting the inherent challenges of enforcing interface continuity conditions through penalty parameter tuning in domain decomposition methods.

We investigate PIKAN's performance in computing von Mises stress for cantilever beam models by analyzing stress profiles along the vertical line $x = 2$, $0 \leq y \leq 2$ (as shown in Fig. 6). PIKAN results are compared with CENN predictions and FEM reference solutions. Fig. 12 demonstrates that PIKAN predictions exhibit consistency with FEM solutions across all interface geometries. As shown in Fig. 12, stress oscillations are observed, which can be attributed to two main factors. First, energy-based methods primarily minimize potential energy through displacement field optimization rather than directly enforcing pointwise stress equilibrium conditions. Second, stress solutions are computed as derivatives of displacement fields combined with constitutive relations, making them highly sensitive to small displacement errors, particularly near boundaries where geometric constraints and material property variations

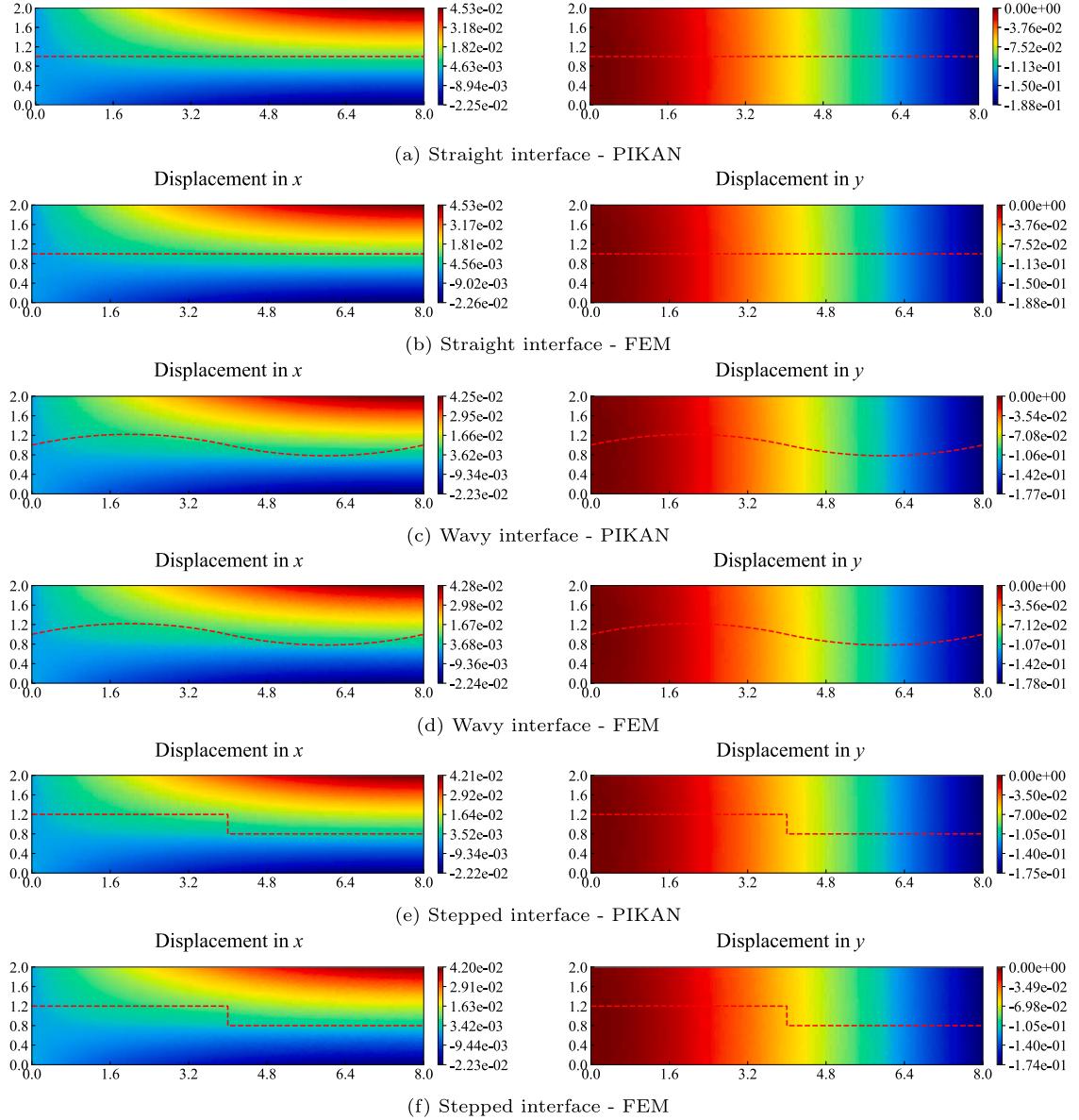


Fig. 9. Displacement fields u_x and u_y for different interface geometries: (a), (c), (e) PIKAN predictions; (b), (d), (f) FEM solutions.

create stress concentrations. Consequently, the numerical accuracy of stress field is inherently lower than that of displacement [51]. Since PIKAN achieves superior displacement accuracy compared to CENN's subdomain approach in multi-material problems, it also yields more accurate stress results. In contrast, CENN exhibits significant stress prediction errors, particularly at material interfaces, due to displacement discontinuities arising from its subdomain decomposition approach. The use of independent neural networks for different material regions results in pronounced stress fluctuations at material boundaries. Across all interface geometries (straight, wavy, and stepped configurations), CENN demonstrates computational instability in non-interface regions with substantial localized errors, confirming the limitations observed in the displacement analysis. Moreover, a comprehensive hyperparameter sensitivity analysis is presented in [Appendix B](#), revealing that optimal performance requires joint optimization of grid size, B-spline order, and network architecture rather than independent tuning.

To systematically evaluate PIKAN's capability in handling multi-material interfaces and high parameter contrast problems, we further analyze a three-layer composite cantilever beam model beyond the previously discussed examples. As shown in [Fig. 6\(d\)](#), the beam has length $L = 8$ and height $3H = 1.8$. The elastic moduli of the three materials are set with significant contrast: $E_1 = 3,000$ MPa, $E_2 = 20,000$ MPa, and $E_3 = 100,000$ MPa, with Poisson's ratio $\nu = 0.3$ for all materials, the shear traction $T = 5$ N/mm. This design creates a rigorous test scenario combining multiple heterogeneous interfaces with extreme parameter jumps to validate the method's performance.

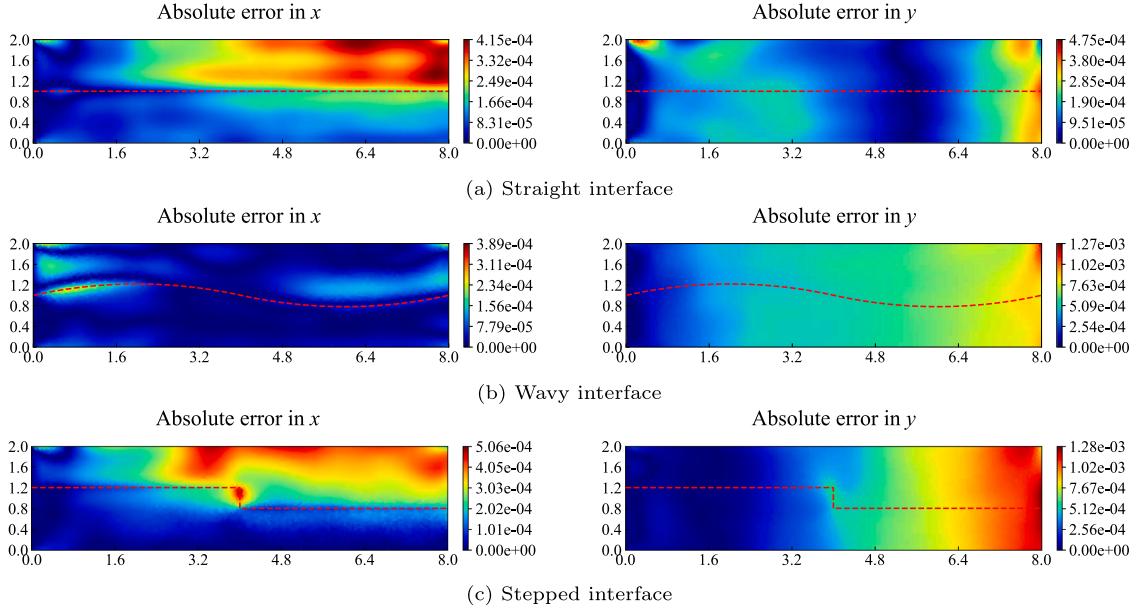
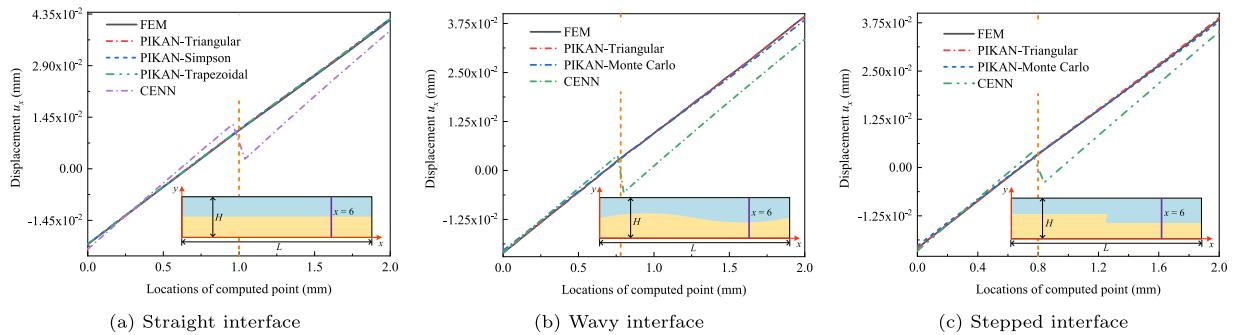
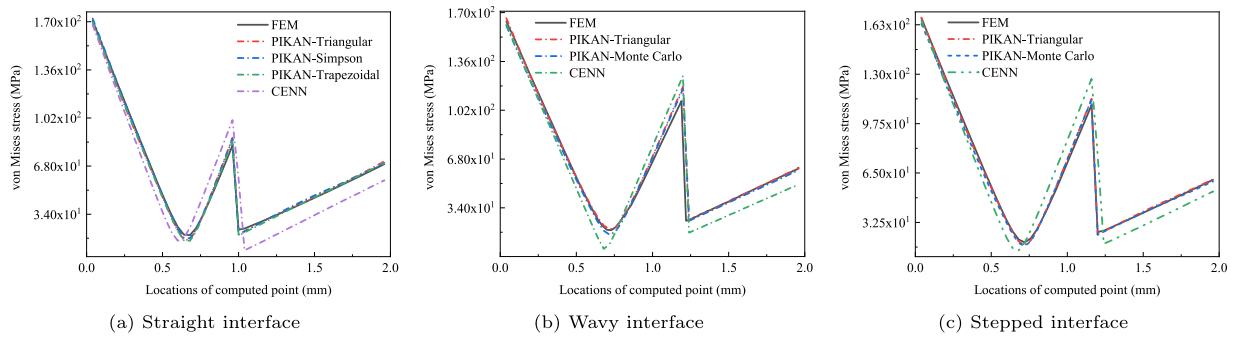
Fig. 10. Absolute error contours between PIKAN and FEM solutions for u_x and u_y .Fig. 11. Displacement (u_x) at $x = 6$ for different interface geometries: PIKAN vs. CENN and FEM solutions. The orange dashed line denotes the interface position.Fig. 12. Von mises stress at $x = 2$: PIKAN vs. CENN and FEM solutions.

Fig. 13(a), (b) compares PIKAN-predicted displacement fields u_x and u_y with reference FEM solutions, showing excellent agreement. More quantitatively, the pointwise absolute error distributions shown in **Fig. 13(c)** demonstrate that even at material interfaces and in regions where elastic modulus spans two orders of magnitude, the maximum error remains well-controlled. This directly validates PIKAN's accurate resolution capability for high-contrast parameters and interface discontinuities.

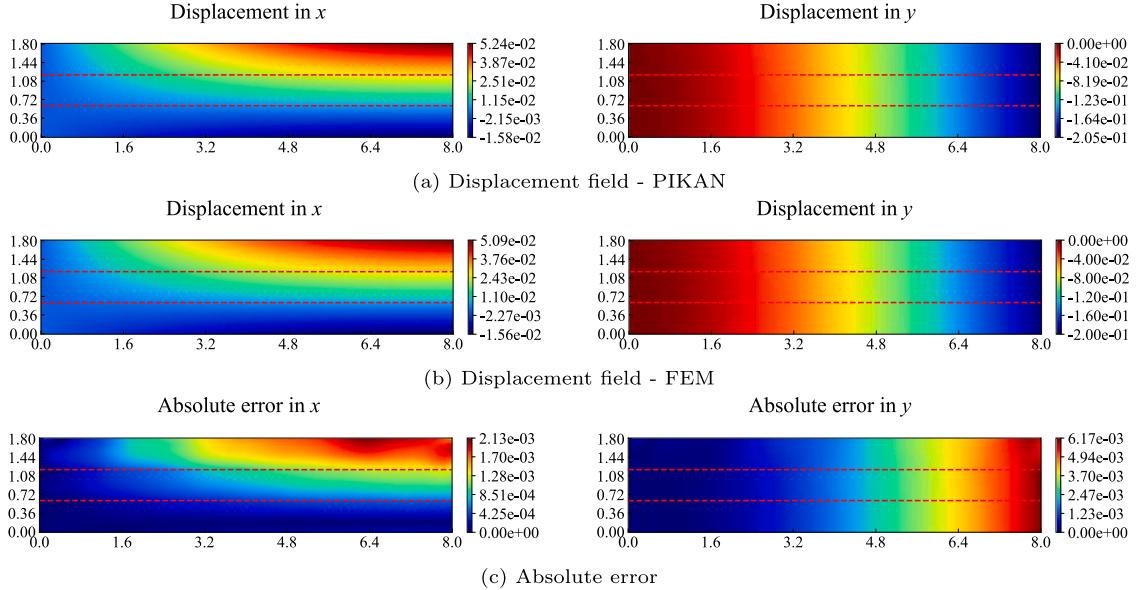


Fig. 13. Displacements and absolute errors (x - and y -directions) in three-layer composite cantilever beam model.

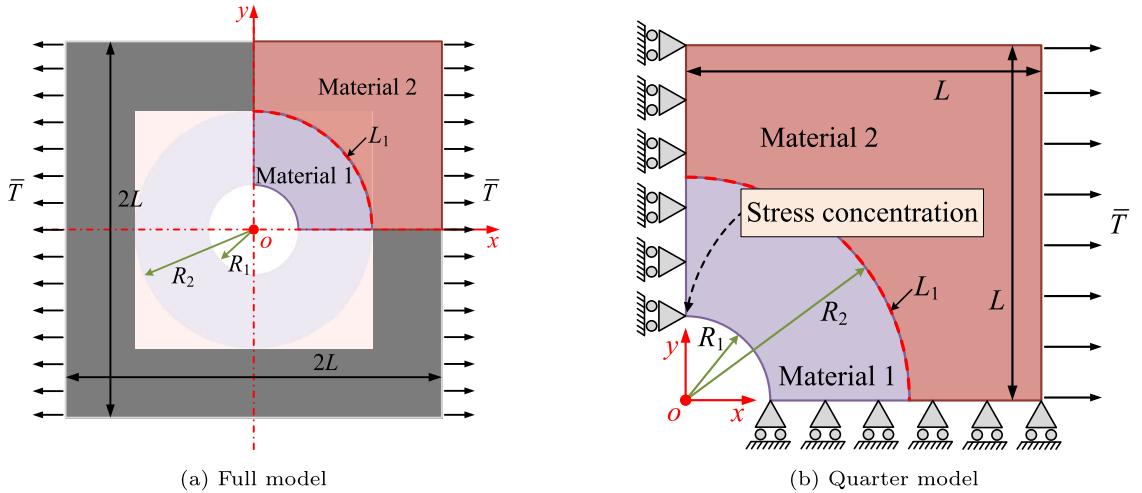


Fig. 14. Heterogeneous plate with central hole model.

5.2. Heterogeneous plate with central hole

We apply PIKAN to analyze a multi-material square plate with a central hole under tensile loading, as shown in Fig. 14(a). Due to symmetry, only a quarter-plate model with dimensions $L \times L$, where $L = 21$, is analyzed. The coordinate origin o is located at the bottom-left corner of the quarter model. As illustrated in Fig. 14(b), the central circular hole has radius $R_1 = 5$ with traction-free boundaries. The material interface (indicated by the red dashed line) is a circular boundary with radius $R_2 = 13$, where material 1 occupies the inner region and material 2 occupies the outer region. A uniform tensile load $\bar{T} = 120$ N/mm is applied to the right boundary of the quarter model. Symmetry boundary conditions are imposed along the left and bottom edges. The heterogeneous plate consists of two materials: Material 1 with Young's modulus $E_1 = 10,000$ MPa and Poisson's ratio $\nu_1 = 0.3$, and Material 2 with $E_2 = 1,000$ MPa and $\nu_2 = 0.4$.

The KAN architecture for this model is configured as [2, 5, 5, 5, 2] with grid size 20, B-spline order $k = 3$, and grid range $[0, 1]$. The neural network contains 1750 trainable parameters. Fig. 15 shows the sample point distribution patterns applied to the multi-material plate model. Purple points represent sample points in Material 1 (inner region), maroon points represent sample points in Material 2 (outer region), and green points along the loaded boundaries are used to apply natural boundary conditions, totaling 211 points. Triangle centroid points comprise 22,786 interior points in Material 1 and 61,624 interior points in Material 2. For uniformly distributed sample points, Material 1 contains 11,574 interior points and Material 2 contains 31,121 interior points. This difference

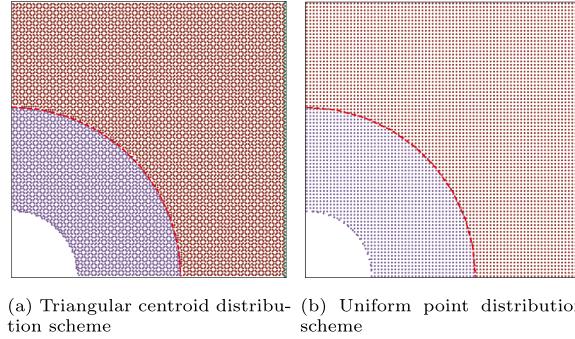


Fig. 15. Sample point distribution for heterogeneous plate model.

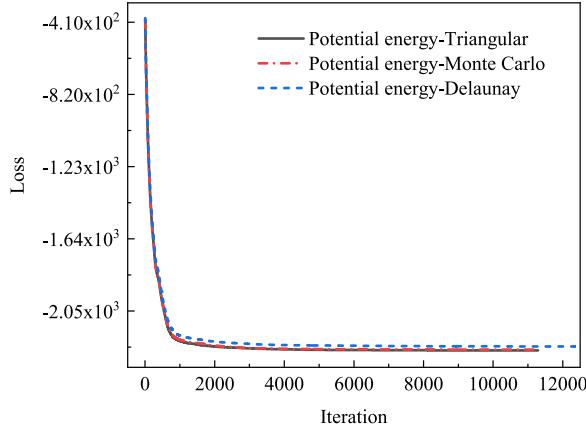


Fig. 16. Loss function evolution during training for heterogeneous plate model.

in point density allows for investigating the effect of sampling resolution on solution accuracy. For comparison, results from CENN are also presented. In CENN, each material domain employs a separate MLP with 6 hidden layers of 30 neurons each, totaling 4802 trainable parameters. The training sample points follow the triangular distribution pattern shown in Fig. 15(a).

Fig. 16 demonstrates the evolution of energy-based loss functions during training for the heterogeneous plate model. As the neural network undergoes iterative optimization, the total potential energy gradually converges to its minimum value, validating the principle of minimum potential energy. All numerical integration schemes exhibit similar convergence behavior, but the Delaunay scheme exhibits marginally weaker convergence in the loss compared to the other two schemes. Fig. 17 presents a detailed comparison between PIKAN predictions (using triangular integration) and FEM reference solutions for displacement magnitude u_{mag} (calculated according to Eq. (25)).

$$u_{\text{mag}} = \sqrt{u_x^2 + u_y^2} \quad (25)$$

The results demonstrate excellent agreement between PIKAN and FEM solutions, with PIKAN successfully capturing the displacement distribution around the central hole and the smooth transition across the material interface. The displacement patterns clearly show the influence of material property differences, with higher displacements occurring in the softer outer material (Material 2). Fig. 18 shows the pointwise absolute error contours for displacement components u_x and u_y compared to FEM reference solutions. The error distribution reveals that PIKAN maintains high accuracy throughout most of the domain, with maximum errors typically below 9.57×10^{-3} .

To examine solution behavior across the entire domain, displacement profiles along the diagonal line $\theta = 45^\circ$ (from the hole edge to the outer boundary, indicated by the red line) are analyzed. Fig. 19 shows the comparison of displacement magnitude u_{mag} along this diagonal line. The results demonstrate that triangular integration and Monte Carlo integration schemes produce solutions that agree well with the FEM reference, maintaining accuracy across both material regions and the interface transition. However, the Delaunay integration scheme exhibits significantly inferior accuracy. This scheme assigns point weights as one-third of the associated triangle area sum, and this discrete integral representation using uniformly sampled points introduces integration errors that prevent the loss function from accurately approximating the true potential energy. Consequently, the network trained on this basis produces relatively high solution errors. This result confirms a critical principle: in variational principle-based physics-informed neural networks, the accuracy of the numerical integration scheme used in the loss function directly affects the solver's final output accuracy. In contrast, CENN predictions show significant deviations from the reference solution throughout the domain, failing to accurately capture the

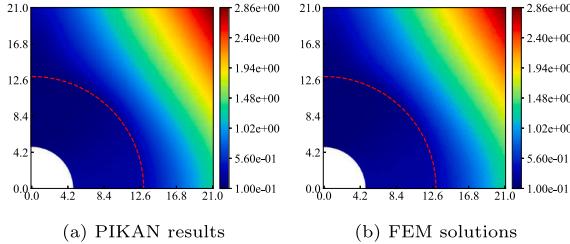


Fig. 17. Displacement u_{mag} for heterogeneous plate model: PIKAN vs. FEM.

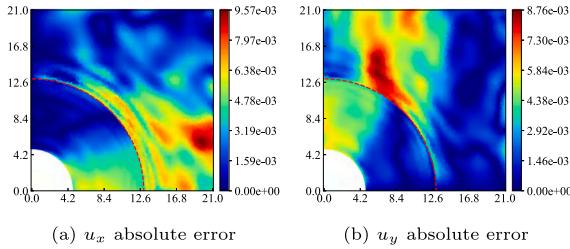


Fig. 18. Absolute error contours for displacement fields in heterogeneous plate model.

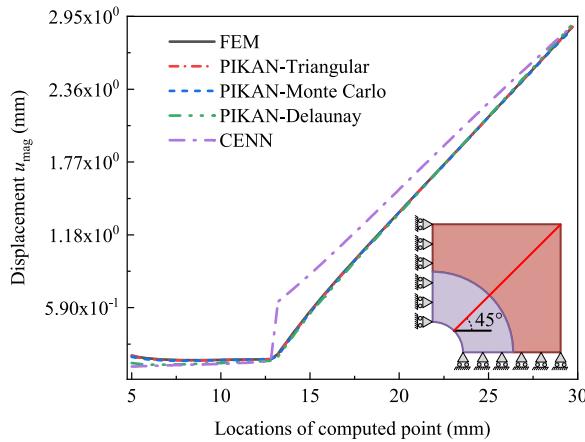


Fig. 19. u_{mag} along diagonal line.

displacement field distribution patterns. This demonstrates the inherent challenges of domain decomposition approaches that rely on penalty terms for interface continuity, particularly for problems with circular interfaces where geometric complexity compounds convergence difficulties.

Next, we evaluate PIKAN's performance by comparing the stress results from PIKAN, FEM and CENN along the vertical line $x = 0$ (from $y = 5$ to $y = 21$). This analysis line extends from the hole boundary through both material regions to the outer boundary, providing insight into stress prediction accuracy across material interfaces and high-gradient regions. The results are presented in Fig. 20. The analysis reveals that triangular integration achieves the highest accuracy, closely converging to the FEM reference solution throughout the domain. In contrast, Monte Carlo and Delaunay integration schemes show poor prediction performance in the Material 1 domain (high-stiffness inner region), where stress concentrations are more pronounced and strain energy density varies dramatically. The numerical approximation errors of these integration schemes are substantial in this region, preventing the loss function from providing sufficiently accurate gradients for network training, thereby degrading final solution accuracy. This differential performance suggests that stress calculations are particularly sensitive to integration scheme choice in regions with steep stress gradients and high material property contrasts. CENN predictions follow the general trend of the reference solution but with substantially larger errors, highlighting the challenges of maintaining stress accuracy in domain decomposition approaches that rely on penalty terms for interface continuity.

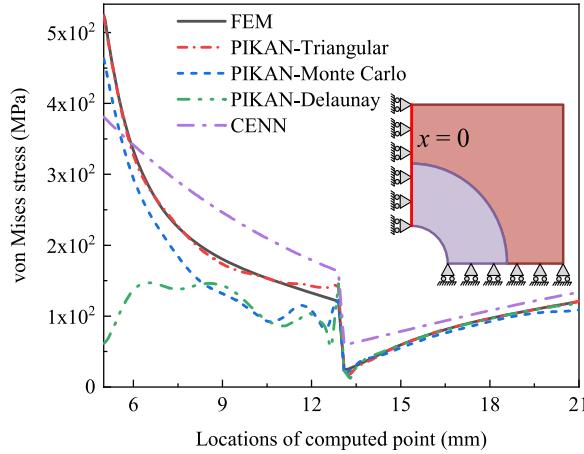
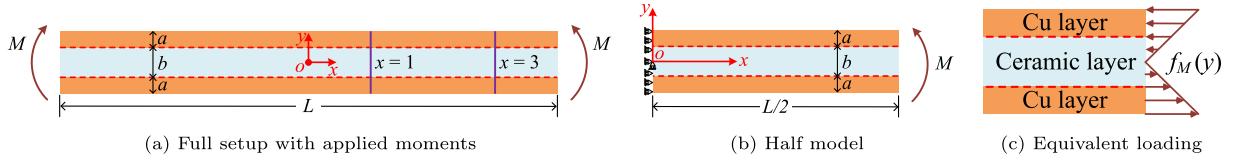
Fig. 20. Stress along the line $x = 0$.

Fig. 21. DBC substrate bending problem.

5.3. PIKAN application to DBC substrate structure

PIKAN is applied to analyze DBC (Direct Bonding Copper) substrates featuring multi-material domains. DBC substrates are widely used in power electronics, high-power modules, and aerospace applications due to their excellent thermal conductivity and mechanical strength. Material property differences and thermal/mechanical stresses during manufacturing or operation cause these substrates to warp, affecting assembly performance and compromising equipment functionality [5].

A simplified DBC substrate warping problem is modeled as shown in Fig. 21(a). The substrate has total length $L = 8$ with an Al_2O_3 ceramic core of thickness $b = 0.48$ sandwiched between upper and lower copper layers, each with thickness $a = 0.26$. The substrate experiences equal and opposite bending moments $M = 100 \text{ N}\cdot\text{mm}$ at both ends, creating pure bending deformation. Due to symmetry about the midplane, only the right half of the model requires analysis, as shown in Fig. 21(b). The coordinate origin is located at the center of the substrate cross-section. Symmetry boundary conditions are applied at the left boundary, while the bending moment is applied as an equivalent distributed force at the right end. The equivalent force distribution, illustrated in Fig. 21(c), is given by

$$f_M(y) = \frac{12M}{h^3}y = 1200y \quad \text{for } y \in (-0.5, 0.5) \quad (26)$$

where $h = 1.0$ is the total substrate thickness. This linear distribution ensures that the resultant force is zero and the moment equals the applied bending moment M . The DBC substrate consists of copper layers with Young's modulus $E_{\text{Cu}} = 128,000 \text{ MPa}$ and Poisson's ratio $\nu_{\text{Cu}} = 0.34$, and an alumina ceramic core with $E_{\text{Al}_2\text{O}_3} = 270,000 \text{ MPa}$ and $\nu_{\text{Al}_2\text{O}_3} = 0.28$.

A KAN with architecture [2, 5, 5, 5, 2] is employed to predict the displacement field (u_x, u_y) with grid size 10, B-spline order $k = 2$, and grid range $[-1, 1]$. The neural network contains 980 trainable parameters. Fig. 22 shows the training point distribution schemes for the DBC substrate model. Due to the layered structure of the DBC substrate, sample points are distributed across three material regions: the upper copper layer, the ceramic core, and the lower copper layer. In the centroid sampling scheme, the two copper layers contain a total of 166,400 sample points (83,200 points per layer) and the ceramic core contains 150,400 sample points. In the uniform sampling strategy, the copper layers contain a total of 84,906 sample points (42,453 points per layer) and the ceramic core contains 76,095 sample points. For both distribution schemes, the right-end natural boundary condition is discretized using 201 uniformly distributed sample points to accurately represent the linear force distribution given by Eq. (26).

Fig. 23 illustrates the evolution of the potential energy loss function for the DBC substrate warpage model. Fig. 24 presents a comprehensive comparison of displacement field solutions between PIKAN predictions (using Simpson integration) and FEM reference solutions for the DBC substrate bending problem. The results demonstrate excellent agreement between PIKAN and FEM in displacement solutions. The horizontal displacement u_x (Figs. 24(a) and 24(b)) shows the expected symmetric distribution about the neutral surface, with the upper and lower surfaces deflecting in opposite directions. For the vertical displacement u_y (Figs. 24(c) and 24(d)), both methods capture the characteristic bending deformation pattern with maximum displacements occurring at the loaded end. The displacement magnitude u_{mag} (Figs. 24(e) and 24(f)) provides a comprehensive view of the overall deformation. Fig. 25 quantifies the prediction accuracy through absolute error contours for both displacement components. The error analysis reveals that

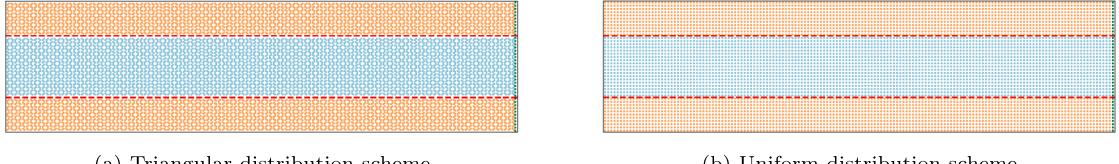


Fig. 22. Training point distribution schemes for DBC substrate model.

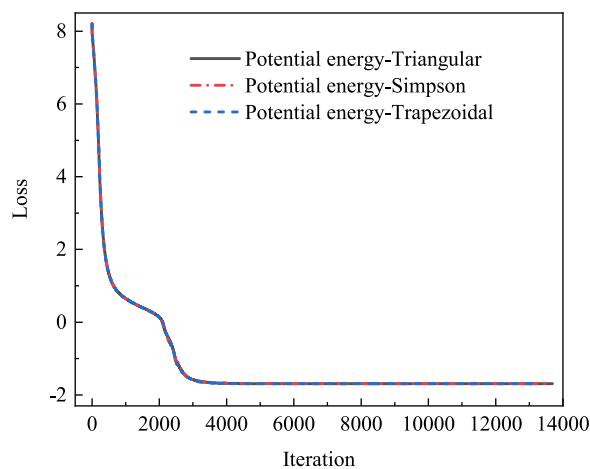


Fig. 23. Evolution of loss functions during training for the DBC substrate model.

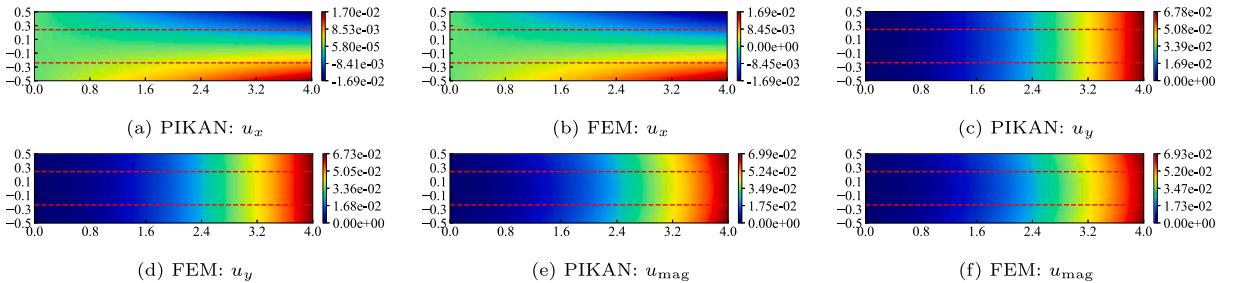


Fig. 24. Displacement field solutions for DBC substrate model: PIKAN vs. FEM.

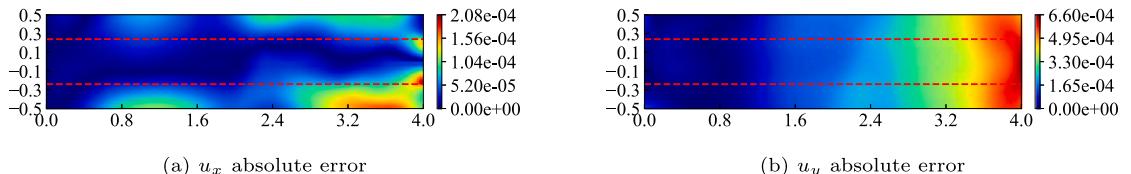


Fig. 25. Absolute error contours for displacement fields in DBC substrate model.

PIKAN maintains high accuracy throughout the domain, with maximum absolute errors typically below 6.6×10^{-4} for both u_x and u_y components. The error distribution shows slightly higher values near the material interfaces and the loaded boundary, which is expected due to the discontinuities in material properties and the complex stress distributions in these regions. Overall, the error magnitudes are several orders smaller than the displacement values, confirming the excellent accuracy of the PIKAN approach for multi-material structural analysis.

To further demonstrate the effectiveness of PIKAN for multi-material problems, we analyze displacement distributions at critical locations within the DBC substrate. The material interfaces L_1 and L_2 represent the upper and lower boundaries between the copper layers and the ceramic core, respectively. These interfaces are critical for understanding stress transfer and potential delamination sites in DBC substrates. Fig. 26 presents displacement distributions at different locations. Fig. 26(a) shows the horizontal displacement u_x along the upper material interface L_1 (copper-ceramic boundary), where the displacement varies linearly due to the pure bending

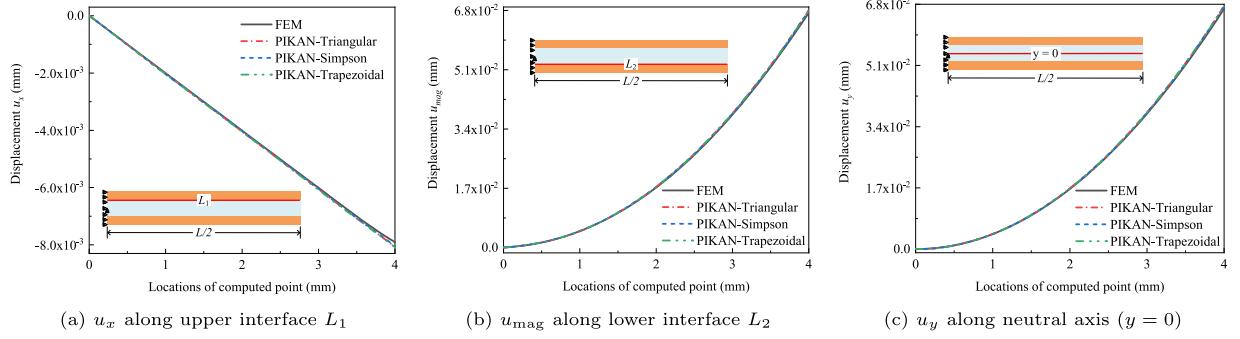


Fig. 26. Displacement distributions at critical locations in DBC substrate.

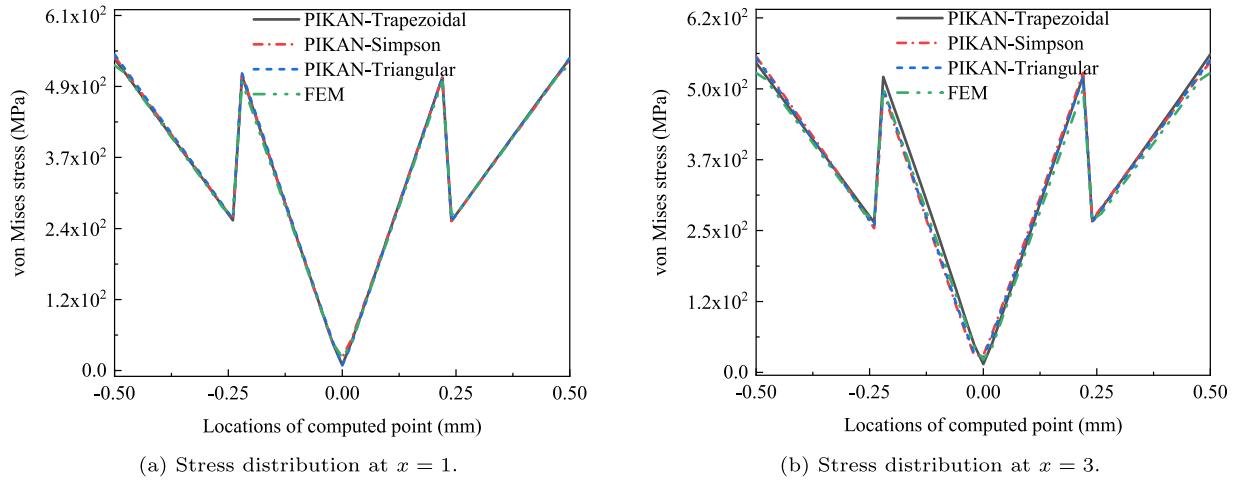


Fig. 27. Stress distributions along vertical lines in the DBC substrate.

deformation. Fig. 26(b) presents the displacement magnitude u_{mag} along the lower material interface L_2 , Fig. 26(c) shows the vertical displacement u_y along the neutral surface (line $y = 0$, where $0 \leq x \leq 4$) within the ceramic layer. The results demonstrate that PIKAN solutions achieve excellent agreement with FEM reference solutions across all analyzed locations. All numerical integration schemes exhibit comparable accuracy levels, indicating the robustness of the PIKAN approach. The smooth displacement distributions at material interfaces confirm that PIKAN effectively handles multi-material problems without requiring explicit interface continuity constraints.

Figs. 27(a) and (b) show stress distributions along lines $x = 1$ and $x = 3$ (where $-0.5 \leq y \leq 0.5$, as shown in Fig. 21) in the DBC substrate. These lines capture stress variations in the interior region and near the loaded end. PIKAN stress results agree well with FEM reference solutions at both locations. The stress distributions show typical multi-material behavior with stress jumps at material interfaces ($y = \pm 0.24$) due to different elastic moduli. Significant stress variations occur in the stiffer ceramic core, while copper layers show lower stress levels. Among these integration schemes, all computational results demonstrate similar accuracy and converge well to the FEM solutions. This indicates that PIKAN can effectively capture stress variations near material interfaces even without employing higher-order polynomial approximation methods (such as Simpson's rule), confirming the effectiveness and robustness of PIKAN for solving multi-material problems.

5.4. PIKAN application to TGV-Cu structure

Glass-substrate interposer packaging enables high-density interconnection of multi-chip systems through Through Glass Via (TGV) technology. TGV structures provide mechanical support and electrical interconnection via electroplated copper pillars fabricated through electroplating processes. After electroplating, residual copper layers on the glass surface must be removed through chemical mechanical polishing (CMP) to achieve smooth interfaces and proper electrical isolation [55]. PIKAN is employed to analyze TGV-Cu structures during the CMP process, as shown in Fig. 28. Four model configurations are investigated: (1) single TGV-Cu structure (Fig. 28(a)); (2) and (3) TGV-Cu structures with different interface roughness patterns (Figs. 28(b) and (c)); and (4) dual-via TGV-Cu structure (Fig. 28(d)).

The geometric dimensions are $L = 1$, $a = 0.4$, and $b = 0.3$. Interface roughness is modeled as semicircular protrusions with radii $R_1 = 0.1$ and $R_2 = 0.05$ at the glass-copper interface. The dual-via configuration features two copper vias with center-to-center spacing

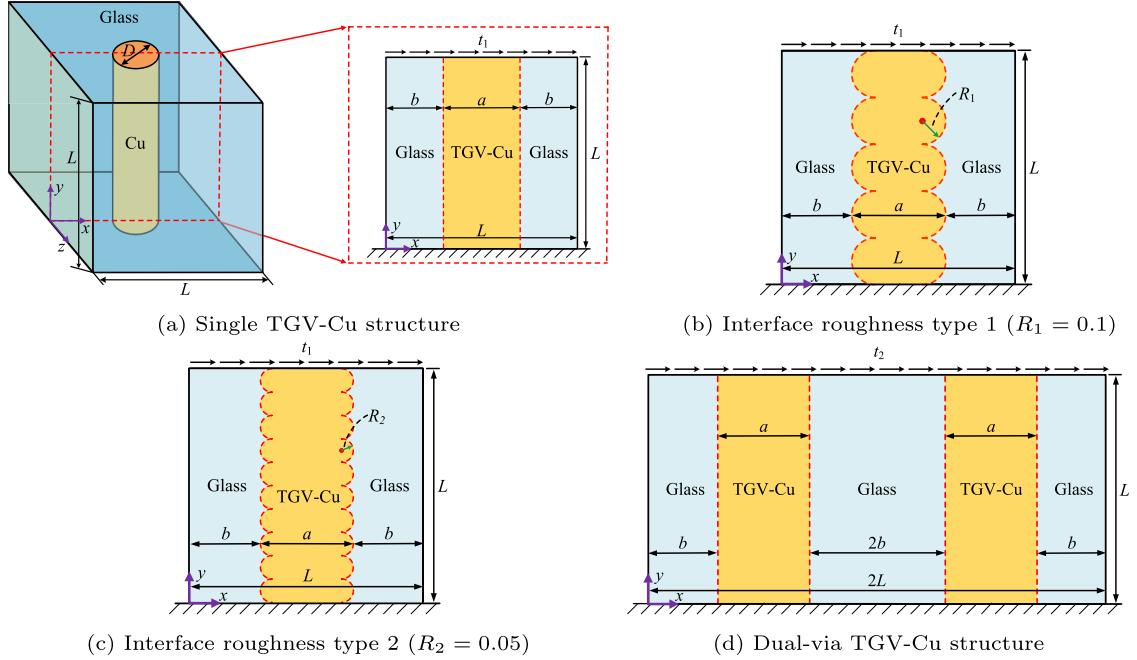


Fig. 28. TGV-Cu structure models for CMP analysis: (a) Baseline single via; (b,c) Single via with interface roughness variations; (d) Dual-via configuration.

Table 2
Material parameters for TGV-Cu models.

Material	Young's Modulus E (MPa)	Poisson's Ratio ν
Glass substrate	77,000	0.24
Cu (electroplated)	150,000	0.3

of 0.6. Boundary conditions simulate the CMP process: the bottom surface is fixed, while horizontal tractions are applied at the top surface. The applied tractions are $t_1 = 200$ N/mm for single-via structures and $t_2 = 250$ N/mm for the dual-via structure. Material properties are listed in **Table 2**.

The KAN architecture for this problem maintains consistency with previous examples: [2, 5, 5, 2] structure with grid size 20, B-spline order 3, grid range [0, 1] and 1750 trainable parameters. **Fig. 29** shows the training sample point distribution schemes for all TGV-Cu structure models. For centroid point sampling, the domain point counts are: single structure - 79,200 points; roughness type 1 - 80,094 points; roughness type 2 - 80,064 points; and dual-via structure - 158,400 points. For uniform distribution sampling patterns, the distributions are: single structure - 40,401 points; both roughness models - 40,761 points each; and dual-via structure - 80,601 points. Natural boundary condition training points are distributed as 201 points for single TGV-Cu structures and 401 points for the dual-via structure, corresponding to the different boundary lengths requiring load application.

Fig. 30 shows the evolution of loss functions for different TGV-Cu structure models during neural network training. Despite varying geometric complexities and structural arrangements, all models exhibit similar convergence behavior with consistent energy reduction trends. The convergence patterns demonstrate PIKAN's robustness across different geometric configurations. All integration schemes achieve stable convergence, with triangular integration typically showing slightly faster convergence. The decreasing loss function indicates the system approaching an equilibrium state. When the loss function stabilizes after a certain number of iterations (approximately 2000 iterations in this case), it indicates the system is approaching a stationary point of the potential energy functional. In practical applications, an early stopping criterion can be adopted: training can be terminated when the relative change in loss function over consecutive iterations falls below a preset tolerance ϵ , thereby balancing accuracy and computational efficiency. In this work, to demonstrate the complete convergence process and ensure result stability, we chose to continue iterations to a sufficiently conservative number.

Fig. 31(a)–(d) compares PIKAN results (using triangular integration) with FEM solutions for single-via TGV-Cu structures with smooth and various roughness interfaces. Similarly, **Fig. 31(e), (f)** shows pointwise absolute error plots for the single-via structure to quantify prediction accuracy. The displacement field distribution patterns predicted by PIKAN show excellent agreement with FEM solutions across u_x and u_y . PIKAN successfully captures the deformation patterns in both glass and copper materials. **Fig. 32** presents the comparison between PIKAN results (using trapezoidal integration) and FEM solutions for the dual-via TGV-Cu structure. The pointwise absolute errors for the dual-via TGV-Cu structure are shown in **Fig. 32(c), (f)**. Despite the increased geometric complexity

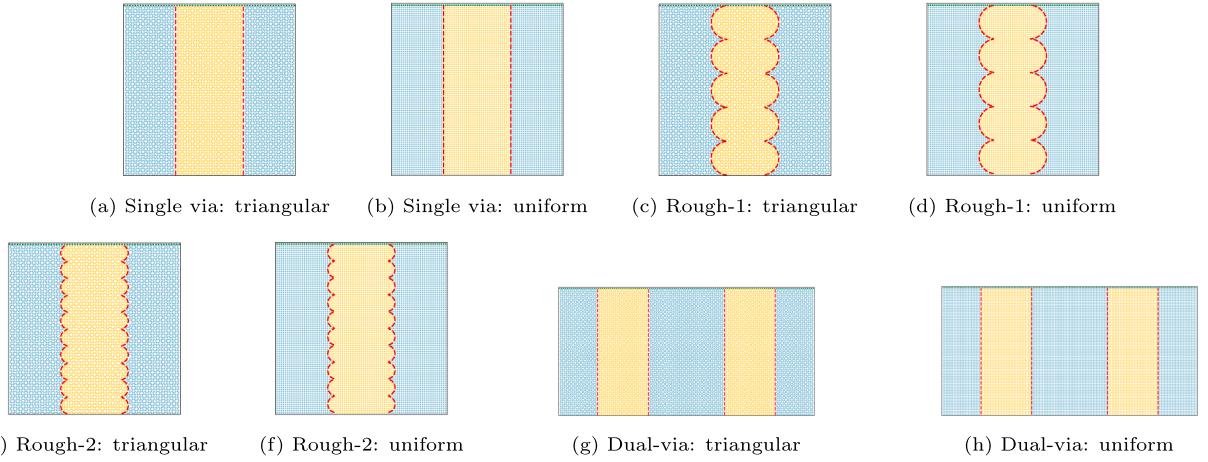


Fig. 29. Training point distribution schemes for TGV-Cu structure models: different sampling strategies applied to single and dual via configurations (Rough-1 and Rough-2 denote roughness 1 and roughness 2, respectively).

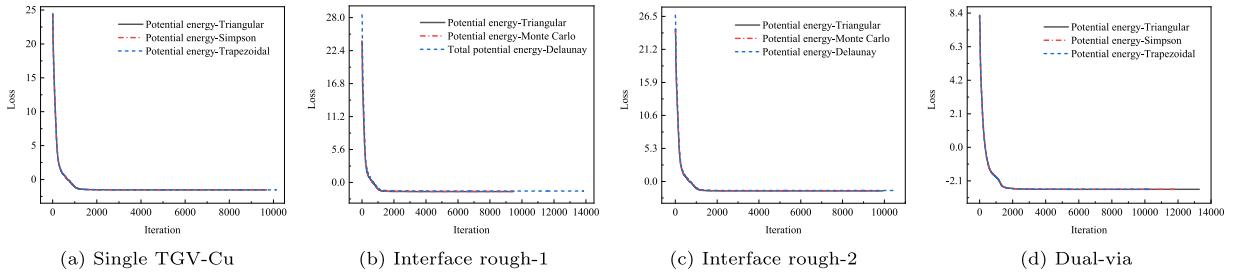


Fig. 30. Loss function evolution for different TGV-Cu structure configurations.

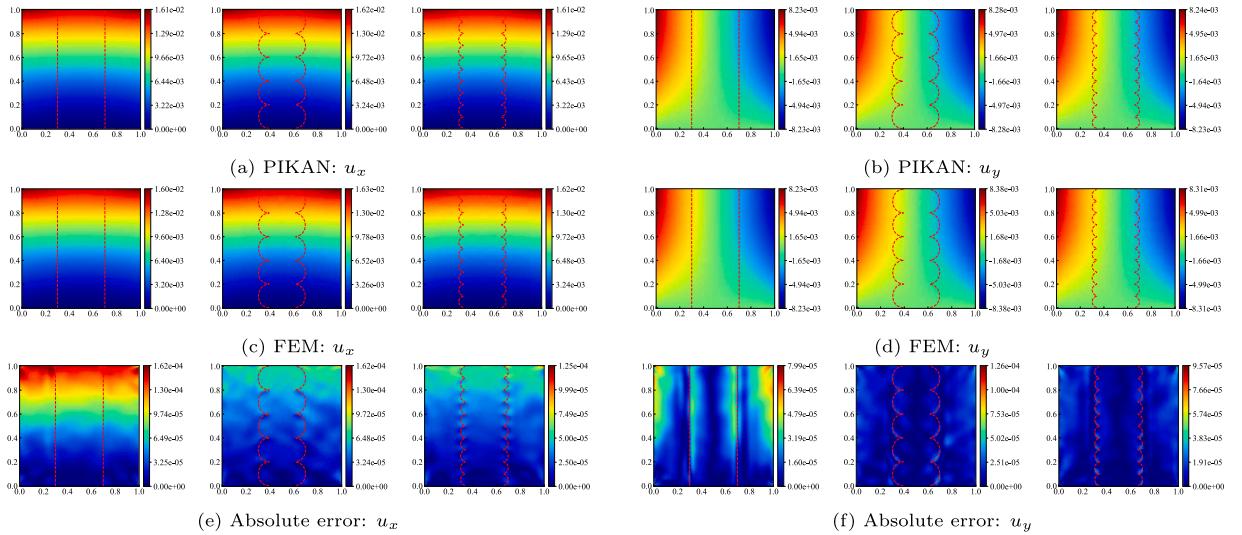


Fig. 31. Comparison of PIKAN and FEM displacement fields and the corresponding absolute errors in single-via TGV-Cu structures.

with multiple material interfaces and via interactions, PIKAN maintains high accuracy in predicting displacement fields. The method effectively captures the coupling effects between adjacent vias and the displacement distribution in the multi-via configuration.

To evaluate PIKAN's performance at critical material interfaces, displacement field u_x values are extracted along the glass-copper interfaces for all TGV-Cu configurations and compared with FEM reference solutions. To evaluate the method's robustness with respect to numerical integration schemes, we systematically compare results obtained using different integration approaches in Fig. 33. These interface analyses are particularly important for understanding stress transfer mechanisms

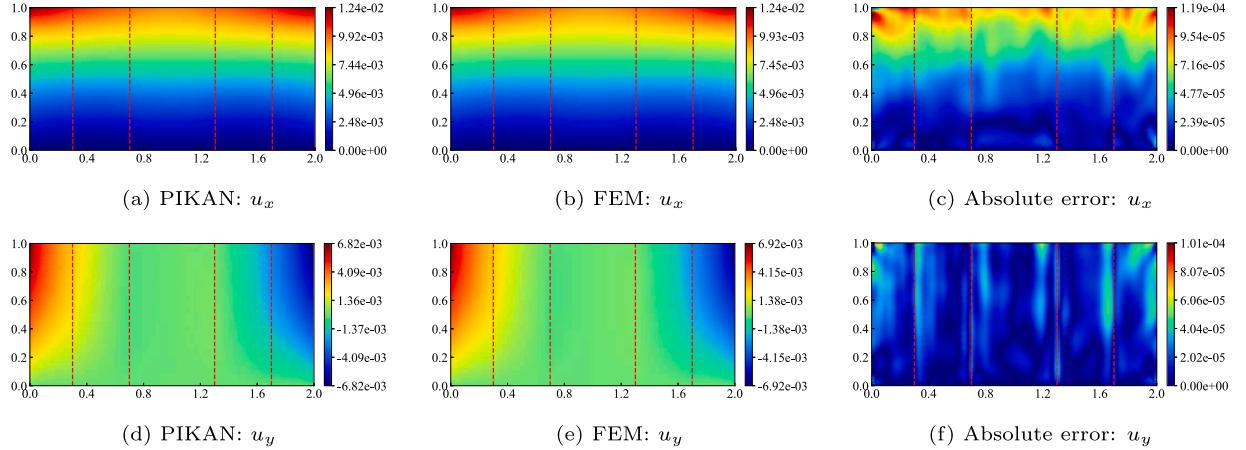


Fig. 32. Comparison of PIKAN and FEM displacement fields and the corresponding absolute errors in dual-via TGV-Cu structure.

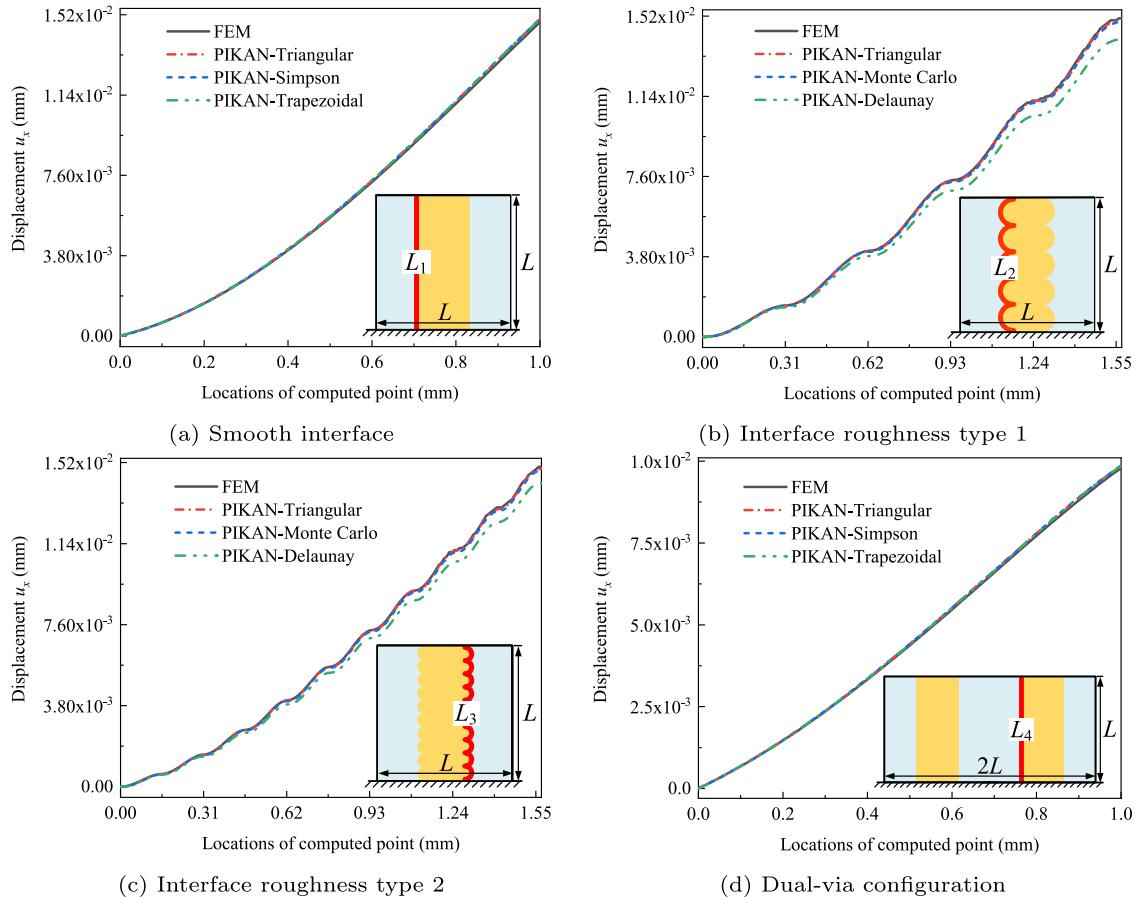


Fig. 33. Displacement u_x along interfaces for different TGV-Cu configurations (comparison of PIKAN predictions using different numerical integration schemes).

and potential failure locations in TGV structures. The comparison results demonstrate that PIKAN maintains excellent agreement with reference solutions across all geometric configurations, from simple smooth interfaces to complex roughness patterns and multi-via arrangements. Triangular, Monte Carlo, Simpson, and trapezoidal integration schemes all achieve high accuracy and converge well to reference solutions. However, Delaunay integration exhibits noticeable deviations from reference solutions across all curved interface geometries. This indicates that this integration scheme has inherent limitations when computing domain-wide strain energy integrals for complex geometries, with the introduced errors directly causing observable

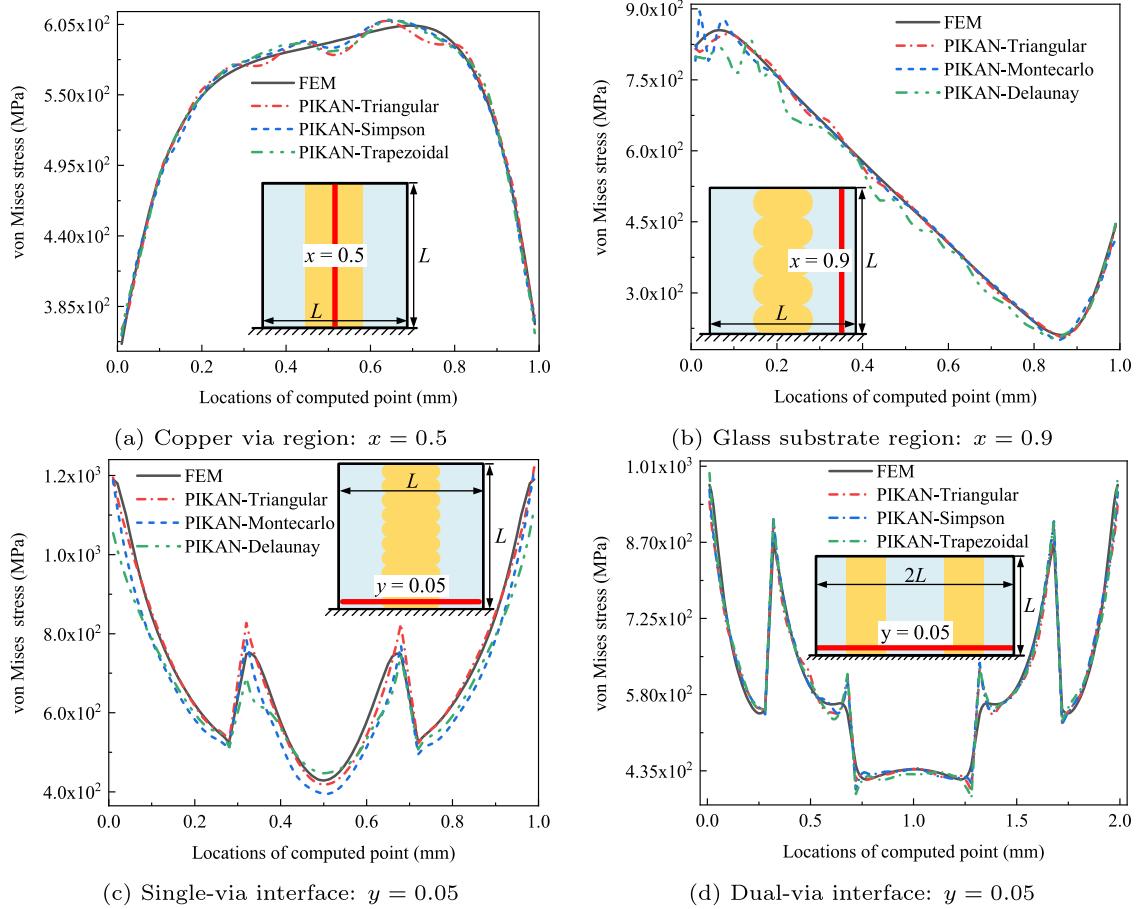


Fig. 34. Von mises stress distributions at critical locations in TGV-Cu structures.

deviations in the trained network solution. The robust performance across different interface geometries confirms PIKAN's effectiveness for complex multi-material structural analysis without requiring specialized interface treatments or domain decomposition approaches.

Fig. 34 presents stress distributions along strategically selected analysis lines to evaluate PIKAN's accuracy in capturing stress concentrations and material interface effects. Fig. 34(a) and (b) show stress distributions along vertical lines at $x = 0.5$ (within the copper via) and $x = 0.9$ (within the glass substrate). These locations capture stress variations in regions of high stress concentration (near the via) and far-field behavior (in the glass matrix). Fig. 34(c) and (d) present stress comparisons along the horizontal line $y = 0.05$, which crosses the glass-copper interface near the bottom boundary for both single-via and dual-via configurations. This analysis line is critical for understanding stress transfer across material interfaces and potential stress concentration effects. The results demonstrate clear performance differences among integration schemes. Triangular integration achieves the highest accuracy due to its adaptive mesh density that better resolves stress gradients near geometric discontinuities. Although Simpson's integration method has higher theoretical precision than the trapezoidal rule, PIKAN results using both methods achieve similar accuracy levels when the sample point density is sufficiently high. Monte Carlo integration performs moderately well but shows some scatter typical of stochastic methods. Delaunay integration exhibits the poorest numerical stability, especially at interface discontinuities. Its suboptimal control area allocation introduces non-negligible discretization errors into the loss function, adversely affecting training process robustness and convergence.

6. Conclusion

This study proposes PIKAN, a novel method combining KAN with the DEM for analyzing multi-material elasticity problems in electronic packaging structures. KAN networks inherently possess piecewise function characteristics that naturally accommodate material property discontinuities, while DEM provides superior computational efficiency and accuracy compared to strong-form PINNs. The PIKAN method effectively integrates these advantages to achieve accurate analysis of complex multi-material structures. Unlike traditional approaches that introduce penalty terms to enforce essential boundary conditions, PIKAN ensures solution validity by constructing admissible displacement fields that automatically satisfy essential boundary conditions. The method requires only a single KAN network to approximate displacement fields across the entire computational domain, eliminating the subdomain partitioning

required in conventional multi-material analysis. This approach avoids hyperparameter tuning for interface continuity conditions, reducing training complexity while achieving superior accuracy compared to domain decomposition methods.

Numerical validation demonstrates PIKAN's accuracy and robustness across diverse multi-material configurations. Comparative analysis with traditional CENN approaches confirms PIKAN's superior performance, particularly for complex interface geometries where penalty-based methods struggle with convergence. The study reveals that numerical integration strategy affects solution accuracy. Additionally, PIKAN requires fewer trainable parameters compared to MLP-based methods. Future work will systematically compare the performance of KANs with different basis functions (such as Chebyshev polynomials) for multi-material problems [33]. Additionally, exploring hybrid MLP-KAN architectures represents a promising direction [56]. For instance, employing MLPs for global approximation while integrating KANs in the output layer to capture local discontinuities with high precision could achieve an optimal balance between efficiency and accuracy. Such hybrid approaches may leverage the complementary strengths of both architectures: MLPs' efficiency in smooth regions and KANs' superior ability to handle sharp discontinuities. At the theoretical level, future research will focus on establishing rigorous mathematical foundations for the proposed PIKAN method, with emphasis on rigorous convergence and stability analysis of PIKAN for solving solid mechanics problems. Furthermore, the energy-based framework has been proven applicable to multiphysics and nonlinear problems [57,58], while the KAN architecture has demonstrated strong potential in solving complex differential equations [33,44,59]. Therefore, the proposed method establishes a solid foundation for addressing broader nonlinear and multiphysics problems, which represents an important future research direction. The combination of DEM's variational formulation with KAN's superior nonlinear representation capability positions PIKAN as a promising approach for tackling increasingly complex engineering problems.

The current study focuses on 2D validation examples. Extension to 3D electronic packaging structures presents significant challenges beyond simply increasing problem dimensionality. To begin with, for analyzing large-scale electronic packaging structures with multi-scale features, further enhancing computational efficiency and scalability will be critical. Future research needs to develop adaptive sampling and training strategies to efficiently capture physical field details in critical regions, and to design specialized optimization techniques for large-scale KAN networks to improve training stability and convergence speed. These advances will enable PIKAN to tackle increasingly complex, real-world electronic packaging problems while maintaining its core advantage of seamless multi-material interface handling. Furthermore, recent research has revealed that standard KANs face theoretical bottlenecks in effectively capturing E(3) transformations, which is essential for representing physical laws in 3D space [60], and their parameter efficiency may degrade poorly with increasing dimensionality [61]. These limitations are particularly concerning for 3D solid mechanics problems where computational efficiency and physical consistency are critical. Despite these challenges, recent studies have begun addressing high-dimensional and symmetry-related issues through improved KAN architectures [62] or by integrating KANs into 3D physical modeling frameworks [63], providing valuable references for future research. However, developing efficient, physically-meaningful PIKAN solvers suitable for 3D solid mechanics problems remains an open research question and represents a key focus for our future work.

In conclusion, this research successfully extends PINNs to heterogeneous material problems, demonstrating significant potential for electronic packaging structure reliability analysis and broader multi-material engineering applications. Source codes are available at <https://github.com/yanpeng-gong/PIKAN-MultiMateria>.

CRediT authorship contribution statement

Yanpeng Gong: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis; **Yida He:** Writing – review & editing, Writing – original draft, Visualization, Validation; **Yue Mei:** Supervision, Writing – review & editing; **Xiaoying Zhuang:** Supervision, Writing – review & editing; **Fei Qin:** Supervision; **Timon Rabczuk:** Writing – review & editing, Supervision.

Data availability

Data will be made available on request.

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.

Acknowledgment

This research was supported by the National Natural Science Foundation of China (Nos. 12002009, 12472198, W2431038), Beijing Natural Science Foundation (L233001) and the Alexander von Humboldt Foundation, Germany. The authors would like to thank Mohammad Sadegh Eshaghi for his valuable suggestions and guidance during the algorithm implementation process of this work.

Supplementary material

Supplementary material associated with this article can be found in the online version at [10.1016/j.apm.2026.116793](https://doi.org/10.1016/j.apm.2026.116793).

Appendix A. PIKAN algorithm for multi-material problems

Algorithm 1 PIKAN algorithm for multi-material problems.

```

1: Input:
2: Physical domains of different materials  $\Omega_1, \Omega_2, \dots, \Omega_n$ 
3: Boundary conditions: essential boundary  $\Gamma_u$  and natural boundary  $\Gamma_\sigma$ 
4: Material parameters:  $E_1, v_1, E_2, v_2, \dots, E_n, v_n$ 
5: Domain sample points  $x$  from all material domains
6: Natural boundary condition points  $x_\sigma$  from  $\Gamma_\sigma$ 
7: KAN architecture and hyperparameters
8: Neural network optimizer
9: Output: Optimized KAN parameters  $\theta^*$ 
10: Initialize: KAN parameters  $\theta = \{c_p^{(j,i)}, m_{j,i}, w_{j,i}\}$ 
11: Normalize all sample point coordinates
12: while Loss function not converged do
13: Obtain displacement predictions  $F(x)$  from single KAN for all sample points
14: Construct admissible displacement field  $u^{pred}(x)$ 
15: Compute displacement gradients  $\nabla u^{pred}$  using automatic differentiation
16: for each material domain  $i = 1, 2, \dots, n$  do
17: Compute strain energy  $\Psi_{in}^{m_i}$  using material parameters  $(E_i, v_i)$ 
18: end for
19: Compute external work potential energy:  $\Psi_{ex}$ 
20: Calculate loss function:  $\mathcal{L}_{PIKAN} = \sum_{i=1}^n \Psi_{in}^{m_i} - \Psi_{ex}$ 
21: Update KAN parameters  $\theta$  using optimizer and backpropagation
22: end while
23: Return: Optimized parameters  $\theta^*$ 

```

Appendix B. Hyperparameter sensitivity analysis of PIKAN

Hyperparameter selection critically affects PINN computational results. This section investigates how different hyperparameter combinations impact PIKAN prediction accuracy using the straight interface cantilever beam model from Fig. 6(a). We analyze the relative L_2 norm error variation for displacement component u_y under different KAN configurations, where each configuration is defined by network depth and width, grid size, and B-spline order. Fig. B.35 presents the hyperparameter sensitivity analysis results. Fig. B.35(a) shows the L_2 norm error of u_y with B-spline order fixed at 3 for three different KAN architectures: [2,5,5,2], [2,5,5,5,2], and [2,5,5,5,5,2]. The results indicate that increasing grid size does not produce consistent accuracy improvements across different network depths, suggesting that optimal grid size depends on the specific network architecture. Fig. B.35(b) demonstrates the effect of varying B-spline order on computational accuracy when grid size is fixed at 10. The results show that the optimal order varies with network architecture. These findings highlight the importance of joint optimization of KAN hyperparameters rather than individual parameter tuning.

The analysis reveals complex relationships between KAN hyperparameters and computational accuracy. While smaller grid sizes and B-spline orders often yield better performance in these specific examples, the optimal hyperparameter combination varies with network architecture and problem complexity. This sensitivity occurs because non-optimal hyperparameter selections can lead to either underfitting (insufficient model capacity to capture solution complexity) or overfitting (excessive model complexity leading to poor generalization), consistent with findings reported in machine learning literature [32]. The non-monotonic relationship between hyperparameters and accuracy highlights a fundamental challenge in KAN optimization: larger grid sizes or higher B-spline orders do not automatically improve performance. Instead, these parameters must be carefully balanced with network depth and width to achieve optimal results. Currently, effective methods for global optimization of KAN hyperparameter combinations are lacking due to the vast search space and extensive computational requirements for systematic exploration. Future work should focus on developing efficient hyperparameter optimization strategies specifically tailored for physics-informed KAN applications.

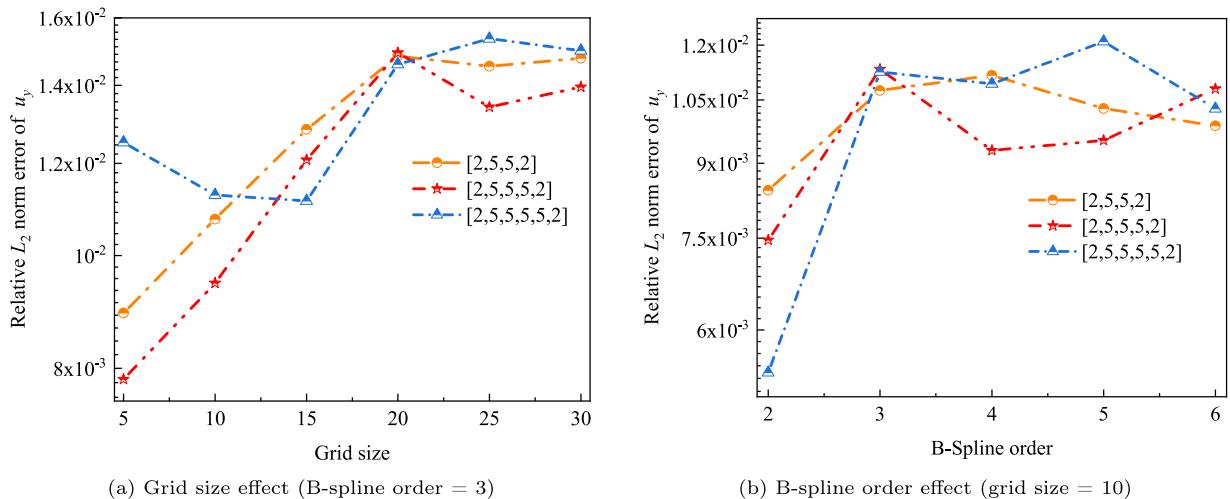


Fig. B.35. Relative L_2 norm error of u_y for different KAN hyperparameter combinations.

References

- [1] W.M. Moreau, Semiconductor lithography: principles, practices, and materials, Springer Science & Business Media, 2012.
- [2] Y.K. Kim, S.M. Lee, D.-S. Hwang, et al., Analyses on the large size PBGA packaging reliability under random vibrations for space applications, *Microelectron. Reliab.* 109 (2020) 113654.
- [3] S. Feng, Y. Guo, G. Królczyk, et al., An engineered solution to multi-physics of insulated gate bipolar transistor module considering electrical-thermal-mechanical coupling effect, *Adv. Eng. Softw.* 175 (2023) 103365.
- [4] Y. Gong, F. Qin, C. Dong, et al., An isogeometric boundary element method for heat transfer problems of multiscale structures in electronic packaging with arbitrary heat sources, *Appl. Math. Model.* 109 (2022) 161–185.
- [5] H. Yu, Y. Guo, Y. Gong, et al., Thermal analysis of electronic packaging structure using isogeometric boundary element method, *Eng. Anal. Bound. Elem.* 128 (2021) 195–202.
- [6] Y. Gong, Y. He, H. Hu, et al., A coupled finite element-boundary element method for transient elastic dynamic analysis of electronic packaging structures, *Eng. Struct.* 326 (2025) 119500.
- [7] F. Qin, Q. He, Y. Gong, et al., The application of FEM-BEM coupling method for steady 2D heat transfer problems with multi-scale structure, *Eng. Anal. Bound. Elem.* 137 (2022) 78–90.
- [8] Y.P. Zhang, Finite-difference time-domain analysis of integrated ceramic ball grid array package antenna for highly integrated wireless transceivers, *IEEE Transp. Antennas Propag.* 52 (2) (2004) 435–442.
- [9] Y. Gong, Y. Kou, Q. Yue, et al., The application of thermomechanically coupled phase-field models in electronic packaging interconnect structures, *Int. Commun. Heat Mass Transf.* 159 (2024) 108033.
- [10] Y. Gong, Y. Kou, Q. Yue, et al., A phase-field study on thermo-mechanical coupled damage evolution and failure mechanisms of sintered silver interconnections, *Eng. Fract. Mech.* 320 (2025) 111039.
- [11] S. Kang, J.H. Lee, S.M. Kim, et al., Machine learning-based solution for thermomechanical analysis of MMIC packaging, *Adv. Mater. Technol.* 8 (5) (2023) 2201479.
- [12] P. Zaghari, S.S. Sinha, D.C. Hopkins, et al., Co-design and ML-based optimization of through-via in silicon and glass interposers for electronic packaging applications, 2025.
- [13] C.-F. Yu, J.-W. Peng, C.-C. Hsiao, et al., Development of GUI-driven AI deep learning platform for predicting warpage behavior of fan-out wafer-level packaging, *Micromachines* 16 (3) (2025) 342.
- [14] L. Yuan, H.S. Park, E. Lejeune, Towards out of distribution generalization for problems in mechanics, *Comp. Methods Appl. Mech. Eng.* 400 (2022) 115569.
- [15] in: T. Rabczuk, K.-J. Bathe (Eds.), *Machine Learning in Modeling and Simulation*, 1st Edition, Cham, Springer, 2023.
- [16] J. Linghu, H. Dong, Y. Nie, et al., Higher-order multi-scale deep ritz method (HOMS-DRM) and its convergence analysis for solving thermal transfer problems of composite materials, *Comput. Mech.* 75 (1) (2025) 71–95.
- [17] J.N. Heidenreich, C. Bonatti, D. Mohr, Transfer learning of recurrent neural network-based plasticity models, *Int. J. Numer. Methods Eng.* 125 (1) (2024) 7357.
- [18] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comp. Phys.* 378 (2019) 686–707.
- [19] S. Cuomo, V.S.D. Cola, F. Giampaolo, et al., Scientific machine learning through physics-informed neural networks: where we are and what's next, *J. Sci. Comp.* 92 (3) (2022) 88.
- [20] E. Samaniego, C. Anitescu, S. Goswami, et al., An energy approach to the solution of partial differential equations in computational mechanics via machine learning: concepts, implementation and applications, *Comput. Method. Appl. M* 362 (2020) 112790.
- [21] E. W. B. Yu, The deep ritz method: a deep learning-based numerical algorithm for solving variational problems, *Commun. Math. Stat.* 6 (1) (2018) 1–12.
- [22] H. Ma, W. Qu, Y. Gu, et al., Spectral Integr. Neural Netw. Large Time Steps 2D 3D Transient Elastodynamic Analys. Neural Netw. 188 (2025) 107559.
- [23] G. Li, L. Qiu, F. Wang, et al., Adaptive multi-domain physics-informed neural networks for coupled thermo-mechanical analysis of thermal barrier coatings, *Eng. Comp.* 41 (2025) 3539–3562.
- [24] Q. Chen, G. Chatzigeorgiou, F. Meraghni, et al., Physics-informed deep homogenization approach for random nanoporous composites with energetic interfaces, *Eng. Appl. Artif. Intell.* 140 (2025) 109735.
- [25] S.A. Niaki, E. Haghigiat, T. Campbell, et al., Physics-informed neural network for modelling the thermochemical curing process of composite-tool systems during manufacture, *Comp. Methods Appl. Mech. Eng.* 384 (2021) 113959.
- [26] Z. Xiong, P. Zhao, A generalized physics-driven neural network for micromechanical and microstructural evolution of heterogeneous materials, *Eur. J. Mech. Solids* 111 (2025) 105551.
- [27] J. Linghu, W. Gao, H. Dong, et al., Higher-order multi-scale physics-informed randomized neural network method for efficient and high-accuracy simulation of dynamic thermo-mechanical coupling problems, *Math. Mech. Solids* 2025, 10812865251362165.

- [28] P. Yao, J. Yang, Y. Zhang, et al., Physics-based nested-ANN approach for fan-out wafer-level package reliability prediction, in: 2022 IEEE 72Nd Electronic Components And Technology Conference (ECTC), IEEE, 2022, pp. 1827–1833.
- [29] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Netw.* 2 (5) (1989) 359–366.
- [30] K. Gao, L. Qiu, G. Li, et al., A physics-informed framework based on radial basis functions for forward and inverse heat conduction problems in anisotropic materials, *Int. J. Heat Mass Transp.* 252 (2025) 127517.
- [31] H. Zhu, Z. Chen, X. Gao, et al., A physics-informed adaptive-wavelet neural network (PIAWNN) for reservoir simulation with two-phase flow, *Adv. Water Resour.* 207 (2026) 105174.
- [32] Z. Liu, Y. Wang, S. Vaidya, et al., Kan: kolmogorov-arnold networks, 2024, [arXiv:2404.19756](#)
- [33] K. Shukla, J.D. Toscano, Z. Wang, et al., A comprehensive and FAIR comparison between MLP and KAN representations for differential equations and operator networks, *Comp. Methods Appl. Mech. Eng.* 431 (2024) 117290.
- [34] Y. Wang, J. Sun, J. Bai, et al., Kolmogorov-Arnold-informed neural network: a physics-informed deep learning framework for solving forward and inverse problems based on Kolmogorov-Arnold networks, *Comp. Methods Appl. Mech. Eng.* 433 (2025) 117518.
- [35] B. Jacob, A.A. Howard, P. Stinis, SPIKANs: separable physics-informed Kolmogorov-Arnold networks, *Mach. Learn. Sci. Technol.* 6 (3) (2025) 035060.
- [36] J. Han, A. Jentzen, E. W. Solving high-dimensional partial differential equations using deep learning, *PNAS* 115 (2018) 8505–8510.
- [37] Y. Wang, J.W. Siegel, Z. Liu, et al., On the expressiveness and spectral bias of KANs, 2025, [arXiv:2410.01803](#)
- [38] J.D. Meshir, A. Palafox, E.A. Guerrero, On the study of frequency control and spectral bias in wavelet-based Kolmogorov Arnold networks, 2025, [arXiv:2502.00280](#)
- [39] J.D. Toscano, L.-L. Wang, G.E. Karniadakis, KKANs: kurková-kolmogorov-arnold networks and their learning dynamics, *Neural Netw.* 191 (2025) 107831.
- [40] Y. Diao, J. Yang, Y. Zhang, et al., Solving multi-material problems in solid mechanics using physics-informed neural networks based on domain decomposition technology, *Comp. Methods Appl. Mech. Eng.* 413 (2023) 116120.
- [41] Y. Wang, J. Sun, W. Li, et al., CENN: conservative energy method based on neural networks with subdomains for solving variational problems involving heterogeneous and complex geometries, *Comp. Methods Appl. Mech. Eng.* 400 (2022) 115491.
- [42] H. Sheng, C. Yang, PFNN: a penalty-free neural network method for solving a class of second-order boundary-value problems on complex geometries, *J. Comp. Phys.* 428 (2021) 110085.
- [43] S. Wang, X. Yu, P. Perdikaris, When and why PINNs fail to train: a neural tangent kernel perspective, *J. Comp. Phys.* 449 (2022) 110768.
- [44] J.D. Toscano, V. Oomen, A.J. Varghese, et al., From PINNs to PIKANs: recent advances in physics-informed machine learning, *Mach. Learn. Comput. Sci. Eng.* 1 (1) (2025) 1–43.
- [45] F. Mostajeran, S.A. Faroughi, Scaled-cPIKANs: domain scaling in Chebyshev-based physics-informed Kolmogorov-Arnold networks, 2025, [arXiv:2501.02762](#)
- [46] S.A. Faroughi, F. Mostajeran, Neural tangent kernel analysis to probe convergence in physics-informed neural solvers: PIKANs vs. PINNs, 2025, [arXiv:2506.07958](#)
- [47] W. Li, M.Z. Bazant, J. Zhu, A physics-guided neural network framework for elastic plates: comparison of governing equations-based and energy-based approaches, *Comp. Methods Appl. Mech. Eng.* 383 (2021) 113933.
- [48] X. Xu, T. Du, W. Kong, et al., Convergence analysis of natural gradient descent for over-parameterized physics-informed neural networks, 2024, [arXiv:2408.00573](#)
- [49] D. Gazoulis, I. Gkanis, C.G. Makridakis, On the stability and convergence of physics informed neural networks, *IMA J. Numer. Anal.* 2025, p. draft090.
- [50] J. Linghu, W. Gao, H. Dong, et al., Higher-order multi-scale physics-informed neural network (HOMS-PINN) method and its convergence analysis for solving elastic problems of authentic composite materials, *J. Comp. Appl. Math.* 456 (2025) 116223.
- [51] J. Bai, H. Jeong, C.P. Batuwatta-Gamage, et al., An introduction to programming physics-informed neural network-based computational solid mechanics, *Int. J. Comp. Methods* 20 (10) (2023) 2350013.
- [52] G. Ye, X. Fan, G. Zhang, Practical aspects of thermomechanical modeling in electronics packaging: a case study with a SiC power package, *Microelectron. Reliab.* 132 (2022) 114514.
- [53] N. Sukumar, A. Srivastava, Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks, *Comp. Methods Appl. Mech. Eng.* 389 (2022) 114333.
- [54] S. Wang, S. Sankaran, H. Wang, et al., An expert's guide to training physics-informed neural networks, 2023, [arXiv:2308.08468](#)
- [55] J.P. Gambino, S.A. Adderly, J.U. Knickerbocker, An overview of through-silicon-via technology and manufacturing challenges, *Microelectron. Eng.* 135 (2015) 73–106.
- [56] A. Kashefi, T. Mukerji, Physics-informed KAN PointNet: deep learning for simultaneous solutions to inverse problems in incompressible flow on numerous irregular geometries, 2025, [arXiv:2504.06327](#)
- [57] K.-C. Lin, C.-H. Hu, K.-C. Wang, Innovative deep energy method for piezoelectricity problems, *Appl. Math. Model.* 126 (2024) 405–419.
- [58] T. Le-Duc, T.N. Vo, H. Nguyen-Xuan, et al., On the Gauss-Legendre quadrature rule of deep energy method for one-dimensional problems in solid mechanics, *Finite Elem. Anal. Des* 242 (2024) 104248.
- [59] S. Patra, S. Panda, B.K. Parida, et al., Physics informed Kolmogorov-Arnold neural networks for dynamical analysis via Efficient-KAN and Wav-KAN, 2024, [arXiv:2407.18373](#)
- [60] F. Alesiani, T. Maruyama, H. Christiansen, et al., Geometric Kolmogorov-Arnold superposition theorem, 2025, [arXiv:2502.16664](#)
- [61] B.C. Koenig, S. Kim, S. Deng, Leankan: a parameter-lean Kolmogorov-Arnold network layer with improved memory efficiency and convergence behavior, *Neural Netw.* 192 (2025) 107883.
- [62] L. Hu, Y. Wang, Z. Lin, Incorporating arbitrary matrix group equivariance into KANs, 2025, [arXiv:2410.00435](#)
- [63] T. Huang, H. Zhang, Y. Zeng, et al., Dreamphysics: learning physics-based 3D dynamics with video diffusion priors, *Proc. AAAI Conf. Artif. Intell.* 39 (2025) 3733–3741.
- [64] F. Wang, X. Li, H. Liu, L. Qiu, X. Yue, An adaptive method of fundamental solutions using physics-informed neural networks, *Eng. Anal. Boundary Elem.* 178 (2025) 106295.