Physics-informed deep learning for three-dimensional transient heat transfer analysis of functionally graded materials

Hongwei Guo[1,2,4] · Xiaoying Zhuang[1,2,4] · Naif Alajlan[3] · Timon Rabczuk[3]

Abstract

We present a physics-informed deep learning model for the transient heat transfer analysis of three-dimensional functionally graded materials (FGMs) employing a Runge–Kutta discrete time scheme. Firstly, the governing equation, associated boundary conditions and the initial condition for transient heat transfer analysis of FGMs with exponential material variations are presented. Then, the deep collocation method with the Runge–Kutta integration scheme for transient analysis is introduced. The prior physics that helps to generalize the physics-informed deep learning model is introduced by constraining the temperature variable with discrete time schemes and initial/boundary conditions. Further the fitted activation functions suitable for dynamic analysis are presented. Finally, we validate our approach through several numerical examples on FGMs with irregular shapes and a variety of boundary conditions. From numerical experiments, the predicted results with PIDL demonstrate well agreement with analytical solutions and other numerical methods in predicting of both temperature and flux distributions and can be adaptive to transient analysis of FGMs with different shapes, which can be the promising surrogate model in transient dynamic analysis.

Keywords Deep learning · Physics-informed · Activation function · Heat transfer · Functionally graded materials · Discontinuous time scheme

✉ Timon Rabczuk
    timon.rabczuk@uni-weimar.de

Hongwei Guo
hongwei.guo@iop.uni-hannover.de

Xiaoying Zhuang
zhuang@iop.uni-hannover.de

Naif Alajlan
najlan@ksu.edu.sa

1   Department of Geotechnical Engineering, Tongji University,
    1239 Siping Road, Shanghai 200092,
    People's Republic of China

2   Key Laboratory of Geotechnical and Underground
    Engineering of Ministry of Education, Tongji University,
    1239 Siping Road, Shanghai 200092,
    People's Republic of China

3   ALISR Laboratory, College of Computer and Information
    Sciences, King Saud University, P. O. Box 51178, 11543
    Riyadh, Saudi Arabia

4   Computational Science and Simulation Technology, Institute
    of Photonics, Leibniz Universität Hannover, Appelstr. 11,
    30167 Hannover, Germany

1 Introduction

Factionally graded materials (FGMs) are innovative materials with a spatial variation in composition and/or microstructure for the purpose of controlling variations in physical properties. Due to their excellent thermal properties, functionally graded materials (FGMs) have been widely used in the high temperature environments such as aerospace engineering (e.g. as thermal barrier coatings for aerospace structures [1]), microelectronics, power generation [2].

Various numerical models have been developed for solving transient heat conduction problems, such as such as finite-difference method (FDM) [3–5], finite-element method (FEM) [6,7], meshless method [8–10], boundary element method (BEM) [11–13], localized Trefftz-based collocation method [14], to mention but a few. Recently, Fu et al. [15] summarized localized collocation methods (LCMs) and introduced the application of LCMs to solve heat conduction problems in those nonhomogeneous materials.

Apart from those traditional numerical methods, machine learning offers another novel opportunity to solve complex partial differential equations (PDEs). Such approaches can be

traced back to the seminal work of Lagaris et al. [16] in 1997. However, such approaches gained in popularity only recently, probably due to advancements in machine learning techniques and associated open-sources tools such as TensorFlow or Pytorch. Some recent and innovative applications with artificial neural networks are summarized as follows: Khatir et al. [17] proposed two-stage approaches to study damage detection, localization and quantification in Functionally Graded Material (FGM) plate structures, with IsoGeometric Analysis (IGA) for modelling while using an improved damage indicator based on Frequency Response Function (FRF) for damage elements identification in first stage and improved Artificial Neural Network using Arithmetic Optimization Algorithm (IANN-AOA) for damage quantification problem in the second stage. Wang et al. [18] proposed a novel and intelligent algorithm based on deep learning to realize the recognition of different types of rail profiles and achieve rapid tracking of the railhead laser stripe. Ho et al. [19] combined feedforward neural networks and marine predator algorithm for structural health monitoring in different scenarios including a simply supported beam, a two-span continuous beam, and a laboratory free-free beam. Based on a coupled model between an artificial neural network (ANN) and antlion optimizer (ALO), Ho et al. managed to localize damages in fixed-free plate structures based on mode shape derivative based damage identification index [20].

Raissi et al. studied physics-informed machine learning by encoding physics with kernel matrix in Gaussian Processes [21,22]. The physics-informed Gaussian Processes were applied in solving linear and nonlinear differential equations. They [23,24] later introduced a physical informed neural networks for supervised learning of nonlinear partial differential equations such as the Burger's equations or Navier–Stokes equations, see also their recent contributions in [25]. Two distinct models were tailored for spatio-temporal datasets: continuous time and discrete time models. Their physical informed neural networks were successfully applied in solving coupled high-dimensional forward-backward stochastic differential equations. The convergence behaviour of physics informed neural networks was studied by Shin et al. [26]. Fu et al. [27] proposed an extrinsic approach based on physics-informed neural networks (PINNs) for solving the partial differential equations on surfaces embedded in high dimensional space which manifested good accuracy and higher efficiency compared with the embedding approach. Karniadakis et al. [25] gave a comprehensive review on physic-informed machine learning framework and summarized the general approaches of introducing physics in machine learning framework and introduced some of the latest applications of physics-informed machine learning.

However, for the transient analysis with physics-informed neural networks, the current works mostly solve the PDEs with a continuous time model and applications are often limited in simple one-dimensional cases [28–30]. Yu et al. [30] applied PINN and extend physical-informed Neural Networks (XPINN) in solving steady and transient heat conduction problems in FGMs based on a continuous time scheme, for the transient analysis of 2D FGMs, however, only the radial coordinate and temporal coordinate are considered. Raissi et al. [21] pointed out that the continuous time model needs a large amount of collocation points in the entire spatio-temporal domain which makes the training prohibitively expensive, and typically fails to handle long-time prediction tasks [31].

In this study, we suggest the physics informed deep learning based collocation method with a discrete time scheme that avoids extra training data from simulations for three dimensional transient heat conduction analysis of FGMs. We also propose fitted activation functions suitable for transient heat transfer analysis. The proposed model will then be validated through several numerical examples. The remainder of this paper is organized as follows. In Sect. 2, we describe the physical model we present our deep collocation method in Sect. 3. Section 4 contains several numerical examples to demonstrate the performance of our approach before the manuscript concludes in Sect. 5.

2 Transient heat transfer analysis in 3D FGMs

The general transient diffusion equation for functional graded materials can be written as:

$$\nabla(k(\boldsymbol{x},t)\nabla T(\boldsymbol{x},t)) = c(\boldsymbol{x},t)\frac{\partial T(\boldsymbol{x},t)}{\partial t}, \boldsymbol{x} \in \Omega, 0 \leq t < \infty \quad (1)$$

where $T(\boldsymbol{x},t)$ is the temperature function, c is the specific heat, k the thermal conductivity and Ω denotes the domain and Γ denotes its boundary. We assume that the thermal conductivity and specific heat vary exponentially in the $z-$direction:

$$\begin{aligned} k(\boldsymbol{x},t) &= k_0 \exp(2\beta z) \\ c(\boldsymbol{x},t) &= c_0 \exp(2\beta z) \end{aligned} \quad (2)$$

where β is the so-called non-homogeneous parameter. After substituting Equation (2) into Equation (1), we obtain

$$\nabla^2 T(\boldsymbol{x},t) + 2\beta T_z = \frac{1}{\alpha}\frac{\partial T(\boldsymbol{x},t)}{\partial t} \quad (3)$$

with $\alpha = k_0/c_o$ and $T_z = \frac{\partial T}{\partial t}$. The Dirichlet boundary Γ_D and von Neumann boundary Γ_N conditions are given as:

$$\begin{aligned} T(\boldsymbol{x},t) &= \bar{T}, \boldsymbol{x} \in \Gamma_D, \\ q(\boldsymbol{x},t) = -k(\cdot)\frac{\partial T(\boldsymbol{x},t)}{\partial n} &= q(\bar{\boldsymbol{x}},t), \boldsymbol{x} \in \Gamma_N \end{aligned} \quad (4)$$

where n is the unit outward normal to Γ_N. In this paper, we assume that the initial temperature to be zero.

3 Physics-informed neural network using collocation method

In this section, the deep learning based collocation method that using physics-informed deep neural networks with Runge–Kutta (RK) integration schemes is introduced. First, a series of collocation points will be generated in the physical domain and at the boundaries denoted by $\boldsymbol{x}_{\Omega} = (x_1, ..., x_{N_{\Omega}})^T$ and $\boldsymbol{x}_{\Gamma} = (x_1, ..., x_{N_{\Gamma}})^T$, respectively, which formulates the dataset of training. Then the time-dependent heat conduction equation will be discretized using the classical Runge–Kutta method – with q stages.

3.1 Collocation points generation

To generate randomly distributed collocation points, various sampling strategies have been developed. The Halton and Hammersley sequences generate random points by constructing the radical inverse [32]. They are both low discrepancy sequences. Another approach is based on the Korobov Lattice [33]. Sobol Sequence is a quasi-random low-discrepancy sequence to generate sampling points [34]. Latin hypercube sampling (LHS) is a statistical method, where a near-random sample of parameter values is generated from a multidimensional distribution [35]. Monte Carlo methods create sampling points by repeated random sampling [36]. We have compared different sampling strategies for the steady state heat conduction equations in nonhonogeneous media in a previous study and found that Latin hypercube sampling (LHS) could yield favourable results with increasing layers [37]. Therefore, Latin hypercube sampling (LHS) is selected to generate collocation points in the transient heat transfer analysis.

3.2 PINNs with discrete time models

3.2.1 Runge–Kutta methods with q stages

For a general time-dependent partial differential equation – as $u'(t) = \mathcal{N}(t, u)$, by applying the general form of Runge-Kutta methods with q stages [21], an update iterative form can be obtained

$$
\begin{aligned}
u^n &= u_i^n, \quad i = 1, \ldots, q, \\
u^n &= u_{q+1}^n,
\end{aligned}
\tag{5}
$$

where

$$
\begin{aligned}
u_i^n &:= u^{n+c_i} - \Delta t \sum_{j=1}^{q} a_{ij} \mathcal{N}[u^{n+c_j}], \quad i = 1, \ldots, q, \\
u_{q+1}^n &:= u^{n+1} - \Delta t \sum_{j=1}^{q} b_j \mathcal{N}[u^{n+c_j}].
\end{aligned}
\tag{6}
$$

with $u^{n+c_j}(x) = u(t^n + c_j \Delta t, x)$ for $j = 1, \ldots, q$. Depending on the choice of the triple parameters $\{a_{ij}, b_j, c_j\}$, an implicit or explicit time-stepping scheme can be obtained. Generally, the matrix $\boldsymbol{A} = [a_{ij}]$ defines the Runge-Kutta matrix and $\boldsymbol{b} = b_i$ and $\boldsymbol{c} = c_i$ indicate the weights and nodes, which can be arranged in Butcher tableau as follows:

$$
\frac{\boldsymbol{c} \mid \boldsymbol{A}}{\boldsymbol{b}^T} =
\begin{array}{c|cccc}
c_1 & a_{11} & a_{12} & \ldots & a_{1q} \\
c_2 & a_{21} & a_{22} & \ldots & a_{2q} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
c_q & a_{q1} & a_{q2} & \ldots & a_{qq} \\
\hline
& b_1 & b_2 & \ldots & b_q
\end{array}
\tag{7}
$$

The theoretical error estimates for Runge–Kutta methods with q stages predict a temporal error accumulation of $\mathcal{O}\left(\Delta t^{2q}\right)$ assuming that $\Delta t < 1$. Otherwise the solution may not converge.

3.2.2 Discrete time approach

In a continuous time scheme, the neural network is used to approximate the mapping $(t, x) \mapsto u(t, x)$. The training data needs to be generated in the spatial-temporal domain, which can be too costly for analysis in high-dimensions, especially in the long time integration. On the other hand, we perform transient analysis in a discrete time scheme by placing a multi-output neural network prior on

$$
\left[u^{n+c_1}(x), \ldots, u^{n+c_q}(x), u^{n+1}(x) \right].
\tag{8}
$$

The neural network is used to approximate the mapping $(x) \mapsto (u^{n+c_1}(x), \ldots, u^{n+c_q}(x), u^{n+1}(x))$. This prior assumption along with equation (6) results in a *physics informed neural network* that takes x as an input and output:

$$
\left[u_1^n(x), \ldots, u_q^n(x), u_{q+1}^n(x) \right],
\tag{9}
$$

with $u(t^n, x) \approx u_{q+1}^n(x)$. The architecture of the neural network in this application can be found in Fig. 1. The colored round circle is the basic computational unit and the purple one on the output layer is the solution at time-step n.

3.3 Activation functions

To introduce the non-linearity regarding material variations into the neural network of Fig. 1 and enable the backpropagation, the activation function σ on hidden layers is

Fig. 1 Basic structure of the deep feed-forward neural network

Input Layer
$\in R^3$

Hidden Layer $\in R^i$

Hidden Layer
$\in R^{q+1}$

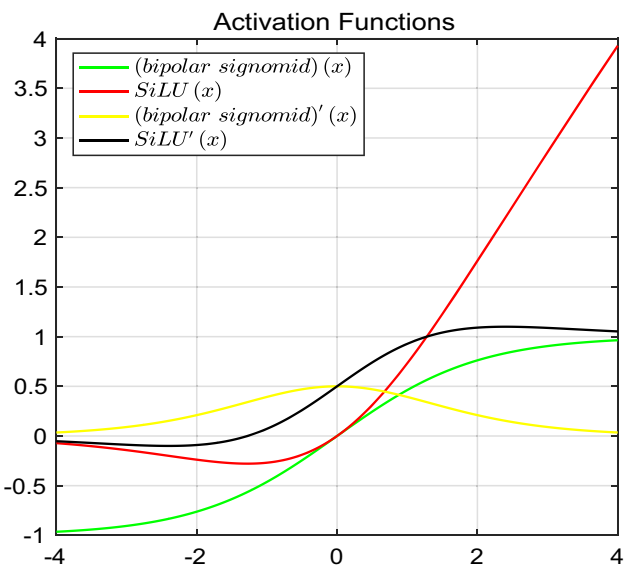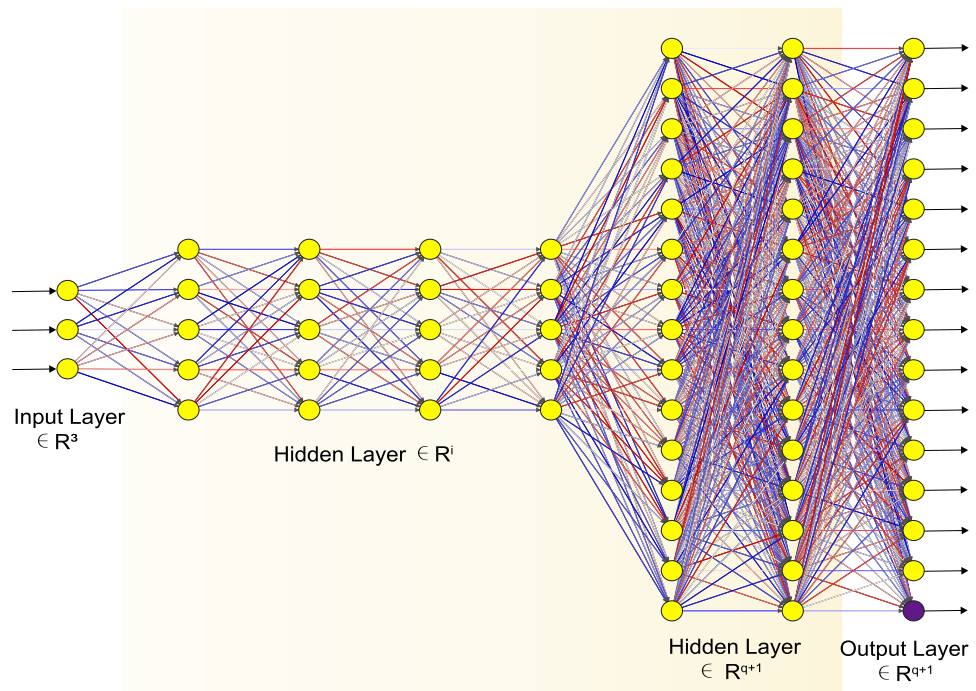Output Layer
$\in R^{q+1}$

Fig. 2 Two suggested activation functions for this dynamic neural network

defined. There are many activation functions σ available such as sigmoids function and hyperbolic tangent function ($Tanh$), to name a few [38]. Selecting the activation function in many cases still remains an open issue and commonly a trade-off between expressivity and trainability of the neural network [39]. For vibration analysis, Raissi et al. [40] report that the sinusoidal activation function is more stable than the hyperbolic tangent function ($Tanh$). For transient analysis with PINNs using discrete time scheme, we also found

that the hyperbolic tangent function ($Tanh$) cannot converge but the sinusoidal activation function is also not stable for many cases. To our experience, the bipolar sigmoid function $f(x) = \frac{e^x - 1}{e^x + 1}$ [41] and sigmoid-weighted linear unit (SiLU) function $f(x) = x \times sigmoid(x)$ [42] yield better results in transient dynamic analysis. The two activation functions that fitted for transient analysis are shown in

The Bipolar sigmoid function is a continuous activation function with a gradual output value in the range $[-1, 1]$, which looks similar to hyperbolic tangent function. However, the hyperbolic tangent function has a steeper slope. The sigmoid-weighted linear unit (SiLU) function resembles the classical ReLU activation but is nevertheless a smooth activation unbounded above but bounded below. Small negative values can capture underlying patterns from data, while large negative values may be filtered out to keep sparsity.

3.4 Physics-informed deep learning formulation

Taking advantage of the Runge Kutta method – substituting Equation (6) to Equation (3), we have

$$
\begin{aligned}
T_i^n &:= T^{n+c_i} - \Delta t \sum_{j=1}^q a_{ij}(\alpha \nabla^2 T^{n+c_j}(\boldsymbol{x}) + 2\alpha\beta T_z^{n+c_j}(\boldsymbol{x})), \quad i = 1, \ldots, q, \\
T_{q+1}^n &:= T^{n+1} - \Delta t \sum_{j=1}^q b_j(\alpha \nabla^2 T^{n+c_j}(\boldsymbol{x}) + 2\alpha\beta T_z^{n+c_j}(\boldsymbol{x})).
\end{aligned}
\tag{10}
$$

Placing a multi-output neural network prior with coordinates as inputs $\boldsymbol{x}$ on temperatures $(T^{n+c_1}(\boldsymbol{x}), \ldots, T^{n+c_q}(\boldsymbol{x}), T^{n+1}(\boldsymbol{x}))$ yields

$$(T^{n+c_1}(\boldsymbol{x}), \ldots, T^{n+c_q}(\boldsymbol{x}), T^{n+1}(\boldsymbol{x})) \approx \boldsymbol{f}(\boldsymbol{x};\theta), \quad (11)$$

Combined with Equation (10), we can devise the physics-informed neural network that outputs $(T_1^n(\boldsymbol{x}), \ldots, T_q^n(\boldsymbol{x}), T_{q+1}^n(\boldsymbol{x}))$ which are approximated by

$$
\begin{aligned}
&(T_1^n(\boldsymbol{x};\theta), \ldots, T_q^n(\boldsymbol{x};\theta), T_{q+1}^n(\boldsymbol{x};\theta)) \\
&= \boldsymbol{f}(\boldsymbol{x};\theta) - \Delta t[\boldsymbol{A};\boldsymbol{b}^T]\mathcal{N}(\boldsymbol{f}(\boldsymbol{x};\theta)).
\end{aligned} \quad (12)
$$

where $\mathcal{N} = \alpha\nabla^2 + 2\alpha\beta$ is the differential operator. The loss function thus constructed from the mean square error is given by

$$Loss = MSE_n + MSE_b, \quad (13)$$

with

$$MSE_n = \frac{1}{N}\sum_{j=1}^{q+1}\sum_{i=1}^{N_n}\|T_j^n(\boldsymbol{x}^{n,i};\theta) - T^{n,i}\|^2, \quad (14)$$

and

$$MSE_b = MSE_{T_{\Gamma_D}} + MSE_{q_{\Gamma_N}}. \quad (15)$$

where $MSE_{T_{\Gamma_D}}$ and $MSE_{q_{\Gamma_N}}$ are defined as:

$$MSE_{T_{\Gamma_D}} = \frac{1}{N_{\Gamma_D}}\sum_{i=1}^{q}\sum_{i=1}^{N_{n_{\Gamma_D}}}\|T^{n+c_i}(\boldsymbol{x}_{\Gamma_D}^{n,i};\theta) - \bar{T}\|^2. \quad (16)$$

and

$$MSE_{q_{\Gamma_N}} = \frac{1}{N_{\Gamma_N}}\sum_{i=1}^{q}\sum_{i=1}^{N_{n_{\Gamma_N}}}\|q^{n+c_i}(\boldsymbol{x}_{\Gamma_N}^{n,i};\theta) - \bar{q}\|^2. \quad (17)$$

The transient analysis with discrete time PINNs model is reduced to an optimization problem:

$$\hat{\theta} = \underset{\theta \in R^K}{\arg\min}\, Loss\,(\theta) \quad (18)$$

One of the most widely used optimization method to train the physics-informed neural network is the combined Adam-L-BFGS-B optimization algorithm. This strategy consists of
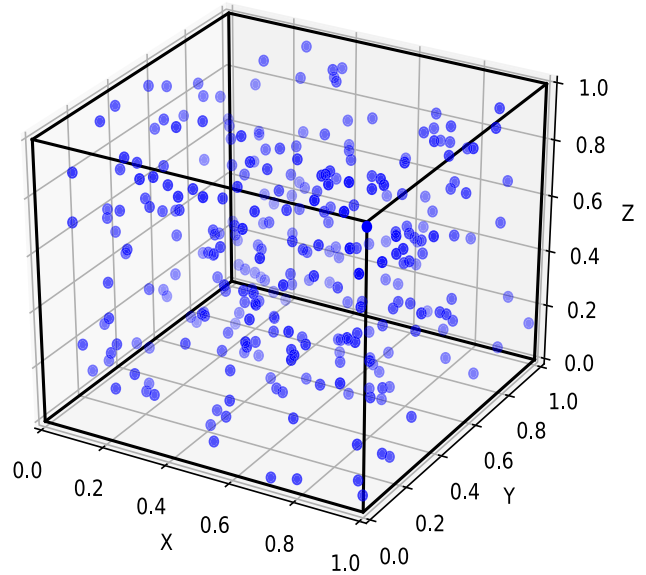

Fig. 4 Random sampling inside the cubic domian

Fig. 3 Schematic of a physics-informed deep learning with discrete time scheme

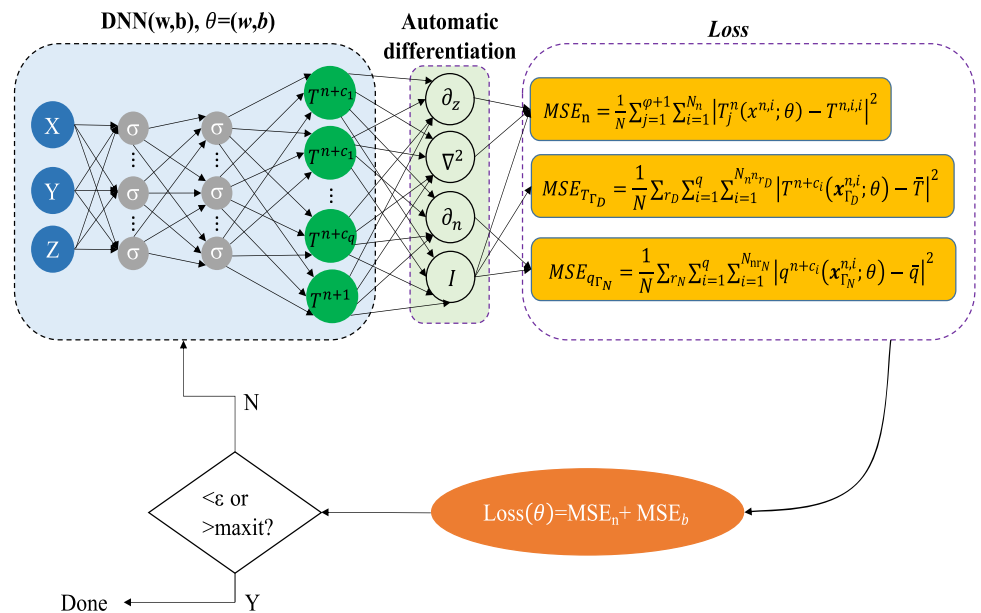**Fig. 5** **a** Predicted temperature and **b** analytical temperature distributions for the functionally graded unit cube at time $t = 0.1s$
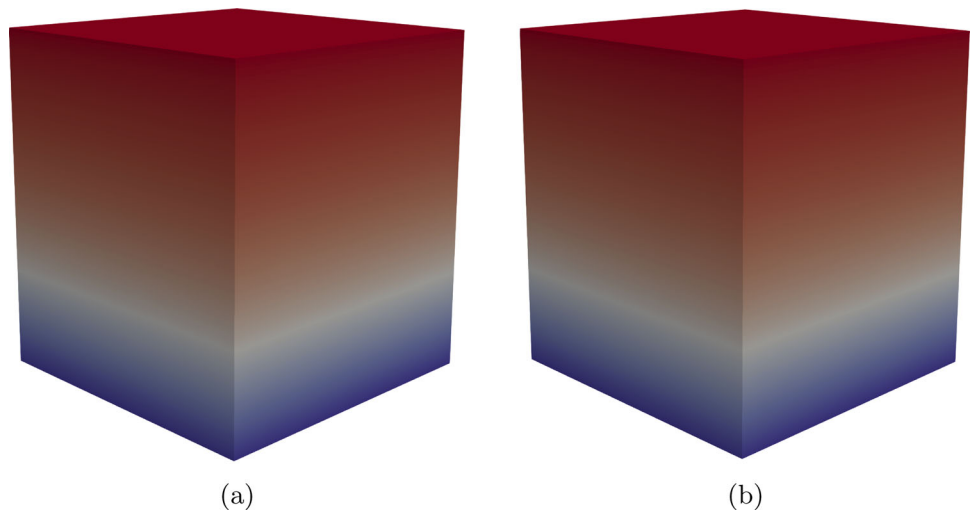

(a)   (b)

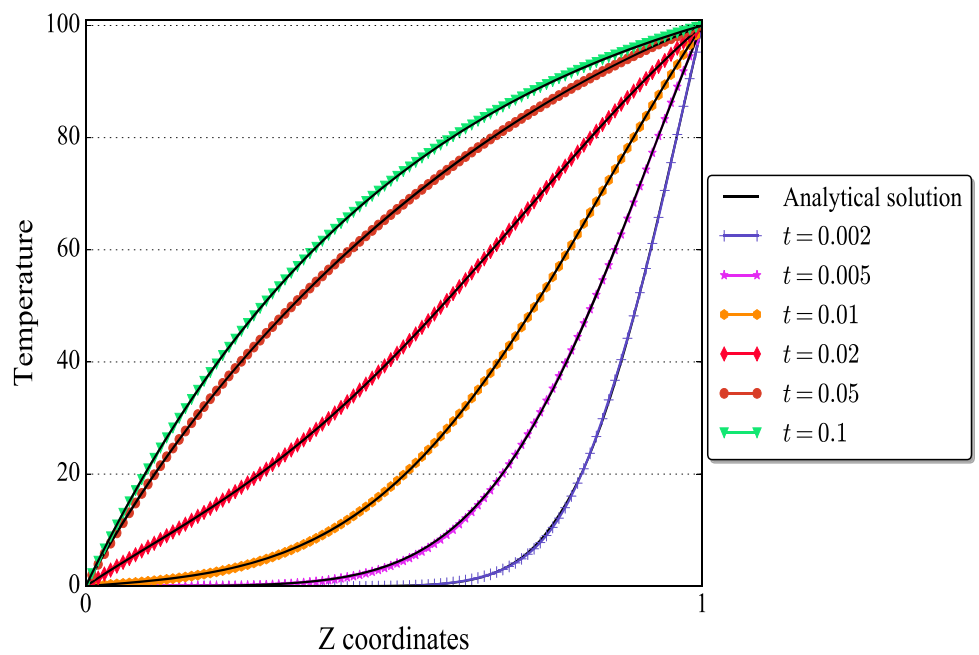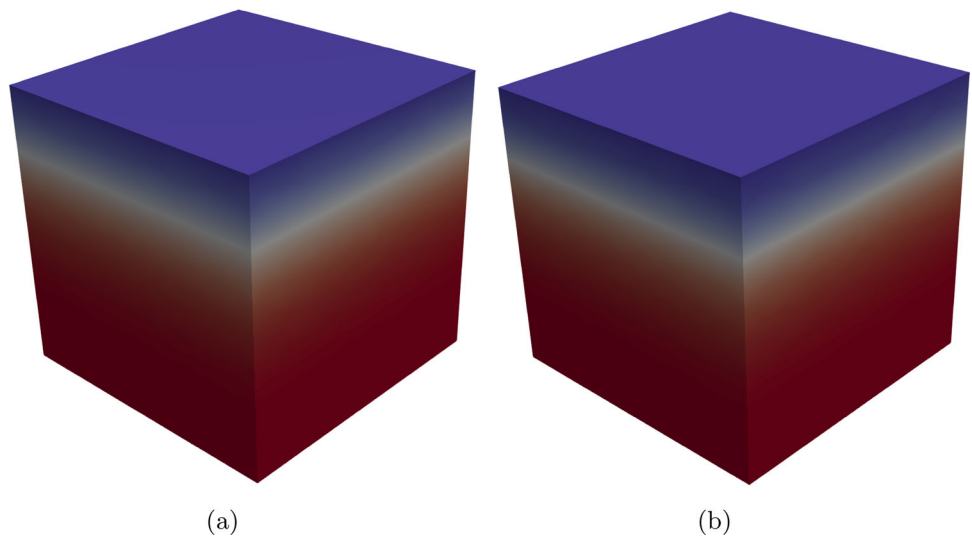**Fig. 6** Temperature profiles in z direction at different time levels for the FGM cube

Legend:
- Analytical solution
- $t = 0.002$
- $t = 0.005$
- $t = 0.01$
- $t = 0.02$
- $t = 0.05$
- $t = 0.1$

Fig. 7 **a** Predicted flux and **b** analytical flux distributions for the functionally graded unit cube at time $t = 0.005s$

(a)   (b)
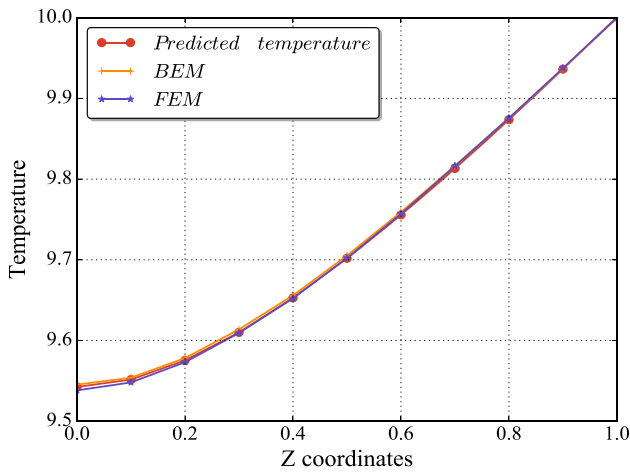
Fig. 8 Temperature profile in z direction at time t=1 for the FGM cube problem with time-dependent boundary condition

training the network first using the Adam algorithm and after a defined number of iterations, we perform the L-BFGS-B optimization of the loss with a small number of executions.

The basic scheme of the 'discrete physics-informed deep learning' is shown in Fig. 3. A fully-connected neural network with space coordinates as inputs is first applied to approximate the temperature at Runge–Kutta nodes and at time step $(n+1)\Delta t$. Then, the derivatives of the temperature outputs are calculated using automatic differentiation (AD), which is then used to formulate the loss. The hyperparameters θ are learnt by minimizing the loss function.

4 Numerical examples

We demonstrate the presented approach through four benchmark problems; 100 stages are employed for the discrete time scheme. We fix the first six hidden layers with 20 neurons per layer and the rest layers are set to be $q+1$ (101) neurons per layer.

4.1 Case 1: FGMs with exponential material gradation

Let us consider a unit cube shown in Fig. 4 where the material properties vary smoothly and continuously in the z-direction. The initial and boundary conditions are given as follows:

$$T(x, y, z; 0) = 0, \tag{19}$$

and

$$
\begin{cases}
q(1, y, z; t) = 0; \ q(0, y, z; t) = 0 \\
q(x, 1, z; t) = 0; \ q(x, 0, z; t) = 0 \\
T(x, y, 1; t) = 100; \ T(x, y, 0; t) = 0
\end{cases} \tag{20}
$$

The thermal conductivity parameter in Eq. (2) is $k_0 = 5$ and specific heat parameter is $c_0 = 1$ and non-homogeneous $\beta = 1$. The analytical solution for this problem is given as:

$$
T(x, y, z; t) = T\frac{1 - e^{-2\beta z}}{1 - e^{-2\beta L}} + \sum_{n=1}^{\infty} B_n sin\left(\frac{n\pi z}{L}\right)
$$
$$
e^{-\beta z}e^{-(n^2\pi^2/L^2 + \beta^2)\alpha t}, \tag{21}
$$

with

$$
B_n = -\frac{2Te^{\beta L}}{\beta^2 L^2 + n^2\pi^2}\left[\beta L sin(n\pi)\frac{1 + e^{-2\beta L}}{1 - e^{-2\beta L}}\right.
$$
$$
\left. -n\pi cos(n\pi)\right] \tag{22}
$$

Fig. 9 **a** Predicted temperature and **b** predicted flux distributions for the functionally graded unit cube at time $t = 1s$
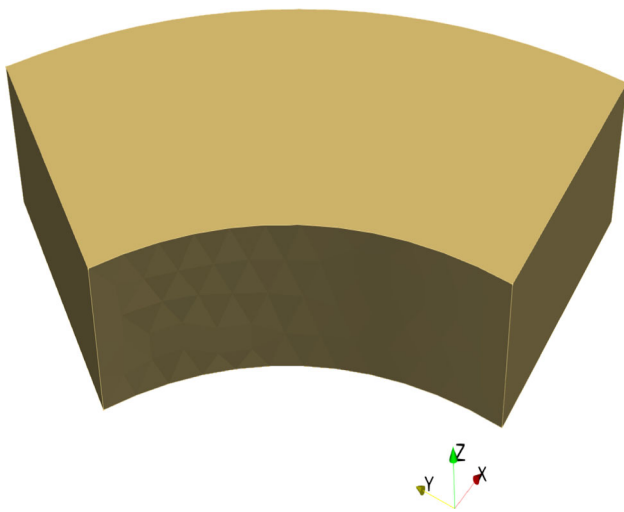

(a)

(b)

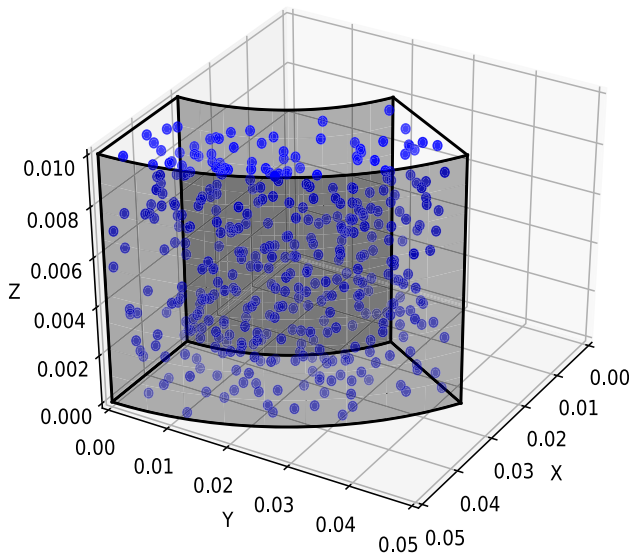**Fig. 10** Geometry of the FGMs with irregular domain

Fig. 11 Random sampling inside the irregular domain

$L = 1$ being the length of the cube and $T = 100$. A random sampling of collocation points as illustrated in Fig. 4 is generated inside the cubic, with 400 collocation points in the domain and 80 collocation points on the boundaries. The learning rate is set to 0.001.

The predicted and analytical temperature at time $t = 0.1s$ are illustrated in Fig. 5. The temperature profiles in z direction at different time levels ranging from 0.002 to 0.1 s for the FGM cube are shown in Fig. 6 and compared with the analytical solution. At each time level the predicted temperature coincides with the analytical solution. The flux distribution at time $t = 0.005s$ is shown in Fig. 7 and agrees well with the analytical solution. The relative error between the predicted temperature and the analytical solution at time $t = 0.1s$ inside the whole functionally graded cube is 3.661e−04.
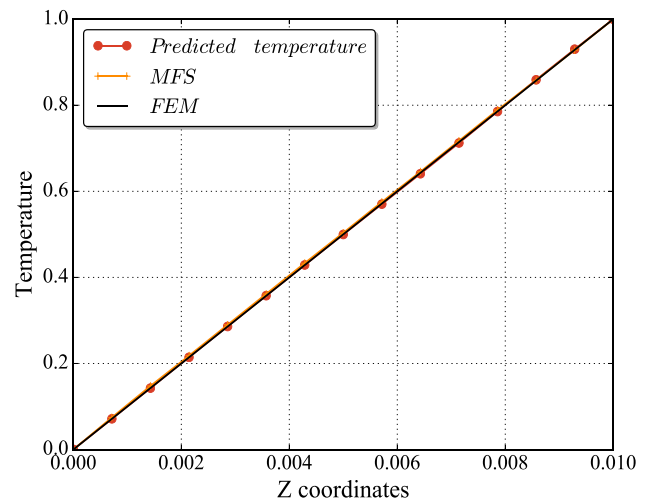

Fig. 12 Temperature profile in z direction at time $t = 0.00001\,s$ for the FGM with irregular domain

4.2 Case 2: FGMs with time-dependent boundary condition

For this unit cube, the top surface is prescribed with a time-dependent boundary condition $T(x, y, 1; t) = 10t$ and all other surfaces are insulated. The initial and boundary conditions are given as

$$T(x, y, z; 0) = 0, \tag{23}$$

and

$$\begin{cases} q(1, y, z; t) = 0; q(0, y, z; t) = 0 \\ q(x, 1, z; t) = 0; q(x, 0, z; t) = 0 \\ T(x, y, 1; t) = 10t; T(x, y, 0; t) = 0 \end{cases} \tag{24}$$

The thermal conductivity and specific heat parameters in Equation (2) are $k_0 = 5$ and $c_0 = 1$. Non-homogeneous parameter $\beta = 1.5$. The temperature profile varies again only in z direction; see Fig. 8 at time $t = 1s$ compared with a BEM solution using 1200 elements and FEM solution with 1000 linear brick elements [43]. The temperature and flux contours at time $t = 1s$ can be found in Fig. 9. Compared with BEM and FEM, deep collocation method is more easy in implementation without the necessity of building elaborate grids and once the deep learning model is trained, it can be deployed to predict the temperature and flux distribution in seconds while maintain the same level of accuracy, which can be a suitable surrogate model for tradition numerical methods.

4.3 Case 3: FGMs with Irregular domain

Let us consider now a problem with an irregular domain as shown in Fig. 10. The inner and outer radii (r_1 and r_2) of

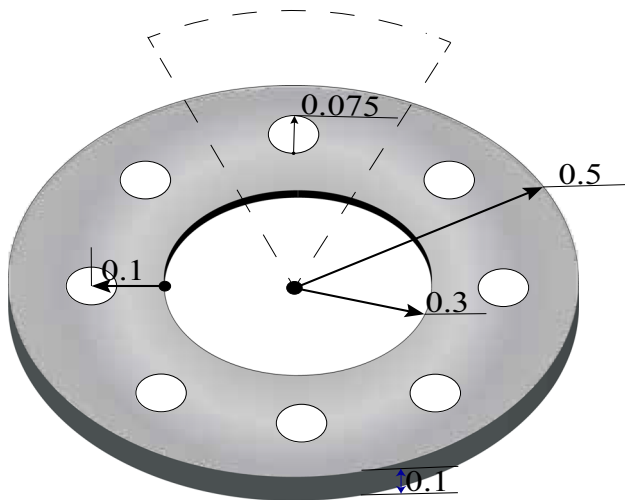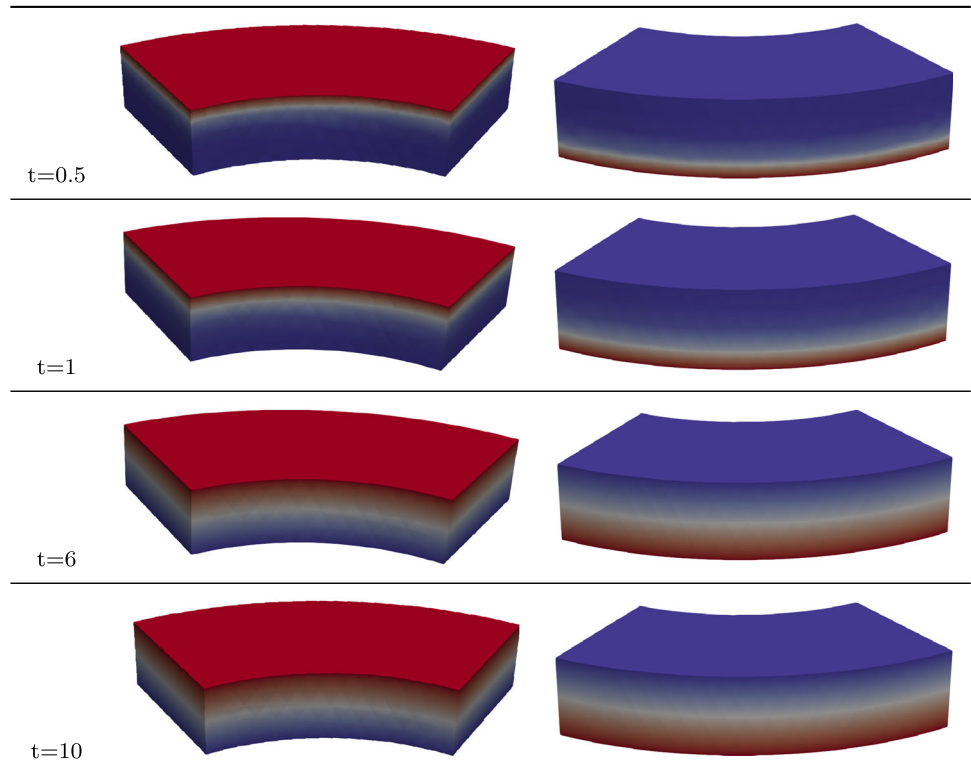**Table 1** Temperature distributions for irregular FGMs at different time levels

t=0.5		
t=1		
t=6		
t=10		

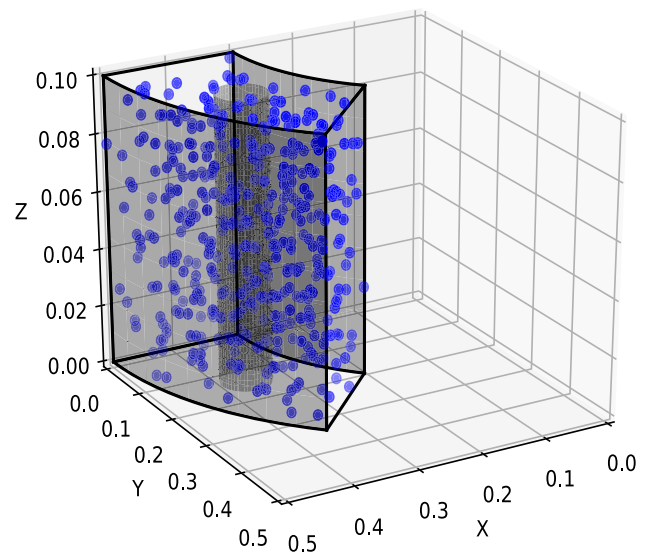**Fig. 13** Geometry of the functionally graded rotor
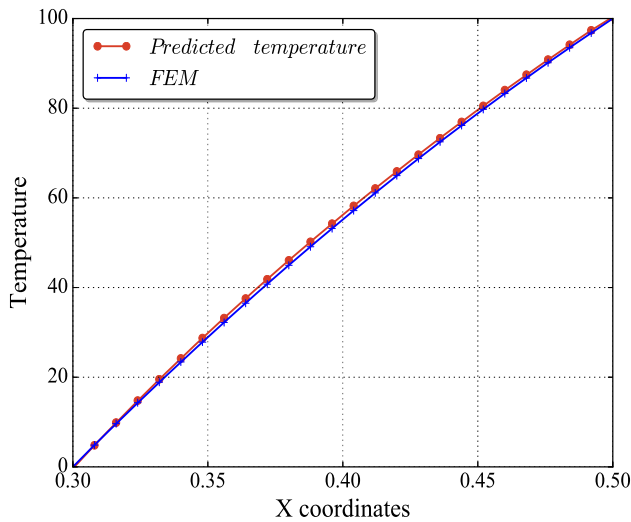
Fig. 14 Random sampling inside the rotor domian

Fig. 15 Temperature profile along the right top edge at time $t = 0.0066s$ for the FGM rotor problem

this annular cylinder are 0.03 and 0.05. The height of this annular cylinder is 0.01. The central angle is $\frac{\pi}{3}$. The material parameters are set as: thermal conductivity parameter $k_0 = 5$, the specific heat parameter $c_0 = 1$ and $\beta = 1.5$. The initial and boundary conditions are

$$T(r, \theta, z; 0) = 0, \tag{25}$$

and

$$\begin{cases} q(r_2, \theta, z; t) = 0; \ q(r_1, \theta, z; t) = 0 \\ q(r, \frac{\pi}{3}, z; t) = 0; \ q(r, 0, z; t) = 0 \\ T(r, \theta, 0.01; t) = 1; \ T(r, \theta, 0; t) = 0 \end{cases} \tag{26}$$

where $r_1 \le r \le r_2$ and $0 \le \theta \le \frac{\pi}{3}$.

First, we generate random collocation points inside the irregular physical domain as shown in Fig. 11, with 400 collocation points inside the domain and 200 collocation points on the boundaries.

Figure 12 depicts the temperature profile along the z direction at $t = 0.00001s$ and compared with the reference solutions from MFS and FEM [43].

Table 1 shows the profile of the overall temperature distribution. The results agree well with the ones from [43]. The gradation of the temperature along z-axis matches the material gradation property of the FGM.

4.4 Case 4: Functionally graded rotor problem

Finally, we study a functionally graded rotor with eight holes presented in Fig. 13. Due to the symmetry, only one-eighth of the rotor is analyzed. The geometric parameters of this rotor are marked in Fig. 13. All lines with arrows imply the length, namely the inner radius is $R_{inner} = 0.5$, outer radius

is $R_{outer} = 0.3$ and the height is 0.1; the diameter of the mounting hole is $Dia_{hole} = 0.075$. The thermal conductivity and specific heat parameters in Equation (2) are $k_0 = 5$ and $c_0 = 1$. Non-homogeneous parameter $\beta = 1.5$. The initial conditions are

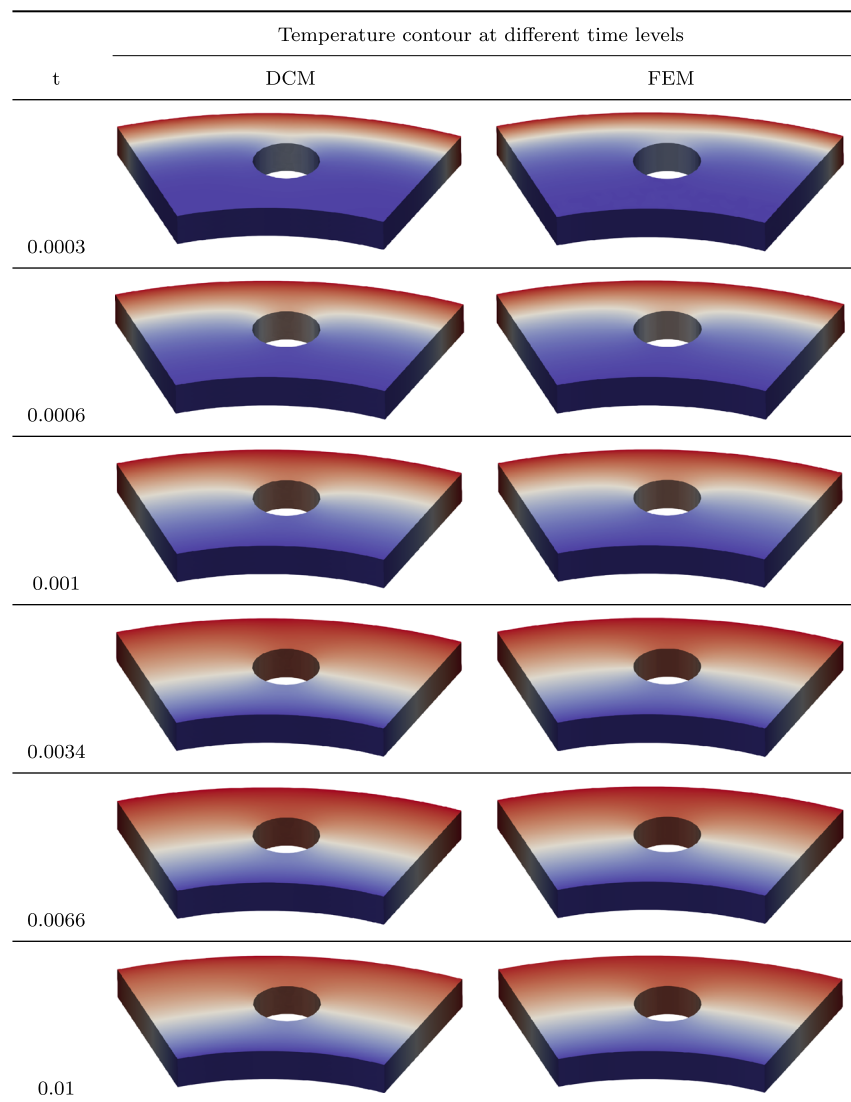$$T(r, \theta, z; 0) = 0 \tag{27}$$

and temperature boundary conditions are imposed on the inner side/surface (0K) and outer side/surface (100K) while other surfaces are adiabatic in which the heat flux is set to 0. Figure 14 shows the collocation points for training. Our results are compared to results of ABAQUS simulations.

The temperature profile along the right top edge at time $t = 0.0066\,s$ is shown in Fig. 15 and the temperature distributions at different time levels is listed in Table 2. It can be observed that predicted temperature at specific locations and time matches well with ABAQUS results. The same can be observed for the evolution of temperature distribution with both numerical methods.

5 Conclusion

We presented a deep learning based collocation method for transient heat transfer analysis of three-dimensional functionally graded materials (FGMs). This deep collocation method combines the classical collocation method and the deep learning method in one framework which is easy in implementation and no necessity to build elaborate grids. For the deep learning model, a physics-informed neural network is combined with a q-stage Runge–Kutta discrete time scheme for transient heat transfer analysis. Nonlinear activation functions are adopted to introduce a nonlinearity into the neural network and fitted for dynamic analysis. We found the bipolar sigmoid function and sigmoid-weighted linear unit (SiLU) function most suitable for the physics-informed neural network for the transient analysis. Based on our previous study, Latin Hypercube sampling is selected for random sampling of collocation points making the proposed truly "meshfree" deep collocation method, such that it can deal with irregular shaped domains easily.

Various numerical examples were studied to validate the performance of proposed method including FGMs with an irregular shape and heat conduction with a variety of boundary conditions. From numerical results, it can be concluded that both temperature and flux inside FGMs predicted by deep collocation method with discrete time scheme and fitted activation function agree well with analytical solutions and other classical numerical methods. The physics-informed deep collocation method can be promising as surrogate models for FEM in dynamic analysis.

Table 2 Temperature contour at different time levels

t	Temperature contour at different time levels	
	DCM	FEM
0.0003		
0.0006		
0.001		
0.0034		
0.0066		
0.01		

References

1. Koizumi M, Niino M (1995) Overview of FMG research in Japan. MRS Bull 20(1):19–21

2. Saleh B, Jiang J, Fathi R, Al-hababi T, Qiong X, Wang L, Song D, Ma A (2020) 30 years of functionally graded materials: an overview of manufacturing methods, applications and future challenges. Compos B Eng 201:108376

3. Wang B-L, Tian Z-H (2005) Application of finite element-finite difference method to the determination of transient temperature field in functionally graded materials. Finite Elem Anal Des 41(4):335–349

4. Wenzhen Q, Fan C-M, Zhang Y (2019) Analysis of three-dimensional heat conduction in functionally graded materials by using a hybrid numerical method. Int J Heat Mass Transf 145:118771

5. Mohebbi F, Evans B, Rabczuk T (2021) Solving direct and inverse heat conduction problems in functionally graded materials using an accurate and robust numerical method. Int J Therm Sci 159:106629

6. Wang B-L, Mai Y-W (2005) Transient one-dimensional heat conduction problems solved by finite element. Int J Mech Sci 47(2):303–317
7. Fu Z-J, Qin Q-H, Chen W (2011) Hybrid-trefftz finite element method for heat conduction in nonlinear functionally graded materials. Eng Comput 5:89
8. Sladek J, Sladek V, Zhang Ch (2003) Transient heat conduction analysis in functionally graded materials by the meshless local boundary integral equation method. Comput Mater Sci 28(3–4):494–504
9. Zhuo-Jia F, Qiang X, Wen C, Cheng Alexander H-D (2018) A boundary-type meshless solver for transient heat conduction analysis of slender functionally graded materials with exponential variations. Comput Math Appl 76(4):760–773
10. Wen H, Yan G, Fan C-M (2020) A meshless collocation scheme for inverse heat conduction problem in three-dimensional functionally graded materials. Eng Anal Boundary Elem 114:1–7
11. Alok S, Paulino Glaucio H, Gray LJ (2002) Transient heat conduction in homogeneous and non-homogeneous materials by the Laplace transform Galerkin boundary element method. Eng Anal Boundary Elem 26(2):119–132
12. Alok S, Paulino Glaucio H (2004) The simple boundary element method for transient heat conduction in functionally graded materials. Comput Methods Appl Mech Eng 193(42–44):4511–4539
13. Abreu AI, Canelas A, Mansur WJ (2013) A cqm-based bem for transient heat conduction problems in homogeneous materials and FMGs. Appl Math Model 37(3):776–792
14. Xi Q, Zhuojia F, Zhang C, Yin D (2021) An efficient localized Trefftz-based collocation scheme for heat conduction analysis in two kinds of heterogeneous materials under temperature loading. Comput Struct 255:106619
15. Zhuojia F, Tang Z, Xi Q, Liu Q, Yan G, Wang F (2022) Localized collocation schemes and their applications. Acta Mech Sin 38(7):422167
16. Lagaris Isaac E, Aristidis L, Fotiadis Dimitrios I (1997) Artificial neural network methods in quantum mechanics. Comput Phys Commun 104(1–3):1–14
17. Samir K, Samir T, Cuong LT, Emad G, Seyedali M, Abdel WM (2021) An improved artificial neural network using arithmetic optimization algorithm for damage assessment in FGM composite plates. Compos Struct 273:114287
18. Wang S, Wang H, Zhou Y, Liu J, Dai Peng D, Magd XW, Abdel (2021) Automatic laser profile recognition and fast tracking for structured light measurement using deep learning and template matching. Measurement 169:108362
19. Viet HL, Huong ND, Mohsen M, Guido DR, Thanh B-T, Gandomi Amir H, Abdel WM (2021) A hybrid computational intelligence approach for structural damage detection using marine predator algorithm and feedforward neural networks. Comput Struct 252:106568
20. Viet HL, Thi TT, Guido DR, Thanh B-T, Long N-N, Abdel WM (2022) An efficient stochastic-based coupled model for damage identification in plate structures. Eng Fail Anal 131:105866
21. Maziar R, Paris P, Karniadakis George E (2019) Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 378:686–707
22. Raissi M (2018) Forward-backward stochastic neural networks: deep learning of high-dimensional partial differential equations. arXiv preprint arXiv:1804.07010
23. Maziar Raissi and George Em Karniadakis (2018) Hidden physics models: machine learning of nonlinear partial differential equations. J Comput Phys 357:125–141
24. Raissi M (2018) Deep hidden physics models: deep learning of nonlinear partial differential equations. J Mach Learn Res 19(1):932–955
25. George EK, Ioannis GK, Lu L, Paris P, Sifan W, Liu Y (2021) Physics-informed machine learning. Nat Rev Phys 3(6):422–440
26. Shin Y, Darbon J, Karniadakis GE (2020) On the convergence and generalization of physics informed neural networks. arXiv preprint arXiv:2004.01806
27. Tang Zhuochao F, Zhuojia RS (2022) An extrinsic approach based on physics-informed neural networks for pdes on surfaces. Mathematics 10(16):87
28. Noakoasteen O, Wang S, Peng Z, Christodoulou C (2020) Physics-informed deep neural networks for transient electromagnetic analysis. IEEE Open J Antennas Propag 1:404–412
29. Deng L, Pan Y (2021) Application of physics-informed neural networks for self-similar and transient solutions of spontaneous imbibition. J Petrol Sci Eng 203:108644
30. Yu BX, Meng Q, Gao Q (2022) Physics-informed neural networks for solving steady and transient heat conduction problems of functionally graded materials. Chin J Comput Mech 5:1–10
31. Sifan W, Paris P (2022) Long-time integration of parametric evolution equations with physics-informed deeponets. J Comput Phys 5:111855
32. Rafajłowicz E, Schwabe R (2006) Halton and hammersley sequences in multivariate nonparametric regression. Stat Prob Lett 76(8):803–812
33. Xiaoqun W, Sloan Ian H, Josef D (2004) On Korobov lattice rules in weighted spaces. SIAM J Numer Anal 42(4):1760–1779
34. Sobol' IM (1967) On the distribution of points in a cube and the approximate evaluation of integrals. USSR Comput Math Math Phys 7(4):86–112
35. Shields Michael D, Jiaxin Z (2016) The generalization of Latin hypercube sampling. Reliab Eng Syst Saf 148:96–108
36. Shapiro A (2003) Monte Carlo sampling methods. Handbooks Oper Res Manag Sci 10:353–425
37. Hongwei G, Xiaoying Z, Pengwan C, Naif A, Timon R (2022) Analysis of three-dimensional potential problems in non-homogeneous media with physics-informed deep collocation method using material transfer learning and sensitivity analysis. Eng Comput 5:17
38. Apicella A, Donnarumma F, Isgrò F, Prevete R (2021) A survey on modern trainable activation functions. Neural Netw 138:14–32
39. Raghu M, Poole B, Kleinberg J, Ganguli S, Sohl-Dickstein J (2017) On the expressive power of deep neural networks. In: international conference on machine learning, pp 2847–2854. PMLR
40. Maziar R, Zhicheng W, Triantafyllou Michael S, Em KG (2019) Deep learning of vortex-induced vibrations. J Fluid Mech 861:119–137
41. Panicker M, Babu C (2012) Efficient FPGA implementation of sigmoid and bipolar sigmoid activation functions for multilayer perceptrons. IOSR J Eng 2:1352–6
42. Elfwing S, Uchibe E, Doya K (2018) Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. Neural Netw 107:3–11
43. Ming L, Chen CS, Chu CC, Young DL (2014) Transient 3D heat conduction in functionally graded materials by the method of fundamental solutions. Eng Anal Bound Elem 45:62–67