Article | Received 18 August 2025; Revised 25 October 2025; Accepted 5 December 2025; Published 22 December 2025
https://doi.org/10.55092/aimat20250016

A physics-constrained and data-driven approach for thermal field inversion in chiplet-based packaging

Yupeng Qi[1], Yue Wu[1,*], Jie Xiong[1,*], Shunbo Hu[1,2,*] and Deng Pan[1,3,*]

[1] Materials Genome Institute, Shanghai University, Shanghai 200444, China
[2] Management and Ministry of Education Key Laboratory of Silicate Cultural Relics Conservation, Institute for the Conservation of Cultural Heritage, School of Cultural Heritage and Information Shanghai University, Shanghai 200444, China
[3] Ministry of Education Key Laboratory of Silicate Cultural Relics Conservation, Shanghai University, Shanghai 200444, China

* Correspondence authors; E-mails: repeatsongyue@gmail.com (Y.W.); xiongjie@shu.edu.cn (J.X.); shunbohu@shu.edu.cn (S.H.); dpan_mgi@shu.edu.cn (D.P.).

Highlights:

- RC-PINN: physics-informed NN for chiplet thermal field inversion with sparse data.
- Material-boundary domain decomposition with temperature/flux continuity.
- Gaussian process uncertainty-guided sampling augments data; $R^2 > 0.99$ with 64 obs.

Abstract: The chiplet-based packaging integrates multiple heterogeneous chiplets with distinct functionalities and materials as a promising alternative to traditional SoC designs, reducing manufacturing costs and improving chip yield. However, increased power density and deteriorating heat dissipation conditions severely affect the functionality and stability of such systems, whereas existing thermal field inversion methods often overlook the correlation between data constraints and physical constraints, leading to insufficient prediction accuracy. This paper proposes a method based on Region-Coupled Physics-Informed Neural Networks (RC-PINNs) to address the thermal field inversion of chiplet packaging using a small dataset. RC-PINN divides the solution domain into multiple regions based on material properties and incorporates physical constraints into the loss functions of each region to ensure the solutions comply with physical laws. Coupling between different regions is achieved through temperature and heat flux continuity conditions to maintain physical consistency. Additionally, a Gaussian Process Uncertainty-Guided Sampling (GP-UGS) method is developed for data augmentation, further improving prediction accuracy in data-sparse scenarios. Experimental results demonstrate that RC-PINN effectively solves the thermal field inversion for chiplet packaging, achieving an average relative error of less than 0.4% under limited data conditions, with a 40% improvement in performance over traditional neural network methods. The GP-UGS method further improves prediction accuracy and generalization, raising the R^2 score to over 0.99 with only 64 observations.

Keywords: chiplet-based packaging; physics-informed neural network; thermal field inversion; small dataset; data augmentation

1. Introduction

The miniaturization of integrated circuits has led to increased production costs and design complexity for System-on-Chip (SoC) designs [1]. Chiplet-based packaging technology offers a promising solution to reduce manufacturing costs and improve chip yield [2]. Chiplet packaging integrates heterogeneous chips into a single system, achieving higher performance and flexibility [3,4]. However, the complex structure increases power density, posing significant thermal management challenges. However, the complex structure increases power density, posing significant thermal management challenges. Uneven heat distribution among chiplets can cause localized overheating, impacting system performance and stability [5–10]. Effective thermal management is crucial for ensuring system reliability [11–14]. Accurate thermal field inversion is vital for monitoring and controlling the thermal state of chips. By inferring the temperature distribution of the entire system from a limited number of temperature measurements, one can locate anomalies in the system, thereby identifying potential reliability issues and improving the system's reliability and lifespan [15–17]. However, accurate thermal field inversion typically requires numerous temperature sensors, and the layout of these sensors is constrained by the chip wiring design [18]. Integrating numerous temperature sensors into a chip not only increases design complexity and cost but also imposes considerable pressure on data processing and transmission, making temperature field reconstruction more challenging [19]. Therefore, achieving precise inversion of the chiplet packaging temperature field under limited observational data conditions is an urgent and complex issue.

The reconstruction of the temperature field of chiplet packaging is a typical inverse problem. In direct problems encountered in engineering and science, the governing equations and sufficient boundary or initial conditions are typically given, allowing for direct solution [20]. Traditional numerical methods, such as the finite difference method and the finite element method, have achieved significant success in solving direct problems [21,22]. Conversely, inverse problems involve inferring the physical quantities of interest in the absence of some governing equations or boundary conditions. In thermal applications in electronic and power equipment, obtaining accurate boundary conditions or heat source information is often extremely challenging. In the absence of complete information, inverse problems exhibit illposed characteristics, leading to solution instability or multiple solutions, which presents challenges for traditional numerical methods [23–25].

While modern processors incorporate on-die power monitoring capabilities (e.g., Intel RAPL, AMD APM) that provide aggregate or per-core power measurements, chiplet-based heterogeneous packages present unique thermal characterization challenges. First, power telemetry is often unavailable for passive package components such as through-silicon vias (TSVs), micro-bumps, and interposer interconnects, which collectively contribute significant heat generation. Second, the spatial distribution of power density within individual chiplets remains difficult to measure at sub-millimeter scales relevant to thermal hotspots, as existing power monitors only provide chip-level or coarse-grained regional data. Third, heterogeneous integration of legacy intellectual property (IP) blocks from different technology nodes often lacks consistent power monitoring infrastructure. Therefore, thermal field inversion with unknown or incompletely characterized heat sources remains a practically important problem for package-level thermal characterization and validation in advanced chiplet systems.

Interpolation algorithms can generate new data points based on the relationships between known data points [26]. However, when the distribution of data points is uneven or sparse, traditional interpolation methods often fail to provide accurate estimates. With the improvement of computational performance, data-driven machine learning methods have emerged as an important tool for inverse problem research [27].

However, these methods face challenges in solving inverse problems, such as handling high-dimensional data, noise, and poor interpretability, which affects accuracy and reliability [28,29].

Physics-Informed Neural Networks (PINNs), introduced by Raissi *et al.* [30], integrate physical laws into neural network training by transforming the solution of governing equations into an optimization task. This approach has demonstrated significant potential in scientific and engineering applications through its inherent enforcement of physically consistent predictions via the systematic incorporation of governing equations into the training process. For instance, Wei *et al.* used PINNs to reconstruct indoor flow fields from sensor data, achieving higher physical consistency compared to traditional neural networks, even with sparse measurement points [31]. Rui *et al.* successfully applied PINNs to three-dimensional wind field reconstruction around buildings, highlighting its capability in capturing complex flow characteristics in wind energy engineering [32]. Similarly, E. Fowler *et al.* demonstrated that PINNs can effectively simulate natural convection-driven cavities without relying on complete datasets [33]. In solid mechanics, PINNs have been used for material deformation and stress analysis, showing advantages in handling diverse boundary conditions and improving computational efficiency [34,35]. For instance, Khatamsaz *et al.* applied PINNs to predict corrosion fatigue crack growth under specific environmental conditions, offering a new tool for structural durability assessment [36]. PINNs have also proven effective in solving thermal conduction problems. Liu *et al.* developed a PINN framework for temperature field reconstruction with unknown heat sources, optimizing hyperparameter selection to enhance performance [37]. Mai *et al.* applied PINNs to turbine blade temperature fields with unknown boundary conditions, achieving higher accuracy and stability than traditional methods [25]. Furthermore, Cai demonstrated that PINNs can address ill-posed heat transfer problems that are challenging for conventional numerical methods, bridging computational and experimental data gaps [38]. In inverse problems, PINNs have been widely utilized. Rasht-Behesht *et al.* applied PINNs to full waveform inversion in acoustic media, improving both inversion accuracy and computational efficiency [39]. Li *et al.* introduced a probabilistic PINN framework for seismic petrophysics inversion, showcasing its flexibility in handling uncertainties [40]. Additionally, Raissi *et al.* demonstrated PINNs' ability to infer parameters of governing equations from sparse and noisy data [41], while Zhang *et al.* highlighted its application in material property estimation, facilitating advancements in material science [42].

Recent advances in domain decomposition for PINNs have shown promising results. Jagtap and Karniadakis [43] introduced the eXtended PINN (XPINN) framework, which decomposes the computational domain into subdomains with independent networks coupled through interface conditions. Similarly, Interface-PINN (I-PINN) methods have addressed multi-region problems. Our RC-PINN builds upon these foundations while specifically addressing the challenges of chiplet thermal field inversion with extreme ma-terial property contrasts and data-sparse conditions.

These studies demonstrate the significant potential of Physics-Informed Neural Networks (PINNs) in extracting physical information and resolving inverse problems, particularly within complex material systems. However, current research on PINNs predominantly focuses on optimization procedures,

including investigations of optimizers, adaptive weighting schemes, equation decoupling techniques, and neural architecture modifications, with the majority of efforts addressing accuracy and computational efficiency as PDE solvers. Recent works have begun exploring the criticality of sampling strategies in PINN methodologies [44–47], though these discussions remain confined to point selection protocols during forward problem solving. Aligned with [48], this work investigates the critical issue of ground truth sampling in PINN-based inverse problem solving. By strategically substituting portions of experimental measurements with Gaussian process regression outputs, we circumvent the computational efficiency degradation inherent in the iterative optimization framework proposed in [49], thereby aiming to achieve cost-effective thermal field monitoring for chiplet packaging applications. To realize this objective in the context of thermal field inversion for chiplet packaging systems, the following challenges must be systematically addressed:

- The irregular alternation of two materials in the same plane creates discontinuities due to differences in material properties, making accurate predictions difficult. Addressing these discontinuities is a critical challenge for achieving accurate predictions.
- Maintaining model accuracy under data scarcity is crucial. How to maximize the accuracy of temperature field reconstruction with limited data and governing equations remains an important topic for in-depth research.
- In practical engineering applications, the diversity of physical information significantly impacts model applicability. Further exploration is needed to assess model applicability in real-world scenarios to ensure stable and reliable predictions across various conditions.

To address these issues, this study proposes the Region-Coupled Physics-Informed Neural Network framework. By embedding physical information into partitioned neural networks, this method resolves material discontinuities in the solution domain. Additionally, a Gaussian process regression-based uncertainty sampling method (GP-UGS) is introduced to enhance model performance under data-scarce conditions. Our main contributions are the following:

- We identified two types of chip thermal inversion problems and designed three experimental cases to simulate various physical conditions.
- We proposed and validated the RC-PINN framework, which extends domain decomposition approaches (e.g., XPINN) to thermal inverse problems with extreme thermal conductivity contrasts (~500:1 between silicon and organic substrates). Unlike XPINN's computational efficiency focus, RC-PINN's decomposition is physically motivated by material boundaries, and our primary contribution lies in the novel integration with GP-UGS for data-sparse inverse problem solving.
- We introduced a Gaussian process regression (GPR)-based uncertainty sampling method for data augmentation, significantly boosting PINN performance in sparse data scenarios.

In this paper, we addressed two types of chip temperature field inversion problems through experiments under three conditions. The results demonstrate the RC-PINN method's superiority in thermal field inversion and verify the GP-UGS method's effectiveness in enhancing PINN performance under low-data conditions.

Following this organizational structure, the remainder of this paper is systematically presented. Section 2 delineates the problem definition and mathematical formulation governing thermal field inversion in chiplet-based packaging systems. Section 3 introduces the proposed Region-Coupled

Physics-Informed Neural Network framework and Gaussian Process Uncertainty-Guided Sampling methodology, including architectural components, coupling mechanisms, and implementation details. Section 4 validates the approach through comparative experiments across three representative thermal scenarios, analyzing prediction accuracy, robustness under sparse data conditions, and computational efficiency. Finally, Section 5 concludes with a synthesis of key findings, discusses practical implications for thermal management in heterogeneous integration, and outlines future research directions.

2. Problem definition

Chiplet-based packaging integrates multiple functional modules, achieving high performance, low cost, and flexibility. For example, the AMD EPYC™ server CPU uses chiplets for modular design, with CPU chiplets for computation and I/O Die for data control, as shown in Figure 1a. The substrate provides physical support and connects modules directly, simplifying manufacturing and reducing costs.

The 2D thermal model captures dominant in-plane heat conduction, reducing system complexity while efficiently representing core heat transfer processes.

In our 2D planar formulation, the term "boundaries" refers to the perimeter edges of the computational domain, which represent the lateral boundaries of the chiplet-substrate assembly. While vertical heat flow through top and bottom surfaces dominates overall heat dissipation in 3D chip stacks, the 2D model focuses on in-plane lateral heat spreading, which is critical for identifying localized hotspots and thermal coupling between adjacent chiplets. The in-plane temperature distributions are the primary measurable quantities through infrared thermography and embedded sensors.

To study the 2D thermal field inversion of chiplet packaging, we simplify the complex multi-chip packaging system into a 2D structural model, as depicted in Figure 1b,c. This model comprises chiplets and the substrate, where heat conduction occurs primarily in the planar direction. The substrate plays a crucial role in efficient heat dissipation, capturing key heat conduction paths and reflecting the system's thermal behavior [49].

In practical applications, chip environments are complex. While complete information may be unavailable, key data can describe the thermal field. Thermocouples or infrared thermometers can measure temperature and heat dissipation as boundary conditions. Under stable conditions, actual power can be assumed equal to rated power, dividing the heat conduction problem into known heat source and known boundary conditions.

For the chiplet packaging thermal field inversion problem, the observation data are given as $\{(x_i, y_i, T_i)\}_{i=1}^{N}$, where (x_i, y_i) are the observation point coordinates, and T_i is the observed temperature. The thermal field inversion can be formulated as the following optimization problem:

$$\min_{\theta} \sum_{i=1}^{N} | T_\theta(x_i, y_i | \phi) - T_i |, \tag{1}$$

where $T_\theta(\cdot)$ is the temperature field predicted by the model with parameters θ, and ϕ represents known physical conditions. To improve the efficiency and accuracy of Physics-Informed Neural Networks, we adopt dimensionless forms of the governing equations.

The dimensionless variables are defined as follows:

$$\xi = \frac{x}{L}, \quad \eta = \frac{y}{L}, \tag{2}$$

where L is the characteristic length. The dimensionless temperature and heat source are expressed as follows:

$$\theta(\xi,\eta) = \frac{T(x,y) - T_{ref}}{\Delta T}, \quad \varphi'(\xi,\eta) = \frac{L^2 Q(x,y)}{k\Delta T}, \tag{3}$$

where T_{ref} is the reference temperature, ΔT is the characteristic temperature, difference, and k is the thermal conductivity. The dimensionless heat conduction equation is the following:

$$\nabla^2\theta(\xi,\eta) = -\varphi'(\xi,\eta), \tag{4}$$

where ∇^2 denotes the Laplace operator, and $\varphi(x,y)$ represents the heat source strength.

Depending on the specific problem, the following types of boundary conditions can be applied:

- Dirichlet Boundary Condition: the boundary temperature is known, and its dimensionless form is the following:

$$\theta(\xi,\eta) = \theta_{bdy}, \tag{5}$$

where $\theta_{bdy} = \frac{T_{bdy} - T_{ref}}{\Delta T}$.

- Neumann Boundary Condition: the heat flux at the boundary is known, and its dimensionless form is the following:

$$-\frac{\partial\theta}{\partial n} = q'_{bdy}, \tag{6}$$

where $q'_{bdy} = \frac{q_{bdy}L}{k\Delta T}$, and $\frac{\partial\theta}{\partial n}$ represents the temperature gradient along the normal direction.

- Robin Boundary Condition: this condition combines temperature and heat flux, often used for convective boundaries, and is written as follows:

$$-\frac{\partial\theta}{\partial n} = h'(\theta - \theta_\infty), \tag{7}$$

where $h' = \frac{hL}{k}$ is the dimensionless convective coefficient, and $\theta_\infty = \frac{T_\infty - T_{ref}}{\Delta T}$.

To maintain continuity of temperature and heat flux at interfaces between adjacent regions, the following conditions must be satisfied:

- Temperature Continuity: the temperatures at the interface must be equal:

$$\theta_i(\xi,\eta) = \theta_j(\xi,\eta) \quad \text{for} \quad (\xi,\eta) \in \Gamma_{ij}, \tag{8}$$

- Heat Flux Continuity: the heat flux must be consistent across the interface:

$$\kappa_i \frac{\partial\theta_i}{\partial n} = -\kappa_j \frac{\partial\theta_j}{\partial n} \quad \text{for} \quad (\xi,\eta) \in \Gamma_{ij}, \tag{9}$$

where κ_i and κ_j are the thermal conductivities of regions i and j, respectively.

In practice, physical information is often incomplete, making it difficult for traditional methods to achieve accurate solutions. By embedding known physical constraints into the neural network's loss function, PINNs offer a promising approach to addressing inverse heat transfer problems, especially when boundary conditions are partially unknown.

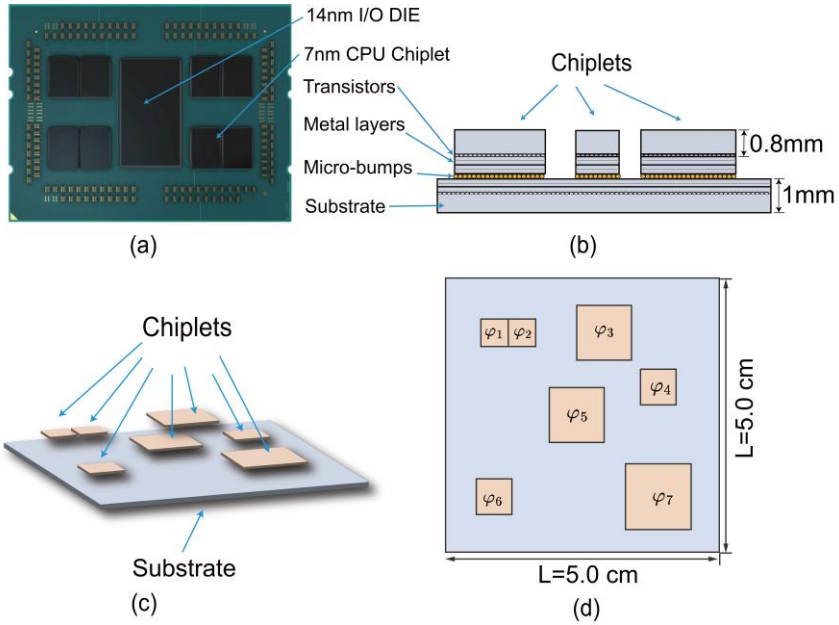


Figure 1. Chiplet examples and simplified Models. **(a)** AMD EPYC chiplet packaging; **(b)** Typical chiplet packaging diagram; **(c, d)** Simplified model for this problem.

3. Methodology

For chip thermal field inversion, where the heat source or boundary conditions are unknown, reconstructing the temperature field becomes a challenging inverse problem. Numerical methods such as FEM and FDM rely on complete physical information and are therefore inapplicable in such cases. The challenge is further exacerbated by data scarcity. In real-world scenarios, it is impractical to deploy extensive sensors to collect temperature data, as this increases costs and negatively impacts chip performance. Existing machine learning methods, which require large datasets, also fail to provide reliable solutions under these constraints.

To address this issue, we propose an RC-PINN method specifically designed for chip thermal field reconstruction with small datasets. To handle the discontinuity in material properties across different regions, RC-PINN divides the chip area into multiple sub-regions, modeling each separately while using physical coupling conditions to reconstruct the global thermal field. Although RC-PINN demonstrates good accuracy under small dataset conditions, to further enhance the model's accuracy and robustness in extremely data-scarce situations, we introduce a Gaussian Process Uncertainty-Guided Sampling method. This method generates new sampling points via Gaussian process regression and selects them based on uncertainty evaluation. The overall architecture of the method is shown in Figure 2.

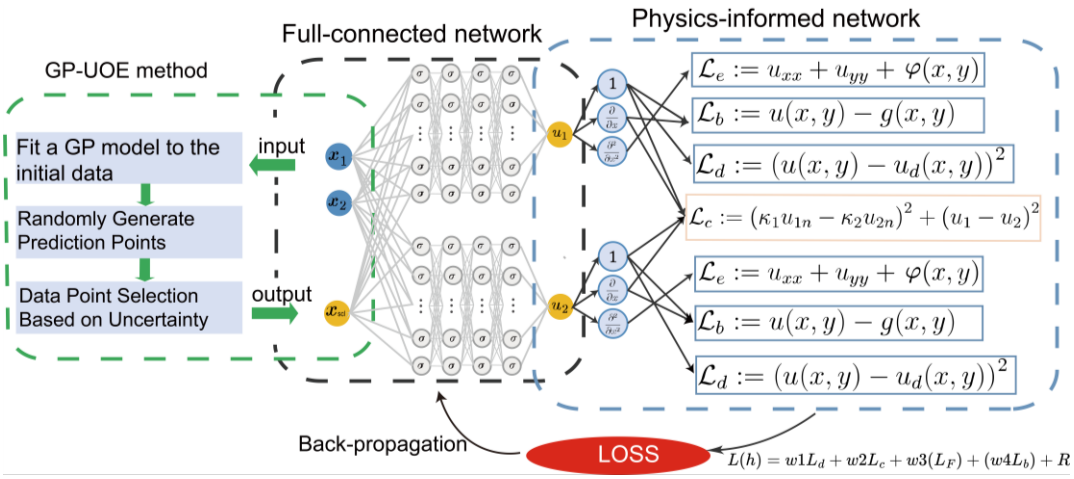


Figure 2. PINN network structure combining partitioned framework with Gaussian process data augmentation.

3.1. Region-Coupled Physics-Informed Neural Networks

In the thermal field inversion task for chiplet architectures, the physical properties of different materials vary, leading to discontinuities in the derivatives (*i.e.*, temperature gradients) at material interfaces. The standard Physics-Informed Neural Network (PINN) framework [29] typically generates models with continuous derivatives, which is unsuitable for such problems.

Our preliminary experiments with standard single-network PINN showed unsatisfactory results for multi-material configurations with extreme thermal conductivity contrasts. The single network struggled to capture sharp temperature gradients at material interfaces, exhibiting significant errors near chip-substrate boundaries. This observation is consistent with comparative studies in XPINN literature, which demonstrated that domain decomposition approaches outperform single-network PINNs for problems with heterogeneous properties. To address this issue, some studies have employed domain decomposition methods. These methods divide the domain into multiple sub-regions using the weak form of partial differential equations (PDEs), solving each sub-region with independent basis functions. While this approach effectively handles discontinuities, it significantly increases model complexity and training difficulty [45]. Other research has attempted to explicitly model material discontinuities within deep neural networks (DNNs) by imposing known boundary conditions in the PINN loss function [31]. Although these studies have advanced the application of PINNs in handling material interfaces, they pay little attention to the challenges of thermal field inversion, particularly the balancing of multiple loss functions in complex multi-interface scenarios.

To address the challenges in thermal field inversion and improve model accuracy and convergence, this paper proposes a Region-Coupled Physics-Informed Neural Network (RC-PINN) framework. The RC-PINN framework integrates domain decomposition and loss function balancing to address the unique demands of chiplet package structures. Its key components include the following:

- Domain Partitioning: the solution domain is divided into multiple sub-regions based on material properties, and each region is solved independently.
- Physical Loss and Boundary Coupling: appropriate physical loss functions are embedded in different regions, and boundary continuity conditions are applied to achieve coupling between regions.

- Uncertainty-Based Weighting: a dynamic weighting mechanism adjusts the contributions of loss functions, improving training stability and convergence.

The structure of the framework is shown in Figure 2. The RC-PINN divides the solution domain into chip and substrate regions, with each region modeled using a separate multilayer perceptron (MLP) to learn the temperature field characteristics. This partitioning strategy enables each MLP model to focus on learning the temperature field distribution within its corresponding sub-region.

For the i-th sub-region, the temperature field $T_i(x)$ is predicted by the corresponding MLP model:

$$T_i(x) = \mathrm{MLP}_i(x; \theta_i), \tag{10}$$

where MLP_i represents the neural network model for the i-th sub-region, and θ_i is the set of parameters for this model. An MLP is a feedforward neural network composed of multiple fully connected layers. The input feature x (e.g., planar coordinates (x, y)) is processed through successive linear transformations and nonlinear activation functions to predict the output $T_i(x)$, expressed as follows:

$$T_i(x) = f^{(L)}\left(W^{(L)} \ldots f^{(1)}\left(W^{(1)}x + b^{(1)}\right) + b^{(L)}\right), \tag{11}$$

where $W^{(l)}$ and $b^{(l)}$ denote the weight matrix and bias vector of the l-th layer, respectively, and $f^{(l)}$ is the activation function of the l-th layer. In this study, the Swish activation function [32] is employed, which is defined as follows:

$$f(x) = x \cdot \frac{1}{1 + e^{-x}} \tag{12}$$

To fully utilize each data point and improve the model's solution accuracy, two independent neural networks are trained separately on the entire training dataset. Since the network models for each sub-region are trained independently, to ensure that the temperature fields in different sub-regions connect seamlessly, the RC-PINN framework introduces continuity conditions for temperature and heat flux as a coupling mechanism. Specifically, these coupling conditions include both temperature continuity and heat flux continuity, ensuring the physical consistency of the global temperature field.

In traditional neural networks, the loss function typically includes only data-driven terms. However, in a Physics-Informed Neural Network, the loss function comprises not only the data-driven loss L_d but also the governing equation loss L_F, boundary condition loss L_b, and our proposed continuity coupling loss L_c. These loss terms are weighted based on their associated physical constraints. By embedding the partial differential equation (PDE) into the loss function, the PINN framework transforms the problem of solving the governing equation into an optimization problem. The specific forms of each loss term are as follows:

Interface collocation points are sampled uniformly along material boundaries, with sufficient density (typically 1000 points per interface) to ensure accurate continuity enforcement. This sampling density has been verified to provide stable and accurate solutions across our test cases. For complex geometries with curved interfaces, the framework can accommodate adaptive sampling strategies such as curvature-based refinement.

- Data-Driven Loss $L_{d,i}(h)$: this term measures the discrepancy between the predicted temperature field and the observed data in sub-region i, defined as follows:

$$L_{d,i}(h) = \frac{1}{|\, P_{d,i} \,|} \sum_{(\xi,\eta)\in P_{d,i}} |\, \theta_i(\xi,\eta) - \theta_{\mathrm{data},i}(\xi,\eta) \,|^2, \tag{13}$$

where $P_{d,i}$ represents the set of known data points in sub-region i.

- Governing Equation Loss $L_F(h)$: this term describes the constraints imposed by the PDE within the solution domain, defined as follows:

$$L_F(h) = \frac{1}{|P_F|} \sum_{(\xi,\eta)\in P_F} |\nabla^2\theta(\xi,\eta) + \varphi(\xi,\eta)|^2, \tag{14}$$

where P_F is the set of sampled points in the domain, and $Q(\xi,\eta)$ represents the heat source term.

- Boundary Condition Loss $L_b(h)$: this term addresses different types of boundary conditions, with specific formulations as follows:

(a) Dirichlet Boundary Condition (known boundary temperature):

$$L_b(h) = \frac{1}{|P_b|} \sum_{(\xi,\eta)\in P_b} |\theta(\xi,\eta) - \theta_{\text{bdy}}|^2, \tag{15}$$

(b) Neumann Boundary Condition (known boundary heat flux density):

$$L_b(h) = \frac{1}{|P_b|} \sum_{(\xi,\eta)\in P_b} |-\frac{\partial\theta}{\partial n} - q_{\text{bdy}}|^2, \tag{16}$$

(c) Robin Boundary Condition (convective condition):

$$L_b(h) = \frac{1}{|P_b|} \sum_{(\xi,\eta)\in P_b} |-\frac{\partial\theta}{\partial n} - h_{\text{bdy}}(\theta - \theta_\infty)|^2, \tag{17}$$

- Continuity Coupling Loss $L_c(h)$: this term ensures the continuity of temperature and heat flux across material interfaces:

$$L_c(h) = \sum_{\Gamma_{ij}} \left(|\theta_i(\xi,\eta) - \theta_j(\xi,\eta)|^2 + |-\kappa_i \frac{\partial\theta_i}{\partial n} - \left(-\kappa_j \frac{\partial\theta_j}{\partial n}\right)|^2 \right), \tag{18}$$

where Γ_{ij} represents the boundaries between all sub-regions.

In these expressions, P_F, P_b, and P_d denote the sets of sampled points within the domain, boundary sampling points, and known data points, respectively. By incorporating these physical constraints as terms in the loss function, the PINN framework can convert the problem of solving governing equations into an optimization problem.

However, as the model complexity grows, balancing fixed weights for loss terms becomes challenging and may hinder convergence. To address this issue, we adopt an uncertainty-based adaptive weighting method, used in computer vision tasks for multi-task learning [50]. By modeling task-specific losses with a Gaussian likelihood function, this method dynamically adjusts the weights of loss terms based on their uncertainty, ensuring balanced contributions to the training process.

In multi-task learning, the loss function for each task is assumed to follow a Gaussian distribution, enabling automatic weight adjustment. For regression tasks, the Gaussian likelihood loss function is as follows:

$$p(y \mid f(x),\sigma) = \mathcal{N}(f(x),\sigma^2), \tag{19}$$

where *f(x)* is the model's predicted output, and σ represents the noise or uncertainty associated with the task. Maximizing the log-likelihood yields the following:

$$L(y,f(x),\sigma) = \frac{1}{2\sigma^2} \| y - f(x) \|^2 + \log\sigma, \tag{20}$$

Tasks with higher uncertainty σ, contribute less to the total loss, while tasks with lower uncertainty are weighted more heavily. This approach eliminates the need for manual weight adjustment, dynamically balancing loss terms based on task difficulty.

Building on this principle, we incorporate the adaptive weighted loss function into the RC-PINN framework to manage physical constraints and temperature field inversion across sub-regions. The total loss function is formulated as follows:

$$L_{\text{total}} = \sum_i \left(\frac{1}{2\sigma_i^2} L_i + \log \sigma_i \right), \tag{21}$$

where L_i denotes the i-th loss term, and σ_i is the uncertainty parameter associated with that loss term. This loss function has the following advantages: the uncertainty parameter σ_i can automatically adjust the weight of each loss term, thus avoiding the cumbersome process of manual parameter tuning. By introducing the uncertainty-based weighting mechanism, RC-PINN can effectively balance the influence of different loss terms when dealing with multi-task, multi-region temperature field inversion problems, significantly improving the convergence and accuracy of the model.

Based on this approach, we design the loss functions for two different types of inverse problems in the RC-PINN framework:

- First Inverse Problem: Known Boundary Conditions, Unknown Heat Source Intensity.

In this scenario, the loss function includes the data-driven losses L_{d_1} and L_{d_2}, boundary condition loss L_b, and continuity coupling loss L_c, while the governing equation loss L_F is not included.

In the unknown heat source inverse problem, the governing equation residual L_F cannot be included in the loss function since the source term Q(x,y) is the inference target. However, the solution remains physics-informed through enforcement of boundary conditions and interface continuity laws, which are fundamental physical constraints derived from energy conservation and Fourier's law. This formulation is consistent with inverse problem methodologies in the PINN literature [25,30,37], where available physical knowledge constrains the solution space while unknown quantities are inferred from data.

The expression for the loss function is as follows:

$$L(h) = \frac{1}{2\sigma_{d_1}^2} L_{d_1}(h) + \frac{1}{2\sigma_{d_2}^2} L_{d_2}(h) + \frac{1}{2\sigma_b^2} L_b(h) + \frac{1}{2\sigma_c^2} L_c(h) + \log\left(\sigma_{d_1} \cdot \sigma_{d_2} \cdot \sigma_b \cdot \sigma_c\right), \tag{22}$$

where σ_d, σ_b, and σ_c are the uncertainty parameters associated with the data-driven, boundary condition, and continuity loss terms, respectively.

- Second Inverse Problem: Known Heat Source Intensity, Unknown Boundary Conditions.

In this scenario, the loss function comprises data-driven losses L_{d_1} and L_{d_2}, the governing equation loss L_F, and the continuity coupling loss L_c, while the boundary condition loss L_b is excluded. The final loss function is expressed as follows:

$$L(h) = \frac{1}{2\sigma_{d_1}^2} L_{d_1}(h) + \frac{1}{2\sigma_{d_2}^2} L_{d_2}(h) + \frac{1}{2\sigma_F^2} L_F(h) + \frac{1}{2\sigma_c^2} L_c(h) + \log\left(\sigma_{d_1} \cdot \sigma_{d_2} \cdot \sigma_F \cdot \sigma_c\right), \tag{23}$$

In this formulation, σ_d, σ_F, and σ_c are the dynamic weighting parameters for the data-driven, governing equation, and continuity loss terms. By automatically adjusting the values of these uncertainty parameters, the weights of the different loss terms can be dynamically balanced, improving the model's adaptability in handling complex tasks.

Implementation Details: Following Kendall *et al.* [49], the uncertainty parameters σ_i are treated as learnable variables initialized to unity ($\sigma_i = 1.0$ for all i) and optimized jointly with the neural network

parameters θ through automatic differentiation during training. This initialization ensures balanced contributions from all loss terms at the beginning of training. To maintain numerical stability, we employ gradient clipping with maximum norm 1.0 and use he Adam optimizer with exponential learning rate decay (initial rate 10^{-3}, decay factor 0.95 every 1000 iterations). These standard techniques prevent potential instabilities during training and ensure robust convergence across all experimental cases presented in Section 4.

Through this approach, we successfully incorporate the dimensionless steady-state 2D heat conduction equation and boundary conditions into the PINN framework's loss function, addressing two different types of inverse problems. This method effectively solves the problem of balancing different loss terms in multi-task learning, significantly enhancing the model's convergence and accuracy.

3.2. Gaussian Process Uncertainty-Guided Sampling

Although PINN models perform well with small datasets, the high cost of labeled data in real-world applications requires achieving higher accuracy with fewer data points. Data augmentation offers an effective solution. We use Gaussian process to fit the original dataset, quantifying prediction uncertainty. Based on this uncertainty, we select optimal data points for training, expanding the dataset and improving model accuracy and robustness.

We employ the standard Radial Basis Function (RBF) kernel for Gaussian process regression, which assumes smooth spatial correlations appropriate for temperature fields governed by the heat conduction equation. While alternative kernels such as the Matern family can accommodate less smooth correlations near material interfaces, our experiments show that RBF performs well for the rectangular chiplet geometries studied. For applications with particularly complex interface geometries, practitioners may explore Matern kernels following standard GP literature.

The data points generated by the Gaussian process carry both potential information and associated predictive uncertainty. This dual nature requires a careful balance between information gain and uncertainty during data point selection. To quantify the value of each candidate data point, we introduce a utility function $I(x)$:

$$I(x) = 1 - \alpha \cdot \sigma(x), \tag{24}$$

where $\sigma(x)$ represents the predictive uncertainty of the Gaussian process at point x, and α is a weighting factor that controls the influence of uncertainty on the utility score. When $I(x) > 0$, the data point is considered beneficial to the model's training. Conversely, if $I(x) \leq 0$, the point may introduce noise or error, potentially harming model performance.

Based on the utility function, we classify data points into three categories and adopt corresponding selection strategies:

(1) Low-uncertainty points ($\sigma(x) < \sigma_{\text{low}}$): These points typically lie in well-understood regions of the model, offering high confidence and minimal error. Selecting these points reinforces the model's precision in already reliable areas.

(2) Moderate-uncertainty points ($\sigma_{\text{low}} \leq \sigma(x) \leq \sigma_{\text{high}}$): These points are found in less-explored regions of the model and, while carrying higher predictive error, provide significant information gain. Adding these points helps the model expand its learning scope, enabling it to discover new patterns and adapt to unexplored areas.

(3) High-uncertainty points ($\sigma(x) > \sigma_{\text{high}}$): Points in this category are typically dominated by noise and exhibit excessive predictive error. Their inclusion is likely to degrade model performance by introducing unnecessary noise. Hence, they are excluded from the training set.

In the GP-UGS framework, the selection of low- and moderate-uncertainty points serves distinct but complementary purposes. Adding low-uncertainty points is akin to in-troducing Gaussian noise, reinforcing the model's understanding of already well-covered regions while improving robustness to data noise. Conversely, moderate-uncertainty points are selected to mimic an active learning strategy, prioritizing exploration of less-studied regions where the model can gain new insights. This dual strategy balances the need for local accuracy enhancement with domain exploration, thereby improving both precision and generalization. By integrating these two types of points, the GP-UGS method ensures that new data points contribute effectively to model improvement while avoiding the drawbacks of excessive noise or insufficient coverage.

3.3. Implementation steps for GP-UGS

In the GP-UGS method, we first fit the existing small dataset with a Gaussian process to obtain the predictive mean $\mu(x)$ and uncertainty $\sigma(x)$. Figure 3 shows the flow of the method. The detailed implementation of the method follows the steps outlined below:

(1) Random Data Generation: a set of random data points $X_{\text{random}} = \{(x_i, y_i)\}_{i=1}^{M}$ is generated within the domain using low-discrepancy sampling to ensure uniform coverage, where M denotes the number of generated points; these points are subsequently evaluated to quantify uncertainty.

(2) Uncertainty Sorting: based on the variance $\sigma(x_i, y_i)$ predicted by the GP model, the points in X_{random} are sorted, resulting in an ordered set X_{sorted}:

$$X_{\text{sorted}} = \{(x_{(1)}, y_{(1)}), (x_{(2)}, y_{(2)}), \dots, (x_{(M)}, y_{(M)})\}, \tag{25}$$

where $\sigma(x_{(1)}, y_{(1)}) \le \sigma(x_{(2)}, y_{(2)}) \le \cdots \le \sigma(x_{(M)}, y_{(M)})$.

(3) Selection of Low-Uncertainty Data Points: the top k_1 points with the lowest uncertainty are selected, forming a low-uncertainty data point set:

$$X_{\text{low}} = \{(x_{(1)}, y_{(1)}), (x_{(2)}, y_{(2)}), \dots, (x_{(k_1)}, y_{(k_1)})\}, \tag{26}$$

These low-uncertainty points, located in high-confidence regions, enhance data coverage density, strengthen the model's learning capacity in core regions, and mitigate the risk of overfitting.

(4) Selection of Moderate-Uncertainty Data Points: data points with moderate uncertainty (e.g., between the p-th and q-th percentiles) are selected from the sorted set, forming a moderate-uncertainty data point set:

$$X_{\text{medium}} = \{(x_{\lceil pM \rceil}, y_{\lceil pM \rceil}), \dots, (x_{\lfloor qM \rfloor}, y_{\lfloor qM \rfloor})\}, \tag{27}$$

These points are located in less-explored regions of the model, possess exploratory potential, and help reduce errors in unknown areas.

(5) Comprehensive Data Point Selection: the final selected data points are as follows:

$$X_{\text{selected}} = X_{\text{low}} \cup X_{\text{m}}, \tag{28}$$

These combined high-confidence and moderate-uncertainty data points constitute the new training dataset and are incorporated into the PINN training process.

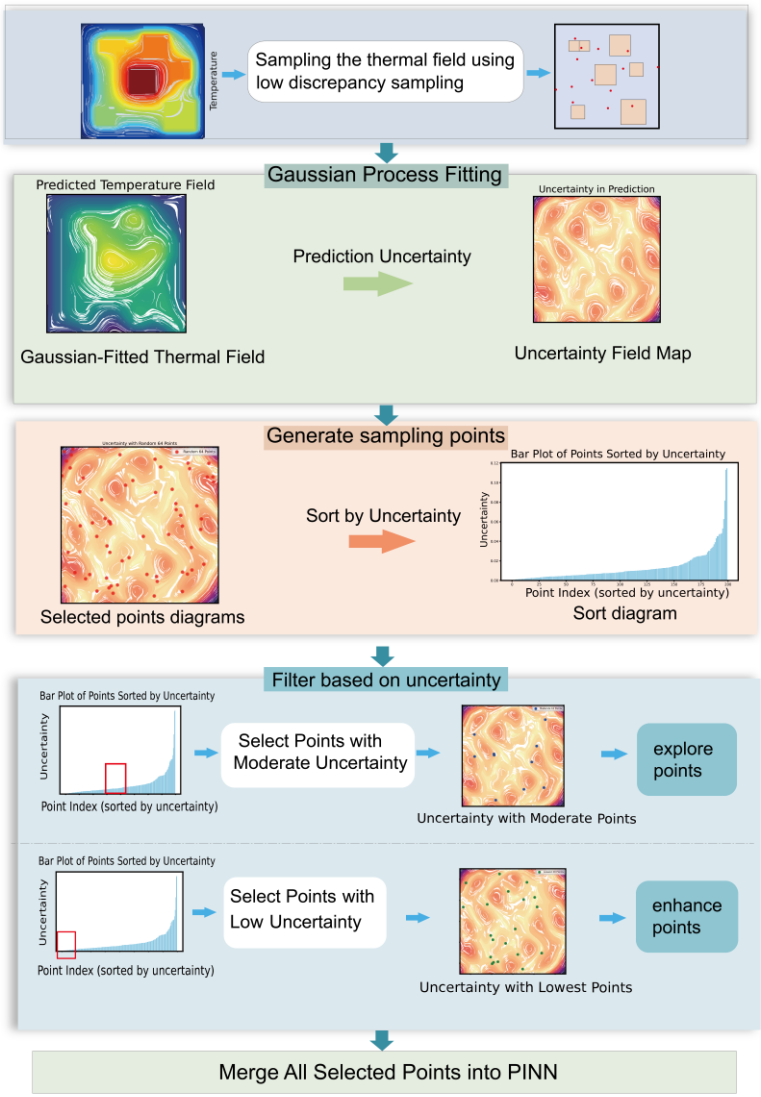


Figure 3. Process of a small data augmentation method based on Gaussian process regression un-certainty.

4. Experiment

Our study investigates two key scenarios in thermal field inversion for chiplet packaging: one where the boundary conditions are known but the heat source intensity is unknown, and another where the heat source intensity is known but the boundary conditions are unknown. Traditional numerical methods struggle in these scenarios, as they rely solely on governing equations and sparse observational data, often leading to inaccurate predictions in the temperature field.

Deep learning techniques, while effective in capturing complex, nonlinear relationships, require large datasets to achieve high predictive accuracy. However, acquiring extensive observational data in thermal field inversion tasks poses significant challenges. Deploying numerous temperature sensors increases costs and risks negatively impacting chip performance. Thus, collecting a large-scale temperature dataset is impractical in real-world applications.

PINN address this limitation by embedding physical laws into the loss function, significantly enhancing the model's predictive accuracy in data-scarce environments. In this study, we investigate the minimum amount of data required for accurate predictions. Specifically, we trained PINN models with

varying numbers of temperature data points (64, 94, 124, and 154) and compared their performance to purely data-driven neural networks to evaluate the benefits of incorporating physics-based constraints.

This is shown in Figure 4, where we assessed these approaches using three representative test cases, each reflecting practical applications in chiplet packaging. The physical setups for the cases were as follows:

- Case 1: all four boundaries of the chip are isothermal at 300 K.
- Case 2: the top and bottom boundaries are isothermal at 300 K, while the left and right boundaries are adiabatic.
- Case 3: the left and bottom boundaries are adiabatic, the top boundary is isothermal, and the right boundary has convective heat transfer conditions.

Case 1 and Case 2 represent scenarios where the boundary conditions are fully known, but the heat source intensity is unknown. In these cases, the PINN model relies on sparse data and the known boundary conditions to infer the heat source intensity distribution. Case 3 represents a scenario where the heat source intensity is known, but the boundary conditions are unknown. Here, the model embeds the heat source information into the physics-informed loss function while inferring the boundary temperature distribution from the available data.

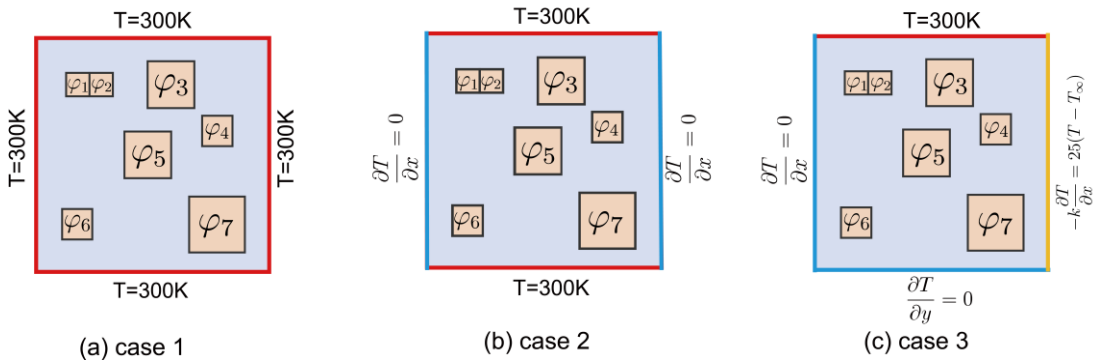


Figure 4. Heat sources and boundary conditions for three cases in the 2D computational domain. The boundaries shown represent the perimeter edges of the planar model. In PINN computation, **(a,b)** have known boundary conditions with unknown heat sources, while **(c)** has a known heat source but unknown boundary conditions.

To ensure uniformity in data distribution and maximize coverage of the solution domain, we adopted low discrepancy sampling (LDS). This method minimizes spatial bias by generating points that are evenly distributed across a region. The uniformity of a sample set $P = \{x_1, \ldots, x_n\}$ is evaluated using the discrepancy metric:

$$D(P) = \sup_{B \in J} \left| \frac{A(B,P)}{N} - \lambda(B) \right|, \tag{29}$$

where $\lambda(B)$ is the Lebesgue measure of sub-region B, A (B, P) is the number of points from P within B, and J is the set of two-dimensional intervals.

By minimizing D(P), LDS ensures that the sample point density in any sub-region is proportional to its area, resulting in a more uniform and representative dataset.

Figure 5 shows the coordinate distribution of temperature data points used in our experiment in Case 1 as an example (including 64, 94, 124, and 154 points). We observed that these sampling points are evenly

distributed across the entire solution domain, with no areas of excessive density or sparsity, validating the effectiveness and appropriateness of the point selection method.

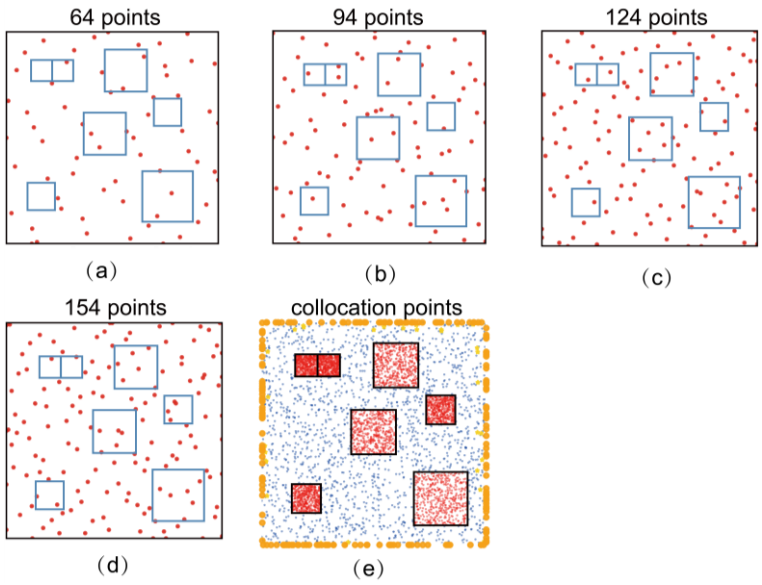


Figure 5. The selection point plot for Case 1 obtained through LDS: 64, 94, 124, and 154 points, The red dots represent sampling points. **(a–d)** Spatial distribution patterns under varying sampling point quantities; **(e)** comprehensive sampling distribution across the entire test set.

After selecting the observation points, we turned our attention to the collocation points used to compute the physical loss functions in the PINN. These collocation points are necessary for solving the governing equations and applying boundary conditions.

- Case 1 and Case 2: In these cases, where the boundary conditions are known but the heat source intensity is unknown, we chose collocation points only along the boundaries of the chip. Specifically, 1000 points were placed on each of the four boundaries to enforce boundary conditions. No collocation points were used within the chip or substrate regions, since the model primarily infers the heat source intensity from boundary information and sparse data.
- Case 3: In this case, where the heat source intensity is known but the boundary conditions are unknown, we selected collocation points within the chip and substrate regions. A total of 1000 points were placed within the chip regions, while 10,000 points were distributed across the substrate regions. This helps the model infer the boundary temperature field while embedding the heat source information directly into the loss function. No boundary points were used in this case, as the model learns the boundary conditions from the data.

Optimization and Hyperparameters: By default, we used the Adam optimizer, with an initial learning rate of 0.008. Each neural network had four hidden layers, each with 64 neurons. The number of iterations was set to 20,000. The computational hardware included an RTX 4090 GPU with 24 GB of memory, paired with a 16-core Xeon® Platinum 8352V CPU.

4.1. Datasets

Both RC-PINN and baseline NN employ fully connected feedforward networks with 4 hidden layers of 64 neurons each. The Swish activation function (defined in Equation 12) is employed throughout

the network, which has been shown to provide smooth gradients beneficial for physics-informed learning. RC-PINN uses separate networks for each material region (chip and substrate), while the baseline NN uses a single unified network covering the entire domain. Training employs the Adam optimizer with initial learning rate of 0.008 for 20,000 iterations. The network architecture follows established practices in PINN literature [30,32], where moderate network depth (4–6 hidden layers) with smooth activation functions provides good balance of approximation capacity and training stability for physics-informed problems.

We used the results from numerical simulations as the training and testing datasets for the PINN model. The chip's dimensions and thermal parameters in the simulations were based on widely used parameters in real-world engineering applications [35]. The substrate material selected was FR4, a commonly used material in electronic chips, while the chip material itself was silicon. The power and heat source intensity for each small chip are listed in Table 1. COMSOL Multiphysics 6.1 was employed to conduct the numerical simulations. Figure 6 shows the temperature field distributions for the three cases after computational convergence. The training datasets consisted of 64, 94, 124, and 154 temperature data points. To clearly demonstrate the model's performance, the test dataset included 10,000 temperature data points, which were selected using a grid-based sampling method to ensure coverage of the entire solution domain.

Table 1. Geometric information and power intensity of the chiplets.

Material No.	1	2	3	4	5	6	7
Side Length (mm)	5	5	10	7.5	10	7.5	12
Power (W)	15	15	35	25	35	25	60
Heat Source Intensity (W/cm^2)	60	60	35	44.4	35	44.4	41.6

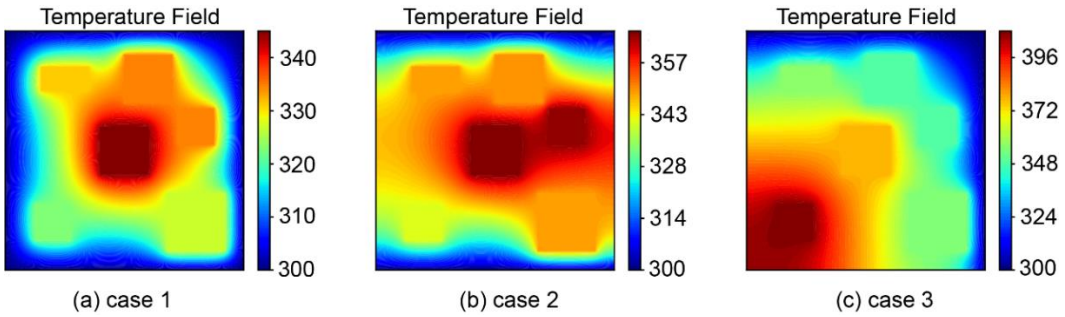


Figure 6. True temperature fields for three cases; **(a–c)** demonstrate the true temperature fields corresponding to Cases 1–3, respectively.

4.2. Results of PINN

During the training process, we observed that the neural network gradually converged after approximately 10,000 to 20,000 iterations. In comparison, standard NN were found to converge more easily than PINN. While incorporating the physical information loss term improved the accuracy of the physical constraints, it also increased the complexity and difficulty of training the network model.

4.2.1. Model performance

Figure 7 illustrates the predicted temperature fields and absolute error distributions for the three experimental cases using PINN. The rows represent results for different numbers of observation points (64, 94, 124, and 154), while the columns correspond to specific cases. The absolute error is calculated as the difference between the PINN predictions and the numerical simulation results.

With as few as 64 temperature data points, PINN effectively captures the overall trend of the chip's temperature field. In Case 1, heat accumulates around Chiplets 3, 4, and 5 due to high heat source intensity under isothermal cooling boundaries, resulting in the highest temperatures. Temperatures within the chiplet region are uniform, while a thermal gradient emerges at the chiplet–substrate interface, driven by thermal conductivity differences. Toward the substrate edges, temperatures gradually decrease.

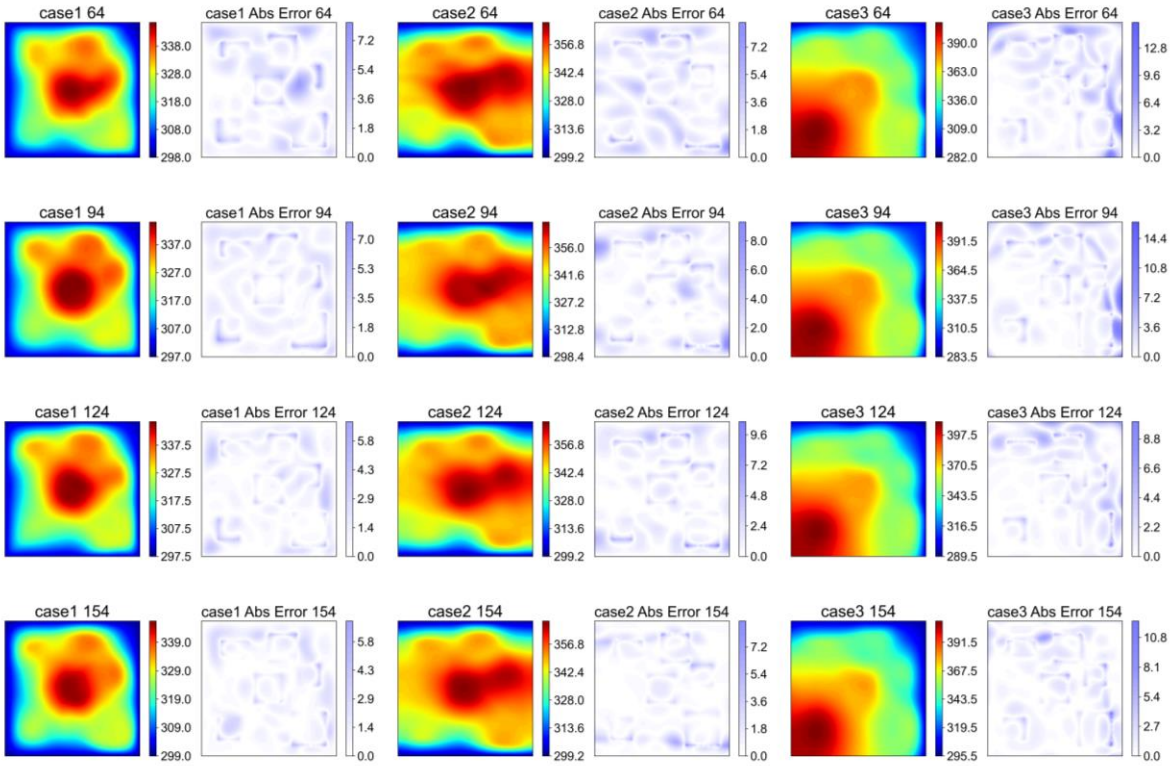


Figure 7. The temperature fields and absolute errors predicted using PINN for the three cases at 64, 94, 124, and 154 observation points.

In Case 2, under adiabatic left and right boundaries and isothermal top and bottom boundaries, heat accumulates on the right side, creating a high-temperature region around Chiplets 4 and 5, aligning with numerical simulations.

In Case 3, with adiabatic left and lower boundaries, poor heat dissipation leads to heat accumulation in the lower-left corner. Conversely, the upper-right area remains cooler due to better heat dissipation conditions.

With only 64 data points, higher prediction errors are primarily observed in regions with sparse temperature data. While precise predictions are more challenging in such areas, the model achieves reasonable accuracy by leveraging physical information. This demonstrates that PINNs, despite

limitations in capturing fine local details with limited data, effectively represent the overall temperature distribution trends in the chip's thermal field.

As the training data increases from 64 to 154 points, the error regions in the prediction maps gradually shrink and lighten, indicating a steady improvement in prediction accuracy. By 124 data points, the absolute error in most regions across all three cases is well controlled, with the maximum error not exceeding 10 K.

Cases 1 and 2 incorporate boundary physical information during the fitting process, resulting in better boundary predictions compared to Case 3. Under 64 data points, the boundary errors for Cases 1 and 2 remain minimal, even in sparse data regions. In contrast, Case 3 exhibits noticeable boundary errors at the upper and right edges with 64 and 94 data points, with some regions exceeding 10 K. These discrepancies are due to the complex boundary conditions in Case 3, which are not embedded in the model.

In Case 3, the known heat source intensity is incorporated into the physical loss, enabling the PINN model to achieve low reconstruction errors in the internal temperature field, even with only 64 data points. For chips 5 and 6, the error distribution is uniform, with no concentrated high-error regions, highlighting the effectiveness of embedding physical information in improving prediction accuracy. By contrast, Cases 1 and 2, which lack accurate heat source integration, exhibit higher errors near the heat source. Additionally, Figure 8 shows that the errors at chips 5 and 6 for Case 3 with 64 and 94 data points are comparable to those in Cases 1 and 2. This indicates that the improved performance in Figure 7 results from leveraging physical information in the governing equation rather than differences between the cases.

Overall, the PINN model exhibits robust performance in all cases, effectively addressing thermal field inversion challenges in chip applications despite minor errors.

To evaluate the role of physical information in improving prediction accuracy, we compared PINN with a baseline NN model. As shown in Figure 8, while NN can capture the general trend of the temperature field, it fails to accurately locate and shape the heat sources. For example, in Case 1, with 94 observation points, NN struggles to distinguish the region between Chip 3 and Chip 4, failing to recognize the heat source boundaries. Even with 154 observation points, NN barely identifies the heat source positions. This limitation stems from NN's inability to handle material interfaces effectively.

The figure also reveals that NN struggles with boundary predictions, showing errors exceeding 20 K across all cases when using 64 observation points. Increasing the data volume does little to alleviate these issues. In Case 3, the boundary error even reaches 40 K with 124 observation points, which is unacceptable in practical applications. By contrast, PINN achieves significantly lower boundary errors in Case 3, even without explicit integration of boundary information. This improvement is due to two factors: (1) embedding heat source information improves overall prediction accuracy, and (2) the partitioned network framework focuses inversion efforts on specific regions, enhancing local performance. While NN's prediction accuracy improves with more observation points, it still shows considerable errors at chip–substrate interfaces and in data-scarce areas.

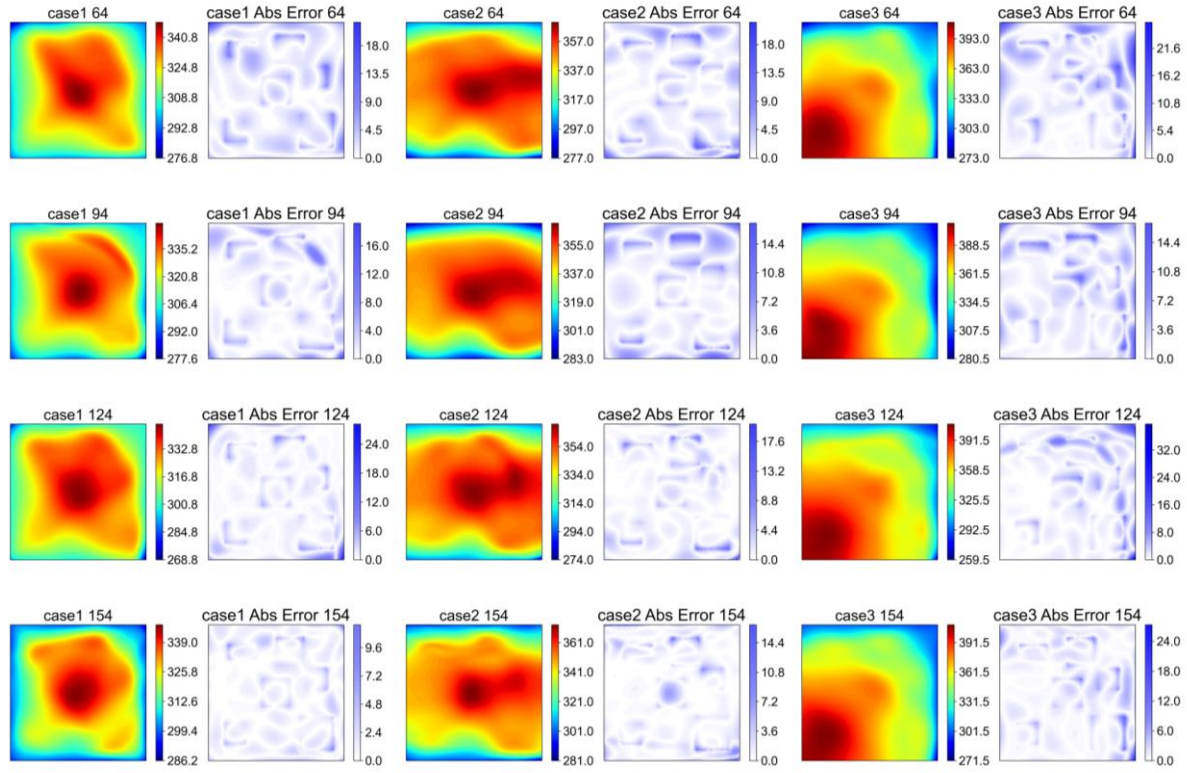
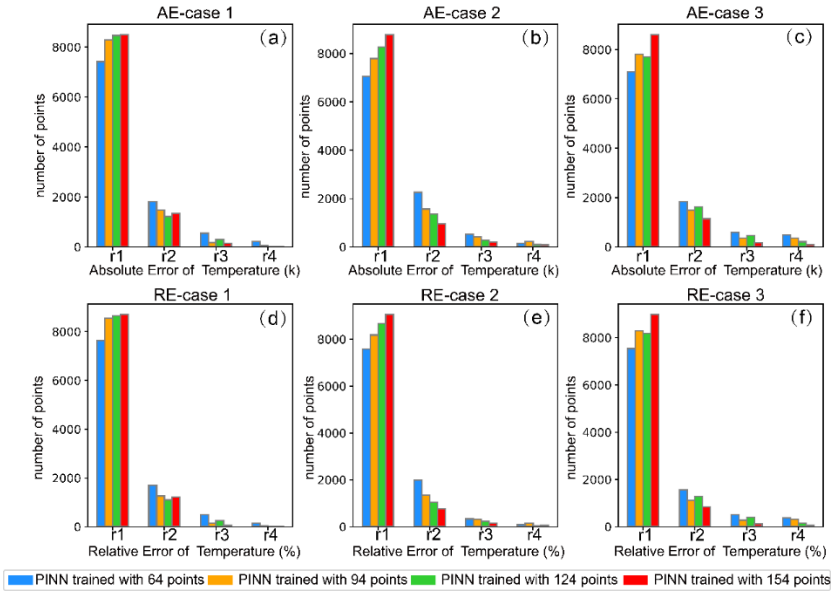


Figure 8. The temperature fields and absolute errors predicted by PINN for the three cases at 64, 94, 124, and 154 observation points, respectively.

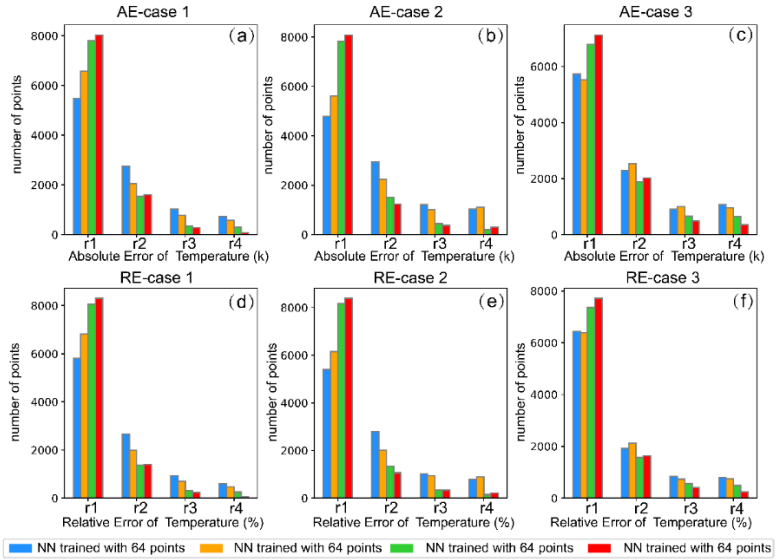
To further evaluate the effects of data volume and physical information on prediction performance, we computed the Absolute Error (AE) and Relative Error (RE) for 10,000 uniformly distributed test points. To avoid data leakage, training points were excluded from the test set. Figures 9 and 10 present the error histograms for PINN and NN, respectively, while Tables 2 and 3 summarize the AE and RE statistics for all three cases under both models.

Figure 9 shows that as the training data volume increases, the number of large-error points decreases, significantly enhancing PINN's accuracy in predicting the chip's temperature field. With more data, the proportion of points with absolute errors below 1.5 K and relative errors under 0.5% steadily grows. In Cases 1 and 2, when the training set reaches 124 points, over 9500 test points have absolute errors below 3 K, and more than 9700 points exhibit relative errors under 1.5%. For the more challenging Case 3, over 9400 points achieve the same accuracy at 124 training points. At 154 training points, more than 9700 points in all three cases have absolute errors below 3 K, with over 9800 points showing relative errors under 1.5%. Notably, in Case 1, this proportion reaches 99%, highlighting PINN's high predictive accuracy.



	r1	r2	r3	r4		r1	r2	r3	r4
AE (k)	0–1.5	1.5–3.0	3.0–4.5	≥ 4.5	RE (%)	0–0.5	0.5–1.5	1–1.5	≥ 1.5

Figure 9. Distribution of Absolute Error (AE) and Relative Error (RE) for PINN on a test set of 10,000 points. (**a**–**c**) AE distributions across the test set sampling points, where intervals r1, r2, r3, and r4 represent the frequency proportions of absolute errors falling within 0–1.5, 1.5–3, 3–4.5, and ≥ 4.5 ranges, respectively. (**d**–**f**) RE distributions across the test set sampling points, with intervals r1, r2, r3, and r4 corresponding to the relative error ranges of 0–0.5, 0.5–1, 1–1.5, and ≥ 1.5, respectively.



	r1	r2	r3	r4		r1	r2	r3	r4
AE (k)	0–1.5	1.5–3.0	3.0–4.5	≥ 4.5	RE (%)	0–0.5	0.5–1.5	1–1.5	≥ 1.5

Figure 10. Distribution plots of Absolute Error (AE) and Relative Error (RE) for PINN on a test set of 10,000 points. (**a**–**c**) AE distributions across the test set sampling points, where intervals r1, r2, r3, and r4 represent the frequency proportions of absolute errors falling within 0–1.5, 1.5–3, 3–4.5, and ≥ 4.5 ranges, respectively. (**d**–**f**) RE distributions across the test set sampling points, with intervals r1, r2, r3, and r4 corresponding to the relative error ranges of 0–0.5, 0.5–1, 1–1.5, and ≥ 1.5, respectively.

As shown in Figure 8, while the NN model performs well in certain regions—such as near chip 5 in Case 2 (154 observation points) and the left boundary in Case 3 (124 observation points)—statistical data in Tables 3 and 4 reveal significant errors across much of the thermal field. This indicates a tendency for NN to overfit, focusing on data-dense regions and improving local accuracy while failing to generalize, resulting in higher errors in data-sparse areas.

Table 2. Absolute Error statistics for PINN and NN models.

Case	Num	PINN				NN			
		0–1.5	1.5–3	3–4.5	≥ 4.5	0–1.5	1.5–3	3–4.5	≥ 4.5
Case 1	64	**7415**	1813	548	224	5473	2758	1040	729
	94	**8290**	1476	180	54	6584	2058	790	568
	124	8460	1226	299	15	7801	1539	363	297
	154	8494	1340	147	19	8083	1610	277	75
Case 2	64	**7062**	2260	533	145	4793	2940	1223	1044
	94	**7791**	1566	423	220	5621	2254	1012	1113
	124	8268	1355	275	102	7829	1518	444	209
	154	8775	951	191	83	8089	1235	372	304
Case 3	64	**7089**	1841	584	486	5732	2285	906	1077
	94	**7814**	1480	358	348	5528	2524	997	951
	124	7705	1628	458	209	6796	1895	655	654
	154	8600	1135	171	94	7129	2018	494	359

When the amount of training data is the same, PINN outperforms NN in temperature prediction, particularly in terms of the number of points with a relative error below 1% and an absolute error below 1.5 K. For all three cases, the number of low-error points is significantly higher for PINN than for NN, especially with 64 data points. Due to the incorporation of physical information, PINN can better compensate for the lack of data in sparse regions. For example, in Case 1, when using 64 data points, the number of points with an absolute error below 1.5 K for PINN exceeds that of NN by 1942; however, as the data increases to 154 points, this gap narrows to 400. This indicates that PINN demonstrates a significant advantage in low-data environments. As the amount of data increases, NN gradually learns more of the underlying relationships among the data, thereby reducing the supplementary effect of physical information on the model. This further highlights the importance and necessity of physical information in low-data conditions.

Table 3. Relative error statistics for PINN and NN models.

Case	Num	PINN				NN			
		0–0.5	0.5–1	1–1.5	≥ 1.5	0–0.5	0.5–1	1–1.5	≥ 1.5
Case 1	64	5807	2657	930	606	5473	2758	1040	729
	94	6828	1988	709	475	6584	2058	790	568
	124	8061	1375	310	254	7801	1539	363	297
	154	8302	1405	240	53	8083	1610	277	75
Case 2	64	5396	2805	1016	783	4793	2940	1223	1044
	94	6152	2011	953	884	5621	2254	1012	1113
	124	8182	1328	339	151	7829	1518	444	209
	154	8389	1065	331	215	8089	1235	372	304
Case 3	64	6448	1920	840	792	5732	2285	906	1077
	94	6378	2129	738	755	5528	2524	997	951
	124	7360	1575	568	497	6796	1895	655	654
	154	7715	1634	410	241	7129	2018	494	359

Table 4. Performance of PINN and NN on MAE, MRE, and R^2 metrics under different data conditions.

Model	Num	Case 1			Case 2			Case 3		
		MAE (k)	MRE (%)	R^2	MAE (k)	MRE (%)	R^2	MAE (k)	MRE (%)	R^2
PINN	64	1.1413	0.3512	0.9829	1.2107	0.3543	0.9904	1.3451	0.3934	0.9940
	94	0.9172	0.2843	0.9904	1.0178	0.2972	0.9913	1.1254	0.3293	0.9953
	124	0.8479	0.2629	0.9912	0.8987	0.2637	0.9940	1.1685	0.3421	0.9959
	154	0.7772	0.2405	0.9923	0.7109	0.2074	0.9955	0.8116	0.2292	0.9981
NN	64	1.9297	0.6023	0.9541	2.1982	0.6504	0.9572	1.9807	0.5752	0.9874
	94	1.4642	0.4565	0.9694	1.9759	0.5877	0.9704	1.9204	0.5516	0.9902
	124	1.1408	0.3566	0.9751	1.0627	0.3107	0.9901	1.5280	0.4451	0.9916
	154	0.9428	0.2929	0.9890	1.0028	0.2945	0.9900	1.2828	0.3660	0.9951

4.2.2. Model performance metrics

To evaluate the performance of NN and PINN in reconstructing the temperature field, we adopted three widely used metrics in regression tasks: mean absolute error (MAE), mean relative error (MRE), and coefficient of determination (R^2). The formulas for these metrics are as follows:

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m} | T_i - \hat{T}_i |, \tag{30}$$

$$\mathrm{MRE} = \frac{1}{m}\sum_{i=1}^{m} \frac{| T_i - \hat{T}_i |}{T_i}, \tag{31}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}\left(T_i - \hat{T}_i\right)^2}{\sum_{i=1}^{m}(T_i - \bar{T})^2}, \tag{32}$$

where m is the total number of test data points, T_i represents the observed (ground truth) temperature at the i-th coordinate, $\hat{T}_i$ is the corresponding predicted temperature, and $\bar{T}$ is the average observed temperature. In this study, the test dataset consists of 10,000 data points.

These three metrics jointly assess the model's performance in temperature field reconstruction, providing a comprehensive evaluation by considering absolute error, relative error, and the coefficient of determination.

As shown in Figure 11, with increasing training data volume, MAE and MRE gradually decrease, while R^2 approaches 1, indicating improved prediction accuracy. Both PINN and NN rely heavily on training data as the primary information source, with larger datasets enabling more accurate temperature field predictions.

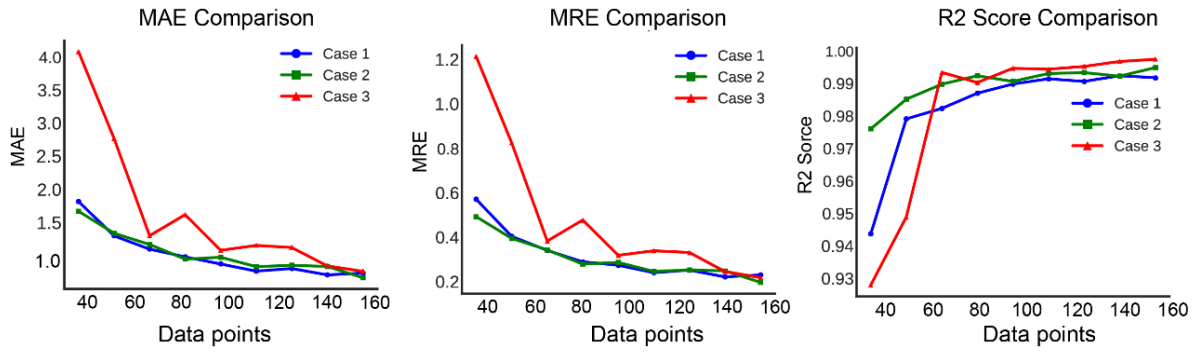


Figure 11. Illustrates the trend of MAE, MRE, and R^2 scores as the data volume increases across the three cases.

Between 64 and 79 data points, slight fluctuations in MAE, MRE, and R^2 were observed in Case 3, caused by randomness in sample point selection and the neural network training process. These variations stem from the stochastic nature of non-convex optimization, where different sampling points may lead to varying regional performance. However, the overall trend of improved accuracy with increasing data remains unaffected.

Additionally, as data points increased from 34 to 64, the model accuracy in Case 3 improved significantly. This can be attributed to the more complex temperature field distribution in Case 3, which requires a greater number of data points to capture its nonlinear characteristics. When data is limited, the model struggles to learn these complex relationships. However, with an increase to 64 data points, the model can better fit these nonlinear distributions, resulting in a rapid reduction in error. This finding also suggests that embedding the governing equations, as compared to boundary conditions, may impose higher data requirements.

Table 4 compares the performance of PINN and NN on the test set across the three cases. Under the same data volume, PINN consistently outperforms NN in MAE, MRE, and R^2. For instance, with 64 data points, PINN achieves a 44.92% lower MAE than NN in Case 2 and reductions of 40.8% and 32.09% in Cases 1 and 3, respectively. Notably, across all three cases, PINN with only 94 data points surpasses NN with 154 data points on most metrics, highlighting its superior efficiency in data-scarce scenarios.

The R^2 score is a comprehensive metric that reflects the overall performance of a model. Given MAE on the test set, a model with smaller error variance is more likely to achieve a higher R^2 score. Thus, both lower average error and smaller error variance contribute to a higher R^2 score. When evaluated under the same data volume, PINN demonstrates a clear advantage in achieving higher R^2 scores on the test set. Specifically, with only 64 data points, the R^2 scores for the three cases using PINN reach 0.9829, 0.9904, and 0.9940, respectively—significantly outperforming NN. Although the Mean Relative Error (MRE) for Case 2 under the PINN model is slightly larger than that of NN at 154 data points, the higher R^2 score indicates that PINN produces fewer large-error points, leading to lower variance in its prediction results. This observation suggests that the inclusion of physics-informed loss functions during training enhances the robustness of the PINN model.

In practical applications of thermal field inversion, the primary cost often stems from data acquisition. Our proposed approach can reduce data acquisition costs by nearly 50% while maintaining high prediction accuracy. For scenarios where minimizing training data is a priority, we find that training PINN with only 64 temperature data points is sufficient to accurately capture the thermal field variations in a chip, achieving an MRE below 0.4% and an R^2 score exceeding 0.98—ensuring highly reliable predictions. For applications requiring even greater accuracy, training with 124 or 154 data points is recommended, as the model's performance stabilizes at a satisfactory level with 124 data points.

4.2.3. Performance of the model after data augmentation using GP-UGS

In previous sections, we demonstrated the effectiveness of the RC-PINN in solving thermal field inversion problems for chiplet packaging. However, given the high cost of data acquisition in real-world applications, achieving superior performance with minimal data remains a critical goal. To address this, we propose a Gaussian process regression-based uncertainty filtering method to strategically select data points for enhancing and exploring the original dataset. This approach not only improves model accuracy but also minimizes additional computational cost.

In this experiment, we applied Gaussian process regression with a radial basis function kernel to a small set of observed data points. From the fitted results, 100 points were uniformly sampled and ranked by uncertainty. The 25 points with the lowest uncertainty were chosen as enhancement points to improve local accuracy, while 5 points with moderate uncertainty (ranked 40th to 60th) were selected as exploration points to expand the model's understanding of less-explored regions. This 5:1 ratio of low uncertainty to exploration points ensures a balance between precision enhancement and domain exploration.

Table 5 shows the performance metrics for GP-UGS on small datasets. Compared to the original PINN model, the GP-UGS approach improves performance overall. The only exception is Case 3 at 94 points, where MAE and MRE slightly increase; however, the R^2 value also improves. As mentioned earlier, when the error variance is lower, the model is more likely to achieve a higher R^2 score, indicating that GP-UGS enhances model stability. This method simply adds data augmentation on top of the original PINN without making any other modifications. With just 34 data points, the prediction accuracy of the GP-UGS method almost reaches that of the NN with 124 data points. Similarly, for Case 2, accuracy at 34 points surpasses that of NN at 94 points. Notably, due to the complexity of the boundary conditions and the absence of embedded boundary information, Case 3 requires higher data accuracy and volume. Yet, even in this scenario, the GP-UGS method still improves prediction accuracy by over 25% compared to the PINN with 34 data points, representing a significant advancement for practical engineering applications.

Table 5. Model performance of RC-PINN after data augmentation using GP-UGS.

Num	MAE	Case 1		Case 2			Case 3		
		MRE	R^2	MAE	MRE	R^2	MAE	MRE	R^2
34	1.1084	0.3426	0.9855	1.3805	0.4055	0.9858	3.0198	0.9042	0.9634
64	0.8977	0.2787	0.9902	1.0178	0.2972	0.9913	1.2895	0.3708	0.9952
94	0.8440	0.2600	0.9918	0.8987	0.2637	0.9940	1.1489	0.3351	0.9955

To better visualize the method's impact on model performance, we compared the results of neural networks, Gaussian processes, and a typical machine learning algorithm, Random Forest, as shown in Figure 12.

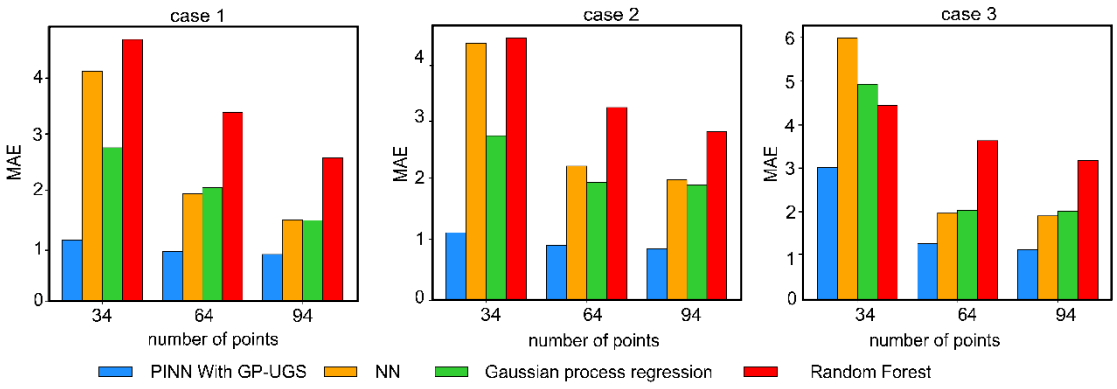


Figure 12. Comparing MAE for three cases at 34, 64, and 94 points.

Figure 12 displays the MAE values for the three cases at different data points, with lower bars indicating better performance. Apart from Case 3 at 34 points, where the MAE exceeds 1.5, all other scenarios remain below this threshold, performing better than other machine learning methods with the same data volume. At 34 points, besides the model with GP-UGS, Gaussian process regression also shows favorable fitting performance. When the data count increases to 64 points or more, the learning outcomes of neural networks and Gaussian processes become comparable, while Random Forest consistently underperforms. Furthermore, the learning accuracy of all three machine learning models improves significantly with more data, as the addition of each data point yields substantial performance gains in low-data settings. In contrast, the GP-UGS method achieves good performance with only 34 data points. The inclusion of physical information and the data augmentation approach enable the model to grasp underlying patterns beyond the dataset. Although accuracy increases with more data, the improvement rate is relatively small, indicating that the method is effective with minimal data, substantially reducing data acquisition costs.

The training duration serves as a critical performance metric for neural network models. In this study, we systematically recorded the training times of NN, RC-PINN, and RC-PINN with GP-UGS (UGS-PINN). The results are shown in Table 6. Experimental results demonstrate that our proposed RC-PINN framework exhibits moderately increased computational time compared with conventional NN implementations. This outcome is anticipated, as RC-PINN architectures inherently require processing additional physics-informed loss terms that encode domain-specific knowledge. The observed computational overhead, while non-negligible, remains justifiable given the framework's significant improvements in prediction accuracy. Furthermore, the UGS-PINN introduces only marginal temporal increments relative to RC-PINN, indicating that the integration of sampling point selection strategy exerts limited impact on training efficiency. Notably, this strategic enhancement enables substantial accuracy gains, suggesting an effective trade-off between computational expenditure and model performance enhancement.

Table 6. Training duration of NN, RC-PINN, and RC-PINN with GP-UGS (UGS-PINN).

Num	NN	Case 1 (min)		Case 2 (min)			Case 3 (min)		
		RC-PINN	UGS-PINN	NN	RC-PINN	UGS-PINN	NN	RC-PINN	UGS-PINN
34			12.3			13.6			19.6
64	11.2	14.5	14.7	12.3	16.0	18.9	11.2	25	27.1
94	12.1	19.3	20.4	13.8	18.2	20.3	12.6	26.3	27.5
124	12.9	21.3		14.2	20.1		14.7	27.9	
154	13.9	23.2		16.6	21.9		15.3	30	

Compared to altering network structures or adopting complex algorithm designs, the GP-UGS method is straightforward, computationally efficient, and flexible. With just a few steps, it can significantly improve performance on existing models.

4.3. Solve the thermal conductivity parameters

In this subsection, we employed RC-PINN to solve the inverse problem of thermal conductivity identification. For this problem, we adopted the boundary conditions and heat source configuration from Case 1, while specifically setting the thermal conductivity of chip 1 as a trainable variable integrated into the optimizer with a learning rate of 1×10^{-3}. All other parameters remained fixed, with the true thermal

conductivity of the chip being 150 W/(m·K). Considering the specific requirements of this inverse problem, we investigated solutions using datasets containing 64, 94, 124, and 154 measurement points. As previously mentioned, low-discrepancy sampling was implemented to ensure uniform spatial coverage across the chip surface. After over 12,000 training iterations, the results presented in Table 7 demonstrate the effectiveness of RC-PINN for parameter inversion. Specifically, the relative errors fell below 5% when using 124 and 154 measurement points. Furthermore, the results indicate that increasing sample size significantly improved parameter estimation accuracy while only moderately increasing computational time, thereby establishing a favorable trade-off between measurement effort and solution precision.

Table 7. The result of solving the thermal conductivity.

Number of Data	K	MAE	MRE (%)	Time (min)
64	164.3	14.3	9.53	19.3
94	158.4	8.4	5.60	22.6
124	143.7	6.3	4.3	24.5
154	**148.6**	**1.4**	**0.93**	**25.7**

5. Conclusions

Chiplet packaging temperature field prediction is critical for thermal management and maintenance. In real-world applications, environmental complexity and measurement constraints make it impossible to obtain accurate and complete physical information. Only a limited number of temperature measurement points can be used for predicting the thermal field, making it a typical inverse problem. To address this challenge, we propose a RC-PINN and Gaussian Process Uncertainty-Guided Sampling. The RC-PINN divides the chip region into multiple sub-regions and introduces a coupling mechanism for material property differences, effectively handling different physical characteristics for precise multi-region thermal field inversion. The GP-UGS method generates new sampling points through Gaussian process regression and enhances data selection by incorporating low and medium uncertainty points, significantly improving model learning ability and generalization in small data settings. Experimental results show superior performance across different inversion scenarios. Data diversity robustness is quantitatively verified through deliberately substituted training protocols: the methodology successfully trains models using Gaussian regression outputs as pseudo-ground truth measurements, with empirical results confirming that physical equation constraints inherently rectify data inaccuracies through multi-physics regularization effects.

Through two problem types and three real-world cases with distinct physical conditions, we validate the model's superior performance and demonstrate the importance of introducing continuity conditions at material boundaries for thermal field inversion. The PINN model with boundary information requires fewer data than a control equation-based PINN because boundary condition losses are more straightforward. The performance improvement of PINN under low data volume is more pronounced since physical information compensates for data deficiencies. While this study employs 2D planar models with clean simulation data to establish fundamental methodologies, wacknowledge that practical deployment requires extensions in three key directions. First, the RC-PINN framework is extensible to 3D geometries through expanded network inputs and 3D surface coupling, where top and bottom boundaries (heat sink and substrate interfaces) will be explicitly represented to enable comprehensive analysis of both lateral and

vertical thermal coupling, though computational cost and data acquisition remain challenges. Second, future work will extend the framework to time-dependent problems by incorporating temporal derivatives into the governin gequations and treating time as an additional network input dimension, enabling the reconstruction of dynamic thermal transients during workload changes essential for comprehensive thermal management of modern chiplet systems. Third, noise robustness validation is crucial for real-world deployment. While the framework's inherent mechanisms—GP noise parameters (σ_n^2) in the likelihood model and physics-based regularization—suggest robustness to measurement noise, comprehensive experimental validation with realistic noise levels (e.g., 1%–5% > Gaussian noise typical of infrared thermography) represents important future work.

Additionally, the uncertainty-based selection method using Gaussian process regression exhibits significant potential for addressing small data problems. Although this investigation employed two-dimensional thermal field simulations and neural network training for validation, this methodological choice stems from the theoretical modeling perspective where 2D configurations provide sufficient rigor for proof-of-concept demonstration while maintaining PINN training tractability, rather than reflecting inherent dimensional limitations of the proposed framework. The principal challenges in extending to three-dimensional scenarios reside in the comprehensive modeling requirements encompassing multi-physics phenomena such as circuit interconnects, material heterogeneities, and interfacial thermal dynamics. Remarkably, the increased incompleteness of governing equations under such complex 3D conditions amplifies the practical value of our hybrid data–physics approach in resolving underdetermined inverse problems. The approach is simple, computationally efficient, and flexible, allowing for easy adjustment of the proportion and number of enhanced and exploration points, regardless of the task. Therefore, this method can be applied to any small data scenario for performance improvement through data augmentation.

In theoretical solutions of partial differential equations systems, incomplete boundary conditions, interface modeling, or equation parameters inevitably induce non-uniqueness of solutions. However, in the engineering context of thermal field monitoring addressed in this work, the introduction of ground truth sampling circumvents this challenge. The incompleteness of the PDE system primarily manifests as increased dependence on data sampling requirements—specifically, the degree of system incompleteness directly correlates with the quantity of required experimental data points to ensure solution uniqueness. Thus, theoretically, the proposed PINN method can be adapted to solve any heat conduction-related task, particularly those involving complex multi-material interactions. Overall, this study indicates that Physical Information Neural Networks have great potential in chip thermal field inversion. By integrating physical constraints and data-driven methods, we effectively address the ill-posed nature of inverse problems. Future work will explore multi-physics coupling issues, more complex boundary conditions, and theoretical considerations regarding the ratio of data points selected through Gaussian process regression to further enhance model applicability and accuracy in engineering applications.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. The critical code is available in the linked GitHub repository (YupengQI99/RC-PINN-YuPeng-DengPan).

Acknowledgments

This work is supported by the National Key R&D Program of China (Grant No. 2023YFB4402600), and the Shanghai Technical Service Center of Science and Engineering Computing, Shanghai University.

Authors' contribution

Conceptualization, Y.Q and D.P; methodology, Y.Q; software, Y.Q; investigation, D.P and J.X; resources, Y.Q; data curation, D.P.; writing—original draft preparation, Y.Q; writing—review and editing, J.X., W.Y. and D.P; supervision, D.P; project administration, D.P.; funding acquisition, D.P and S.H. All authors have read and agreed to the published version of the manuscript.

Conflicts of interests

The authors declare no conflict of interest.

References

[1] Kim J, Murali G, Park H, Qin E, Kwon H, *et al.* Architecture, chip, and package codesign flow for interposer-based 2.5-D chiplet integration enabling heterogeneous IP reuse. *IEEE Trans. VLSI Syst.* 2020, 28:2424–2437.

[2] Ravikumar CP. Chiplets for integration of electronic systems. *IETE J. Educ.* 2024, 65:92–108.

[3] Li T, Hou J, Yan J, Liu R, Yang H, *et al.* Chiplet heterogeneous integration technology—status and challenges. *Electronics* 2020, 9:670.

[4] Nie C, Xu Q, Chen L. Equivalent thermal model of through silicon via and bump for advanced packaging of integrated circuits. *Microelectron. Reliab.* 2022, 137:114790.

[5] Song Y, Fu R, Chen C, Wang Q, Su M, *et al.* Case-embedded cooling for high heat flux microwave multi-chip array. *Appl. Therm. Eng.* 2022, 214:118852.

[6] Chen C, Hou F, Ma R, Su M, Li J, *et al.* Design, integration and performance analysis of a lid-integral microchannel cooling module for high-power chip. *Appl. Therm. Eng.* 2021, 198:117457.

[7] Ye Y, Jiao B, Kong Y, Liu R, Du X, *et al.* Experimental investigations on the thermal superposition effect of multiple hotspots for embedded microfluidic cooling. *Appl. Therm. Eng.* 2022, 202:117849.

[8] Jang KW. A study on thermal cycling (T/C) reliability of anisotropic conductive film (ACF) flip chip assembly for thin chip-on-board (COB) packages. *Microelectron. Reliab.* 2012, 52:1174–1181.

[9] Hatakeyama T, Ishizuka M, Nakagawa S. Estimation of maximum temperature in 3D-integrated package by thermal network method. In *Proceedings of the 2010 12th Electronics Packaging Technology Conference*, Singapore, December 8–10, 2010, pp. 68–71.

[10] Jiang X, Lin P, Song Y, Huang Y, Lian B, *et al.* Effect of temperature cycling on reliability of flip chip solder joint. In *Proceedings of the 2014 15th International Conference on Electronic Packaging Technology*, Chengdu, China, August 12–15, 2014, pp. 989–991.

[11] Liu Z, Li J, Liu X. Novel functionalized BN nanosheets/epoxy composites with advanced thermal conductivity and mechanical properties. *ACS Appl. Mater. Interfaces* 2020, 12:6503–6515.

[12] Chen Y, Xie B, Long J, Kuang Y, Chen X, *et al.* Interfacial laser-induced graphene enabling high-performance liquid−solid triboelectric nanogenerator. *Adv. Mater.* 2021, 33:2104290.
[13] Parameswaran S, Balakrishnan S, Ang B. On-die thermal evaluation system. In *Proceedings of the 2017 IEEE International Conference on Microelectronic Systems Education (MSE)*, Lake Louise, Canada, May 11–12, 2017, pp. 55–58.
[14] Lee Y. Thermal-aware design and management of embedded real-time systems. In *Proceedings of the 2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, Grenoble, France, February 1–5, 2021, pp. 1252–1255.
[15] Evans MN, Kaplan A, Cane MA. Pacific sea surface temperature field reconstruction from coral $\delta^{18}O$ data using reduced space objective analysis. *Paleoceanography* 2002, 17(1):7-1–7-13.
[16] Lu D, Wang C. Three-dimensional temperature field inversion calculation based on an artificial intelligence algorithm. *Appl. Therm. Eng.* 2023, 225:120237.
[17] Yan M, Wei H, Onabajo M. On-chip thermal profiling to detect malicious activity: system-level concepts and design of key building blocks. *IEEE Trans. VLSI Syst.* 2021, 29:530–543.
[18] Nath V, Mandal JK, Eds. *Proceedings of the Second International Conference on Microelectronics, Computing & Communication Systems (MCCS 2017)*. Lecture Notes in Electrical Engineering; Singapore: Springer, 2019, volume 476.
[19] Zhang Y, Srivastava A, Zahran M. Chip level thermal profile estimation using on-chip temperature sensors. In *Proceedings of the 2008 IEEE International Conference on Computer Design*, Lake Tahoe, USA, October 12–15, 2008, pp. 432–437.
[20] Kunze H, Torre DL, Galán MR. Optimization methods in inverse problems and applications to science and engineering. *Optim. Eng*. 2021, 22:2151–2158.
[21] Sun W, Ma H, Qu W. A hybrid numerical method for non-linear transient heat conduction problems with temperature-dependent thermal conductivity. *Appl. Math. Lett.* 2024, 148:108868.
[22] Sun W, Qu W, Gu Y, Li PW. An arbitrary order numerical framework for transient heat conduction problems. *Int. J. Heat Mass Transf.* 2024, 218:124798.
[23] Ku CY, Liu CY, Xiao JE, Hsu SM, Yeih W. A collocation method with space–time radial polynomials for inverse heat conduction problems. *Eng. Anal. Bound. Elem.* 2021, 122:117–131.
[24] Li S, Eds. *Computational and Experimental Simulations in Engineering: Proceedings of ICCES 2023—Volume 3*. Mechanisms and Machine Science; Cham: Springer Nature, 2024, volume 146.
[25] Mai J, Li Y, Long L, Huang Y, Zhang H, *et al*. Two-dimensional temperature field inversion of turbine blade based on physics-informed neural networks. *Phys. Fluids* 2024, 36:037114.
[26] Habermann C, Kindermann F. Multidimensional spline interpolation: theory and applications. *Comput. Econ.* 2007, 30:153–169.
[27] Parish EJ, Duraisamy K. A paradigm for data-driven predictive modeling using field inversion and machine learning. *J. Comput. Phys.* 2016, 305:758–774.
[28] Kabanikhin SI. Definitions and examples of inverse and ill-posed problems. *J. Inverse Ill-Posed Probl.* 2008, 16:317–357.
[29] Bai Y, Chen W, Chen J, Guo W. Deep learning methods for solving linear inverse problems: research directions and paradigms. *Signal Process.* 2020, 177:107729.

[30] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* 2019, 378:686–707.

[31] Wei C, Ooka R. Indoor airflow field reconstruction using physics-informed neural network. *Build. Environ.* 2023, 242:110563.

[32] Rui EZ, Chen ZW, Ni YQ, Yuan L, Zeng GZ. Reconstruction of 3D flow field around a building model in wind tunnel: A novel physics-informed neural network framework adopting dynamic prioritization self-adaptive loss balance strategy. *Eng. Appl. Comput. Fluid Mech.* 2023, 17:2238849.

[33] Fowler E, McDevitt CJ, Roy S. Physics-informed neural network simulation of thermal cavity flow. *Sci. Rep*. 2024, 14:15203.

[34] Franklin TS, Souza LS, Fontes RM, Martins MA. A Physics-Informed Neural Networks (PINN) oriented approach to flow metering in oil wells: an ESP lifted oil well system as a case study. *Digit. Chem. Eng.* 2022, 5:100056.

[35] Zhang Z, Lin C, Wang B. Physics-informed shape optimization using coordinate projection. *Sci. Rep*. 2024, 14:6537.

[36] Khatamsaz D, Neuberger R, Roy AM, Zadeh SH, Otis R, *et al*. A physics informed bayesian optimization approach for material design: Application to NiTi shape memory alloys. *npj Comput. Mater.* 2023, 9:221.

[37] Liu X, Peng W, Gong Z, Zhou W, Yao W. Temperature field inversion of heat-source systems via physics-informed neural networks. *Eng. Appl. Artif. Intell.* 2022, 113:104902.

[38] Cai S, Wang Z, Wang S, Perdikaris P, Karniadakis GE. Physics-informed neural networks for heat transfer problems. *J. Heat Transf.* 2021, 143:060801.

[39] Rasht‐Behesht M, Huber C, Shukla K, Karniadakis GE. Physics-informed neural networks (PINNs) for wave propagation and full waveform inversions. *JGR Solid Earth* 2022, 127:e2021JB023120.

[40] Li P, Liu M, Alfarraj M, Tahmasebi P, Grana D. Probabilistic physics-informed neural network for seismic petrophysical inversion. *Geophysics* 2024, 89:M17–M32.

[41] Haghighat E, Raissi M, Moure A, Gomez H, Juanes R. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. *Comput. Methods Appl. Mech. Eng.* 2021, 379:113741.

[42] Zhang E, Dao M, Karniadakis GE, Suresh S. Analyses of internal structures and defects in materials using physics-informed neural networks. *Sci. Adv.* 2022, 8:eabk0644.

[43] Jagtap AD, Karniadakis GE. Extended physics-informed neural networks (XPINNs): a generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. *Commun. Comput. Phys.* 2020, 28(5):2002–2041.

[44] Liu Y, Chen L, Ding J, Chen Y. An adaptive sampling method based on expected improvement function and residual gradient in PINNs. *IEEE Access* 2024, 12:92130–92141.

[45] Zhou C, Chen J, Yang Z, Matyasko A, Png CE. Enhanced physics-informed neural networks with optimized sensor placement via multi-criteria adaptive sampling. In *Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN)*, Yokohama, Japan, June 30–July 5, 2024, pp. 1–8.

[46] Bauduin V, Cuomo S, Di Cola VS. Impact of collocation point sampling techniques on PINN performance in groundwater flow predictions. *J. Comput. Math. Data Sci.* 2025, 14:100107.
[47] Celaya A, Fuentes D, Riviere B. An adaptive collocation point strategy for physics informed neural networks via the QR discrete empirical interpolation method. *arXiv* 2025, arXiv:2501.07700.
[48] Park KV. AL-PINN: active learning-driven physics-informed neural networks for efficient sample selection in solving partial differential equations. *arXiv* 2025, arXiv:2502.03963.
[49] Feng Y, Wu L. Analysis of interfacial thermal stresses of chip-substrate structure. *Int. J. Solids Struct.* 2001, 38:1551–1562.
[50] Kendall A, Gal Y, Cipolla R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. *arXiv* 2018, arXiv:1705.07115.